

# GEOPHYSICAL FLUID MECHANICS

## VOLUME 3

SHALLOW WATER FLOWS, VORTICITY, NEARLY GEOSTROPHIC FLOWS,  
VERTICAL COORDINATES, AND SCALAR FIELDS

Stephen.M.Griffies@gmail.com

Princeton University Atmospheric and Oceanic Sciences

Draft from February 19, 2026



COPYRIGHT ©2026 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	vii
GUIDE TO THIS BOOK	xvii
<b>Part I. Shallow water flows</b>	<b>1</b>
1 SHALLOW WATER MODELS	3
2 SHALLOW WATER DYNAMICS	39
<b>Part II. Vorticity</b>	<b>87</b>
3 VORTICITY AND CIRCULATION	89
4 TWO-DIMENSIONAL NON-DIVERGENT FLOW	115
5 SHALLOW WATER VORTICITY AND POTENTIAL VORTICITY	163
6 VORTICITY AND CIRCULATION MECHANICS	209
7 POTENTIAL VORTICITY MECHANICS	263
8 POTENTIAL VORTICITY BUDGETS	289
<b>Part III. Nearly geostrophic flows</b>	<b>311</b>
9 MODELS OF NEARLY GEOSTROPHIC FLOWS	313
10 PLANETARY GEOSTROPHIC VORTICITY ANALYSIS	351
11 FOUNDATIONS OF QUASI-GEOSTROPHY	375
<b>Part IV. Generalized vertical coordinates</b>	<b>403</b>
12 FOUNDATIONS IN TENSOR CALCULUS	407
13 MECHANICAL EQUATIONS	433
14 CONTINUOUS AND LAYERED ISOPYCNAL MODELS	471
<b>Part V. Scalar fields</b>	<b>507</b>
15 DIFFUSIVE PROCESSES	511

16 TRACER ADVECTION AND DIFFUSION	543
17 EDDY-MEAN INTERACTION FOR SCALARS	591
18 PARAMETERIZED OCEAN TRACER TRANSPORT	623
19 OCEAN DENSITY AND SEA LEVEL	659
20 WATER MASS TRANSFORMATION THEORY	689
 <b>Part VI. Solutions to Exercises</b>	 <b>737</b>
 <b>Part VII. End matter</b>	 <b>741</b>
A GLOSSARY OF CONCEPTS AND TERMS	743
B LIST OF ACRONYMS	785
C LIST OF SYMBOLS	787
BIBLIOGRAPHY	797
INDEX	819

## 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 \(2025\)](#)

## Disclaimer

Reference has been made to the published and unpublished literature to support the development of this book, as well as consultations with large language models to verify citations, refine exercise solutions, and clarify glossary entries. All material from such sources has been extensively reworked and refined into the author's voice for purposes of uniformity, pedagogy, and clarity.





## 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 multiple volumes, each of which comprises parts with 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.

## Generalized vertical coordinates

The chapters on [generalized vertical coordinate \(GVC\)s \(GVC\)](#)s 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.



## Volume 4

### 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.

### 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.



### Solution manual

A solution manual is available for students (with partial solutions provided) and instructors (with full solutions). Solutions aim 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, exposing details helps to identify when an incorrect result follows from a physical/conceptual error (e.g., incorrect setup of the problem solution) or from a mathematical error (e.g., sign error).



## **Part I**

### **Shallow water flows**

The ocean and atmosphere are continuously stratified fluids, meaning that density varies in the vertical through action of the Earth's gravity field. But for many conceptual and practical purposes, we find it useful to make the [shallow water approximation](#), in which the continuous fluid is vertically discretized into constant density homogeneous fluid layers. If the layer interfaces are material (no matter or thermodynamic properties are transferred between the immiscible layers), then thermodynamic processes are absent, in which we refer to the adiabatic [shallow water model](#). Our goal in this part of the book is to develop an understanding of the shallow water model fundamentals and to study some of its dynamical properties.

Pressure is hydrostatic in the shallow water model, which, as we show, means that motion occurs in vertical columns whereby horizontal velocity is independent of the vertical position within a layer, whereas the column can expand and contract so that there is vertical motion within a layer. Horizontal momentum is transferred between undulating shallow water layers via the pressure [form stress](#) studied in [VOLUME 2](#). To incorporate planetary rotation, we make approximations arising from the [hydrostatic primitive equations](#) studied in [VOLUME 2](#), so that the Coriolis acceleration is based on just the local vertical component to the planetary rotation.

The shallow water model provides a versatile theoretical toolbox for deducing how flow is affected by rotation and (when multiple layers are used) discretized vertical stratification. Consequently, the shallow water model is featured in many areas of geophysical fluid mechanics as well as in ocean and atmosphere applications. Throughout our studies, we make particular use of shallow water models for the study of form stress, vorticity, potential vorticity, waves, instabilities, wave-mean flow interaction, vertical coordinates, and circulation. Indeed, as noted in Section 1.6 of [Bühler \(2014a\)](#), the shallow water model can be viewed as describing a shallow layer of an incompressible and homogeneous fluid (which is how we make use of the model), or as a two-dimensional compressible flow with direct connections to the study of gas dynamics and acoustics. Namely, the acoustic waves of compressible flow are directly analogous to the gravity waves of shallow water flow. However, shallow water flow is somewhat simpler than compressible three-dimensional flow, since shallow water flow is horizontal within a layer. For these reasons, the shallow water model has found great use both inside and outside of geophysical fluid mechanics.

The shallow water model is a vertically discrete realization of a continuously stratified fluid described by the [isopycnal](#) coordinates of [VOLUME 4](#). However, the mathematical formalism of [generalized vertical coordinates](#), and the need to formulate the fluid equations using isopycnal coordinates, is unnecessary when working with the shallow water model. The reason for the simplification is that columnar motion within a shallow water layer means that lateral gradients of properties need not be projected along the slope of the layer. In contrast, this projection is needed for a continuously stratified fluid described by generalized vertical coordinates. So although there is beauty and power in the methods of generalized vertical coordinates for studying continuously stratified flows, it is liberating to avoid that formalism while still capturing much of the underlying physics associated with stratification (albeit discretely stratified). This feature of the shallow water model greatly adds to its physical allure and mathematical accessibility.

#### 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](#).

# Chapter 1

## SHALLOW WATER MODELS

In this chapter we formulate the mechanical equations for a suite of shallow water models. In later chapters in this book, we make use of these models for a variety of geophysical fluid applications, such as the study of form stresses, vorticity, potential vorticity, waves, instabilities, wave-mean flow decomposition, vertical coordinates, and circulation. Indeed, the shallow water model can be viewed as describing a shallow layer of an incompressible and homogeneous fluid, which is how we make use of the model, or as a two-dimensional compressible flow with direct connections to the study of gas dynamics and acoustics. Evidently, the shallow water model offers a rich and elegant home for many questions of both conceptual and technical nature, and as such we return to it frequently in this book. For this chapter, we focus on establishing the basis for these equations, both for the single shallow water layer as well as multiple shallow water layers (stacked shallow water). We also introduce the reduced gravity models commonly used for ocean circulation studies, which are models with one layer that is dynamically inactive.

### READER'S GUIDE TO THIS CHAPTER

We make use of fluid kinematics and dynamics described in a variety of earlier chapters, with the presentation inspired by Chapter 3 of [Vallis \(2017\)](#) as well as various sections in [Salmon \(1998\)](#). We make use of the formulation for further study of the shallow water fluid mechanics, including dynamical balances in Chapter 2, vorticity mechanics in Chapter 5, and wave mechanics in VOLUME 4. [This video](#) offers a pedagogical introduction to shallow water flows. We commonly make use of ocean language and refer to the fluid as water. Even so, the shallow water model has many applications to the study of large-scale atmospheric flows.

The fluid density is constant within a shallow water layer, so that mass conservation is the same as volume conservation. Hence, the terms “mass conservation” and “volume conservation” are commonly used interchangeably when working with shallow water models. The horizontal velocity is vertically uniform within a shallow water layer, whereas the vertical velocity and hydrostatic pressure are linear functions of vertical position within the layer. When acting on a property that is vertically uniform within a layer, the gradient operator,  $\nabla$ , results in a horizontal vector. To minimize notational clutter, we typically write  $\nabla$  for brevity, rather than  $\nabla^k$  or  $\nabla_h$  (with  $k$  the layer index). The meaning of the resulting vector equations are clear from the functional dependencies of the fields present in the equations.

1.1	Loose threads	4
1.2	A single shallow water layer	4
1.2.1	Pressure gradient force within the fluid layer	5
1.2.2	Effective sea level and the inverse barometer sea level	7

1.2.3	Further comments on pressure in a homogeneous layer . . . . .	8
1.2.4	Momentum equation . . . . .	8
1.2.5	Thickness equation . . . . .	9
1.2.6	Bottom kinematic boundary condition . . . . .	11
1.2.7	Surface kinematic boundary condition . . . . .	12
1.2.8	Column stretching and the vertical velocity . . . . .	13
1.2.9	Tracer concentration equation . . . . .	14
1.2.10	Summary comments . . . . .	14
<b>1.3</b>	<b>Reduced gravity model for the upper ocean</b> . . . . .	<b>15</b>
1.3.1	Momentum and thickness equations for the active layer . . . . .	15
1.3.2	Relating undulations of the top and bottom layer interfaces . . . . .	17
1.3.3	Momentum equation with reduced gravity . . . . .	17
1.3.4	Reduced gravity and relative buoyancy . . . . .	18
1.3.5	Further study . . . . .	19
<b>1.4</b>	<b>Stacked shallow water equations</b> . . . . .	<b>19</b>
1.4.1	Formulation of a 2-layer model . . . . .	20
1.4.2	$N$ -layer model equations . . . . .	22
1.4.3	Dynamic pressure and the Montgomery potential . . . . .	24
1.4.4	Properties of the vertical velocity . . . . .	24
1.4.5	Rigid lid shallow water models . . . . .	26
1.4.6	Comments on vanishing layers . . . . .	27
<b>1.5</b>	<b>Vector-invariant velocity equation</b> . . . . .	<b>27</b>
1.5.1	Basic manipulations . . . . .	28
1.5.2	Dynamical pressure and the Magnus acceleration . . . . .	28
<b>1.6</b>	<b>Non-conservative processes</b> . . . . .	<b>29</b>
1.6.1	Volume transfer across layer interfaces (dia-surface transport) . . . . .	29
1.6.2	Subgrid scale advective volume tranport within layers . . . . .	31
1.6.3	Tracer equation . . . . .	31
1.6.4	Viscous frictional stresses acting within the layer . . . . .	34
1.6.5	Parameterized interfacial stresses . . . . .	34
1.6.6	Transfer of horizontal momentum from inter-layer volume transfer	35
1.6.7	Inclusion of subgrid along-layer volume transport . . . . .	36
1.6.8	Further study . . . . .	37
<b>1.7</b>	<b>Exercises</b> . . . . .	<b>38</b>

---

## 1.1 Loose threads

- Internal and external mode as per Raymond's notes.

## 1.2 A single shallow water layer

Consider a homogeneous layer of fluid in a uniform effective gravitational field (gravity plus planetary centrifugal,<sup>1</sup> bounded from below by solid walls. If there are no lateral force imbalances, then the fluid remains static. Now perturb the fluid so that it has a nonuniform layer thickness, say with a bump in a particular region. Conservation of fluid mass (which translates into volume conservation for a uniform density layer) means that thicker fluid regions must be exactly compensated by thinner fluid regions. Furthermore, layer thickness gradients create pressure differences (thicker fluid layer has larger hydrostatic pressure than thinner layer),

---

<sup>1</sup>For details of the effective gravity, see the discussion of a particle moving around a rotating planet in VOLUME 1.

which in turn drives fluid motion. If the fluid has much larger lateral extent than vertical, then the lateral motion occurs as an expanding and contracting column with no vertical dependence to the horizontal pressure forces and thus the horizontal motion is vertically independent.

The essence of a perfect fluid (i.e., no irreversible processes such as mixing) shallow water flow concerns the motion of fluid columns accelerated by pressure gradients created by layer thickness undulations, and the associated conservation of mass ensuring that the accumulation of fluid in one region is balanced by the depletion of fluid in another. Pressure gradients act to homogenize the layer thickness. However, planetary rotation and the corresponding Coriolis acceleration allows for layer thickness to be non-constant even in a steady state.

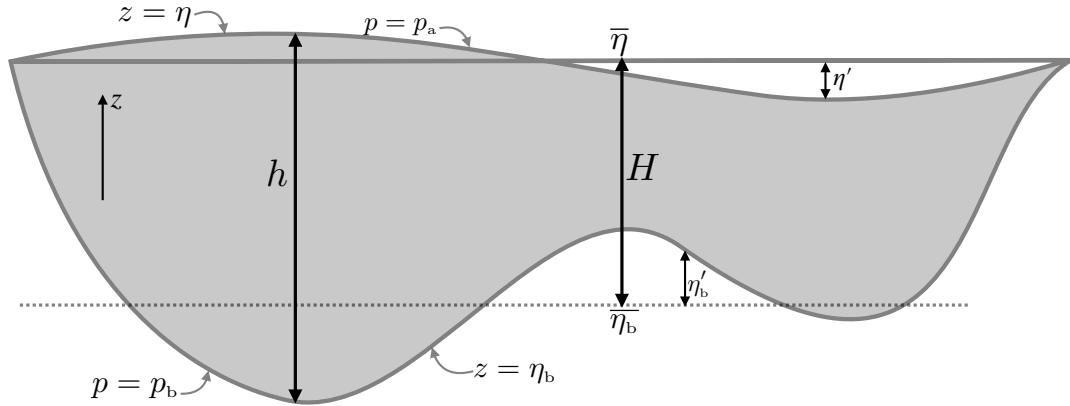


FIGURE 1.1: A single layer of shallow water fluid with thickness  $h = \eta - \eta_b$  that extends from the bottom at  $z = \eta_b$  to the free surface at  $z = \eta$ . The area averaged thickness is defined by  $H = \bar{h} = A^{-1} \int h \, dx \, dy = A^{-1} \int (\eta - \eta_b) \, dx \, dy = \bar{\eta} - \bar{\eta}_b$ , where  $A = \int dx \, dy$  is the horizontal area of the layer. The deviation of the free surface from  $\bar{\eta}$  is given by  $\eta' = \eta - \bar{\eta} = \eta - (\bar{\eta}_b + H)$  so that  $\eta(x, y, t) = \eta_b(x, y) + h(x, y, t) = \bar{\eta}_b + H + \eta'(x, y, t)$ . Likewise, the deviation of the bottom from  $\bar{\eta}_b$  is given by  $\eta'_b = \eta_b - \bar{\eta}_b$ , so that  $h = \eta - \eta_b = H + \eta' - \eta'_b$ . Volume conservation for the layer is maintained in the absence of volume boundary fluxes, in which case  $\bar{\eta}' = 0$ . Note that the position of the reference height,  $z = 0$ , is arbitrary. Atmospheric conventions typically set  $z = 0$  so that  $\bar{\eta}_b = 0$ ,  $\eta = H + \eta'$ , and  $\bar{\eta} = H$ . Oceanic conventions typically choose  $\bar{\eta} = 0$  so that  $\eta = \eta'$  and  $\bar{\eta}_b = -H$ . We are only concerned with fluctuations that leave the free surface monotonic; i.e., we do not consider overturns or breaking waves. This assumption is implied by assuming that each column extending from  $\eta_b \leq z \leq \eta$  maintains hydrostatic balance.

### 1.2.1 Pressure gradient force within the fluid layer

Figure 1.1 shows a single shallow water layer with a generally non-flat bottom and an undulating free surface height. We assume that each column of fluid within the layer is in hydrostatic balance, so that the vertical momentum equation reduces to

$$\frac{\partial p}{\partial z} = -\rho g. \quad (1.1)$$

In our study of hydrostatic pressure in VOLUME 2, we noted that the hydrostatic balance in moving fluid is consistent with lateral flow length scales that are much larger than vertical flow scales (small vertical to horizontal aspect ratio). This scaling is satisfied by large-scale geophysical fluid motion. Hence, a **shallow water approximation** is relevant for studies of fluid motion where horizontal flow scales are very large relative to vertical flow scales.

Since the fluid density is assumed constant (i.e., the fluid is a homogeneous layer), we can integrate the hydrostatic balance from the surface to an arbitrary vertical position within the

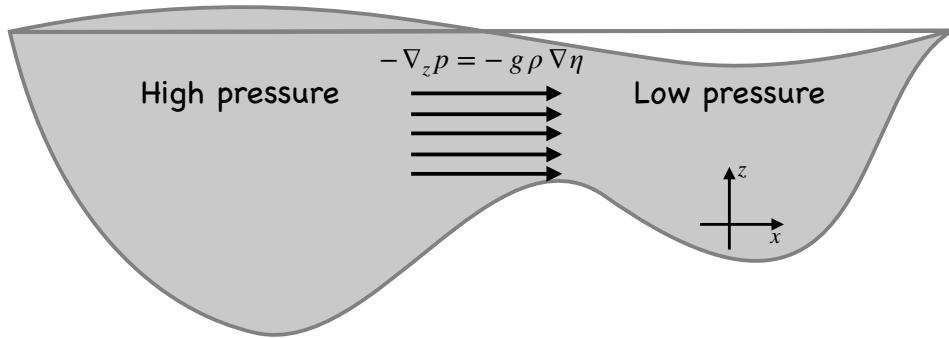


FIGURE 1.2: The horizontal pressure gradient within a shallow water layer is independent of the vertical position within that layer. In this schematic we assume here are zero horizontal gradients in the applied surface pressure,  $\nabla p_a = 0$ , in which case the horizontal acceleration from pressure within a single shallow water fluid layer is determined solely by the surface height,  $-\nabla_h p = -g \rho \nabla_h \eta$ . The acceleration is uniform throughout the layer and points from regions of high free surface height to regions of low free surface height (e.g., sea level highs towards sea level lows). Although bottom topography interacts with the flow and thus affects the shape of the free surface, the topography does not appear explicitly in the horizontal pressure gradient (though the bottom shape does affect bottom pressure form stresses as per the discussion in VOLUME 2 as well as Section 2.4). Instead, we only need to know the shape of the free surface (and the applied pressure  $p_a$ ) to know the horizontal pressure force throughout the layer.

layer

$$p(x, y, z, t) = p_a(x, y, t) + g \rho \int_z^\eta dz = p_a(x, y, t) + g \rho [\eta(x, y, t) - z], \quad (1.2)$$

where  $p_a(x, y, t)$  is the pressure applied to the layer free surface, say from the overlying atmosphere. Furthermore, the horizontal pressure gradient thus takes the form

$$\nabla_h p = \nabla_h p_a + g \rho \nabla_h \eta. \quad (1.3)$$

This pressure gradient is vertically independent within the layer, as depicted in Figure 1.2.

For the left hand side of the pressure gradient in equation (1.3), it is useful to write  $\nabla_h$  since the pressure within a layer is a function of  $z$  (equation (1.2)), whereas for the horizontal momentum equation we only want the horizontal pressure gradient. For the right hand side,  $p_a$  and  $\eta$  are independent of  $z$ , so that there is no need to expose the  $z$  subscript on the gradient operator. We thus drop the subscript when no ambiguity results, as per our convention noted at the start of the chapter.

### The pressure gradient force is vertically independent within a layer

Although hydrostatic pressure within a shallow water layer is vertically dependent as per equation (1.2), the horizontal pressure gradient, as given by equation (1.3), has no vertical dependence within the layer. The acceleration from this hydrostatic pressure force points from highs in the effective sea level to lows in the effective sea level (see Figure 1.2). The vertical independence of the pressure gradient within a shallow water layer holds also for multiple shallow water layers as discussed in Section 1.4, in which case the horizontal velocity has no vertical structure within any layer.

### Pressure gradient force only depends on layer interface gradients

It is notable that the horizontal pressure gradient is solely determined by properties at the upper interface of the layer. That is, we only need to know the shape of the free surface,  $z = \eta$ , and the applied pressure,  $p_a$ , to know the horizontal pressure force acting throughout the layer. There is no explicit dependence on the shape of the bottom topography. The result also holds for the stacked shallow water model discussed in Section 1.4, whereby horizontal pressure gradients are determined by gradients in the layer interfaces and with no appearance of the bottom topography. This characteristic is specific to the use of a pressure gradient body force to describe the role of pressure on the layer momentum. As a complement, we saw in VOLUME 2 how to formulate the pressure contact force. It is through the contact force perspective that we see, in Section 2.4, how bottom topography appears in the momentum balance of a shallow water layer. In particular, this perspective exposes the topographic form stresses that mechanically exchange momentum between the layer and the solid earth bottom.

#### 1.2.2 Effective sea level and the inverse barometer sea level

When considering a nonzero atmospheric pressure, we sometimes find it useful to introduce an effective free surface height

$$\eta^{\text{eff}} = \eta + p_a / (\rho g), \quad (1.4)$$

with gradients in the effective sea level leading to horizontal motion. However, when considering motions on time scales longer than a few days, the ocean free surface adjusts under atmospheric loading towards an [inverse barometer](#) sea level. An inverse barometer sea level compensates the atmospheric pressure so that there is no net horizontal pressure gradient, and so there is no motion induced by the atmospheric pressure. Hence, all that matters for motion are deviations from the inverse barometer sea level.<sup>2</sup>

We expose details to support the above comments by writing pressure within the shallow water layer as

$$p(z) = p_a + \rho g (\eta - z) \quad (1.5a)$$

$$= \bar{p}_a + (p_a - \bar{p}_a) + \rho g (\eta - z) \quad (1.5b)$$

$$= \bar{p}_a + \rho g (\eta - \eta^{\text{ib}} - z). \quad (1.5c)$$

In these equations we introduced the area mean atmospheric pressure,  $\bar{p}_a$ , as well as the inverse barometer sea level

$$\rho g \eta^{\text{ib}} = \bar{p}_a - p_a. \quad (1.6)$$

The inverse barometer sea level is defined according to deviations of the atmospheric pressure from its area mean. Hence, when the atmospheric pressure is higher than the area mean,  $p_a > \bar{p}_a$ , then the inverse barometer sea level is negative,  $\eta^{\text{ib}} < 0$ , reflecting the downward depression of the inverse barometer sea level. In contrast,  $\eta^{\text{ib}} > 0$  for anomalously low pressures,  $p_a < \bar{p}_a$ .

Introducing the inverse barometer sea level brings the horizontal pressure gradient to the form

$$\nabla_h p = \nabla p_a + g \rho \nabla \eta = \rho g \nabla \eta^{\text{eff}} = \rho g \nabla(\eta - \eta^{\text{ib}}). \quad (1.7)$$

---

<sup>2</sup>The discussion in this section follows Appendix C to [Griffies and Greatbatch \(2012\)](#), where further details are provided. In particular, they allow for continuously stratified density rather than the shallow water model considered here. Even so, the key points about the inverse barometer gleaned from the shallow water discussion also hold for the case of a continuously stratified fluid.

If sea level adjusts to the atmospheric pressure so that  $\nabla\eta = \nabla\eta^{\text{ib}}$ , then there is no horizontal pressure gradient in the shallow water layer, in which case there is no induced motion from the atmospheric pressure. It follows that for dynamical purposes, we can seamlessly incorporate atmospheric pressure into the formalism by working with deviations of sea level from the inverse barometer sea level,  $\eta - \eta^{\text{ib}}$ .

### 1.2.3 Further comments on pressure in a homogeneous layer

Vertical independence of the horizontal pressure gradient within the shallow water layer is a direct result of the assumed hydrostatic nature of pressure within the layer. To emphasize this point, we certainly can imagine a homogeneous fluid layer in which the horizontal velocity has a vertical shear. For example, in VOLUME 4 we study surface gravity waves in a homogeneous fluid layer. Such waves have an amplitude that exponentially decays moving downward from the ocean surface, and so the horizontal and vertical fluid motion associated with the waves have a non-zero vertical shear. Such motion cannot be caused by a horizontal gradient in the hydrostatic pressure since this gradient has no vertical dependence throughout the homogeneous layer

$$\partial_z(\nabla_h p_{\text{hydro}}) = 0. \quad (1.8)$$

Hence, in a homogeneous fluid layer, hydrostatic pressure gradients can only drive a horizontal flow that is vertically independent within that layer. So if the vertically sheared horizontal flow is found within a homogeneous fluid layer, and if pressure gradients cause this flow, then it can only be through gradients in the non-hydrostatic pressure. As discussed in our study of surface gravity waves in VOLUME 4, such waves indeed involve non-hydrostatic pressure forces that drive the vertical dependence to the wave amplitude.

Moving beyond the homogeneous layer assumption, we saw in VOLUME 2 that a horizontal gradient in the density leads to a vertically dependent hydrostatic pressure gradient  $\partial_z(\nabla_h p_{\text{hydro}}) \neq 0$ . This hydrostatic pressure force can impart vertical shears to the horizontal flow. Thermal wind shear is the canonical example whereby vertical shears in the horizontal velocity are present in geostrophically balanced fluids as driven by horizontal density gradients (see VOLUME 2). We encounter the shallow water version of thermal wind in Section 2.2.2.

### 1.2.4 Momentum equation

If there is no friction anywhere in the fluid, including at the upper and lower boundaries, then the horizontal momentum is effected only by the Coriolis and pressure forces. Following our discussion of the [Traditional Approximation](#) in VOLUME 2, we retain only the local vertical component to the Coriolis acceleration, which is compatible with the hydrostatic approximation. We are thus led to the horizontal velocity equation

$$\frac{D\mathbf{u}}{Dt} + f\hat{\mathbf{z}} \times \mathbf{u} = -\nabla(g\eta + p_a/\rho), \quad (1.9)$$

where

$$\mathbf{v} = \mathbf{u} + w\hat{\mathbf{z}} \quad (1.10)$$

splits out the horizontal velocity vector,  $\mathbf{u}$ , from the vertical velocity component,  $w$ .

The Coriolis parameter,  $f = f\hat{\mathbf{z}}$ , is vertically independent, as is the horizontal pressure force from gradients in the free surface and applied pressure. Consequently, if the horizontal velocity for the initial flow state is vertically independent, it remains so for all time. The

material time derivative thus only has contributions from the local time derivative and from horizontal advection

$$\frac{D\mathbf{u}}{Dt} = \left[ \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \right] \mathbf{u} \quad (1.11)$$

so that the shallow water velocity equation (1.9) takes on the form

$$(\partial_t + u \partial_x + v \partial_y) \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla(g \eta + p_a/\rho). \quad (1.12)$$

### 1.2.5 Thickness equation

The mass of a shallow water layer is constant in the absence of sources, sinks, or boundary fluxes. Hence, changes in mass at a particular region in the fluid must arise from mass fluxed across the region boundaries, leaving one region and accumulating in another. For simplicity, we assume that no mass crosses the fluid top (the free surface) or the bottom (the solid earth). We consider the more general case of boundary mass transport in Section 1.6.

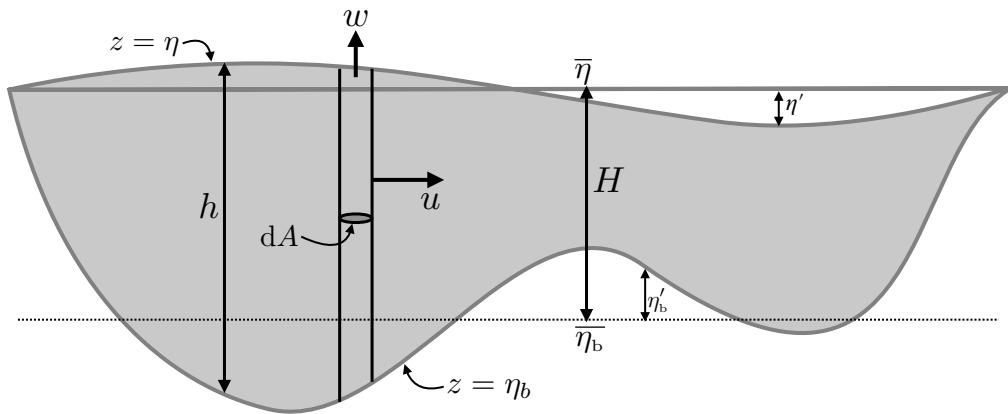


FIGURE 1.3: Mass budget for a column of shallow water fluid with fixed cross-sectional area,  $dA$ , constant density,  $\rho$ , and layer thickness,  $h(x, y, t)$ . The column mass is affected only by horizontal transport (transport within the layer) in the absence of boundary mass fluxes through the top,  $z = \eta$ , or bottom,  $z = \eta_b$ . Note that since the density of the layer is constant, then mass equals to the constant density times the volume.

Consider an infinitesimally thin (in horizontal cross-section) vertical column of shallow water fluid that is fixed in space and extending from  $\eta_b \leq z \leq \eta$ . Let the horizontal cross-sectional area be written as  $dA$  and the thickness be  $h = \eta - \eta_b$  (see Figure 1.3). The total mass of fluid in this column is given by

$$M = \int_{\text{column}} \left[ \int_{\eta_b}^{\eta} \rho dz \right] dA = \rho \int_{\text{column}} (\eta - \eta_b) dA = \rho \int_{\text{column}} h dA. \quad (1.13)$$

Time changes in the column mass thus arise from time changes in the layer thickness integrated over the horizontal area of the column

$$\frac{dM}{dt} = \rho \int_{\text{column}} \frac{\partial h}{\partial t} dA, \quad (1.14)$$

where

$$\frac{\partial h}{\partial t} = \frac{\partial (\eta - \eta_b)}{\partial t} = \frac{\partial \eta}{\partial t}, \quad (1.15)$$

since the bottom topography at  $z = \eta_b(x, y)$  is static.

### General derivation

The mass within a fluid column changes due to mass crossing the column boundaries. Again, we assume here that no mass crosses the top or bottom interfaces. Hence, we only consider mass moving horizontally across the vertical boundaries of the column

$$\text{mass per time entering column} = -\rho \oint_{\text{column}} \mathbf{u} \cdot \hat{\mathbf{n}} dS, \quad (1.16)$$

where  $\hat{\mathbf{n}}$  is the outward normal at the column boundary, and  $dS$  is the area element along the column boundary. The area integral computed over the column boundary involves a vertical integral and a circumferential line integral

$$\text{mass per time entering column} = -\rho \oint_{\text{column}} \left[ \int \mathbf{u} \cdot \hat{\mathbf{n}} dz \right] dl, \quad (1.17)$$

where  $dl$  is the infinitesimal line element around the column circumference. Since  $\hat{\mathbf{n}} \cdot \mathbf{u}$  is vertically independent, we can perform the vertical integral to render

$$-\rho \oint_{\text{column}} \left[ \int \mathbf{u} \cdot \hat{\mathbf{n}} dz \right] dl = -\rho \oint_{\text{column}} h \mathbf{u} \cdot \hat{\mathbf{n}} dl = -\rho \int_{\text{column}} \nabla \cdot (h \mathbf{u}) dA, \quad (1.18)$$

where the second equality follows from the divergence theorem applied to the horizontal cross-sectional area of the column. Equating this result to the mass time tendency (1.14), and noting that the horizontal cross-sectional area is arbitrary, yields an equation for the layer thickness

$$\partial_t h + \nabla \cdot (h \mathbf{u}) = 0. \quad (1.19)$$

This result means that the thickness of fluid at a fixed location increases if there is a convergence of thickness onto that location, and decreases if thickness diverges from the location.

We may also write the thickness equation (1.19) using the material time derivative

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (1.20)$$

Hence, thickness of a material fluid column increases in regions where the horizontal flow converges and it decreases where the horizontal flow diverges. As for the horizontal velocity in equation (1.11), the material time derivative arises from the local time tendency plus horizontal advection

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y}, \quad (1.21)$$

### Special case with a rectangular column

To further our understanding of the second step in equation (1.18), consider the special case of a rectangular column, for which the mass per time of fluid entering the column is given by

mass per time entering column

$$= -\rho \int_{\text{column}} [(u h)_{\text{east}} - (u h)_{\text{west}}] dy - \rho \int_{\text{column}} [(v h)_{\text{north}} - (v h)_{\text{south}}] dx. \quad (1.22)$$

Taking the limit as the horizontal cross-section of the column becomes infinitesimal leads to

mass per time entering column

$$= -\rho \int_{\text{column}} \left[ \frac{\partial(u h)}{\partial x} + \frac{\partial(v h)}{\partial y} \right] dx dy = -\rho \int_{\text{column}} \nabla \cdot (h \mathbf{u}) dA, \quad (1.23)$$

thus recovering the result (1.18).

### 1.2.6 Bottom kinematic boundary condition

Kinematic boundary conditions arise from geometric constraints placed on the fluid system. We consider here the kinematic boundary condition at the bottom interface in the case where there is no flow through this interface, and follow in Section 1.2.7 with the surface kinematic boundary condition.<sup>3</sup>

The ocean bottom is located at a vertical position,  $z = \eta_b(x, y)$ . This location can equivalently be specified mathematically by the surface

$$s(x, y, z) = z - \eta_b(x, y) = 0. \quad (1.24)$$

The outward normal (pointing from the fluid into the rock) at this surface is given by

$$\hat{\mathbf{n}} = -\frac{\nabla s}{|\nabla s|} = \frac{\nabla \eta_b - \hat{\mathbf{z}}}{\sqrt{1 + \nabla \eta_b \cdot \nabla \eta_b}}. \quad (1.25)$$

If the bottom is impenetrable to fluid then the velocity field is constrained to satisfy the no-normal flow boundary condition

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta_b. \quad (1.26)$$

Making use of the bottom outward normal (1.25) leads to

$$w = \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \eta_b. \quad (1.27)$$

For a flat bottom, with  $\nabla \eta_b = 0$ , the no-normal flow condition means that  $w(\eta_b) = 0$ . For the case of nontrivial bottom topography,  $w(\eta_b) = 0$  remains if flow occurs along lines of constant topography; i.e., along isobaths, in which case  $\mathbf{u} \cdot \nabla \eta_b = 0$ . But more generally, sloping bottoms lead to a nonzero vertical velocity component. Dynamically, a nonzero bottom vertical velocity arises from forces at the bottom that cause the horizontal velocity to cross isobaths,  $\mathbf{u} \cdot \nabla \eta_b \neq 0$ .

The kinematic result (1.27) is written in an Eulerian sense, with the velocity constrained to satisfy this relation at each point along the bottom interface. It has a dual material interpretation based on acknowledging that the bottom interface is a material surface. A fluid element on the bottom at  $s = z - \eta_b = 0$  will thus remain there; it does not cross the bottom interface. Rather, it can at most move tangentially to the bottom.<sup>4</sup> We can ensure the

<sup>3</sup>From the discussion of fluid kinematics in VOLUME 2, we use the term **material surface** for any continuous surface or interface that is impenetrable to the flow of matter or thermal energy (mechanical energy can be transferred via pressure forces). In VOLUME 2, we derive the kinematic boundary conditions for a fluid at interfaces. We here apply those ideas to the shallow water system, so that the presentation in Sections 1.2.6 and 1.2.7 offer a review of the kinematic boundary conditions from VOLUME 2 as applied to a shallow water layer.

<sup>4</sup>Details of the tangential motion along a material boundary require dynamical information such as boundary stresses (see VOLUME 2). We are not concerned with dynamical information here, rather our concern is solely with kinematics.

no-normal flow constraint by setting

$$\frac{Ds}{Dt} = \frac{D(z - \eta_b)}{Dt} = 0 \quad \text{at } z = \eta_b. \quad (1.28)$$

Rearrangement of this result leads to the Eulerian constraint (1.27). Equivalently, we can write this boundary condition in the form

$$w = \frac{D\eta_b}{Dt} \quad \text{at } z = \eta_b. \quad (1.29)$$

Since  $\eta_b = \eta_b(x, y)$ , this expression of the kinematic boundary condition is identical to equation (1.27).

### 1.2.7 Surface kinematic boundary condition

We here assume the surface boundary is a material interface and thus derive the surface kinematic boundary condition. In Section 1.6 we consider the slightly more general case of volume crossing this surface. As a material surface, the surface kinematic boundary condition follows analogously to the bottom. However, there is a fundamentally new feature in that the layer's upper free surface is a time dependent moving boundary. We study such boundaries in VOLUME 2 when detailing the kinematic boundary conditions for a material surface. We here review some of that discussion.

The free surface is located at a vertical position  $z = \eta(x, y, t)$ . Equivalently, the free surface can be specified by a surface of constant  $s$ , where

$$s(x, y, z, t) = z - \eta(x, y, t) = 0. \quad (1.30)$$

The outward normal to the free surface is thus given by

$$\hat{\mathbf{n}} = \frac{\nabla s}{|\nabla s|} = \frac{\hat{\mathbf{z}} - \nabla \eta}{\sqrt{1 + \nabla \eta \cdot \nabla \eta}}. \quad (1.31)$$

We must account for motion of the surface when formulating the no-normal flow condition. To do so, write the no-normal flow condition as

$$(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta(x, y, t), \quad (1.32)$$

where  $\mathbf{v}^{(s)}$  is the velocity of a point on the ocean surface. The velocity of a point fixed on an arbitrary surface with specified  $s$  satisfies

$$\frac{\partial s}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla s = 0. \quad (1.33)$$

As defined,  $\mathbf{v}^{(s)}$  advects a fluid element in a manner to always keep the element fixed on the constant  $s$  surface. With  $\hat{\mathbf{n}} = \nabla s / |\nabla s|$ , we have

$$\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = -\frac{\partial_t s}{|\nabla s|} = \frac{\partial_t \eta}{\sqrt{1 + \nabla \eta \cdot \nabla \eta}}. \quad (1.34)$$

Hence, if the surface remains static, then  $\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = 0$ . But more generally, the surface is moving, and that movement is fundamental to the surface kinematic boundary condition.

Making use of the result (1.34) in the no-normal flow constraint (1.32) then leads to the surface kinematic boundary condition

$$\partial_t \eta = w - \mathbf{u} \cdot \nabla \eta \quad \text{at } z = \eta. \quad (1.35)$$

As for the bottom kinematic boundary condition written as (1.28), we can interpret the surface kinematic condition (1.35) materially, in which case

$$\frac{Ds}{Dt} = \frac{D(z - \eta)}{Dt} = 0 \quad \text{at } z = \eta. \quad (1.36)$$

That is, in the absence of flow across the surface boundary, that surface remains material. We can write this boundary condition in the equivalent form

$$w = \frac{D\eta}{Dt} = (\partial_t + \mathbf{u} \cdot \nabla) \eta \quad \text{at } z = \eta. \quad (1.37)$$

### 1.2.8 Column stretching and the vertical velocity

Since the fluid has constant density, we know that the velocity has zero three-dimensional divergence

$$\nabla \cdot \mathbf{u} + \partial_z w = 0 \implies \partial_z w = -\nabla \cdot \mathbf{u}. \quad (1.38)$$

This result also follows since material fluid elements in the constant density shallow water layer maintain a constant volume (see the fluid kinematics studied in VOLUME 2). Furthermore, since the horizontal velocity has no vertical dependence, we can vertically integrate the continuity equation (1.38) from the bottom to an arbitrary vertical position within the layer to render

$$w(z) = w(\eta_b) - (z - \eta_b) \nabla \cdot \mathbf{u}, \quad (1.39)$$

so that the vertical velocity is a linear function of the vertical position within a shallow water layer. Applying this equation at the upper surface,  $z = \eta(x, y, t)$ , yields

$$w(\eta) = w(\eta_b) - (\eta - \eta_b) \nabla \cdot \mathbf{u}. \quad (1.40)$$

Eliminating the horizontal convergence between equations (1.39) and (1.40) leads to

$$w(z) - w(\eta_b) = \left[ \frac{z - \eta_b}{\eta - \eta_b} \right] [w(\eta) - w(\eta_b)]. \quad (1.41)$$

Making use of the surface kinematic boundary condition (1.37) and bottom kinematic boundary condition (1.29) renders the material form

$$\frac{1}{z - \eta_b} \left[ \frac{D(z - \eta_b)}{Dt} \right] = \frac{1}{\eta - \eta_b} \left[ \frac{D(\eta - \eta_b)}{Dt} \right]. \quad (1.42)$$

Finally, introducing the layer thickness  $h = \eta - \eta_b$  yields the material conservation law

$$\frac{D}{Dt} \left[ \frac{z - \eta_b}{h} \right] = 0. \quad (1.43)$$

Again,  $h = \eta - \eta_b$  is the layer thickness and  $z - \eta_b$  is the height of a fluid element from the bottom interface (see Figure 1.1). Consequently, equation (1.43) means that the ratio of

the fluid element height above the bottom to the layer thickness remains constant as the fluid element moves through the shallow water fluid. That is, a vertical column of shallow water fluid stretches or squeezes coherently within a shallow water layer, so that the relative vertical position remains fixed for a point within the column. Shallow water mechanics thus comprises the mechanics of coherent vertical fluid columns moving within a fluid layer. This constrained behavior results from the linear  $z$  dependence of the vertical velocity, which itself is a result of the vertical independence of the horizontal velocity, and which follows from the vertical independence of the horizontal gradient of hydrostatic pressure.

### 1.2.9 Tracer concentration equation

Suppose there is a material substance, a tracer, contained within a shallow water layer.<sup>5</sup> We expect the tracer concentration,  $\psi$ , to have a non-uniform vertical structure within the layer, in addition to having horizontal structure:  $\psi = \psi(x, y, z, t)$ . If the tracer is advected through the layer without any diffusion, then the concentration satisfies the perfect fluid tracer equation (i.e., the [advection equation](#))

$$\partial_t \psi + \mathbf{u} \cdot \nabla_h \psi + w \partial_z \psi = 0. \quad (1.44)$$

For a shallow water layer, where the horizontal velocity has no vertical dependence within a layer, we find it sufficient to study the vertically averaged tracer concentration within the layer,

$$C(x, y, t) \equiv \frac{1}{h} \int_{\eta_b}^{\eta} \psi(x, y, z, t) dz. \quad (1.45)$$

To develop the evolution equation for  $C$ , we vertically integrate the tracer equation (1.44) over the layer and make use of Leibniz's rule from [VOLUME 2](#)

$$\int_{\eta_b}^{\eta} \frac{\partial \psi}{\partial t} dz = \partial_t(h C) - \psi(\eta) \partial_t \eta \quad (1.46a)$$

$$\int_{\eta_b}^{\eta} \mathbf{u} \cdot \nabla \psi dz = \nabla \cdot (h C \mathbf{u}) - \psi(\eta) \mathbf{u}(\eta) \cdot \nabla \eta + \psi(\eta_b) \mathbf{u}(\eta_b) \cdot \nabla \eta_b - h C \nabla \cdot \mathbf{u} \quad (1.46b)$$

$$\int_{\eta_b}^{\eta} w \partial_z \psi dz = w(\eta) \psi(\eta) - w(\eta_b) \psi(\eta_b) - h C \partial_z w. \quad (1.46c)$$

For these equations we made use of the  $z$  independence of  $\mathbf{u}$  and  $\partial_z w$  within the layer. Use of the kinematic boundary conditions from [Sections 1.2.6](#) and [1.2.7](#), and the three dimensional continuity equation,  $\nabla \cdot \mathbf{u} + \partial_z w = 0$ , renders the equation for the layer vertically averaged tracer concentration

$$\frac{\partial(h C)}{\partial t} + \nabla \cdot (h C \mathbf{u}) = 0 \iff \frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C = \frac{DC}{Dt} = 0. \quad (1.47)$$

As a self-consistency check, note that if the tracer concentration has a horizontally uniform value, then the flux-form of the tracer equation (1.47) reduces to the thickness equation (1.19).

### 1.2.10 Summary comments

#### Key physical assumptions for the shallow water fluid

The shallow water fluid model is based on the following two assumptions.

---

<sup>5</sup>We develop the theory of material tracers in [VOLUME 2](#).

- The fluid layer has a uniform density, which then means the fluid is incompressible.
- The pressure is hydrostatic.

Since the pressure gradient is vertically independent within the layer, an initial horizontal velocity that is vertically independent will remain vertically independent. That is, the fluid moves as columns within the layer, with the columns stretching and squashing depending on the divergence or convergence of volume towards the column.

The term “shallow” refers to the small vertical to horizontal aspect ratio,  $H/L \ll 1$ , which in turn is consistent with the hydrostatic approximation studied in VOLUME 2. The term “water” refers to the incompressible nature of each shallow water fluid layers, which is a physically more relevant approximation for the ocean than for the atmosphere (see our discussion of the Boussinesq ocean in VOLUME 2). Nonetheless, the shallow water model has direct applications to many features of both the atmosphere and ocean circulation, and as such it is widely used across the atmosphere and ocean sciences.

### Shallow water fluid columns are not Taylor columns

The columnar motion of fluid within a shallow water layer is reminiscent of the [Taylor column](#) encountered in our study of geostrophy in VOLUME 2. However, the columnar motion of fluid within a Taylor column holds for homogeneous fluids undergoing rapid rotation. Furthermore, the horizontal fluid velocity within a Taylor column is non-divergent so that there is no vertical motion of the fluid. These properties allow one to interpret a Taylor column as a fluid mechanical realization of a column of rigid matter much like a solid body.

In contrast, shallow water fluid columns do not rely on rotation, but instead arise from the hydrostatic balance (small aspect ratio flow) maintained within each homogeneous layer. Additionally, shallow water columns are not rigid. Rather, they stretch and squash in the presence of nonzero divergence in the horizontal flow, thus leading to vertical motion of fluid within the column. Finally, shallow water columns remain coherent even as they move over topography, and yet, again, they can stretch and compress. In VOLUME 4, we further discuss the connections and distinctions between vertically coherent flow present in the small aspect ratio shallow water layer versus that for the rapidly rotating flow leading to the [Taylor-Proudman effect](#).

## 1.3 Reduced gravity model for the upper ocean

The [reduced gravity model](#) describes a dynamically active layer of uniform density,  $\rho_1$ , above a stagnant layer of density,  $\rho_2$ , and below a fluid of zero density,  $\rho_0 = 0$ . It is often referred to as the 1.5 layer model. This theoretical model has been used, to some success, as an idealization of the upper ocean circulation whereby an active layer (e.g., the region above the ocean pycnocline), sits above an inactive layer (the abyss) of much smaller motion (here assumed to be zero motion). In this way, we introduce the [level of no motion](#), below which (baroclinic) currents vanish.

### 1.3.1 Momentum and thickness equations for the active layer

We develop the momentum equations for the [reduced gravity model](#) by making use of the hydrostatic balance, in which pressure at a vertical position,  $z$ , in the upper layer is computed

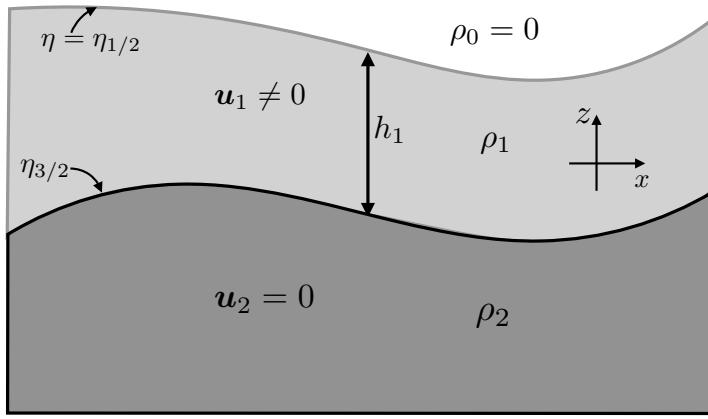


FIGURE 1.4: Reduced gravity model of shallow water fluid. The lower layer with density  $\rho_2$  is dynamically inactive and thus has a zero velocity. The upper layer is dynamically active with thickness,  $h = \eta_{1/2} - \eta_{3/2}$ , and density,  $\rho_1$ . The dynamically active layer is bounded above by a zero density atmosphere,  $\rho_0 = 0$ . The lower inactive layer is assumed to be infinitely deep so that its continuity equation can be ignored; i.e., even though there are zero currents within this layer, the layer thickness can still undulate. The reduced gravity between the two layers is defined by  $g'_{3/2} = g(\rho_2 - \rho_1)/\rho_{\text{ref}} \ll g$ , whereas the reduced gravity at the top interface is given by  $g'_{1/2} = g$ .

as (see Figure 1.4)

$$p_1(x, y, z, t) = p_{1/2}(x, y, t) + g \rho_1 [\eta_{1/2}(x, y, t) - z], \quad (1.48)$$

where we denote an interface value by a half-index, so that  $\eta_{1/2}$  and  $p_{1/2}$  are the interface height and pressure at the upper layer interface.<sup>6</sup> Since the fluid above the upper layer is assumed to have zero density, we set

$$p_{1/2} = 0. \quad (1.49)$$

The horizontal momentum equation for the upper (active) layer is given by

$$\rho_{\text{ref}} \left[ \frac{D\mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 \right] = -g \rho_1 \nabla \eta_{1/2}, \quad (1.50)$$

where the  $z$  dependent term in the pressure (1.48) drops out when computing the horizontal pressure gradient. Setting the reference density equal to the top layer density,  $\rho_{\text{ref}} = \rho_1$ , leads to the more tidy equation

$$\frac{D\mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 = -g \nabla \eta_{1/2}. \quad (1.51)$$

The equations for the upper layer are completed by volume conservation in the form of the thickness equation

$$\frac{Dh_1}{Dt} = -h_1 \nabla \cdot \mathbf{u}_1. \quad (1.52)$$

<sup>6</sup>There is no general consensus on this notation, with some treatments, such as [Vallis \(2017\)](#) and [Cushman-Roisin and Beckers \(2011\)](#) using an integer to label an interface quantity, whereas some numerical model papers (e.g., [Griffies et al. \(2020\)](#)) use the half-index. We prefer the half-index since it removes any ambiguity concerning the ordering of the interfaces relative to the layer.

### 1.3.2 Relating undulations of the top and bottom layer interfaces

The pressure in the lower stagnant layer is given by the weight per horizontal area of fluid above it, and it can be written

$$p_2(x, y, z, t) = g \rho_1 (\eta_{1/2} - \eta_{3/2}) + g \rho_2 (\eta_{3/2} - z) \quad (1.53a)$$

$$= g \rho_1 \eta_{1/2} + g (\rho_2 - \rho_1) \eta_{3/2} - g \rho_2 z \quad (1.53b)$$

$$= \rho_{\text{ref}} (g_{1/2}^r \eta_{1/2} + g_{3/2}^r \eta_{3/2}) - g \rho_2 z. \quad (1.53c)$$

In this equation we introduced the reference density and reduced gravity

$$\rho_{\text{ref}} = \rho_1 \quad \text{and} \quad g_{1/2}^r = g \quad \text{and} \quad g_{3/2}^r = g (\rho_2 - \rho_1) / \rho_{\text{ref}} \ll g. \quad (1.54)$$

Taking the reference density as the top layer density is common for Boussinesq shallow water models, and will be assumed in our formulations with multiple layers in Section 1.4. We employ the half-index notation for the reduced gravity since it is computed by differencing the densities between two adjacent layers, and as such reduced gravities are considered an interface property. Additionally, the reduced gravities multiply a corresponding interface gradient, which also uses the half-integer notation.

For the reduced gravity model we assume the lower layer is motionless. To maintain zero motion in the lower layer requires the lower layer horizontal pressure gradient to vanish

$$\rho_{\text{ref}}^{-1} \nabla_h p_2 = g_{1/2}^r \nabla \eta_{1/2} + g_{3/2}^r \nabla \eta_{3/2} = 0. \quad (1.55)$$

This constraint links the undulations of the top and bottom interfaces of the dynamically active layer

$$\nabla \eta_{1/2} = -\frac{g_{3/2}^r}{g_{1/2}^r} \nabla \eta_{3/2} = -[(\rho_2 - \rho_1) / \rho_{\text{ref}}] \nabla \eta_{3/2} \implies \eta_{1/2} = -\eta_{3/2} [(\rho_2 - \rho_1) / \rho_{\text{ref}}] + \text{constant}. \quad (1.56)$$

The density ratio,  $(\rho_2 - \rho_1) / \rho_{\text{ref}}$ , is positive but typically much smaller than unity. Hence, the relation (1.56) means that undulations of the free surface,  $\eta_{1/2}$ , are of opposite sign to and of much smaller amplitude than the undulations in the lower interface,  $\eta_{3/2}$ . This behavior is typical for undulations of the pycnocline region of the ocean and the free surface as depicted in Figure 1.5.

### 1.3.3 Momentum equation with reduced gravity

Relation (1.55) can be used to write the momentum equation (1.51) in the form

$$\frac{D\mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 = g_{3/2}^r \nabla \eta_{3/2}. \quad (1.57)$$

It is common to make use of the momentum equation in the form (1.57), rather than the original form (1.51). The reason is that historically, ocean hydrography measurements<sup>7</sup> have allowed for an estimate of the pycnocline slope,  $\nabla \eta_{3/2}$ , whereas it was not until precise satellite altimetry measurements starting in the 1990s that we could estimate the far smaller sea level slope,  $\nabla \eta_{1/2}$ .

<sup>7</sup>In oceanography, hydrography refers to measurements of temperature, salinity, and pressure; see [Talley et al. \(2011\)](#).

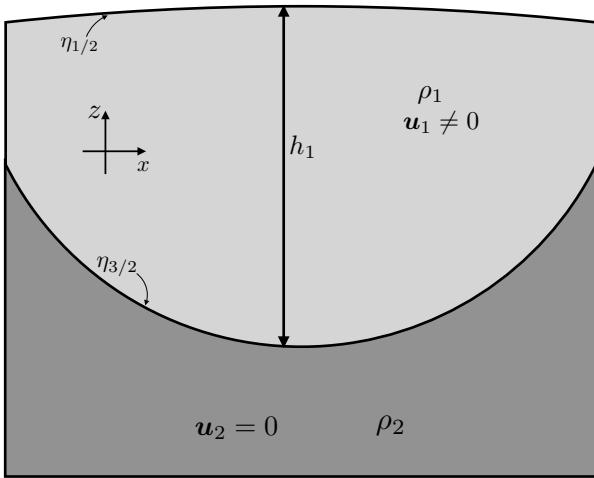


FIGURE 1.5: A vertical slice through a reduced gravity, or 1.5 layer, ocean in hydrostatic balance. Shown here is a plug of dynamically active light water, as may occur in a warm core eddy to the subtropical gyres, sitting on top of heavy water of zero motion. The free surface corresponds to  $\eta_{1/2}$  in Figure 1.4, whereas the pycnocline (heavy black line) corresponds to the lower interface  $\eta_{3/2}$  of Figure 1.4. The sea surface experiences an applied pressure  $p = p_s$ , assumed to be uniform for this idealized situation. Note how sea level is a maximum above the pycnocline minimum, with this geometry reflected in equation (1.56). In the ocean, the slope of the pycnocline is about 100-300 times larger than the slope of the sea level. That is, sea level may show undulations on the order of a metre, whereas the pycnocline undulations are on the order of 100-300 m.

### 1.3.4 Reduced gravity and relative buoyancy

Equation (1.55)

$$g \nabla \eta_{1/2} = -g_{3/2}^r \nabla \eta_{3/2}, \quad (1.58)$$

says that with  $g_{3/2}^r \ll g$ , the free surface slopes are much smaller than interior slopes

$$|\nabla \eta_{1/2}| \ll |\nabla \eta_{3/2}|. \quad (1.59)$$

We thus infer that the interior interface has less resistance to vertical motion than the free surface. To help understand this result, recall the study of [Archimedean buoyancy](#) from VOLUME 2. We see that the reduced gravity,  $g_{3/2}^r$ , is the Archimedean buoyancy of layer 1 relative to layer 2, with normalization given by the reference density as per the Boussinesq approximation

$$\text{buoyancy layer 1 relative to layer 2} = -g(\rho_1 - \rho_2)/\rho_{\text{ref}} = g_{3/2}^r. \quad (1.60)$$

A small relative buoyancy between two density layers renders little resistance for motion of the layer interface. Indeed, as the density difference vanishes, so too does the buoyant resistance to motion of the layer interface. For the upper free surface, the buoyancy of layer 0 (a zero mass atmosphere) relative to layer 1 equals to the full gravitational acceleration.

$$\text{buoyancy layer 0 relative to layer 1} = -g(\rho_0 - \rho_1)/\rho_{\text{ref}} = g, \quad (1.61)$$

where we assumed that  $\rho_{\text{ref}} = \rho_1$  and  $\rho_0 = 0$ . Evidently, the relative buoyancy between the top shallow water layer and the atmosphere is given by the full gravity acceleration, which indicates the large buoyant resistance to vertical motion of the free surface.

It is for this reason that the interior interface is more flexible than the free surface, as depicted in Figure 1.5. Even for an atmosphere with mass, so that  $\rho_0 > 0$ , the upper interface's reduced gravity is close to  $g$  since the atmosphere is roughly 1000 times less dense than seawater.

This result holds in general, whereby increasing the reduced gravity between density layers, and thus increasing the vertical density stratification, increases the resistance to motion of the layer interface and thus reduces the interface's flexibility.

### 1.3.5 Further study

The material in this section is inspired by Section 3.2 of [Vallis \(2017\)](#). [Tomczak and Godfrey \(1994\)](#) make use of the reduced gravity model for interpreting aspects of the observed ocean. Additional use is made by [Griffies et al. \(2014\)](#) for interpreting large-scale sea level patterns.

## 1.4 Stacked shallow water equations

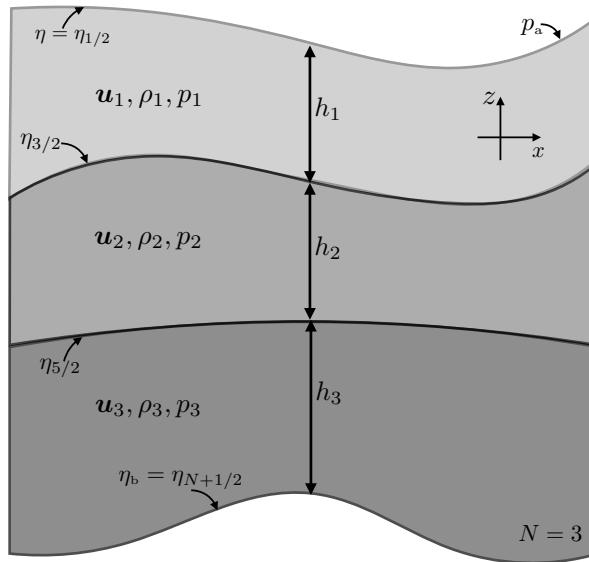


FIGURE 1.6: Three dynamically active layers of stacked shallow water fluid ( $N = 3$ ). The notation corresponds to that for the reduced gravity model of Figure 1.4, yet here with three dynamically active layers. In particular,  $\eta_{1/2}$  is the free surface,  $\eta_{1/2} = \eta$ , whereas  $\eta_{N+1/2} = \eta_b$  is the bottom interface. Hence, the total thickness of a column is  $h_1 + h_2 + h_3 = \eta_{1/2} - \eta_b$ . The “atmosphere” above the layers is assumed to apply a pressure,  $p_a$ , to the upper surface. The horizontal velocity is vertically independent within each layer, as is the horizontal pressure gradient, so that both are discontinuous across a layer interface. In contrast, the pressure is a linear function of  $z$  within a layer and is continuous across a layer interface. Finally, the vertical velocity is also a linear function of  $z$  within a layer but it is discontinuous across a layer interface. The reduced gravity defined between each layer is given by  $g_{k+1/2}^r = g(\rho_{k+1} - \rho_k)/\rho_{ref}$ . We take the reference density as  $\rho_{ref} = \rho_1$ , which results in a tidy set of layer equations.

In studies of shallow water fluids, much of the formalism developed for a single layer can be readily extended to an arbitrary number of layers. We here pursue this extension and thereby expose the underlying kinematics and dynamics of stacked shallow water models. We assume the layers are immiscible so that matter and thermal properties are not exchanged between the layers. Consequently, the layers couple only through mechanical forces arising from the pressure [form stress](#) studied in VOLUME 2. Furthermore, we continue to assume that the horizontal velocity has no vertical dependence within each shallow water layer, which in turn means the horizontal pressure gradient is vertically independent within each layer. The notation for our derivations is depicted in Figure 1.6 in the case of three active layers.

In Chapter 14, we develop the equations for a continuously stratified Boussinesq fluid making use of isopycnal vertical coordinates. Although the vertical stratification is continuous in that case, we find that the isopycnal equations are isomorphic to the stacked shallow water equations. Hence, besides being of intrinsic interest as a versatile theoretical model, the stacked shallow water model offers a suitable step towards studies of a continuously stratified fluid using isopycnal coordinates.

### 1.4.1 Formulation of a 2-layer model

We here display the equations for two layers, thus offering the seeds for an extension to  $N$  layers in Section 1.4.2.

#### Thickness and tracer equations

Each shallow water layer satisfies its own independent thickness equation and tracer equation, representing the conservation of volume and tracer content for each layer

$$\frac{\partial h_1}{\partial t} + \nabla \cdot (h_1 \mathbf{u}_1) = 0 \quad (1.62a)$$

$$\frac{\partial h_2}{\partial t} + \nabla \cdot (h_2 \mathbf{u}_2) = 0 \quad (1.62b)$$

$$\frac{\partial(h_1 C_1)}{\partial t} + \nabla \cdot (h_1 C_1 \mathbf{u}_1) = 0 \quad (1.62c)$$

$$\frac{\partial(h_2 C_2)}{\partial t} + \nabla \cdot (h_2 C_2 \mathbf{u}_2) = 0. \quad (1.62d)$$

We emphasize that there is no explicit coupling between these equations, as each layer separately must satisfy volume conservation and tracer conservation. However, the velocities are coupled through the pressure force, as we now discuss.

#### Pressure within a layer

To compute the pressure within a layer, we make use of the hydrostatic balance and integrate down from the surface, which results in the pressure fields

$$p_1 = \rho_1 g (\eta_{1/2} - z) + p_a \quad (1.63)$$

$$p_2 = \rho_1 g (\eta_{1/2} - \eta_{3/2}) + \rho_2 g (\eta_{3/2} - z) + p_a. \quad (1.64)$$

As for the reduced gravity model in equation (1.53c), it is convenient to write pressure in layer-two using the reduced gravity, which leads to

$$p_2 - p_a = \rho_1 g (\eta_{1/2} - \eta_{3/2}) + \rho_2 g (\eta_{3/2} - z) \quad (1.65a)$$

$$= g \eta_{1/2} \rho_1 + g (\rho_2 - \rho_1) \eta_{3/2} - g \rho_2 z \quad (1.65b)$$

$$= \rho_{ref} (g'_{1/2} \eta_{1/2} + g'_{3/2} \eta_{3/2}) - g \rho_2 z \quad (1.65c)$$

with the Boussinesq reference density and reduced gravities given by

$$\rho_{ref} = \rho_1 \quad \text{and} \quad g'_{1/2} = g \quad \text{and} \quad g'_{3/2} = g(\rho_2 - \rho_1)/\rho_{ref} \ll g. \quad (1.66)$$

As for the reduced gravity model in equation (1.54), we set the reference density equal to the top layer density. As per the discussion around equation (1.76),  $\rho_{ref} = \rho_1$  results in a somewhat

tidier set of layer velocity equations.

### Boussinesq reference density and the reduced gravity

We are formulating the shallow water model according to Boussinesq ocean equations from VOLUME 2, whereby fluid elements conserve their volume rather than their mass. According to our discussion of the Boussinesq momentum equation, again in VOLUME 2, density is set to a constant reference density when measuring the inertial mass of a fluid element, yet density remains the *in situ* density when measuring the weight of a fluid element. For the shallow water model, we multiply the inertial acceleration and Coriolis acceleration with a reference density,

$$\rho_{\text{ref}} = \text{shallow water reference density}, \quad (1.67)$$

whereas pressure and potential energy maintain the density,  $\rho_k$ , of the shallow water layer.

We further make use of the Boussinesq ocean when computing the buoyancy frequency. Namely, with  $\varrho$  the potential density, the Boussinesq form of the squared buoyancy frequency is

$$N^2 = -g \left[ \frac{1}{\rho_{\text{ref}}} \frac{\partial \varrho}{\partial z} \right] \approx -\frac{g}{\rho_{\text{ref}}} \frac{\Delta \rho}{\Delta z} = -g^r / \Delta z. \quad (1.68)$$

It is for this reason that we take  $\rho_{\text{ref}}$  for the denominator of the reduced gravity throughout all of the shallow water layers, thus defining

$$g_{k-1/2}^r = g (\rho_k - \rho_{k-1}) / \rho_{\text{ref}} \quad \text{and} \quad g_{1/2}^r = g. \quad (1.69)$$

### Horizontal velocity equations

The horizontal velocity equations for the two layers take the form

$$\rho_{\text{ref}} \left[ \frac{D_1 \mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 \right] = -\nabla p_1 \quad (1.70a)$$

$$\rho_{\text{ref}} \left[ \frac{D_2 \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_2 \right] = -\nabla p_2, \quad (1.70b)$$

where we introduced the layer material time derivatives

$$\frac{D_k}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}_k \cdot \nabla \quad \text{for } k = 1, 2. \quad (1.71)$$

Making use of expressions (1.63) and (1.65c) for layer pressures leads to the horizontal momentum equations

$$\rho_{\text{ref}} \left[ \frac{D_1 \mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 \right] = -g \rho_1 \nabla [\eta_{1/2} + p_a / (g \rho_1)] \quad (1.72a)$$

$$\rho_{\text{ref}} \left[ \frac{D_2 \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_2 \right] = -g \rho_1 \nabla [\eta_{1/2} + p_a / (g \rho_1)] - g_{3/2}^r \rho_{\text{ref}} \nabla \eta_{3/2}. \quad (1.72b)$$

It is convenient to express the interface heights in terms of layer thicknesses,  $h_1$  and  $h_2$ , since the layer thicknesses are the prognostic fields determined by time stepping the thickness equations (1.62a) and (1.62b). We thus write

$$\eta_{1/2} = \eta_b + h_1 + h_2 \quad \text{and} \quad \eta_{3/2} = \eta_b + h_2, \quad (1.73)$$

so that

$$p_1 - p_a = \rho_1 g (\eta_b + h_1 + h_2) - g \rho_1 z \quad (1.74a)$$

$$p_2 - p_a = \rho_1 g (\eta_b + h_1 + h_2) + \rho_{ref} g_{3/2}^r (\eta_b + h_2) - g \rho_2 z, \quad (1.74b)$$

thus resulting in the horizontal momentum equations

$$\rho_{ref} \left[ \frac{D_1 \mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 \right] = -\rho_1 g \nabla [\eta_b + h_1 + h_2 + p_a / (g \rho_1)] \quad (1.75a)$$

$$\rho_{ref} \left[ \frac{D_2 \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_2 \right] = -\rho_1 g \nabla [\eta_b + h_1 + h_2 + p_a / (g \rho_1)] - \rho_{ref} g_{3/2}^r \nabla (\eta_b + h_2). \quad (1.75b)$$

Notice how layer thickness from one layer appears in the other layer's pressure gradient. In this way, changes in the thickness of one layer have a direct impact on pressure forces and flow in the adjacent layer. Also notice how the bottom topography appears in the bottom pressure gradient, which arises due to our switch from layer interfaces to layer thicknesses.

As already noted, it is common for a stacked shallow water model to choose

$$\rho_{ref} = \rho_1. \quad (1.76)$$

We here see why taking this choice is desirable, since doing so brings the layer velocity equations (1.75a) and (1.75b) into the more tidy forms

$$\frac{D_1 \mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 = -g \nabla [\eta_b + h_1 + h_2 + p_a / (g \rho_1)] \quad (1.77a)$$

$$\frac{D_2 \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_2 = -g \nabla [\eta_b + h_1 + h_2 + p_a / (g \rho_1)] - g_{3/2}^r \nabla (\eta_b + h_2). \quad (1.77b)$$

### Vertical shear in horizontal velocities between layers

The difference in layer velocities,  $\mathbf{u}_1 - \mathbf{u}_2$ , represents the vertical shear in the layers. This difference is affected by a pressure gradient arising just from bottom topography and the interior layer thickness,  $h_2$

$$\frac{D_1 \mathbf{u}_1}{Dt} - \frac{D_2 \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times (\mathbf{u}_1 - \mathbf{u}_2) = g_{3/2}^r \nabla (\eta_b + h_2). \quad (1.78)$$

That is, the vertical shear does not directly feel undulations of the free surface,  $\eta_{1/2}$ , or the applied pressure,  $p_a$ . Rather, it feels these surface undulations only indirectly via nonlinear terms appearing in the advection on the left hand side. We further discuss this result in Section 2.2.2 by introducing thermal wind and the Margules' Relation.

## 1.4.2 *N*-layer model equations

The 2-layer equations from Section 1.4.1 can be generalized to  $N$ -layers. The thickness equation and tracer equation represent volume and tracer conservation for each layer

$$\frac{\partial h_k}{\partial t} + \nabla \cdot (\mathbf{u}_k h_k) = 0 \quad (1.79a)$$

$$\frac{\partial (h_k C_k)}{\partial t} + \nabla \cdot (\mathbf{u}_k h_k C_k) = 0, \quad (1.79b)$$

where  $\kappa = 1, N$  is the discrete layer index and there is no implied summation on this label.<sup>8</sup>

### Expressions for the pressure

Some work is needed to generalize the pressure gradient appearing in the velocity equation. For that purpose, write the interface height as

$$\eta_{\kappa+1/2} = \eta_b + \sum_{j=\kappa+1}^N h_j \quad \text{with } \eta_{N+1/2} = \eta_b \text{ and } \eta_{1/2} = \eta. \quad (1.80)$$

For example, the layer interfaces with  $N = 3$  layers are given by

$$\eta_{1/2} = \eta_b + h_1 + h_2 + h_3 \quad \eta_{3/2} = \eta_b + h_2 + h_3 \quad \eta_{5/2} = \eta_b + h_3 \quad \eta_{7/2} = \eta_b. \quad (1.81)$$

In turn, the hydrostatic pressure within layer- $\kappa$  is given by

$$p_\kappa = -g \rho_\kappa z + p_a + \rho_{ref} \sum_{j=0}^{\kappa-1} g_{j+1/2}^r \eta_{j+1/2} = p_{\kappa-1/2} + g \rho_\kappa (\eta_{\kappa-1/2} - z), \quad (1.82)$$

where the reduced gravities are defined according to equation (1.69)

$$g_{j-1/2}^r = g (\rho_j - \rho_{j-1}) / \rho_{ref} > 0 \quad \text{with} \quad g_{1/2}^r = g. \quad (1.83)$$

As an example, the layer pressures for  $N = 3$  are given by

$$p_1 = p_a + g \rho_1 (\eta_{1/2} - z) \quad p_2 = p_{3/2} + g \rho_2 (\eta_{3/2} - z) \quad p_3 = p_{5/2} + g \rho_3 (\eta_{5/2} - z). \quad (1.84)$$

The half-integer pressures are evaluated on the corresponding interface, and the hydrostatic balance yields the pressure difference

$$p_{\kappa+1/2} - p_{\kappa-1/2} = g \rho_\kappa h_\kappa = g \rho_\kappa (\eta_{\kappa-1/2} - \eta_{\kappa+1/2}). \quad (1.85)$$

### Summary of the thickness weighted velocity equation

The velocity equation for an arbitrary layer- $\kappa$  is given by

$$\rho_{ref} \left[ \frac{D_\kappa \mathbf{u}_\kappa}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_\kappa \right] = -\nabla p_\kappa \iff [\partial_t + (\mathbf{u}_\kappa \cdot \nabla)] \mathbf{u}_\kappa + f \hat{\mathbf{z}} \times \mathbf{u}_\kappa = -\rho_{ref}^{-1} \nabla p_\kappa. \quad (1.86)$$

Use of the layer thickness equation (1.79a) readily leads to the thickness-weighted momentum equation

$$\partial_t (h_\kappa \mathbf{u}_\kappa) + \nabla \cdot (h_\kappa \mathbf{u}_\kappa \otimes \mathbf{u}_\kappa) + f \hat{\mathbf{z}} \times h_\kappa \mathbf{u}_\kappa = -(h_\kappa / \rho_{ref}) \nabla p_\kappa. \quad (1.87)$$

We study the thickness-weighted momentum equation in Section 2.3, where we find it offers a more suitable framework than the velocity equation for studying the pressure form stresses acting at layer interfaces.

<sup>8</sup>To help distinguish the layer index,  $\kappa$ , from a tensor index,  $k$ , we write the layer index using an upright Roman font whereas a tensor index is slanted.

### 1.4.3 Dynamic pressure and the Montgomery potential

The term  $g \rho_k z$  appearing in the layer pressures (1.82) has a zero horizontal gradient. Hence, it does not contribute to the horizontal pressure gradient acceleration acting on a layer. As a result, it is a “do nothing” pressure that motivates us to introduce the shallow water **dynamical pressure** (e.g., equation (3.44) of [Vallis \(2017\)](#))

$$p_k + g \rho_k z = p_a + \rho_{\text{ref}} \sum_{j=0}^{k-1} g_{j+1/2}^r \eta_{j+1/2} \equiv p_a + p_k^{\text{dyn}} \implies \nabla p_k = \nabla p_a + \nabla p_k^{\text{dyn}}, \quad (1.88)$$

The term “dynamical” is motivated since  $p_k^{\text{dyn}}$  gradients directly lead to accelerations. Notably, the dynamical pressure is vertically independent within a layer, so that it has a jump moving across a layer interface. Also note that  $p_k^{\text{dyn}}$  is related to the shallow water version of the **Montgomery potential** via

$$M_k^{\text{dyn}} = \rho_{\text{ref}}^{-1} p_k^{\text{dyn}} = \sum_{j=0}^{k-1} g_{j+1/2}^r \eta_{j+1/2} \quad (1.89)$$

(e.g., Section 12.2 of [Cushman-Roisin and Beckers \(2011\)](#)), with the continuous Montgomery potential defined as part of our discussion of isopycnal models in Chapter 14. When studying the mechanical energy budget in Section 2.6, we find that  $p_k^{\text{dyn}}$  naturally appears since the energy budget is derived using the pressure gradient body force. The difference of the dynamic pressure between two layers is given by

$$p_k^{\text{dyn}} - p_{k+1}^{\text{dyn}} = -\rho_{\text{ref}} g_{k+1/2}^r \eta_{k+1/2} \iff M_k^{\text{dyn}} - M_{k+1}^{\text{dyn}} = -g_{k+1/2}^r \eta_{k+1/2}, \quad (1.90)$$

which is an expression of the hydrostatic balance for shallow water layers.

We often find it useful to study pressure contributions to the momentum equation via the duality between the pressure gradient body force and the pressure contact force, as when studying pressure **form stress** in VOLUME 2, and further applied to the shallow water in Section 2.4 and when studying isopycnal models in Chapter 14. For those purposes, we do not use  $p_k^{\text{dyn}}$  since it is not the hydrostatic pressure measured within a shallow water layer, so that  $p_k^{\text{dyn}}$  cannot be directly used to convert the pressure gradient body force to the pressure contact force. Instead, we must use the pressure,  $p_k$ .

### 1.4.4 Properties of the vertical velocity

From the kinematics of VOLUME 2, we know that

$$w = \frac{D\eta}{Dt} \quad (1.91)$$

provides an expression for the vertical velocity of a fluid parcel at a point on an impermeable surface,  $z = \eta(x, y, t)$ . We made use of this equation in Section (1.2.7) when studying the surface kinematic boundary condition for a single shallow water layer. We make use of this equation in this section to further an understanding of the vertical velocity in a shallow water model. In particular, we find that the jump in horizontal velocity across the interface leads to a jump in the vertical velocity.

**Jump condition for the vertical velocity across a layer interface**

Consider two shallow water layers labelled by  $k$  and  $k+1$  that are separated by an interface at  $z = \eta_{k+1/2}(x, y, t)$ . Define the vertical velocity at the  $z = \eta_{k+1/2}$  interface according to the material kinematic condition

$$w^{(k)}(\eta_{k+1/2}) = w_{k+1/2}^{(k)} = (\partial_t + \mathbf{u}_k \cdot \nabla) \eta_{k+1/2}. \quad (1.92)$$

Evidently,  $w_{k+1/2}^{(k)}$  is the vertical velocity for a fluid parcel on the top side of the  $z = \eta_{k+1/2}$  interface. Similarly, define the vertical velocity for a fluid parcel on the lower side of the  $z = \eta_{k+1/2}$  interface according to

$$w_{k+1/2}^{(k+1)} = (\partial_t + \mathbf{u}_{k+1} \cdot \nabla) \eta_{k+1/2}. \quad (1.93)$$

Taking the difference leads to the jump condition

$$w^{(k)}(\eta_{k+1/2}) - w^{(k+1)}(\eta_{k+1/2}) = (\mathbf{u}_k - \mathbf{u}_{k+1}) \cdot \nabla \eta_{k+1/2} \neq 0. \quad (1.94)$$

As advertised, the vertical velocity has a jump across the interface that arises from the jump in the horizontal velocity. In Section 2.2.2 we derive the shallow water expression (2.9) for thermal wind balance maintained by geostrophic flow, known as the Margules' relation. In that special case there is no jump in the vertical velocity since  $\hat{\mathbf{z}} \times (\mathbf{u}_k - \mathbf{u}_{k+1})$  is parallel to  $\nabla \eta_{k+1/2}$ , so that  $(\mathbf{u}_k - \mathbf{u}_{k+1}) \cdot \nabla \eta_{k+1/2} = 0$ . We thus conclude that the vertical velocity jump arises from ageostrophic contributions to the horizontal flow.

It is tempting to introduce a sub-region that smoothly connects the horizontal velocity between the layers, thus removing the jump condition for both  $\mathbf{u}$  and  $w$ . Doing so enhances the realism of the stacked layers, since a physical realization of fluid layers will have a finite sized region that interpolates between the layer properties. However, adding this transition zone moves us beyond the stacked shallow water model, and so requires analysis that falls outside shallow water theory. Hence, we do not pursue that avenue. Instead, in this chapter we remain within the shallow water theory while acknowledging it has limitations, including jumps in the horizontal velocity across layer interfaces.

**Preservation of the relative vertical position within a column**

Following the approach from Section 1.2.8, we compute the vertical velocity within a layer by vertically integrating the non-divergence condition,  $\nabla \cdot \mathbf{v} = 0$ . For layer  $k$  we have

$$w^{(k)}(z) = w_{k+1/2}^{(k)} - (z - \eta_{k+1/2}) \nabla \cdot \mathbf{u}_k, \quad (1.95)$$

where  $w^{(k)}(z)$  is the vertical velocity at a vertical position,  $z$ , that is located within layer  $k$  so that  $\eta_{k+1/2} \leq z \leq \eta_{k-1/2}$ . Equation (1.95) generalizes the single-layer equation (1.39), thus revealing that the vertical velocity within an arbitrary shallow water layer is a linear function of the vertical position. Now evaluate equation (1.95) at the upper interface for the layer,  $z = \eta_{k-1/2}$ , in which case

$$w_{k-1/2}^{(k)} = w_{k+1/2}^{(k)} - h_k \nabla \cdot \mathbf{u}_k, \quad (1.96)$$

where

$$h_k = \eta_{k-1/2} - \eta_{k+1/2} \quad (1.97)$$

is the layer thickness. Eliminating the horizontal convergence between equations (1.95) and (1.96) renders

$$\frac{w_{k-1/2}^{(k)} - w_{k+1/2}^{(k)}}{h_k} = \frac{w^{(k)}(z) - w_{k+1/2}^{(k)}}{z - \eta_{k+1/2}}. \quad (1.98)$$

Making use of the kinematic boundary condition at the two interfaces allows us to write

$$w_{k-1/2}^{(k)} = \frac{D^{(k)}\eta_{k-1/2}}{Dt} = (\partial_t + \mathbf{u}_k \cdot \nabla) \eta_{k-1/2} \quad (1.99a)$$

$$w_{k+1/2}^{(k)} = \frac{D^{(k)}\eta_{k+1/2}}{Dt} = (\partial_t + \mathbf{u}_k \cdot \nabla) \eta_{k+1/2}, \quad (1.99b)$$

which then brings equation (1.98) to the form

$$\frac{1}{h_k} \frac{D^{(k)}h_k}{Dt} = \frac{1}{z - \eta_{k+1/2}} \frac{D^{(k)}(z - \eta_{k+1/2})}{Dt} \implies \frac{D^{(k)}}{Dt} \left[ \frac{z - \eta_{k+1/2}}{h_k} \right] = 0 \quad (1.100)$$

Just like for a single layer of shallow water fluid, vertical columns within each layer are stretched and squeezed in a manner that preserves the relative vertical position for a point within the column.

### Averaged vertical velocity for a layer

Since the vertical velocity is a linear function of  $z$  within a layer, its layer averaged value is the average of the vertical velocity at the upper and lower layer interfaces. It is a useful exercise of the formalism to prove this result

$$\bar{w}_k = \frac{1}{h_k} \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} w^{(k)}(z) dz \quad (1.101a)$$

$$= \frac{1}{h_k} \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} \left[ w_{k+1/2}^{(k)} - (z - \eta_{k+1/2}) \nabla \cdot \mathbf{u}_k \right] dz \quad (1.101b)$$

$$= w_{k+1/2}^{(k)} + \eta_{k+1/2} \nabla \cdot \mathbf{u}_k - 1/(2h_k) (\eta_{k-1/2}^2 - \eta_{k+1/2}^2) \nabla \cdot \mathbf{u}_k \quad (1.101c)$$

$$= w_{k+1/2}^{(k)} - (h_k/2) \nabla \cdot \mathbf{u}_k \quad (1.101d)$$

$$= (w_{k+1/2}^{(k)} + w_{k-1/2}^{(k)})/2, \quad (1.101e)$$

where the final step made use of the identity (1.96). Note that we can also write the layer averaged vertical velocity as

$$\bar{w}_k = (w_{k+1/2}^{(k)} + w_{k-1/2}^{(k)})/2 = \frac{D^{(k)}\bar{\eta}_k}{Dt}, \quad (1.102)$$

where

$$\bar{\eta}_k = (\eta_{k+1/2} + \eta_{k-1/2})/2 \quad (1.103)$$

is the average of the interface height for the layer.

### 1.4.5 Rigid lid shallow water models

Throughout this section we formulated the equations for an N-layer shallow water model where  $\eta_{1/2}$  is the undulating free surface. For some applications of large-scale oceanography, it is

useful to remove the external gravity waves from the model, where these gravity waves are associated with linear fluctuations of the upper free surface.<sup>9</sup> To remove these gravity waves we can, by fiat, set the upper ocean interface to a rigid constant, conventionally taken as  $\eta_{1/2} = 0$ . This assumption is known as the [rigid lid approximation](#). There is good justification for this approximation given that undulations of the free surface are far smaller than undulations of interior interfaces, as seen for the reduced gravity model in Section 1.3.

If there is no applied pressure,  $p_a = 0$ , then there is no horizontal pressure gradient in the upper layer, much like an inverted reduced gravity model (see Exercise 1.5). A dynamically more interesting case arises when there is an applied pressure,  $p_a \neq 0$ , so that motion is generated in layer one, which in turn induces motion throughout all layers. Indeed, as seen when studying the horizontally non-divergent barotropic model in Chapter 4, there must be a nonzero applied surface pressure in order to constrain the free surface to be flat and fixed. This pressure is referred to as the [lid pressure](#).

A rigid lid constrains the vertically integrated flow to be horizontally non-divergent. We see this property by summing the thickness equation over all of the shallow water layers

$$\partial_t \eta_{1/2} + \sum_{k=1}^N \nabla \cdot (\mathbf{u}_k h_k) = 0 \quad \text{where} \quad \sum_{k=1}^N \partial_t h_k = \partial_t (\eta_{1/2} - \eta_b) = \partial_t \eta_{1/2}. \quad (1.104)$$

For the rigid lid,  $\partial_t \eta_{1/2} = 0$ , which then leads to the non-divergence condition

$$\sum_{k=1}^N \nabla \cdot (\mathbf{u}_k h_k) = \nabla \cdot \sum_{k=1}^N \mathbf{u}_k h_k = \nabla \cdot \mathbf{U} = 0 \quad \text{with} \quad \mathbf{U} = \sum_{k=1}^N \mathbf{u}_k h_k. \quad (1.105)$$

Note that velocity in the upper layer remains horizontally divergent since the  $\eta_{3/2}$  interface is not generally rigid. Even so, fixing  $\eta_{1/2}$  to be rigid serves to remove the relatively fast external gravity waves from the stacked shallow water model. We return to the rigid lid assumption when discussing the horizontal non-divergent barotropic model in Chapter 4.

#### 1.4.6 Comments on vanishing layers

Interior layer interfaces can intersect either the surface, referred to as an [outcrop](#), or the solid bottom, referred to as an [incrop](#). In this manner, any particular layer may only exist over a sub-region of the full horizontal domain. Treatment of such geometries is actually quite subtle when numerically discretizing the equations of motion, since it requires methods to allow for layers to vanish and appear as a function of time. The thickness weighted equations developed in this section offer one means to handle these situations. They do so by ensuring that terms properly vanish as  $h \rightarrow 0$  (see the end of Section 1.6.1 for more on this point). We also considered some of the related conceptual subtleties in Vtwo when studying [available potential energy](#) in a continuously stratified Boussinesq ocean. [Griffies et al. \(2020\)](#) present a primer on finite volume numerical methods that support vanishing layers.

## 1.5 Vector-invariant velocity equation

We here derive the [vector invariant](#) form of the shallow water velocity equation. This formulation proves useful in the shallow water kinetic energy budget of Section 2.6.3, as well as for the

---

<sup>9</sup>In VOLUME 4, we study gravity waves in a single shallow water layer.

shallow water vorticity in Chapter 5.

### 1.5.1 Basic manipulations

The following manipulations hold for each shallow water layer, so it is convenient to drop layer indices to reduce clutter. Start by introducing vorticity for the full velocity field

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} \quad (1.106a)$$

$$= \hat{\mathbf{x}} (\partial_y w - \partial_z v) + \hat{\mathbf{y}} (\partial_z u - \partial_x w) + \hat{\mathbf{z}} (\partial_x v - \partial_y u) \quad (1.106b)$$

$$= \hat{\mathbf{x}} \partial_y w - \hat{\mathbf{y}} \partial_x w + \hat{\mathbf{z}} (\partial_x v - \partial_y u) \quad (1.106c)$$

$$= -\hat{\mathbf{z}} \times \nabla w + \hat{\mathbf{z}} (\partial_x v - \partial_y u), \quad (1.106d)$$

where we set

$$\partial_z u = \partial_z v = 0, \quad (1.107)$$

which holds for the horizontal velocity within a shallow water layer. We also find it convenient to introduce the vorticity associated with the horizontal velocity field

$$\boldsymbol{\omega}^* = \nabla \times \mathbf{u} = -\hat{\mathbf{x}} \partial_z v + \hat{\mathbf{y}} \partial_z u + \hat{\mathbf{z}} (\partial_x v - \partial_y u) = \hat{\mathbf{z}} (\partial_x v - \partial_y u). \quad (1.108)$$

The vertical component to vorticity is particularly important for geophysical flows, in which case we write

$$\zeta = \hat{\mathbf{z}} \cdot \boldsymbol{\omega} = \hat{\mathbf{z}} \cdot \boldsymbol{\omega}^* = \partial_x v - \partial_y u. \quad (1.109)$$

The vector identity from VOLUME 1

$$(\mathbf{u} \cdot \nabla) \mathbf{u} = (1/2) \nabla(\mathbf{u} \cdot \mathbf{u}) - \mathbf{u} \times (\nabla \times \mathbf{u}) = (1/2) \nabla(\mathbf{u} \cdot \mathbf{u}) + \boldsymbol{\omega}^* \times \mathbf{u} \quad (1.110)$$

brings the inviscid shallow water velocity equation (1.86)

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\rho_{\text{ref}}^{-1} \nabla p \quad (1.111)$$

into its vector invariant form

$$\frac{\partial \mathbf{u}}{\partial t} + (f + \zeta) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla(p/\rho_{\text{ref}} + \mathbf{u} \cdot \mathbf{u}/2). \quad (1.112)$$

Again, this equation holds separately for each layer, so that by reintroducing layer indices we have

$$\frac{\partial \mathbf{u}_k}{\partial t} + (f + \zeta_k) \hat{\mathbf{z}} \times \mathbf{u}_k = -\nabla(p_k/\rho_{\text{ref}} + \mathbf{u}_k \cdot \mathbf{u}_k/2). \quad (1.113)$$

### 1.5.2 Dynamical pressure and the Magnus acceleration

As in our discussion in VOLUME 2 of the vector invariant velocity equation for continuously stratified flows, the velocity equation (1.113) exposes two physical processes that lend insight into the motion in the shallow water fluid.

#### Dynamical pressure from kinetic energy per mass

The kinetic energy per mass,  $\mathbf{u} \cdot \mathbf{u}/2$ , adds the dynamic pressure to the mechanical pressure associated with the free surface undulations. Gradients in the kinetic energy thus drive

accelerations towards regions of smaller kinetic energy; i.e., down the kinetic energy gradient.<sup>10</sup>

### Magnus acceleration

As discussed in VOLUME 2, the **Magnus acceleration** is a body acceleration defined by the nonlinear term

$$\mathbf{A}_{\text{magnus}} = -\boldsymbol{\omega}^* \times \mathbf{u} = \zeta (\mathbf{u} \times \hat{\mathbf{z}}), \quad (1.114)$$

appearing in the vector-invariant velocity equation (1.113). There is a non-zero Magnus acceleration when a shallow water fluid column spins while it moves, with this acceleration acting to deflect the spinning column perpendicular to its trajectory. Consider a shallow water fluid column moving zonally, in which case the Magnus acceleration is

$$\mathbf{A}_{\text{magnus}} = u \zeta (\hat{\mathbf{x}} \times \hat{\mathbf{z}}) = -\hat{\mathbf{y}} u \zeta. \quad (1.115)$$

With a positive relative vorticity,  $\zeta > 0$ , the Magnus acceleration is directed to the right of the motion, which is in the same direction as the Coriolis acceleration in the northern hemisphere. For large-scale geophysical flows, the Magnus acceleration is sub-dominant to the Coriolis acceleration. However, the Magnus acceleration is a crucial facet of large (order unity or larger) Rossby number motions in which relative vorticity is sizable.

## 1.6 Non-conservative processes

Much of the focus on shallow water mechanics in this book concerns the inviscid and adiabatic case in which shallow water layers are immiscible and there are no frictional or boundary stresses within or between layers. Even so, we find occasion to consider the transfer of volume between layers via irreversible dia-surface fluxes, as well as momentum transfer associated with viscous friction and boundary stresses, and tracer transfer via diffusion. Such non-conservative exchanges are central to many applications of shallow water models, thus prompting us to formulate the stacked shallow water equations that include such processes. The discussion offers a vertically discrete version of the continuous isopycnal model primitive equations detailed in Chapter 14.

### 1.6.1 Volume transfer across layer interfaces (dia-surface transport)

Consider the case of dia-surface transfer as occurs across the ocean surface through evaporation, precipitation, and river runoff, or as occurs for interior layers in the presence of irreversible mixing processes. Let  $w^{(\eta)}$  measure the volume per time per horizontal area of fluid crossing the surface interface of the shallow water layer, with  $w^{(\eta)}$  having dimensions of length per time. As shown in Figure 1.7, we choose a sign convention so that  $w_{k-1/2}^{(\eta)} > 0$  means volume leaves shallow layer  $k$  through its upper interface, whereas  $w_{k+1/2}^{(\eta)} > 0$  means that volume enters the layer through its lower interface.

#### Surface kinematic boundary condition for a single shallow water layer

The **kinematic boundary condition** (1.36) expresses the material nature of the free surface in the absence of boundary volume flux. In the presence of a non-zero boundary flux,  $w^{(\eta)} \neq 0$ ,

---

<sup>10</sup>The dynamic pressure afforded by the kinetic energy should not be confused with the shallow water dynamical pressure introduced by equation (1.88). They are distinct.

we follow the formulation in VOLUME 2 of the kinematic boundary condition for the ocean free surface with a mass flux, which renders the boundary condition

$$\frac{D(z - \eta)}{Dt} = w^{(\dot{\eta})} \quad \text{at } z = \eta. \quad (1.116)$$

In effect, this relation defines the surface boundary flux, and this boundary condition can be written in the equivalent form

$$\frac{D\eta}{Dt} = w(\eta) - w^{(\dot{\eta})}. \quad (1.117)$$

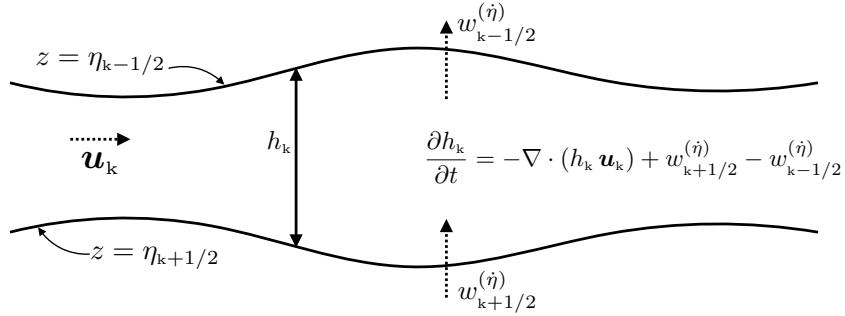


FIGURE 1.7: Dia-surface transport across the upper interface,  $w_{k-1/2}^{(\dot{\eta})}$ , and lower interface,  $w_{k+1/2}^{(\dot{\eta})}$ , of shallow water layer  $k$ . The sign convention is that  $w_{k+1/2}^{(\dot{\eta})} > 0$  for fluid entering layer  $k$  through its lower interface, and  $w_{k-1/2}^{(\dot{\eta})} > 0$  for fluid leaving layer  $k$  through its upper interface. The evolution of layer thickness at a horizontal position is given by equation (1.120), with a thickness time tendency due to the convergence of horizontal thickness transport within a layer, and the net transport of fluid moving across the layer interfaces.

### Layer thickness equation for a single shallow water layer

Following through the derivation of the thickness equation in Section 1.2.5, we are trivially led to the following generalization in the presence of volume moving through the upper interface

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = -w^{(\dot{\eta})}, \quad (1.118)$$

or equivalently

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} - w^{(\dot{\eta})}. \quad (1.119)$$

As volume leaves the layer through its upper interface,  $w^{(\dot{\eta})} > 0$ , equation (1.118) says that the layer thickness decreases, and conversely the layer thickness increases when volume is added via  $w^{(\dot{\eta})} < 0$ .

### Thickness equation for a stacked shallow water model

It is straightforward to generalize the thickness equation (1.118) to the case of a stacked shallow water model, in which

$$\frac{\partial h_k}{\partial t} + \nabla \cdot (h_k \mathbf{u}_k) = -(w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}). \quad (1.120)$$

In this way, a positive volume flux coming through the lower layer interface,  $w_{k+1/2}^{(\dot{\eta})} > 0$ , leads to an increase in  $h_k$ , whereas a positive flux leaving through the upper interface,  $w_{k-1/2}^{(\dot{\eta})} >$

0, decreases the layer thickness. The right hand side is written as the thickness weighted convergence of the cross-layer transport. Finally, the flux-form equation (1.120) takes on its material evolution form

$$\frac{D_k h_k}{Dt} = -h_k \nabla \cdot \mathbf{u}_k - (w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}). \quad (1.121)$$

Evidently, the material evolution of layer thickness is affected by the convergence of the horizontal flow as well as the transfer of volume across the layer interfaces.

### Vanishing layers and transfer across the top and bottom of the stacked layers

A shallow water layer can vanish at any horizontal position, so that it need not be defined for all horizontal positions in the domain. Such vanishing layers are elegantly handled by working with the thickness weighted tracer and momentum budgets, as doing so avoids dividing by a zero layer thickness when computing vertical transfer across layer interfaces. It also means that the transfer of material and momentum across the top and bottom of a stacked shallow water model occurs via the layer adjacent to these boundaries. That is, boundary fluxes are incorporated via cross-layer fluxes entering or leaving the layer that sits at the surface layer or bottom.

### 1.6.2 Subgrid scale advective volume tranport within layers

In the presence of subgrid scale processes, there can be an additional advective transport of volume within a layer beyond that arising from the resolved flow field,  $\mathbf{u}$ . Such subgrid scale transport from a velocity,  $\mathbf{u}_k^{sgs}$ , is commonly found when studying the effects of turbulent eddies on ocean and atmospheric flows, with the papers by [Gent and McWilliams \(1990\)](#) and [Gent et al. \(1995\)](#) offering the canonical ocean example (see VOLUME 4 for more on subgrid scale eddy parameterizations). For present purposes, we are not concerned with how to compute this subgrid flow. Rather, we are concerned with how to account for this flow in a manner that conserves volume, tracer, and momentum. For the thickness equation the extra transport simply represents a term added to the resolved flow, so that the thickness equation takes on the form

$$\partial_t h_k + \nabla \cdot [(h_k (\mathbf{u}_k + \mathbf{u}_k^{sgs})] = -(w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}), \quad (1.122)$$

where  $\mathbf{u}_k^{sgs}$  is the subgrid scale flow. Volume for the layer is conserved in the presence of  $\mathbf{u}_k^{sgs}$  so long as  $\mathbf{u}_k^{sgs}$  satisfies the same no-flow kinematic boundary conditions as  $\mathbf{u}$ . Bringing the subgrid scale contribution to the right hand side renders

$$\partial_t h_k + \nabla \cdot (h_k \mathbf{u}_k) = -(w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}) - \nabla \cdot (h_k \mathbf{u}_k^{sgs}). \quad (1.123)$$

As shown in the following sections, the presence of both the dia-surface transport and the subgrid scale along-layer transport lead to corresponding terms in the tracer and momentum equations, with such terms required for conservation of tracer and momentum.

### 1.6.3 Tracer equation

Tracer content is transferred into or out of layer  $k$  in the presence of a nonzero volume flux,  $w_{k\pm 1/2}^{(\dot{\eta})} \neq 0$ . Likewise, tracer is transferred within a layer due to advection by  $\mathbf{u}_k + \mathbf{u}_k^{sgs}$ . To formulate the tracer budget admitting such transfer, we work with the thickness weighted tracer

equation so to work with an extensive quantity. Namely, the tracer content per horizontal area,  $h_k C_k$ , which satisfies a relatively straightforward budget.

### Tracer equation with cross-layer transport and subgrid advection

We ensure self-consistency between the tracer equation and thickness equation (1.120) by writing the tracer budget, in the presence of cross-layer transfer and subgrid scale along-layer advection, as

$$\partial_t(h_k C_k) + \nabla \cdot (h_k C_k \mathbf{u}_k^{\text{eff}}) = -[(w^{(\dot{\eta})} C)_{k-1/2} - (w^{(\dot{\eta})} C)_{k+1/2}], \quad (1.124)$$

where we introduced the shorthand for the effective flow

$$\mathbf{u}_k^{\text{eff}} = \mathbf{u}_k + \mathbf{u}_k^{\text{sgs}}. \quad (1.125)$$

The tracer fluxes on the right hand side,  $(w^{(\dot{\eta})} C)_{k\pm 1/2}$ , are evaluated at the layer interfaces. Evidently, a positive flux at the lower interface,  $(w^{(\dot{\eta})} C)_{k+1/2} > 0$ , contributes to an increase in the tracer content for layer  $k$ , whereas  $(w^{(\dot{\eta})} C)_{k-1/2} > 0$  signals tracer leaving through the upper interface.

To estimate the tracer concentration at the layer interface, we make use of the **upwind advective flux**. For the flux  $(w^{(\dot{\eta})} C)_{k+1/2}$ , the upwind method uses  $C_{k+1}$  if  $w_{k+1/2}^{(\dot{\eta})} > 0$ , whereas the flux uses  $C_k$  if  $w_{k+1/2}^{(\dot{\eta})} < 0$

$$(w^{(\dot{\eta})} C)_{k+1/2} = w_{k+1/2}^{(\dot{\eta})} \begin{cases} C_{k+1} & \text{if } w_{k+1/2}^{(\dot{\eta})} > 0 \\ C_k & \text{if } w_{k+1/2}^{(\dot{\eta})} < 0, \end{cases} \quad (1.126)$$

with the same approach used for other interfaces. This method ensures that integrated tracer content is conserved for the full domain, since the tracer content that enters one layer leaves an adjacent layer.

### Consistency of the thickness weighted tracer equation and tracer concentration equation

As a consistency check, note that the thickness weighted tracer equation (1.124) reduces to the thickness equation (1.122) upon setting the tracer concentration to a space-time constant. As a second consistency check, use the product rule in equation (1.124) to have

$$C_k [\partial_t h_k + \nabla \cdot (h_k \mathbf{u}_k^{\text{eff}})] + h_k (\partial_t C_k + \mathbf{u}_k^{\text{eff}} \cdot \nabla C_k) = -[(w^{(\dot{\eta})} C)_{k-1/2} - (w^{(\dot{\eta})} C)_{k+1/2}]. \quad (1.127)$$

Use of the thickness equation (1.122) and rearrangement renders the advective form tracer concentration equation

$$h_k (\partial_t + \mathbf{u}_k^{\text{eff}} \cdot \nabla) C_k = -[[w^{(\dot{\eta})} C)_{k-1/2} - C_k w_{k-1/2}^{(\dot{\eta})}] - [(w^{(\dot{\eta})} C)_{k+1/2} - C_k w_{k+1/2}^{(\dot{\eta})}]. \quad (1.128)$$

The tracer concentration equation (1.128) is self-consistent since both sides vanish when  $C_k$  is a space-time constant.

The particular case of  $w_{k-1/2}^{(\dot{\eta})} > 0$  and  $w_{k+1/2}^{(\dot{\eta})} > 0$ , and use of upwind tracer fluxes (1.126), leads to the tracer concentration equation

$$h_k (\partial_t + \mathbf{u}_k^{\text{eff}} \cdot \nabla) C_k = -w_{k+1/2}^{(\dot{\eta})} (C_k - C_{k+1}). \quad (1.129)$$

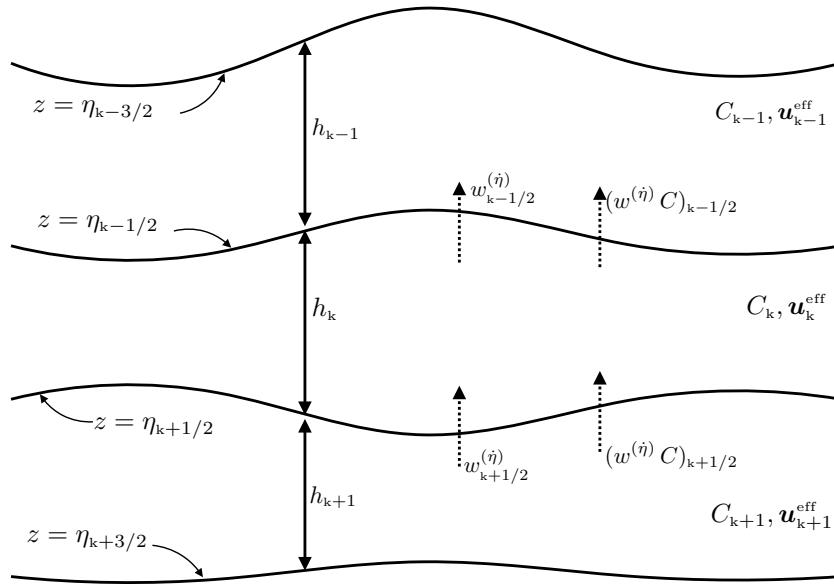


FIGURE 1.8: Depicting the tracer budget within a shallow water layer according to the thickness weighted tracer equation (1.124) and in the presence of dia-layer transport as in Figure 1.7. The cross-layer tracer fluxes,  $(w^{(\dot{\eta})} C)_{k\pm 1/2}$ , are computed according to an upwind method (1.126).

Evidently, again with  $w_{k+1/2}^{(\dot{\eta})} > 0$ , the tracer concentration in layer  $k$  increases if  $C_{k+1} > C_k$  and decreases if  $C_{k+1} < C_k$ . This is the expected behavior, thus further supporting self-consistency of both the thickness weighted tracer equation and the tracer concentration equation.

### Including tracer diffusion

The addition of a cross-layer diffusive flux, with a kinematic diffusivity  $\kappa_{cr} > 0$ , and an along-layer diffusive flux, with a kinematic diffusivity  $\kappa_{al} > 0$ , leads to the shallow water thickness-weighted advection-diffusion equation

$$\begin{aligned} \partial_t(h_k C_k) + \nabla \cdot [h_k (C_k \mathbf{u}_k^{\text{eff}} - \kappa_{al} \nabla C_k)] \\ = -[(w^{(\dot{\eta})} C - \kappa_{cr} \partial_z C)_{k-1/2} - (w^{(\dot{\eta})} C - \kappa_{cr} \partial_z C)_{k+1/2}]. \quad (1.130) \end{aligned}$$

In this equation,  $(-\kappa_{cr} \partial_z C)_{k\pm 1/2}$  is the downgradient cross-layer diffusive tracer flux evaluated at the  $k \pm 1/2$  interface, and  $-\kappa_{al} \nabla C_k$  is the downgradient along-layer diffusive flux within layer  $k$ . Both diffusivities,  $\kappa_{cr}$  and  $\kappa_{al}$ , can be functions of space and time. Finally, notice that the diffusive fluxes arise from gradients in the tracer concentration, not gradients in the thickness weighted tracer. Consequently, the presence of tracer diffusion does not affect self-consistency with the thickness equation, since the diffusive fluxes all vanish when the tracer concentration is a space-time constant.

### Thickness weighted tracer equation or tracer concentration equation?

As suggested at the start of this section, the thickness weighted tracer equation offers a straightforward means to ensure that tracer content is conserved when including further physical processes, such as cross-layer transport, along-layer subgrid scale advection, and tracer diffusion. The reason is that this equation provides a budget for the tracer content per horizontal area within the shallow water layer. In contrast, tracer concentration, as an intrinsic property,

does not satisfy a budget equation. We thus recommend using the thickness weighted tracer budget as the foundation for self-consistently including new physical processes. The tracer concentration is then simply diagnosed through division

$$C_k = \frac{(h_k C_k)}{h_k}. \quad (1.131)$$

Additionally, the thickness weighted tracer equation can be converted to the tracer concentration equation through use of the product rule and the thickness equation.

#### 1.6.4 Viscous frictional stresses acting within the layer

As discussed in VOLUME 2, accelerations from friction appear in the momentum equation via the divergence of the **friction stress tensor**. For a shallow water fluid we may choose to include a frictional stress proportional to lateral shears within each layer, much like the friction operators studied in VOLUME 2. The thickness weighted velocity equation (1.87) is given for a layer in the form

$$\partial_t(h_k \mathbf{u}_k) + \nabla \cdot (h_k \mathbf{u}_k \otimes \mathbf{u}_k) + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k = -(h_k / \rho_{\text{ref}}) \nabla p_k + h_k \mathbf{F}_k, \quad (1.132)$$

where  $\mathbf{F}_k$  is an acceleration arising from viscous friction. For a Laplacian operator arising from within-layer strains, we make use of the **generalized vertical coordinate** discussion in Section 12.15 to write

$$h_k \mathbf{F}_k = \nabla \cdot (h_k \nu_k \nabla \mathbf{u}_k), \quad (1.133)$$

where  $\nu_k > 0$  is the along-layer kinematic viscosity (dimensions  $L^2 T^{-1}$ ), which can generally be a function of the flow. Notice how the thickness appears inside the divergence operator on the right hand side of equation (1.133). It accords with the treatment of along-layer tracer diffusion in equation (1.130). Like for the tracer, we note that the friction arises from shears in the velocity, not shears in the thickness weighted velocity.<sup>11</sup> Treatment of the Laplacian viscous operator (1.133) is directly analogous to treatment of Laplacian tracer diffusion in Section 1.6.3. However, when the viscosity is a function of space, the direct analog holds only for the Cartesian coordinates used here. More general coordinates (e.g., spherical) require a distinct treatment to ensure symmetries are respected by the friction operator so to conserve **angular momentum** as studied in VOLUME 2. The appendix to *Griffies and Hallberg (2000)* and Part 5 of *Griffies (2004)* offer details for both the Laplacian and biharmonic friction operators.

#### 1.6.5 Parameterized interfacial stresses

We now consider parameterized stresses at the layer interfaces. One application concerns the treatment of winds at the top of the shallow water column, and bottom friction at the interface with the solid-earth bottom. For this purpose, consider the Boussinesq form (see VOLUME 2) of the vertical stress divergence appearing in the momentum equation

$$\mathbf{F} = \frac{1}{\rho_0} \frac{\partial \boldsymbol{\tau}}{\partial z}, \quad (1.134)$$

<sup>11</sup>Placement of thickness inside the divergence operator in equation (1.133) is consistent with the continuum treatment of **generalized vertical coordinates** in Section 12.15. As noted by *Gent (1993)*, this placement of the thickness is often ignored in the shallow water literature. Instead, one often finds  $\mathbf{F}_k = \nabla \cdot (\nu_k \nabla \mathbf{u}_k)$ , which is appropriate only when the layer thickness is a constant, as for the two-dimensional non-divergent barotropic model in Chapter 4. For the shallow water model, in which thickness is generally not constant, then use of a friction operator in the form  $\mathbf{F}_k = \nabla \cdot (\nu_k \nabla \mathbf{u}_k)$  does *not* correspond to the divergence of a symmetric frictional stress tensor, and as such it is physically inconsistent.

where  $\boldsymbol{\tau}$  is the horizontal stress vector acting on the layer interface. We discretize this stress divergence for the shallow water layer as

$$\mathbf{F}_k = \frac{1}{\rho_{\text{ref}}} \frac{\boldsymbol{\tau}_{k-1/2} - \boldsymbol{\tau}_{k+1/2}}{h_k}. \quad (1.135)$$

In this expression,  $h_k$  is the layer thickness,  $\boldsymbol{\tau}_{k-1/2}$  is the stress vector acting at the upper layer interface and  $\boldsymbol{\tau}_{k+1/2}$  is the stress vector at the lower layer interface. Since a shallow water layer is homogeneous, the interface stress is, in effect, applied uniformly throughout the layer as a layer body stress.

### Wind stress

Stress at the top interface of the top layer arises from winds, so that

$$\boldsymbol{\tau}_{k_{\text{top}}} = \boldsymbol{\tau}_{\text{wind}}, \quad (1.136)$$

where  $k_{\text{top}}$  is the index for the layer that sits at the top of the column at the particular horizontal position. We have  $k_{\text{top}} = 1$  for horizontal positions where that top layer exists. Yet it is possible for this layer to vanish, such as when a lower layer outcrops, in which case  $k_{\text{top}} \neq 1$ . Furthermore, it is important to apply this boundary stress over a nontrivial layer thickness, which for thin upper layers can mean the boundary stress is applied to more than one layer. Doing so ensures that the boundary stress does not over-accelerate a layer that happens to be very thin.

### Bottom drag

At the bottom interface of the lowest layer,  $\kappa = k_{\text{bot}}$ , bottom drag is commonly applied as a means to include dissipation, in which case<sup>12</sup>

$$\boldsymbol{\tau}_{k_{\text{bot}}} = \boldsymbol{\tau}^{\text{bot}} = \rho_{\text{ref}} C_D |\mathbf{u}|_{k_{\text{bot}}} \mathbf{u}_{k_{\text{bot}}}, \quad (1.137)$$

where  $C_D$  is a dimensionless drag coefficient (typically with values on the order of  $10^{-3}$ ), and  $\mathbf{u}_{k_{\text{bot}}}$  is the velocity in the layer adjacent to the bottom. Generally we expect  $k_{\text{bot}} = \kappa$ , yet it will differ in regions of incropping layers along the bottom.

### Viscous stress

Finally, we may wish to consider a viscous stress that acts between layers and takes on the form

$$\boldsymbol{\tau}_{\kappa \pm 1/2} = -\rho_{\text{ref}} (\mu \partial_z \mathbf{u})_{\kappa \pm 1/2}, \quad (1.138)$$

where  $\mu > 0$  is a vertical kinematic viscosity (dimensions  $L^2 T^{-1}$ ) and  $(\mu \partial_z \mathbf{u})_{\kappa \pm 1/2}$  is the viscous stress evaluated at the  $\kappa \pm 1/2$  interface.

## 1.6.6 Transfer of horizontal momentum from inter-layer volume transfer

A transfer of volume across the layer interface gives rise to a transfer of horizontal momentum. In a manner directly akin to the treatment of tracers in Section 1.6.3, we have the thickness

---

<sup>12</sup>Be aware that the bottom drag coefficient,  $C_D$ , is distinct from tracer concentration,  $C_\kappa$ .

weighted velocity equation given by

$$\begin{aligned} \partial_t(h_k \mathbf{u}_k) + \nabla \cdot (h_k \mathbf{u}_k \otimes \mathbf{u}_k) + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k + \nabla \cdot (h_k \nu_k \nabla \mathbf{u}_k) \\ &- [(w^{(\dot{\eta})} \mathbf{u} - \mu \partial_z \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u} - \mu \partial_z \mathbf{u})_{k+1/2}] + \delta_{k,k_{\text{top}}} \boldsymbol{\tau}_{\text{wind}} - \delta_{k,k_{\text{bot}}} \boldsymbol{\tau}^{\text{bot}}. \end{aligned} \quad (1.139)$$

In this equation,  $(w^{(\dot{\eta})} \mathbf{u})_{k+1/2}$  is the transfer of horizontal momentum across the lower layer interface, and  $(w^{(\dot{\eta})} \mathbf{u})_{k-1/2}$  is the transfer across the upper layer interface. To estimate the horizontal velocity at the interface, we can take an upwind approach just as for the tracers in equation (1.126)

$$(w^{(\dot{\eta})} \mathbf{u})_{k+1/2} = w_{k+1/2}^{(\dot{\eta})} \begin{cases} \mathbf{u}_{k+1} & \text{if } w_{k+1/2}^{(\dot{\eta})} > 0 \\ \mathbf{u}_k & \text{if } w_{k+1/2}^{(\dot{\eta})} < 0, \end{cases} \quad (1.140)$$

Gent (1993) notes that some realizations of the shallow water equations in the presence of  $w^{(\dot{\eta})} \neq 0$  fail to incorporate the transfer of horizontal momentum between the layers present in equation (1.139). As such, these realizations are not self-consistent and thus do not correspond to a discrete realization of a continuous isopycnal model.<sup>13</sup>

### 1.6.7 Inclusion of subgrid along-layer volume transport

In Section 1.6.2, we introduced a subgrid scale advective volume tranport within layers, and then included that transport in the tracer equation in Section 1.6.3. To see how to do so for the linear momentum equation, it is sufficient to focus on the case without friction or viscosity, Unlike the thickness weighted tracer equation (1.124), here we do not have a straightforward consistency check with the thickness equation (1.122) to determine a unique form for including the effects from  $\mathbf{u}_k^{\text{sgs}}$  into the momentum equation. We thus consider the following formulations.

#### Absence of $\mathbf{u}_k^{\text{sgs}}$ in the thickness weighted velocity equation

In this formulation we assume there is no appearance of  $\mathbf{u}_k^{\text{sgs}}$  in the thickness weighted velocity equation, so that

$$\begin{aligned} \partial_t(h_k \mathbf{u}_k) + \nabla \cdot (h_k \mathbf{u}_k \otimes \mathbf{u}_k) + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k \\ &- [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}]. \end{aligned} \quad (1.141)$$

Use of the product rule gives

$$\begin{aligned} h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k \cdot \nabla) \mathbf{u}_k] + \mathbf{u}_k [\partial_t h_k + \nabla \cdot (h_k \mathbf{u}_k)] + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k \\ &- [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}], \end{aligned} \quad (1.142)$$

with the thickness equation (1.123) then rendering

$$\begin{aligned} h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k \cdot \nabla) \mathbf{u}_k] + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k \\ &- [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}] + \mathbf{u}_k [\nabla \cdot (h_k \mathbf{u}_k^{\text{sgs}}) + w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}]. \end{aligned} \quad (1.143)$$

This formulation is unsatisfying since the velocity is advected by the velocity,  $\mathbf{u}_k$ , whereas tracer is advected by  $\mathbf{u}_k^{\text{eff}}$  as seen in equations (1.124) and (1.128).

<sup>13</sup>We formulate the continuous isopycnal model equations in Chapter 14.

### Absence of $u_k^{\text{sgs}}$ in the velocity equation

As noted by [Jansen et al. \(2024\)](#), certain layered ocean models (such as MOM6 [Adcroft et al. \(2019\)](#)), are formulated with no appearance of  $u_k^{\text{sgs}}$  in the velocity equation, so that

$$h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k \cdot \nabla) \mathbf{u}_k] + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k = -(h_k / \rho_{\text{ref}}) \nabla p_k - [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}]. \quad (1.144)$$

Use of the product rule and thickness equation (1.123) leads to the thickness weighted velocity equation

$$\begin{aligned} \partial_t(h_k \mathbf{u}_k) + \nabla \cdot (h_k \mathbf{u}_k \otimes \mathbf{u}_k) + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k \\ &- [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}] - \mathbf{u}_k [\nabla \cdot (h_k \mathbf{u}_k^{\text{sgs}}) + w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}]. \end{aligned} \quad (1.145)$$

This equation is unsatisfying for the same reason as the velocity equation (1.143). Namely, both formulations have the advection of velocity determined by the velocity,  $\mathbf{u}_k$ , whereas tracer is advected by  $\mathbf{u}_k^{\text{eff}}$ , as seen in equations (1.124) and (1.128).

### Advection by $u_k^{\text{eff}}$ in the velocity equation and thickness weighted velocity equation

Next consider the case where advection occurs with  $u_k^{\text{eff}}$ , just like for tracers as in equation (1.124), so that the thickness weighted velocity equation is written in the form

$$\partial_t(h_k \mathbf{u}_k) + \partial_i[h u_i^{\text{eff}} \mathbf{u}]_k + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k = -(h_k / \rho_{\text{ref}}) \nabla p_k - [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}], \quad (1.146)$$

where we exposed the Cartesian tensor index,  $i$ , on the left hand side. The thickness equation (1.122) leads to the identity

$$\partial_t(h_k \mathbf{u}_k) + \partial_i[h u_i^{\text{eff}} \mathbf{u}]_k = h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k^{\text{eff}} \cdot \nabla) \mathbf{u}_k] + \mathbf{u}_k [\partial_t h_k + \nabla \cdot (h_k \mathbf{u}_k^{\text{eff}})] \quad (1.147a)$$

$$= h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k^{\text{eff}} \cdot \nabla) \mathbf{u}_k] - \mathbf{u}_k (w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})}), \quad (1.147b)$$

which, when inserted to equation (1.146), gives the velocity equation

$$\begin{aligned} h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k^{\text{eff}} \cdot \nabla) \mathbf{u}_k] + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k \\ &- [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - w_{k-1/2}^{(\dot{\eta})} \mathbf{u}_k] + [(w^{(\dot{\eta})} \mathbf{u})_{k+1/2} - w_{k+1/2}^{(\dot{\eta})} \mathbf{u}_k]. \end{aligned} \quad (1.148)$$

This form for the velocity equation corresponds to the tracer concentration equation (1.128), with both equations having advection determined by the effective velocity,  $\mathbf{u}^{\text{eff}}$ . Rearrangement renders the equivalent form

$$\begin{aligned} h_k [\partial_t \mathbf{u}_k + (\mathbf{u}_k \cdot \nabla) \mathbf{u}_k] + f \hat{\mathbf{z}} \times h_k \mathbf{u}_k &= -(h_k / \rho_{\text{ref}}) \nabla p_k \\ &- [(w^{(\dot{\eta})} \mathbf{u})_{k-1/2} - (w^{(\dot{\eta})} \mathbf{u})_{k+1/2}] + [w_{k-1/2}^{(\dot{\eta})} - w_{k+1/2}^{(\dot{\eta})} - h_k (\mathbf{u}_k^{\text{sgs}} \cdot \nabla)] \mathbf{u}_k. \end{aligned} \quad (1.149)$$

#### 1.6.8 Further study

Elements of the discrete stacked shallow water model formulation given in this section are considered by [Jansen et al. \(2024\)](#) for the case of a continuous vertical coordinate.



## 1.7 Exercises

EXERCISE 1.1: RELATIONS FOR VERTICAL VELOCITY (EXERCISE (3.2) OF [Vallis \(2006\)](#))

Show that the vertical velocity within a shallow water system is given by

$$w = \left[ \frac{z - \eta_b}{h} \right] \frac{Dh}{Dt} + \frac{D\eta_b}{Dt}, \quad (1.150)$$

where  $\eta_b$  is the position of the bottom topography (see Figure 1.1). Interpret the result, showing that it gives sensible answers at the top and bottom of the fluid layer.

EXERCISE 1.2: STRETCHING OF A VERTICAL COLUMN WITH LAYER VOLUME EXCHANGE

Show that if there is transport across the surface interface of a single shallow water layer, as per the thickness equation (1.118), then the column stretching equation (1.43) becomes

$$\frac{D}{Dt} \left[ \ln \left( \frac{z - \eta_b}{h} \right) \right] = \frac{w^{(\eta)}}{h}. \quad (1.151)$$

Evidently, in the presence of a surface boundary volume flux, a column of shallow water fluid no longer stretches or squeezes uniformly. Instead, for  $w^{(\eta)} < 0$ , a fluid parcel moves down within the column as more fluid is added to the top of the layer, and conversely when volume leaves the layer.

EXERCISE 1.3: DERIVING THE SHALLOW WATER TRACER EQUATION (1.47)

Show all steps needed to derive equations (1.46a)-(1.46c) and then show the steps leading to the shallow water tracer equation (1.47). Hint: use the  $z$  independence of  $\mathbf{u}$  and  $\partial_z w$  within the shallow water layer.

EXERCISE 1.4: SHALLOW WATER EQUATIONS WITH TIDES

In VOLUME 2, we derived the ocean primitive equations in the presence of astronomical forcing that leads to tides. Specialize the general results from that discussion to derive the thickness and momentum equations for a single layer of shallow water fluid in the presence of astronomical tidal forcing. Assume the perturbation geopotential is vertically independent.

EXERCISE 1.5: INVERTED REDUCED GRAVITY MODEL

Derive the shallow water equations for a single moving layer of fluid of density  $\rho_2$  above a rigid floor, with this moving layer below a stagnant fluid of density  $\rho_1$ , with  $\rho_1 < \rho_2$ , and with the upper stagnant layer assumed to have infinite thickness (as per the upper ocean reduced gravity model in Figure 1.4). Assume  $\nabla p_a = 0$ . Discuss the constraint placed on the interface  $\eta_{1/2}$  to maintain a stagnant upper layer. Show that as  $\rho_1/\rho_2 \rightarrow 0$  with  $\rho_{ref} = \rho_2$ , then the single layer shallow water equations emerge. Make use of notation from the three-layer system shown in Figure 1.6. This model might be used to study flow in the atmosphere well above the boundary layer, or the abyssal ocean well below the pycnocline. Hint: invert the approach taken in Section 1.3 for the reduced gravity model of the upper ocean.



## Chapter 2

# SHALLOW WATER DYNAMICS

In this chapter we further our dynamical understanding of the shallow water model. The study includes geostrophy, thermal wind (as expressed by Margules' relation), momentum budgets, pressure form stress, kinetic energy, gravitational potential energy, available potential energy, and mechanical energy. We also develop the following case studies to exemplify the fundamentals: the steady force balance in a zonally reentrant channel, and the dynamics of angular momentum in a rotating tank of shallow water fluid. We offer many details to support those aiming to become nimble with the equations describing the momentum and energy dynamics of a stacked shallow water model.

### READER'S GUIDE TO THIS CHAPTER

This chapter builds from the formulations in Chapter 1, as well as the mechanics of geostrophic flow and pressure form stress from VOLUME 2. We make use of the results in this chapter for many of the subsequent chapters. Notationally, we follow the same convention for the gradient operator noted at the start of Chapter 1.

<b>2.1</b>	<b>Loose threads</b>	<b>40</b>
<b>2.2</b>	<b>Geostrophic balance and thermal wind</b>	<b>40</b>
2.2.1	Geostrophy for a single layer	40
2.2.2	Margules' relation for two layers	41
2.2.3	Geostrophic transport within layers	42
2.2.4	Geostrophic eddying flow	43
<b>2.3</b>	<b>Thickness weighted momentum equation</b>	<b>45</b>
2.3.1	Single layer equations	45
2.3.2	Geostrophic and ageostrophic contributions	46
2.3.3	Form stresses acting on a shallow water column	47
2.3.4	Comments on the two pressure force formulations	49
<b>2.4</b>	<b>Contact pressure forces in shallow water layers</b>	<b>49</b>
2.4.1	Pressure contact force and pressure body force	50
2.4.2	$N$ -layer equations	50
2.4.3	Contact pressure force along vertical sides	50
2.4.4	Contact pressure force along the top and bottom interfaces	52
2.4.5	Form stress	53
2.4.6	Net contact pressure force on a shallow water column	53
2.4.7	Contact pressure force summed over all layers	54
2.4.8	Horizontal pressure force with potential energy gradients	54
2.4.9	Momentum equation with contact pressure forces	55
2.4.10	Further reading	58

---

<b>2.5</b>	<b>Energetics for a single layer</b>	<b>58</b>
2.5.1	Gravitational potential energy	59
2.5.2	Kinetic energy and work from pressure gradients	60
2.5.3	Kinetic energy and buoyancy work	61
2.5.4	Kinetic energy and pressure form stress	62
2.5.5	Mechanical energy budget	63
2.5.6	Available potential energy	66
<b>2.6</b>	<b>Energetics for <math>N</math> layers</b>	<b>67</b>
2.6.1	Potential energy and available potential energy	67
2.6.2	Potential energy and buoyancy work	69
2.6.3	Kinetic energy	69
2.6.4	Comments and further study	70
<b>2.7</b>	<b>Momentum balance in a zonal channel</b>	<b>71</b>
2.7.1	Volume transport for steady flow	71
2.7.2	Steady meridional balances	72
2.7.3	Steady zonal balance	73
2.7.4	The role of frictional bottom drag	74
2.7.5	Correlation between surface height and topographic slope	74
2.7.6	Connection to meridional geostrophic transport	75
2.7.7	Sinusoidal example	76
2.7.8	Comments and further study	78
<b>2.8</b>	<b>Angular momentum in a rotating tank</b>	<b>80</b>
2.8.1	Angular momentum for a column of shallow water fluid	81
2.8.2	Material time evolution of the angular momentum	81
2.8.3	Materially invariant angular momentum	82
2.8.4	Comments	82
<b>2.9</b>	<b>Exercises</b>	<b>83</b>

---

## 2.1 Loose threads

- Some more interpretation of the  $N$ -layer mechanical energy equations and their connections to the continuous case.
- More schematics
- Consider adding a diabatic term to the energy budgets in Sections 2.5 and 2.6.

## 2.2 Geostrophic balance and thermal wind

As described in VOLUME 2, a geostrophic balance arises when the flow balances are such that the material time derivative plays a subdominant role in the inviscid horizontal momentum equation. This is a balance that holds when the Rossby number is small. The resulting balance between Coriolis and pressure accelerations constitutes the geostrophic balance. We consider here the implications of geostrophy for one and two-layer shallow water systems.

### 2.2.1 Geostrophy for a single layer

Ignoring the applied pressure ( $p_a = 0$ ) leads to the geostrophic balance for a single shallow water layer

$$f \hat{z} \times \mathbf{u}_g = -g \nabla \eta \implies f \mathbf{u}_g = g \hat{z} \times \nabla \eta, \quad (2.1)$$

or in component form

$$u_g = -\frac{g}{f} \frac{\partial \eta}{\partial y} \quad \text{and} \quad v_g = \frac{g}{f} \frac{\partial \eta}{\partial x}. \quad (2.2)$$

Consequently, the shallow water layer geostrophic current is balanced by the gradient of the free surface (sea level). In the northern hemisphere, where  $f > 0$ , geostrophic shallow water currents flow counter-clockwise around negative sea level anomalies (low pressure) and clockwise around positive sea level anomalies (high pressure). The opposite orientation holds in the southern hemisphere, where  $f < 0$ . Figure 2.1 shows a schematic of geostrophic balance for a single shallow water layer.

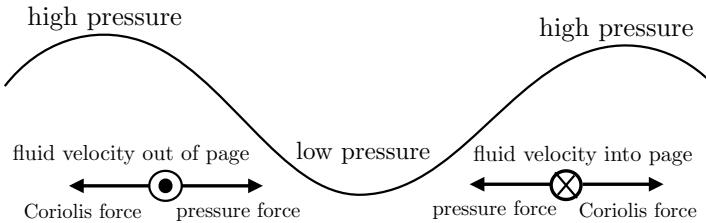


FIGURE 2.1: Side view of geostrophic balance for a single shallow water layer, here shown with two high pressure centers surrounding a low pressure center. A fluid particle path follows isolines of the free surface (constant pressure surface), with the Coriolis force balancing the pressure gradient force. In the northern hemisphere, where  $f > 0$ , geostrophic flow is counter-clockwise around a low pressure center and clockwise around a high pressure center, so that the Coriolis force acts to the right of the flow and it is balanced by a pressure gradient acting to the left. For the southern hemisphere, geostrophic flow is in the opposite direction since  $f < 0$ , so that the pressure gradient acts to the right of the flow and Coriolis to the left.

## 2.2.2 Margules' relation for two layers

Now consider two shallow water layers as in Figure 1.6. Recall the layer pressure equations (1.63) and (1.64), which leads to the pressure difference

$$p_1 - p_2 = g \eta_{3/2} (\rho_1 - \rho_2) + g z (\rho_2 - \rho_1) = g_{3/2}^r \rho_{\text{ref}} (z - \eta_{3/2}), \quad (2.3)$$

where the reduced gravity is given by equation (1.69)

$$g_{3/2}^r = g (\rho_2 - \rho_1) / \rho_{\text{ref}} > 0, \quad (2.4)$$

where  $\rho_{\text{ref}}$  is the shallow water Boussinesq reference density (1.67). The density difference,  $\rho_2 - \rho_1$ , is generally much smaller than either density, so that  $g_{3/2}^r \ll g$ . For the Boussinesq shallow water system, the momentum equations are given by

$$\frac{D^{(1)} \mathbf{u}_1}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_1 = -\rho_{\text{ref}}^{-1} \nabla p_1 \quad (2.5a)$$

$$\frac{D^{(2)} \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u}_2 = -\rho_{\text{ref}}^{-1} \nabla p_2, \quad (2.5b)$$

where we introduced the material time derivatives for each layer (we introduced this notation in Section 1.4.4)

$$\frac{D^{(k)}}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}_k \cdot \nabla. \quad (2.6)$$

Making use of the pressure difference (2.3) renders

$$\frac{D_1 \mathbf{u}_1}{Dt} - \frac{D_2 \mathbf{u}_2}{Dt} + f \hat{\mathbf{z}} \times \Delta \mathbf{u} = -\rho_{ref}^{-1} \nabla(p_1 - p_2) = g'_{3/2} \nabla \eta_{3/2}, \quad (2.7)$$

where

$$\Delta \mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2 \quad (2.8)$$

is the vertical difference of the layer horizontal velocities. We see that the difference in the geostrophic velocities for the two layers is proportional to the slope of the interface between the two layers

$$f \hat{\mathbf{z}} \times \Delta \mathbf{u}_g = g'_{3/2} \nabla \eta_{3/2} \implies \Delta u_g = +(g'_{3/2}/f) \partial_y \eta_{3/2} \quad \text{and} \quad \Delta v_g = -(g'_{3/2}/f) \partial_x \eta_{3/2}. \quad (2.9)$$

These equations are known as the Margules' relation. It applies at any interface between two shallow water fluid layers. It says that the vertical difference between the layer geostrophic velocities is proportional to the interface slope. When the slope is large, the vertical difference in the geostrophic velocity is large. Also, the velocity difference is large when the reduced gravity is large; i.e., when the density difference is large. We illustrate this relation in Figure 2.2. The Margules relation is a discrete (two-layer) version of the thermal wind balance studied in VOLUME 2.

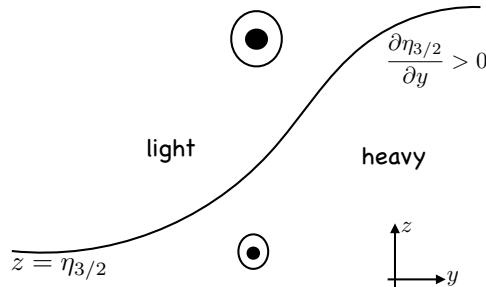


FIGURE 2.2: Illustrating Margules' relation for the northern hemisphere ( $f > 0$ ). Here we show the interface between a two-layer shallow water model with a heavy layer to the right and a light layer to the left. The slope of the interface is positive,  $\partial \eta_{3/2} / \partial y > 0$ , thus leading to an increase in the eastward zonal geostrophic velocity moving upward, as depicted by the circles with a dot. This orientation corresponds to the northern hemisphere atmospheric jet stream, whereby the interface between cold/heavy air to the north and warm/light air to the south leads to a zonal thermal wind jet. This figure is directly comparable to the continuously stratified thermal wind balance studied in VOLUME 2.

### 2.2.3 Geostrophic transport within layers

We are often interested in computing the net volume transport within a layer of fluid in order to measure how much the fluid is moving across a particular region. For an  $N$ -layer shallow water fluid this transport is written

$$\mathbf{U} = \int \mathbf{u} dz = \sum_{k=1}^N \mathbf{u}_k h_k, \quad (2.10)$$

where  $\mathbf{u}_k$  is the layer horizontal velocity and  $h_k$  the layer thickness. For many purposes it is sufficient to compute the transport due to the geostrophic motion, in which case

$$\mathbf{u}_k = (\rho_{\text{ref}} f)^{-1} \hat{\mathbf{z}} \times \nabla p_k, \quad (2.11)$$

so that the geostrophic transport is

$$\mathbf{U}_g = \int_{\eta_b}^{\eta} \mathbf{u}_g dz = (\rho_{\text{ref}} f)^{-1} \hat{\mathbf{z}} \times \sum_{k=1}^N h_k \nabla p_k. \quad (2.12)$$

For the pressure gradient we can make use of the expression (1.88)

$$\nabla p_k = \nabla p_a + \nabla p_k^{\text{dyn}} = \nabla p_a + \rho_{\text{ref}} \sum_{j=1}^k g_{j+1/2}^r \nabla \eta_{j+1/2} = \nabla p_{k-1} + \rho_{\text{ref}} g_{k-1/2}^r \nabla \eta_{k-1/2}, \quad (2.13)$$

thus revealing the cascade of contributions from each of the layer interfaces.

As an exercise, let us write the geostrophic transport for  $N = 3$  layers, in which the layer pressure gradients are

$$\nabla p_1 = \nabla p_a + \rho_{\text{ref}} g \nabla \eta_{1/2} \quad (2.14a)$$

$$\nabla p_2 = \nabla p_1 + \rho_{\text{ref}} g_{3/2}^r \nabla \eta_{3/2} \quad (2.14b)$$

$$\nabla p_3 = \nabla p_2 + \rho_{\text{ref}} g_{5/2}^r \nabla \eta_{5/2}. \quad (2.14c)$$

We see here the utility of setting

$$\rho_{\text{ref}} = \rho_1, \quad (2.15)$$

in which case the geostrophic transport within the three layers is

$$h_1 \mathbf{u}_{1g} = \frac{h_1}{f} \hat{\mathbf{z}} \times \nabla(g_{1/2}^r \eta_{1/2}) \quad (2.16a)$$

$$h_2 \mathbf{u}_{2g} = \frac{h_2}{f} \hat{\mathbf{z}} \times \nabla(g_{1/2}^r \eta_{1/2} + g_{3/2}^r \eta_{3/2}) \quad (2.16b)$$

$$h_3 \mathbf{u}_{3g} = \frac{h_3}{f} \hat{\mathbf{z}} \times \nabla(g_{1/2}^r \eta_{1/2} + g_{3/2}^r \eta_{3/2} + g_{5/2}^r \eta_{5/2}), \quad (2.16c)$$

so that the vertically integrated geostrophic transport is

$$\mathbf{U}_g = f^{-1} \hat{\mathbf{z}} \times \left[ h_1 \nabla(g \eta) + h_2 \nabla(g \eta + g_{3/2}^r \eta_{3/2}) + h_3 \nabla(g \eta + g_{3/2}^r \eta_{3/2} + g_{5/2}^r \eta_{5/2}) \right]. \quad (2.17)$$

This expression for  $\mathbf{U}_g$  displays the cascade of contributions from each of the layer interfaces and their corresponding reduced gravities. Evidently, the geostrophic transport is directly related to the slopes for the layer interfaces, with more transport associated with larger magnitudes in the slopes as well as larger reduced gravities.

## 2.2.4 Geostrophic eddying flow

The ocean and atmosphere are highly turbulent fluids, with turbulent features extending from the small scales (millimeters) to large scales (hundreds to thousands of kilometers). The larger scale macro-turbulent features feel the earth's rotation and thus maintain a force balance

close to geostrophic.<sup>1</sup> We here outline some features of an eddying ocean **geostrophic flow** as idealized using the reduced gravity model of Section 1.3. Figure 2.3 shows a vertical-zonal slice through the upper portion of an ocean eddy in the middle latitude northern hemisphere ( $f > 0$ ). The central region consists of a geostrophic eddy, sometimes also referred to as an ocean **mesoscale** eddy. The signature of the eddy is a depression in the free surface height and upward deformation of the pycnocline. The lateral scale of the eddy is on the order of the internal deformation scale (see Exercise 2.10).

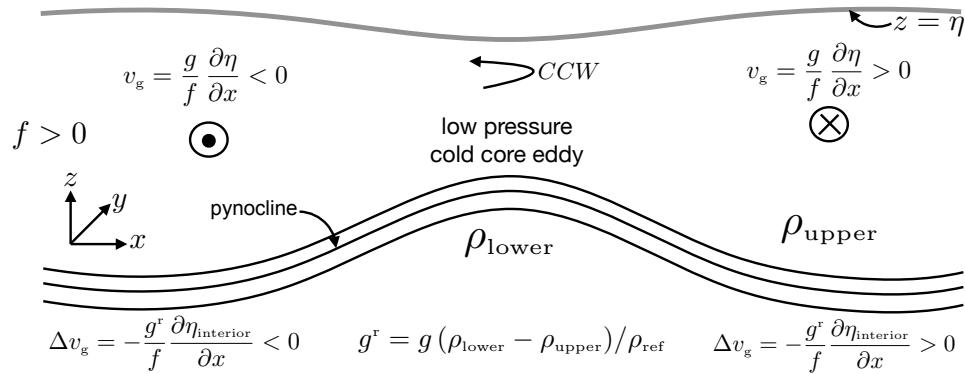


FIGURE 2.3: Vertical-zonal slice through a northern hemisphere mid-latitude cold-core ocean eddy looking from the south to the north (north is into the page). The ocean free surface is depressed down in the middle so that geostrophic flow is cyclonic (counter-clockwise in north) within the upper portion of the eddy where  $\rho = \rho_{\text{upper}}$ . The pycnocline (region of enhanced vertical density gradient) is deformed upward, and the baroclinic flow implied by Margules' relation (2.9) is indicated, making use of the reduced gravity  $g^r = g(\rho_{\text{lower}} - \rho_{\text{upper}})/\rho_{\text{ref}}$ . Note that  $\Delta v_g = v_{\text{upper}} - v_{\text{lower}}$ , so that  $\Delta v_g > 0$  means that the meridional flow increases in the  $+\hat{y}$ -direction when moving from the lower to upper layer, and conversely when  $\Delta v_g < 0$ .

The ocean eddy in Figure 2.3 is an anomalously dense cyclonic mesoscale eddy with the dense water causing the pycnocline to deviate upward. If density is dominated by temperature, as it typically is within the middle to lower latitude oceans, then the eddy is a cold core eddy, meaning that the core of the eddy is cold. Under geostrophic balance, water circulates counter-clockwise in the northern hemisphere within the upper portion of the cold core eddy, in the region where the pressure gradient force is dominated by the free surface undulation. In this case we see say the eddy is cyclonic. According to the reduced gravity model from Section 1.3 (see in particular Figure 1.5), the ratio of the free surface undulation to the pycnocline undulation scales like the reduced gravity, so that a meter undulation of the free surface corresponds to roughly 100 m undulation of the pycnocline. The same ideas hold for a warm core eddy, such as that depicted in Figure 2.4, with undulations complementing those in the cold core and thus supporting anti-cyclonic thermal wind flow.

In presenting the idealized rendition of an ocean eddy in Figure 2.3, we are assuming a reduced gravity model is sufficient and that the atmosphere has no significant horizontal pressure gradients over the scale of the eddy. Under these assumptions, we make use of the Margules' relation (2.9) to deduce the thermal wind flow in the upper layer relative to the layer below; i.e., the vertical shear in the geostrophic flow. For the left side of the eddy, where  $\partial\eta_{\text{interior}}/\partial x > 0$ , the vertical shear in the meridional geostrophic velocity is southward, consistent

<sup>1</sup>As noted in VOLUME 2, the **gradient wind** balance provides a more accurate approximation to flows in ocean and atmospheric eddies by also including the **centrifugal** acceleration associated with the curved motion. Even so, the geostrophic balance provides a sufficient approximation for many purposes and it will be used here, along with thermal wind.

with orientation of the flow implied by the sea surface gradient. Conversely, on the right side of the eddy, where  $\partial\eta_{\text{interior}}/\partial x < 0$ , the vertical shear in the meridional flow is northward.

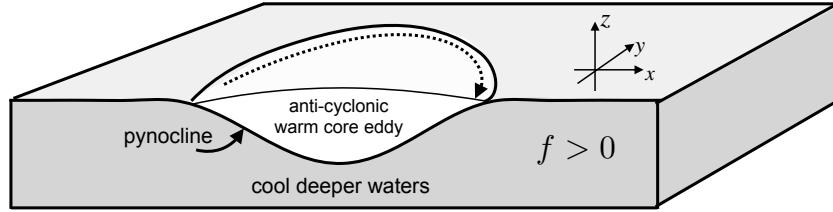


FIGURE 2.4: Schematic of a warm core (light water) geostrophic ocean eddy as idealized by a reduced gravity model. The geostrophic/thermal wind flow is anti-cyclonic within the eddy, which contrasts to the cyclonic flow for a cold core eddy as depicted in Figure 2.3. The eddy is characterized by a slight expansion of the free surface (high pressure) and a relatively larger depression of the pycnocline.

## 2.3 Thickness weighted momentum equation

Throughout our discussion of the shallow water model, we made use of the prognostic equation for the velocity of a layer. Here, we study the momentum equation as determined by the vertically integrated velocity within a shallow water layer. This formulation proves particularly useful when studying forces acting on the layer, such as those from pressure contact forces (including form stresses) as well as kinetic stresses due to the fluid motion.<sup>2</sup>

For a shallow water model with just a single layer, the water column extends from the surface to the bottom of the layer (see Figure 2.5)

$$\int_{\eta_b}^{\eta} \mathbf{u} dz = \mathbf{u} h, \quad (2.18)$$

so that the column momentum equals to  $\mathbf{u} h \rho dx dy$ . The resulting momentum equation is written in its flux form. In Section 2.7 we illustrate the momentum budget for a zonal channel. We also show in Section 2.4.2 that the  $N$ -layer equations are isomorphic to the single layer, thus allowing for concepts developed for a single layer to be readily extended to multiple layers.

### 2.3.1 Single layer equations

Recall the velocity and thickness equations written using the material time operator

$$\frac{D\mathbf{u}}{Dt} + f \hat{z} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (2.19)$$

Combining these two equations allows us to write the thickness weighted material acceleration as

$$h \frac{D\mathbf{u}}{Dt} = h \frac{D\mathbf{u}}{Dt} + \mathbf{u} \left[ \frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} \right] = \partial_t(h \mathbf{u}) + \nabla \cdot [h \mathbf{u} \otimes \mathbf{u}], \quad (2.20)$$

so that the thickness weighted equation takes the vector form

$$\partial_t(h \mathbf{u}) + \nabla \cdot [h \mathbf{u} \otimes \mathbf{u}] + f \hat{z} \times (h \mathbf{u}) = -g h \nabla \eta. \quad (2.21)$$

<sup>2</sup>This formulation is of particular use for studies of rotating hydraulics such as pursued in the book by [Pratt and Whitehead \(2008\)](#).

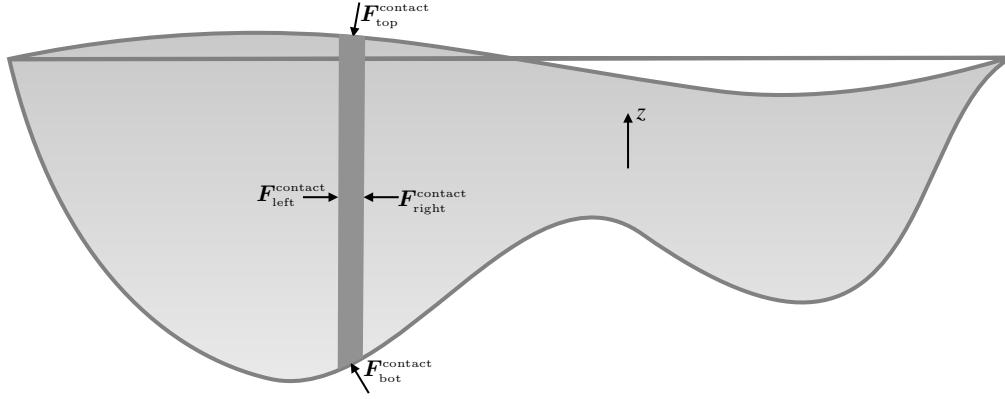


FIGURE 2.5: Momentum of a column of a single layer shallow water fluid is affected by contact forces at the column boundaries (pressure and friction), as well as body forces acting throughout the column (Coriolis and gravity).

The Cartesian tensor form of the [tensor product](#) (also called the *outer product* in linear algebra) is<sup>3</sup>

$$[\mathbf{u} \otimes \mathbf{u}]_{mn} = u_m u_n \quad \text{for } m, n = 1, 2, \quad (2.22)$$

with  $-h \rho \mathbf{u} \otimes \mathbf{u}$  the specialization to the shallow water system of the [kinetic stress tensor](#) appearing in the continuously stratified momentum equation ([VOLUME 2](#)). The component form of the thickness weighted momentum equation (2.21) is

$$\partial_t(h u) + \partial_x(h u^2) + \partial_y(h u v) - v h f = -g h \partial_x \eta \quad (2.23a)$$

$$\partial_t(h v) + \partial_x(h u v) + \partial_y(h v^2) + u h f = -g h \partial_y \eta. \quad (2.23b)$$

Note that when the bottom is flat then  $\nabla h = \nabla \eta$  so that

$$\partial_t(h u) + \partial_x(h u^2 + g h^2/2) + \partial_y(h u v) - v h f = 0 \quad (2.24a)$$

$$\partial_t(h v) + \partial_x(h u v) + \partial_y(h v^2 + g h^2/2) + u h f = 0. \quad (2.24b)$$

### 2.3.2 Geostrophic and ageostrophic contributions

Bringing the Coriolis terms to the right hand side of equations (2.23a) and (2.23b) renders

$$\partial_t(h u) + \partial_x(h u^2) + \partial_y(h u v) = h(-g \partial_x \eta + v f) \quad (2.25a)$$

$$\partial_t(h v) + \partial_x(h u v) + \partial_y(h v^2) + u h f = h(-g \partial_y \eta - u f). \quad (2.25b)$$

In the absence of rotation, the right hand side has contributions only from the thickness weighted pressure gradient. For the case of rotation it sometimes proves useful to decompose velocity into its geostrophic and ageostrophic components

$$f u = f(u_a + u_g) = f u_a - g \partial \eta / \partial y \quad (2.26a)$$

$$f v = f(v_a + v_g) = f v_a + g \partial \eta / \partial x, \quad (2.26b)$$

---

<sup>3</sup>Equation (2.22) is the only place in this chapter where a subscript refers to a tensor label. Otherwise, subscripts refer to a shallow water layer index as in Section 2.4.2.

in which case equations (2.25a) and (2.25b) become

$$\partial_t(h u) + \partial_x(h u^2) + \partial_y(h u v) = h f v_a \quad (2.27a)$$

$$\partial_t(h v) + \partial_x(h u v) + \partial_y(h v^2) = -h f u_a. \quad (2.27b)$$

One should be careful not to take the  $f = 0$  limit of these equations since one might spuriously conclude there is no free surface contribution. Instead, equations (2.25a) and (2.25b) should be the basis for the  $f = 0$  limit.

### 2.3.3 Form stresses acting on a shallow water column

The kinetic stress contributes to momentum evolution in equations (2.23a) and (2.23b) via its divergence. In contrast, the pressure stress contributes as a thickness weighted pressure gradient body stress. In this subsection, and in all of Section 2.4, we formulate pressure as a contact stress, in which case it also contributes to momentum evolution as a divergence. In so doing, we provide a flux-form conservation law for momentum that supports analysis and interpretation.

#### Reintroducing atmospheric pressure to symmetrize the forces acting on the layer

To expose both the surface and bottom form stresses, we reintroduce the atmospheric pressure,  $p_a$ , and thus make use of the effective sea level (1.4)

$$\eta^{eff} = \eta + p_a/(\rho g) = \eta_b + h + p_a/(\rho g), \quad (2.28)$$

with the corresponding thickness weighted horizontal momentum equation

$$\partial_t(h \mathbf{u}) + \nabla \cdot [h \mathbf{u} \otimes \mathbf{u}] + f \hat{\mathbf{z}} \times (h \mathbf{u}) = -g h \nabla \eta^{eff}. \quad (2.29)$$

#### Exposing the contact pressure stresses

The free surface height equals  $\eta = \eta_b + h$ , in which case the momentum equation (2.29) is

$$\partial_t(h \mathbf{u}) + \nabla \cdot [h \mathbf{u} \otimes \mathbf{u}] + f \hat{\mathbf{z}} \times (h \mathbf{u}) = -(g/2) \nabla h^2 - g h \nabla [\eta_b + p_a/(\rho g)]. \quad (2.30)$$

To help interpret this equation it is convenient to write the boundary terms on the right hand side as

$$-g h \nabla [\eta_b + p_a/(\rho g)] = -\nabla(h p_a/\rho) + (p_a/\rho) \nabla(\eta - \eta_b) - g h \nabla \eta_b \quad (2.31a)$$

$$= -\nabla(h p_a/\rho) + (p_a/\rho) \nabla \eta - (g h + p_a/\rho) \nabla \eta_b \quad (2.31b)$$

$$= -\nabla(h p_a/\rho) + \rho^{-1} (p_a \nabla \eta - p_b \nabla \eta_b), \quad (2.31c)$$

so that

$$-(g/2) \nabla h^2 - g h \nabla [\eta_b + p_a/(\rho g)] = -\nabla[(g/2) h^2 + h p_a/\rho] + \rho^{-1} (p_a \nabla \eta - p_b \nabla \eta_b). \quad (2.32)$$

The first term on the right hand side is the gradient of the layer integrated hydrostatic pressure

$$P \equiv \int_{\eta_b}^{\eta} [p_a + \rho g (\eta - z)] dz = h (\rho g h/2 + p_a), \quad (2.33)$$

and the second term exposes the form stresses acting at the surface and bottom of the layer. With these expressions, the horizontal thickness weighted momentum equation (2.30) becomes

$$\frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot [h \mathbf{u} \otimes \mathbf{u} + \mathbf{I} P/\rho] + f \hat{\mathbf{z}} \times (h \mathbf{u}) = (p_a \nabla \eta - p_b \nabla \eta_b)/\rho, \quad (2.34)$$

where  $\mathbf{I}$  is the unit tensor. Exposing the zonal and meridional components renders

$$\partial_t(h u) + \partial_x(h u^2 + P/\rho) + \partial_y(h u v) - v h f = (p_a \partial_x \eta - p_b \partial_x \eta_b)/\rho \quad (2.35a)$$

$$\partial_t(h v) + \partial_x(h u v) + \partial_y(h v^2 + P/\rho) + u h f = (p_a \partial_y \eta - p_b \partial_y \eta_b)/\rho. \quad (2.35b)$$

The horizontal pressure gradient appears as a continuous operator since we assumed an infinitesimal horizontal cross-sectional area for the fluid column. In contrast, the pressure form stresses appear as a vertical finite difference across the layer interfaces, which results since we are integrating over the thickness of a finite layer. Furthermore, note how the vertically integrated pressure contributions appear in a flux-form, which contrasts to the body force version that appears as thickness weighted pressure gradient.

### Kinetic stresses and contact pressure stresses combined into a momentum flux

To anticipate the thickness weighted momentum equation for the stacked shallow water model in Section 2.4, write the finite difference of the form stresses as

$$p_a \nabla \eta - p_b \nabla \eta_b = p_{1/2} \nabla \eta_{1/2} - p_{3/2} \nabla \eta_{3/2} \equiv \delta_k(p_{k-1/2} \nabla \eta_{k-1/2}). \quad (2.36)$$

We here introduced the layer interface difference operator

$$\delta_k(\Psi_{k-1/2}) = \Psi_{k-1/2} - \Psi_{k+1/2} = -(\Psi_{k+1/2} - \Psi_{k-1/2}), \quad (2.37)$$

with the backward difference motivated since  $k$  increases downward whereas  $\hat{\mathbf{z}}$  points upward. In the following, we choose to define the difference operator to only act on interface fields. Hence, any layer quantity, such as the layer thickness, commutes with the interface operator

$$\delta_k(h A_{k-1/2}) = h (A_{k-1/2} - A_{k+1/2}). \quad (2.38)$$

Also note that the thickness itself is the difference between the layer interfaces

$$h = \eta - \eta_b = \delta_k \eta_{k-1/2}, \quad (2.39)$$

where  $\eta_{1/2} = \eta$  and  $\eta_{3/2} = \eta_b$ .

With the above notation, the component momentum equations (2.35a) and (2.35b) take on the matrix-vector form

$$\begin{bmatrix} \partial_t(h u) - h f v \\ \partial_t(h v) + h f u \end{bmatrix} = - \begin{bmatrix} \partial_x & \partial_y & h^{-1} \delta_k \end{bmatrix} \begin{bmatrix} D_1^{(u)} & D_1^{(v)} & 0 \\ D_2^{(u)} & D_2^{(v)} & 0 \\ D_3^{(u)} & D_3^{(v)} & 0 \end{bmatrix}. \quad (2.40)$$

The  $3 \times 3$  matrix is a second order tensor with the first and second columns consisting of the layer thickness weighted momentum fluxes

$$\rho \mathbf{D}^{(u)} = (\rho h u^2 + P) \hat{\mathbf{x}} + \rho h u v \hat{\mathbf{y}} - p_{k-1/2} \partial_x \eta_{k-1/2} h \hat{\mathbf{z}} \quad (2.41a)$$

$$\rho \mathbf{D}^{(v)} = \rho h u v \hat{\mathbf{x}} + (\rho h v^2 + P) \hat{\mathbf{y}} - p_{k-1/2} \partial_y \eta_{k-1/2} h \hat{\mathbf{z}}, \quad (2.41b)$$

where we suppressed unnecessary layer indices. The horizontal flux components are given by minus the thickness weighted kinetic stress,  $\rho h \mathbf{u} \otimes \mathbf{u}$ , plus the vertically integrated contact pressure acting on the vertical sides of the shallow water column. The vertical flux component contains the pressure form stresses acting on the top and bottom interfaces, with these interfacial form stresses leading to the vertical transfer of horizontal form stresses across the layer boundaries. These fluxes allow us to write the thickness weighted zonal and meridional momentum equations as

$$\partial_t (h u) - v h f = -(\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} h^{-1} \delta_k) \cdot \mathbf{D}^{(u)} \quad (2.42a)$$

$$\partial_t (h v) + u h f = -(\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} h^{-1} \delta_k) \cdot \mathbf{D}^{(v)}. \quad (2.42b)$$

In this form of the momentum equation, contributions from contact stresses (kinetic stresses and form stresses) appear as the convergence of these stresses. Note that the divergence operator is built as combination of the continuous horizontal gradient operator along with a finite difference vertical operator. The third column of the tensor (2.40) is identically zero and so it can be readily dropped. However, we include it to connect with the [Eliassen-Palm flux](#) tensor as detailed by [Maddison and Marshall \(2013\)](#). We return to equations (2.42a) and (2.42b) in Section 2.4.9 for the stacked shallow water model.

### 2.3.4 Comments on the two pressure force formulations

The momentum equations (2.42a) and (2.42b) are written as a flux-form conservation law, with only the Coriolis force appearing as a body force. This formulation follows that for [Cauchy equation of motion](#) as studied in VOLUME 2. We make use of these flux-form momentum equations in Section 2.7 when discussing force balances in a zonally periodic channel, as well as in Chapter 14 where we formulate the thickness weighted averaged shallow water equations. Before doing so, we focus in Section 2.4 by further unpacking the contact force version of pressure as it appears in the shallow water model.

What has been gained by writing the momentum equation as the thickness weighted forms (2.42a) and (2.42b) versus the non-flux form velocity equation (2.19)? Indeed, the thickness weighted formulation is arguably less elegant and takes more effort to derive. A key reason we consider the thickness weighted equations is that they provide a venue to study how pressure contact forces alter momentum of a shallow water fluid column through interactions with the bottom, the surface, and adjacent vertical columns. We put this perspective to use in Section 2.7 when studying the force balances on a zonally reentrant channel. Additionally, in Section 2.4 we further pursue the contact force perspective by studying how pressure form stresses appear within a stacked shallow water model.

## 2.4 Contact pressure forces in shallow water layers

For a finite region of fluid, if the boundary area integrated contact pressure stress is nonzero, then pressure accelerates the region. In this section we study the physics and maths of contact pressure forces as they appear in the stacked shallow water model. As revealed by this study, the columnar motion of the shallow water fluid is fundamental to the analysis. Namely, the contact pressure approach is realized by studying the thickness weighted velocity equations of motion, which determine evolution of the momentum per horizontal area of a shallow water

fluid column. We introduced the thickness weighted approach in Section 2.3 for a single shallow water layer, and it led to the flux-form momentum equations (2.42a) and (2.42b). The single layer results are reproduced here for the stacked shallow water model, yet only after furthering our understanding of how pressure forces act to move momentum through shallow water layers.

### 2.4.1 Pressure contact force and pressure body force

As discussed in VOLUME 2, the connection between pressure body forces and pressure contact forces arise through an application of Gauss's divergence theorem to scalar fields<sup>4</sup>

$$\mathbf{F}_{\mathcal{R}}^{\text{press}} = - \int_{\mathcal{R}} \nabla p \, dV = - \oint_{\partial\mathcal{R}} p \hat{\mathbf{n}} \, d\mathcal{S}. \quad (2.43)$$

The first expression on the right hand side is a volume integral of the pressure gradient over the fluid region,  $\mathcal{R}$ . This expression provides the body force version of the pressure force. The second expression is a surface area integral over the region boundary,  $\partial\mathcal{R}$ , whose outward normal is  $\hat{\mathbf{n}}$ . This second expression provides the contact force version of the pressure force. Neither expression is more or less fundamental. Instead, they offer complementary insights into how pressure acts to modify the momentum of a fluid, with general notions of this complementarity the topic of a chapter in VOLUME 2. We here pursue the contact force perspective as a means to understand pressure form stress or interfacial form stress acting between layers of a shallow water fluid. There is also a pressure form stress acting between a fluid layer and the solid earth (topographic form stress), as well as between a fluid layer and the overlying atmosphere when the atmosphere has a non-zero mass (i.e., atmospheric form stress).

### 2.4.2 $N$ -layer equations

We start this section with the  $N$ -layer shallow water thickness and velocity equations derived in Section 1.4.2

$$\frac{\partial h_k}{\partial t} + \nabla \cdot (h_k \mathbf{u}_k) = 0 \quad \text{and} \quad [\partial_t + (\mathbf{u}_k \cdot \nabla)] \mathbf{u}_k + f \hat{z} \times \mathbf{u}_k = -(1/\rho_{\text{ref}}) \nabla p_k, \quad (2.44)$$

where  $k = 1, N$  is the layer index with no implied summation over this index,  $\rho_{\text{ref}}$  is the Boussinesq reference density (often chosen as  $\rho_{\text{ref}} = \rho_1$ ), and equation (1.82) gives the horizontal pressure gradient acceleration. Equations (2.44) are isomorphic to the single layer equations considered in Section 2.3.1. Hence, the thickness weighted velocity equation is a simple generalization of the single layer equation (2.21)

$$\frac{\partial(h_k \mathbf{u}_k)}{\partial t} + \nabla \cdot [h_k \mathbf{u}_k \otimes \mathbf{u}_k] + f \hat{z} \times (h_k \mathbf{u}_k) = -(h_k / \rho_{\text{ref}}) \nabla p_k, \quad (2.45)$$

where, again, there is no implied summation over the layer index,  $k$ . We commonly refer to the thickness weighted equation (2.45) as the momentum equation since  $\rho dx dy h_k \mathbf{u}_k$  is the horizontal momentum of a shallow water fluid column,

### 2.4.3 Contact pressure force along vertical sides

We now build up our understanding of pressure form stresses acting in a stacked shallow water fluid, with the essence of this discussion following that encountered for the single layer in

---

<sup>4</sup>See the vector calculus chapter in VOLUME 1.

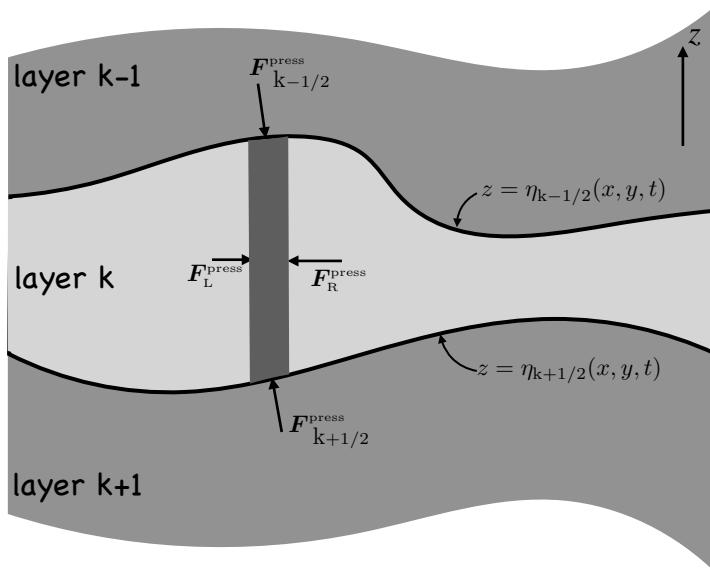


FIGURE 2.6: A schematic of the contact pressure force per area acting on the boundaries of a vertical column region within a shallow water layer of density  $\rho_k$ . Since fluid moves as vertical columns in a shallow water layer, we focus on the pressure forces acting on this column. The horizontal cross-sectional area of the column is vertically independent. The interface at the lower boundary is at the vertical position  $z = \eta_{k+1/2}$ , and the upper interface is at  $z = \eta_{k-1/2}$ . In accordance with Newton's third law, pressures are continuous across each of the  $\eta_{k\pm 1/2}$  layer interfaces so that the pressure forces are equal in magnitude yet oppositely directed on the opposite sides to the interfaces. The layer thickness is the difference between the interface positions,  $h_k = \eta_{k-1/2} - \eta_{k+1/2}$ . The boundaries of the columnar region feel a contact pressure force from the surrounding fluid that acts inward. The left side of the column experiences a pressure  $p_L$ ; the right side experiences  $p_R$ ; the upper interface has a pressure  $p_{k-1/2}$  acting between the layer  $k-1$  and layer  $k$ , and the lower interface has a pressure  $p_{k+1/2}$  acting between the layer  $k+1$  and layer  $k$ . The net pressure force acting on the column is computed as the area integral of the pressure acting around the full extent of the column boundaries. The [interfacial form stress](#) is the horizontal component of the stress.

Section 2.3.3. Our interest concerns the pressure acting on the boundaries of a fluid column within a shallow water layer, such as shown in Figure 2.6.

The pressure at a vertical position within the shallow water layer- $k$  is given by

$$p_k(z) = \rho_k g (\eta_{k-1/2} - z) + p_{k-1/2}. \quad (2.46)$$

Integrating this pressure over the layer thickness yields

$$P_k \equiv \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} p_k(z) dz \quad (2.47a)$$

$$= g \rho_k [\eta_{k-1/2} h_k - (1/2) (\eta_{k-1/2}^2 - \eta_{k+1/2}^2)] + p_{k-1/2} h_k \quad (2.47b)$$

$$= h_k (g \rho_k h_k / 2 + p_{k-1/2}). \quad (2.47c)$$

Since pressure is a linear function of  $z$  within a layer, the vertically averaged hydrostatic pressure within a layer,  $P_k/h_k$ , equals to the pressure at the upper interface,  $p_{k-1/2}$ , plus one-half the weight per area of the layer,  $g \rho_k h_k / 2$ .

The zonal pressure force acting on the column sides is the difference between the pressure integrated across the left and right zonal faces of the column. Assuming the fluid column to have an infinitesimal horizontal cross-sectional area  $dx dy$ , we find the zonal pressure force is

given by

$$dy \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} (p_L - p_R) dz = -dx dy \left[ (g/2) \rho_k \frac{\partial h_k^2}{\partial x} + \frac{\partial(h_k p_{k-1/2})}{\partial x} \right] \quad (2.48a)$$

$$= -dx dy \partial_x \left[ (g/2) \rho_k h_k^2 + h_k p_{k-1/2} \right] \quad (2.48b)$$

$$= -dx dy \partial_x P_k. \quad (2.48c)$$

The analogous result holds for the meridional direction, thus rendering the net contact pressure force acting on the vertical sides of the column

$$\mathbf{F}_{\text{sides}}^{\text{press}} = -dx dy \nabla P_k. \quad (2.49)$$

Hence, the contact force on the vertical sides of the column is given by the gradient of the layer vertically integrated pressure, with the vertical integral given by equation (2.47c). It is notable that this semi-discrete exercise reveals no more information than already contained within the integral theorem (2.43). Nonetheless, it is useful to see how the integral theorem manifests within discrete shallow water layers.

#### 2.4.4 Contact pressure force along the top and bottom interfaces

Now consider the contact pressure force acting on the top interface. This interface is generally sloped, so that the contact force has a component in both the vertical and horizontal directions. The vertical component to the pressure force maintains hydrostatic balance with the contact pressure at the lower boundary interface. The horizontal component provides a horizontal acceleration, with this acceleration (sign and magnitude) determined by the slope of the interface. Following our study of pressure form stress in VOLUME 2, the **interfacial form stress** is the name given to the horizontal component of the pressure stress acting on the sloped interface.

To mathematically characterize the pressure force on the top interface,  $z = \eta_{k-1/2}$ , requires the outward normal

$$\hat{\mathbf{n}}_{k-1/2} = \frac{\nabla(z - \eta_{k-1/2})}{|\nabla(z - \eta_{k-1/2})|} = \frac{\hat{\mathbf{z}} - \nabla\eta_{k-1/2}}{\sqrt{1 + (\nabla\eta_{k-1/2})^2}}. \quad (2.50)$$

Temporarily assume the interface slope to have a zero projection in the  $\hat{\mathbf{y}}$  direction. In this case, the outward normal is

$$\hat{\mathbf{n}}_{k-1/2} = \frac{\hat{\mathbf{z}} - \hat{\mathbf{x}} \partial_x \eta_{k-1/2}}{\sqrt{1 + (\partial_x \eta_{k-1/2})^2}} \quad (2.51a)$$

$$= \frac{\hat{\mathbf{z}} - \hat{\mathbf{x}} \tan \varphi_{k-1/2}}{\sqrt{1 + \tan^2 \varphi_{k-1/2}}} \quad (2.51b)$$

$$= (\hat{\mathbf{z}} - \hat{\mathbf{x}} \tan \varphi_{k-1/2}) \cos \varphi_{k-1/2}, \quad (2.51c)$$

where we defined the interface slope as

$$\frac{\partial \eta_{k-1/2}}{\partial x} = \tan \varphi_{k-1/2}, \quad (2.52)$$

with  $\varphi_{k-1/2}$  the angle between the horizontal plane and the interface. Trigonometry leads to an

expression for the area of the top of the column<sup>5</sup>

$$dS_{k-1/2} = \frac{dx dy}{\cos \varphi_{k-1/2}}, \quad (2.53)$$

so that the product of the area and the outward normal is given by

$$\hat{n}_{k-1/2} dS_{k-1/2} = dx dy (\hat{z} - \hat{x} \partial_x \eta_{k-1/2}). \quad (2.54)$$

This result generalizes to an interface slope that projects into both horizontal directions

$$\hat{n}_{k-1/2} dS_{k-1/2} = dx dy (\hat{z} - \nabla \eta_{k-1/2}), \quad (2.55)$$

so that the contact pressure force acting on layer-k at its top interface is given by

$$\mathbf{F}_{\text{top}}^{\text{press}} = -dx dy (\hat{z} - \nabla \eta_{k-1/2}) p_{k-1/2}. \quad (2.56)$$

Analogous considerations lead to the contact pressure force acting on layer-k at the bottom of the column

$$\mathbf{F}_{\text{bot}}^{\text{press}} = dx dy (\hat{z} - \nabla \eta_{k+1/2}) p_{k+1/2}. \quad (2.57)$$

#### 2.4.5 Form stress

As studied in VOLUME 2, pressure **form stress** is the horizontal projection of the contact pressure stress acting on the sloped top or bottom interface of the fluid layer. The corresponding forces acting on layer-k is the horizontal area element multiplied by the form stress

$$\mathbf{F}_{\text{top}}^{\text{form}} = dx dy (p_{k-1/2} \nabla \eta_{k-1/2}) \quad (2.58)$$

$$\mathbf{F}_{\text{bot}}^{\text{form}} = -dx dy (p_{k+1/2} \nabla \eta_{k+1/2}). \quad (2.59)$$

These forces render a mechanically reversible vertical exchange of horizontal momentum. This momentum exchange occurs without any exchange of matter. Rather, it an inviscid exchange that occurs according to Newton's third law (the action/reaction law).

For a specific case, consider a  $k - 1/2$  interface that slopes upward in the  $\hat{x}$  direction (e.g., see Figure 2.6). Form stress acting at the interface provides a  $+\hat{x}$  directed acceleration on the column. For the  $k + 1/2$  interface, a negatively sloped interface also experiences a  $+\hat{x}$  directed acceleration.

#### 2.4.6 Net contact pressure force on a shallow water column

Summing the contact pressure forces (2.49), (2.56), and (2.57), and dividing by the horizontal area, leads to the net pressure force per horizontal area acting on a column within layer-k

$$\frac{\mathbf{F}_{\text{net},k}^{\text{press}}}{dx dy} = -\nabla P_k - (\hat{z} - \nabla \eta_{k-1/2}) p_{k-1/2} + (\hat{z} - \nabla \eta_{k+1/2}) p_{k+1/2} \quad (2.60a)$$

$$= \rho_k g h_k \hat{z} - \nabla P_k + p_{k-1/2} \nabla \eta_{k-1/2} - p_{k+1/2} \nabla \eta_{k+1/2}. \quad (2.60b)$$

---

<sup>5</sup>Equation (2.53) was also found in VOLUME 2, as part of our derivation of the **kinematic boundary condition** for a material interface.

To reach this result we made use of the hydrostatic relation for the vertical pressure difference across a layer

$$p_{k+1/2} - p_{k-1/2} = \rho_k g h_k. \quad (2.61)$$

The vertical component of the net contact pressure force balances the weight of the column within the layer, which is expected since the shallow water fluid is in hydrostatic balance. The horizontal contact pressure force arises from a horizontal gradient plus the form stress at the surface and bottom interfaces. The gradient term is removed when integrating horizontally over the full domain given that the thickness of the layer vanishes upon reaching the coastlines. The resulting net force on the full domain arises just from the weight of the fluid acting in the vertical, plus form stress at the surface and bottom. We further discuss this point in Section 2.4.7.

#### 2.4.7 Contact pressure force summed over all layers

Summing the contact pressure force (2.60b) over all layers reveals the contact forces on the interior layer interfaces vanish, as per Newton's third law, thus leaving just the form stress at the surface and bottom and the contact pressure force acting on the vertical sides. Dividing by the horizontal area of the column leads to the net pressure force per area

$$\frac{1}{dx dy} \sum_{k=1}^N \mathbf{F}_{\text{net},k}^{\text{press}} = \hat{\mathbf{z}} g \sum_{k=1}^N \rho_k h_k + p_a \nabla \eta_{1/2} - p_b \nabla \eta_b - \sum_{k=1}^N \nabla P_k \quad (2.62a)$$

$$= (p_b - p_a) \hat{\mathbf{z}} + p_a \nabla \eta_{1/2} - p_b \nabla \eta_b - \sum_{k=1}^N \nabla P_k, \quad (2.62b)$$

where we wrote the total weight per area within the column as the difference between the bottom pressure and applied surface pressure

$$g \sum_{k=1}^N \rho_k h_k = p_b - p_a. \quad (2.63)$$

The horizontal components to the applied and bottom pressure terms in equation (2.62b) arise from pressure form stresses applied to the interfaces at the top and bottom of the column. The vertical component arises from the net weight per area of the fluid. The summation term is the horizontal gradient of the vertically integrated contact pressure applied along the vertical sides of the column.

#### 2.4.8 Horizontal pressure force with potential energy gradients

There is another means to express the horizontal pressure force. Here, we expose the gravitational potential energy per horizontal area for a column of fluid within a shallow water layer

$$\mathcal{P}_k = g \rho_k \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} z dz = (g \rho_k / 2) (\eta_{k-1/2}^2 - \eta_{k+1/2}^2). \quad (2.64)$$

Use of the layer gravitational potential energy brings the layer vertical integral of the hydrostatic pressure from Section 2.4.3 into

$$\int_{\eta_{k+1/2}}^{\eta_{k-1/2}} p_k(z) dz = g \rho_k h_k^2 / 2 + h_k p_{k-1/2} = \mathcal{P}_k - g \rho_k h_k \eta_{k+1/2} + h_k p_{k-1/2}. \quad (2.65a)$$

Making use of this result in equation (2.60b), along with a few lines of algebra, yields the net horizontal contact pressure force acting on a shallow water column

$$-\nabla P_k + \delta_k(p_{k-1/2} \nabla \eta_{k-1/2}) = -\nabla \mathcal{P}_k - \delta_k(\eta_{k-1/2} \nabla p_{k-1/2}). \quad (2.66)$$

To reach the identity (2.66) requires the hydrostatic relation,  $p_{k+1/2} - p_{k-1/2} = g \rho_k h_k$ , and the layer thickness,  $h_k = \eta_{k-1/2} - \eta_{k+1/2}$ . A consistency check notes that the curl of both sides to equation (2.66) are the same. The identity (2.66) suggests we define the form stress and its dual

$$\mathbf{F}^{\text{form}} = p \nabla \eta \quad \text{and} \quad \mathbf{F}^{\text{dual form}} = -\eta \nabla p, \quad (2.67)$$

with both  $\mathbf{F}^{\text{form}}$  and  $\mathbf{F}^{\text{dual form}}$  defined on layer interfaces. These two stresses have the same curl, and thus impart the same pressure torque on a column of fluid<sup>6</sup>

$$\nabla \times (p \nabla \eta) = \nabla \times (-\eta \nabla p). \quad (2.68)$$

However, these stresses are distinct and as such cannot be arbitrarily interchanged.<sup>7</sup>

## 2.4.9 Momentum equation with contact pressure forces

### Comparing the body force version and the contact force version

Recall that the thickness weighted velocity equation (2.45), as written in terms of the pressure gradient body force, is given by<sup>8</sup>

$$\frac{\partial(h_k \mathbf{u}_k)}{\partial t} + \nabla \cdot [h_k \mathbf{u}_k \otimes \mathbf{u}_k] + f \hat{z} \times (h_k \mathbf{u}_k) = -(h_k / \rho_{\text{ref}}) \nabla_h p_k, \quad (2.69)$$

again with no implied summation over the layer label,  $k$ . Alternatively, we can make use of the net contact pressure force (2.60b) so that

$$\frac{\partial(h_k \mathbf{u}_k)}{\partial t} + \nabla \cdot (h_k \mathbf{u}_k \otimes \mathbf{u}_k + \mathbf{I} P_k / \rho_{\text{ref}}) + f \hat{z} \times (h_k \mathbf{u}_k) = \delta_k(p_{k-1/2} \nabla \eta_{k-1/2}) / \rho_{\text{ref}}, \quad (2.70)$$

where  $\mathbf{I}$  is the identity tensor,  $P_k$  is the layer integrated pressure given by equation (2.47c), and  $\delta_k$  is the difference operator defined by equation (2.37). Choosing the contact pressure force as in equation (2.66) to expose the potential energy brings the momentum equation (2.70) into

<sup>6</sup>See Chapter 5 on shallow water vorticity for more details.

<sup>7</sup>As noted in our discussion of pressure form stress in VOLUME 2, much of the literature refers to  $-\eta \nabla p$  as the form stress rather than the dual form stress. This usage presumably originates from the common application of zonal averages for studying atmospheric motions, whereby  $\eta \partial_x p^x = -p \partial_x \eta^x$ . But this identity does not hold for arbitrary averaging operators, such as the ensemble averages commonly used for turbulence studies. So it is generally necessary to distinguish the form stress from the dual form stress.

<sup>8</sup>In equation (2.69) we wrote the gradient on the pressure as  $\nabla_h$  since we are only interested in the horizontal gradient acting on  $p_k(x, y, z)$ . All other objects in equation (2.69) are just a function of horizontal position within a layer, so that  $\nabla$  acting on them reduces to  $\nabla_h$ . Hence, the subscript on the gradient operator,  $\nabla_h$ , is exposed only when it acts on a function of  $z$ , such as for  $p_k(x, y, z)$ . Since  $p_k(x, y, z)$  is a linear function of  $z$ , its horizontal gradient is vertically independent within the layer, as illustrated in Figure 1.1.

the alternative form

$$\frac{\partial(h_k \mathbf{u}_k)}{\partial t} + \nabla \cdot [h_k \mathbf{u}_k \otimes \mathbf{u}_k + \mathbb{P}_k / \rho_{\text{ref}}] + f \hat{\mathbf{z}} \times (h_k \mathbf{u}_k) = -\delta_k (\eta_{k-1/2} \nabla p_{k-1/2}) / \rho_{\text{ref}}. \quad (2.71)$$

Equations (2.69), (2.70), and (2.71) allow us to identify the body force and contact force versions of the thickness weighted horizontal pressure acceleration

$$-h_k \nabla_h p_k = -\nabla P_k + \delta_k (p_{k-1/2} \nabla \eta_{k-1/2}) = -\nabla \mathcal{P}_k - \delta_k (\eta_{k-1/2} \nabla p_{k-1/2}). \quad (2.72)$$

### The balance of pressure torques acting on a shallow water column

A necessary (but not sufficient) check of the identity (2.72) can be found by verifying that the curl agrees for each expression

$$-\nabla \times \delta_k (\eta_{k-1/2} \nabla p_{k-1/2}) = \nabla \times \delta_k (p_{k-1/2} \nabla \eta_{k-1/2}) \quad (2.73a)$$

$$= \delta_k [\nabla \times (p_{k-1/2} \nabla \eta_{k-1/2})] \quad (2.73b)$$

$$= \delta_k [\nabla p_{k-1/2} \times \nabla \eta_{k-1/2}] \quad (2.73c)$$

$$= \nabla p_{k-1/2} \times \nabla \eta_{k-1/2} - \nabla (p_{k-1/2} + g \rho_k h_k) \times \nabla \eta_{k+1/2} \quad (2.73d)$$

$$= \nabla p_{k-1/2} \times \nabla h_k - g \rho_k \nabla h_k \times \nabla \eta_{k+1/2} \quad (2.73e)$$

$$= \nabla (p_{k-1/2} + g \rho_k \eta_{k+1/2}) \times \nabla h_k \quad (2.73f)$$

$$= \nabla (p_{k-1/2} - g \rho_k h_k + g \rho_k \eta_{k-1/2}) \times \nabla h_k \quad (2.73g)$$

$$= \nabla_h [p_{k-1/2} + g \rho_k (\eta_{k-1/2} - z)] \times \nabla h_k \quad (2.73h)$$

$$= \nabla_h p_k \times \nabla h_k \quad (2.73i)$$

$$= -\nabla \times (h_k \nabla_h p_k), \quad (2.73j)$$

which concurs with the curl of the left hand side of equation (2.72). To reach this result we set  $h_k = \eta_{k-1/2} - \eta_{k+1/2}$  and used equation (2.46) for the pressure within a shallow water layer:  $p_k(z) = \rho_k g (\eta_{k-1/2} - z) + p_{k-1/2}$ .

Anticipating our discussion of vorticity for the shallow water fluid in Chapter 5, we observe that the identity derived above,

$$\nabla \times \delta_k (p_{k-1/2} \nabla \eta_{k-1/2}) = -\nabla \times (h_k \nabla_h p_k), \quad (2.74)$$

says that the difference between the interfacial pressure torques acting on the top and bottom of a shallow water layer precisely balances minus the torque arising from the thickness weighted horizontal pressure gradient acting within the layer. This rather remarkable fine tuning of the interfacial and interior pressure torques is a direct consequence of assuming that the fluid motion is restricted to extensible vertical columns within each shallow water layer. This balance is not maintained within a three dimensional fluid, where fluid columns can generally tilt and bend.<sup>9</sup>

---

<sup>9</sup>See the discussion of vortex mechanics in Chapter 6.

### Layer summed momentum equation

Taking the vertical sum of the layer- $k$  momentum equation (2.70) leads to the column integrated horizontal momentum equation

$$\frac{\partial \mathbf{U}}{\partial t} + f \hat{\mathbf{z}} \times \mathbf{U} + \nabla \cdot \left[ \sum_{k=1}^N (h_k \mathbf{u}_k \otimes \mathbf{u}_k + \mathbf{I} P_k / \rho_{\text{ref}}) \right] = [p_a \nabla \eta_{1/2} - p_b \nabla \eta_b] / \rho_{\text{ref}}, \quad (2.75)$$

where  $\mathbf{U}$  is the vertically integrated horizontal velocity given by equation (2.10). The same vertical sum for equation (2.71) leads to

$$\frac{\partial \mathbf{U}}{\partial t} + f \hat{\mathbf{z}} \times \mathbf{U} + \nabla \cdot \left[ \sum_{k=1}^N (h_k \mathbf{u}_k \otimes \mathbf{u}_k + \mathbf{I} \mathcal{P}_k / \rho_1) \right] = [-\eta_{1/2} \nabla p_a + \eta_b \nabla p_b] / \rho_{\text{ref}}. \quad (2.76)$$

The right hand side of equation (2.75) exposes the pressure form stresses acting on the ocean surface and bottom, whereas the right hand side of equation (2.76) exposes the dual form stress acting on the ocean surface and bottom.

### Decomposing into vertically averaged and vertical deviation velocities

For detailed analyses of the vertically integrated (layer summed) momentum and vorticity budgets, it is of interest to introduce the vertical averaging operator along with the deviation from the average,

$$\bar{\Phi}^z = \frac{\sum_{k=1}^N h_k \Phi_k}{\sum_{k=1}^N h_k} \quad \text{and} \quad \Phi'_k = \Phi_k - \bar{\Phi}^z, \quad (2.77)$$

so that the vertically integrated kinetic stress in equation (2.75) is

$$\sum_{k=1}^N h_k \mathbf{u}_k \otimes \mathbf{u}_k = D [\bar{\mathbf{u}}^z \otimes \bar{\mathbf{u}}^z + \overline{\mathbf{u}' \otimes \mathbf{u}'}^z] \quad \text{where } D = \sum_{k=1}^N h_k. \quad (2.78)$$

The velocity,  $\mathbf{u}'_k$ , is the deviation of the layer- $k$  velocity from the vertically average velocity, and we refer to it as the [internal velocity](#), whereas the vertically averaged velocity,  $\bar{\mathbf{u}}^z$ , is the [external velocity](#).<sup>10</sup> The identity (2.78) reveals that the vertically integrated kinetic stress can be decomposed into a stress arising from internal-internal velocity interactions plus external-external velocity interactions. By construction, there are no cross-terms (i.e., no internal-external terms) appearing in this vertically integrated stress.

### Momentum fluxes

Following the single layer discussion in Section 2.3.3, we write the momentum equation (2.70) in the form

$$\frac{\partial(h u)}{\partial t} - v h f = -(\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} h^{-1} \delta_k) \cdot \mathbf{D}^{(u)} \quad (2.79a)$$

$$\frac{\partial(h v)}{\partial t} + u h f = -(\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} h^{-1} \delta_k) \cdot \mathbf{D}^{(v)}, \quad (2.79b)$$

<sup>10</sup>It is also common in the ocean physics literature to refer to  $\mathbf{u}'_k$  as the [baroclinic velocity](#) and  $\bar{\mathbf{u}}^z$  as the [barotropic velocity](#).

where we dropped the  $k$  label for brevity and introduced the layer momentum fluxes

$$\mathbf{D}^{(u)} = (h u^2 + P/\rho_{\text{ref}}) \hat{\mathbf{x}} + h u v \hat{\mathbf{y}} - (p_{k-1/2} \partial_x \eta_{k-1/2}/\rho_{\text{ref}}) h \hat{\mathbf{z}} \quad (2.80\text{a})$$

$$\mathbf{D}^{(v)} = h u v \hat{\mathbf{x}} + (h v^2 + P/\rho_{\text{ref}}) \hat{\mathbf{y}} - (p_{k-1/2} \partial_y \eta_{k-1/2}/\rho_{\text{ref}}) h \hat{\mathbf{z}}. \quad (2.80\text{b})$$

Likewise, we can write the momentum equation (2.71) in the component form

$$\frac{\partial(h u)}{\partial t} - v h f = -(\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} h^{-1} \delta_k) \cdot \mathbf{E}^{(u)} \quad (2.81\text{a})$$

$$\frac{\partial(h v)}{\partial t} + u h f = -(\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} h^{-1} \delta_k) \cdot \mathbf{E}^{(v)}, \quad (2.81\text{b})$$

where the dual layer momentum fluxes are given by

$$\mathbf{E}^{(u)} = (h u^2 + \mathcal{P}/\rho_{\text{ref}}) \hat{\mathbf{x}} + h u v \hat{\mathbf{y}} + (\eta_{k-1/2} \partial_x p_{k-1/2}/\rho_{\text{ref}}) h \hat{\mathbf{z}} \quad (2.82\text{a})$$

$$\mathbf{E}^{(v)} = h u v \hat{\mathbf{x}} + (h v^2 + \mathcal{P}/\rho_{\text{ref}}) \hat{\mathbf{y}} + (\eta_{k-1/2} \partial_y p_{k-1/2}/\rho_{\text{ref}}) h \hat{\mathbf{z}}. \quad (2.82\text{b})$$

Besides swapping the vertically integrated pressure,  $P$ , for the potential energy,  $\mathcal{P}$ , the dual momentum fluxes,  $\mathbf{E}^{(u)}$  and  $\mathbf{E}^{(v)}$ , make use of the dual form stress,  $\eta_{k-1/2} \nabla p_{k-1/2}$ , rather than the form stress,  $-p_{k-1/2} \nabla \eta_{k-1/2}$ . Upon performing an eddy-mean flow decomposition,<sup>11</sup> the eddy correlation portion of the fluxes (2.82a) and (2.82b) lead to the shallow water Eliassen-Palm flux, which fill the rows in the Eliassen-Palm flux tensor.

#### 2.4.10 Further reading

*Ward and Hogg* (2011) and *Barthel et al.* (2017) offer pedagogical treatments of the stacked shallow water equations in the context of idealized simulations that lend insight into the dynamical balances. *Maddison and Marshall* (2013) study the Eliassen-Palm flux tensor for continuously stratified quasi-geostrophy as well as the Boussinesq hydrostatic equations.

## 2.5 Energetics for a single layer

In this section we develop budgets for gravitational potential energy, kinetic energy, and mechanical energy for a single shallow water layer sitting on top of a non-flat bottom. Since the shallow water model has no internal energy, the total energy of the fluid is just that arising from the mechanical energy of the macroscopic motion. As part of this discussion we also consider the available potential energy (APE).

Motion within a shallow water layer occurs in vertical columns, so that we consider the energy of the layer integrated motion. The gravitational potential energy of a shallow water column is affected by vertical movement of the top and bottom of the column within the gravitational field. The kinetic energy is affected by pressure work, with this work, as seen in Sections 2.3 and 2.4, expressed either as a gradient body force or a contact force. Furthermore, pressure work leading to vertical motion manifests as buoyancy work. Our goal in this section is to study these energetic transformations and their mathematical expressions. We then extend the energetic analysis from the single layer to multiple layers in Section 2.6, though note that much of our work for the single layer is sufficient for multiple layers.

<sup>11</sup>We perform such decomposition, for example, when studying thickness weighted averaging of the shallow water equations in Chapter 14.

Before diving into details, we note that shallow water energetics can be derived from a layer integration of the continuously stratified Boussinesq ocean energy equations in VOLUME 2. That approach offers a somewhat more telescopic presentation relative to that given here. However, we choose to present the derivations in a manner that supports skills in manipulating the shallow water equations, and further exposes the physical concepts arising from the motion of shallow water fluid columns.

### 2.5.1 Gravitational potential energy

The gravitational potential energy per horizontal area of a shallow water fluid is given by<sup>12</sup>

$$\mathcal{P}^{\text{sw}} = g \rho \int_{\eta_b}^{\eta} z \, dz = \frac{g \rho}{2} (\eta^2 - \eta_b^2) = \rho g h \bar{\eta}, \quad (2.83)$$

where the final equality introduced the layer thickness,  $h$ , and average of the layer interface heights

$$h = \eta - \eta_b \quad \text{and} \quad \bar{\eta} = (\eta + \eta_b)/2. \quad (2.84)$$

Notice how the gravitational potential energy vanishes when  $\eta^2 = \eta_b^2$ . For the case  $\eta = \eta_b$ , there is no fluid in the column and so we expect the potential energy to vanish. For the case  $\eta = -\eta_b > 0$ , there is the same amount of fluid above  $z = 0$  as below, in which case the potential energy for the column vanishes since we are computing it relative to the  $z = 0$  reference state. Furthermore, in the flat bottom case,  $\eta_b = 0$  so that  $h = \eta - \eta_b = \eta$ , in which case the potential energy (2.83) reduces to  $\mathcal{P}_{\text{flat}}^{\text{sw}} = g \rho \eta^2 / 2$ .

#### Material time derivative of gravitational potential energy

Taking the material time derivative of the gravitational potential energy in equation (2.83) yields

$$\frac{D\mathcal{P}^{\text{sw}}}{Dt} = g \rho \left[ \eta \frac{D\eta}{Dt} - \eta_b \frac{D\eta_b}{Dt} \right] \quad (2.85a)$$

$$= g \rho [\eta w(\eta) - \eta_b w(\eta_b)] \quad (2.85b)$$

$$= g \rho h [w(\eta_b) - \eta \nabla \cdot \mathbf{u}] \quad (2.85c)$$

$$= g \rho h [w(\eta) - \eta_b \nabla \cdot \mathbf{u}], \quad (2.85d)$$

where we used equations for the vertical velocity component from Section 1.2.8, and for the final equality we used equation (1.96) to write

$$w(\eta) - w(\eta_b) = -h \nabla \cdot \mathbf{u}. \quad (2.86)$$

Evidently, the potential energy changes according to how the thickness of the layer increases through vertical motion along the top and bottom interfaces, and as weighted by the position of these interfaces relative to  $z = 0$ . Finally, it is useful to write equation (2.85d) in its flux-form, which is given by

$$\partial_t \mathcal{P}^{\text{sw}} + \nabla \cdot (\mathbf{u} \mathcal{P}^{\text{sw}}) = \mathcal{P}^{\text{sw}} \nabla \cdot \mathbf{u} + g \rho h [w(\eta) - \eta_b \nabla \cdot \mathbf{u}]. \quad (2.87)$$

<sup>12</sup>We include the “sw” superscript to distinguish the shallow water energetic terms, which we consider in their thickness weighted form so their dimensions are energy per area. In other areas of this book, we consider the energy per mass, such as in our study of energetics of the continuously stratified fluid in VOLUME 2.

The source term on the right hand side can be written as a buoyancy work term, which we show next.

### Exposing the role of buoyancy work

An alternative expression for the potential energy evolution is found by working with the final expression for potential energy in equation (2.83), whose material time derivative is

$$\frac{D\mathcal{P}^{sw}}{Dt} = g \rho \bar{\eta} \frac{Dh}{Dt} + g \rho h \frac{D\bar{\eta}}{Dt} = -\mathcal{P}^{sw} \nabla \cdot \mathbf{u} + \rho g h \bar{w}, \quad (2.88)$$

where we introduced the averaged vertical velocity for the layer according to equation (1.102)

$$\bar{w} = \frac{D\bar{\eta}}{Dt} = \frac{1}{2} \frac{D(\eta + \eta_b)}{Dt}. \quad (2.89)$$

The flux-form version of the potential energy equation (2.88) thus takes on the form

$$\partial_t \mathcal{P}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{P}^{sw}) = g \rho h \bar{w}. \quad (2.90)$$

This equation is the shallow water analog to the gravitational potential energy budget studied in VOLUME 2 for a continuously stratified fluid. In particular, we see that the buoyancy work term,  $g \rho h \bar{w}$ , alters potential energy when there is vertical motion through the gravity field. As a check on the manipulations, we verify that the potential energy budgets (2.87) and (2.90) are indeed self-consistent by noting that

$$\mathcal{P}^{sw} \nabla \cdot \mathbf{u} + \rho g h [w(\eta) - \eta_b \nabla \cdot \mathbf{u}] = \rho g h [(\bar{\eta} - \eta_b) \nabla \cdot \mathbf{u} + w(\eta)] \quad (2.91a)$$

$$= \rho g h [\frac{1}{2}h \nabla \cdot \mathbf{u} + w(\eta)] \quad (2.91b)$$

$$= \rho g h [-\frac{1}{2}Dh/Dt + w(\eta)] \quad (2.91c)$$

$$= \rho g h \bar{w}. \quad (2.91d)$$

### 2.5.2 Kinetic energy and work from pressure gradients

The kinetic energy per horizontal area is

$$\mathcal{K}^{sw} = \frac{1}{2} \int_{\eta_b}^{\eta} \rho \mathbf{u} \cdot \mathbf{u} dz = \frac{1}{2} \rho h \mathbf{u} \cdot \mathbf{u}. \quad (2.92)$$

Its material time derivative is given by

$$\frac{D\mathcal{K}^{sw}}{Dt} = \rho h \mathbf{u} \cdot \frac{D\mathbf{u}}{Dt} + \frac{1}{2} \rho \mathbf{u} \cdot \mathbf{u} \frac{Dh}{Dt} \quad (2.93a)$$

$$= -h \mathbf{u} \cdot \nabla p + \rho h \mathbf{u} \cdot \mathbf{F} + \frac{\mathcal{K}^{sw}}{h} \frac{Dh}{Dt} \quad (2.93b)$$

$$= -h \mathbf{u} \cdot \nabla p + \rho h \mathbf{u} \cdot \mathbf{F} - \mathcal{K}^{sw} \nabla \cdot \mathbf{u}, \quad (2.93c)$$

where we made use of the velocity equation (1.9) along with the addition of a frictional acceleration,  $\mathbf{F}$  (see Section 1.6.5), and we used the thickness equation (1.20). Rearrangement of equation (2.93c) leads to the flux-form budget for layer integrated kinetic energy

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{K}^{sw}) = -h \mathbf{u} \cdot \nabla p + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.94)$$

The first term on the right hand side is the projection of the horizontal velocity onto the horizontal pressure gradient acceleration, thus indicating that kinetic energy for the fluid column increases if the velocity has a component that is directed down the horizontal pressure gradient. This term arises from the work done by the horizontal pressure gradient force acting on the moving fluid columns. The second right hand side term is the projection of the velocity onto the thickness weighted horizontal friction, which accounts for work done by friction and/or boundary stresses on the moving fluid.

For a slight modification to the budget equation (2.94), write the layer pressure as

$$p = p_a + \rho g (\eta - z) = \rho g (\eta^{\text{eff}} - z), \quad (2.95)$$

where we introduced the effective free surface from equation (1.4)

$$\eta^{\text{eff}} = \eta + p_a / (\rho g). \quad (2.96)$$

The thickness weighted pressure work thus takes the form

$$-h \mathbf{u} \cdot \nabla p = -h \rho g \mathbf{u} \cdot \nabla \eta^{\text{eff}}, \quad (2.97)$$

so that the kinetic energy equation (2.94) becomes

$$\partial_t \mathcal{K}^{\text{sw}} + \nabla \cdot [\mathbf{u} (\mathcal{K}^{\text{sw}} + \rho g h \eta^{\text{eff}})] = \rho g \eta^{\text{eff}} \nabla \cdot (h \mathbf{u}) + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.98)$$

### 2.5.3 Kinetic energy and buoyancy work

Following the second formulation of gravitational potential energy in Section 2.5.1, we here expose the buoyancy work that is contained in the term  $\rho g \eta^{\text{eff}} \nabla \cdot (h \mathbf{u})$  appearing in equation (2.98). For this purpose, make use of the thickness equation,  $\partial_t h = -\nabla \cdot (\mathbf{u} h)$ , to write

$$\rho g \eta^{\text{eff}} \nabla \cdot (h \mathbf{u}) = -\rho g \eta^{\text{eff}} \partial_t h = -\rho g \eta^{\text{eff}} \partial_t (\delta_k \eta_{k-1/2}) = -\delta_k (\rho g \eta^{\text{eff}} \partial_t \eta_{k-1/2}). \quad (2.99)$$

For the final two equations we introduced the layer index, with  $k = 1$  for the single layer and with the layer interfaces  $\eta_{1/2} = \eta$  and  $\eta_{3/2} = \eta_b$ . Additionally,  $\delta_k$  is the finite difference operator (2.37) so that

$$h = \eta - \eta_b = \eta_{1/2} - \eta_{3/2} = \delta_k \eta_{k-1/2}. \quad (2.100)$$

We inserted  $\rho g \eta^{\text{eff}}$  into the difference operator in equation (2.99) since this term is vertically independent across the layer.

For the next step, we make use of equation (2.95) for the layer pressure,  $p = \rho g (\eta^{\text{eff}} - z)$ , in which case<sup>13</sup>

$$-\delta_k (\rho g \eta^{\text{eff}} \partial_t \eta_{k-1/2}) = -\delta_k [(p + \rho g z) \partial_t \eta_{k-1/2}] \quad (2.101a)$$

$$= -\delta_k (p_{k-1/2} \partial_t \eta_{k-1/2}) - \delta_k (\rho g \eta_{k-1/2} \partial_t \eta_{k-1/2}). \quad (2.101b)$$

This step is somewhat subtle since we replaced  $\rho g \eta^{\text{eff}}$ , which is vertically independent within a layer, with the sum  $p + \rho g z$ , where both  $p$  and  $\rho g z$  are functions of  $z$ . For equation (2.101b) we replaced  $p$  and  $z$  with their interface values since the argument of the difference operator is evaluated on the layer interfaces. Observe that  $\partial_t \eta_b = 0$ , and yet it is convenient to carry this

<sup>13</sup>Equations (2.101a) and (2.101b) are inspired by some unpublished notes from Christopher Wolfe.

term through the manipulations to retain symmetry of the equations, and to anticipate the same formulation for the stacked shallow water energetics in Section 2.6.

We now make use of the potential energy in Section 2.5.1 by writing

$$-\delta_k(\rho g \eta_{k-1/2} \partial_t \eta_{k-1/2}) = -(\rho g/2) \partial_t (\eta^2 - \eta_b^2) = -\partial_t \mathcal{P}^{sw} = \nabla \cdot (\mathbf{u} \mathcal{P}^{sw}) - g \rho h \bar{w}, \quad (2.102)$$

where the final step used the potential energy equation (2.90). As advertised, this formulation exposes the buoyancy work term,  $-g \rho h \bar{w}$ , which then allows us to bring the kinetic energy equation (2.98) to the form

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot [\mathbf{u} (\mathcal{K}^{sw} - \mathcal{P}^{sw} + \rho g h \eta^{\text{eff}})] + \delta_k(p_{k-1/2} \partial_t \eta_{k-1/2}) = -g \rho h \bar{w} + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.103)$$

We can simplify the advective flux by writing

$$-\mathcal{P}^{sw} + \rho g h \eta^{\text{eff}} = -\rho g h \bar{w} + \rho g \eta + h p_a = h (\rho g h/2 + p_a) = P, \quad (2.104)$$

where the final equality introduced the layer integrated pressure

$$P = \int_{\eta_b}^{\eta} p(z) dz = \rho g \int_{\eta_b}^{\eta} (\eta^{\text{eff}} - z) dz = h (\rho g h/2 + p_a). \quad (2.105)$$

The kinetic energy equation (2.103) thus takes the form

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot [\mathbf{u} (\mathcal{K}^{sw} + P)] + \delta_k(p_{k-1/2} \partial_t \eta_{k-1/2}) = -g \rho h \bar{w} + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.106)$$

The term  $\delta_k(p_{k-1/2} \partial_t \eta_{k-1/2})$  on the left hand side is a vertical transfer that, as seen in Section 2.5.4, arises in part from pressure form stresses acting on the boundary of the fluid column.

#### 2.5.4 Kinetic energy and pressure form stress

As a final form of the kinetic energy equation, we recombine the vertical transfer term and the buoyancy work term in equation (2.106) to have

$$-g \rho h \bar{w} - \delta_k(p_{k-1/2} \partial_t \eta_{k-1/2}) = -g \rho h \bar{w} - p_a [w(\eta) - \mathbf{u} \cdot \nabla \eta] + p_b [w(\eta_b) - \mathbf{u} \cdot \nabla \eta_b]. \quad (2.107)$$

Introducing the bottom pressure and vertically integrated pressure

$$p_b = p_a + \rho g h \quad \text{and} \quad P = h (p_a + \frac{1}{2} \rho g h), \quad (2.108)$$

leads to

$$-g \rho h \bar{w} - \delta_k(p_{k-1/2} \partial_t \eta_{k-1/2}) = (P/h) [w(\eta_b) - w(\eta)] + \mathbf{u} \cdot (p_a \nabla \eta - p_b \nabla \eta_b) \quad (2.109a)$$

$$= \frac{P}{h} \frac{D(\eta_b - \eta)}{Dt} + \mathbf{u} \cdot (p_a \nabla \eta - p_b \nabla \eta_b) \quad (2.109b)$$

$$= -\frac{P}{h} \frac{Dh}{Dt} + \mathbf{u} \cdot (p_a \nabla \eta - p_b \nabla \eta_b) \quad (2.109c)$$

$$= P \nabla \cdot \mathbf{u} + \mathbf{u} \cdot (p_a \nabla \eta - p_b \nabla \eta_b), \quad (2.109d)$$

so that the kinetic energy equation (2.106) can be written

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{K}^{sw}) = \mathbf{u} \cdot (-\nabla P + p_a \nabla \eta - p_b \nabla \eta_b) + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.110)$$

We have thus exposed the work done by pressure form stresses acting on the vertical side of an expanding or contracting shallow water column, plus those form stresses acting on the top and bottom layer interfaces.

A somewhat more direct way to derive the kinetic energy equation (2.110) is to return to the kinetic energy equation (2.94) and make use of the identity (2.72). This identity equates the thickness weighted horizontal pressure gradient acting on a shallow water column, to the pressure form stresses acting over the boundary of the column, and for a single shallow water layer this identity is

$$-h \nabla p = -\nabla P + p_a \nabla \eta - p_b \nabla \eta_b. \quad (2.111)$$

Nonetheless, we chose the more tedious derivation as it provides a check on the correctness of the transfer term and the buoyancy work term in the kinetic energy budget (2.106).

### 2.5.5 Mechanical energy budget

The mechanical energy per horizontal area for the shallow water layer is given by

$$\mathcal{M}^{sw} = \mathcal{K}^{sw} + \mathcal{P}^{sw} = (\rho/2) [h \mathbf{u} \cdot \mathbf{u} + g (\eta^2 - \eta_b^2)] = \rho h (\frac{1}{2} \mathbf{u} \cdot \mathbf{u} + g \bar{\eta}). \quad (2.112)$$

To form a budget equation for the mechanical energy we simply add the budgets for the gravitational potential energy and kinetic energy.

#### Summary of the potential and kinetic energy budgets

We here summarize the flux-form budgets for potential energy and kinetic energy (equations (2.90), (2.94), (2.98), (2.106), and (2.110)):

$$\partial_t \mathcal{P}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{P}^{sw}) = g \rho h \bar{w} \quad (2.113a)$$

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{K}^{sw}) = -h \mathbf{u} \cdot \nabla p + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.113b)$$

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot [\mathbf{u} (\mathcal{K}^{sw} + \rho g h \eta^{\text{eff}})] = \rho g \eta^{\text{eff}} \nabla \cdot (h \mathbf{u}) + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.113c)$$

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot [\mathbf{u} (\mathcal{K}^{sw} + P)] + \delta_k (p_{k-1/2} \partial_t \eta_{k-1/2}) = -g \rho h \bar{w} + \rho h \mathbf{u} \cdot \mathbf{F} \quad (2.113d)$$

$$\partial_t \mathcal{K}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{K}^{sw}) = \mathbf{u} \cdot (-\nabla P + p_a \nabla \eta - p_b \nabla \eta_b) + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.113e)$$

#### Example forms of the mechanical energy budget

Adding equations (2.113a) and (2.113b) leads to the mechanical energy budget

$$\partial_t \mathcal{M}^{sw} + \nabla \cdot (\mathbf{u} \mathcal{M}^{sw}) = g \rho h \bar{w} - h \mathbf{u} \cdot \nabla p + \rho h \mathbf{u} \cdot \mathbf{F}, \quad (2.114)$$

whereas the sum of equations (2.113a) and (2.113d) leads to

$$\partial_t \mathcal{M}^{sw} + \nabla \cdot [\mathbf{u} (\mathcal{M}^{sw} + P)] + \delta_k (p_{k-1/2} \partial_t \eta_{k-1/2}) = \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.115)$$

As we discuss in Section 2.6, equations (2.114) and (2.115) hold also for the mechanical energy in a  $N$ -layer shallow water model.

#### Specializing to the single layer

Making use of the identity (see equation (2.104))

$$\mathcal{M}^{sw} + P = \rho h (\frac{1}{2} \mathbf{u} \cdot \mathbf{u} + g \eta^{\text{eff}}), \quad (2.116)$$

brings the mechanical energy budget (2.115) to the form

$$\partial_t \mathcal{M}^{\text{sw}} + \nabla \cdot [h \mathbf{u} (\frac{1}{2} \rho \mathbf{u} \cdot \mathbf{u} + \rho g \eta + p_a)] + \delta_k (p_{k-1/2} \partial_t \eta_{k-1/2}) = \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.117)$$

Further note that  $\partial_t \eta_b = 0$  so that

$$\nabla \cdot (h \mathbf{u} p_a) + \delta_k (p_{k-1/2} \partial_t \eta_{k-1/2}) = \nabla \cdot (h \mathbf{u} p_a) + p_a \partial_t \eta = h \mathbf{u} \cdot \nabla p_a, \quad (2.118)$$

where  $\partial_t \eta + \nabla \cdot (h \mathbf{u}) = 0$ . We are thus led to single layer mechanical energy budget

$$\partial_t \mathcal{M}^{\text{sw}} + \nabla \cdot [h \mathbf{u} (\frac{1}{2} \mathbf{u} \cdot \mathbf{u} + g \eta)] = -h \mathbf{u} \cdot \nabla p_a + \rho h \mathbf{u} \cdot \mathbf{F}. \quad (2.119)$$

which proves of use when studying the mechanical energy of shallow water waves in VOLUME 4.

### Shallow water form of the Bernoulli theorem

It is notable that in the steady state of an unforced perfect shallow water fluid (so with  $p_a = 0$  and  $\mathbf{F} = 0$ ), equation (2.117) becomes

$$\mathbf{u} \cdot \nabla (\frac{1}{2} \mathbf{u} \cdot \mathbf{u} + g \eta) = 0, \quad (2.120)$$

which is an expression of [Bernoulli's theorem](#) for steady shallow water flow (see VOLUME 2 for the continuously stratified version).

### Domain integrated mechanical energy

The budget equations (2.114) or (2.115) have already been layer integrated. So to study the domain integrated energetics requires only an area integral. The budget equation (2.115) is ideally suited for this purpose since all terms, except the non-conservative acceleration (e.g., friction), are written as a flux divergence and so they represent transfer processes. The domain integral of the flux-form mechanical energy equation (2.115) is

$$\int \partial_t \mathcal{M}^{\text{sw}} d\mathcal{S} = - \oint_{\partial\mathcal{S}} (\mathcal{M}^{\text{sw}} + P) \mathbf{u} \cdot \hat{\mathbf{n}} dl + \int_{\mathcal{S}} (-p_a \partial_t \eta + h \rho \mathbf{u} \cdot \mathbf{F}) d\mathcal{S}, \quad (2.121)$$

where we set  $\partial_t \eta_b = 0$ . The boundary integral on the right hand side is computed as a line integral around the edge of the layer. We consider three options for the layer geometry as illustrated in Figure 2.7. First, the thickness vanishes at the edge of the domain as in the case of shorelines, in which case the boundary integral vanishes since each term in the integral is thickness weighted (and thickness vanishes at the shoreline edge). Second, we assume the layer is bounded by vertical sidewalls, in which case  $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$  at the sidewall boundaries. Third, the domain has periodicity in one or both directions (e.g., a zonal channel), in which case the boundary integral again vanishes in the periodic directions.

Evidently, for either of the three types of domain boundaries considered above, the boundary integral in equation (2.121) vanishes, in which case the domain integrated mechanical energy budget reduces to

$$\int_{\mathcal{S}} \partial_t \mathcal{M}^{\text{sw}} d\mathcal{S} = \int_{\mathcal{S}} (-p_a \partial_t \eta + h \rho \mathbf{u} \cdot \mathbf{F}) d\mathcal{S} = - \int_{\mathcal{S}} h \mathbf{u} \cdot (\nabla p_a - \rho \mathbf{F}) d\mathcal{S}. \quad (2.122)$$

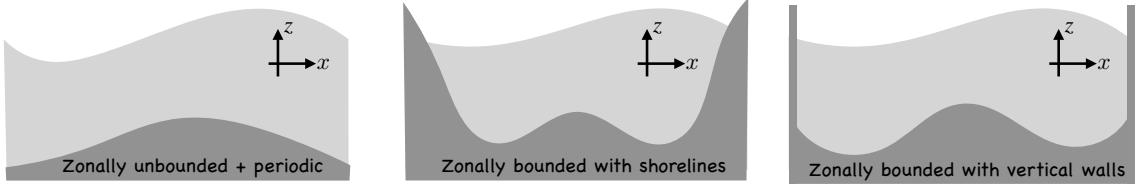


FIGURE 2.7: Three canonical zonal topologies/geometries considered in the study of fluid flow, particularly ocean flows. The light portion of each panel represents the fluid whereas the darker portion is the solid earth bottom topography. Left panel: zonally unbounded and periodic channel. Here, the topography, surface boundary forcing, and flow are zonally periodic. Middle panel: zonally bounded region where the zonal bounds occur along sloping shorelines at which the fluid thickness vanishes. The horizontal position of the vanishing thickness is time dependent since the fluid can move up and down the shoreline. Right panel: zonally bounded region where the fluid encounters a vertical sidewall so that the horizontal position of the fluid boundary is fixed, and so there is no horizontal position where the fluid thickness vanishes. Fixed vertical sidewall boundaries are commonly found in numerical model simulations and tacitly assumed in many theoretical treatments. Even so, vertical sidewall boundaries are uncommon in Nature.

To reach the second equality we wrote

$$-\int_S p_a \partial_t \eta \, dS = \int_S p_a \nabla \cdot (h \mathbf{u}) \, dS = -\int_S h \mathbf{u} \cdot \nabla p_a \, dS, \quad (2.123)$$

where we set

$$\int \nabla \cdot (p_a h \mathbf{u}) \, dS = 0, \quad (2.124)$$

which follows from the same reasoning used for the boundary integral in equation (2.121). Evidently, there is a nonzero domain integrated atmospheric pressure work only if there is a nonzero time tendency for the layer thickness. Furthermore, if  $p_a$  is a spatial constant, volume conservation for the full layer means that

$$\int_S \partial_t \eta \, dS = -\int_S \nabla \cdot (h \mathbf{u}) \, dS = 0, \quad (2.125)$$

in which case we find that equation (2.122) is indeed self-consistent.

If the domain is bounded by sloping sides, then the mechanical energy vanishes at the horizontal boundaries, so that the time derivative acting on  $\mathcal{M}^{sw}$  commutes with the area integral on the left hand side of equation (2.121). Alternatively, if the layer is bounded by fixed vertical walls, or is periodic in one or both directions, then the horizontal domain boundaries are static, again meaning that the time derivative commutes with the horizontal integral. In each case we can write the domain integrated mechanical energy equation as

$$\frac{d}{dt} \int_S \mathcal{M}^{sw} \, dS = -\int_S h \mathbf{u} \cdot (\nabla p_a - \rho \mathbf{F}) \, dS. \quad (2.126)$$

In the absence of work on the layer from atmospheric pressure, then the total mechanical energy is affected only via non-conservative forces, such as those from viscous friction. It then follows that for a perfect and unforced shallow water fluid, then the domain integrated mechanical energy remains constant, in which case there is an exact exchange between the domain integrated gravitational potential energy and kinetic energy.

### 2.5.6 Available potential energy

As discussed in VOLUME 2, a huge portion of the gravitational potential energy is not realizable as kinetic energy, merely because the minimum potential energy state is when the fluid is at rest with some fluid parcels sitting above others. [Available potential energy](#) measures the gravitational potential energy that can be converted to kinetic energy through a reversible rearrangement of the fluid. We here display the available potential energy for a single shallow water layer in a simply connected domain, thus specializing the available potential energy considered in VOLUME 2 for the continuously stratified Boussinesq ocean.

Taking  $z = 0$  as the reference level, the domain integrated gravitational potential energy for a single shallow water layer is

$$\int \mathcal{P}^{\text{sw}} d\mathcal{S} = g \rho \int d\mathcal{S} \int_{\eta_b}^{\eta} z dz = \frac{g \rho}{2} \int (\eta^2 - \eta_b^2) d\mathcal{S}. \quad (2.127)$$

We define a background or reference state potential energy as the potential energy contained in the fluid at rest, so that the free surface interface has its uniform area average value. Write the area average free surface height as

$$\langle \eta \rangle = \frac{1}{A} \int \eta d\mathcal{S}, \quad (2.128)$$

where

$$A = \int d\mathcal{S} \quad (2.129)$$

is the horizontal area integral over the full domain of the fluid. Hence, the reference state gravitational potential energy is realized when the surface height is flat at  $z = \langle \eta \rangle$ , so that

$$\int \mathcal{P}_{\text{ref}}^{\text{sw}} d\mathcal{S} = \frac{g \rho}{2} \int (\langle \eta \rangle^2 - \eta_b^2) d\mathcal{S}. \quad (2.130)$$

The available potential energy is defined by the difference

$$E_{\text{APE}} = \int (\mathcal{P}^{\text{sw}} - \mathcal{P}_{\text{ref}}^{\text{sw}}) d\mathcal{S} = \frac{g \rho}{2} \int (\eta^2 - \langle \eta \rangle^2) d\mathcal{S} = \frac{g \rho}{2} \int (\eta')^2 d\mathcal{S} \geq 0, \quad (2.131)$$

where

$$\eta' = \eta - \langle \eta \rangle \quad (2.132)$$

is the anomalous free surface. Note how the bottom topography cancelled out from  $E_{\text{APE}}$  since  $\eta_b^2$  appears in both  $\mathcal{P}^{\text{sw}}$  and  $\mathcal{P}_{\text{ref}}^{\text{sw}}$ . Also note that to reach the final equality in equation (2.131) required the identity

$$\int (\eta - \langle \eta \rangle)^2 d\mathcal{S} = \int (\eta^2 - 2 \langle \eta \rangle \eta + \langle \eta \rangle^2) d\mathcal{S} \quad (2.133a)$$

$$= \int \eta^2 d\mathcal{S} + A \langle \eta \rangle^2 - 2 \langle \eta \rangle \int \eta d\mathcal{S} \quad (2.133b)$$

$$= \int \eta^2 d\mathcal{S} - \langle \eta \rangle^2 A \quad (2.133c)$$

$$= \int (\eta^2 - \langle \eta \rangle^2) d\mathcal{S}. \quad (2.133d)$$

Equation (2.131) shows that the available potential energy is non-negative for the shallow water layer. That is, any slope to the shallow water layer represents a store of positive available potential energy. We derive the available potential energy in Section 2.6.1 for the  $N$ -layer shallow water model, showing that it too is non-negative.

## 2.6 Energetics for $N$ layers

We here generalize the single layer mechanical energy analysis from Section 2.5 to  $N$ -layers, making use of the  $N$ -layer equations from Section 1.4.

### 2.6.1 Potential energy and available potential energy

Here we develop the equation for potential energy in a form that naturally leads to the expression for the available potential energy.

#### Potential energy for $N$ -layers

To derive the gravitational potential energy per horizontal area in a stacked shallow water model, first consider the case with  $N = 3$  (Figure 1.6), in which the potential energy per horizontal area is

$$\mathcal{P}^{\text{sw}} = g \int_{\eta_b}^{\eta} z \rho dz = g \rho_3 \int_{\eta_b}^{\eta_{5/2}} z dz + g \rho_2 \int_{\eta_{5/2}}^{\eta_{3/2}} z dz + g \rho_1 \int_{\eta_{3/2}}^{\eta} z dz, \quad (2.134)$$

which then leads to

$$2 \mathcal{P}^{\text{sw}} = g \rho_3 (\eta_{5/2}^2 - \eta_b^2) + g \rho_2 (\eta_{3/2}^2 - \eta_{5/2}^2) + g \rho_1 (\eta^2 - \eta_{3/2}^2) \quad (2.135a)$$

$$= g \eta_{5/2}^2 (\rho_3 - \rho_2) + g \eta_{3/2}^2 (\rho_2 - \rho_1) + g \eta_{1/2}^2 \rho_1 - g \rho_3 \eta_b^2 \quad (2.135b)$$

$$= \rho_{\text{ref}} (g_{5/2}^r \eta_{5/2}^2 + g_{3/2}^r \eta_{3/2}^2 + g_{1/2}^r \eta_{1/2}^2) - g \rho_3 \eta_b^2, \quad (2.135c)$$

where  $g_{1/2}^r = g$  and  $\eta_{1/2} = \eta$ . Generalizing this expression to an arbitrary number of layers leads to

$$\mathcal{P}^{\text{sw}} = \frac{1}{2} \left[ -g \rho_N \eta_b^2 + \rho_{\text{ref}} \sum_{k=0}^{N-1} g_{k+1/2}^r \eta_{k+1/2}^2 \right], \quad (2.136)$$

with  $\eta_{N+1/2} = \eta_b$  (see Figure 1.6). The first term is a constant and so does not contribute to time changes of the potential energy, and so it is commonly ignored in the literature.

#### Available potential energy for $N$ -layers

The time dependent terms in equation (2.136) are positive definite, so that potential energy increases when the reduced gravity increases and/or the layer interfaces deviate in either direction from their resting values. This behavior motivates us to introduce the available potential energy by decomposing the layer interface heights into their area mean and deviations, just as for the single layer in Section 2.5.6

$$\eta'_{k\pm 1/2} = \eta_{k\pm 1/2} - \langle \eta_{k\pm 1/2} \rangle. \quad (2.137)$$

Volume conservation for each layer implies that the area mean,  $\langle \eta_{k+1/2} \rangle$ , is a space and time constant. Substituting  $\eta_{k+1/2} = \langle \eta_{k+1/2} \rangle + \eta'_{k+1/2}$  into equation (2.136) leads to

$$\begin{aligned} \mathcal{P}^{\text{sw}} &= \frac{1}{2} \left[ -g \rho_N \eta_b^2 + \rho_{\text{ref}} \sum_{k=0}^{N-1} g_{k+1/2}^r \langle \eta_{k+1/2} \rangle^2 \right] \\ &\quad + \frac{\rho_{\text{ref}}}{2} \left[ 2 \sum_{k=0}^{N-1} g_{k+1/2}^r \langle \eta_{k+1/2} \rangle \eta'_{k+1/2} + \sum_{k=0}^{N-1} g_{k+1/2}^r (\eta'_{k+1/2})^2 \right]. \end{aligned} \quad (2.138)$$

The first bracketed term is a constant in time, and it measures the potential energy of the system when all interfaces sit at their area mean values, for which  $\eta'_{k+1/2} = 0$ . When performing an area integral over the full domain, the term with  $\langle \eta_{k+1/2} \rangle \eta'_{k+1/2}$  vanishes since the area integral of  $\eta'_{k+1/2}$  vanishes. We thus define the available potential energy for the  $N$ -layer shallow water fluid

$$E_{\text{APE}} = \frac{\rho_{\text{ref}}}{2} \int \sum_{k=0}^{N-1} g_{k+1/2}^r (\eta'_{k+1/2})^2 dS \geq 0. \quad (2.139)$$

This expression generalizes the single layer form given by equation (2.131). As expected, the available potential energy increases whether a layer interface moves up or down, and it vanishes when all interfaces heights equal to their area mean values.

### Time evolution of the potential energy

To derive an evolution equation for the potential energy, we take the time derivative of equation (2.136)

$$\partial_t \mathcal{P}^{\text{sw}} = \rho_{\text{ref}} \sum_{k=0}^{N-1} g_{k+1/2}^r \eta_{k+1/2} \partial_t \eta_{k+1/2} = \rho_{\text{ref}} \sum_{k=0}^{N-1} g_{k+1/2}^r \eta_{k+1/2} \sum_{j=k+1}^N \partial_t h_j, \quad (2.140)$$

where the second equality made use of equation (1.80) to relate the interface height and the layer thickness. We can collapse the double sum through the following identities

$$\begin{aligned} &\sum_{k=0}^{N-1} g_{k+1/2}^r \eta_{k+1/2} \sum_{j=k+1}^N \partial_t h_j \\ &= \partial_t h_N \sum_{j=0}^{N-1} g_{j+1/2}^r \eta_{j+1/2} + \partial_t h_{N-1} \sum_{j=0}^{N-2} g_{j+1/2}^r \eta_{j+1/2} + \dots + \partial_t h_1 g_{1/2}^r \eta_{1/2}. \end{aligned} \quad (2.141)$$

The product of reduced gravity and interface height can be written in terms of the shallow water dynamic pressure from equation (1.88)

$$p_k^{\text{dyn}} = \rho_{\text{ref}} \sum_{j=0}^{k-1} g_{j+1/2}^r \eta_{j+1/2}, \quad (2.142)$$

in which case we have the potential energy budget

$$\partial_t \mathcal{P}^{\text{sw}} = \sum_{k=1}^N p_k^{\text{dyn}} \partial_t h_k = - \sum_{k=1}^N p_k^{\text{dyn}} \nabla \cdot (\mathbf{u}_k h_k), \quad (2.143)$$

where the second equality made use of the layer thickness equation (1.79a). Rearrangement brings this equation to its flux-form expression

$$\partial_t \mathcal{P}^{\text{sw}} + \nabla \cdot \sum_{k=1}^N p_k^{\text{dyn}} h_k \mathbf{u}_k = \sum_{k=1}^N h_k \mathbf{u}_k \cdot \nabla p_k^{\text{dyn}}. \quad (2.144)$$

## 2.6.2 Potential energy and buoyancy work

### Potential energy for $N$ -layers

Following the method for the single layer in Section 2.5.1, we write the gravitational potential energy for a column of shallow water fluid as

$$\mathcal{P}^{\text{sw}} = g \int_{\eta_b}^{\eta} z \rho dz = g \sum_{k=1}^N \rho_k h_k \bar{\eta}_k = \sum_{k=1}^N \mathcal{P}_k^{\text{sw}}, \quad (2.145)$$

where, again,

$$h_k = \eta_{k-1/2} - \eta_{k+1/2} \quad \text{and} \quad \bar{\eta}_k = (\eta_{k-1/2} + \eta_{k+1/2})/2 \quad (2.146)$$

makes use of the layer thickness,  $h_k$ , and the average interface height,  $\bar{\eta}_k$ . We can think of equation (2.145) as building the potential energy using layer properties, whereas the alternative expression (2.136) is built using interface properties.

### Time evolution of the potential energy

Taking the time derivative of the potential energy (2.145) leads to

$$\partial_t \mathcal{P}_k^{\text{sw}} = g \rho_k \partial_t (h_k \bar{\eta}_k) \quad (2.147a)$$

$$= g \rho_k (\partial_t h_k \bar{\eta}_k + h_k \partial_t \bar{\eta}_k) \quad (2.147b)$$

$$= g \rho_k [-\bar{\eta}_k \nabla \cdot (\mathbf{u}_k h_k) + h_k (\bar{w}_k - \mathbf{u}_k \cdot \nabla \bar{\eta}_k)] \quad (2.147c)$$

$$= g \rho_k h_k \bar{w}_k - g \nabla \cdot (\rho_k \bar{\eta}_k h_k \mathbf{u}_k) \quad (2.147d)$$

$$= g \rho_k h_k \bar{w}_k - \nabla \cdot (\mathbf{u}_k \mathcal{P}_k^{\text{sw}}), \quad (2.147e)$$

so that the flux-form budget is given by

$$\partial_t \mathcal{P}_k^{\text{sw}} + \nabla \cdot (\mathbf{u} \mathcal{P}_k^{\text{sw}}) = g \rho_k h_k \bar{w}_k. \quad (2.148)$$

This budget is identical to equation (2.90) that we discussed for a single layer, with the buoyancy work term exposed on the right hand side.

## 2.6.3 Kinetic energy

### Kinetic energy and work from pressure gradients

Following our treatment in Section 2.5.2, consider the kinetic energy per horizontal area contained in a shallow water layer

$$\mathcal{K}_k^{\text{sw}} = \frac{1}{2} \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} \rho_{\text{ref}} \mathbf{u}_k \cdot \mathbf{u}_k dz = \frac{1}{2} \rho_{\text{ref}} h_k \mathbf{u}_k \cdot \mathbf{u}_k. \quad (2.149)$$

Note that we set the density equal to the reference density as per the Boussinesq ocean studied in VOLUME 2, whereby the inertial mass is determined by the reference density.

To derive an evolution equation for kinetic energy, multiply the thickness equation (1.79a) by  $\mathbf{u}_k \cdot \mathbf{u}_k$ , and take the dot product of the vector-invariant velocity equation (1.113) with  $h_k \mathbf{u}_k$ ,

$$(\mathbf{u}_k \cdot \mathbf{u}_k) \partial_t h_k = -(\mathbf{u}_k \cdot \mathbf{u}_k) \nabla \cdot (h_k \mathbf{u}_k) \quad (2.150a)$$

$$h_k \mathbf{u}_k \cdot \partial_t \mathbf{u}_k = -h_k \mathbf{u}_k \cdot \nabla(p_k/\rho_{\text{ref}} + \mathbf{u}_k \cdot \mathbf{u}_k/2) + h_k \mathbf{u}_k \cdot \mathbf{F}_k, \quad (2.150b)$$

where we included an acceleration,  $\mathbf{F}_k$ , arising from friction or boundary stresses, as discussed in Section 1.6.5. Making use of these expressions leads to the time derivative of the kinetic energy per area,  $\mathcal{K}_k^{\text{sw}} = \rho_{\text{ref}} h_k \mathbf{u}_k \cdot \mathbf{u}_k / 2$ , of a shallow water layer

$$(2/\rho_{\text{ref}}) \partial_t \mathcal{K}_k^{\text{sw}} = \partial_t(h_k \mathbf{u}_k \cdot \mathbf{u}_k) \quad (2.151a)$$

$$= (\mathbf{u}_k \cdot \mathbf{u}_k) \partial_t h_k + 2 h_k \mathbf{u}_k \cdot \partial_t \mathbf{u}_k \quad (2.151b)$$

$$= -(\mathbf{u}_k \cdot \mathbf{u}_k) \nabla \cdot (h_k \mathbf{u}_k) - 2 h_k \mathbf{u}_k \cdot \nabla(p_k/\rho_{\text{ref}} + \mathbf{u}_k \cdot \mathbf{u}_k/2) + 2 h_k \mathbf{u}_k \cdot \mathbf{F}_k \quad (2.151c)$$

$$= -2 \nabla \cdot [\mathbf{u}_k \mathcal{K}_k^{\text{sw}} / \rho_{\text{ref}}] - 2 h_k \mathbf{u}_k \cdot \nabla(p_k/\rho_{\text{ref}}) + 2 h_k \mathbf{u}_k \cdot \mathbf{F}_k, \quad (2.151d)$$

which then leads to the equivalent forms for the layer integrated kinetic energy budget

$$\partial_t \mathcal{K}_k^{\text{sw}} + \nabla \cdot (\mathbf{u}_k \mathcal{K}_k^{\text{sw}}) = -h_k \mathbf{u}_k \cdot [\nabla p_k - \rho_{\text{ref}} \mathbf{F}_k], \quad (2.152a)$$

$$\partial_t \mathcal{K}_k^{\text{sw}} + \nabla \cdot [\mathbf{u}_k (\mathcal{K}_k^{\text{sw}} + h_k p_k)] = p_k \nabla \cdot (h_k \mathbf{u}_k) + \rho_{\text{ref}} h_k \mathbf{u}_k \cdot \mathbf{F}_k. \quad (2.152b)$$

Recall that the pressure gradient is given by equation (1.88)

$$\nabla p_k = \nabla p_a + \nabla p_k^{\text{dyn}} = \nabla p_a + \rho_{\text{ref}} \nabla M_k^{\text{dyn}}, \quad (2.153)$$

which arises from gradients in the applied atmospheric pressure plus the dynamic pressure, and where  $M_k^{\text{dyn}} = \rho_{\text{ref}} p_k^{\text{dyn}}$  is the Montgomery potential from Section 1.4.3. The budget equation (2.152b) corresponds to the single layer equation (2.98), where we identify the dynamic pressure for shallow water model with just a single layer

$$p_{k=1}^{\text{dyn}} = \rho g \eta^{\text{eff}} = p_a + \rho g \eta. \quad (2.154)$$

### Kinetic energy and buoyancy work

The single layer discussion in Section 2.5.3 fully anticipated the  $N$ -layer case, so that we can immediately translate equation (2.106) to an arbitrary layer

$$\partial_t \mathcal{K}_k^{\text{sw}} + \nabla \cdot [\mathbf{u} (\mathcal{K}_k^{\text{sw}} + P_k)] + \delta_k(p_{k-1/2} \partial_t \eta_{k-1/2}) = -g \rho_k h_k \bar{w}_k + \rho_k h_k \mathbf{u}_k \cdot \mathbf{F}_k. \quad (2.155)$$

#### 2.6.4 Comments and further study

The discussion of mechanical energy for the single shallow water layer in Section 2.5.5 directly transfers to the  $N$ -layer case, thus making it unnecessary for us to develop the theory any further.

Elements of the discussion in this section are motivated by the energetic analysis of *Loose et al. (2022)*, who detail the mechanical energy in a stacked shallow water model and decompose the energy budget into mean and transient eddy contributions.

## 2.7 Momentum balance in a zonal channel

In this section we study the vertically integrated steady momentum budget for a single shallow water layer of density  $\rho$  in a zonally re-entrant channel, such as depicted in Figure 2.8. Such shallow water models have been used to garner insights into adiabatic aspects of Southern Ocean circulation, and we keep this application in mind for the following (so that  $f < 0$ ). We are particularly interested in the force balances required to reach a steady flow in the presence of a prescribed constant wind stress acceleration,  $\tau/\rho$ . The channel has arbitrary topography, including northern and southern bounds with sloping shelves and shorelines. Applying a zonal surface stress inserts zonal momentum to the fluid through the ocean surface. For simplicity we set the atmospheric pressure to zero,  $p_a = 0$ , so that the bottom pressure is  $p_b = \rho g h$ , and there is no form stress acting on the layer's upper surface.

A similar analysis was presented in VOLUME 2 for the vertically integrated axial angular momentum budget in a continuously stratified fluid. Following from that analysis, we ignore the role of internal viscous friction. However, we consider bottom frictional stresses written as a quadratic bottom drag

$$\mathbf{F}^{\text{drag}} = -C_d \mathbf{u} |\mathbf{u}|, \quad (2.156)$$

where  $C_d > 0$  is a dimensionless bottom drag coefficient.

The horizontal areal extent of the domain is a function of space and time since the shallow water layer rises up and down the northern and southern shorelines as motion occurs. Even so, since layer thickness vanishes at the shoreline edge, the horizontal boundary conditions for the shallow water layer are easy to apply when working with the thickness weighted equations. That is, all thickness weighted fields vanish at the shoreline edge merely since the thickness vanishes at the edge. Indeed, working with thickness weighted fields allows us to handle any degree of vanishing layer thicknesses, including if the topography in the center of the channel becomes an island rather than a submerged seamount. We also made use of this property of thickness weighted budgets in VOLUME 2, where we also considered sloping sides rather than the commonly considered, yet less realistic, vertical sides.

### 2.7.1 Volume transport for steady flow

Before considering the steady force balance, we establish a constraint based on volume conservation by considering the steady thickness equation

$$\nabla \cdot (h \mathbf{u}) = 0. \quad (2.157)$$

As discussed for non-divergent flow in VOLUME 2, this non-divergence condition means that there is zero net steady transport crossing any simply connected closed contour in the fluid. A particularly interesting closed contour is one that is periodic and extends across the full zonal extent of the channel (see Figure 2.8), in which case

$$\oint h \mathbf{u} \cdot \hat{\mathbf{n}} ds = 0, \quad (2.158)$$

where  $\oint$  denotes a periodic line integral across the zonal extent of the channel,  $\hat{\mathbf{n}}$  is a unit vector normal to the contour, and  $ds$  is the arc-length line element along the contour. The constraint (2.158) reflects the inability of the steady flow to build up or deplete the fluid on one region of the channel at the expense of another. In particular, if the contour follows a constant

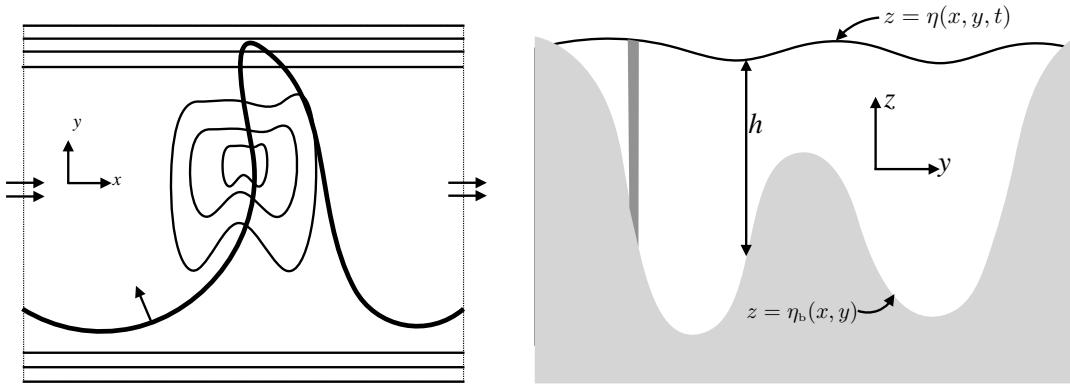


FIGURE 2.8: A zonally periodic/re-entrant southern hemisphere channel with northern and southern shelves and arbitrary seamount topography. Flow that leaves either the east or west boundary is assumed to re-enter the other side, so that the topology is periodic zonally. Left panel: horizontal (plan) view, showing the contours of the shelves and the topography, with flow leaving one of the zonal ends re-entering the other. The arbitrary dark solid contour extends across the zonal extent of the channel and is periodic, with a unit vector,  $\hat{n}$ , depicted normal to a point along the contour. In a steady state, the net fluid transport crossing this contour vanishes:  $\oint h \mathbf{u} \cdot \hat{n} ds = 0$ , meaning that there is no accumulation of fluid within any region of the channel. Right panel: meridional-vertical view through an arbitrary longitude, along with a sample vertical column of water extending from the bottom to the surface. The shoreline edges occur where the layer thickness vanishes on the northern and southern shelves. So although the position of the shoreline edge is a function of space and time (since the fluid moves up and down the shoreline slope), the vanishing layer thickness found at the edge renders a simple treatment of boundary conditions for the thickness weighted equations.

latitude line, then we see that for a steady state there is no net meridional transport across a latitude circle

$$\oint h \mathbf{u} \cdot \hat{\mathbf{y}} dx = \oint h v dx = 0 \quad \text{steady flow.} \quad (2.159)$$

This result means that the thickness weighted Coriolis acceleration appearing in the zonal momentum equation vanishes when integrated zonally

$$\oint f h v dx = f \oint h v dx = 0 \quad \text{steady flow.} \quad (2.160)$$

We make use of this result in Section 2.7.3.

## 2.7.2 Steady meridional balances

Consider the thickness weighted meridional momentum equation (2.35b) in the presence of a wind stress and bottom drag

$$\partial_t(h v) + \partial_x(h u v) + \partial_y(h v^2 + g h^2/2) + u h f = -(p_b/\rho) \partial_y \eta_b + \tau^y/\rho - C_d v |\mathbf{u}|. \quad (2.161)$$

Integrating zonally over the channel removes the zonal transport term,  $\partial_x(h u v)$ , due to periodicity

$$\oint [\partial_t(h v) + \partial_y(h v^2 + g h^2/2) + u h f] dx = \oint [-(p_b/\rho) \partial_y \eta_b + \tau^y/\rho - C_d v |\mathbf{u}|] dx. \quad (2.162)$$

We can pull the time derivative outside of the zonal integral since the domain is zonally periodic, in which case the time changes to the net meridional transport across a latitude circle are given

by

$$\frac{d}{dt} \oint v h dx = \oint [-\partial_y(h v^2 + g h^2/2) - u h f - (p_b/\rho) \partial_y \eta_b + \tau^y/\rho - C_d v |\mathbf{u}|] dx. \quad (2.163)$$

Correspondingly, a steady state along a latitude circle is realized by the balance

$$\oint [\partial_y(h v^2 + g h^2/2) + u h f] dx = \oint [-(p_b/\rho) \partial_y \eta_b + \tau^y/\rho - C_d v |\mathbf{u}|] dx. \quad (2.164)$$

The right hand side represents forcing by the topographic form stress, meridional wind stress, and bottom drag. That forcing, integrated over a latitude circle, balances the left hand side, which is the Coriolis acceleration arising from zonal motion, plus the meridional divergence of meridional momentum advection plus layer integrated pressure.

We can eliminate the nonlinear term on the left hand side of the balance (2.164) by integrating meridionally across the channel. Since  $h = 0$  at the shoreline edges we know that  $h v^2 + g h^2/2$  vanishes at the boundaries, thus leaving

$$\int \left[ \oint u h f dx \right] dy = \int \left[ \oint [-(p_b/\rho) \partial_y \eta_b + \tau^y/\rho - C_d v |\mathbf{u}|] dx \right] dy. \quad (2.165)$$

This is a balance between the integrated meridional Coriolis force on the left hand side with pressure form stress, winds, and bottom drag on the right hand side.

### 2.7.3 Steady zonal balance

Now consider the thickness weighted zonal momentum equation (2.35a), here with bottom drag and wind stress contributions

$$\partial_t(h u) + \partial_x(h u^2 + g h^2/2) + \partial_y(h u v) - v h f = -(p_b/\rho) \partial_x \eta_b + \tau^x/\rho - C_d u |\mathbf{u}|. \quad (2.166)$$

Assuming a steady state and integrating along a latitude circle leads to

$$\rho \oint \partial_y(h u v) dx = \oint [-(p_b \partial_x \eta_b + \tau^x - C_d \rho u |\mathbf{u}|)] dx, \quad (2.167)$$

where we dropped the Coriolis acceleration as per volume conservation in equation (2.160). For flows that are quasi-linear, the nonlinear term  $\partial_y(h u v)$  will be subdominant to the wind stress and topographic form stress, thus leading to the approximate balance along each latitude circle

$$\oint p_b \partial_x \eta_b dx \approx \oint [\tau^x - C_d \rho u |\mathbf{u}|] dx \quad \text{nonlinear term small.} \quad (2.168)$$

We realize an exact balance over the full channel domain by meridionally integrating the latitude balance (2.167), in which the nonlinear term  $\partial_y(h u v)$  drops out since  $h = 0$  at the northern and southern shoreline edges

$$\int \left[ \oint \tau^x dx \right] dy = \int \left[ \oint [p_b \partial_x \eta_b + C_d \rho u |\mathbf{u}|] dx \right] dy. \quad (2.169)$$

Again, this is an exact steady state balance realized by integrating the zonal thickness weighted momentum equation over the full domain channel. It is a balance between the zonal momentum input from the winds (left hand side) to the full domain, and the integrated bottom form stress

plus bottom frictional drag (right hand side).

#### 2.7.4 The role of frictional bottom drag

Consider a flat bottom channel, in which case the area integrated balance (2.169) is between winds and bottom drag

$$\int \left[ \oint \tau^x dx \right] dy = \int \left[ \oint C_d \rho u |\mathbf{u}| dx \right] dy \implies \rho^{-1} \langle \tau^x \rangle = C_d \langle u |\mathbf{u}| \rangle, \quad (2.170)$$

where the angle brackets denote a channel area mean. Typical empirical values for the dimensionless bottom drag coefficient are

$$C_d \approx 2 \times 10^{-3}. \quad (2.171)$$

An area mean eastward zonal wind stress of  $\langle \tau^x \rangle = 0.1 \text{ N m}^{-2}$  leads to a root-mean-square zonal velocity of

$$\sqrt{\langle u |\mathbf{u}| \rangle} \approx \sqrt{\langle u^2 \rangle} \approx 0.2 \text{ m s}^{-1}, \quad (2.172)$$

where the first approximation follows from assuming the zonal velocity dominates over the meridional velocity. How realistic is this number for the Southern Ocean? Field measurements from the Southern Ocean suggest that vertically and area averaged velocities are far smaller than this value. Furthermore, if this value occurred in a channel 4000 m deep and 2000 km wide (a rough idealization of the Antarctic Circumpolar Current), then this vertically averaged velocity would yield a zonal volume transport of  $\approx 1500 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ , which is about ten times larger than the measured transport through the Drake Passage.

*Munk and Palmén (1951)* identified the problematic aspect of assuming a bottom frictional stress balance for the Southern Ocean. In brief, the field measurements do not support a frictional balance, either from bottom drag or from internal turbulent viscous friction. By inference, they proposed that topographic form stress is the dominant term that balances wind stress in the Southern Ocean. They supported that inference through estimates based on topographic features encountered by the Antarctic Circumpolar Current in its transit of the Southern Ocean. Subsequent studies using theory, field measurements, and numerical models support their conclusion. Indeed, in numerical models one finds that so long as there is only a modest degree of bottom slope, the bottom topographic form stress dominates over bottom drag. Given these considerations, we dispense with bottom drag for the remainder of this section.

#### 2.7.5 Correlation between surface height and topographic slope

Given the minor role for bottom drag in establishing a steady channel flow in the Southern Ocean, the balance (2.169) says that an eastward area integrated wind stress must be balanced by a westward topographic form stress

$$\int \left[ \oint \tau^x dx \right] dy = \int \left[ \oint p_b \partial_x \eta_b dx \right] dy. \quad (2.173)$$

What is required to establish a westward topographic form stress? Quite simply, in the area mean, there must be an anomalously large bottom pressure in regions where  $\partial_x \eta_b > 0$  and an anomalously small bottom pressure in regions where  $\partial_x \eta_b < 0$ . Bottom pressure in a shallow

water layer is determined by the column thickness. Hence, to establish the anomalous bottom pressures requires an anomalously thick fluid column upstream of topographic bumps and thin fluid column downstream. This situation is illustrated in Figure 2.9 described in Section 2.7.7.

To further reveal the correlation between surface height and bottom topography, write  $p_b = \rho g h$  and use  $\eta = h + \eta_b$  so that

$$p_{\text{bot}} \partial_x \eta_b = \rho g (\eta - \eta_b) \partial_x \eta_b = \rho g \eta \partial_x \eta_b - (g \rho / 2) \partial_x \eta_b^2. \quad (2.174)$$

The balance (2.173) thus becomes

$$\langle \tau^x \rangle = \rho g \langle \eta \partial_x \eta_b \rangle. \quad (2.175)$$

Furthermore, due to zonal periodicity, it is only the zonal anomalies in  $\eta$  and  $\eta_b$  that contribute so that

$$\langle \tau^x \rangle = \rho g \langle \eta' \partial_x \eta'_b \rangle, \quad (2.176)$$

where

$$\eta = \eta' + L^{-1} \oint \eta \, dx \quad \text{and} \quad \eta_b = \eta'_b + L^{-1} \oint \eta_b \, dx, \quad (2.177)$$

with  $L = \oint dx$  the zonal length of the channel. With  $\langle \tau^x \rangle > 0$ , we see that surface height anomalies must be positively correlated with the bottom slope,  $\int \eta' \partial_x \eta'_b \, dx \, dy > 0$ . That is, the surface height is high where topography slopes are positive and low where topography slopes are negative.

As noted above, we must have a positive correlation between surface height anomalies and topographic slope, as in equation (2.176). It follows that a nonzero zonal integrated topographic form stress requires a nonzero phase shift between surface height anomalies and the bottom topography anomalies. That is, if the surface height and bottom topography were perfectly aligned along a latitude circle, then  $\oint \eta' \partial_x \eta'_b \, dx = 0$ , in which case there is a zero zonal integrated topographic form stress along that latitude. The required phase shift between the free surface and bottom topography has the free surface shifted ahead (i.e., to the west) of the topography. We consider an explicit example in Section 2.7.7 to help in our understanding.

## 2.7.6 Connection to meridional geostrophic transport

Zonal periodicity allows us to swap the zonal derivative in the balance (2.176) so that

$$\langle \tau^x \rangle = -\rho g \langle \partial_x \eta' \eta'_b \rangle \quad (2.178)$$

For the large scale flows under consideration here, we can assume that  $g \partial_x \eta'$  is associated with an anomalous meridional geostrophic velocity

$$g \partial_x \eta' = f v'_g \quad (2.179)$$

so that the balance (2.178) is

$$\langle \tau^x \rangle = -\rho \langle f v'_g \eta'_b \rangle. \quad (2.180)$$

Hence, the steady balance is realized with anomalous meridional geostrophic transport correlated with topographic anomalies. Note that periodicity means that the steady meridional geostrophic

transport has a zero zonal integral

$$\oint v'_g dx = (g/f) \oint \partial_x \eta' dx = 0, \quad (2.181)$$

which follows from the steady volume balance discussed in Section 2.7.1.

### 2.7.7 Sinusoidal example

To help further our understanding of the balance (2.176), and the phase shift required to develop nonzero zonal integrated topographic form stress, consider a sinusoidal topography that is a function only of the zonal direction. Also assume that the free surface has a sinusoidal shape (though we do not specify the dynamical mechanism for it to reach this shape). With these assumptions the anomalous surface height and bottom topography can be written

$$\eta' = \eta_o \sin(\kappa x + \varphi) \quad \text{and} \quad \eta'_b = D \sin(\kappa x) \quad (2.182)$$

where  $\eta_o > 0$  is a constant amplitude for the free surface undulations,  $D > 0$  is the constant amplitude for the bottom undulations,  $\kappa = 2\pi n/L$  is the wavenumber for the undulations,  $n > 0$  is an integer,  $L$  is the size of the zonal channel, and  $\varphi$  is a phase shift between the topography and the free surface. The corresponding meridional geostrophic flow is given by

$$f v'_g = g \partial_x \eta' = g \kappa \eta_o \cos(\kappa x + \varphi). \quad (2.183)$$

Plugging into the balance (2.176) leads to

$$\sin \varphi = \left[ \frac{L \langle \tau^x \rangle}{D \rho g \eta_o \pi n} \right]. \quad (2.184)$$

For  $\langle \tau^x \rangle > 0$  we see that the free surface undulations are, as expected, shifted to the west of the bottom topography undulations. As an explicit example from an idealized channel configuration, set

$$L = 10^7 \text{ m} \quad D = 10^3 \text{ m} \quad \eta_o = 1 \text{ m} \quad \tau^x = 1 \text{ N m}^{-2}, \quad (2.185)$$

in which case

$$\sin \varphi \approx (\pi n)^{-1}. \quad (2.186)$$

For  $n = 1$ , which corresponds to just one topographic bump, then we have a phase shift of  $\varphi \approx 18^\circ$ , and this case is depicted in Figure 2.9 for the southern hemisphere. If there are two bumps, then the phase shift is reduced to  $\varphi \approx 9^\circ$  since each bump shares half the burden of balancing the wind stress.

#### Summarizing the dependencies

We here highlight the various dependencies in the phase equation (2.184).

- The phase shift increases with both larger wind stress and larger zonal extent to the domain. This dependency arises since with enhanced wind stress and an enhanced zonal **fetch** (distance over which the winds blow), there is an increase in zonal momentum inserted to the ocean that must be absorbed by the bottom. The larger phase shift increases this topographic form stress, thus enabling the balance.

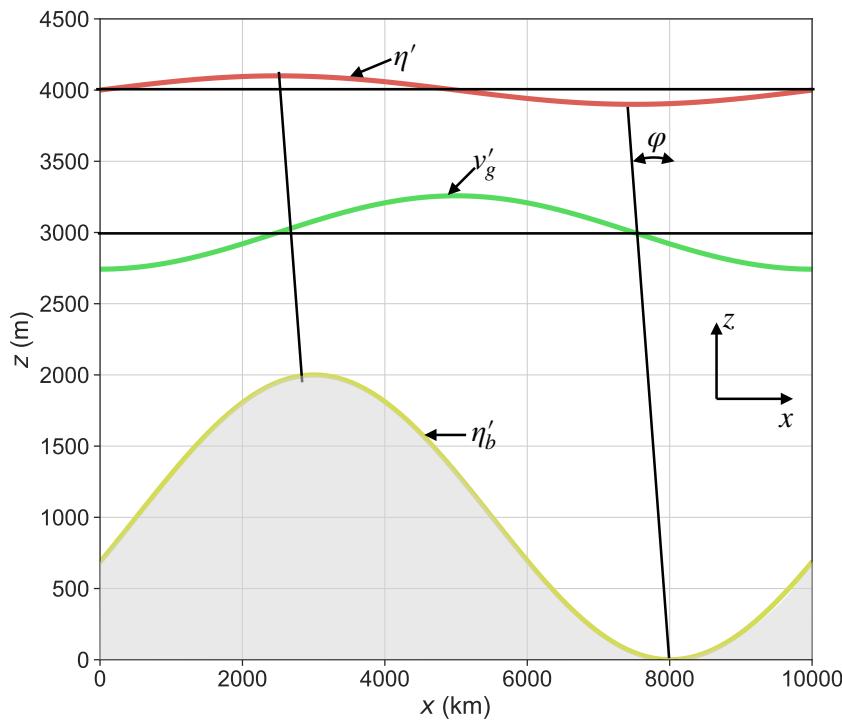


FIGURE 2.9: Zonal-vertical view of the wind stress and bottom pressure acting on a steady state layer of shallow water fluid flowing over a sinusoidal bump (north is directed into the page and east is to the right). The eastward surface stress acts in the  $+\hat{x}$  direction and it is balanced by a westward topographic form stress. To establish this form stress when integrated over the zonal extent of the channel requires the sea surface undulations to be phase shifted to the west of the bottom topography undulations, with the phase shift,  $\varphi$ , given by equation (2.184). To better visualize the sea surface, the surface undulations are in units of cm whereas the bottom topography undulations are in units of m. In general, for the topographic form stress to balance the wind stress, the bottom pressure must be anomalously large where  $\partial\eta_b/\partial x > 0$  and small where  $\partial\eta_b/\partial x < 0$ , thus leading to the anomalously thick fluid column upstream of the bump and thin column downstream. This correlation also leads to a corresponding meridional geostrophic flow as shown here by the green curve for the southern hemisphere where  $f < 0$ , with  $v_g > 0$  (northward) when shown above the horizontal line and  $v_g < 0$  (southward) when below the line.

- Conversely, the phase shift decreases for larger topography  $D$ , and larger undulations in the free surface,  $\eta_o$ , in which case the free surface becomes more aligned with the topographic ridges. This result follows since the topographic form stress is larger for larger topography, thus requiring less phase shift in the surface wave patterns to affect a bottom pressure anomaly.
- Phase shifts decrease when there are more topographic bumps in the channel, with  $n$  the parameter setting the number of bumps. For the Southern Ocean, [Munk and Palmén \(1951\)](#) identified around four or five large-scale topographic features that provide the dominant balance for the zonal wind stress.
- What if the parameters are such that the right hand side of equation (2.184) has a magnitude larger than unity (e.g., huge winds, very long fetch, small topography, small surface height amplitude)? This situation signals that topographic form stress is insufficient to balance the winds. In a numerical model, one can merely increase the topographic wavenumber,  $n$ , to increase the topographic form stress. Yet where topography is fixed, such as in Nature, then bottom frictional stresses come into play to help reach a force

balance (see Section 2.7.4).

### Distinguishing steady motion from zero motion

Why is the phase shift (2.184) independent of the fluid depth? One might suspect that to reach a force balance would require more form stress if there is more fluid. Instead, the force balance, as reflected in the phase shift, depends on the zonal anomalies of the surface height and bottom topography. The depth of the fluid is absent.

The answer to this question is that we are seeking a force balance. When forces are balanced there is no acceleration and thus, as per Newton's second law, the fluid maintains a constant velocity relative to the laboratory reference frame. If we instead wished to stop the fluid, then we would need to decelerate all fluid elements to zero velocity. Determining the forces needed to stop the fluid requires the total fluid mass and thus its depth (as well as the time over which the fluid is to stop). If the fluid is in motion, then halting the motion requires a net force, and that is a very different consideration than the case of zero net force. So in brief, a steady state refers to the absence of time dependence in the fluid from the perspective of an Eulerian (laboratory) observer, with a steady state not necessarily a static state.

## 2.7.8 Comments and further study

### Gyres and channels

The dominance of topographic form stress for the steady Southern Ocean balance contrasts to many of the frictional theories for gyre circulations, such as the solutions discussed in Chapter 4. In those theories, based on flat bottom and vertical side configurations, the curl of the wind stress is balanced by torques created by friction.<sup>14</sup> One is thus led to conclude that zonally re-entrant channels exhibit a fundamentally distinct steady force balance from gyres. However, as discussed in VOLUME 2, *Hughes and de Cueves* (2001) showed how topographic form stresses associated with sloping sides can lead to an inviscid balance for gyres. That is, friction is far less important so long as the bottom can support topographic torques. In this manner, gyres and channels share much in common so long as they both contain topography and sloping sides.

### The case with vertical stratification

The analysis of a single shallow water layer has direct relevance to flow in a stratified fluid. The reason is that when integrating over the full depth of the fluid, internal interfacial form stresses cancel pairwise.<sup>15</sup> The resulting net balance for contact forces is concerned with just those acting on the boundaries at the surface and the bottom. This property of contact forces was also implicit in our study in VOLUME 2 of angular momentum in a continuously stratified fluid. In that discussion we encountered the correlation between bottom pressure and bottom topography slope for steady flow in a channel as realized by a balance between wind stress and bottom form stress. This correlation also holds for the single shallow water layer.

Although the single layer provides a direct connection to the vertically integrated momentum in a continuously stratified fluid, the direct connection between undulations in the sea surface height and bottom pressure is more nuanced when allowing for stratification. We here outline some of the considerations that arise with flow in a stratified channel with a topographic bump.

---

<sup>14</sup>We encounter such balances for the shallow water vorticity in Chapter 5.

<sup>15</sup>This property of interfacial form stresses follows from Newton's third law.

Our presentation is rather incomplete, with a more thorough analysis supported by numerical simulations.

Following the analysis in VOLUME 2 of hydrostatic pressure scaling, we decompose the horizontal gradient of bottom pressure according to

$$\nabla_h p_b = \underbrace{\nabla_h p_a + g \rho(\eta) \nabla_h \eta}_{\text{external contribution}} - \underbrace{\rho_o \int_{\eta_b}^{\eta} \nabla_h b dz'}_{\text{internal contribution}} \equiv \nabla_h p_{\text{bext}} + \nabla_h p_{\text{bint}}, \quad (2.187)$$

where  $b = -g(\rho - \rho_o)/\rho_o$  is the [Archimedean buoyancy](#) from VOLUME 2. The external pressure gradient is all that is available for a single shallow water layer, so that there is a direct correlation between bottom topography and surface pressure in the steady channel flow where bottom form stress balances wind stress. In contrast, for a continuously stratified fluid, or for a stacked shallow water fluid, the internal contribution to the pressure gradient is nonzero since buoyancy generally has a horizontal gradient.

As discussed in Section 16.4 of [Olbers et al. \(2012\)](#), watermasses in the Southern Ocean generally align themselves with lighter water on the upstream side of topographic features (to the west), and heavy water on the downstream side. This configuration means that the internal pressure is lower upstream of a topographic bump and higher downstream, thus leading to a westward internal pressure gradient force. To realize a steady flow with a balance between bottom pressure form stress and wind stress requires the external contribution to the bottom pressure gradient to counteract the internal contribution. Hence, the free surface height must have larger undulations in the presence of vertical stratification than without, with high values upstream of the bump and low values downstream. We depict this configuration in Figure 2.10, with the caption offering further details. [Zhang et al. \(2024\)](#) further detail the facets of the dynamical balances, including transient adjustments towards the steady balances discussed here. They offer a particularly clear distinction between the barotropic dynamics (as realized by a single shallow water layer) and baroclinic dynamics. Furthermore, they emphasize the central role of the barotropic dynamics for maintaining the momentum balance even for stratified fluid.

### Meridional overturning circulation

The balances in this section are modified when allowing for the vertical transfer of volume between the layers as required to admit a meridional-vertical overturning circulation. In this case, there can be net meridional motion along a latitude circle to add the Coriolis force to the steady force balance, including transport within an [Ekman boundary layer](#) (balance between Coriolis and surface stress as studied in VOLUME 2) for the layer feeling the zonal surface stress. Section 21.7 of [Vallis \(2017\)](#) as well as Chapter 16 of [Olbers et al. \(2012\)](#) provide pedagogical discussions of flow in the Antarctic Circumpolar Current, in which interfacial pressure form stress developed from baroclinic eddies provides a mechanism for vertically redistributing horizontal momentum. The studies from [Webb and de Cueves \(2007\)](#) and [Zhang et al. \(2024\)](#) detail how the barotropic dynamics, which act over days and weeks, continue to play the dominant role in balancing the surface wind stress and topographic form stress, even in the case of a stratified ocean with eddy form stresses, which have a time scale of years.

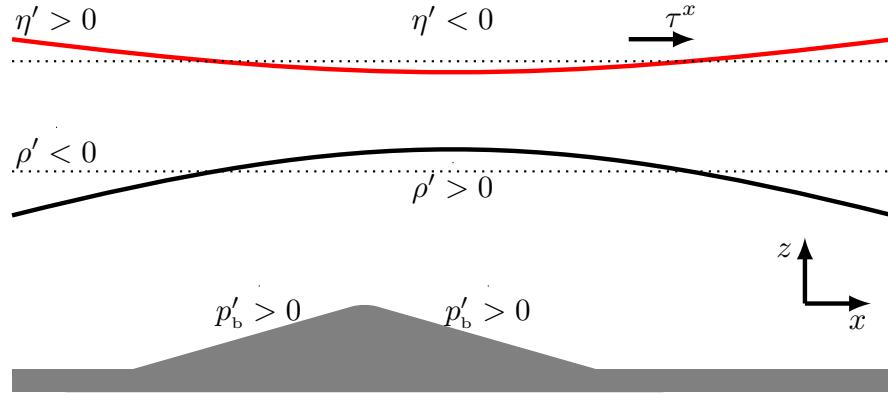


FIGURE 2.10: A schematic of steady flow in a two-layer zonally periodic channel with a topographic bump and with bottom pressure form stress balancing zonal wind stress. To realize this balance requires anomalously high bottom pressure on the upstream side of the bump and anomalously low bottom pressure on the downstream side, where anomalies are relative to the zonal mean. Such anomalous bottom pressure is just as for a single shallow water layer. However, for a stratified fluid the bottom pressure is established by the sum of effects from the external and internal pressure fields. The internal pressure field arises from density, here shown with anomalously light water on the upstream side of the bump and heavy water downstream, such as occurs in the Southern Ocean. This density field leads to a westward contribution to the bottom pressure gradient force; i.e., anomalously low bottom pressure on the upstream side of the bump and high bottom pressure on the downstream. The external pressure field arises from the free surface undulations (red line), with a high upstream of the bump and low downstream. This free surface field leads to an eastward contribution to the bottom pressure gradient force; i.e., anomalously high bottom pressure on the upstream side of the bump and low bottom pressure on the downstream. For the bottom pressure form stress to balance the wind stress, we must have the external pressure gradient dominate the internal pressure gradient. Note that undulations of the free surface height are roughly 100-300 times smaller than those of the density field, with the relative undulations set according to the reduced gravity as described in Section 1.3.

## 2.8 Angular momentum in a rotating tank

As our second case study for this chapter, we study **angular momentum** for a layer of shallow water fluid in a rotating cylindrical tank. This system was discussed in VOLUME 2, where we developed the horizontal equation of motion

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla (p/\rho + g_e z - \Omega^2 r^2/2), \quad (2.188)$$

where  $r^2 = x^2 + y^2$  is the radial distance from the rotational axis,

$$\Omega = \frac{1}{2}f \quad (2.189)$$

is the constant angular rotation rate, and the vertical component to the right hand side is the hydrostatic balance,  $\partial p / \partial z = -\rho g_e$ . Where convenient, we make use of the polar coordinates (see VOLUME 1) in the following, whereby

$$x = r \cos \vartheta \quad (2.190a)$$

$$y = r \sin \vartheta, \quad (2.190b)$$

with the polar angle  $\vartheta$  measured counter-clockwise from the positive  $x$ -axis.

### 2.8.1 Angular momentum for a column of shallow water fluid

The angular momentum for a column of shallow water fluid, computed with respect to the vertical rotational axis, is given by

$$L^z = \delta M [\mathbf{x} \times (\mathbf{u} + \mathbf{U}_{\text{rigid}})] \cdot \hat{\mathbf{z}}, \quad (2.191)$$

where  $\mathbf{x} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}} = r \hat{\mathbf{r}}$  is the position vector relative to the rotational axis,  $\delta M = \rho h \delta A$  is the constant mass for the fluid column, and the rigid-body rotation velocity is

$$\mathbf{U}_{\text{rigid}} = (f/2) \hat{\mathbf{z}} \times \mathbf{x} = r \Omega \hat{\boldsymbol{\vartheta}}, \quad (2.192)$$

where  $\hat{\mathbf{z}} \times \hat{\mathbf{r}} = \hat{\boldsymbol{\vartheta}}$  is the azimuthal unit vector pointing counter-clockwise around the origin.

We can further massage the expression for the angular momentum by writing

$$\mathbf{x} \times \mathbf{u} = (x v - y u) \hat{\mathbf{z}} = r^2 \dot{\vartheta} \hat{\mathbf{z}}, \quad (2.193)$$

where  $\dot{\vartheta} = D\vartheta/Dt$  is the angular velocity. Likewise, we have

$$\mathbf{x} \times \mathbf{U}_{\text{rigid}} = r^2 \Omega \hat{\mathbf{z}}, \quad (2.194)$$

so that the angular momentum can be written

$$L^z = \delta M [\mathbf{x} \times (\mathbf{u} + \mathbf{U}_{\text{rigid}})] \cdot \hat{\mathbf{z}} = \delta M r^2 (\dot{\vartheta} + \Omega). \quad (2.195)$$

### 2.8.2 Material time evolution of the angular momentum

The material time evolution for the angular momentum is given by

$$\frac{DL^z}{Dt} = \delta M [\mathbf{u} \times (\mathbf{u} + \mathbf{U}_{\text{rigid}})] \cdot \hat{\mathbf{z}} + \delta M \left[ \mathbf{x} \times \left( \frac{D\mathbf{u}}{Dt} + \frac{D\mathbf{U}_{\text{rigid}}}{Dt} \right) \right] \cdot \hat{\mathbf{z}}. \quad (2.196)$$

Note that we set

$$\frac{D(\delta M)}{Dt} = \rho D(h \delta A) Dt = 0 \quad (2.197)$$

since the shallow water fluid columns each have constant volume as they move with the horizontal flow (Section 1.2.5). Using the rigid-body rotation velocity given by equation (2.192), and with a constant rotation rate, yields

$$\mathbf{u} \times \mathbf{U}_{\text{rigid}} + \mathbf{x} \times \frac{D\mathbf{U}_{\text{rigid}}}{Dt} = \mathbf{u} \times (\boldsymbol{\Omega} \times \mathbf{x}) + \mathbf{x} \times (\boldsymbol{\Omega} \times \mathbf{u}) \quad (2.198a)$$

$$= (\mathbf{x} \cdot \mathbf{u}) f \hat{\mathbf{z}}. \quad (2.198b)$$

Making use of the material evolution of the horizontal velocity given by equation (2.188) renders

$$\left[ \mathbf{x} \times \frac{D\mathbf{u}}{Dt} \right] \cdot \hat{\mathbf{z}} = (\mathbf{x} \times [-f \hat{\mathbf{z}} \times \mathbf{u} - \nabla(p/\rho + g_e z - \Omega^2 r^2/2)]) \cdot \hat{\mathbf{z}} \quad (2.199a)$$

$$= -f (\mathbf{x} \cdot \mathbf{u}) - (\mathbf{x} \times g \nabla \eta) \cdot \hat{\mathbf{z}}. \quad (2.199b)$$

The centrifugal term dropped out since

$$\mathbf{x} \times \nabla r^2 = 2\mathbf{x} \times r\hat{\mathbf{r}} = 2\mathbf{x} \times \mathbf{x} = 0. \quad (2.200)$$

The gravitational term dropped out since

$$(\mathbf{x} \times \nabla z) \cdot \hat{\mathbf{z}} = (\mathbf{x} \times \hat{\mathbf{z}}) \cdot \hat{\mathbf{z}} = 0, \quad (2.201)$$

as does the vertical component to the pressure gradient. We are thus left with

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g(\mathbf{x} \times \nabla \eta) \cdot \hat{\mathbf{z}}. \quad (2.202)$$

Consequently, the axial angular momentum for a fluid column is modified by the torque from the horizontal pressure gradient caused by undulations in the free surface height. Note how there is no contribution from the Coriolis acceleration, so that the evolution of angular momentum is the same whether viewed in the laboratory frame or rotating frame.

We can bring the expression (2.202) into a more transparent form by switching to polar coordinates

$$\mathbf{x} \times \nabla \eta = r\hat{\mathbf{r}} \times \left[ \hat{\mathbf{r}} \frac{\partial \eta}{\partial r} + \hat{\vartheta} \frac{1}{r} \frac{\partial \eta}{\partial \vartheta} \right] = \frac{\partial \eta}{\partial \vartheta} \hat{\mathbf{z}}, \quad (2.203)$$

so that

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g \frac{\partial \eta}{\partial \vartheta}. \quad (2.204)$$

This result is directly analogous to the angular momentum evolution for a fluid moving around a sphere as derived in VOLUME 2. Namely, in the presence of angular pressure gradients, the fluid experiences a torque that in turn leads to a change in the angular momentum relative to the vertical rotation axis.

### 2.8.3 Materially invariant angular momentum

The angular momentum for a fluid column is materially invariant (i.e., a constant on a material fluid parcel) if

$$\frac{DL^z}{Dt} = 0 \iff \frac{\partial \eta}{\partial \vartheta} = 0. \quad (2.205)$$

For the flat bottom homogeneous fluid studied in VOLUME 2, we find that the free surface takes on a radial parabolic shape when the fluid is in rigid-body rotation. In this case,  $\nabla \eta$  is in the radial direction, in which case  $\mathbf{x} \times \nabla \eta = 0$ . Consequently, when the fluid is in rigid-body rotation, the angular momentum for each fluid column remains materially constant.

### 2.8.4 Comments

The material evolution equation (2.202) also holds for a fluid on the  $f$ -plane tangent to a sphere. The  $f$ -plane formulation is slightly simpler than the tank since the centrifugal term is absorbed into the geopotential<sup>16</sup>. However, the tank is arguably more pedagogical as it is simpler to visualize and to conduct laboratory experiments. See Section 6.6.4 of [Marshall and Plumb \(2008\)](#) for more discussion of rotating tank experiments. Also, we return to this physical system in VOLUME 4 when studying centrifugal instability of cyclostrophically balanced flow.

---

<sup>16</sup>For details of the effective gravitational potential, see the discussion of a particle moving around a rotating planet in VOLUME 1.



## 2.9 Exercises

### EXERCISE 2.1: POTENTIAL TEMPERATURE SLOPES IN ATMOSPHERE AND OCEAN

Use the two-layer thermal wind relations from Section 2.2.2, also known as Margules' relation, to estimate the slope of the potential temperature surfaces in the atmosphere and ocean. This exercise is based on exercise 3.2 of [Vallis \(2006\)](#).

- Model the atmosphere as two immiscible shallow water layers of different density stacked one above the other. Using reasonable values for any required physical parameters, estimate the vertical displacement of the interfacial surface associated with a pole-to-equator temperature difference of 40K. You may wish to consult [Wallace and Hobbs \(2006\)](#) or [Marshall and Plumb \(2008\)](#) for physical scales.
- Estimate a vertical interfacial displacement in the ocean thermocline associated with a temperature difference of 20K over a horizontal distance of 4000 km. The interface between the two shallow water layers offers a crude representation of the main oceanic thermocline. Ignore salinity effects so that temperature and density are directly proportional.

Double-check your results by examining some atmosphere and ocean latitude-height profiles for potential temperature (e.g., Figure 5.8 of [Marshall and Plumb \(2008\)](#)).

### EXERCISE 2.2: CIRCULAR STEADY GEOSTROPHIC FLOW

Consider a single layer of shallow water fluid in steady geostrophic balance on a *f*-plane so that

$$f \hat{z} \times \mathbf{u}_g = -g \nabla \eta. \quad (2.206)$$

Assume  $f > 0$  and that the free surface has a circular Gaussian shape

$$\eta = \eta_0 e^{-r^2/(2\sigma^2)} \quad (2.207)$$

where  $r^2 = x^2 + y^2$  is the squared radial position and  $\sigma$  is the standard deviation of the Gaussian.

- Determine the horizontal geostrophic velocity components corresponding to this free surface undulation. Write the solution in both Cartesian coordinates and polar coordinates. Is the flow oriented cyclonic or anti-cyclonic?
- Determine the equation for flow streamlines. What is the geometric shape of a streamline?  
Hint: recall the discussion of streamlines in VOLUME 2.

### EXERCISE 2.3: STEADY STATE MOMENTUM AND GEOSTROPHY

Consider a single layer of shallow water fluid with zero boundary mass fluxes through the surface. Assume the lateral boundaries are solid. All boundaries are thus material. The domain integrated horizontal momentum (within the rotating reference frame) is defined by

$$\mathbf{P} = \int \rho \mathbf{u} dV = \int \rho h \mathbf{u} d\mathcal{S}. \quad (2.208)$$

Show that for tangent plane motion

$$\frac{d\mathbf{P}}{dt} = 0 \quad (2.209)$$

can be realized either by (A) zero flow everywhere, (B) flow that is in geostrophic balance at

each point, or (C) flow that is in geostrophic balance as a global integral.

**EXERCISE 2.4: THICKNESS WEIGHTED MOMENTUM FOR TWO LAYERS**

Following the methods from Section 2.3, derive the thickness weighted momentum equation for an inviscid two-layer stacked shallow water fluid. That is, derive the evolution equation for  $h_1 \mathbf{u}_1 + h_2 \mathbf{u}_2$ , thus providing the two-layer version of equation (2.21).

**EXERCISE 2.5: CONTACT PRESSURE FORCE ON A SINGLE LAYER**

As a check on our calculation of the contact pressure force (2.60b), consider a single shallow water layer under a massless atmosphere. Show that the contact pressure force per mass is given by

$$\frac{\mathbf{F}_{\text{net}}^{\text{press}}}{M} = g \hat{\mathbf{z}} - g \nabla \eta. \quad (2.210)$$

As expected, the horizontal component of this force equals to the pressure gradient body force per mass detailed in Section 1.2.1. The vertical pressure force balances the weight of the fluid as per the hydrostatic balance.

**EXERCISE 2.6: TOPOGRAPHIC FORM STRESS FOR A RIDGE**

As in Section 2.7, apply a constant eastward zonal wind to a zonally reentrant channel with a single shallow water layer. Let the layer flow over a topographic ridge of height  $H$  above the surrounding flat bottom, and let the ridge be a function just of zonal position,  $\eta'(x)$ . Furthermore, assume the ridge has a constant slope on both the upstream (west) side,  $S_{\text{up}}$ , and downstream (east) side,  $S_{\text{dn}}$ . An example is depicted in Figure 2.11. Following the force balance (2.176), derive an expression for the free surface height zonally averaged over the upstream side of the ridge, minus the free surface height zonally averaged over the downstream sides of the ridge,

$$\Delta\eta' = (\bar{\eta}')_{\text{up}} - (\bar{\eta}')_{\text{dn}}, \quad (2.211)$$

where

$$(\bar{\eta}')_{\text{up}} = \frac{\int_{x_{\text{up}}}^{x_0} \eta' dx}{L_{\text{up}}} \quad \text{and} \quad (\bar{\eta}')_{\text{dn}} = \frac{\int_{x_0}^{x_{\text{dn}}} \eta' dx}{L_{\text{dn}}}. \quad (2.212)$$

Show that the expression for  $\Delta\eta'$  is independent of the two slopes. Instead, the only geometric property that determines  $\Delta\eta'$  is the ridge height,  $H$ . Discuss this result.

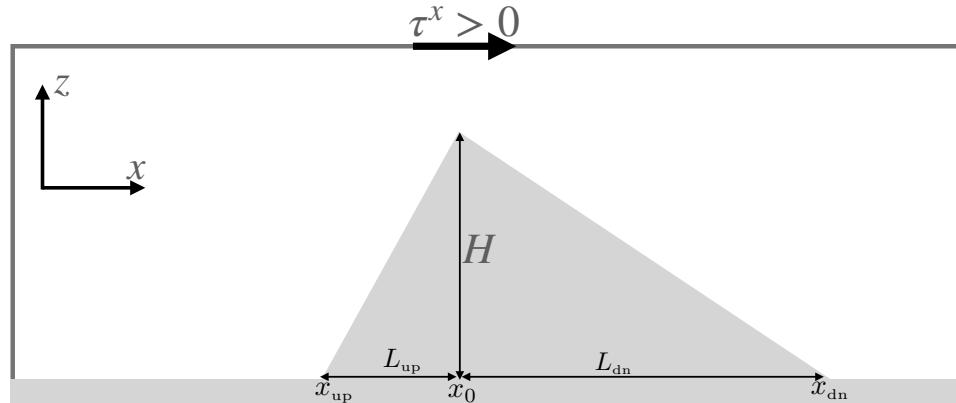


FIGURE 2.11: Zonal-vertical view of a single layer of shallow water fluid moving in a zonally re-entrant channel for use in Exercise 2.6. The domain has a topographic ridge that is a function just of the zonal direction and has constant slopes on its western and eastern sides. There is a constant eastward zonal wind stress.

## EXERCISE 2.7: GEOSTROPHIC TRANSPORT

Consider a zonal-vertical section of shallow water flow in the middle latitude northern hemisphere. Let the section be 1000 m deep and away from side and bottom boundaries. Assume the sea level is 1 cm higher at the eastern end of the section than the western end. Estimate the mass transport (kg/sec) of constant density seawater going through the section. What direction is the transport? Hint: Assume geostrophic balance; choose a representative constant seawater density; and note that the zonal width of the section cancels out so it is not needed.

## EXERCISE 2.8: AVAILABLE POTENTIAL ENERGY FOR TWO SHALLOW WATER LAYERS

Compute the available potential energy for two shallow water layers using the notation from Figure 1.6 with nontrivial bottom topography,  $z = \eta_b(x, y)$ . Show that the APE is non-negative. Assume the domain is simply connected. Hint: The answer is given by specializing the arbitrary  $N$  results in equation (2.139) to the special case of the  $N = 2$ . However, for this exercise you should not merely quote that result. Instead, show all steps starting from the potential energy for an  $N = 2$  layer model.

EXERCISE 2.9: RATIO OF KE TO APE FOR SINGLE LAYER  $f$ -PLANE GEOSTROPHY

Consider a single layer of shallow water fluid in geostrophic balance on an  $f$ -plane with a flat bottom. Show that the ratio of kinetic energy to available potential energy scales like

$$\frac{E_{\text{KE}}}{E_{\text{APE}}} \sim \left[ \frac{L_d^{\text{ext}}}{L} \right]^2. \quad (2.213)$$

In this equation,  $L$  is the horizontal length scale for the fluctuation of the free surface  $\eta$  (i.e.,  $\nabla \eta \sim \eta'/L$ ), and

$$L_d^{\text{ext}} = \frac{\sqrt{g H}}{f} \quad (2.214)$$

is the external deformation radius. The scaling (2.213) means that for scales larger than the external deformation radius,  $L_d$ , the available potential energy is larger than the kinetic energy. The converse holds for scales smaller than  $L_d$ .

EXERCISE 2.10: RATIO OF KE TO APE FOR 1.5 LAYER  $f$ -PLANE GEOSTROPHY

Consider a reduced gravity system (Section 1.3) in geostrophic balance on an  $f$ -plane. Show that the ratio of kinetic energy to available potential energy scales like

$$\frac{E_{\text{KE}}}{E_{\text{APE}}} \sim \left[ \frac{L_d^{\text{int}}}{L} \right]^2. \quad (2.215)$$

In this equation,  $L$  is the horizontal length scale for the fluctuation of the internal interface  $\eta_{3/2}$  (i.e.,  $\nabla \eta_{3/2} \sim \eta'_{3/2}/L$ ), and

$$L_d^{\text{int}} = \frac{\sqrt{g_{3/2}^r \bar{h}}}{f} \quad (2.216)$$

is the internal deformation radius with  $g_{3/2}^r = g(\rho_2 - \rho_1)/\rho_{\text{ref}}$  the reduced gravity and  $\bar{h} = A^{-1} \int (\eta_{1/2} - \eta_{3/2}) dS$  the area averaged layer thickness (see Figure 1.4). The scaling (2.215) means that for scales larger than the internal deformation radius,  $L_d^{\text{int}}$ , the available potential energy is larger than the kinetic energy, and conversely for scales smaller than  $L_d^{\text{int}}$ .

To solve this exercise you must make use of the following.

- Derive the APE for two layers with a flat bottom.

- Assume the contribution to the APE from free surface undulations, is much smaller than from the interior interface. So that the APE is roughly due just to undulations of the interior interface. This assumption follows from Figure 1.5.
- The scaling (2.215) is identical to that found for the quasi-geostrophic system in Chapter 11. However, to solve this exercise it is not sufficient to merely reproduce the scaling as pursued in Chapter 11. Instead, use here the expressions for APE and KE appropriate for the shallow water system.



## **Part II**

# **Vorticity**

**Vorticity** locally measures the spin of a fluid element. For geophysical flows, external forces, ultimately due to differential heating over the planet, resupply vorticity in the face of dissipation. The addition of planetary vorticity, arising from motion on a rotating spherical planet, also renders a nonzero vorticity to geophysical fluids even when the fluid is at rest in the rotating Earth reference frame. Vorticity thus plays a central role in characterizing geophysical fluid motions, even for relatively sluggish and laminar flow, since motion on a rotating planet always involves vorticity.

Besides offering a key method for flow characterization, vorticity evolution and its steady balances provide the means to mechanically understand how flows respond under changes to forces. Surprisingly, it does so even without needing to directly compute forces acting on the fluid. We thus encounter examples where vorticity mechanics offers a more direct and focused explanation for flow behavior than momentum or energy mechanics. This practical feature of vorticity mechanics represents the central reason it is so essential to the theoretical machinery of geophysical fluid mechanics.

Potential vorticity is a strategically chosen component of vorticity whose evolution is simpler than the full vorticity vector, thus helping to identify key facets of geophysical flows, their forcing, and constraints. Indeed, under certain assumptions of balance, knowledge of potential vorticity offers the means to deduce all prognostic information about certain rotating and stratified flows. For these and other reasons explored in this part of the book, potential vorticity has found great use for understanding and predicting geophysical fluid flows. The central importance of potential vorticity for the study of atmospheric and oceanic flows helps to distinguish geophysical fluid mechanics from other areas of fluid mechanics.

#### OUTLINE FOR THIS PART OF THE BOOK

We start this part of the book by introducing vorticity and circulation in Chapter 3, making use of **Stokes' theorem** to show that the area integral of vorticity over a finite region yields the circulation around the region's boundary. In Chapter 4 we study vorticity in a horizontal flow that is non-divergent, thus leading to the study of non-divergent barotropic flow. This flow is fully described by the vorticity field, and it offers many insights into large-scale vortical flows in the atmosphere and ocean. Chapter 5 then introduces the mechanics of vorticity and potential vorticity within a shallow water fluid. It was for the **shallow water model** that *Rossby* (1940) revealed the power of potential vorticity conservation for understanding geophysical fluid flow patterns.

In Chapter 6 we fully dive into the fundamentals of vorticity and circulation. It is here that we encounter **Kelvin's circulation theorem**, which states that circulation around an arbitrary simply closed loop remains materially conserved for a perfect barotropic flow. In Chapter 7 we explore the foundations of **potential vorticity** and then in Chapter 8 develop differential and integral potential vorticity budget equations. Our study of potential vorticity budgets exposes the remarkable **impermeability theorem** that constrains the evolution of potential vorticity.

#### MATHEMATICS IN THIS PART

The mathematics in this chapter rely mostly on the Cartesian tensor analysis and vector calculus from VOLUME 1. We make extensive use of the integral theorems from vector calculus.

## Chapter 3

# VORTICITY AND CIRCULATION

Vorticity measures the angular motion contained in a fluid flow at each point within the fluid; i.e., it is a measure of spin. Vorticity generalizes to continuum mechanics the notion of angular momentum that is central to the study of rigid body mechanics. We here relate the two, showing that flows with a nonzero strain lead to distinctions between vorticity and angular momentum. Circulation measures the fluid flow computed over a closed line integral (circuit) within the fluid. Helmholtz was an early proponent of vorticity whereas Kelvin introduced circulation to help understand vorticity. These two flow properties are connected through Stokes' theorem, with the study of vortex lines and vortex tubes clearly exposing the connections.

### CHAPTER GUIDE

We here study kinematic properties of vorticity and circulation, making use of vector calculus from VOLUME 1. The concepts and methods introduced in this chapter are fundamental to the study of fluid mechanics.

<b>3.1</b>	<b>Loose threads</b>	<b>90</b>
<b>3.2</b>	<b>Vorticity</b>	<b>90</b>
3.2.1	Rotation of line elements	91
3.2.2	Rotating reference frame	91
3.2.3	There are no vorticity sources	92
3.2.4	Further study	92
<b>3.3</b>	<b>Irrational flows</b>	<b>92</b>
3.3.1	Characterizing irrotational flows	93
3.3.2	Comments	93
<b>3.4</b>	<b>Circulation of the velocity field</b>	<b>94</b>
<b>3.5</b>	<b>The free vortex</b>	<b>95</b>
3.5.1	Velocity	95
3.5.2	Vorticity	96
3.5.3	Angular momentum	96
3.5.4	Circulation	97
<b>3.6</b>	<b>Translation and rigid-body rotation</b>	<b>97</b>
3.6.1	Absolute vorticity	97
3.6.2	Rigid-body rotation on a plane	98
3.6.3	Circulation for rigid-body rotation	98
3.6.4	Comments	99
<b>3.7</b>	<b>Kinematics of vortex lines and vortex tubes</b>	<b>99</b>
3.7.1	Vortex lines and vortex tubes	99

---

3.7.2	Kinematic properties . . . . .	100
3.7.3	Helmholtz's theorems . . . . .	101
3.7.4	Further study . . . . .	102
<b>3.8</b>	<b>Relative vorticity from curvature and shear . . . . .</b>	<b>103</b>
3.8.1	Circular flow . . . . .	103
3.8.2	Generalization to natural coordinates . . . . .	104
3.8.3	Example vorticities . . . . .	104
<b>3.9</b>	<b>Relating angular momentum to vorticity and strain . . . . .</b>	<b>106</b>
3.9.1	Linear momentum . . . . .	107
3.9.2	Angular momentum . . . . .	108
3.9.3	Taylor expanding the velocity . . . . .	109
3.9.4	Angular momentum, strain rate, and vorticity . . . . .	109
3.9.5	Comments and further reading . . . . .	110
<b>3.10</b>	<b>Exercises . . . . .</b>	<b>111</b>

---

## 3.1 Loose threads

- Write up the solution to Exercise 3.7.

## 3.2 Vorticity

Vorticity is the curl of the velocity field

$$\boldsymbol{\omega} = \nabla \times \mathbf{v}. \quad (3.1)$$

Vorticity measures the rotation or spin of fluid flow at each point and, unlike angular momentum, it does so without reference to an origin. In this manner, vorticity is an intrinsic property of the flow. In addition to writing vorticity as the curl of the velocity, we may choose to use the equivalent expression

$$\boldsymbol{\omega} = [\nabla \cdot (\mathbf{v} \times \hat{\mathbf{x}})] \hat{\mathbf{x}} + [\nabla \cdot (\mathbf{v} \times \hat{\mathbf{y}})] \hat{\mathbf{y}} + [\nabla \cdot (\mathbf{v} \times \hat{\mathbf{z}})] \hat{\mathbf{z}}. \quad (3.2)$$

That is, a vorticity component in a particular coordinate direction is the divergence of the velocity field after being rotated by  $-\pi/2$  around the coordinate axis direction. For example, the vector  $\mathbf{v} \times \hat{\mathbf{z}}$  is the result of rotating the velocity by  $-\pi/2$  radians around the  $\hat{\mathbf{z}}$  axis, with the identity

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{v}) = \nabla \cdot (\mathbf{v} \times \hat{\mathbf{z}}) = \partial_x v - \partial_y u \quad (3.3)$$

leading to the vertical component of the vorticity in equation (3.2).

The vorticity transforms as a vector under coordinate rotations. However, vorticity changes sign under a mirror symmetry transformation (i.e., right hand axes transformed to left hand axes), thus making it an **axial vector** (VOLUME 1). A simple means to understand this property is to note that the spinning earth rotates counter-clockwise when viewed from above the north pole and clockwise when viewed from below the south pole.

Figure 3.1 provides an example zonal flow with a meridional strain (shear). The vertical component to the vorticity is negative for this flow, as per the right hand rule

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{v}) = \partial_x v - \partial_y u < 0. \quad (3.4)$$

Furthermore, an imaginary test “paddle wheel” placed anywhere within this flow spins clockwise

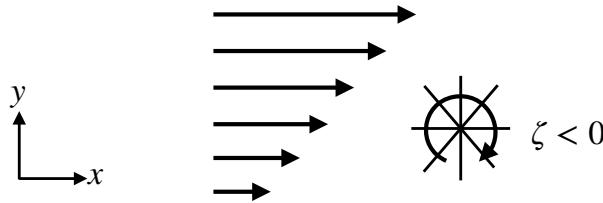


FIGURE 3.1: An example zonal flow with a meridional shear,  $\mathbf{v} = u(y)\hat{x}$ , and a corresponding vertical vorticity component that is negative:  $\zeta = \hat{z} \cdot (\nabla \times \mathbf{v}) = \partial v / \partial x - \partial u / \partial y = -\partial u / \partial y < 0$ . The clockwise arrow surrounds a test “paddle wheel” that exhibits a clockwise spin about its axis when placed in this flow. Such test paddle wheels only spin when there is nonzero vorticity. The right hand rule determines the sign of the vorticity, which for this example is into the page (negative  $\hat{z}$ ).

about its axis. The nonzero spin of a test paddle wheel is a fundamental property of fluid flow with nonzero vorticity.

### 3.2.1 Rotation of line elements

In VOLUME 2, we examined considered the kinematics of a material line element,  $\delta\mathbf{x}$ , whose evolution is given by

$$\frac{D(\delta x_m)}{Dt} = \delta x_n \frac{\partial v_m}{\partial x_n} \implies \frac{D(\delta \mathbf{x})}{Dt} = (\delta \mathbf{x} \cdot \nabla) \mathbf{v}. \quad (3.5)$$

This equation says that the material line element evolves according to the velocity gradient tensor  $\partial_n v_m$ . The symmetric portion of this tensor is the strain rate tensor,

$$S_{mn} = \frac{1}{2}(\partial_n v_m + \partial_m v_n), \quad (3.6)$$

whose action generates changes in the distance between the fluid particles (VOLUME 2). The anti-symmetric portion to the velocity gradient tensor is known as the rotation tensor,

$$\mathbb{R}_{mn} = \frac{1}{2}(\partial_n v_m - \partial_m v_n). \quad (3.7)$$

The rotation tensor is related to vorticity via

$$\mathbb{R}_{mn} = -\frac{1}{2}\epsilon_{mnp}\omega_p \iff \mathbb{R} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}, \quad (3.8)$$

so that

$$2\mathbb{R}_{mn}\delta x_n = -\epsilon_{mnp}\omega_p\delta x_n \implies 2\mathbb{R} \cdot \delta \mathbf{x} = \boldsymbol{\omega} \times \delta \mathbf{x}. \quad (3.9)$$

From our discussion of rotation in VOLUME 1, this equation means that vorticity in a fluid generates a rigid rotation of a material line element around the instantaneous axis defined by the vorticity. This result accords with Figure 3.1, whereby vorticity leads to the spin of a test paddle wheel; i.e., the rotation of line elements.

### 3.2.2 Rotating reference frame

For another means to understand the kinematics of vorticity, view the flow field from a reference frame that rotates with a constant angular velocity,  $\boldsymbol{\Gamma}$ , analogous to the case of observing

geophysical flows from the non-inertial terrestrial reference frame. Following our discussion of rotating reference frames in VOLUME 1, we know that the velocity observed in the non-rotating or absolute reference frame,  $\mathbf{v}_a$ , is related to the rotating reference frame velocity,  $\mathbf{v}$ , via

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Gamma} \times \mathbf{x}. \quad (3.10)$$

The vorticity measured in the absolute reference frame,  $\boldsymbol{\omega}_a$ , is related to the relative vorticity measured in the rotating reference frame,  $\boldsymbol{\omega}$ , via

$$\boldsymbol{\omega}_a = \nabla \times \mathbf{v}_a = \nabla \times \mathbf{v} + \nabla \times (\boldsymbol{\Gamma} \times \mathbf{x}) = \boldsymbol{\omega} + 2\boldsymbol{\Gamma}. \quad (3.11)$$

If there is a point in the fluid whereby the rotating reference frame's angular velocity equals to one-half the absolute vorticity at that point,  $\boldsymbol{\Gamma} = \boldsymbol{\omega}_a/2$ , then the rotating reference frame's vorticity (the relative vorticity) vanishes at that point

$$\boldsymbol{\Gamma} = \boldsymbol{\omega}_a/2 \implies \boldsymbol{\omega} = 0. \quad (3.12)$$

Hence, we may interpret  $\boldsymbol{\omega}_a/2$  as twice the local and instantaneous angular velocity of the fluid. Correspondingly, if the absolute vorticity,  $\boldsymbol{\omega}_a$ , is spatially constant, then we can move to a rotating reference frame in which the relative vorticity vanishes everywhere, with such flow referred to as **irrotational**.

### 3.2.3 There are no vorticity sources

Vorticity has zero divergence

$$\nabla \cdot \boldsymbol{\omega} = \nabla \cdot (\nabla \times \mathbf{v}) = 0. \quad (3.13)$$

This property is akin to the non-divergent nature of the velocity vector in an **incompressible flow**. However, vorticity is non-divergent for both compressible and incompressible flow. Consequently, there are no interior sources or sinks of vorticity for any fluid. This very basic kinematic property plays an important role in developing some further properties of vorticity in Chapter 6.

### 3.2.4 Further study

[This video from 3Blue1Brown](#) provides some compelling graphics to help develop intuition for the divergence and curl of a vector, with examples drawn from fluid flow.

## 3.3 Irrotational flows

Most geophysical flows have nonzero vorticity. Indeed, even when at rest on the earth, a geophysical fluid carries the vorticity of the rotating planet. However, if we can ignore the planetary vorticity component, as when focused on motions too small to feel the Coriolis acceleration, we can find some geophysically relevant flows with vanishing vorticity. Linear gravity waves in the absence of planetary rotation provide a particularly relevant example (VOLUME 4). There are also many examples from engineering flows.

### 3.3.1 Characterizing irrotational flows

Irrotational fluid flow is characterized by a zero vorticity

$$\boldsymbol{\omega} = 0 \text{ = irrotational flow.} \quad (3.14)$$

Since the curl of a gradient vanishes, irrotational flow has a velocity field equal to the gradient of a velocity potential

$$\nabla \times \mathbf{v} = 0 \implies \mathbf{v} = \nabla \Psi. \quad (3.15)$$

Irrotational flow is therefore sometimes called **potential flow**. Figure 3.2 illustrates a two-dimensional flow field generated by taking the gradient of a scalar potential so that the flow has zero vorticity. In this case, the vertical component of the vorticity vanishes at each point since  $\partial v / \partial x = \partial u / \partial y$ .

If the flow is non-divergent, as in a **Boussinesq ocean**, then the velocity potential is a **harmonic function** since it satisfies Laplace's equation

$$\nabla \cdot \mathbf{v} = 0 \implies \nabla^2 \Psi = 0. \quad (3.16)$$

The study of harmonic functions is a very mature area of mathematical physics, thus providing a great deal of analytic power towards the study of potential / non-divergent flows.

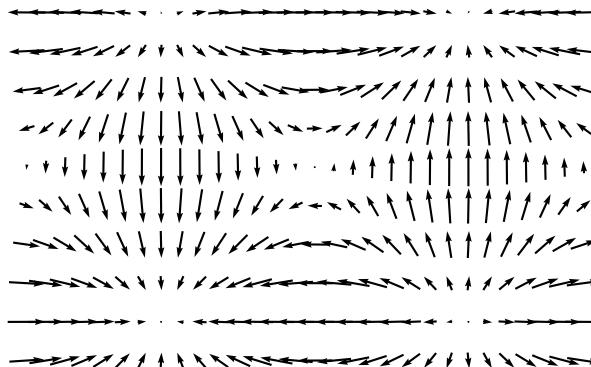


FIGURE 3.2: An example horizontal flow based on a potential,  $\Psi = \sin(x/5) \sin(y/5)$ . The flow has zero vorticity,  $\boldsymbol{\omega} \cdot \hat{z} = \zeta = \partial v / \partial x - \partial u / \partial y = 0$ , since the flow is based on a scalar potential:  $\boldsymbol{\omega} = \nabla \times \mathbf{v} = \nabla \times \nabla \Psi = 0$ . This example illustrates how irrotational flow may have nontrivial structure even though a test paddle wheel will not spin anywhere in the flow, since there is zero vorticity given that  $\partial v / \partial x = \partial u / \partial y$ .

### 3.3.2 Comments

This book does not discuss turbulence in any depth. Nevertheless, we here note that three dimensional turbulence fundamentally relies on vorticity. Hence, irrotational flows, though they may exhibit chaotic motions, are not turbulent since they do not allow for the nonlinear cascade of energy to small spatial scales, with this cascade a fundamental characteristic of three dimensional turbulence. As we see in Section 6.3, vorticity evolves from sources that tilt and stretch vortex tubes. Vortex stretching is the key source for the turbulent cascade in three dimensional turbulence. Section 3.3 of [Tennekes and Lumley \(1972\)](#) provides a pedagogical discussion of vorticity in the context of three-dimensional turbulence.

## 3.4 Circulation of the velocity field

The velocity circulation, or more briefly the **circulation**, is defined as the oriented closed loop line integral of velocity as projected onto the unit tangent of the path

$$\mathcal{C} \equiv \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x}, \quad (3.17)$$

with Figure 3.3 offering a schematic. The line element,  $d\mathbf{x}$ , is oriented in the counter-clockwise direction around the circuit  $\partial S$ . More precisely, let  $\mathbf{x}(\varphi)$  be an expression for the position of a point on the circuit, with  $\varphi(x, y, z, t)$  a parameter that measures the distance along the closed circuit.<sup>1</sup> The difference between two very close positions along the circuit defines the increment

$$d\mathbf{x} = \mathbf{x}(\varphi + \delta\varphi) - \mathbf{x}(\varphi). \quad (3.18)$$

By construction,  $d\mathbf{x}$  is tangent to the circuit so that  $\mathbf{v} \cdot d\mathbf{x}$  picks out the component of the velocity that is tangent to the path.

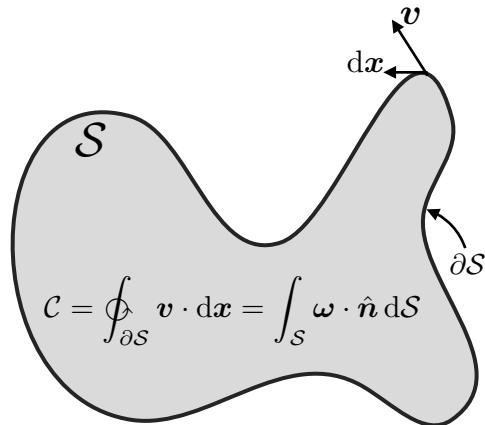


FIGURE 3.3: The velocity circulation around the boundary of a surface,  $\partial S$ , is determined by the line integral of the velocity projected into the direction of the line integral,  $\mathbf{v} \cdot d\mathbf{x}$ . Stokes' theorem shows that the velocity circulation computed as a line integral is identical to the normal projection of the vorticity integrated over the area of the region,  $\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS$ .

Stokes' theorem renders the very important identity

$$\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_S (\nabla \times \mathbf{v}) \cdot \hat{\mathbf{n}} dS = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (3.19)$$

where  $\hat{\mathbf{n}}$  is the outward unit normal vector orienting the area according to the right-hand rule applied to the bounding circuit. The area integral expression motivates interpreting velocity circulation as the “integrated flux of vorticity” that penetrates the surface. Stokes' theorem provides the means to connect the vorticity theories promoted by Helmholtz to the circulation theories of Kelvin.

---

<sup>1</sup>See our discussion of line integrals in VOLUME 1 for more about how to parameterize a curve for integration.

## 3.5 The free vortex

Consider a two-dimensional rotating fluid in the  $x$ - $y$  plane with angular velocity given by

$$\boldsymbol{\Omega} = \frac{\mathbf{x} \times \mathbf{v}}{r^2} = \frac{K \hat{\mathbf{z}}}{r^2}. \quad (3.20)$$

The constant  $K$  has dimensions  $L^2 T^{-1}$ , and  $r^2 = x^2 + y^2$  is the squared distance from the axis of rotation with  $\hat{\mathbf{z}}$  the unit vector normal to the  $x$ - $y$  plane. The angular velocity falls off as the squared distance from the center, whereas it is singular at the origin. As shown in this section, the fluid flow associated with this [free vortex](#) has zero vorticity and zero circulation for all points except the origin. Yet the same points with zero vorticity and zero circulation have a constant angular momentum relative to the origin. As shown by [Exercise 3.5](#), and pursued in more detail in [Section 3.9](#), nonzero angular momentum can arise in a fluid with zero circulation so long as there is a nonzero strain within the fluid, such as the flow arising from the free vortex.

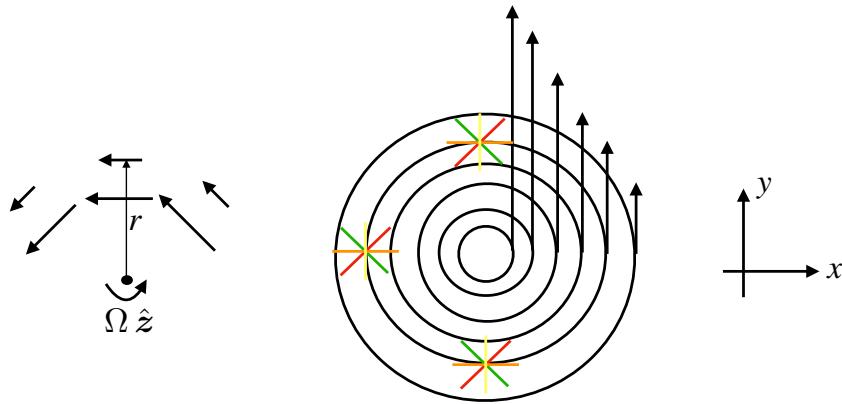


FIGURE 3.4: Irrotational counter-clockwise planar flow in the presence of a free vortex with velocity  $\mathbf{v} = (K/r) \hat{\boldsymbol{\vartheta}}$ . The tangential velocity decays as  $1/r$  from the origin and the vorticity,  $\nabla \times \mathbf{v}$ , vanishes for all points except the origin. Test paddle wheels (colored line segments) do not spin when placed anywhere except at the origin. The free vortex has constant angular momentum per mass (computed relative to the origin), since the tangential velocity falls off as  $1/r$  thus canceling the moment-arm distance  $r$ .

### 3.5.1 Velocity

Fluid flows in a circular orbit when in the free vortex flow field. Hence, the velocity of a fluid particle is perpendicular to its position vector,  $\mathbf{x} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}}$ , with respect to the origin

$$\mathbf{v} \cdot \mathbf{x} = 0. \quad (3.21)$$

The velocity for pure rotational flow is given by<sup>2</sup>

$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{x} = \frac{K(-y \hat{\mathbf{x}} + x \hat{\mathbf{y}})}{r^2} = \frac{K \hat{\boldsymbol{\vartheta}}}{r}, \quad (3.22)$$

where  $\hat{\boldsymbol{\vartheta}}$  is the polar angle unit vector oriented in the counter-clockwise direction. We illustrate the velocity field (3.22) in Figure 3.4, which reveals the  $1/r$  behavior with a singularity at the origin.

<sup>2</sup>See our study of rotational kinematics in [VOLUME 1](#).

### 3.5.2 Vorticity

Away from the origin the vorticity vector vanishes

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} = 0, \quad (3.23)$$

whereas it is singular at the origin. It is useful to expose a few details of this calculation by considering the vertical component to the vorticity, as computed using the polar coordinate curl (VOLUME 1), in which

$$\hat{z} \cdot (\nabla \times \mathbf{v}) = r^{-1} \partial_r(r v^\vartheta) - r^{-1} \partial_\vartheta v^r. \quad (3.24)$$

Since  $v^r = 0$  there is a contribution only from the first term. Yet for the velocity (3.22) we have  $r v^\vartheta = K$ , so that  $\partial_r(r v^\vartheta) = 0$ . It is further insightful to perform the product rule to render

$$\hat{z} \cdot (\nabla \times \mathbf{v}) = v^\vartheta/r + \partial_r v^\vartheta = K/r^2 - K/r^2 = 0, \quad (3.25)$$

which reveals that zero vorticity arises from an exact compensation between the curvature induced vorticity,  $v^\vartheta/r$ , and the normal shear induced vorticity,  $\partial_r v^\vartheta$

$$v^\vartheta/r = -\partial_r v^\vartheta = K/r^2. \quad (3.26)$$

We return to this result in Section 3.8, where we present a general means to decompose the vorticity into that arising from curvature in the flow plus that arising from normal shears (see Figure 3.10). The free vortex is a special case where these two contributions exactly counteract one another, thus leaving zero net vorticity.

In Figure 3.4 we exhibit a test paddle wheel in various positions around the free vortex. As the paddle wheel center moves counter-clockwise with the flow, the marked paddle wheel blades remain oriented at the same fixed angle. That is, the paddle wheel orbits around the vortex center but it does not spin since the vorticity vanishes in the region bounded away from the origin. Again, the vorticity vanishes in this case since the curvature induced vorticity exactly counteracts the normal shear induced vorticity, as per equation (3.26).

### 3.5.3 Angular momentum

Although vorticity is zero everywhere, except at the origin, the angular momentum (computed relative to the origin) is nonzero, as expected since the fluid is rotating around the vortex center. The angular momentum arises just from the nonzero strain in the flow field (see Exercise 3.5), with the strain causing fluid particles to move relative to one another. The angular momentum per unit mass, relative to the center of the vortex, is constant and pointed vertically

$$\mathbf{x} \times \mathbf{v} = r \hat{\mathbf{r}} \times (K/r) \hat{\vartheta} = K \hat{z}. \quad (3.27)$$

This result follows since the velocity falls off as  $1/r$  to cancel the moment-arm distance,  $r$ . Hence, the angular momentum per mass is the same for all fluid particles in the presence of a free vortex, no matter what radial distance the particles have from the vortex center.

### 3.5.4 Circulation

The circulation vanishes for any circuit bounded away from the origin since vorticity vanishes away from the origin. However, the circulation is nonzero for any circuit enclosing the origin

$$\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_0^{2\pi} \mathbf{v} \cdot \hat{\boldsymbol{\vartheta}} r d\vartheta = 2\pi K. \quad (3.28)$$

To reach this result, we set the line element to

$$d\mathbf{x} = \hat{\boldsymbol{\vartheta}} r d\vartheta \quad (3.29)$$

and inserted the velocity (3.22) represented in cylindrical polar coordinates,  $\mathbf{v} \cdot \hat{\boldsymbol{\vartheta}} = K/r$ . Hence, the singular point vortex at  $r = 0$  induces a nonzero circulation for all circuits that enclose the origin.

## 3.6 Translation and rigid-body rotation

Rigid-body fluid motion occurs when all fluid particles are rigidly locked into their relative positions. There are two kinds of rigid body motion: translation and rotation. The velocity field for this motion is given by

$$\mathbf{v} = \mathbf{U} + \boldsymbol{\Gamma} \times \mathbf{x}, \quad (3.30)$$

where  $\mathbf{x}$  is the position vector relative to the origin,  $\mathbf{U}$  is a translation velocity, and  $\boldsymbol{\Gamma}$  is an angular velocity. For rigid body motion, both  $\mathbf{U}$  and  $\boldsymbol{\Gamma}$  are spatially uniform, but can in general be time dependent. The [strain rate tensor](#) vanishes for uniform translation or rigid-body motion

$$S_{mn} = \frac{1}{2}(\partial_m v_n + \partial_n v_m) = 0. \quad (3.31)$$

A zero strain rate tensor is expected since strain measures the relative motion between fluid particles, and for a rigid-body motion there is no such motion. Even so, the vorticity for rigid-body flow is nonzero (see Exercise 3.2)

$$\boldsymbol{\omega} = \nabla \times (\boldsymbol{\Gamma} \times \mathbf{x}) = 2\boldsymbol{\Gamma}. \quad (3.32)$$

We encountered this vorticity in Section 3.2.2 when connecting vorticity and angular velocity. In the remainder of this section, we set the rotation rate to that of the planet,  $\boldsymbol{\Gamma} = \boldsymbol{\Omega}$ , and assume it to be constant in space and time.

### 3.6.1 Absolute vorticity

For planetary fluid mechanics, planetary rotation imparts [planetary vorticity](#) to fluids. Hence, the total or [absolute vorticity](#) of a fluid is the vector sum of the [relative vorticity](#),  $\boldsymbol{\omega}$ , plus the planetary vorticity

$$\boldsymbol{\omega}_a = \boldsymbol{\omega}_{\text{planet}} + \boldsymbol{\omega}. \quad (3.33)$$

In this equation,

$$\boldsymbol{\omega}_{\text{planet}} = 2\boldsymbol{\Omega}_{\text{planet}} \quad (3.34)$$

is the planetary vorticity associated with rigid-body motion of a fluid particle stationary with respect to the planet, and

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} \quad (3.35)$$

is the relative vorticity. The relative vorticity measures the vorticity of the fluid due to motion relative to the rotating sphere, with  $\mathbf{v}$  the velocity relative to the rotating sphere.

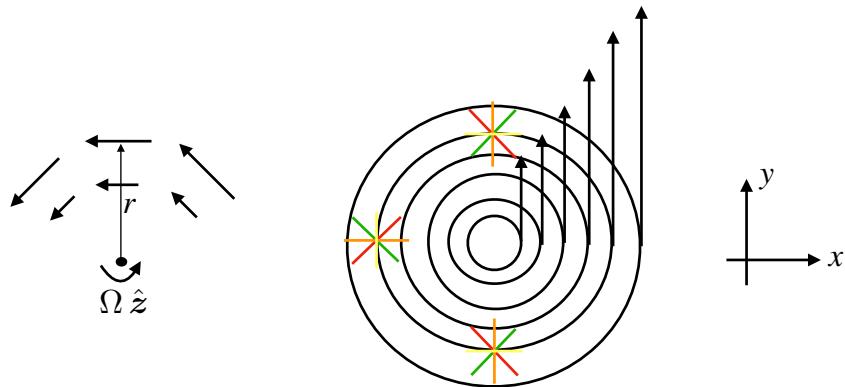


FIGURE 3.5: Rigid body fluid motion, whereby the fluid velocity is: (A) purely tangential and linearly proportional to the radial distance from the vortex center,  $\mathbf{v} = |\Omega| r \hat{\vartheta}$ ; (B) fluid particles maintain a fixed relative position; (C) and vorticity is constant and points perpendicular to the page,  $\boldsymbol{\omega} = 2\boldsymbol{\Omega} = 2|\Omega| \hat{z}$ . Test paddle wheels rigidly move around the center, and they exhibit a spin about their axis that manifests the nonzero vorticity.

### 3.6.2 Rigid-body rotation on a plane

Consider the circular rigid-body rotation on a plane shown in Figure 3.5, in which the velocity is purely tangential and linearly proportional to the distance from the center

$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{x} = |\Omega| (-y \hat{\mathbf{x}} + x \hat{\mathbf{y}}) = |\Omega| r \hat{\vartheta}. \quad (3.36)$$

Assuming the center of mass to be at the circle center, the angular momentum for the flow is the same as that for a rigid-body. Even though the motion of each fluid particle is rigidly fixed relative to all other particles, there is a nonzero vorticity in this flow as illustrated by the spin of colored test paddle wheels in Figure 3.5.

### 3.6.3 Circulation for rigid-body rotation

For rigid-body rotation, the velocity circulation around a circular path of radius  $R$  is given by

$$\mathcal{C} = \oint \mathbf{v} \cdot d\mathbf{x} = \oint (\boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{x} = R^2 |\Omega| \oint d\vartheta = 2\pi R^2 |\Omega| = 2A |\Omega|, \quad (3.37)$$

where  $A = \pi R^2$  is the area of the circle. Hence, the circulation per area for rigid-body rotating fluid flow is twice the angular rotation rate, which is the magnitude of the vorticity

$$\mathcal{C}/A = |\boldsymbol{\omega}| = 2 |\boldsymbol{\Omega}|. \quad (3.38)$$

### 3.6.4 Comments

As seen in Section 3.5, fluid flow in the presence of a free vortex has zero vorticity for all points except the origin of the vortex. However, the same points also have a constant angular momentum relative to the origin, and they experience a nonzero strain. In contrast, constant rigid-body rotating fluid flow has a nonzero vorticity, nonzero angular momentum, yet a zero strain. Section 3.9 details the connection between vorticity, strain, and angular momentum, where we see that angular momentum can be nonzero if either vorticity or strain are nonzero. These ideas are illustrated in [this 3-minute video](#) as well as in this [10 minute video from the Open University](#).

## 3.7 Kinematics of vortex lines and vortex tubes

We here develop the basics of vortex kinematics, with this discussion closely following from the kinematics of material line elements discussed in VOLUME 2.

### 3.7.1 Vortex lines and vortex tubes

A **vortex line** is a curve in the fluid that is instantaneously tangent to the vorticity at each point along the curve.<sup>3</sup> That is, the collection of vortex lines provides the collection of integral curves for the vorticity field. A vortex line is mathematically parameterized just like any other curve, whereby we write the spatial coordinates along the curve as a function of a suitable parameter  $\varphi$  (e.g., the arc-length)

$$\mathbf{x}(\varphi) = x(\varphi) \hat{\mathbf{x}} + y(\varphi) \hat{\mathbf{y}} + z(\varphi) \hat{\mathbf{z}}. \quad (3.39)$$

Correspondingly, the tangent vector for the curve is given by

$$\frac{d\mathbf{x}(\varphi)}{d\varphi} = \frac{dx(\varphi)}{d\varphi} \hat{\mathbf{x}} + \frac{dy(\varphi)}{d\varphi} \hat{\mathbf{y}} + \frac{dz(\varphi)}{d\varphi} \hat{\mathbf{z}}. \quad (3.40)$$

The three coordinates of the vortex line are constrained so that the tangent is parallel to vorticity at each point

$$\frac{d\mathbf{x}(\varphi)}{d\varphi} \times \boldsymbol{\omega} = 0, \quad (3.41)$$

which is satisfied by the following constraint

$$\frac{dx/d\varphi}{\omega_x} = \frac{dy/d\varphi}{\omega_y} = \frac{dz/d\varphi}{\omega_z}. \quad (3.42)$$

These equations are directly analogous to those satisfied by velocity **streamlines** as discussed in VOLUME 2. Notably, the velocity is not constant along a velocity streamline, nor is vorticity constant along a vortex line. In a steady state, streamlines are the same as fluid particle trajectories. However, a vortex line does not offer an interpretation in terms of particle trajectories.

A **vortex tube** is the accumulation of a bundle of vortex lines that pass through a simple closed curve such as that illustrated in Figure 3.6. By definition, the sides of the vortex tube

---

<sup>3</sup>As noted in Section 1.4 of [Saffman \(1992\)](#), a **vortex filament** is a **vortex tube** surrounded by irrotational fluid, which contrasts to the more general concept of a vortex line.

are parallel to the vorticity field, since the sides are constructed from vortex lines.<sup>4</sup>

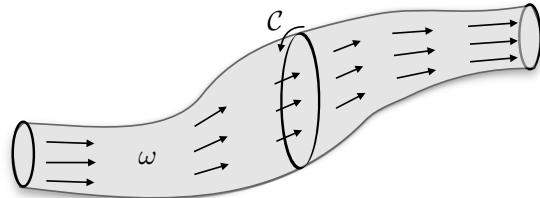


FIGURE 3.6: A vortex line is a line in the fluid that is everywhere tangent to the vorticity vector. A **vortex tube** is the accumulation of vortex lines passing through a closed loop. A related, but distinct, concept arises when a vortex tube is surrounded by irrotational flow, with such tubes referred to as **vortex filaments** (see Section 1.4 of [Saffman \(1992\)](#)). We here depict a vortex tube and circulation computed around the tube  $C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int \omega \cdot \hat{\mathbf{n}} dS$ . Since vorticity has zero divergence, the circulation is the same for any loop embracing the vortex tube (Helmholtz's first theorem from Section 3.7.3). A uniform circulation along the tube means that the magnitude of the vorticity is larger in regions where the tube has a small area and conversely the circulation magnitude is smaller where the tube has a large area. Note that orientation of the circulation integral must accommodate the oppositely directed outward normals on the tube end caps. That is, circulation is computed around a counter-clockwise orientated path, with this orientation determined relative to the outward normal.

### 3.7.2 Kinematic properties

Vorticity has zero divergence

$$\nabla \cdot \boldsymbol{\omega} = \nabla \cdot (\nabla \times \mathbf{v}) = 0, \quad (3.43)$$

which follows since the divergence of a curl vanishes. Integrating the non-divergence relation over an arbitrary closed volume within the fluid leads to

$$\int_{\mathcal{R}} \nabla \cdot \boldsymbol{\omega} dV = \oint_{\partial \mathcal{R}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0, \quad (3.44)$$

where we made use of Gauss's divergence theorem to reach the surface integral expression, with  $\hat{\mathbf{n}} dS$  the oriented area element on the boundary of the volume,  $\partial \mathcal{R}$ , and  $\hat{\mathbf{n}}$  the outward normal on the boundary. This result means there is no net vorticity entering or leaving an arbitrary closed region. That is, there is a vanishing net integrated “flux” of vorticity across the surface bounding a closed region. Consequently, there are no sources or sinks of vorticity within the fluid, meaning there is no accumulation nor depletion of vorticity within any arbitrary closed region within the fluid.

Now specialize the surface integral in equation (3.44) to a volume along a chosen vortex tube such as in Figure 3.6. The two ends of the tube generally have different cross-sectional areas. The integral over the sides of the vortex tube vanishes since vorticity is parallel to the sides of the tube. Hence, the surface integral only picks up contributions from the two ends of the tube<sup>5</sup>

$$\int_A \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS_A + \int_B \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS_B = 0. \quad (3.45)$$

<sup>4</sup>We defined a similar notion, the streamtube, for a non-divergent velocity in VOLUME 2.

<sup>5</sup>In VOLUME 2, we develop a similar set of results for a streamtube in a non-divergent velocity field.

The outward normals point in the opposite direction so that the flux of vorticity is independent of position along the tube. Stoke's theorem transfers the vorticity constraint to a constraint on the circulation around the circumference of the tube

$$\oint_A \mathbf{v} \cdot d\mathbf{x} + \oint_B \mathbf{v} \cdot d\mathbf{x} = 0. \quad (3.46)$$

Hence, the circulation around the vortex tube is the same no matter where it is computed. The circulation constraints (3.45) and (3.46) are kinematic, holding for any vorticity field. We now consider some consequences of this constraint.

### 3.7.3 Helmholtz's theorems

There are a few basic properties of vorticity that follow from its vanishing divergence. These properties are known as Helmholtz's theorems.

#### Helmholtz's first theorem

Since the cross-sectional slices used to derive the circulation constraint (3.46) are arbitrary, the constraint holds throughout the full extent of the vortex tube. Hence, as noted following equation (3.46), the circulation is the same for any position along the vortex tube; i.e., the strength of a vortex tube is the same value along its length (see Figure 3.6). This result is known as [Helmholtz's first theorem](#).

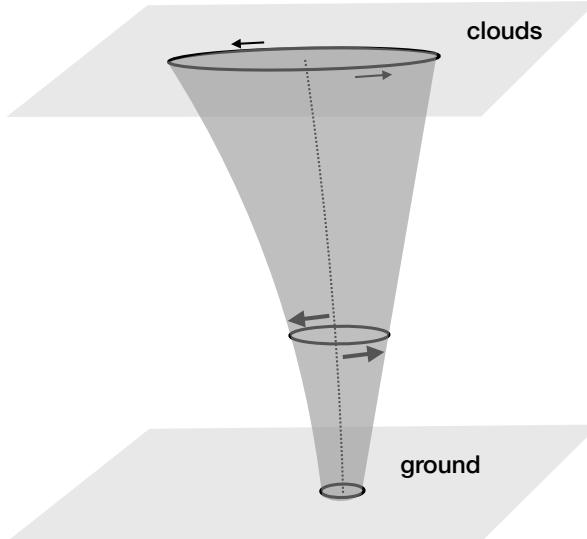


FIGURE 3.7: A vortex tube idealization of a tornado. Since the circulation around the tube is uniform (Helmholtz's first theorem), the tangential velocity of a fluid particle has a larger magnitude in regions where the vortex area is smaller, such as near the ground. As the tornado reaches into the clouds, it generally has a larger cross-sectional area and thus a smaller magnitude for the tangential velocity.

As a corollary, we refer to the vorticity constraint (3.45) to note that any process that changes in the vortex tube cross-sectional area is compensated by changes in vorticity. For example, let the vortex tube shrink over some region. To maintain constant circulation along the tube, the vorticity magnitude must increase where the area decreases, which in turn means that the velocity circulating around the tube increases in magnitude as the area reduces. Think of a tornado as in Figure 3.7, which is a natural expression of a vortex tube. Near the ground,

the cross-sectional area of the tornado is small, with the tangential velocity of a fluid particle within the tube relatively large. Near the tornado top, the cross-sectional area is large so the tangential velocity is relatively small.

### Helmholtz's second theorem

The vorticity constraint (3.45) cannot be satisfied by a finite vorticity if the area of a vortex tube vanishes anywhere. Hence, a vortex tube cannot begin or end within the fluid. This result follows from the absence of vortex sources and sinks within the fluid. Hence, a vortex tube can only loop with itself (e.g., a smoke ring as in Figure 3.8), or intersect a boundary (as for a tornado in Figure 3.7, where the ground and clouds form the boundary).

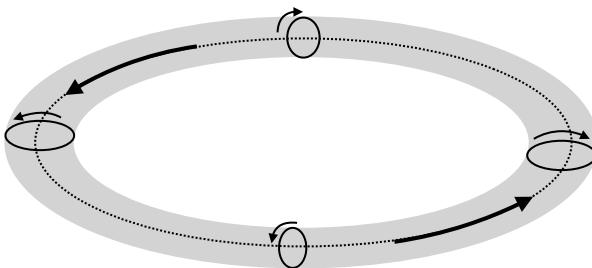


FIGURE 3.8: A vortex ring (torus) is a vortex tube that closes on itself. We here depict a vortex ring with vorticity pointing counter-clockwise around the ring. The tangential velocity is oriented as shown so that the vorticity points according to the right hand rule. That is, orient the fingers on the right hand according to the tangential velocity. The thumb of the right hand then points in the direction of the vorticity vector.

### Helmholtz's third theorem

Helmholtz's third theorem states that an unforced inviscid barotropic fluid that has zero vorticity remains irrotational forever. This theorem is a special case of Kelvin's circulation theorem that is studied in Section 6.2.

#### 3.7.4 Further study

A particularly insightful and pedagogical discussion of the ideas in this section can be found in Chapter 5 of [Acheson \(1990\)](#). Additionally, the following videos offer laboratory demonstrations of vorticity in non-rotating and rotating flows.

- Helmholtz's theorems are vividly exhibited by [this video](#) from the *Physics Girl* of flow generated by a paddle in a swimming pool. She also discusses vortex rings in [this video](#).
- Vorticity and Helmholtz's theorems are also described by [this classic video](#) from Prof. Shapiro.
- A rotating tank experiment shown near the 15 minute mark of [this video](#) from Prof. Fultz shows how vorticity is affected by vortex stretching.
- [This video](#) offers a classic tutorial on vorticity in non-rotating fluids from Prof. Shapiro.

## 3.8 Relative vorticity from curvature and shear

In this section we decompose the vorticity into two terms: one arising from curvature in the flow and another arising from shears in the direction normal to the flow. This decomposition is formulated for horizontal flows, but can be generalized to arbitrary flow. It offers yet another means to understand the kinematic properties of vorticity.

### 3.8.1 Circular flow

Before treating the general case, consider a two-dimensional velocity that locally takes the form of an angular flow

$$\mathbf{u} = u^\vartheta(r, \vartheta) \hat{\boldsymbol{\vartheta}}, \quad (3.47)$$

where  $\vartheta$  is the polar angle. Circulation around the circular wedge shown in Figure 3.9 has zero contributions from the two radial segments since these segments are perpendicular to the angular flow. The circulation is thus given by

$$\mathcal{C} = \oint_{\text{wedge}} \mathbf{u} \cdot d\mathbf{x} \quad (3.48a)$$

$$= \int_{\vartheta}^{\vartheta + \delta\vartheta} u^\vartheta(r + \delta r, \vartheta') (r + \delta r) d\vartheta' + \int_{\vartheta + \delta\vartheta}^{\vartheta} u^\vartheta(r, \vartheta') r d\vartheta' \quad (3.48b)$$

$$= \int_{\vartheta}^{\vartheta + \delta\vartheta} [u^\vartheta(r + \delta r, \vartheta') (r + \delta r) - r u^\vartheta(r, \vartheta')] d\vartheta' \quad (3.48c)$$

$$\approx \int_{\vartheta}^{\vartheta + \delta\vartheta} \left[ u^\vartheta(r, \vartheta') \delta r + \frac{\partial u^\vartheta}{\partial r} r \delta r \right] d\vartheta' \quad (3.48d)$$

$$= r \delta r \int_{\vartheta}^{\vartheta + \delta\vartheta} \left[ \frac{u^\vartheta}{r} + \frac{\partial u^\vartheta}{\partial r} \right] d\vartheta', \quad (3.48e)$$

where the approximation holds when  $\delta r \rightarrow 0$ . Taking the further limit  $\delta\vartheta \rightarrow 0$  renders

$$\mathcal{C} \approx \zeta \delta A = \zeta r \delta r \delta\vartheta = \left[ \frac{u^\vartheta}{r} + \frac{\partial u^\vartheta}{\partial r} \right] r \delta r \delta\vartheta \implies \zeta = \frac{u^\vartheta}{r} + \frac{\partial u^\vartheta}{\partial r}. \quad (3.49)$$

The first term in the vorticity arises from the nonzero radius of curvature of the circular flow whereas the second term arises from radial shear.

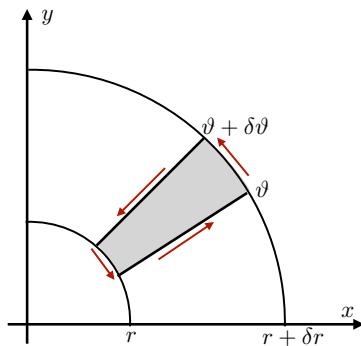


FIGURE 3.9: Circulation around the circular wedge  $[r, r + \delta r] \otimes [\vartheta, \vartheta + \delta\vartheta]$ .

### 3.8.2 Generalization to natural coordinates

The decomposition (3.49) can be generalized to arbitrary horizontal flow by making use of the [natural coordinates](#) from [VOLUME 2](#). Here, we introduce the locally orthogonal coordinates,  $(s, n)$ , with  $s$  the arc-length defined along the trajectory of a fluid element and  $n$  measuring the distance normal to the trajectory. We make the convention that the unit tangent direction,  $\hat{u}$ , is aligned along the local flow direction whereas the unit normal direction,  $\hat{n}$ , is to the left facing downstream. Furthermore, the radius of curvature at a point along a trajectory (i.e., the radius of the [curvature circle](#) constructed along the trajectory) is positive if the flow turns into the positive  $\hat{n}$  direction (left turn) and negative for oppositely curved flow (right turn).<sup>6</sup> Finally, the radius of curvature is infinite for straight flow.

For the counter-clockwise circuit in Figure 3.9, a left turn occurs with  $\hat{n} = -\hat{r}$  so that equation (3.49) takes on the general form

$$\zeta = \underbrace{\frac{|\mathbf{u}|}{R}}_{\text{curv}} - \underbrace{\frac{\partial |\mathbf{u}|}{\partial n}}_{\text{shear}} = \zeta_{\text{curv}} + \zeta_{\text{shear}}. \quad (3.50)$$

Again, the first term arises from curvature in the flow, with  $R$  the radius of curvature. This [curvature vorticity](#) is sometimes also called the [orbital vorticity](#). A trajectory turning to the left has  $R > 0$  and this curved trajectory contributes to a positive vorticity; conversely for a trajectory turning to the right. The second term in equation (3.50) arises from shears computed normal to the flow direction. If the flow speed decreases in the normal direction (e.g., towards the center of the circle in Figure 3.9), then that too contributes to a positive vorticity. Furthermore, flow with  $\zeta = 0$  arises if there is an exact compensation between the curvature-induced vorticity with the shear-induced vorticity

$$\zeta = 0 \implies \frac{|\mathbf{u}|}{R} = \frac{\partial |\mathbf{u}|}{\partial n}. \quad (3.51)$$

### 3.8.3 Example vorticities

#### Rigid body vortex and free vortex

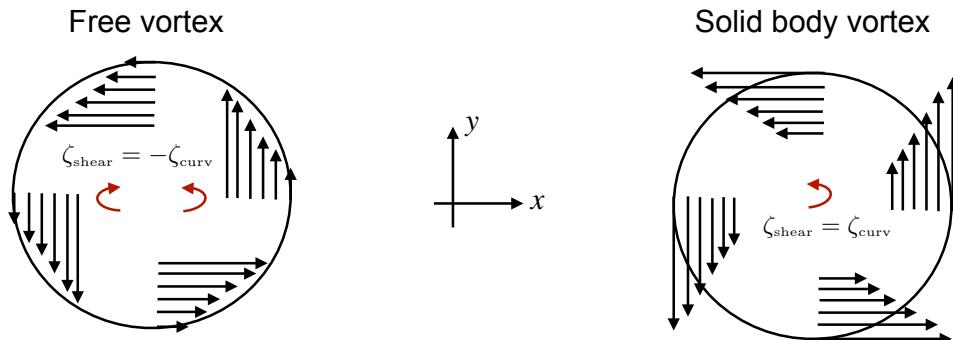


FIGURE 3.10: Decomposing the relative vorticity into its contributions from curvature and normal shear. Left panel: the free vortex from Figure 3.4 has  $\zeta_{\text{curv}} = -\zeta_{\text{shear}} = K/r^2$  so that  $\zeta = \zeta_{\text{curv}} + \zeta_{\text{shear}} = 0$ . Right panel: the rigid-body vortex from Figure 3.5 has  $\zeta_{\text{curv}} = \zeta_{\text{shear}} = \Omega$  so that  $\zeta = 2\Omega$ .

<sup>6</sup>See the discussion of natural coordinates for horizontal flow in [VOLUME 2](#).

To help further understand the decomposition (3.50), consider the case of rigid-body rotation (Figure 3.5) where  $\mathbf{v} = \Omega r \hat{\boldsymbol{\vartheta}}$  and  $\hat{\mathbf{n}} = -\hat{\mathbf{r}}$ . For a circle the radius of curvature equals to the radius of the circle, so that

$$\zeta_{\text{curv}} = \frac{\Omega r}{r} = \Omega \quad \text{and} \quad \zeta_{\text{shear}} = -\partial_n |\mathbf{u}| = \partial_r (\Omega r) = \Omega, \quad (3.52)$$

in which case

$$\zeta = \zeta_{\text{curv}} + \zeta_{\text{shear}} = 2\Omega = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{v}), \quad (3.53)$$

as depicted in the right panel of Figure 3.10. Likewise, for the free vortex (Figure 3.4) we have  $\mathbf{v} = (K/r) \hat{\boldsymbol{\vartheta}}$  so that

$$\zeta_{\text{curv}} = K/r^2 \quad \text{and} \quad \zeta_{\text{shear}} = \partial |\mathbf{u}| / \partial r = -K/r^2, \quad (3.54)$$

which yields zero relative vorticity

$$\zeta = \zeta_{\text{curv}} + \zeta_{\text{shear}} = 0, \quad (3.55)$$

as depicted in the left panel of Figure 3.10.

### Gaussian jet moving along a line and around a circle

Next consider a Gaussian jet moving along a straight line in the meridional direction with velocity field

$$\mathbf{v}(x) = v(x) \hat{\mathbf{y}} = v_0 \exp[-(x - x_{\max})^2/L^2] \hat{\mathbf{y}}, \quad (3.56)$$

where  $v_0$  is the velocity scale,  $L$  is the e-folding length scale for the jet, and  $x_{\max}$  is the position of the jet maximum. The corresponding vorticity of the jet is given by

$$\zeta_{\text{line}} = \partial_x v = -2[(x - x_{\max})/L^2] v(x). \quad (3.57)$$

We depict the velocity (3.56) and vorticity (3.57) in the top row of Figure 3.11. Note the symmetry of the vorticity around the jet maximum at  $x = x_{\max}$ , with the vorticity extrema corresponding to inflection points of the jet.<sup>7</sup>

Now assume the same jet is moving counter-clockwise around a circle so that the velocity field is given by

$$\mathbf{v}(r) = v(r) \hat{\boldsymbol{\vartheta}} = v_0 \exp[-(r - r_{\max})^2/L^2] \hat{\boldsymbol{\vartheta}}, \quad (3.58)$$

where we made use of the polar coordinates from VOLUME 1, with  $r^2 = x^2 + y^2$  the squared radial distance from the center, and  $\hat{\boldsymbol{\vartheta}}$  the unit vector pointing in the counter-clockwise direction from the  $\hat{\mathbf{x}}$  axis. Making use of the curl operation in polar coordinates (again, see VOLUME 1) renders the vorticity

$$\zeta = r^{-1} \partial_r (r v) = \frac{v}{r} - \frac{2(r - r_{\max}) v}{L^2} = \zeta_{\text{curve}} + \zeta_{\text{shear}}. \quad (3.59)$$

We depict these terms in the third row of Figure 3.11. Note how the vorticity,  $\zeta$ , has the amplitude of its maximum increased, whereas the amplitude of its minimum is decreased. Correspondingly, the vorticity is not symmetric about the jet maximum, with its zero crossing

<sup>7</sup>Inflection points are where the curvature vanishes and has opposite signs on either side. For a function of a single variable, an inflection point is where the second derivative vanishes and changes sign when moving to either side.

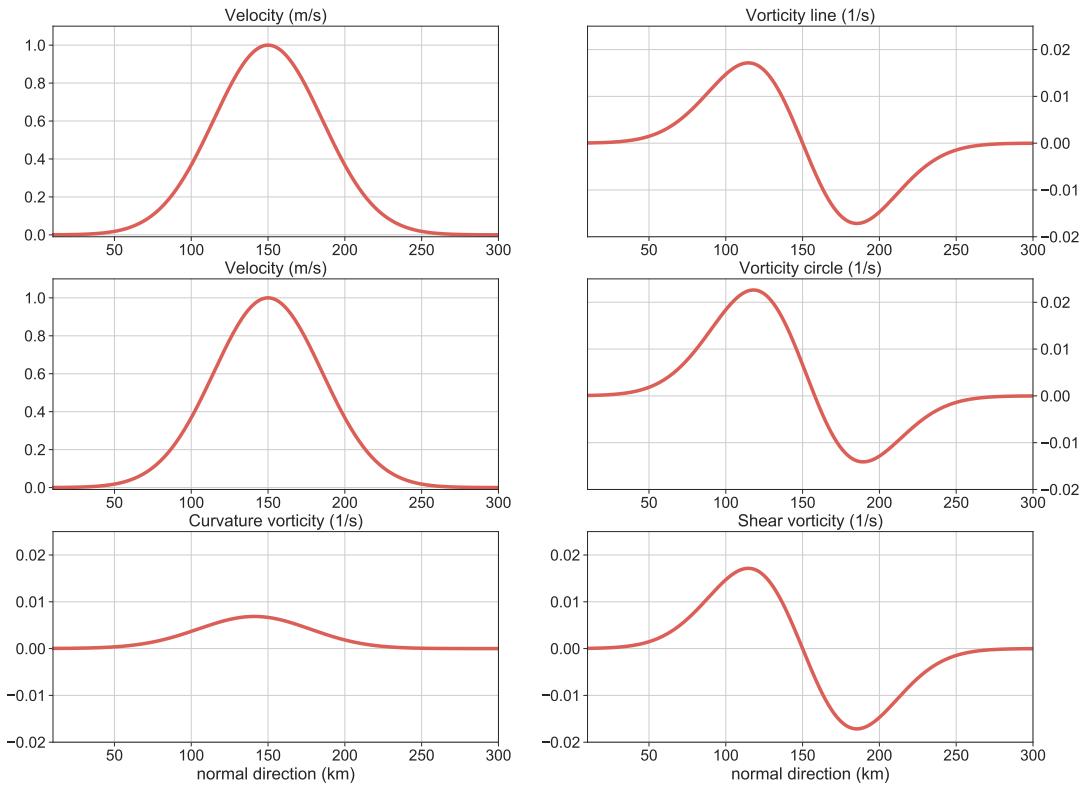


FIGURE 3.11: Top row: velocity and vorticity within a Gaussian jet that is moving along a straight line according to equations (3.56) and (3.57). We set the jet e-folding scale as  $L = 50$  km, jet maximum at  $x_{\max} = 150$  km, and velocity scale as  $v_0 = 1 \text{ m s}^{-1}$ . Note the symmetry of the vorticity relative to the jet maximum, with the vorticity extrema corresponding to the inflection points of the Gaussian velocity field. Middle left: the same Gaussian velocity profile only now for a jet that is moving counter-clockwise around a circle so that  $\mathbf{v} = v(r)\hat{\theta}$  according to equation (3.58). To avoid a singularity of the curvature vorticity at  $r = 0$  (infinite curvature) we assume the velocity field vanishes within a small distance from the origin. Middle right: vorticity for the Gaussian jet moving around the circle, with  $\zeta = r^{-1}\partial_r(rv)$ . Notice how the vorticity is not symmetric relative to the jet maximum. Rather, the zero vorticity occurs to the right of the jet maximum, and the vorticity on the inside of the jet maximum is larger in magnitude than the outer vorticity. Lower left: vorticity due to the curvature of the jet as it moves around the circle,  $\zeta_{\text{curve}} = v(r)/r$ . Lower right: vorticity due to the radial shear in the Gaussian jet,  $\zeta_{\text{shear}} = -2(r - r_{\max})v(r)/L^2$ .

to the outside of the jet maximum. Asymmetry of the vorticity arises from the curvature vorticity,  $\zeta_{\text{curve}} = v/r$ , which has its peak inside the jet maximum. So although the shear vorticity,  $\zeta_{\text{curve}}$ , is symmetric around the jet axis, the curvature vorticity causes a movement of vorticity towards the center of the circle.

## 3.9 Relating angular momentum to vorticity and strain

As noted in Section 3.6.4, fluid flow in the presence of a free vortex (Section 3.5) has zero vorticity for all points except the origin of the vortex. However, the same points with zero vorticity also have a constant angular momentum relative to the origin, and they experience a nonzero strain. In contrast, rigid-body fluid flow (Section 3.6) has a nonzero vorticity, nonzero angular momentum, yet a zero strain rate. In this section we study the connection between vorticity, strain rate, and angular momentum for a fluid. We here show that angular momentum and vorticity are fundamentally distinct in a fluid that has straining motion between fluid

elements. In contrast, angular momentum and vorticity are directly proportional (through the moment of inertia tensor) when the fluid exhibits rigid-body motion in which the strain rate tensor vanishes.

In this section we make use of the [Lagrangian reference frame](#) from [VOLUME 2](#), whereby we write the trajectory of a fluid particle using the motion field. Namely, by fixing the material coordinate,  $\mathbf{a}$ , the motion field provides the spatial position,  $\mathbf{x}$ , of the fluid particle as a function of the Lagrangian time,  $T$

$$\mathbf{x} = \mathbf{X}(\mathbf{a}, T). \quad (3.60)$$

### 3.9.1 Linear momentum

Consider the velocity and linear momentum of a simply connected material fluid region denoted by  $\mathcal{R}(\mathbf{v})$ , with each point of the region moving with the fluid velocity. Let an arbitrary fluid particle within this region be marked with the material label,  $\mathbf{a}$ , so that its position vector is  $\mathbf{X}(\mathbf{a}, T)$  and its velocity is

$$\mathbf{v}^L(\mathbf{a}, T) = \partial_T \mathbf{X}(\mathbf{a}, T), \quad (3.61)$$

where the time derivative is computed holding the material label fixed, and where the “L” superscript signals a Lagrangian velocity. Since the fluid particle is within a finite material region, we can decompose its motion into the sum of the region’s center of mass motion plus motion of the particle relative to the center of mass

$$\mathbf{v}^L(\mathbf{a}, T) = \partial_T \mathbf{X}(\mathbf{a}, T), \quad (3.62a)$$

$$= \partial_T [\bar{\mathbf{X}}(T) + \mathbf{X}'(\mathbf{a}, T)] \quad (3.62b)$$

$$= \bar{\mathbf{v}}^L(T) + \mathbf{v}'^L(\mathbf{a}, T). \quad (3.62c)$$

The velocity,  $\mathbf{v}'^L(\mathbf{a}, T)$ , is defined relative to the region’s center of mass velocity,  $\bar{\mathbf{v}}^L(T)$ . Furthermore, the center of mass velocity is given by

$$\bar{\mathbf{v}}^L = \frac{d\bar{\mathbf{X}}}{dT} \quad (3.63a)$$

$$= \frac{d}{dT} \left[ \frac{\int_{\mathcal{R}(\mathbf{v})} \mathbf{x} \rho dV}{\int_{\mathcal{R}(\mathbf{v})} \rho dV} \right] \quad (3.63b)$$

$$= \frac{1}{M} \int_{\mathcal{R}(\mathbf{v})} \frac{D\mathbf{x}}{Dt} \rho dV \quad (3.63c)$$

$$= \frac{1}{M} \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV. \quad (3.63d)$$

The identity (3.63c) follows since the material region maintains a constant mass,

$$M = \int_{\mathcal{R}(\mathbf{v})} \rho dV \implies \frac{dM}{dt} = 0, \quad (3.64)$$

allowing the denominator to come outside the Lagrangian time derivative. Additionally, each of the fluid parcels in the region maintains constant mass. As per [Reynolds transport theorem](#), the time derivative moves across the integral to act materially on the position vector. The final equality, (3.63d), follows since the material time derivative of a particle trajectory, when

evaluated at a point,  $\mathbf{x}$ , equals to the velocity field at that point

$$\mathbf{v}(\mathbf{x}, t) = \frac{D\mathbf{x}}{Dt}. \quad (3.65)$$

It follows that the linear momentum for the material fluid region is given by

$$\mathbf{P} = \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV = M \bar{\mathbf{v}}^L. \quad (3.66)$$

We conclude that the total linear momentum of an material fluid region equals to that of a point particle of mass,  $M = \int_{\mathcal{R}(\mathbf{v})} \rho dV$ , moving with the center of mass velocity,  $\bar{\mathbf{v}}^L$ . The analogous result was derived in VOLUME 1 for a system of discrete point particles.

### 3.9.2 Angular momentum

Consider angular momentum for a material fluid region, which is determined by the material integral of the angular momentum for each fluid parcel

$$\mathbf{L} = \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x} \times \mathbf{v}) \rho dV. \quad (3.67)$$

Our goal is to expose how physically distinct aspects of the fluid motion contribute to the angular momentum. To proceed, decompose the position vector of a point within the region into the center of mass position plus a deviation,  $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{x}'$ , where  $\bar{\mathbf{x}} = \bar{\mathbf{X}}$  is the instantaneous position of the moving center of mass. The angular momentum thus takes the form

$$\mathbf{L} = \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x} \times \mathbf{v}) \rho dV \quad (3.68a)$$

$$= \int_{\mathcal{R}(\mathbf{v})} [(\bar{\mathbf{x}} + \mathbf{x}') \times \mathbf{v}] \rho dV \quad (3.68b)$$

$$= \bar{\mathbf{X}} \times \left[ \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV \right] + \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \times \mathbf{v}) \rho dV \quad (3.68c)$$

$$= (\bar{\mathbf{X}} \times \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \times \mathbf{v}) \rho dV. \quad (3.68d)$$

The final equality introduced the linear momentum for the fluid region,  $\mathbf{P}$ , in the form of equation (3.66). The first term in equation (3.68d) is the angular momentum of the region with respect to the position of the center of mass. The second term arises from deviations of fluid particle positions relative to the center of mass.

We now focus on how the deviation term,  $\int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \times \mathbf{v}) \rho dV$ , contributes to the angular momentum (3.67). As we will see, this analysis exposes how angular momentum of the extended material fluid region is affected by vorticity and strain rate in the fluid flow. To facilitate some of the manipulations, we make use of basic Cartesian tensor analysis from VOLUME 1, including the Einstein summation convention whereby repeated indices are summed over their range.

### 3.9.3 Taylor expanding the velocity

We perform a Taylor expansion of the velocity  $\mathbf{v}(\mathbf{x})$  around the instantaneous center of mass position,  $\bar{\mathbf{x}} = \bar{\mathbf{X}}$ , and truncate the expansion to the leading order term<sup>8</sup>

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}(\bar{\mathbf{x}} + \mathbf{x}') \approx \mathbf{v}(\bar{\mathbf{x}}) + (\mathbf{x}' \cdot \nabla) \mathbf{v}|_{\mathbf{x}=\bar{\mathbf{x}}}. \quad (3.69)$$

We are thus left with

$$\mathbf{L} = (\bar{\mathbf{X}} \times \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \times \mathbf{v}) \rho dV \quad (3.70a)$$

$$= (\bar{\mathbf{X}} \times \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \times \mathbf{v}(\bar{\mathbf{x}})] \rho dV + \int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \times (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})] \rho dV. \quad (3.70b)$$

The velocity,  $\mathbf{v}(\bar{\mathbf{x}})$ , can be removed from the integration since it is evaluated at the center of mass point. Hence, the second term in equation (3.70b) vanishes

$$\int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \times \mathbf{v}(\bar{\mathbf{x}})] \rho dV = \left[ \int_{\mathcal{R}(\mathbf{v})} \mathbf{x}' \rho dV \right] \times \mathbf{v}(\bar{\mathbf{x}}) = 0, \quad (3.71)$$

where  $\int_{\mathcal{R}(\mathbf{v})} \mathbf{x}' \rho dV = 0$  by definition of the center of mass. The angular momentum is thus given by the two terms

$$\mathbf{L} = (\bar{\mathbf{X}} \times \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \times (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})] \rho dV. \quad (3.72)$$

The  $m'$ th component of the second right hand side term can be written

$$\int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \times (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})]_m \rho dV = \epsilon_{mnp} \int_{\mathcal{R}(\mathbf{v})} x'_n [(\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})]_p \rho dV \quad (3.73a)$$

$$= \epsilon_{mnp} \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \partial_q v_p(\bar{\mathbf{x}}) \rho dV \quad (3.73b)$$

$$= \epsilon_{mnp} \left[ \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \partial_q v_p(\bar{\mathbf{x}}). \quad (3.73c)$$

We removed the velocity derivatives

$$\partial_q v_p(\bar{\mathbf{x}}) = \left[ \frac{\partial v_p}{\partial x_q} \right]_{\mathbf{x}=\bar{\mathbf{x}}} \quad (3.74)$$

from the integral, since they are evaluated at the center of mass point and so do not participate in the integration.

### 3.9.4 Angular momentum, strain rate, and vorticity

Following from the discussion of flow kinematics in VOLUME 2, we know that the velocity derivatives,  $\partial_q v_p$ , appearing in equation (3.73c) form the components to a second order tensor known as the **velocity gradient tensor**, which can be decomposed into the **strain rate tensor**,  $\mathbf{S}$ ,

---

<sup>8</sup>The velocity field evaluated at the center of mass position,  $\mathbf{v}(\bar{\mathbf{x}})$ , is not equal to the center of mass velocity:  $\mathbf{v}(\bar{\mathbf{x}}) \neq \bar{\mathbf{v}}^L$ .

and the rotation tensor,  $\mathbf{R}$ . Introducing these two tensors brings the angular momentum for a continuum fluid region into the form

$$L_m = (\bar{\mathbf{X}} \times \mathbf{P})_m + \epsilon_{mnp} \left[ \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \mathbb{S}_{qp} + \epsilon_{mnp} \left[ \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \mathbb{R}_{qp} \quad (3.75a)$$

$$= (\bar{\mathbf{X}} \times \mathbf{P})_m + \epsilon_{mnp} \left[ \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \mathbb{S}_{qp} + \frac{1}{2} \epsilon_{mnp} \epsilon_{sqp} \left[ \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \omega_s \quad (3.75b)$$

$$= \underbrace{(\bar{\mathbf{X}} \times \mathbf{P})_m}_{\text{center of mass}} + \underbrace{\epsilon_{mnp} \left[ \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \mathbb{S}_{qp}}_{\text{strain rate contribution}} + \underbrace{\frac{1}{2} \left[ \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \cdot \mathbf{x}' \delta_{ms} - x'_m x'_s) \rho dV \right] \omega_s}_{\text{vorticity contribution}}. \quad (3.75c)$$

Since each point in the fluid continuum can be considered the center of mass for an arbitrary material region, the decomposition (3.75c) holds in general.

- **CENTER OF MASS ANGULAR MOMENTUM:** The first term on the right hand side of equation (3.75c) arises from the angular momentum of the material region as measured with respect to the center of mass position. It has the form of angular momentum for a point particle studied in VOLUME 1.
- **STRAINS:** The second contribution is proportional to fluid deformations studied in VOLUME 2. At each point of the fluid, deformations are measured by the strain rate tensor,  $S_{qp}$ . A rigid body moves by uniform translations and/or rigid-body rotations, with the strain rate tensor vanishing for rigid-body motions (see Section 3.6). The contribution from strain rates is weighted by an integral of fluid particle position relative to the center of mass position. A closed form expression for this integral is available only for special shapes.
- **VORTICITY:** The third contribution to angular momentum (3.75c) contains the vorticity as weighted by the moment of inertia tensor

$$I_{ms} \equiv \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \cdot \mathbf{x}' \delta_{ms} - x'_m x'_s) \rho dV. \quad (3.76)$$

Since the material region is evolving and is not rigid, the moment of inertia tensor is time dependent. Even so, the contribution

$$L_m^{\text{vorticity}} \equiv \frac{1}{2} I_{ms} \omega_s \quad (3.77)$$

has the same form as angular momentum for a rigid body, with one-half the vorticity playing the role of angular velocity. Evidently, vorticity in fluid flow contributes to angular momentum for a material region via its product with the moment of inertia tensor for the region.

### 3.9.5 Comments and further reading

Angular momentum is computed relative to a chosen origin, whereas vorticity is an intrinsic property measuring the spin of the fluid at a point. So although they both offer measures of the rotational properties of fluid motion, they are distinct when the fluid experiences a nonzero

strain rate. It is only for the special case of a rigid-body motion that the strain rate vanishes, in which case the angular momentum of a fluid region is directly related to vorticity.

Further discussion of the material in this section can be found in [Chatwin \(1973\)](#), Section 2.3.1 of [Davidson \(2015\)](#), and the online notes “The Vorticity Equation and Conservation of Angular Momentum” from A.J. DeCaria. See also Section 1.2 of the vorticity monograph by [Saffman \(1992\)](#).



## 3.10 Exercises

### EXERCISE 3.1: VORTICITY AT A STATIC NO-SLIP BOUNDARY

Show that  $\boldsymbol{\omega} \cdot \hat{\mathbf{n}} = 0$  with a static no-slip boundary condition, such as a solid static wall. Here,  $\hat{\mathbf{n}} = 0$  is the outward normal at the boundary. Hint: make use of Stoke's theorem and the adherence condition,  $\mathbf{v} = 0$ , at a static no-slip boundary.

### EXERCISE 3.2: VORTICITY FOR RIGID-BODY ROTATION

A fluid in rigid-body rotation has an angular velocity

$$\mathbf{v}_{\text{rigid-body}} = \boldsymbol{\Omega} \times \mathbf{x}, \quad (3.78)$$

with  $\mathbf{x}$  the position vector of a point in the fluid. Show that this fluid velocity has a corresponding vorticity given by

$$\nabla \times \mathbf{v}_{\text{rigid-body}} = 2\boldsymbol{\Omega}. \quad (3.79)$$

There are a few methods to prove this result. Display one method and be sure it is fully explained.

### EXERCISE 3.3: PLANETARY ROTATION IS NON-DIVERGENT

Show that a fluid in rigid-body rotation with angular velocity

$$\mathbf{v}_{\text{rigid-body}} = \boldsymbol{\Omega} \times \mathbf{x}, \quad (3.80)$$

has zero divergence

$$\nabla \cdot \mathbf{v}_{\text{rigid-body}} = 0. \quad (3.81)$$

Consequently, rotation of the planet imparts zero divergence to fluid motion. We make use of this result in part to justify our study of non-rotating fluid kinematics in **VOLUME 2**.

### EXERCISE 3.4: VELOCITY POTENTIAL FOR THE FREE VORTEX

What is the velocity potential (3.15) for the free vortex whose velocity field is given by equation (3.22)? Hint: The problem is two-dimensional and rotationally symmetric, so it is convenient to make use of polar coordinates  $x = r \cos \vartheta$  and  $y = r \sin \vartheta$  from **VOLUME 1**.

### EXERCISE 3.5: STRAIN TENSOR FOR THE FREE VORTEX

Determine all components to the strain rate tensor from **VOLUME 2**

$$S_{pq} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{1}{2} \left[ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right] \\ \frac{1}{2} \left[ \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right] & \frac{\partial v}{\partial y} \end{bmatrix} \quad (3.82)$$

for the free vortex as specified by the velocity field (3.22). Present the answer in the form of a

$2 \times 2$  matrix.

**EXERCISE 3.6: VANISHING VISCOSIVE FRICTION FOR RIGID-BODY MOTION**

As discussed in VOLUME 2, viscous effects from molecular viscosity in a non-divergent flow appear in the momentum equation as a Laplacian weighted by a constant molecular viscosity

$$\text{viscous force per mass} = \nu \nabla^2 \mathbf{v}. \quad (3.83)$$

In this equation,  $\nu > 0$  is the molecular kinematic viscosity, which is a constant that is a property of the fluid. Show that the viscous operator vanishes for a non-divergent flow in rigid-body rotation. That is, rigid-body motion engenders no frictional dissipation. This result reflects the lack of frictional interaction in a strain-free flow.

**EXERCISE 3.7: MOMENT OF INERTIA FOR CYLINDER AND SPHERE**

The moment of inertia for a continuous distribution of matter is defined by equation (3.76). In this exercise we compute the moment of inertia for two matter distributions with constant density but with different shapes. The results can be found through Google, so be sure to check your answer.

- Determine the moment of inertia for a right circular cylinder, where the moment is computed relative to the axis through the center of the cylinder. For simplicity, use cylindrical-polar coordinates and orient the vertical axis through the center of the cylinder.
- Determine the moment of inertia for a sphere computed around an axis through the center of the sphere.

**EXERCISE 3.8: VORTICITY FOR A C-GRID NUMERICAL MODEL**

Vorticity is commonly diagnosed in numerical model simulations. In this exercise we consider how one might determine a discrete equation for the vertical vorticity when the horizontal velocity is arranged according to the Arakawa C-grid (*Arakawa and Lamb, 1977*) commonly used in ocean models, and as depicted in Figure 3.12. Derive an expression for the area averaged vorticity over the shaded grid cell region centered at the vorticity point,  $q_{i,j}$ . Make use of Stokes' theorem with the surrounding C-grid velocity components and the corresponding grid distances.

**EXERCISE 3.9: PARTICLE TRAJECTORIES FROM A POINT VORTEX AND POINT DIVERGENCE**

Consider a two-dimensional (horizontal) flow resulting from a point vortex as well as a point divergence, both at  $\mathbf{x} = 0$ ,

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \mathcal{C} \delta(\mathbf{x}) \quad (3.84a)$$

$$\nabla \cdot \mathbf{u} = \mathcal{D} \delta(\mathbf{x}). \quad (3.84b)$$

We here introduced the constants,  $\mathcal{C} > 0$  and  $\mathcal{D} > 0$ , as well as the Dirac delta,  $\delta(\mathbf{x}) = \delta(x) \delta(y)$ . From VOLUME 1, recall that the Dirac delta satisfies the normalization condition,  $\int_S \delta(\mathbf{x}) d\mathcal{S} = 1$ , for any region,  $S$ , that includes the origin,  $\mathbf{x} = 0$ . Consequently, the Dirac delta has physical dimensions of inverse area.

Throughout this exercise we ignore boundaries, so that the flow is considered on an infinite plane. Also, the reference frame is not rotating. Furthermore, the flow resulting from each point source is axially symmetric, so that the only functional dependence is radial; i.e., distance from the origin. Use polar coordinates (VOLUME 1) throughout this exercise.

- Given the dimensions of the Dirac delta and those for the vorticity and divergence, then

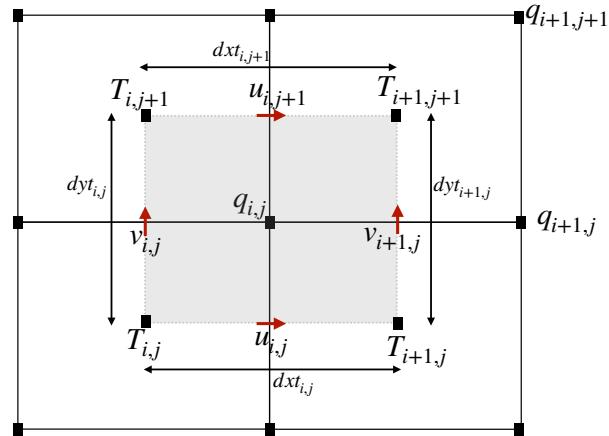


FIGURE 3.12: Layout for velocity on a discrete Arakawa C-grid for use in Exercise 3.8. The central T-point is labeled  $T_{i,j}$  and its corresponding vorticity point,  $q_{i,j}$ , is located to its northeast. This exercise aims to determine the area averaged vorticity for the shaded region. The zonal velocity,  $u_{i,j}$ , is arranged on the east face of the T-cell, whereas the meridional velocity,  $v_{i,j}$ , is on the north face. The zonal and meridional grid distances are indicated, thus measuring distances between adjacent tracer points and so measuring the sides of the shaded region.

what are the physical dimensions of  $\mathcal{C}$  and  $\mathcal{D}$ ?

- What is the horizontal velocity field resulting from the point vortex ( $\mathcal{C} > 0$  and  $\mathcal{D} = 0$ )? Sketch this field.
- What is the expression for the fluid particle trajectory within the velocity field generated by the point vortex?
- What is the horizontal velocity field resulting from the point divergence ( $\mathcal{C} = 0$  and  $\mathcal{D} > 0$ )? Sketch this field.
- What is the expression for the fluid particle trajectory within the velocity field generated by the point divergence?
- Compute the fluid particle trajectories when the divergence source picks up an oscillatory time dependence, so that

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \mathcal{C} \delta(\mathbf{x}) \quad (3.85a)$$

$$\nabla \cdot \mathbf{u} = \mathcal{D} \delta(\mathbf{x}) \sin(\omega_d t). \quad (3.85b)$$

What parameter settings ensure that radial oscillations of the particle position are small relative to the initial radial distance? Only derive the angular position,  $\vartheta(t)$ , assuming these parameter settings.

- Now also include an oscillatory behavior to the vortex source, so that

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \mathcal{C} \delta(\mathbf{x}) \cos(\omega_c t) \quad (3.86a)$$

$$\nabla \cdot \mathbf{u} = \mathcal{D} \delta(\mathbf{x}) \sin(\omega_d t). \quad (3.86b)$$

Compute the fluid particle trajectories assuming the condition from the previous part of this exercise is met. That is, assume the parameter settings ensure that radial oscillations of the particle position are small relative to the initial radial distance.

You may choose to consider the following hints.

- We are only concerned with horizontal motion in this exercise.

- In Section 4.3.8 we study point vortices in the horizontally non-divergent fluid. The velocity field for the point vortex is derived there.
- When both sources are turned on, the velocity field is the linear superposition of the velocity from the point vortex and that from the point divergence. The reason is that we are ignoring any back-reaction of the flow from one source onto the other source.
- It is useful to recall the discussion of fluid particle trajectories from VOLUME 2.



## Chapter 4

# TWO-DIMENSIONAL NON-DIVERGENT FLOW

A single layer of homogeneous shallow water fluid is among the simplest conceptual models available for the study of fluid motion. In the language of vorticity as described in Chapter 6, a single shallow water layer is a **barotropic fluid** since it has zero **baroclinicity** (Section 6.4). Even so, the layer thickness fluctuates as a result of a nonzero horizontal divergence ( $\nabla_h \cdot \mathbf{u} = -\partial_z w \neq 0$ ). We say that the single shallow water layer is a horizontally divergent barotropic model, with this divergence supporting gravity waves as studied in VOLUME 4.

For some geophysical fluid studies, we are primarily interested in the vortical (Rossby wave) motions (see VOLUME 4), with these motions having lower frequency than the gravity wave motions associated with horizontal divergence. The two-dimensional **non-divergent barotropic model** focuses on vortical motion by assuming the horizontal velocity has zero divergence, with this assumption serving to filter out gravity waves. In this chapter we study the non-divergent barotropic model, in which the flow is two-dimensional and so it moves in vertical columns with a fixed thickness. As we find here, this model offers many insights into the nature of geophysical fluid motion. We thus develop some kinematic and dynamic properties of the non-divergent barotropic model, and use this model to exemplify basic features of geophysical flows constrained by conservation of absolute vorticity.<sup>1</sup>

### CHAPTER GUIDE

Besides offering a wealth of insights towards the nature of geophysical flows, the horizontally non-divergent barotropic model offers an ideal platform for the study of vorticity waves (e.g., Rossby waves) and barotropic instability in VOLUME 4. Since all fields in this chapter are a spatial function only of the horizontal position, the gradient operator is itself two-dimensional; e.g.,  $\nabla\psi = \nabla_h\psi$ . For much of the chapter, we use **tangent plane Cartesian coordinates** following from the **tangent plane** approximation. However, in Section 4.9 we consider two-dimensional flow on a sphere and so make use of the spherical coordinates detailed in VOLUME 1.

4.1	Loose threads . . . . .	117
4.2	Basic equations and their properties . . . . .	117
4.2.1	Equations for velocity . . . . .	117
4.2.2	Kinetic energy of the flow . . . . .	118
4.2.3	Kinematics of rigid fluid columns . . . . .	118

<sup>1</sup>This model is the basis for the pioneering numerical weather prediction model of [Charney et al. \(1950\)](#). It is also useful for studies of coherent vortex structures, with Chapter 3 of [McWilliams \(2006\)](#) exploring analytical vortex solutions.

---

4.2.4	Vertical velocity . . . . .	120
4.2.5	Velocity self-advection and the kinetic stress tensor . . . . .	121
4.2.6	Further reading . . . . .	121
<b>4.3</b>	<b>Vorticity . . . . .</b>	<b>121</b>
4.3.1	Vorticity equation . . . . .	122
4.3.2	Constraints from material invariance of absolute vorticity . . . . .	122
4.3.3	Rossby potential vorticity . . . . .	123
4.3.4	Jacobian forms of vorticity advection . . . . .	123
4.3.5	Taylor-Bretherton identity . . . . .	124
4.3.6	Poisson equation for the streamfunction . . . . .	125
4.3.7	Zonal flow as an exact geostrophic solution . . . . .	125
4.3.8	A point vortex and the free space Green's function . . . . .	125
4.3.9	Green's function solution . . . . .	126
<b>4.4</b>	<b>Connection to equivalent barotropic flow . . . . .</b>	<b>127</b>
4.4.1	Vorticity equation for the depth averaged flow . . . . .	127
4.4.2	Comments . . . . .	128
<b>4.5</b>	<b>The lid pressure . . . . .</b>	<b>128</b>
4.5.1	Poisson equation for pressure . . . . .	129
4.5.2	Gradient wind balance . . . . .	130
4.5.3	Pressure source from self-advection . . . . .	130
4.5.4	Pressure source from circular rigid-body flow . . . . .	131
4.5.5	Pressure source from irrotational (pure strain) flow . . . . .	133
4.5.6	Pressure source from Coriolis acceleration . . . . .	134
4.5.7	Pressure source from friction . . . . .	134
4.5.8	Comments and further study . . . . .	135
<b>4.6</b>	<b>Kinematics of vorticity gradients . . . . .</b>	<b>135</b>
4.6.1	Material evolution of the vorticity gradient . . . . .	135
4.6.2	A two-time scale assumption . . . . .	136
4.6.3	Evolution of the squared vorticity gradient . . . . .	137
4.6.4	Comments and further study . . . . .	137
<b>4.7</b>	<b>Constraints from material invariance of absolute vorticity . . . . .</b>	<b>137</b>
4.7.1	Relative vorticity from curvature and planetary beta . . . . .	138
4.7.2	Relative vorticity from curvature and normal shears . . . . .	139
4.7.3	Curvature, shear, and planetary contributions . . . . .	140
4.7.4	Beta drift . . . . .	140
4.7.5	Understanding and prediction . . . . .	143
<b>4.8</b>	<b>Steady flow and the <math>\beta</math>-plume . . . . .</b>	<b>143</b>
4.8.1	The Rossby potential vorticity equation . . . . .	143
4.8.2	Steady flow balances . . . . .	143
4.8.3	Planetary geostrophic flow and the effective beta . . . . .	144
4.8.4	The beta plume Green's function . . . . .	145
4.8.5	Further study . . . . .	146
<b>4.9</b>	<b>Barotropic circulation on a rotating sphere . . . . .</b>	<b>147</b>
4.9.1	Zonal and meridional velocity components . . . . .	147
4.9.2	Vector-invariant velocity equation . . . . .	149
4.9.3	Vorticity equation . . . . .	149
4.9.4	Area conservation . . . . .	150
4.9.5	Circulation around a zonal loop . . . . .	150
4.9.6	Stokes' theorem and the zonal averaged zonal flow . . . . .	150
4.9.7	Planetary circulation . . . . .	151
4.9.8	Connecting circulation and axial angular momentum . . . . .	151
4.9.9	Connecting absolute vorticity and axial angular momentum . . . . .	154
4.9.10	Further study . . . . .	155

## 4.1 Loose threads

- Look at [Belmadani et al. \(2013\)](#) for further insights into the beta plume in Section 4.8.4.
- In the Rossby wave chapter, check the convergence of momentum as per the discussion in Section 4.9.8.
- Notation for strain and rotation tensors.
- New section on PV cascades just after Section 4.5. Follow discussion on pages 11 and 12 of McWilliams' turbulence notes.
- Make use of proper tensor notation and covariant derivatives. It would provide a very nice contribution for those aiming to work with this model on the sphere or elsewhere.

## 4.2 Basic equations and their properties

The horizontally non-divergent barotropic model arises from the following assumptions.

- The flow occurs in a single homogeneous fluid layer of constant density.
- The horizontal velocity is non-divergent:  $\nabla \cdot \mathbf{u} = 0$ , which means that the thickness of the layer is time independent.

The second property provides rather strict constraints on the flow relative to flow in the horizontally divergent shallow water layer. In this section we exhibit the governing equations and derive some of their properties.

### 4.2.1 Equations for velocity

The velocity equation for the non-divergent barotropic model follows in a manner akin to the shallow water model from Chapter 1. Namely, the fluid is a homogeneous (uniform and constant density) layer so that the horizontal velocity on a tangent plane satisfies

$$\frac{D\mathbf{u}}{Dt} + f\hat{\mathbf{z}} \times \mathbf{u} = -\nabla\varphi \quad \text{and} \quad \nabla \cdot \mathbf{u} = 0, \quad (4.1)$$

where the pressure is normalized according to

$$\varphi = p/\rho \quad (4.2)$$

with  $\rho$  the constant layer density, and where material evolution occurs with the two-dimensional non-divergent flow

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla. \quad (4.3)$$

All fields are depth independent and there is no vertical motion ( $w = 0$ ), so that the flow occurs in rigid fluid columns. Making use of the vector identity from VOLUME 1

$$(\mathbf{u} \cdot \nabla)\mathbf{u} = \zeta\hat{\mathbf{z}} \times \mathbf{u} + \nabla(\mathbf{u}^2/2), \quad (4.4)$$

brings the velocity equation (4.1) into its vector-invariant form

$$\partial_t \mathbf{u} + (f + \zeta) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla(\varphi + \mathbf{u} \cdot \mathbf{u}/2), \quad (4.5)$$

which is a useful starting point in Section 4.3 for deriving the equation for the vorticity

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \partial_x v - \partial_y u. \quad (4.6)$$

The horizontal non-divergent flow can be described by a streamfunction

$$\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi \implies u = -\partial_y \psi \quad \text{and} \quad v = \partial_x \psi. \quad (4.7)$$

Making use of the identities

$$\nabla \psi = -\hat{\mathbf{z}} \times \mathbf{u} \implies \mathbf{u} \cdot \nabla \psi = 0 \quad (4.8)$$

brings the vector-invariant velocity equation (4.5) into the form

$$\partial_t \mathbf{u} - (f + \zeta) \nabla \psi = -\nabla(\varphi + \mathbf{u} \cdot \mathbf{u}/2). \quad (4.9)$$

### 4.2.2 Kinetic energy of the flow

Taking the scalar product of the velocity with the velocity equation (4.9) yields the kinetic energy equation

$$\partial_t \mathcal{K} = -\nabla \cdot [\mathbf{u} (\mathcal{K} + \varphi)] \implies \frac{D\mathcal{K}}{Dt} = -\nabla \cdot (\mathbf{u} \varphi), \quad (4.10)$$

where we introduced the kinetic energy per mass

$$\mathcal{K} = \mathbf{u} \cdot \mathbf{u}/2 = (u^2 + v^2)/2 = \nabla \psi \cdot \nabla \psi / 2. \quad (4.11)$$

The gravitational potential energy is constant since the fluid density is a uniform constant (so the buoyancy vanishes) and the vertical velocity vanishes. Hence, mechanical energy transformation only involves kinetic energy. Furthermore, the domain integrated kinetic energy is a constant (for the inviscid case), as shown in Exercise 4.5.

### 4.2.3 Kinematics of rigid fluid columns

Recall the thickness equation (1.20) for a shallow water layer

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}, \quad (4.12)$$

where  $h = \eta - \eta_b$  is the column thickness,  $z = \eta(x, y, t)$  is the upper layer interface (the free surface), and  $z = \eta_b(x, y)$  is the lower interface (the bottom topography) (see Figure 1.1). With zero divergence in the horizontal velocity, the thickness of a fluid column is constant

$$(\partial_t + \mathbf{u} \cdot \nabla) h = \partial_t(\eta - \eta_b) + \mathbf{u} \cdot \nabla(\eta - \eta_b) = 0. \quad (4.13)$$

**Choosing  $\eta = 0$  to satisfy the thickness equation**

We choose to satisfy the thickness equation (4.13) by setting the upper surface to be static and flat

$$\eta = 0, \quad (4.14)$$

so that the column thickness is constant at each point in space

$$\partial_t h = \partial_t(\eta - \eta_b) = 0. \quad (4.15)$$

By setting  $\eta = 0$  we constrain the horizontal flow to follow the bottom topography so that

$$\mathbf{u} \cdot \nabla \eta_b = (\nabla \psi \times \nabla \eta_b) \cdot \hat{\mathbf{z}} = 0. \quad (4.16)$$

The case of a flat bottom offers no constraint, since  $\nabla \eta_b = 0$ . However, for the case with nontrivial bottom topography, the constraint (4.16) is satisfied only if the flow streamfunction,  $\psi$ , is a function of the bottom topography,

$$\psi = \psi(\eta_b), \quad (4.17)$$

so that the streamfunction is a constant along isolines of constant topography. Taking the limit as the sides become vertical, this boundary condition means that  $\psi$  is a constant along vertical sidewalls so that

$$\mathbf{u} \cdot \hat{\mathbf{n}} = (\hat{\mathbf{z}} \times \nabla \psi) \cdot \hat{\mathbf{n}} = \hat{\mathbf{t}} \cdot \nabla \psi = 0, \quad (4.18)$$

where  $\hat{\mathbf{n}}$  is the horizontal outward unit vector at the boundary, and  $\hat{\mathbf{t}} = \hat{\mathbf{n}} \times \hat{\mathbf{z}}$  is the unit tangent vector along the boundary.

**Streamfunction for the thickness-weighted velocity**

With rigid columns, the thickness equation (4.12) reduces to the non-divergence condition,

$$\nabla \cdot (h \mathbf{u}) = 0. \quad (4.19)$$

We can thus introduce a streamfunction,  $\Psi$ , for the thickness weighted horizontal velocity

$$h \mathbf{u} = \hat{\mathbf{z}} \times \nabla \Psi = \hat{\mathbf{z}} \times h \nabla \psi. \quad (4.20)$$

$\Psi$  has physical dimensions of  $L^3 T^{-1}$ , whereas the velocity streamfunction,  $\psi$ , has physical dimensions of  $L^2 T^{-1}$ .

**Summary**

We formulate the two-dimensional non-divergent flow within a homogeneous fluid layer, for which a fluid column does not expand or contract, vertical motion vanishes (as explored in Section 4.2.4), and the flow is constrained to move along lines of constant topography so that the streamfunction depends only on the bottom topography.

#### 4.2.4 Vertical velocity

##### Vanishing vertical velocity

With zero horizontal divergence and with a flat free surface ( $\eta = 0$ ), there is identically zero vertical motion within the layer

$$w = 0. \quad (4.21)$$

Another manner to deduce this property is by noting that the surface kinematic boundary condition for a static and flat free surface leads to

$$w(\eta) = \partial_t \eta + \mathbf{u} \cdot \nabla \eta = 0 + 0, \quad (4.22)$$

so that the vertical velocity at  $z = \eta = 0$  vanishes. With  $w(0) = 0$ , and with  $\nabla \cdot \mathbf{u} = 0$  for the horizontal velocity, then  $w = 0$  throughout the layer. Correspondingly, this constraint means that the no-normal flow bottom kinematic condition renders a horizontal velocity that is aligned with the topography,  $\mathbf{u} \cdot \nabla \eta_b = 0$ , which is a property we already encountered in Section 4.2.3. Since the vertical velocity is zero, the gravitational potential energy is a uniform constant, so that the mechanical energy budget involves only the kinetic energy.

##### But can there be a non-vanishing vertical velocity?

In Section 4.2.3 we chose to satisfy the thickness equation by setting  $\partial_t \eta = 0$  and  $\nabla \eta = 0$ , in which case the vertical velocity vanishes so long as the horizontal flow follows the bottom topography. A static and flat upper boundary is a sufficient condition, yet it is not necessary for satisfying the thickness equation. Is there a viable alternative?

Another approach to satisfying the thickness equation (4.13) is to consider a static yet non-flat free surface ( $\eta \neq 0$ ) with horizontal flow constrained so that

$$\mathbf{u} \cdot \nabla(\eta - \eta_b) = 0. \quad (4.23)$$

The surface and bottom kinematic boundary conditions then result in a vertical velocity at the surface and bottom

$$w(\eta) = -\mathbf{u} \cdot \nabla \eta \quad \text{and} \quad w(\eta_b) = -\mathbf{u} \cdot \nabla \eta_b. \quad (4.24)$$

These two boundary velocities must be equal to ensure  $\partial_z w = 0$ . If the bottom topography is flat,  $\nabla \eta_b = 0$ , then  $w(\eta_b) = w(\eta) = 0$ , in which case  $\nabla \eta = 0$ , so that the upper surface is flat. If the bottom is not flat, then the kinematic constraints lead to

$$\nabla \eta = \nabla \eta_b, \quad (4.25)$$

so that the upper surface slope equals to the slope of the bottom topography. In the ocean, it is generally the case that the upper ocean surface undulates far less than the bottom topography. Hence,  $\nabla \eta = \nabla \eta_b$  is not generally a realistic behavior, though one may study this case particularly with weakly sloping bottom topography.

For the remainder of this chapter we follow the conventional approach whereby  $w = 0$  everywhere. Hence, the upper boundary is static and flat, and we set  $\eta = 0$ . In the presence of a non-flat bottom, then the horizontal flow follows lines of constant bottom topography so that  $\mathbf{u} \cdot \nabla \eta_b = 0$ .

### 4.2.5 Velocity self-advection and the kinetic stress tensor

The velocity self-advection appearing in the velocity equation (4.1) can be written as the divergence of a  $2 \times 2$  symmetric tensor

$$-(\mathbf{u} \cdot \nabla) \mathbf{u} = \nabla \cdot \mathbb{E} \iff -u_m \partial_m u_n = -\partial_m(u_m u_n) = \partial_m \mathbb{E}_{mn}, \quad (4.26)$$

where

$$\mathbb{E}_{mn} = -u_m u_n \longleftrightarrow \mathbb{E} = -\mathbf{u} \otimes \mathbf{u} = \begin{bmatrix} -u^2 & -uv \\ -uv & -v^2 \end{bmatrix}, \quad (4.27)$$

thus bringing the momentum equation (4.1) to the Eulerian form

$$\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla \varphi + \nabla \cdot \mathbb{E}. \quad (4.28)$$

We refer to

$$\rho \mathbb{E} = -\rho \mathbf{u} \otimes \mathbf{u} \quad (4.29)$$

as the **kinetic stress tensor**, with its three-dimensional form introduced in VOLUME 2.

In anticipation of our study of vorticity in Section 4.3, we find it useful to decompose the kinetic stress tensor into its horizontally isotropic and horizontally anisotropic components<sup>2</sup>

$$\mathbb{E} = \begin{bmatrix} -u^2 & -uv \\ -uv & -v^2 \end{bmatrix} = -\mathcal{K} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -(u^2 - v^2)/2 & -uv \\ -uv & (u^2 - v^2)/2 \end{bmatrix} \equiv -\mathcal{K} \mathbb{I} + \mathbb{F}, \quad (4.30)$$

where we introduced the kinetic energy per mass according to equation (4.11), as well as the trace-free anisotropic portion of the kinetic stress tensor

$$\mathbb{F} = \mathcal{K} \mathbb{I} + \mathbb{E} = \begin{bmatrix} -(u^2 - v^2)/2 & -uv \\ -uv & (u^2 - v^2)/2 \end{bmatrix} \iff \mathbb{F}_{mn} = \mathcal{K} \delta_{mn} - u_m u_n. \quad (4.31)$$

Making use of the decomposition (4.30) brings the velocity equation (4.32) to the form

$$\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = \nabla \cdot [\mathbb{F} - \mathbb{I}(\mathcal{K} + \varphi)]. \quad (4.32)$$

Comparing to the vector-invariant velocity equation (4.5) allows us to infer the identity

$$\nabla \cdot \mathbb{F} = -\zeta (\hat{\mathbf{z}} \times \mathbf{u}). \quad (4.33)$$

We provide an alternative derivation of this identity in Exercise (4.1).

### 4.2.6 Further reading

As discussed by [Hoskins et al. \(1983\)](#), [Waterman and Hoskins \(2013\)](#) and [Waterman and Lilly \(2015\)](#), the decomposition (4.30) is useful for developing a geometric interpretation of eddying flow features.

## 4.3 Vorticity

For the two-dimensional non-divergent flow with a vanishing vertical velocity, the vertical component of the relative vorticity is the only nonzero vorticity component, and it is given by

<sup>2</sup>Recall our discussion of isotropy in VOLUME 1.

the Laplacian of the streamfunction

$$\zeta = \hat{z} \cdot (\nabla \times \mathbf{u}) = \partial_x v - \partial_y u = \nabla^2 \psi = \nabla \cdot (h^{-1} \nabla \Psi). \quad (4.34)$$

We here derive basic features of the vorticity for this fluid on the  $\beta$ -plane whereby the Coriolis parameter is

$$f = f_0 + \beta (y - y_0). \quad (4.35)$$

### 4.3.1 Vorticity equation

To form the dynamical equation for the vorticity, take the zonal derivative of the meridional momentum equation (see equation (4.1)), the meridional derivative of the zonal momentum equation, and then subtract

$$\frac{\partial}{\partial t} \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) + \frac{\partial}{\partial x} [\nabla \cdot (\mathbf{u} v)] - \frac{\partial}{\partial y} [\nabla \cdot (\mathbf{u} u)] + f \nabla \cdot \mathbf{u} + \beta v = 0. \quad (4.36)$$

The pressure gradient dropped out since there is zero **baroclinicity** for a barotropic flow.<sup>3</sup> We now make use of the identity

$$\frac{\partial}{\partial x} [\nabla \cdot (\mathbf{u} v)] - \frac{\partial}{\partial y} [\nabla \cdot (\mathbf{u} u)] = \frac{\partial \mathbf{u}}{\partial x} \cdot \nabla v - \frac{\partial \mathbf{u}}{\partial y} \cdot \nabla u + \mathbf{u} \cdot \nabla \zeta = \mathbf{u} \cdot \nabla \zeta, \quad (4.37)$$

where we used the non-divergence condition,  $\partial_x u + \partial_y v = 0$ , thus rendering

$$\frac{\partial \mathbf{u}}{\partial x} \cdot \nabla v - \frac{\partial \mathbf{u}}{\partial y} \cdot \nabla u = \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} \quad (4.38a)$$

$$= -\frac{\partial v}{\partial y} \frac{\partial v}{\partial x} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} - \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} \quad (4.38b)$$

$$= 0. \quad (4.38c)$$

We are thus led to the vorticity equation

$$\frac{\partial \zeta}{\partial t} + \nabla \cdot (\mathbf{u} \zeta) = -\beta v \quad \text{and} \quad \frac{D\zeta}{Dt} = -\beta v. \quad (4.39)$$

Evidently, the material evolution of relative vorticity in a horizontally non-divergent barotropic fluid is only affected by meridional advection of planetary vorticity. Since  $\beta > 0$  over the globe, northward flow ( $v > 0$ ) produces a negative source (clockwise tendency) for relative vorticity following a fluid parcel. This source term is the **beta effect** ( $\beta$ -effect) that we study in Section 6.6.

### 4.3.2 Constraints from material invariance of absolute vorticity

Since  $f$  is time independent, we can write the vorticity equation (4.39) in the form

$$(\partial_t + \mathbf{u} \cdot \nabla) \zeta_a = 0 \iff \frac{D\zeta_a}{Dt} = 0, \quad (4.40)$$

where

$$\zeta_a = \zeta + f \quad (4.41)$$

---

<sup>3</sup>We discuss baroclinicity in Section 6.4.

is the vertical component of the absolute vorticity. Equation (4.40) is a relatively simple version of the vorticity equation encountered in this book. It is simple because there are no sources on the right hand side, with the variety of vorticity sources studied in Chapter 6. In particular, there is no stretching or tilting of fluid columns in the non-divergent barotropic fluid.<sup>4</sup>

To maintain a materially constant absolute vorticity requires the relative vorticity to change oppositely to that of the planetary vorticity. For example, in the northern hemisphere the relative vorticity must decrease in value ( $\zeta \downarrow$ ) for fluid particles moving northward. This change in the relative vorticity is needed to counteract the increasing planetary vorticity ( $f \uparrow$ ) when moving northward. This result accords with the  $-\beta v$  source found in the relative vorticity equation (4.39). Furthermore, it is a reflection of the beta effect studied in Section 6.6.2. We further pursue these invariance properties in Section 4.7.

### 4.3.3 Rossby potential vorticity

According to the kinematic boundary condition (4.16), horizontal flow is aligned with isobaths (lines of constant bathymetry or topography). Flow moving along constant isobaths generally crosses latitude lines, and in so doing the relative vorticity must change precisely to maintain  $f + \zeta$  materially constant. As shown here, we can combine the absolute vorticity with the bottom topography to render a materially conserved object, the Rossby potential vorticity, which further helps to understand constraints on the flow (see Section 4.8 for an example).

As noted in Section 7.2.6 (see equation (7.24)), the absolute vorticity is the form of Ertel's potential vorticity for the horizontally non-divergent barotropic model. Additionally, since fluid columns are rigid so that  $Dh/Dt = 0$ , the Rossby potential vorticity is materially conserved (in the absence of non-conservative processes)

$$\frac{DQ}{Dt} = 0 \quad \text{with } Q = (\zeta + f)/h. \quad (4.42)$$

As discussed in Section 5.3, this form of potential vorticity conservation also holds for the inviscid shallow water equations, yet where  $h$  is no longer rigid so that the flow is horizontally non-divergent and can thus cross isobaths.

### 4.3.4 Jacobian forms of vorticity advection

In some contexts it is convenient to write the advection operator acting on relative vorticity as

$$\mathbf{u} \cdot \nabla \zeta = u \partial_x \zeta + v \partial_y \zeta \quad (4.43a)$$

$$= -\partial_y \psi \partial_x \zeta + \partial_x \psi \partial_y \zeta \quad (4.43b)$$

$$= \hat{\mathbf{z}} \cdot (\nabla \psi \times \nabla \zeta) \quad (4.43c)$$

$$\equiv J(\psi, \zeta), \quad (4.43d)$$

where  $J$  is the Jacobian operator

$$J(A, B) = \hat{\mathbf{z}} \cdot (\nabla A \times \nabla B) = \frac{\partial A}{\partial x} \frac{\partial B}{\partial y} - \frac{\partial A}{\partial y} \frac{\partial B}{\partial x}. \quad (4.44)$$

---

<sup>4</sup>In Section 6.5, we see how stretching and tilting of fluid columns provides a sources of vorticity for more general fluid flow.

We can also make use of the Jacobian for the advection of absolute vorticity<sup>5</sup>

$$\mathbf{u} \cdot \nabla \zeta_a = \hat{\mathbf{z}} \cdot (\nabla \psi \times \nabla \zeta_a) \equiv J(\psi, \zeta + f). \quad (4.45)$$

### 4.3.5 Taylor-Bretherton identity

An equivalent means to write the vorticity equation is to start from the velocity equation in the form (4.32). Taking the curl and projecting onto the vertical direction then leads to

$$\partial_t \zeta = -\beta v + \hat{\mathbf{z}} \cdot [\nabla \times \nabla \cdot \mathbb{F}]. \quad (4.46)$$

The nonlinear forcing from the anisotropic portion of the kinetic stress can be written

$$\hat{\mathbf{z}} \cdot [\nabla \times \nabla \cdot \mathbb{F}] = \hat{\mathbf{z}}_n \epsilon_{npq} \partial_p (\partial_m \mathbb{F}_{mq}) \quad (4.47a)$$

$$= \partial_m \partial_p (\epsilon_{npq} \hat{\mathbf{z}}_n \mathbb{F}_{mq}) \quad (4.47b)$$

$$= -\partial_m \partial_p (\epsilon_{pnq} \hat{\mathbf{z}}_n \mathbb{F}_{qm}) \quad (4.47c)$$

$$= -\partial_m \partial_p (\hat{\mathbf{z}} \times \mathbb{F})_{pm} \quad (4.47d)$$

$$= -\nabla \cdot [\nabla \cdot (\hat{\mathbf{z}} \times \mathbb{F})], \quad (4.47e)$$

where we used symmetry of the anisotropic kinetic tensor,  $\mathbb{F}_{mq} = \mathbb{F}_{qm}$ .

#### Taylor-Bretherton identity for relative vorticity

Comparing to the vorticity equation in the form (4.39) reveals the identity

$$\nabla \cdot (\mathbf{u} \zeta) = \nabla \cdot [\nabla \cdot (\hat{\mathbf{z}} \times \mathbb{F})] \implies \mathbf{u} \zeta = \nabla \cdot (\hat{\mathbf{z}} \times \mathbb{F}). \quad (4.48)$$

This equation says that the advective vorticity flux equals to the divergence of the counter-clockwise rotated anisotropic kinetic stress tensor. Equation (4.48) is a special form of the **Taylor-Bretherton identity** that provides a connection between the vorticity flux and the momentum flux. We encounter the shallow water form of this identity in VOLUME 4 when studying the decomposition of eddy and mean flows.

#### Verifying the Taylor-Bretherton identity

The divergence expression on the left hand side of equation (4.48) can be satisfied by  $\hat{\mathbf{z}} \times \nabla \Upsilon + \nabla \cdot (\hat{\mathbf{z}} \times \mathbb{F})$ , with  $\Upsilon$  an arbitrary gauge function. However,  $\Upsilon = 0$  is zero for the anisotropic kinetic tensor (4.31), as seen by

$$[\nabla \cdot (\hat{\mathbf{z}} \times \mathbb{F})]_1 = \partial_m (\epsilon_{mnp} \hat{\mathbf{z}}_n \mathbb{F}_{p1}) \quad (4.49a)$$

$$= \epsilon_{m3p} \partial_m \mathbb{F}_{p1} \quad (4.49b)$$

$$= -\epsilon_{3mp} \partial_m \mathbb{F}_{p1} \quad (4.49c)$$

$$= -\epsilon_{312} \partial_1 \mathbb{F}_{21} - \epsilon_{321} \partial_2 \mathbb{F}_{11} \quad (4.49d)$$

$$= -\partial_x \mathbb{F}_{21} + \partial_y \mathbb{F}_{11} \quad (4.49e)$$

$$= \partial_x (u v) + \partial_y (-u^2 + v^2)/2 \quad (4.49f)$$

$$= v \partial_x u + u \partial_x v - u \partial_y u + v \partial_y v \quad (4.49g)$$

$$= u \zeta, \quad (4.49h)$$

---

<sup>5</sup>The Jacobian operator corresponds to the Poisson bracket used in Hamiltonian mechanics.

and likewise

$$[\nabla \cdot (\hat{z} \times \mathbb{F})]_2 = \partial_y \mathbb{F}_{12} - \partial_x \mathbb{F}_{22} = v \zeta. \quad (4.50)$$

### Taylor-Bretherton identity for potential vorticity

Building from the development for relative vorticity, we can readily connect the potential vorticity flux,  $\mathbf{u} q = \mathbf{u} \zeta_a$ , to the anisotropic kinetic stress. We do so by considering the two equivalent forms for the potential vorticity equation

$$\partial_t q = -\nabla \cdot (\mathbf{u} q) \quad \text{and} \quad \partial_t q = -\nabla \cdot [\mathbf{u} f + \nabla \cdot (\hat{z} \times \mathbb{F})]. \quad (4.51)$$

Hence, the Taylor-Bretherton identity for potential vorticity in the two dimensional non-divergent flow is given by

$$\mathbf{u} q = \mathbf{u} f + \nabla \cdot (\hat{z} \times \mathbb{F}). \quad (4.52)$$

### 4.3.6 Poisson equation for the streamfunction

Given boundary conditions, the barotropic vorticity equation (4.40) allows us to determine the evolution of vorticity. We can in turn invert Poisson's equation,

$$\nabla^2 \psi = \zeta, \quad (4.53)$$

to determine the streamfunction and then the velocity field,  $\mathbf{u} = \hat{z} \times \nabla \psi$ . This inversion requires information about the boundary conditions for the streamfunction, as discussed in Section 4.2.3. We present a Green's function approach in Section 4.3.9 for the streamfunction. By this method, time integration of the absolute vorticity equation is sufficient to fully specify time evolution of the horizontal velocity. Notably, we do not need to explicitly determine pressure to determine the flow.

### 4.3.7 Zonal flow as an exact geostrophic solution

An arbitrary zonal velocity with a meridional shear,  $\mathbf{u} = U(y) \hat{x}$ , is an exact solution of the inviscid non-divergent barotropic model. We see this property by plugging into the velocity equation (4.1) and noting that  $D\mathbf{u}/Dt = 0$ . Hence, this flow is an exact geostrophic solution whose pressure field is itself also just a function of latitude and whose meridional gradient is determined by

$$f \hat{z} \times \mathbf{u} = -\nabla \varphi \implies \partial_y \varphi = -f(y) U(y). \quad (4.54)$$

Furthermore, each term in the vorticity equation (4.39) identically vanishes, so that the vorticity

$$\zeta = -\partial_y U \quad (4.55)$$

remains constant in time at each point in space. These properties make  $\mathbf{u} = U(y) \hat{x}$  a common choice for a background flow in the study of Rossby waves (VOLUME 4).

### 4.3.8 A point vortex and the free space Green's function

Consider an axially symmetric vortex centered at the origin,  $\mathbf{x} = 0$ , and in an infinite free space so there are no boundaries. Making use of polar coordinates (VOLUME 1), we know that the

relative vorticity for this vortex is related to the velocity via

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = r^{-1} \partial_r(r u^\vartheta), \quad (4.56)$$

with  $r$  the distance from the origin and  $\vartheta$  the polar angle. The identity (4.56) holds since all of the flow fields have axial symmetry, meaning that all fields are a function only of the radial distance from the origin. Now further assume the vorticity is given by

$$\zeta^{\text{point}} = -\alpha \delta(\mathbf{x}), \quad (4.57)$$

where  $\alpha$  is a constant, and we write the Dirac delta in polar coordinates (see VOLUME 1)

$$\delta(\mathbf{x}) = r^{-1} \delta(r) \delta(\vartheta). \quad (4.58)$$

In this manner we assume the vortex has zero extent yet infinite strength, which are properties of the Dirac delta. We can connect the constant,  $\alpha$ , to the circulation by considering an arbitrary circuit that encloses the origin

$$\mathcal{C} = \oint_{\partial\mathcal{S}} \mathbf{u} \cdot d\mathbf{r} = \int_{\mathcal{S}} \zeta^{\text{point}} d\mathcal{S} = -\alpha, \quad (4.59)$$

so that

$$\zeta^{\text{point}} = \mathcal{C} \delta(\mathbf{x}) = (\mathcal{C}/r) \delta(r) \delta(\vartheta). \quad (4.60)$$

Referring to the Green's function discussion in VOLUME 1, we see that the streamfunction for the point vortex (again, in the absence of boundaries) is the free-space Green's function for the Laplacian operator. This Green's function in the two-dimensional space of the barotropic model is given by the point vortex streamfunction

$$\nabla^2 \psi^{\text{point}} = \mathcal{C} \delta(\mathbf{x}) \iff \psi^{\text{point}} = (\mathcal{C}/2\pi) \ln(r/r_0). \quad (4.61)$$

We introduced the arbitrary constant,  $r_0$ , to ensure the argument to the natural log is dimensionless. But  $r_0$  merely adds a constant to the streamfunction, and so its precise value is physically irrelevant and so it is ignored in the following. The flow associated with the point vortex is given by

$$\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi^{\text{point}} = \frac{\mathcal{C} \hat{\vartheta}}{2\pi r}. \quad (4.62)$$

Evidently, the flow is a purely angular swirling motion around the point vortex that falls off as the inverse distance from the vortex.

Though highly idealized, the flow field arising from the point vortex is aligned with that in the far field for realistic localized vorticities, thus making the point vortex a physically relevant theoretical idealization. The discussion here briefly touched upon the theoretical richness of point vorticities in fluid mechanics, with chapter 3 of [McWilliams \(2006\)](#) offering a more thorough study, also in the context of two dimensional non-divergent flow.

### 4.3.9 Green's function solution

Poisson's equation for the streamfunction admits a Green's function solution (VOLUME 1), which makes use of the superposition principle to have the point vortex solution (4.61) build up the streamfunction arising from an arbitrary distribution of vorticity. For brevity, we only

consider the case without boundaries, with the more general case including boundaries detailed in VOLUME 1.

We seek a streamfunction that satisfies the Poisson equation

$$\nabla^2 \psi = \zeta, \quad (4.63)$$

with  $\zeta$  a given vorticity that is localized in space. Making use of the free-space Green's function from VOLUME 1 leads to

$$\psi(\mathbf{x}) = \frac{1}{2\pi} \int \zeta(\mathbf{y}) \ln(|\mathbf{x} - \mathbf{y}|) dS_y \quad (4.64)$$

where  $dS_y$  is the area element for the horizontal integral over  $\mathbf{y}$ . This streamfunction is built by convolving the vorticity source with the free space Green's function. The corresponding velocity is given by

$$\mathbf{u}(\mathbf{x}) = \hat{\mathbf{z}} \times \nabla \psi(\mathbf{x}) = \frac{1}{2\pi} \int \zeta(\mathbf{y}) \frac{\hat{\mathbf{z}} \times (\mathbf{x} - \mathbf{y})}{|\mathbf{x} - \mathbf{y}|^2} dS_y. \quad (4.65)$$

As a check, we see that if the vorticity source is chosen to be a point vortex at the origin,  $\zeta(\mathbf{y}) = C \delta(\mathbf{y})$ , then  $\mathbf{u}$  correctly reduces to the point vortex velocity (4.62). The expression (4.65) is sometimes referred to as the two-dimensional **Biot-Savart law**, which arises from the analog in electromagnetism.<sup>6</sup>

## 4.4 Connection to equivalent barotropic flow<sup>7</sup>

In many cases, flows respecting the assumptions of quasi-geostrophy (Section 9.5) possess a vertical profile that can be separated from the horizontal. In this case we write the horizontal velocity as

$$\mathbf{u}(x, y, z, t) = \Gamma(z) \mathbf{u}^{eb}(x, y, t), \quad (4.66)$$

where  $\Gamma > 0$  is a single-signed non-dimensional structure function that has a unit vertical average,  $\langle \Gamma \rangle = 1$ , when computed over the fluid layer thickness, and where  $\mathbf{u}^{eb}(x, y, t)$  carries the horizontal spatial dependence of the flow. With  $\Gamma > 0$ , the horizontal flow remains in the same direction throughout the fluid column; i.e., eastward flow at the top of the column remains eastward at the bottom. This orientation of the flow is generally referred to as **equivalent barotropic**.

### 4.4.1 Vorticity equation for the depth averaged flow

To connect the very particular form (4.66) for the flow with the non-divergent barotropic model, we anticipate a discussion in Section 11.2.2 in which the quasi-geostrophic vorticity equation is shown to be

$$\frac{\partial \zeta_g}{\partial t} + \mathbf{u}_g \cdot \nabla \zeta_g = -\beta \zeta_g + f_0 \frac{\partial w}{\partial z}, \quad (4.67)$$

---

<sup>6</sup>See, for example, Section 5.2 of [Jackson \(1975\)](#) or Section 5.2 of [Griffiths \(1981\)](#) for an electromagnetism discussion of the Biot-Savart law. Note, however, that the more common expression of the Biot-Savart law is given in three-dimensions, where the magnetic field has a  $(\mathbf{x} - \mathbf{y})/|\mathbf{x} - \mathbf{y}|^3$  dependence rather than the  $(\mathbf{x} - \mathbf{y})/|\mathbf{x} - \mathbf{y}|^2$  dependence found in two-dimensions as in equation (4.65).

<sup>7</sup>Section 4.4 requires a basic understanding of quasi-geostrophic theory as discussed in Section 9.5. So this discussion here can be readily skipped on first reading, and returned to after studying quasi-geostrophy in Chapter 11.

where  $\mathbf{u}_g$  is the horizontally non-divergent geostrophic velocity, and  $\zeta_g = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}_g)$  is the geostrophic relative vorticity. We see that the quasi-geostrophic vorticity is affected by both the beta-effect and by vertical stretching of fluid columns (the  $\partial w / \partial z$  term), whereas vertical stretching is absent from the vorticity equation (4.39) for the horizontally non-divergent barotropic fluid. There are occasions in which it is sensible to assume the vertical velocity for a quasi-geostrophic flow vanishes at the top and bottom of the fluid domain, such as when considering flow in the absence of topography. In this case, performing the decomposition (4.66) for the horizontal flow and then taking a vertical average of the quasi-geostrophic vorticity equation (4.67) leads to

$$\frac{\partial \zeta_g^{eb}}{\partial t} + \langle \Gamma^2 \rangle \mathbf{u}_g^{eb} \cdot \nabla \zeta_g^{eb} = -\beta v_g^{eb}. \quad (4.68)$$

This equation motivates us to define

$$\mathbf{u}^* = \langle \Gamma^2 \rangle \mathbf{u}_g^{eb} \quad \text{and} \quad \zeta^* = \langle \Gamma^2 \rangle \zeta_g^{eb}, \quad (4.69)$$

which are the original geostrophic fields,  $\mathbf{u}_g(x, y, z, t)$  and  $\zeta_g(x, y, z, t)$ , when evaluated at a depth where  $\Gamma(z^*) = \langle \Gamma^2 \rangle$ . The depth,  $z^*$ , is known as the [equivalent barotropic depth](#). Introduction of the starred fields then brings the vorticity equation (4.68) into the form of the non-divergent barotropic vorticity equation

$$\frac{\partial \zeta^*}{\partial t} + \mathbf{u}^* \cdot \nabla \zeta^* = -\beta v^*. \quad (4.70)$$

#### 4.4.2 Comments

[Charney et al. \(1950\)](#) justified their study of the non-divergent barotropic vorticity model by noting its connection to the commonly observed equivalent barotropic structure of the large-scale middle latitude atmosphere. The equivalent barotropic model has been a very useful analysis and prediction tool for meteorologists, and it formed the basis of many numerical weather prediction models into the 1980s. Section 7.1 of [Haltiner and Williams \(1980\)](#) offers further details on such numerical models.

One hypothesis for why quasi-geostrophic flows tend towards an equivalent barotropic profile relates to movement of energy in rotationally dominant turbulent flows, whereby energy cascades to the larger scales. As discussed in [Smith and Vallis \(2001\)](#) and [Venaille et al. \(2012\)](#), among others, this *inverse* energy cascade pumps mechanical energy into a vertically uniform or “barotropic” structure. In a realistic flow, including topography, stratification, and variable forcing, this cascade is never realized completely, thus rendering a flow that approaches an equivalent barotropic structure but never quite gets there fully.

The case of  $f = 0$  with flat bottom is referred to as [two dimensional turbulence](#) ([Kraichnan and Montgomery, 1980](#)). This model has a history of key theoretical results that presaged their analog in quasi-geostrophic turbulence. See chapters 11 and 12 of [Vallis \(2017\)](#) for further discussion of this model and its relevance to geostrophic turbulence of the ocean and atmosphere.

## 4.5 The lid pressure

Recall the velocity equation (4.1) for the two-dimensional non-divergent flow

$$[\partial_t + (\mathbf{u} \cdot \nabla)] \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla \varphi. \quad (4.71)$$

As noted in Section 4.3.6, we do not need to determine the lid pressure,  $\varphi$ , to determine the evolution of the flow. Instead, we can determine the flow by time stepping vorticity and then inverting the elliptic problem to get the streamfunction. Furthermore, the free surface is absolutely flat even in the presence of topography. Hence, there is no pressure generated by undulations of the free surface. So how is there flow in this model??!

We answer this question by studying the lid pressure. It is the lid pressure that maintains a flat upper surface in the two-dimensional non-divergent barotropic model (rigid lid approximation). The lid pressure is imposed by some external means (whose details are not important), thus ensuring that the flow remains horizontally non-divergent. In turn, it is the lid pressure that provides the force to drive flow. We here expose details to help understand the fundamental, yet somewhat hidden, role for the lid pressure.

To further motivate the analysis, we again emphasize the strong constraints placed on the horizontal flow, which must satisfy the following conditions at each space and time point

$$\nabla \cdot \mathbf{u} = 0 \quad \text{and} \quad \mathbf{u} \cdot \nabla \eta_b = 0 \quad \text{and} \quad w = 0. \quad (4.72)$$

Kinematics of the constrained flow induce a depth-independent pressure that enforces these constraints. Examining the resulting pressure field furthers our understanding of the forces acting in the moving fluid. This role for pressure as the enforcer of non-divergence is shared by the three-dimensional non-divergent flow found in a Boussinesq ocean. It is simpler to visualize fields in the two dimensional non-divergent barotropic model, thus facilitating understanding and insights that are also useful for the three dimensional case.

### 4.5.1 Poisson equation for pressure

We derive the pressure equation by using the two-dimensional non-divergence property of the horizontal flow and then developing a diagnostic relation. We can eliminate the time derivative from equation (4.1) by taking  $\partial/\partial x$  on the zonal equation and  $\partial/\partial y$  on the meridional equation, then adding. The result is a diagnostic relation for the Laplacian of the pressure<sup>8</sup>

$$-\nabla^2 \varphi = \partial_x [\nabla \cdot (\mathbf{u} u)] + \partial_y [\nabla \cdot (\mathbf{u} v)] - f \zeta + \beta u, \quad (4.73)$$

where we set

$$\nabla \cdot \partial_t \mathbf{u} = \partial_t \nabla \cdot \mathbf{u} = 0 \implies \partial_x (\partial_t u) = -\partial_y (\partial_t v). \quad (4.74)$$

Making use of the boundary conditions discussed in Section 4.2.3, the elliptic partial differential equation (4.73) can be inverted to find the pressure field (VOLUME 1). As for the three-dimensional Boussinesq ocean, or for a three-dimensional incompressible fluid, pressure is the force that instantaneously constrains the velocity to remain non-divergent.

Numerically inverting an elliptic operator in equation (4.73) is straightforward on simple domains, such as flat bottom rectangular regions or a smooth sphere. However, when the bottom is not flat, or when there are islands (i.e., the domain is not simply connected), then the elliptic inversion can be numerically difficult to perform. This algorithmic complexity is one reason numerical barotropic models are less commonly used for realistic numerical experimentation than the more general shallow water models. Even so, as pursued in the

---

<sup>8</sup>We maintain the minus sign on the left hand side of equation (4.73) so that a positive source on the right hand side leads to a positive  $\varphi$ . We can readily understand the sign by taking a Fourier transform, whereby the Laplacian operator picks up a minus sign when converted to Fourier space. We follow the same sign convention when studying the Green's function for the Poisson equation in VOLUME 1.

remainder of this section, we can make use of idealized configurations to garner insight into how pressure maintains the flow constraints.

### 4.5.2 Gradient wind balance

For two-dimensional non-divergent flow we can write the contribution to the pressure equation (4.73) from self-advection in terms of the streamfunction, or equivalently as the Jacobian of the velocity field. We do so through the following manipulations

$$\partial_x[\nabla \cdot (\mathbf{u} u)] + \partial_y[\nabla \cdot (\mathbf{u} v)] = \nabla \cdot (\partial_x \mathbf{u} u) + \nabla \cdot (\mathbf{u} \partial_x u) + \nabla \cdot (\partial_y \mathbf{u} v) + \nabla \cdot (\mathbf{u} \partial_y v) \quad (4.75a)$$

$$= \partial_x \mathbf{u} \cdot \nabla u + \partial_y \mathbf{u} \cdot \nabla v \quad (4.75b)$$

$$= (\partial_x u)^2 + (\partial_y v)^2 + 2 \partial_x v \partial_y u \quad (4.75c)$$

$$= 2 [(\partial_x u)^2 + \partial_x v \partial_y u] \quad (4.75d)$$

$$= 2 [(\partial_{xy}\psi)^2 - \partial_{xx}\psi \partial_{yy}\psi] \quad (4.75e)$$

$$= 2 \hat{\mathbf{z}} \cdot (\partial_x \nabla \psi \times \partial_y \nabla \psi) \quad (4.75f)$$

$$= -2 J(\partial_x \psi, \partial_y \psi), \quad (4.75g)$$

$$= 2 J(v, u), \quad (4.75h)$$

where we introduced the Jacobian operator from equation (4.39). The pressure equation (4.73) thus takes on the form

$$-\nabla^2 \varphi = 2 J(u, v) - \nabla \cdot (f \nabla \psi), \quad (4.76)$$

where we also wrote

$$-\nabla \cdot (f \nabla \psi) = -f \nabla^2 \psi - \beta \partial_y \psi = -f \zeta + \beta u. \quad (4.77)$$

We refer to equation (4.76) as a **gradient wind** balance (see VOLUME 2 for more on gradient wind balance). Here, the Jacobian term accounts for the **centrifugal** acceleration of the curved fluid motion, and the  $\nabla \cdot (f \nabla \psi)$  term accounts for the Coriolis acceleration, both for the  $f$ -plane and  $\beta$ -plane. Equation (4.76) thus offers a more accurate diagnostic relation for the pressure field than provided by assuming a geostrophically balanced flow.

### 4.5.3 Pressure source from self-advection

We introduce yet another way to examine the self-advection source appearing in the pressure equation (4.73), and do so by writing it as

$$\partial_x[\nabla \cdot (\mathbf{u} u)] + \partial_y[\nabla \cdot (\mathbf{u} v)] = S_{mn} S_{mn} - R_{mn} R_{mn}, \quad (4.78)$$

where the strain rate tensor,  $\mathbf{S}$ , and rotation tensor,  $\mathbf{R}$ , are second order tensors that have components given by<sup>9</sup>

$$S_{mn} = (\partial_n v_m + \partial_m v_n)/2 = S_{nm} \quad \text{strain rate tensor} \quad (4.79)$$

$$R_{mn} = (\partial_n v_m - \partial_m v_n)/2 = -R_{nm} \quad \text{rotation tensor.} \quad (4.80)$$

---

<sup>9</sup>We study the strain rate tensor and rotation tensor as part of the kinematics in VOLUME 1.

For two-dimensional flow the rotation tensor is related to the relative vorticity via

$$R_{mn} = -\epsilon_{mn} \zeta / 2 = -\epsilon_{mn} \nabla^2 \psi / 2, \quad (4.81)$$

where  $\epsilon_{mn}$  is the anti-symmetric **permutation symbol** (VOLUME 1)

$$\epsilon_{mn} = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{21} & \epsilon_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad (4.82)$$

which leads to the double contraction of the rotation tensor

$$R_{mn} R_{mn} = \zeta^2 / 2. \quad (4.83)$$

The identity (4.78) indicates that the squared strain rate (sometimes referred to as the **splat**) provides a positive source to the Poisson equation (4.73) whereas squared vorticity provides a negative source. Bringing these results together leads to the pressure equation

$$-\nabla^2 \varphi = S_{mn} S_{mn} - \zeta^2 / 2 - \nabla \cdot (f \nabla \psi) = 2 [(\partial_{xy} \psi)^2 - \partial_{xx} \psi \partial_{yy} \psi] - \nabla \cdot (f \nabla \psi), \quad (4.84)$$

where the second equality made use of equation (4.75e) to write the pressure source fully in terms of the streamfunction.

In the following sections, we present two examples to help understand how pressure gradients arise in a two-dimensional barotropic flow, with the pressure gradients maintaining a non-divergent two-dimensional velocity field. We start by assuming a non-rotating reference frame, so that it is sufficient to examine how pressure responds to the nonlinear source term  $S_{mn} S_{mn} - \zeta^2 / 2$ . Thereafter we include sources from the rotating reference frame.

#### 4.5.4 Pressure source from circular rigid-body flow

Consider a velocity field in a flat domain that is initialized in circular rigid-body motion (Figure 4.1)

$$\mathbf{u} = \boldsymbol{\Omega} \times \mathbf{x} = \Omega (-y \hat{\mathbf{x}} + x \hat{\mathbf{y}}), \quad (4.85)$$

where  $\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}$  is a constant angular velocity. This flow has zero strain (as do all rigid-body flows) yet constant vorticity

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \zeta = 2 \Omega \implies -R_{mn} R_{mn} = -2 \Omega^2. \quad (4.86)$$

The velocity time tendency from the self-advection acceleration is

$$-(\mathbf{u} \cdot \nabla) \mathbf{u} = \Omega^2 \mathbf{x}, \quad (4.87)$$

which is the outward directed centrifugal acceleration associated with the circular rigid-body motion.<sup>10</sup> The velocity equation (4.71) thus takes the form (recall we are considering  $f = 0$  for now)

$$\partial_t \mathbf{u} = \Omega^2 \mathbf{x} - \nabla \varphi. \quad (4.88)$$

---

<sup>10</sup>We considered the more general case in VOLUME 2 when decomposing the material acceleration for two-dimensional flow into **natural coordinates**.

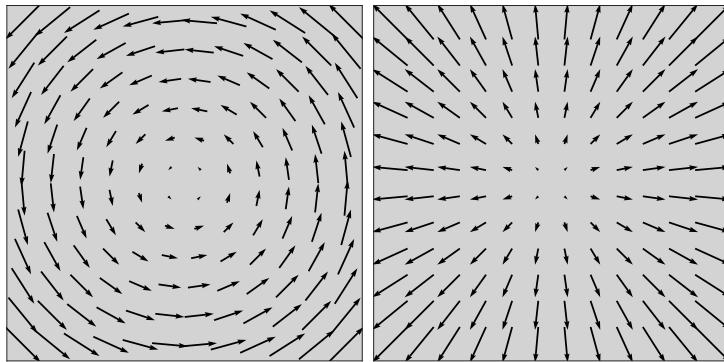


FIGURE 4.1: Left panel: example flow field from rigid-body rotation,  $\mathbf{u} = \boldsymbol{\Omega} \times \mathbf{x} = \Omega(-y, x)$ , which has vorticity  $\nabla \times \mathbf{u} = 2\boldsymbol{\Omega}$  and zero strain. Right panel: corresponding acceleration from self-advection,  $-(\mathbf{u} \cdot \nabla) \mathbf{u} = \Omega^2 \mathbf{x}$ , which is a centrifugal acceleration. The centrifugal acceleration from self-advection is exactly compensated by the pressure gradient force:  $-(\mathbf{u} \cdot \nabla) \mathbf{u} - \nabla \varphi = 0$ , thus allowing for the rigid-body motion to be an exact steady solution to two-dimensional non-divergent flow. The units are arbitrary.

### Pressure gradient acceleration exactly balances centrifugal acceleration

If acting alone, the centrifugal acceleration would create a velocity field that diverges from the origin. However, the velocity is constrained to remain non-divergent at each instance, so the centrifugal acceleration cannot be the full story. Indeed, to maintain a non-divergent flow we find a pressure gradient that exactly balances the centrifugal acceleration. This pressure gradient is established instantaneously as per the solution to the elliptic pressure equation. We thus infer that there is a low pressure at the origin so that the pressure gradient force points inward, thus balancing the centrifugal acceleration.

We can extrapolate from this example to conclude that vorticity provides a source of low pressure in a non-divergent flow. Furthermore, since the centrifugal acceleration from the velocity self-advection is exactly balanced by the pressure gradient acceleration, the rigid-body flow is an exact steady solution for the non-divergent and non-rotating (zero Coriolis) barotropic system. This result holds even in the presence of viscosity, since the rigid-body flow has no strain and hence it does not support viscous stresses.<sup>11</sup>

### Rotationally symmetric pressure field

We can support the above general statements by deriving an explicit expression for the pressure field, and we do so by solving the pressure Poisson equation (4.73). Since we have assumed zero planetary rotation ( $f = 0$ ), equation (4.73) reduces to

$$-\nabla^2 \varphi = -2\Omega^2, \quad (4.89)$$

which also follows from taking the divergence of equation (4.88) and noting that  $\nabla \cdot \mathbf{x} = 2$ . Furthermore, the solid-body flow is assumed to be rotationally symmetric so that all fields have only radial dependence. In this case, pressure satisfies the ordinary differential equation

$$r^{-1} \partial_r(r \partial_r \varphi) = 2\Omega^2, \quad (4.90)$$

<sup>11</sup>We discuss the connection between strain rates and stresses in VOLUME 2.

where  $r$  is the radial distance from the origin and we used the polar coordinate version of the Laplacian from VOLUME 1. The solution to the pressure equation (4.90) is given by

$$\varphi = (\Omega r)^2/2 \quad \text{and} \quad -\nabla\varphi = -\Omega^2 \mathbf{x} = (\mathbf{u} \cdot \nabla) \mathbf{u}, \quad (4.91)$$

where we set  $\varphi(r=0)=0$ . Evidently, the pressure grows parabolically moving radially away from the center.

### The pressure field in relation to rigid-body rotating shallow water

To help understand the physics of the pressure field (4.91), recall the analysis from VOLUME 2, in which we studied the motion of a rigid-body rotating homogeneous fluid layer in a cylindrical tank. In contrast to the barotropic system, the horizontal velocity in a homogeneous fluid layer, such as in a shallow water fluid, is divergent so that the layer thickness is not constrained to remain flat. Hence, the centrifugal acceleration causes the velocity to diverge from the center so that the layer thickness increases radially outward, with the layer bounded by the tank wall. At steady state, the homogeneous fluid layer has a parabolic free surface with a minimum at the center. The parabolic free surface creates a pressure field that precisely corresponds to the pressure field (4.91) in the non-divergent barotropic system. Note that the dynamical adjustment of a homogeneous shallow water fluid layer contain linear fluctuations in the form of gravity waves such as discussed in Chapter 2. In contrast, the adjustment required to reach a steady state occurs instantaneously in the non-divergent barotropic fluid.

#### 4.5.5 Pressure source from irrotational (pure strain) flow

Consider the following pure strain flow (Figure 4.2)

$$\mathbf{u} = \Omega (-x \hat{\mathbf{x}} + y \hat{\mathbf{y}}), \quad (4.92)$$

whose vorticity vanishes and whose self-advection acceleration is given by

$$-(\mathbf{u} \cdot \nabla) \mathbf{u} = -\Omega^2 \mathbf{x} \implies S_{mn} S_{mn} = 2\Omega^2, \quad (4.93)$$

thus leading to a velocity equation (again,  $f=0$  is assumed)

$$\partial_t \mathbf{u} = -\Omega^2 \mathbf{x} - \nabla \varphi. \quad (4.94)$$

The acceleration (4.93) is exactly the opposite of that produced by the rigid rotation studied in the previous example (equation (4.87)). Hence, to maintain a horizontally non-divergent barotropic flow, a pressure field is established with a high pressure at the center that exactly counteracts the converging self-advection acceleration present in the pure strain flow

$$\varphi = -(\Omega r)^2/2 \quad \text{and} \quad -\nabla\varphi = \Omega^2 \mathbf{x} = (\mathbf{u} \cdot \nabla) \mathbf{u}. \quad (4.95)$$

This example illustrates how strain provides a source for high pressure in a non-divergent flow. Furthermore, we see that this flow, in the absence of viscosity, is an exact steady solution for non-divergent barotropic flow in a non-rotating reference frame. However, in contrast to the rigid-body flow in Section 4.5.4, the purely strained flow (4.92) supports viscous friction, so that this flow does not remain steady in the presence of viscosity.

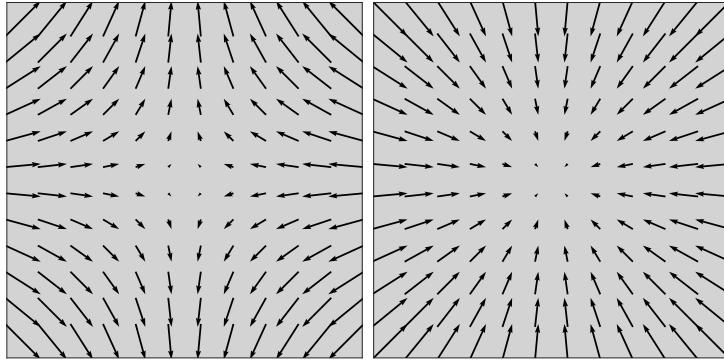


FIGURE 4.2: Left panel: example purely strained flow field with zero vorticity,  $\mathbf{u} = \Omega(-x\hat{\mathbf{x}} + y\hat{\mathbf{y}})$ . Right panel: corresponding converging acceleration from self-advection,  $-(\mathbf{u} \cdot \nabla) \mathbf{u} = -\Omega^2 \mathbf{x}$ . The units are arbitrary.

#### 4.5.6 Pressure source from Coriolis acceleration

In addition to the self-advection source, pressure is affected by a source from the Coriolis acceleration

$$-\nabla^2 \varphi_{\text{geostrophy}} \equiv \nabla \cdot (f \hat{\mathbf{z}} \times \mathbf{u}) = -\nabla \cdot (f \nabla \psi) = \beta u - \zeta f. \quad (4.96)$$

As such, we can write the pressure gradient as<sup>12</sup>

$$-\nabla \varphi_{\text{geostrophy}} = f \hat{\mathbf{z}} \times \mathbf{u}. \quad (4.97)$$

Consider cyclonic flow around an arbitrary point. What is the pressure field induced by this flow? As in our discussion in Section 4.5.4 of the rigid-body rotating flow, a cyclonic flow has an associated centrifugal acceleration that points outward. To counteract the centrifugal acceleration, and thus to maintain a non-divergent flow, requires an inward pointing pressure force; i.e., a low pressure center. Hence, cyclonic circulation induces, through the Coriolis acceleration, a negative pressure source whereas anti-cyclonic circulation induces a positive pressure source. This is indeed an interesting perspective on geostrophic balance!

#### 4.5.7 Pressure source from friction

Consider flow with friction, in which case the velocity equation (4.1) takes the form

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla \varphi + \mathbf{F}, \quad (4.98)$$

with  $\mathbf{F}$  a frictional acceleration. In this case we have yet another source for pressure given by

$$-\nabla^2 \varphi_{\text{friction}} \equiv -\nabla \cdot \mathbf{F}. \quad (4.99)$$

As discussed in our study of stress in VOLUME 2, viscous friction is generally associated with a nonzero **strain rate tensor**. We see that the frictional acceleration induces a high pressure source in regions where frictional accelerations converge,  $-\nabla \cdot \mathbf{F} > 0$ , with this source acting to maintain non-divergent flow in the presence of converging frictional acceleration.

<sup>12</sup>Formally, we can add an arbitrary gauge function to the right hand side of equation (4.97), with this term of the form  $\hat{\mathbf{z}} \times \nabla \chi$ . However, since we derived the pressure Poisson equation from the velocity equation, then we know there is no gauge function arising in equation (4.97).

### 4.5.8 Comments and further study

- The pressure equation (4.73) is elliptic, and elliptic equations need boundary conditions. In the presence of topography the boundary conditions are modified relative to the flat bottom case. Hence, pressure knows about topography through its boundary conditions. The resulting pressure force keeps the flow non-divergent and the flow aligned with topography as per the kinematic conditions in Section 4.2.3.
- The discussion of pressure induced by self-advection in Sections 4.5.4 and 4.5.5 is based on a similar presentation in Appendix B of *Jeevanjee and Romps* (2015).
- *Bryan* (1969) provided the first working numerical algorithm to simulate the ocean general circulation. Bryan's method made use of the rigid lid approximation of Section 4.5 so that the depth integrated velocity is non-divergent. However, the vorticity in Bryan's ocean model is affected by more than just the beta-effect. The reason is that the depth integrated velocity equation includes contributions from baroclinic processes, and such processes affect the barotropic vorticity in a baroclinic fluid. We detail such effects in Sections 10.6 and 10.7.

The rigid lid method was used for large-scale ocean circulation modeling until the late 1990s. Free surface methods, allowing divergence in the depth integrated flow, have largely displaced the rigid lid as a practical method for time stepping ocean models (e.g., see chapter 12 of *Griffies* (2004)).

## 4.6 Kinematics of vorticity gradients

In this section we examine some kinematic properties of vorticity gradients. These properties are of direct relevance to understanding facets of turbulence in two-dimensional non-divergent flows.

### 4.6.1 Material evolution of the vorticity gradient

To derive the evolution equation for  $\nabla \zeta_a$ , take the gradient of the vorticity equation

$$\frac{D\zeta_a}{Dt} = 0 \implies \nabla \frac{D\zeta_a}{Dt} = 0. \quad (4.100)$$

Performing the product rule and rearranging leads to

$$\frac{D(\partial_1 \zeta_a)}{Dt} = -(G^1{}_1 \partial_1 \zeta_a + G^2{}_1 \partial_2 \zeta_a) \quad (4.101a)$$

$$\frac{D(\partial_2 \zeta_a)}{Dt} = -(G^1{}_2 \partial_1 \zeta_a + G^2{}_2 \partial_2 \zeta_a), \quad (4.101b)$$

where we introduced the Cartesian representation of the velocity gradient tensor<sup>13</sup>

$$G^i{}_j = \partial_j u^i \iff \mathbf{G} = \begin{bmatrix} \partial_1 v^1 & \partial_2 v^1 \\ \partial_1 v^2 & \partial_2 v^2 \end{bmatrix} = \begin{bmatrix} -\partial_{xy}\psi & -\partial_{yy}\psi \\ \partial_{xx}\psi & \partial_{xy}\psi \end{bmatrix}, \quad (4.102)$$

---

<sup>13</sup>In VOLUME 1, we study the velocity gradient tensor, and its decomposition into the strain rate tensor and the rotation tensor.

with

$$(v^1, v^2) = (u, v). \quad (4.103)$$

We can write the evolution in the following succinct manners

$$\frac{D(\partial_j \zeta_a)}{Dt} = -G^i{}_j \partial_i \zeta_a \iff \frac{D(\nabla \zeta_a)}{Dt} = -\mathbf{G}^T \cdot \nabla \zeta_a. \quad (4.104)$$

### 4.6.2 A two-time scale assumption

The gradient of the relative vorticity,  $\nabla \zeta = \nabla(\nabla^2 \psi)$ , has spatial structure determined by three derivatives acting on the streamfunction, whereas the velocity gradient tensor is built from only two derivatives. We thus expect  $\nabla \zeta$  to have finer spatial structure, and corresponding shorter time scale, than the velocity gradient tensor. Hence, to garner an approximate sense for how the vorticity gradient evolves, assume the velocity gradient tensor is fixed in time while the gradient of the vorticity evolves. Evolution can be characterized by determining the eigenvalues of  $\mathbf{G}$ , with its characteristic polynomial given by

$$\lambda^2 - \lambda G^i{}_i + \det(\mathbf{G}) = 0 \implies \lambda^2 = -\hat{\mathbf{z}} \cdot (\nabla u \times \nabla v), \quad (4.105)$$

where the trace of the velocity gradient tensor vanishes given the non-divergent nature of the flow

$$G^i{}_i = \nabla \cdot \mathbf{u} = 0. \quad (4.106)$$

There are either two real eigenvalues or two imaginary eigenvalues, depending on the sign of the right hand side of equation (4.105). We garner physical insight into what determines this sign by decomposing the velocity gradient tensor into the strain rate tensor,  $\mathbf{S}$ , and the rotation tensor,  $\mathbf{R}$ ,

$$\mathbf{G} = (\mathbf{G} + \mathbf{G}^T)/2 + (\mathbf{G} - \mathbf{G}^T)/2 = \mathbf{S} + \mathbf{R}, \quad (4.107)$$

with the Cartesian coordinate representations given by

$$\mathbf{S} = \begin{bmatrix} \partial_1 v^1 & (\partial_2 v^1 + \partial_1 v^2)/2 \\ (\partial_2 v^1 + \partial_1 v^2)/2 & \partial_2 v^2 \end{bmatrix}, \quad (4.108)$$

and

$$\mathbf{R} = \begin{bmatrix} 0 & -(\partial_1 v^2 - \partial_2 v^1)/2 \\ (\partial_1 v^2 - \partial_2 v^1)/2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\zeta/2 \\ \zeta/2 & 0 \end{bmatrix}. \quad (4.109)$$

The double contraction of these tensors

$$\mathbf{S} : \mathbf{S} = S^i{}_j S^j{}_i = (\partial_1 v^1)^2 + (\partial_2 v^2)^2 + (\partial_2 v^1 + \partial_1 v^2)^2/2 \quad (4.110a)$$

$$\mathbf{R} : \mathbf{R} = R^i{}_j R^j{}_i = (\partial_1 v^2 - \partial_2 v^1)^2/2, \quad (4.110b)$$

reveals that the eigenvalues of the velocity gradient tensor are determined by a competition between the strain rate tensor and the rotation tensor

$$\lambda = \pm \sqrt{(\mathbf{S} : \mathbf{S} - \mathbf{R} : \mathbf{R})/2}. \quad (4.111)$$

For flow dominated by the strain rate, there is exponential stretching along one of the eigendirections and exponential compression in the other direction. A hyperbolic (saddle) flow pattern respects this behavior, with the streamfunction having a local behavior of the form  $\psi \propto xy$ ,

which corresponds to the pure strain fields discussed as part of the flow kinematics chapter in VOLUME 1. Flow dominated by rotation (i.e., vorticity) generates two imaginary eigenvalues that reflect an oscillatory/recurrent pattern without growth of the vorticity gradients. Notably, the domain integral of the squared eigenvalue vanishes when ignoring boundary terms, which can be seen by writing

$$\lambda^2 = -\hat{\mathbf{z}} \cdot (\nabla u \times \nabla v) = -\hat{\mathbf{z}} \cdot \nabla \times (u \nabla v) \implies \int \lambda^2 dA = 0. \quad (4.112)$$

### 4.6.3 Evolution of the squared vorticity gradient

Taking the scalar product of equation (4.104) with the gradient of the absolute vorticity leads to the evolution equation

$$\frac{1}{2} \frac{D|\nabla \zeta_a|^2}{Dt} = -\nabla \zeta_a \cdot \mathbf{G}^T \cdot \nabla \zeta_a = -\nabla \zeta_a \cdot \mathbf{S} \cdot \nabla \zeta_a, \quad (4.113)$$

where the second equality noted that

$$\nabla \zeta_a \cdot \mathbf{R} \cdot \nabla \zeta_a = \partial_i \zeta_a \delta^{ij} R^k_j \partial_k \zeta_a = 0, \quad (4.114)$$

given that the rotation tensor is anti-symmetric. Hence, evolution of the squared gradient of the absolute vorticity is determined by the strain rate tensor alone, with the rotation tensor only acting to rotate the gradient direction and not altering its magnitude.

### 4.6.4 Comments and further study

The characterization of flow as a competition between strain rate (stretching and compressing) versus rotation (vorticity) is useful in practice, even for flows more complicated than the two-dimensional non-divergent flows considered here. For these reasons, the studies from [Okubo \(1970\)](#) and [Weiss \(1991\)](#) have been emulated by many others, with the Okubo-Weiss parameter

$$W \equiv 4\lambda^2 = 2(\mathbf{S} : \mathbf{S} - \mathbf{R} : \mathbf{R}), \quad (4.115)$$

a common diagnostic used to characterize fluid flows.

## 4.7 Constraints from material invariance of absolute vorticity

Following from Section 4.3.2, we here examine constraints on the flow imposed by material invariance of absolute vorticity, which holds for two-dimensional non-divergent flow when there is no dissipation

$$\frac{D(\zeta + f)}{Dt} = 0. \quad (4.116)$$

These constraints offer insights into the flow behavior and allow us to predict responses to perturbations. Notably, these predictions arise even without direct information about the forces giving rise to the responses. Rather, we deduce the responses based on vorticity constraints alone.

We frame the discussion in terms of the decomposition (3.50) of relative vorticity into a

curvature (or orbital) term and normal shear term

$$\zeta = \frac{|\mathbf{u}|}{R} - \frac{\partial |\mathbf{u}|}{\partial n} \equiv \zeta_{\text{curv}} + \zeta_{\text{shear}}, \quad (4.117)$$

so that

$$\frac{D(\zeta_{\text{curv}} + \zeta_{\text{shear}} + f)}{Dt} = 0. \quad (4.118)$$

When the flow turns, the curvature term contributes with  $R$  the radius of curvature.<sup>14</sup> If the flow develops shears in the direction normal to the flow, then the shear term contributes. In the presence of  $\beta = \partial_y f$ , meridional motion through the planetary vorticity field requires a compensating response from relative vorticity. In general each of the three terms contribute to the relative vorticity, yet for pedagogical purposes we consider cases where one term is subdominant and so can be ignored.

#### 4.7.1 Relative vorticity from curvature and planetary beta

According to the decomposition (4.117), fluid flow that curves to the left (facing downstream) picks up a positive relative vorticity from flow curvature,  $R > 0 \Rightarrow \zeta > 0$ , as depicted in Figure 4.3. The oppositely curved flow has a negative radius of curvature so flow curving to the right picks up a negative relative vorticity,  $R < 0 \Rightarrow \zeta < 0$ . We focus here on the case where the normal shear induced relative vorticity can be neglected so that we are only concerned with curvature induced vorticity plus planetary vorticity (beta effect).

Consider a flow that is initially zonal with zero relative vorticity. If the flow turns meridionally then it experiences a change in relative vorticity both through the curvature term plus a change in planetary vorticity since  $f$  changes. To maintain constant absolute vorticity, a fluid column that moves meridionally requires the relative vorticity induced by the curved flow path to counteract the change in planetary vorticity. As we now discuss, the constraint of fixed absolute vorticity, in the absence of induced normal shears, means that eastward flow (westerly winds) cannot turn meridionally while maintaining fixed absolute vorticity, whereas westward flow (easterly winds) can turn (see Figure 4.3).

Consider westward flow in the northern hemisphere ( $f > 0$ ). If the flow turns to the north (to the right facing downstream) then this flow picks up a curvature-induced negative relative vorticity,  $\zeta < 0$ , and an increase in the planetary vorticity ( $f$  increases). Likewise, a westward flowing fluid column that turns equatorward (to the left) has a positive curvature-induced relative vorticity ( $\zeta > 0$ ) and a reduction in planetary vorticity ( $f$  decreases). Hence, westward flow in the northern hemisphere can turn either poleward (to the north) or equatorward (to the south) and still maintain constant absolute vorticity, so long as the curved motion induces the proper relative vorticity to counteract the changes to  $f$ . The same arguments also hold in the southern hemisphere, so that the general scenarios are depicted in Figure 4.3.

The situation is different for eastward flow. Consider again flow in the northern hemisphere. A poleward (to the left) turning fluid column is associated with a positive curvature-induced relative vorticity,  $\zeta > 0$ , as well as an increase in the planetary vorticity. Hence, this motion changes the absolute vorticity and as such it is not allowed if the absolute vorticity is constrained to remain constant. Likewise, an equatorward (to the right) turning eastward fluid column

<sup>14</sup>As discussed in our study of horizontal motion using natural coordinates (VOLUME 2), we take the convention whereby the normal direction is to the left of flow when facing downstream. Flow turning into the normal direction (to the left) has a positive radius of curvature,  $R > 0$ , and flow turning opposite to the normal direction (to the right) has  $R < 0$ .

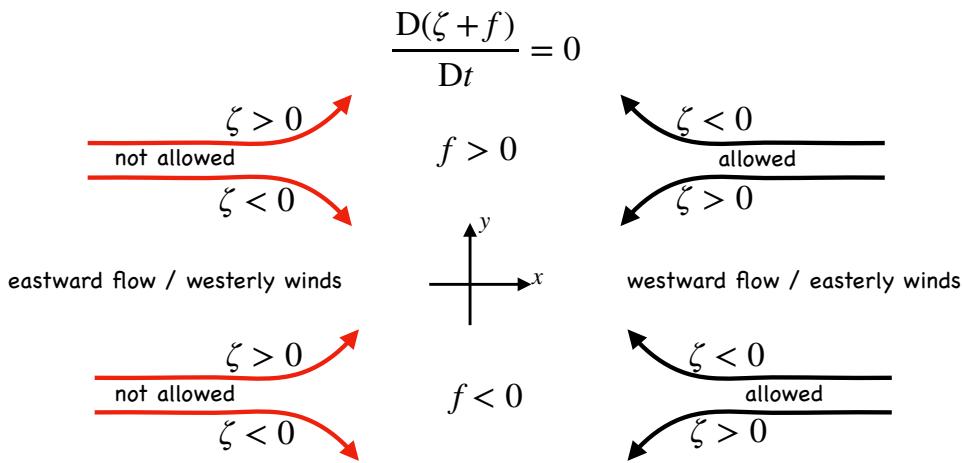


FIGURE 4.3: Illustrating the constraints on a two dimensional non-divergent flow imposed by material invariance of absolute vorticity:  $\zeta + f = \text{constant}$ . We assume flow is over a flat region and assume there is only curvature induced relative vorticity (no shear-induced relative vorticity;  $\partial|\mathbf{u}|/\partial n = 0$ ) as per the decomposition in equation (4.117). In each of the four cases depicted, the entering flow has zero relative vorticity, which means that absolute vorticity must remain constant at the initial Coriolis parameter,  $\zeta + f = f_{\text{initial}}$ . The red eastward flow (westerly winds) that turns meridionally picks up a curvature vorticity that supports the change in planetary vorticity, thus precluding material invariance of absolute vorticity. Hence, the meridional turning of eastward flow is not allowed so that flow must remain zonal for absolute vorticity to remain invariant. In contrast, the oppositely directed westward flow (easterly winds) can deviate either to the north or south and still retain a constant absolute vorticity. We illustrate flows for both the northern and southern hemispheres. This figure is adapted from Figure 4.13 of [Holton and Hakim \(2013\)](#).

induces a negative curvature-induced relative vorticity,  $\zeta < 0$ , and a decrease in planetary vorticity, again leading to a change in absolute vorticity. Hence, eastward flow (westerly winds) in either hemisphere must remain zonal to maintain a constant absolute vorticity.

As an application of these results, consider the situation depicted in Figure 4.4, whereby inviscid flow in the interior of an ocean domain moves westward into a frictional western boundary layer. The constraints imposed by absolute vorticity invariance allow for this flow to occur, whereas the opposite is disallowed whereby eastward inviscid flow cannot enter an eastern boundary. We return to this example in Section 5.7 when discussing western intensification of ocean gyres.

#### 4.7.2 Relative vorticity from curvature and normal shears

Now consider the case in which the meridional displacements are small so that the beta effect can be neglected. In this case there is an exchange between relative vorticity arising from curvature and relative vorticity from normal shears, thus leaving their sum materially invariant

$$\frac{D(\zeta_{\text{curv}} + \zeta_{\text{shear}})}{Dt} = 0. \quad (4.119)$$

We depict an example in Figure 4.5 whereby a vortex undergoes a left turn facing downstream. While on the curve, the relative vorticity of the vortex is in part due to the positive curvature vorticity,  $\zeta_{\text{curv}} > 0$ . If the relative vorticity is positive on the straight portion of the trajectory, then when on the curve the shear vorticity must lose some of its strength in order to compensate for the curvature vorticity. Conversely, if the relative vorticity is negative on the straight portion of the trajectory, then when on the curve the shear vorticity gains in strength to allow

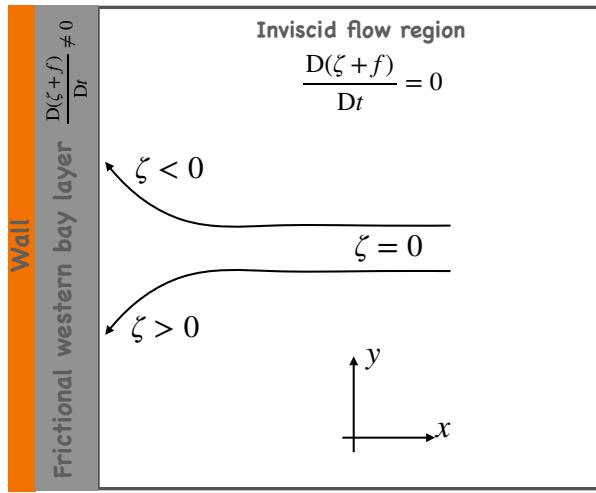


FIGURE 4.4: Illustrating the constraints on a homogeneous and constant thickness fluid layer imposed by material invariance of absolute vorticity:  $\zeta + f = \text{constant}$ . As per the results from Figure 4.3, inviscid flow with initially zero relative vorticity can enter a western boundary layer as depicted here, whereas it cannot enter an eastern boundary layer. This conclusion assumes that there is no shear-induced relative vorticity ( $\partial|\mathbf{u}|/\partial n = 0$ ) that can overcome changes in the vorticity induced by changes to  $f$  and by curvature-induced relative vorticity (see Section 3.8.2). This figure is adapted from Figure 19.12 of [Vallis \(2017\)](#).

for the positive curvature vorticity.

### 4.7.3 Curvature, shear, and planetary contributions

We now consider all three terms appearing in the vorticity equation (4.118). Let us consider again the eastward flow that turns to the north in the northern hemisphere. Such flow is not allowed if the only source for relative vorticity is curvature. However, if the eastward flow, as it turns, picks up a shear that induces a nonzero negative relative vorticity, then such flow can turn so long as the shear-induced negative relative vorticity balances the positive absolute vorticity from increases in  $f$  and the curvature-induced vorticity. Writing this condition for the shear-induced relative vorticity yields

$$\zeta_{\text{shear}} = -\Delta f - \zeta_{\text{curv}} = -(f_{\text{final}} - f_{\text{init}}) - \zeta_{\text{curv}} < 0, \quad (4.120)$$

where  $f_{\text{init}}$  and  $f_{\text{final}}$  are the initial and final Coriolis parameters. Conversely, if the flow deviates towards the equator then it can do so only if there is a positive shear-induced relative vorticity

$$\zeta_{\text{shear}} = -(f_{\text{final}} - f_{\text{init}}) - \zeta_{\text{curv}} > 0. \quad (4.121)$$

### 4.7.4 Beta drift

In VOLUME 2, we encounter an exercise in which we introduce the Rossby effect ([Rossby, 1948](#)), in which a circular cyclonic vortex experiences an area integrated Coriolis acceleration that is directed poleward, and with the integrated acceleration vanishing on the  $f$ -plane. Hence, this poleward drift arises from the beta effect. Following [Rossby \(1948\)](#), we did not consider the pressure field associated with the vortex, so it is unclear whether such a vortex would actually drift due northward. Indeed, subsequent studies showed that motion of an initially circular vortex sets up a secondary circulation that renders a poleward+westward beta-drift ( $\beta$ -drift);

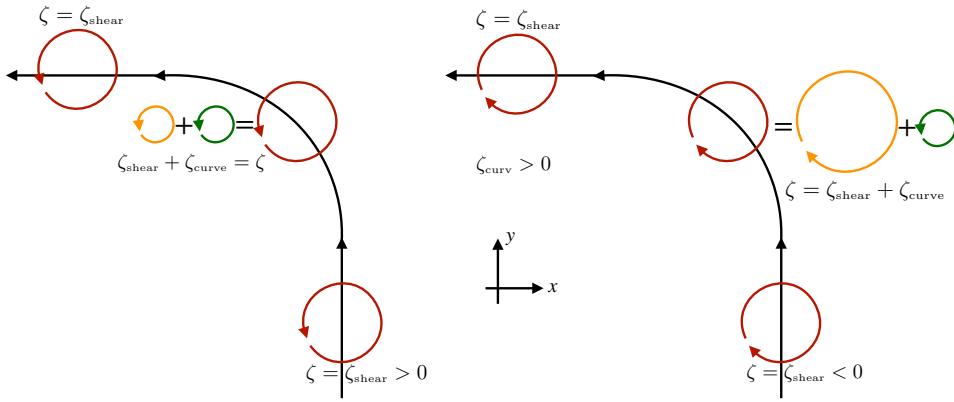


FIGURE 4.5: Material invariance of relative vorticity,  $D\zeta/Dt = D(\zeta_{curv} + \zeta_{shear})/Dt = 0$ , means that as a vortex moves around a curve its shear vorticity is modified to keep the total relative vorticity invariant. On the straight portion of these trajectories, the relative vorticity is due only to shear vorticity,  $\zeta = \zeta_{shear}$  since  $\zeta_{curv} = 0$ . However, when the vortex enters the curve, maintaining a constant relative vorticity requires an exchange of shear vorticity with the curvature vorticity. In this example we illustrate a steady flow that turns to the left so that the vortex picks up a positive curvature vorticity when on the curve,  $\zeta_{curv} > 0$ . Left panel: a vortex that enters the left turn with a positive relative vorticity must give some of its shear vorticity to the curvature vorticity in order to maintain  $\zeta$  constant along the trajectory. Right panel: a vortex that enters the curve with a negative relative vorticity sees its shear vorticity increase in magnitude to compensate for the positive curvature vorticity.

i.e., northwestward in the northern hemisphere and southwestward in the southern hemisphere. As for the other motions considered in this section, we describe the mechanism for beta drift by invoking conservation of absolute vorticity respected by an inviscid non-divergent barotropic flow. This discussion reflects similar ideas encountered when studying [Rossby waves](#) in [VOLUME 4](#).

Consider a circularly symmetric northern hemisphere cyclonic monopole as shown in Figure 4.6. The monopole flow has positive circulation and thus positive relative vorticity. On an  $f$ -plane this circulation is stationary, whereas parcels moving around the monopole on the  $\beta$ -plane pick up anomalous relative vorticity according to the beta effect:  $D\zeta/Dt = -\beta v$ . On the west side of the monopole, fluid elements are moving southward and thus pick up a positive anomalous relative vorticity ( $-\beta v > 0$ ), whereas on the east side the northward flow picks up a negative relative vorticity anomaly. We note that the material evolution of relative vorticity is also reflected in the local time changes, since for an initially circular monopole, the only contribution to the local evolution is given by the beta effect. We see this property by writing the vorticity equation using polar coordinates

$$\partial_t \zeta = -\beta v - \mathbf{u} \cdot \nabla \zeta = -\beta v - (\dot{r} \partial_r + \dot{\vartheta} \partial_\vartheta) \zeta, \quad (4.122)$$

where  $r$  is the radial coordinate and  $\vartheta$  is the angular coordinate measured counter-clockwise from the  $x$ -axis. By assumption, the flow is initially moving only in the angular direction, so that  $\dot{r} = 0$ . Additionally, the monopole is symmetric in the angular direction, so that  $\partial_\vartheta \zeta = 0$ . As a result,  $\partial_t \zeta = -\beta v$ .

From the above analysis, we see that throughout the western side of the monopole, beta induces a positive anomalous vorticity, whereas beta induces a negative vorticity anomaly throughout the eastern side. When combined with the vorticity from the monopole, we see that the beta induced anomalous vorticity leads to a westward drift of the location for the maximum vorticity; i.e., the monopole maximum drifts to the west. Yet that is not the full story.

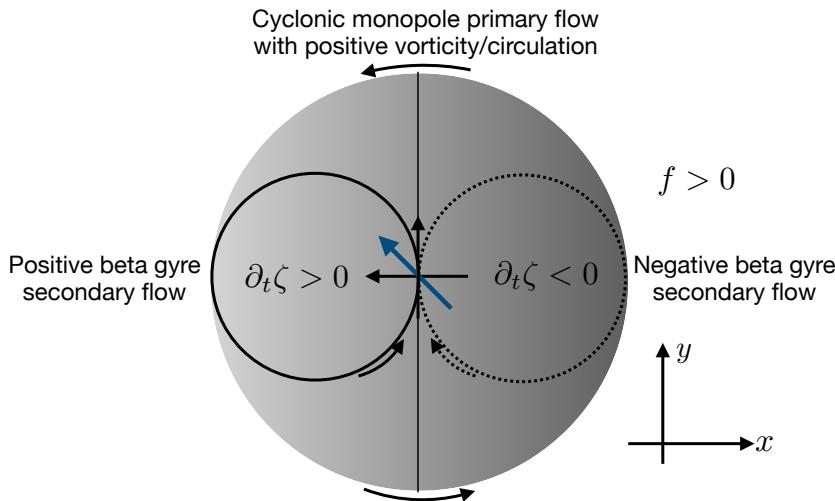


FIGURE 4.6: Schematic cyclonic and circularly symmetric northern hemisphere ( $f > 0$ ) monopole flow in a two-dimensional non-divergent barotropic flow. The monopole has positive circulation and thus positive relative vorticity. On the  $\beta$ -plane, parcels moving around the monopole pick up anomalous relative vorticity according to the beta effect:  $D\zeta/Dt = -\beta v$ . On the western side, parcels move southward and thus pick up a positive relative vorticity anomaly ( $-\beta v > 0$ ), whereas on the eastern side the northward flow picks up a negative relative vorticity anomaly. The beta effect thus induces a westward drift of the monopole maximum, towards where the relative vorticity is increasing. Additionally, the positive anomaly on the western side of the monopole induces a **secondary circulation** known as a **beta gyre**, with this gyre rotating counter-clockwise, whereas there is an oppositely oriented beta gyre on the eastern side. The secondary circulation from the counter-rotating beta gyres induces a northward drift to the monopole. The combined effect of the westward beta induced drift and the northward drift from the beta gyres leads to a net northwestward **beta-drift** ( **$\beta$ -drift**) for the monopole.

In addition to the westward drift of the monopole maximum, the positive vorticity anomaly on the western side induces a positive gyre-like circulation referred to as a **beta gyre**. Likewise, the negative vorticity on the eastern side induces a negative beta gyre. The beta gyre circulation is an example of a **secondary circulation**, since it arises in response to the anomalies induced by motion through the primary monopole circulation. Furthermore, the counter-rotating beta gyres induce a northward drift of the monopole. The combined westward drift induced by beta acting on the primary monopole circulation, plus the northward drift from the secondary beta gyre circulations, leads to an overall northwestward drift of the monopole. More generally, a cyclonic monopole experiences a poleward and westward beta drift, whereas for anti-cyclonic monopoles the beta drift is equatorward and westward.

The extent to which beta drift is respected by more realistic monopoles depends on many factors, such as the strength and radius of the monopole, strength of the background planetary vorticity gradient, and stability of the monopole. The literature on these topics makes use of numerical models to probe the nonlinearities associated with these relatively strong, and sometimes unstable, flow regimes. Some of the papers are motivated by motion of coherent ocean eddies (e.g., [McWilliams and Flierl, 1979](#); [Carnevale et al., 1991](#)), and others are motivated by motion of atmospheric tropical cyclones (e.g., [Holland, 1983](#); [Smith, 1993](#)). For tropical cyclones born off the coast of Africa in the tropical Atlantic, beta drift gives the cyclones a general tendency to move northwestward toward North America (absent environmental flows that can counteract the beta drift). More recently, [Gavriel and Kaspi \(2021\)](#) employed these concepts to help understand vortices found in the polar regions of the Jovian atmosphere.

### 4.7.5 Understanding and prediction

The examples in this section illustrate the power of vorticity constraints for the purpose of predicting flow responses. The power largely rests on our ability to determine flow responses without directly determining forces causing the response. Even so, without determining the forces acting in the fluid, our understanding of the dynamics remains incomplete even if our ability to predict is complete. So when one can determine the forces (it is not always as simple as the examples in Section 4.5), then doing so offers further physical insights into the nature of the flow.

## 4.8 Steady flow and the $\beta$ -plume

In this section we consider the steady solution to the horizontally non-divergent barotropic model in the planetary geostrophic regime studied in VOLUME 2 (see also Chapter 10). For this purpose, we return to the Rossby potential vorticity (4.42) and add a Rayleigh drag along with a frictional stress vector,  $\tau$ .

### 4.8.1 The Rossby potential vorticity equation

Introducing the Rayleigh drag and frictional stresses into the velocity equation (4.1) leads to

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla\varphi - \gamma \mathbf{u} + \mathbf{F}. \quad (4.123)$$

In this equation,  $\gamma \geq 0$  is the constant Rayleigh drag coefficient (with dimensions of inverse time), and

$$\mathbf{F} = \frac{\boldsymbol{\tau}^{\text{wind}} - \boldsymbol{\tau}^{\text{bot}}}{h \rho_0} \quad (4.124)$$

is the acceleration arising from difference between the surface wind stress,  $\boldsymbol{\tau}^{\text{wind}}$ , and bottom stress,  $\boldsymbol{\tau}^{\text{bot}}$  (see equation (1.135)). Carrying the Rayleigh drag and boundary stresses through the derivation of the Rossby potential vorticity equation (4.42) yields

$$h \frac{DQ}{Dt} = -\gamma \zeta + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) \quad \text{with } Q = (f + \zeta)/h. \quad (4.125)$$

Evidently, potential vorticity is no longer materially conserved in the presence of either Rayleigh drag or boundary stresses.

### 4.8.2 Steady flow balances

For steady flow, the potential vorticity equation (4.125) reads

$$h \mathbf{u} \cdot \nabla Q = -\gamma \zeta + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}), \quad (4.126)$$

which takes on the following form in terms of the transport streamfunction,  $h \mathbf{u} = \hat{\mathbf{z}} \times \nabla \Psi$  (equation (4.20)),

$$(\hat{\mathbf{z}} \times \nabla Q) \cdot \nabla \Psi = -\gamma \nabla \cdot (h^{-1} \nabla \Psi) + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}), \quad (4.127)$$

where the potential vorticity is

$$h Q = f + \zeta = f + \nabla \cdot (h^{-1} \nabla \Psi). \quad (4.128)$$

Use of the product rule on the Rayleigh drag term, and rearrangement, leads to

$$(h \hat{z} \times \nabla Q + \gamma \nabla \ln h) \cdot \nabla \Psi = \gamma \nabla^2 \Psi + h \hat{z} \cdot (\nabla \times \mathbf{F}). \quad (4.129)$$

To help intrepret this streamfunction equation, introduce

$$\mathbf{A} \equiv h \hat{z} \times \nabla Q + \gamma \nabla \ln h, \quad (4.130)$$

which is a horizontally divergent vector ( $\nabla \cdot \mathbf{A} \neq 0$ ) with physical dimensions  $L^{-1} T^{-1}$  (the same dimensions as  $\beta = \partial_y f$ ). In this case the streamfunction equation (4.129) takes on the form of a steady advective-diffusive-source equation<sup>15</sup>

$$\underbrace{\mathbf{A} \cdot \nabla \Psi}_{\text{advection}} = \underbrace{\gamma \nabla^2 \Psi}_{\text{diffusion}} + \underbrace{h \hat{z} \cdot (\nabla \times \mathbf{F})}_{\text{source}}. \quad (4.131)$$

The vector,  $\mathbf{A}$ , serves as an advection “velocity” that acts to align the streamfunction along integral paths defined by  $\mathbf{A}$ ; i.e., the streamfunction is “advected” by  $\mathbf{A}$ . For example, in the absence of non-conservative process and boundary stresses, isolines of constant streamfunction and Rossby potential vorticity are aligned,

$$(\hat{z} \times \nabla Q) \cdot \nabla \Psi = (\nabla Q \times \nabla \Psi) \cdot \hat{z} = 0 \quad \text{if } \gamma = 0 \text{ and } \mathbf{F} = 0. \quad (4.132)$$

In this case, the streamfunction functionally depends only on the potential vorticity,  $\Psi = \Psi(Q)$ , which is another way of stating that the steady unforced and inviscid flow is along lines of constant  $Q$ . The presence of boundary stresses,  $\hat{z} \cdot (\nabla \times \mathbf{F}) \neq 0$ , causes the steady flow to deviate from  $Q$  isolines, with the stresses providing a source to the streamfunction equation (4.131). Finally, the presence of Rayleigh drag ( $\gamma > 0$ ) acts to diffuse or spread the streamfunction isolines.

### 4.8.3 Planetary geostrophic flow and the effective beta

We develop more insights into the steady flow by linearizing the streamfunction equation (4.131), which is done by assuming the flow maintains **planetary geostrophy** (see VOLUME 2 or Chapter 10). In this case the potential vorticity is independent of the streamfunction and takes on the form

$$Q^{pg} = f/h. \quad (4.133)$$

Correspondingly,  $\mathbf{A}$  is now independent of the streamfunction and is given by

$$\mathbf{A} = -\beta \hat{x} - Q^{pg} \hat{z} \times \nabla h + \gamma \nabla \ln h \quad (4.134a)$$

$$= \hat{x} [-\beta + Q^{pg} \partial_y h + (\gamma/h) \partial_x h] + \hat{y} [-Q^{pg} \partial_x h + (\gamma/h) \partial_y h]. \quad (4.134b)$$

In the special case of a uniform layer thickness ( $\nabla h = 0$ ), we find  $\mathbf{A} = -\beta \hat{x}$ , so that the streamlines are advected to the west according to planetary beta. The more general advection is somewhat more complex. Even so, below we find interesting cases in which the advection remains zonal.

The advection velocity,  $\mathbf{A}$ , is purely zonal if

$$h Q^{pg} \partial_x h = \gamma \partial_y h \implies f \partial_x h = \gamma \partial_y h, \quad (4.135)$$

---

<sup>15</sup>We study the physics of advection and diffusion in VOLUME 4. Here, we only require a few basic features.

which then leads to

$$\mathbf{A} = -\beta^{\text{eff}} \hat{\mathbf{x}} \quad \text{with } \beta^{\text{eff}} = \beta + \partial_y h (f^2 + \gamma^2) / (h f). \quad (4.136)$$

Stated in terms of the topographic slopes, the advective streamfunction transport is zonal if the topography satisfies

$$\nabla \ln h = \frac{(\beta^{\text{eff}} - \beta)(\gamma \hat{\mathbf{x}} + f \hat{\mathbf{y}})}{f^2 + \gamma^2}. \quad (4.137)$$

The effective beta parameter,  $\beta^{\text{eff}}$ , is comprised of three contributions:

$$\beta^{\text{planetary}} = \beta = \partial_y f \quad (4.138a)$$

$$\beta^{\text{topog}} = f \partial_y \ln h = Q^{\text{pg}} \partial_y h \quad (4.138b)$$

$$\beta^{\text{Rayleigh}} = (\gamma^2/f) \partial_y \ln h = (\gamma^2/f^2) Q^{\text{pg}} \partial_y h. \quad (4.138c)$$

The first term,  $\beta^{\text{planetary}}$ , is the planetary beta that arises from meridional dependence of the planetary Coriolis parameter,  $\beta = \partial_y f$ . The second term,  $\beta^{\text{topog}}$ , arises from meridional dependence of the bottom topography in the presence of planetary rotation. The third term,  $\beta^{\text{Rayleigh}}$ , arises from meridional dependence of the bottom topography in the presence of Rayleigh drag and planetary rotation.

Note that  $\beta^{\text{eff}}$  is not sign-definite, and it passes through zero if

$$Q^{\text{pg}} \partial_y h^{\text{zero}} = -\frac{\beta f^2}{f^2 + \gamma^2} < 0. \quad (4.139)$$

In this case the planetary potential vorticity has an opposite sign from the meridional topographic slope, with  $Q^{\text{pg}} \partial_y h^{\text{zero}} = -\beta$  for case of vanishing Rayleigh drag.

#### 4.8.4 The beta plume Green's function

With a zonal advective transport, the streamfunction equation (4.131) reduces to the linear partial differential equation<sup>16</sup>

$$-(\beta^{\text{eff}} \partial_x + \gamma \nabla^2) \Psi = h \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}). \quad (4.140)$$

It is notable that the linear differential operator on the left hand side is not **self-adjoint**, in a manner akin to the diffusion operator studied in VOLUME 1. . We can make use of **Green's function** methods (VOLUME 1) to write an expression for the streamfunction. Ignoring boundaries (i.e., assume an infinite  $\beta$ -plane) allows us to make direct use of a solution from VOLUME 1

$$\Psi(\mathbf{x}) = \int h \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) \mathcal{G}^\ddagger(\mathbf{x}|\mathbf{x}_0) d\mathcal{S}_0, \quad (4.141)$$

where  $\mathcal{G}^\ddagger(\mathbf{x}|\mathbf{x}_0)$  is the adjoint free-space Green's function that satisfies

$$-(-\beta^{\text{eff}} \partial_x + \gamma \nabla^2) \mathcal{G}^\ddagger(\mathbf{x}|\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0), \quad (4.142)$$

where  $\delta(\mathbf{x} - \mathbf{x}_0)$  is the **Dirac delta** with source at  $\mathbf{x} = \mathbf{x}_0$ . Following the derivation of the diffusion equation **reciprocity condition** in VOLUME 1, we can relate the adjoint free space

<sup>16</sup>We encounter equation (4.140) again in Section 5.7.5 when studying the western intensification of ocean gyres. Namely, the analysis of [Stommel \(1948\)](#) leads to the Stommel equation (5.100) for the streamfunction, which is the same as the beta plume equation (4.140).

Green's function,  $\mathcal{G}^\dagger(\mathbf{x}|\mathbf{x}_0)$  to the free space Green's function,  $\mathcal{G}(\mathbf{x}|\mathbf{x}_0)$ , through

$$\mathcal{G}^\dagger(\mathbf{x}|\mathbf{x}_0) = \mathcal{G}(\mathbf{x}_0|\mathbf{x}), \quad (4.143)$$

where  $\mathcal{G}(\mathbf{x}|\mathbf{x}_0)$  satisfies

$$-(\beta^{\text{eff}} \partial_x + \gamma \nabla^2) \mathcal{G}(\mathbf{x}|\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0). \quad (4.144)$$

Since the Dirac delta is a positive point source, it provides a positive point source to for the Green's function equation (4.144). We refer to the Green's function,  $\mathcal{G}(\mathbf{x}|\mathbf{x}_0)$ , as the **beta plume**, with this name motivated by the sketch in Figure 4.7. That is, consider the case of  $\beta^{\text{eff}} > 0$ , so that the beta plume streamlines extend to the west of the Dirac delta source. The Rayleigh drag,  $\gamma > 0$ , causes the streamlines to spread both zonally and meridionally westward away from the source.

We can determine an analytical expression for the Green's function,  $\mathcal{G}(\mathbf{x}|\mathbf{x}_0)$ , through the substitution

$$\mathcal{G}(\mathbf{x}|\mathbf{x}_0) = e^{-(x-x_0)/L_s} \Phi(\mathbf{x}|\mathbf{x}_0), \quad (4.145)$$

where we introduced the length scale

$$L_s \equiv 2\gamma/\beta^{\text{eff}}. \quad (4.146)$$

It is notable that this length scale increases when increasing the Rayleigh drag, which reflects the contribution of the Rayleigh drag to the westward spreading of the streamlines. The substitution (4.145) brings the Green's function problem (4.144) into the form

$$-(\nabla^2 - L_s^{-2}) \Phi(\mathbf{x}|\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0)/\gamma, \quad (4.147)$$

which is the two-dimensional version of the **screened Poisson equation** from VOLUME 1. Note that for the right hand side we set  $\delta(\mathbf{x} - \mathbf{x}_0) e^{-(x-x_0)/L_s} = \delta(\mathbf{x} - \mathbf{x}_0)$ , since the exponential factor equals unity at  $x = x_0$  and so it does not alter the Dirac delta.<sup>17</sup>

The solution to equation (4.147) is proportional to the Hankel function of the first kind and zeroth order,  $H_0^{(1)}$ , with the argument a pure imaginary number.<sup>18</sup> Reintroducing the exponential scaling from equation (4.145) then renders the beta plume Green's function

$$\mathcal{G}(\mathbf{x}|\mathbf{x}_0) = \frac{i}{4\gamma} H_0^{(1)}(ir/L_s) e^{-(x-x_0)/L_s}, \quad (4.148)$$

where  $i = \sqrt{-1}$ . We plot the magnitude of this Green's function in Figure 4.7. The exponential scaling makes the beta plume highly asymmetric, with more amplitude to the west of the Dirac source. The Dirac source in equation (4.144) provides a positive point source, so that the circulation is counter-clockwise around the source point.

### 4.8.5 Further study

Much from this section follows [Rhines \(1980\)](#) and [Haine and Fuller \(2016\)](#). [Welander \(1968\)](#) suggested interpreting the steady streamfunction equation (4.129) in terms of a steady advective-diffusive-source balance, and we pursue a similar interpretation for the steady shallow water planetary geostrophic flow in Section (5.7.8).

<sup>17</sup>We detail this property of the Dirac delta in VOLUME 1.

<sup>18</sup>The Hankel function with an imaginary argument is sometimes referred to as MacDonalds' function.

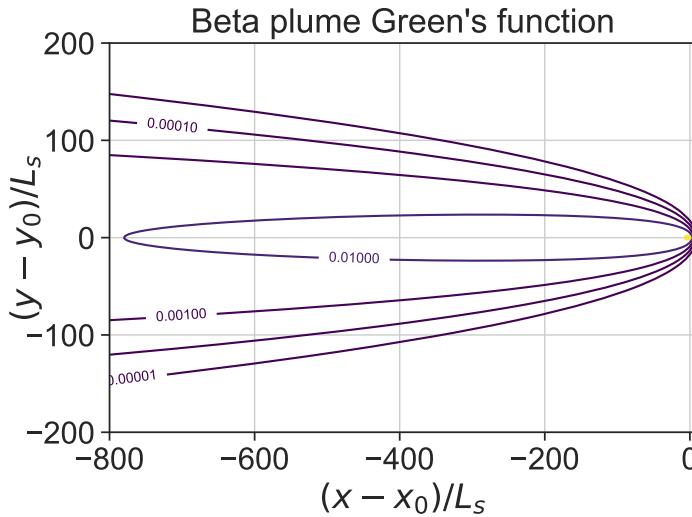


FIGURE 4.7: Sketch of the magnitude for the non-dimensional beta plume Green's function,  $\gamma \mathcal{G}(\mathbf{x}|\mathbf{x}_0)$ , given by equation (4.148) and here shown with  $\beta^{\text{eff}} = 1.62 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$ , which is equal to the planetary  $\beta$  at latitude  $\pi/4$ . We furthermore choose the Rayleigh drag of  $\gamma = (2 \text{ year})^{-1}$ , and scale the coordinate axes by the length,  $L_s = 2\gamma/\beta^{\text{eff}} \approx 2 \text{ km}$ , from equation (4.146).

## 4.9 Barotropic circulation on a rotating sphere

Thus far in this chapter, we have made use of tangent plane Cartesian coordinates to study flow on the tangent plane. This model, particularly the  $\beta$ -plane, provides a rich theoretical framework that serves many purposes in geophysical fluid mechanics. However, it is important to recognize that a more faithful rendition of planetary flows, especially atmospheric flows, considers motion on the sphere. Hence, in this section we study the barotropic model on a smooth sphere (Figure 4.8), making use of the spherical coordinates summarized in VOLUME 1 and with all flow taking place on a constant geopotential denoted by  $r = R$ .

Before starting, it is useful to highlight a central distinction between flows on the tangent plane versus flows on the sphere. Namely, flows on the tangent plane lose their physical relevance when extending the boundary of the plane too far in the meridional direction, thus prompting the use either of walls (e.g., zonal channels) or double periodicity (periodicity in both the zonal and meridional directions). In contrast, barotropic flow on a smooth sphere experiences no lateral boundaries. Hence, when making use of the tangent plane, results that depend on lateral boundaries should be discounted for purposes of connecting to motion on the sphere.

### 4.9.1 Zonal and meridional velocity components

Divergence of the angular (“horizontal”) flow vanishes

$$\nabla \cdot \mathbf{u} = \frac{1}{r \cos \phi} \frac{\partial u}{\partial \lambda} + \frac{1}{r \cos \phi} \frac{\partial (v \cos \phi)}{\partial \phi} = r_{\perp}^{-1} [\partial_{\lambda} u + \partial_{\phi} (v \cos \phi)] = 0, \quad (4.149)$$

with the zonal and meridional velocity components<sup>19</sup>

$$u = (r \cos \phi) D\lambda/Dt \quad \text{zonal velocity component} \quad (4.150a)$$

<sup>19</sup>For a discussion of these velocity components, see the chapter in VOLUME 1 concerned with particle motion around a sphere.

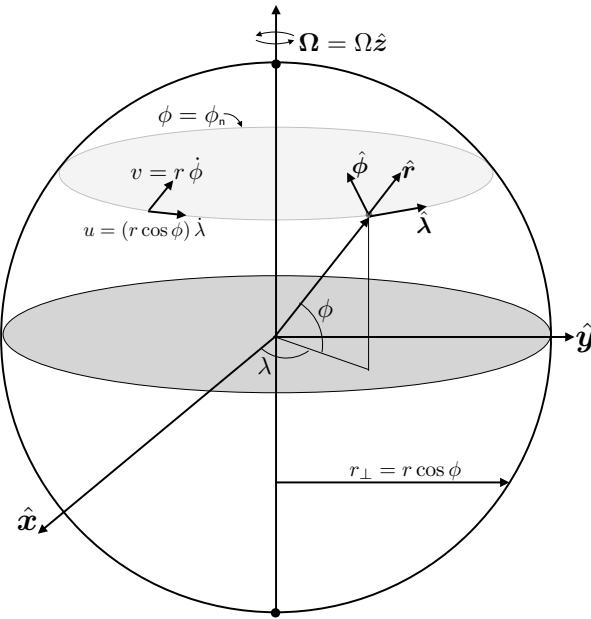


FIGURE 4.8: Depicting the geometry of flow on a smooth geopotential on a rotating planet. Here we show the zonal velocity component,  $u = (r \cos \phi) D\lambda / Dt = (r \cos \phi) \dot{\lambda}$ , and meridional velocity component,  $v = r D\phi / Dt = r \dot{\phi}$ , and examine circulation around a constant latitude loop,  $\phi = \phi_n$ . The distance to the rotational axis,  $r_\perp = r \cos \phi$ , is the moment-arm for use in computing the axial angular momentum.

$$v = r D\phi / Dt \quad \text{meridional velocity component.} \quad (4.150b)$$

Since the velocity is non-divergent, we can introduce a streamfunction,

$$\mathbf{u} = \hat{\mathbf{r}} \times \nabla \psi, \quad (4.151)$$

which takes on the following form

$$\hat{\mathbf{r}} \times \nabla \psi = (\hat{\mathbf{r}} \times \hat{\boldsymbol{\lambda}}) (r \cos \phi)^{-1} \partial_\lambda \psi + (\hat{\mathbf{r}} \times \hat{\boldsymbol{\phi}}) r^{-1} \partial_\phi \psi \quad (4.152a)$$

$$= \hat{\boldsymbol{\phi}} (r \cos \phi)^{-1} \partial_\lambda \psi - \hat{\boldsymbol{\lambda}} r^{-1} \partial_\phi \psi, \quad (4.152b)$$

so that

$$u = -r^{-1} \partial_\phi \psi \quad \text{and} \quad v = (r \cos \phi)^{-1} \partial_\lambda \psi. \quad (4.153)$$

The radial component to the relative vorticity is given by

$$\zeta = \hat{\mathbf{r}} \cdot (\nabla \times \mathbf{u}) = r_\perp^{-1} [\partial_\lambda v - \partial_\phi (u \cos \phi)], \quad (4.154)$$

which equals to the Laplacian of the streamfunction

$$\zeta = \nabla^2 \psi = \frac{1}{(r \cos \phi)^2} \frac{\partial^2 \psi}{\partial \lambda^2} + \frac{1}{r^2 \cos \phi} \frac{\partial}{\partial \phi} \left[ \cos \phi \frac{\partial \psi}{\partial \phi} \right]. \quad (4.155)$$

### 4.9.2 Vector-invariant velocity equation

To help become familiar with spherical coordinates, we here derive the vector-invariant velocity equation starting from the spherical coordinate version of the velocity equation<sup>20</sup>

$$\hat{\lambda} Du/Dt + \hat{\phi} Dv/Dt + r^{-1} \tan \phi (u + 2\Omega r_\perp) (\hat{r} \times \mathbf{u}) = -\nabla \varphi, \quad (4.156)$$

which has the component expressions

$$\partial_t u + \nabla \cdot (\mathbf{u} u) - r^{-1} u v \tan \phi - 2\Omega v \sin \phi = -r_\perp^{-1} \partial_\lambda \varphi \quad (4.157a)$$

$$\partial_t v + \nabla \cdot (\mathbf{u} v) + r^{-1} u^2 \tan \phi + 2\Omega u \sin \phi = -r^{-1} \partial_\phi \varphi. \quad (4.157b)$$

In Exercise 4.13 we prove the identity

$$[(\mathbf{u} \cdot \nabla) u - r_\perp^{-1} u v \tan \phi] \hat{\lambda} + [(\mathbf{u} \cdot \nabla) v + r_\perp^{-1} u^2 \tan \phi] \hat{\phi} = \nabla \mathcal{K} + \zeta \hat{r} \times \mathbf{u}, \quad (4.158)$$

where

$$\mathcal{K} = \mathbf{u} \cdot \mathbf{u}/2 \quad (4.159)$$

is the kinetic energy per mass. The identity (4.158) then brings the velocity equation (4.156) into its vector invariant form

$$\partial_t \mathbf{u} + (f + \zeta) \hat{r} \times \mathbf{u} = -\nabla(\varphi + \mathcal{K}), \quad (4.160)$$

where we introduced twice the radial component to the planetary angular rotation rate (i.e., the Coriolis parameter)

$$f = 2\Omega \sin \phi. \quad (4.161)$$

As expected, the spherical coordinate vector-invariant velocity equation (4.160) has the same form as the Cartesian coordinate version (4.5).

### 4.9.3 Vorticity equation

To derive the vorticity equation, we can proceed just like in Section 4.3.1 by taking the curl of the vector-invariant velocity equation (4.160). We can also invoke coordinate invariance of the vorticity equation (4.39), thus allowing us to write down, by inspection, the spherical coordinate version

$$\partial_t \zeta + \nabla \cdot (\mathbf{u} \zeta) = -\beta v \iff \frac{D(\zeta + f)}{Dt} = 0. \quad (4.162)$$

In this equation, the velocity, vorticity, and divergence operator take their spherical coordinate forms so that

$$\partial_t \zeta + r_\perp^{-1} [\partial_\lambda(u \zeta) + \partial_\phi(v \zeta \cos \phi)] = -\beta v, \quad (4.163)$$

with the  $\beta$  parameter given by

$$\beta = r^{-1} \partial_\phi f = (2\Omega/r) \cos \phi. \quad (4.164)$$

---

<sup>20</sup>The velocity equation (4.156) arises from setting  $w = 0$ ,  $\dot{w} = 0$ , and  $\partial_z = 0$  in the geopotential coordinate velocity equations studied in VOLUME 2.

#### 4.9.4 Area conservation

We are particularly interested in establishing some identities connecting properties along a constant latitude circle to properties integrated over the region poleward of this latitude. Without loss of generality we focus on the region north of  $\phi = \phi_n$  (Figure 4.8). The first property we establish follows from noting that  $\nabla \cdot \mathbf{u} = 0$  leads to a zero zonally integrated meridional flow. Namely, the divergence theorem implies that for any simply connected region

$$0 = \int_{\mathcal{S}} \nabla \cdot \mathbf{u} d\mathcal{S} = \oint_{\partial\mathcal{S}} \mathbf{u} \cdot \hat{\mathbf{n}} d\ell. \quad (4.165)$$

Letting  $\mathcal{S}$  be the polar cap, so that its only boundary is along the latitude circle  $\phi = \phi_n$ , renders

$$\int_0^{2\pi} v r \cos \phi_n d\lambda = r \cos \phi_n \int_0^{2\pi} v d\lambda = 0, \quad (4.166)$$

where we noted that the distance to the rotational axis,  $r \cos \phi_n$ , is constant for the zonal integral. The constraint (4.166) means that flow moving north across any latitude circle is exactly compensated by flow moving south across the same latitude. The constraint is a result of area conservation holding for the two-dimensional non-divergent flow, whereby the area of any region is conserved by the flow.<sup>21</sup>

#### 4.9.5 Circulation around a zonal loop

Consider the circulation around a constant latitude (zonal) loop ( $\phi = \phi_n$ )

$$\mathcal{C} = \oint_{\partial\mathcal{S}} u r \cos \phi d\lambda = r \cos \phi_n \int_0^{2\pi} u d\lambda. \quad (4.167)$$

We oriented the circulation loop in a counter-clockwise direction looking down from above the north pole, which accords with the direction of the Earth's rotation. Furthermore, we see from equation (4.167) that the zonally averaged zonal flow is related to the circulation via

$$\bar{u} = \frac{1}{\int_0^{2\pi} r \cos \phi_n d\lambda} \int_0^{2\pi} u r \cos \phi_n d\lambda = \frac{1}{2\pi} \int_0^{2\pi} u d\lambda = \frac{\mathcal{C}}{2\pi r \cos \phi_n}. \quad (4.168)$$

#### 4.9.6 Stokes' theorem and the zonal averaged zonal flow

Stokes' theorem says the circulation around a latitude circle equals to the radial component of the vorticity integrated over the area north of the latitude used for computing the circulation

$$\mathcal{C} = \iint_{\mathcal{S}} \zeta r^2 \cos \phi d\phi d\lambda = r^2 \int_0^{2\pi} \int_{\phi_n}^{\pi/2} \zeta \cos \phi d\phi d\lambda. \quad (4.169)$$

Equating the circulation in equation (4.167) to the vorticity integral (4.169) leads to

$$\cos \phi_n \int_0^{2\pi} u d\lambda = r \int_0^{2\pi} \int_{\phi_n}^{\pi/2} \zeta \cos \phi d\phi d\lambda. \quad (4.170)$$

---

<sup>21</sup>We study kinematic properties of non-divergent flows in VOLUME 1.

Evidently, Stokes' theorem connects any changes to the area integrated relative vorticity (right hand side) to changes to the zonal circulation (left hand side). We use this result to express the zonal mean zonal flow at latitude,  $\phi_n$ , to the vorticity integrated over the polar cap

$$\bar{u} = \frac{r \int_0^{2\pi} \int_{\phi_n}^{\pi/2} \zeta \cos \phi d\phi d\lambda}{2\pi \cos \phi_n}. \quad (4.171)$$

Evidently, there are two ways to examine processes leading to changes in the zonal mean zonal winds. First, we can study momentum balances along the latitude  $\phi_n$ , as per the discussion in Section 4.9.8. Alternatively, we can examine processes leading to changes in the area integrated vorticity in the region north of  $\phi_n$ . As we see with the thought experiment at the end of Section 4.9.8, the kinematic result (4.171) has broad implications for dynamical processes, either linear or nonlinear, that affect a transfer of momentum.

### 4.9.7 Planetary circulation

The velocity of a point fixed to the rotating sphere is purely zonal,

$$\mathbf{u}_p = \boldsymbol{\Omega} \times \mathbf{r} = (r \Omega \cos \phi) \hat{\lambda}, \quad (4.172)$$

and this velocity is non-divergent

$$\nabla \cdot \mathbf{u}_p = 0. \quad (4.173)$$

Referring to equation (4.152b) we have the corresponding streamfunction

$$\partial_\phi \psi_p = -r^2 \Omega \cos \phi \implies \psi_p = -r^2 \Omega \sin \phi = -r^2 f/2, \quad (4.174)$$

where we introduced the Coriolis parameter,  $f = 2\Omega \sin \phi$ .

With  $\mathbf{u}_p = (r \Omega \cos \phi) \hat{\lambda}$ , the planetary circulation around the  $\phi = \phi_n$  latitude circle is

$$\mathcal{C}_{\text{planet}} = \int_0^{2\pi} \mathbf{u}_p \cdot \hat{\lambda} r \cos \phi_n d\lambda = 2\pi \Omega (r \cos \phi_n)^2, \quad (4.175)$$

so that the absolute circulation is

$$\mathcal{C}_a = 2\pi r \cos \phi_n (r \Omega \cos \phi_n + \bar{u}). \quad (4.176)$$

Equating the planetary circulation, through Stokes' theorem, to the area integrated planetary vorticity yields

$$\mathcal{C}_{\text{planet}} = 2\Omega \int_0^{2\pi} \int_{\phi_n}^{\pi/2} r^2 \cos \phi d\phi d\lambda = 4\pi r^2 \Omega \int_{\phi_n}^{\pi/2} d(\sin^2 \phi) d\phi = 2\pi \Omega (r \cos \phi_n)^2, \quad (4.177)$$

which agrees with the circulation computed in equation (4.175).

### 4.9.8 Connecting circulation and axial angular momentum

As studied in VOLUME 2, the axial angular momentum per mass is given by

$$l^z = r_\perp (u + \Omega r_\perp), \quad (4.178)$$

which has a zonal average

$$\bar{l}^z = r_{\perp} (\bar{u} + \Omega r_{\perp}) = \mathcal{C}_a / 2\pi. \quad (4.179)$$

From the chapter on momentum in VOLUME 2, we know that for the inviscid barotropic fluid the axial angular momentum satisfies the conservation law

$$\partial_t l^z + \nabla \cdot (\mathbf{u} l^z) = -\partial_{\lambda} \varphi \iff \partial_t l^z + r_{\perp}^{-1} [\partial_{\lambda}(u l^z) + \partial_{\phi}(v l^z \cos \phi)] = -\partial_{\lambda} \varphi. \quad (4.180)$$

Integrating around a constant latitude circle,  $\int_0^{2\pi} r_{\perp} d\lambda$ , eliminates the zonal pressure gradient as well as the zonal convergence of axial angular momentum, thus leaving<sup>22</sup>

$$\frac{\partial}{\partial t} \int_0^{2\pi} l^z r_{\perp} d\lambda = - \int_0^{2\pi} \partial_{\phi}(v l^z \cos \phi) d\lambda. \quad (4.181)$$

If we further integrate over the latitude domain, from pole to pole, we find that the right hand side vanishes since  $\cos \phi = 0$  at the poles, thus showing that the total axial angular momentum remains constant in the absence of external torques. But instead of integrating over latitudes, we focus on the zonally integrated budget (4.181), with the aim of establishing an understanding of what drives zonal mean zonal flow.

On the left hand side of equation (4.181) we can remove the moment arm,  $r_{\perp} = r \cos \phi$ , from the zonal integral as well as from the time derivative, and for the right hand side we can exchange the zonal integral and the meridional derivative, thus leading to

$$r_{\perp} \frac{\partial}{\partial t} \int_0^{2\pi} l^z d\lambda = - \frac{\partial}{\partial \phi} \int_0^{2\pi} v l^z \cos \phi d\lambda. \quad (4.182)$$

Evidently, the zonally integrated axial angular momentum changes in time according to the meridional convergence of axial angular momentum transported by the meridional flow. In effect, the meridional convergence acts as a torque to modify the axial angular momentum. To find an expression for the zonally averaged velocity, we make use of the definition (4.178) for the axial angular momentum per mass, and note that the only time dependence arises from the zonal velocity, so that

$$2\pi r_{\perp}^2 \partial_t \bar{u} = - \frac{\partial}{\partial \phi} \int_0^{2\pi} v (u + r_{\perp} \Omega) r_{\perp} \cos \phi d\lambda. \quad (4.183)$$

For the right hand side, recall the area conservation result (4.166), which allows us to drop the  $r_{\perp} \Omega$  term to find

$$r_{\perp}^2 \partial_t \bar{u} = -r \partial_{\phi}(\bar{u} \bar{v} \cos^2 \phi) \iff \partial_t \bar{u} = -(r \cos^2 \phi)^{-1} \partial_{\phi}(\bar{u} \bar{v} \cos^2 \phi). \quad (4.184)$$

Now decompose the nonlinear product according to the zonal mean and deviations from the mean,

$$uv = (\bar{u} + u')(\bar{v} + v') = (\bar{u} + u')v', \quad (4.185)$$

where the second equality follows since  $\bar{v} = 0$  from equation (4.166). With  $\int_0^{2\pi} \bar{u} v' d\lambda = \bar{u} \int_0^{2\pi} v' d\lambda = 0$ , equation (4.184) reduces to

$$\partial_t \bar{u} = -(r \cos^2 \phi)^{-1} \partial_{\phi}(\bar{u}' \bar{v}' \cos^2 \phi). \quad (4.186)$$

---

<sup>22</sup>We write  $\partial_t$ , rather than  $d/dt$ , when acting on the integral in equation (4.182) since the derivative is computed holding the latitude fixed.

This result states that the time tendency for the zonally averaged zonal flow (left hand side) is driven by meridional convergence of the zonally averaged correlation between fluctuations in the zonal and meridional velocity (right hand side). That is, the zonal mean flow is driven by meridional convergence of the momentum flux. The non-negative weighting from  $\cos^2 \phi$  acts to scale down the correlations moving towards the poles. Expanding the product rule leads to

$$r \partial_t \bar{u} = -\partial_\phi(\bar{u}' \bar{v}') + 2\bar{u}' \bar{v}' \tan \phi, \quad (4.187)$$

which reveals a singularity at the pole arising from the vanishing moment-arm and the corresponding zero length of a latitude circle.

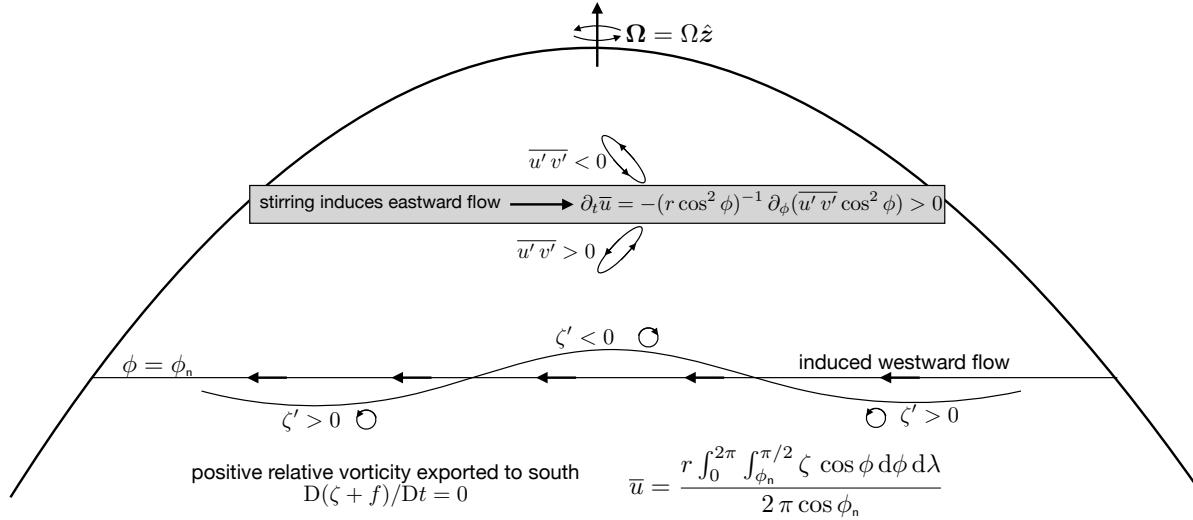


FIGURE 4.9: This figure depicts a region of localized stirring in the northern hemisphere, where it is assumed it imparts zero net angular momentum to the flow. The stirring sends signals meridionally (e.g., Rossby wave packets) that cause fluid particles originally on constant latitude circles to fluctuate meridionally. Consider a region outside the stirring region at some arbitrary latitude  $\phi = \phi_n$ , and assume the flow is initially in a near rigid body motion (i.e., small Rossby number). Since the absolute vorticity,  $\zeta + f$ , is materially conserved, and with  $\beta > 0$ , any meridional fluctuations cause positive relative vorticity to be exported southward across a latitude circle outside of the stirring region. As a result, the area integrated relative vorticity decreases in the region north of  $\phi = \phi_n$ . This decrease, through the Stokes' theorem result (4.171), renders a westward acceleration at latitudes outside of the stirring region. Since there is no net angular momentum introduced through the stirring (by assumption), and since Stokes' theorem renders a westward flow at all latitudes outside of the stirring region, then we infer that there is an eastward acceleration within the stirring region. From the angular momentum budget analysis leading to equation (4.186), an eastward acceleration requires convergence of angular momentum onto the stirring region:  $\partial_t \bar{u} = \partial_t \bar{u} = -(r \cos^2 \phi)^{-1} \partial_\phi(\bar{u}' \bar{v}' \cos^2 \phi) > 0$ .

Note that equation (4.186) can be found through manipulations of the zonal momentum equation (4.157a). We chose to present the derivation starting from axial angular momentum given its connection to circulation around a latitude circle.

Figure 4.9, inspired by a thought experiment from [Held \(2000\)](#), summarizes the Stokes' theorem result (4.171) and the zonal acceleration (4.186) arising from the angular momentum budget. The arguments detailed in the figure caption hold for any dynamical fluctuation that is induced in the region outside of the stirring. Most notably, the linear Rossby waves studied in VOLUME 4 manifest the angular momentum convergence as per equation (4.186).

### 4.9.9 Connecting absolute vorticity and axial angular momentum

The meridional derivative of the axial angular momentum (4.178) is given by

$$\partial_\phi l^z = -u r \sin \phi - r^2 \Omega \sin \phi \cos \phi + r \partial_\phi u \cos \phi - r^2 \Omega \cos \phi \sin \phi \quad (4.188a)$$

$$= \partial_\phi(r u \cos \phi) - f r^2 \cos \phi \quad (4.188b)$$

$$= r^2 \cos \phi \left[ \frac{1}{r \cos \phi} \frac{\partial(u \cos \phi)}{\partial \phi} - f \right] \quad (4.188c)$$

$$= -r^2 \cos \phi (\zeta + f - r_\perp^{-1} \partial_\lambda v), \quad (4.188d)$$

where we introduced the radial component to the relative vorticity,  $\zeta = \hat{r} \cdot (\nabla \times \mathbf{u}) = r_\perp^{-1} [\partial_\lambda v - \partial_\phi(u \cos \phi)]$  from equation (4.154). We thus find the meridional convergence of axial angular momentum is given by

$$-\frac{1}{r \cos \phi} \frac{\partial l^z}{\partial \phi} = r (\zeta + f - r_\perp^{-1} \partial_\lambda v). \quad (4.189)$$

The area integral of this equation between two latitudes,  $\phi_s$  and  $\phi_n$ , leads to the identity

$$\int_0^{2\pi} [l^z(\phi_n) - l^z(\phi_s)] d\lambda = - \int_{\phi_s}^{\phi_n} \int_0^{2\pi} (\zeta + f) r^2 \cos \phi d\lambda d\phi. \quad (4.190)$$

Evidently, the difference in axial angular momentum, as integrated along two latitude circles, equals to minus the area integrated absolute vorticity. We depict this balance in Figure 4.10.

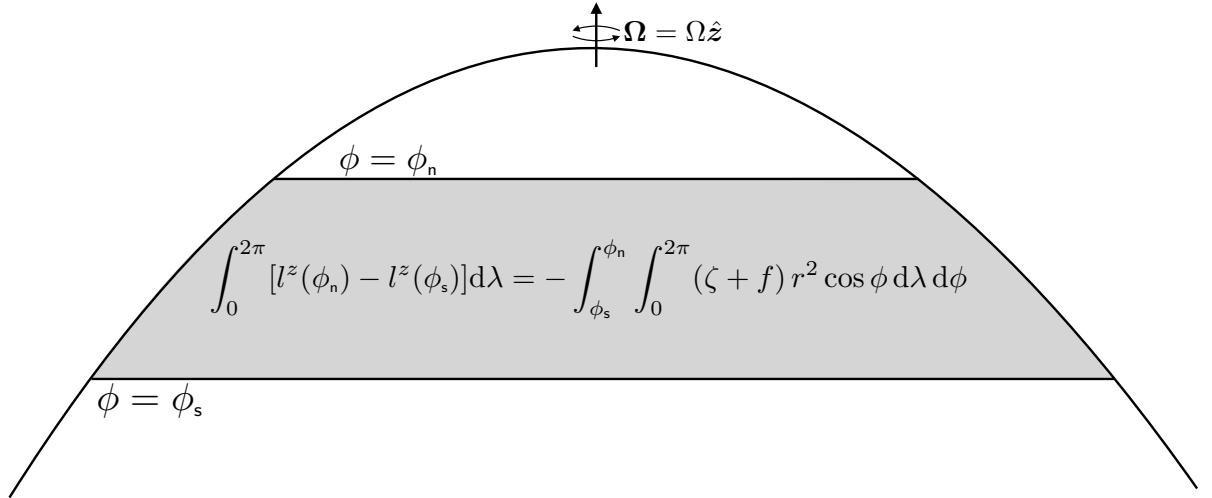


FIGURE 4.10: Illustrating the balance (4.190) between area integrated absolute vorticity and the angular momentum integrated over two latitude circles, here depicted in the northern hemisphere. Implications for this balance are detailed in Section 4.9.9.

Let us consider some thought experiments to help understand the result (4.190). First, consider an integral that extends to the poles, in which case the left hand side vanishes since there is no axial angular momentum at the poles. We thus find that the global area integrated radial component to the absolute vorticity vanishes

$$\int_{-\pi/2}^{\pi/2} \int_0^{2\pi} (\zeta + f) r^2 \cos \phi d\lambda d\phi = 0. \quad (4.191)$$

Now the global integral of the Coriolis parameter vanishes since  $\sin \phi$  and  $\cos \phi$  are orthogonal

$$\int_{-\pi/2}^{\pi/2} f \cos \phi d\phi = 2\Omega \int_{-\pi/2}^{\pi/2} \sin \phi \cos \phi d\phi = 0. \quad (4.192)$$

We thus find a vanishing global area integrated radial component to the relative vorticity

$$\int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \zeta r^2 \cos \phi d\lambda d\phi = 0. \quad (4.193)$$

This result can be verified mathematically through use of equation (4.154) for relative vorticity, in which case

$$\int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \zeta r_\perp r d\lambda d\phi = \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} [\partial_\lambda v - \partial_\phi(u \cos \phi)] r d\lambda d\phi = 0, \quad (4.194)$$

where zonal periodicity eliminates the  $\partial_\lambda v$  term, and  $\cos \phi = 0$  at the poles eliminates the  $\partial_\phi(u \cos \phi)$  term.

Next, consider the case of rigid body rotation ( $\mathbf{u} = 0$ ) so that  $\zeta = 0$ . Here, the axial angular momentum decreases going poleward, as expected due to the decreasing moment-arm length (distance from the axis of rotation). We thus find the left hand side of equation (4.190) is negative for the northern hemisphere and positive for the southern hemisphere. This sign change is reflected on the right hand side by a sign change for the Coriolis parameter when crossing the equator.

Finally, allow for a nonzero relative vorticity. For large-scale flows, the relative vorticity is much smaller in magnitude than the planetary vorticity, so that the axial angular momentum continues to decrease moving poleward. Indeed, if the relative vorticity becomes large enough to overcome the effects from planetary rotation by making  $f(f + \zeta) < 0$ , then the flow becomes inertially unstable as per the discussion in VOLUME 4. In this case, [inertial instability](#) rapidly returns the flow towards a stable state.<sup>23</sup>

#### 4.9.10 Further study

Elements of this section were inspired by Section 1 from the Woods Hole Geophysical Fluid Dynamics summer program lectures by [Held \(2000\)](#) on the atmospheric general circulation.



## 4.10 Exercises

**EXERCISE 4.1: VORTICITY IDENTITY IN A TWO-DIMENSIONAL NON-DIVERGENT BAROTROPIC FLOW**

Directly prove the identity (4.33) holding for the two-dimensional non-divergent barotropic flow.

**EXERCISE 4.2: ALTERNATIVE FORM OF THE VORTICITY FLUX**

---

<sup>23</sup>In stratified baroclinic flows, the related criterion for symmetric instability involves negative Ertel potential vorticity and slantwise displacements.

Show that the vorticity flux can be written

$$\mathbf{u} \cdot \boldsymbol{\zeta} = u \nabla v - v \nabla u + \hat{\mathbf{z}} \times \nabla \mathcal{K}, \quad (4.195)$$

so that the vorticity equation (4.39) can be written in the alternative form

$$\frac{\partial \zeta}{\partial t} + \nabla \cdot (u \nabla v - v \nabla u) = -\beta v. \quad (4.196)$$

As a corollary, we see that steady  $f$ -plane flow satisfies the constraint

$$\nabla \cdot (u \nabla v - v \nabla u) = u \nabla^2 v - v \nabla^2 u = 0. \quad (4.197)$$

#### EXERCISE 4.3: EXAMPLE TWO-DIMENSIONAL NON-DIVERGENT FLOW

Consider a perfect two-dimensional non-divergent flow in a non-rotating reference frame

$$\frac{D\mathbf{u}}{Dt} = -\nabla \varphi \quad \text{and} \quad \nabla \cdot \mathbf{u} = 0. \quad (4.198)$$

Let the velocity be given by the steady flow

$$\mathbf{u} = U [\hat{\mathbf{x}} \sin(k y) + \hat{\mathbf{y}} \sin(k x)], \quad (4.199)$$

where  $U$  is a constant with dimensions  $L T^{-1}$  and  $k$  is a wavenumber with dimensions  $L^{-1}$ . We provide a sketch of this flow in Figure 4.11.

- (a) Compute the streamfunction,  $\psi$ , so that  $\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi$ .
- (b) Compute the self-advection,  $(\mathbf{u} \cdot \nabla) \mathbf{u}$  and show that  $\nabla \times [(\mathbf{u} \cdot \nabla) \mathbf{u}] = \mathbf{0}$ .
- (c) Compute the vorticity,  $\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u})$ .
- (d) Compute the pressure, to within an arbitrary constant.

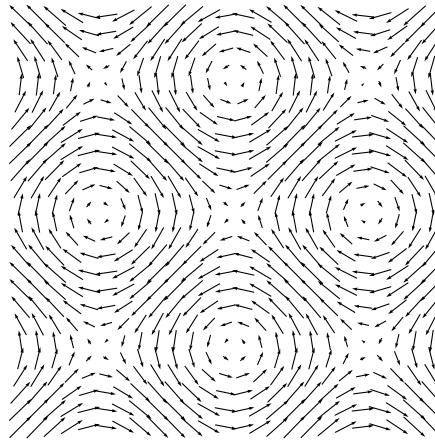


FIGURE 4.11: A sketch of the steady two-dimensional non-divergent sinusoidal flow (4.199), as given by  $\mathbf{u} = U [\hat{\mathbf{x}} \sin(k y) + \hat{\mathbf{y}} \sin(k x)]$ . The units are arbitrary. This flow is considered in Exercise 4.3.

#### EXERCISE 4.4: VELOCITY ARISING FROM A GIVEN VORTICITY

Following the discussion in Section 4.3.8, consider an axially symmetric two-dimensional non-

divergent flow with a single vortex of the form

$$\zeta(r) = \begin{cases} \zeta_0 & \text{for } r < r_0 \\ 0 & \text{for } r > r_0. \end{cases} \quad (4.200)$$

- (a) What is the velocity field corresponding to this vorticity?
- (b) What is the circulation around a circular circuit with radius  $r < r_0$ ? Assume the velocity is non-singular at the origin.
- (c) What is the circulation around a circular circuit with radius  $r > r_0$ ?
- (d) For both the circular circuits with  $r < r_0$  and  $r > r_0$ , write the circulation in terms of the velocity.

#### EXERCISE 4.5: INTEGRAL PROPERTIES OF THE INVISCID 2D NON-DIVERGENT FLOW

In this exercise, we establish some domain integrated conservation properties for inviscid two-dimensional non-divergent flow on a  $\beta$ -plane. Assume the geometry is a flat plane defined over a finite region,  $\mathcal{S}$ , with static material boundary,  $\partial\mathcal{S}$ . Many of the properties derived here are discussed in Section 3.1 of [McWilliams \(2006\)](#).

- (a) Show that the domain integrated kinetic energy per mass remains constant in time

$$\frac{d}{dt} \int_{\mathcal{S}} \mathcal{K} d\mathcal{S} = \frac{1}{2} \frac{d}{dt} \int_{\mathcal{S}} \mathbf{u} \cdot \mathbf{u} d\mathcal{S} = 0, \quad (4.201)$$

where the horizontal integral extends over the full fluid domain  $\mathcal{S}$ .

- (b) Why is the mechanical energy budget only associated with kinetic energy? What about the gravitational potential energy?
- (c) Show that the domain integrated relative vorticity (equal also to the relative circulation) is constant in time

$$\frac{dC}{dt} = \frac{d}{dt} \int_{\mathcal{S}} \zeta d\mathcal{S} = 0. \quad (4.202)$$

- (d) Show that the domain integrated enstrophy is constant in time for  $f$ -plane motion ( $\beta = 0$ )

$$\frac{dZ^{(\zeta)}}{dt} = \frac{d}{dt} \int_{\mathcal{S}} \zeta^2 d\mathcal{S} = 0. \quad (4.203)$$

- (e) Show that the domain integrated potential enstrophy is constant in time even with  $\beta \neq 0$

$$\frac{dZ^{(q)}}{dt} = \frac{d}{dt} \int_{\mathcal{S}} q^2 d\mathcal{S} = 0. \quad (4.204)$$

#### EXERCISE 4.6: ALTERNATIVE EXPRESSION FOR THE DOMAIN INTEGRATED KINETIC ENERGY

For a simply connected region,  $\mathcal{S}$ , with static material boundary,  $\partial\mathcal{S}$ , show that the globally integrated kinetic energy per mass can be written

$$\int_{\mathcal{S}} \mathcal{K} d\mathcal{S} = \frac{1}{2} \int_{\mathcal{S}} \mathbf{u} \cdot \mathbf{u} d\mathcal{S} = \frac{1}{2} \int_{\mathcal{S}} \zeta (\psi_b - \psi) d\mathcal{S} = \frac{1}{2} [\psi_b C - \int_{\mathcal{S}} \psi \zeta d\mathcal{S}], \quad (4.205)$$

where  $\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi$  is the horizontally non-divergent velocity,  $\psi$  is the streamfunction,  $\psi_b$  is the streamfunction on the boundary, and  $\zeta = \nabla^2 \psi$  is the vorticity. Hint: from our study of [non-divergent flow](#) kinematics in VOLUME 2, the streamfunction for two-dimensional non-divergent flow is a constant on material boundaries.

#### EXERCISE 4.7: CIRCULATION IN A 2D BAROTROPIC FLOW

Consider a non-divergent barotropic flow on a  $\beta$ -plane in the presence of a biharmonic friction operator, where the governing vorticity equation is

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta + \beta y) = -\nu \nabla^4 \zeta, \quad (4.206)$$

with  $\nu > 0$  a constant biharmonic viscosity with dimensions of  $L^4 T^{-1}$ . Show that the circulation around a fixed material area,  $\mathcal{S}$ , in the fluid evolves according to

$$\frac{d\mathcal{C}}{dt} = - \oint_{\partial\mathcal{S}} \left[ \psi \frac{\partial q}{\partial s} + \nu \frac{\partial(\nabla^2 \zeta)}{\partial n} \right] ds, \quad (4.207)$$

where  $s$  is the arc-length along the boundary of the region and  $n$  is a coordinate normal to the boundary.

#### EXERCISE 4.8: ANGULAR MOMENTUM FOR BAROTROPIC FLOW IN A BASIN

The exercise derives some equations presented in [Holloway and Rhines \(1991\)](#), who offer a specialized example of the shallow water angular momentum discussed in Section 2.8.

As in Section 2.8.1, the relative angular momentum for a region of fluid is given by

$$\mathbf{L} = \int d\mathcal{S} \int (\mathbf{x} \times \mathbf{v}) \rho dz, \quad (4.208)$$

where  $\mathbf{x}$  is the position vector and the relative angular momentum is that due to the motion of the fluid with respect to the rigid-body. For a barotropic fluid of constant density,  $\rho$ , and constant thickness,  $H$ , and correspondingly a zero vertical velocity, the relative angular momentum reduces to

$$\mathbf{L} = \rho H \int_{\mathcal{S}} (\mathbf{x} \times \mathbf{u}) d\mathcal{S}, \quad (4.209)$$

with  $\mathbf{u}$  the horizontal velocity and  $\mathcal{S}$  the horizontal region. For barotropic motion on a tangent plane we are interested in the vertical component of the relative angular momentum

$$L^z = \rho H \int_{\mathcal{S}} \hat{\mathbf{z}} \cdot (\mathbf{x} \times \mathbf{u}) d\mathcal{S}. \quad (4.210)$$

Show for a simply connected and bounded region,  $L^z$  can be written

$$L^z = 2 \rho H \int_{\mathcal{S}} (\psi_b - \psi) d\mathcal{S}, \quad (4.211)$$

where  $\psi$  is the streamfunction satisfying  $\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi$ , and  $\psi_b$  is the value of the streamfunction evaluated on the region boundary. Hint: note that  $\nabla \cdot \mathbf{x} = 2$  for a horizontal position vector. Also recall from our study of kinematics of non-divergent flows in VOLUME 1 that the streamfunction equals to a spatial constant when evaluated along a material boundary.

#### EXERCISE 4.9: STEADY AXIALLY SYMMETRIC FLOW

Consider a two-dimensional non-divergent velocity

$$\mathbf{v} = \hat{\mathbf{z}} \times \nabla \psi. \quad (4.212)$$

Assume the streamfunction is static and depends only on the radial distance from an arbitrary origin,

$$\psi = \psi(r), \quad (4.213)$$

where  $r = \sqrt{x^2 + y^2}$ , and assume the velocity is a solution to the steady inviscid non-divergent barotropic dynamics on an  $f$ -plane.

- (a) Show that the velocity only has an angular component

$$\mathbf{v} = v^\varphi \hat{\boldsymbol{\varphi}}, \quad (4.214)$$

where  $\hat{\boldsymbol{\varphi}}$  is the angular unit vector oriented counter-clockwise from the  $\hat{\mathbf{x}}$  axis.<sup>24</sup> Express  $v^\varphi$  in terms of the streamfunction  $\psi$ . Hint: recall our discussion in VOLUME 1 of polar coordinates.

- (b) Write the relative vorticity in terms of the streamfunction using polar coordinates.  
(c) Consider the circulation

$$\mathcal{C} = \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r}, \quad (4.215)$$

where  $\mathcal{S}$  is a circular region in the  $x$ - $y$  plane centered at  $r = 0$ . Express the circulation in terms of  $v^\varphi$  and the radius of the circle.

- (d) Write the pressure gradient acceleration in terms of  $v^\varphi$ ,  $f$ , and  $r$ . Hint: remember that  $\hat{\boldsymbol{\varphi}}$  is a function of the polar angle  $\varphi$ .  
(e) Interpret the steady balance of accelerations in terms of the balanced horizontal flow dynamics studied in VOLUME 2.  
(f) Why is this axial symmetric solution only valid for an  $f$ -plane? Hint: show that if  $\beta \neq 0$  that there is an inconsistency in the velocity equation.

**EXERCISE 4.10: GALILEAN TRANSFORMATION OF PV ADVECTION AND THE APV METHOD**  
In our study of kinematics in VOLUME 1, we established the invariance of the material time derivative operator under a Galilean transformation

$$\bar{\mathbf{x}} = \mathbf{x} + \mathbf{U} t \quad \text{and} \quad \bar{\mathbf{u}} = \mathbf{u} + \mathbf{U}, \quad (4.216)$$

where  $\mathbf{U}$  is a constant. Here we study the Galilean transformation properties of the non-divergent barotropic model on a  $\beta$ -plane.

- (a) Determine the Galilean transformation properties of the potential vorticity equation

$$\frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q = \frac{\partial q}{\partial t} + J(\psi, q) = 0, \quad (4.217)$$

where  $q = \zeta + f$ ,  $\hat{\mathbf{z}} \times \psi = \mathbf{u}$ , and  $J$  is the Jacobian operator.

- (b) Determine the Galilean transformation properties of the relative vorticity equation (4.39)

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \cdot \zeta = -\beta v. \quad (4.218)$$

Discuss why there is Galilean invariance only for zonal Galilean boosts,  $\mathbf{U} = \hat{\mathbf{x}} U$ .

- (c) An Euler forward time stepping scheme for the PV equation leads to

$$q^{n+1} = q^n - \Delta t \mathbf{u}^n \cdot \nabla q^n, \quad (4.219)$$

where  $\Delta t$  is the discrete time step and the integer  $n$  represents the discrete time label. Inspired by this time discrete expression, [Sadourny and Basdevant \(1985\)](#) proposed the

---

<sup>24</sup>The azimuthal angular coordinate,  $\varphi$ , that appears in  $\mathbf{v} = v^\varphi \hat{\boldsymbol{\varphi}}$ , is not the same as the density normalized pressure,  $\varphi = p/\rho$ , defined by equation (4.2) and used throughout this chapter.

anticipated potential vorticity method for parameterizing subgrid scale processes. The simplest form of APV is given by

$$\frac{\partial q}{\partial t} = -\mathbf{u} \cdot \nabla[q - \tau \mathbf{u} \cdot \nabla q] = -J[\psi, q - \tau J(\psi, q)], \quad (4.220)$$

with  $\tau$  a constant time scale. From the time discrete expression (4.219), we see that the anticipated potential vorticity method makes use of an estimate for the future value of potential vorticity in computing the advection operator, thus motivating the term “anticipated” in the method’s name.

Show that  $\tau \neq 0$  breaks Galilean invariance for the equation (4.220). Provide a discussion of why invariance is broken. Hint: [Vallis and Hua \(1988\)](#) offer a technical reason for why Galilean invariance is broken, making use of the streamfunction and Jacobian form. You do not necessarily need to follow their approach. Rather, it is sufficient to merely note how velocity appears in the anticipated potential vorticity operator.

#### EXERCISE 4.11: ELEMENTS OF THE FOFONOFF GYRE

A Fofonoff gyre is an unforced inviscid solution in a flat bottom bounded domain with a rigid lid. For a single layer of homogeneous fluid with constant thickness, the absolute vorticity is materially invariant,  $D(\zeta + f)/Dt = 0$ . An explicit solution is derived in Section 19.5.3 of [Vallis \(2017\)](#) for quasi-geostrophic flow using the method of matched asymptotics. We depict elements of a double Fofonoff gyre in Figure 4.12. Provide a narrative for this flow based on material conservation of absolute vorticity. In particular, discuss how the flow enters and leaves the side boundaries and conversely how it leaves and enters the interior region. We are not concerned with how this flow is established. Instead, assume the flow exists and discuss how its existence is consistent with  $D(\zeta + f)/Dt = 0$ . Hint: recall our discussion of Figure 4.3.

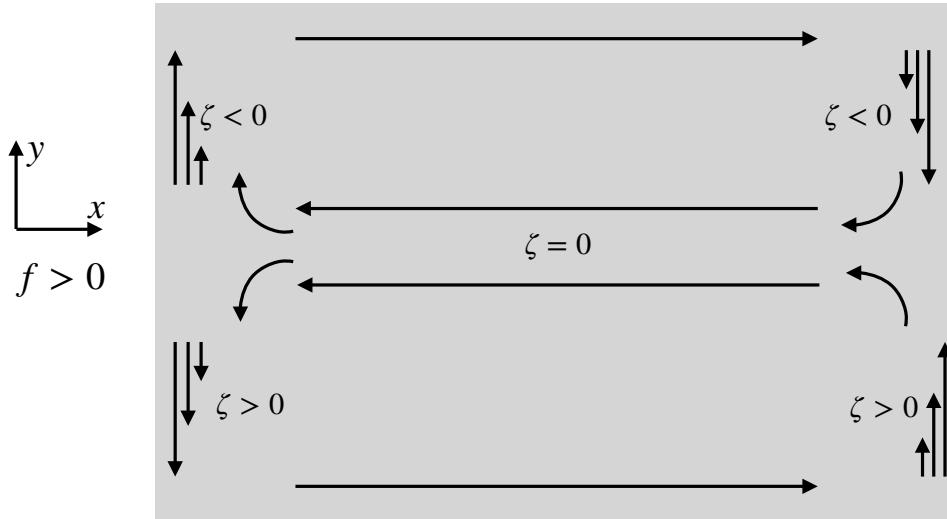


FIGURE 4.12: A Fofonoff gyre is an unforced inviscid flow in a bounded domain where  $D(\zeta + f)/Dt = 0$ . We here depict elements of this double-gyre flow in the northern hemisphere as part of Exercise 4.11.

#### EXERCISE 4.12: PRESSURE EQUATION WITH $w \neq 0$

Equation (4.73) or equation (4.76) provide equivalent expressions for the pressure Poisson equation with  $\nabla \eta = 0$  and, correspondingly, with  $w = 0$ . However, in Section 4.2.4 we considered the possibility of  $\nabla \eta = \nabla \eta_b$ , thus providing a solution with  $w \neq 0$ . In this case,

derive the Poisson equation for pressure as decomposed according to  $\varphi = g \eta + \varphi'$ .

EXERCISE 4.13: VECTOR-INVARIANT IN SPHERICAL COORDINATES

Prove, using spherical coordinates, the identity (4.158).





## Chapter 5

# SHALLOW WATER VORTICITY AND POTENTIAL VORTICITY

In this chapter we study vorticity and potential vorticity within the shallow water system. We start by deriving the evolution equation for vorticity by taking the curl of the velocity equation. Combining vorticity evolution with mass continuity then renders the evolution equation for potential vorticity. Potential vorticity is a material invariant for inviscid shallow water motion, thus providing a mechanical constraint on the fluid flow. After developing the basic concepts and equations, we consider a variety of flow regimes and case studies, mostly with an ocean focus. These case studies illustrate where the study of vorticity, potential vorticity, and circulation enhances our understanding of geophysical fluid mechanics.

### CHAPTER GUIDE

The shallow water fluid offers a fruitful conceptual model to introduce the dynamics of vorticity and potential vorticity while requiring a relatively modest level of mathematical sophistication that makes use of vector calculus identities from VOLUME 1. We also require an understanding of shallow water mechanics from Chapters 1 and 2, as well as the vorticity kinematics introduced in Chapter 3. The concepts and methods developed in this chapter are fundamental to the remaining vorticity chapters in this part of the book.

As anticipated in Section 4.3.3, the form of potential vorticity encountered here is sometimes referred to as [shallow water potential vorticity](#) (also the [Rossby potential vorticity](#)). Its connection to the more general [Ertel potential vorticity](#) (Chapter 7) is postponed until VOLUME 4, where we study the [Boussinesq ocean](#) equations using isopycnal vertical coordinates. As we see there, the shallow water equations provide a discrete representation of the isopycnal equations. Correspondingly, the Ertel potential vorticity expressed using isopycnal coordinates has its discrete form given by shallow water potential vorticity studied in the present chapter.

<b>5.1 Shallow water vorticity equation</b>	<b>164</b>
5.1.1 Vorticity equation for a single layer	165
5.1.2 Vorticity equation for $N$ -layers	166
5.1.3 Vorticity flux divergence and curl of nonlinear advection	166
<b>5.2 Potential vorticity for a rotating cylinder</b>	<b>167</b>
5.2.1 Mass conservation	167
5.2.2 Angular momentum conservation	167
5.2.3 Material invariance of potential vorticity	168
5.2.4 Connecting angular momentum and vorticity	169

5.2.5	Comments and further study . . . . .	169
<b>5.3</b>	<b>Shallow water (Rossby) potential vorticity . . . . .</b>	<b>169</b>
5.3.1	Mass conservation plus the vorticity equation . . . . .	170
5.3.2	Motivating the name . . . . .	171
5.3.3	Mass conservation + Kelvin's circulation theorem . . . . .	171
5.3.4	A fluid column with constant $f \neq 0$ . . . . .	173
5.3.5	Material invariance of an arbitrary function of PV . . . . .	174
5.3.6	$N$ -layer potential vorticity . . . . .	174
5.3.7	Further study . . . . .	175
<b>5.4</b>	<b>Potential vorticity with non-conservative processes . . . . .</b>	<b>175</b>
5.4.1	Material time evolution of potential vorticity . . . . .	175
5.4.2	The potential vorticity flux . . . . .	175
<b>5.5</b>	<b>Example implications of material PV invariance . . . . .</b>	<b>177</b>
5.5.1	Topographic beta effect . . . . .	177
5.5.2	Planetary geostrophic potential vorticity and $f/H$ contours . . . . .	178
5.5.3	Spin up of converging flow . . . . .	180
5.5.4	Further study . . . . .	180
<b>5.6</b>	<b>Circulation with non-conservative processes . . . . .</b>	<b>181</b>
5.6.1	Circulation around a closed streamline in steady flow . . . . .	182
5.6.2	Circulation from wind stress and Rayleigh drag . . . . .	182
<b>5.7</b>	<b>A primer on steady ocean gyres . . . . .</b>	<b>183</b>
5.7.1	Steady and large-scale vorticity balance . . . . .	183
5.7.2	Planetary geostrophic flow and the Sverdrup balance . . . . .	184
5.7.3	Taylor-Proudman, Sverdrup balance, and $f/h$ invariance . . . . .	185
5.7.4	Sverdrup flow in a closed domain with anti-cyclonic wind stress . . . . .	186
5.7.5	Western intensification and the role of beta . . . . .	187
5.7.6	A role for bottom pressure torques . . . . .	189
5.7.7	Properties of area integrated bottom pressure torques . . . . .	191
5.7.8	Advection-diffusion of the steady streamfunction . . . . .	192
5.7.9	Comments and further study . . . . .	193
<b>5.8</b>	<b>Column vorticity . . . . .</b>	<b>194</b>
5.8.1	Formulating the column vorticity equation . . . . .	194
5.8.2	Summary of the column vorticity equation . . . . .	196
5.8.3	Steady linear column vorticity balance and the island rule . . . . .	196
5.8.4	Further study . . . . .	199
<b>5.9</b>	<b>Free surface patterns in steady ocean gyres . . . . .</b>	<b>199</b>
5.9.1	Formulating the free surface equation . . . . .	199
5.9.2	Advecting the free surface . . . . .	199
5.9.3	Rayleigh drag and free surface diffusion . . . . .	201
5.9.4	Further study . . . . .	201
<b>5.10</b>	<b>Exercises . . . . .</b>	<b>202</b>

## 5.1 Shallow water vorticity equation

In this section we formulate the vorticity equation for the shallow water fluid, starting with a single layer and then extending to multiple layers. We sometimes make use of the vertical component to the absolute vorticity from equation (1.108)

$$\boldsymbol{\omega}^* = (\zeta + f) \hat{z} = \zeta_a \hat{z}, \quad (5.1)$$

which is the sum of the relative vorticity of the horizontal flow,  $\boldsymbol{\omega}^* = \zeta \hat{z}$ , plus the rigid-body vorticity,  $f \hat{z}$ , due to motion of the rotating reference frame (recall Section 3.6.1). The absolute

vorticity appears in the vector-invariant velocity equation (1.113), which is valid for each of the layers in a shallow water fluid

$$\partial_t \mathbf{u} + \boldsymbol{\omega}_a^* \times \mathbf{u} = -\nabla(p/\rho_{ref} + \mathbf{u} \cdot \mathbf{u}/2). \quad (5.2)$$

This equation forms the starting point for deriving the shallow water vorticity equation.

### 5.1.1 Vorticity equation for a single layer

We make use of the following vector identity from VOLUME 1 to express the curl of the Magnus acceleration plus Coriolis acceleration in the form

$$\nabla \times (\boldsymbol{\omega}_a^* \times \mathbf{u}) = \boldsymbol{\omega}_a^* (\nabla \cdot \mathbf{u}) + (\mathbf{u} \cdot \nabla) \boldsymbol{\omega}_a^* - \mathbf{u} (\nabla \cdot \boldsymbol{\omega}_a^*) - (\boldsymbol{\omega}_a^* \cdot \nabla) \mathbf{u} \quad (5.3a)$$

$$= \boldsymbol{\omega}_a^* (\nabla \cdot \mathbf{u}) + (\mathbf{u} \cdot \nabla) \boldsymbol{\omega}_a^*, \quad (5.3b)$$

so that

$$\hat{\mathbf{z}} \cdot [\nabla \times (\boldsymbol{\omega}_a^* \times \mathbf{u})] = \nabla \cdot (\mathbf{u} \zeta_a) \quad (5.4)$$

Equation (5.3b) required setting

$$\nabla \cdot \boldsymbol{\omega}_a^* = \nabla \cdot \boldsymbol{\omega}^* + \nabla \cdot (f \hat{\mathbf{z}}) = 0, \quad (5.5)$$

which follows since this expression involves the divergence of a curl (first right hand side term) and since  $f$  has no  $z$  dependence. We furthermore set

$$(\boldsymbol{\omega}_a^* \cdot \nabla) \mathbf{u} = \zeta_a \partial_z \mathbf{u} = 0, \quad (5.6)$$

which follows since the horizontal velocity in a shallow water fluid is depth independent within a layer (see Section 1.2).

Applying the curl operator,  $\hat{\mathbf{z}} \cdot (\nabla \times)$ , onto the vector-invariant velocity equation (5.2) annihilates the gradient of pressure and kinetic energy, with the identity (5.4) leading to the flux-form evolution equation for absolute vorticity

$$\partial_t \zeta_a + \nabla \cdot (\mathbf{u} \zeta_a) = 0. \quad (5.7)$$

This equation says that the vertical component to the absolute vorticity,  $\zeta_a$ , at a point in the inviscid shallow water fluid changes according to the horizontal convergence of vorticity advected to that point

$$\partial_t \zeta_a = -\nabla \cdot (\mathbf{u} \zeta_a). \quad (5.8)$$

We can write the vorticity equation (5.7) in the material form

$$\frac{D \zeta_a}{Dt} = -\zeta_a \nabla \cdot \mathbf{u}, \quad (5.9)$$

where the material time derivative for the shallow water fluid includes advection just by the horizontal flow

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla = \frac{\partial}{\partial t} + u \partial_x + v \partial_y. \quad (5.10)$$

The material evolution equation (5.9) means that the absolute vorticity of a shallow water fluid column, moving with the horizontal flow, changes according to the horizontal convergence of the fluid flow as multiplied by the absolute vorticity. For comparison, recall the horizontally non-

divergent barotropic fluid has  $\nabla \cdot \mathbf{u} = 0$ , so that the absolute vorticity in that flow is materially invariant (see equation (4.40)). In contrast, the shallow water fluid supports horizontal flow convergence, and with the flow convergence providing a source to the vorticity.

### 5.1.2 Vorticity equation for $N$ -layers

The previous results for a single layer are readily extended to  $N$ -layers, simply because the velocity for layer- $k$  evolves according to equation (5.2), now with a subscript  $k$  to denote the layer

$$\partial_t \mathbf{u}_k + (f + \zeta_k) \hat{\mathbf{z}} \times \mathbf{u}_k = -\nabla(p_k/\rho_{\text{ref}} + \mathbf{u}_k \cdot \mathbf{u}_k/2), \quad (5.11)$$

where  $\zeta_k = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}_k)$  is the vertical component to the layer- $k$  relative vorticity. Taking the curl and making use of the mathematical identities used for single layer in Section 5.1.1 renders the vorticity equation for layer- $k$

$$\partial_t \zeta_{ak} + \nabla \cdot (\mathbf{u}_k \zeta_{ak}) = 0 \iff \frac{D_k \zeta_{ak}}{Dt} = -\zeta_{ak} \nabla \cdot \mathbf{u}_k \quad (5.12)$$

where

$$\zeta_{ak} = f + \zeta_k \quad (5.13)$$

is the vertical component to the absolute vorticity of layer- $k$ . Hence, the vorticity equation for an arbitrary layer in a stacked shallow water model is the same as that for a single shallow water layer.

### 5.1.3 Vorticity flux divergence and curl of nonlinear advection

We revisit the manipulations from Section 5.1.1 to explicitly identify a connection between the nonlinear terms in the vorticity equation. Start by writing the velocity equation in the advective form and the vector invariant form

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta \quad (5.14a)$$

$$\partial_t \mathbf{u} + (f + \zeta) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla(g \eta + \mathbf{u} \cdot \mathbf{u}/2). \quad (5.14b)$$

Taking their curl yields two expressions of the vorticity equation

$$\partial_t \boldsymbol{\omega}^* + \nabla \times [(\mathbf{u} \cdot \nabla) \mathbf{u}] + \nabla \times [f \hat{\mathbf{z}} \times \mathbf{u}] = 0 \quad (5.15a)$$

$$\frac{\partial \boldsymbol{\omega}^*}{\partial t} + \nabla \times [(f + \zeta) \hat{\mathbf{z}} \times \mathbf{u}] = 0, \quad (5.15b)$$

whose equality leads to

$$\nabla \times [(\mathbf{u} \cdot \nabla) \mathbf{u} - \zeta \hat{\mathbf{z}} \times \mathbf{u}] = 0. \quad (5.16)$$

Making use of the identity

$$\hat{\mathbf{z}} \cdot [\nabla \times (\zeta \hat{\mathbf{z}} \times \mathbf{u})] = \nabla \cdot (\mathbf{u} \zeta) \quad (5.17)$$

renders the relation

$$\hat{\mathbf{z}} \cdot \nabla \times [(\mathbf{u} \cdot \nabla) \mathbf{u}] = \nabla \cdot (\mathbf{u} \zeta). \quad (5.18)$$

We thus see that the divergence of the advective vorticity flux (right hand side) equals to the curl of the nonlinear advection (left hand side). This identity holds for each layer in an  $N$ -layer shallow water model.

## 5.2 Potential vorticity for a rotating cylinder

To introduce the concept of shallow water potential vorticity, consider a fluid cylinder of constant mass  $M$ , constant density  $\rho$ , variable radius  $R$ , and variable height  $h$ , and assume the cylinder rotates about its central axis. Furthermore, assume the fluid particles within the cylinder rotate as a rigid-body, meaning there are no strains in the fluid, and yet allow the cylinder radius and height to change. This analysis offers a useful (albeit incomplete) conceptual picture for a coherently rotating column of a shallow water layer, in which time derivatives in the following are interpreted as material derivatives.

### 5.2.1 Mass conservation

With a constant density, mass conservation for the material cylinder means that its volume is fixed. Hence, mass conservation constrains the relative changes to the radius and height of the cylinder. Namely, a materially constant cylinder mass

$$M = \pi R^2 h \rho \quad (5.19)$$

implies

$$\frac{2}{R} \frac{DR}{Dt} = -\frac{1}{h} \frac{Dh}{Dt}. \quad (5.20)$$

That is, mass conservation means that the relative height decreases as twice the relative radius increases. So if the cylinder is squashed ( $h$  decreases) then it thickens ( $R$  increases). Conversely, the cylinder thins ( $R$  decreases) as it extends ( $h$  increases).

### 5.2.2 Angular momentum conservation

A second constraint arises from angular momentum conservation. Choose the center of mass coordinate axes through the center of the cylinder, with the  $z$ -axis along the central line of the cylinder and with  $z = 0$  at the cylinder mid-point. The angular rotation vector is thus given by

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}. \quad (5.21)$$

With this axis orientation, rotation occurs about the center of mass so that the angular momentum of the center of mass vanishes. The moment of inertia tensor for a cylinder with this axis orientation is given<sup>1</sup>

$$I_{mn} = \delta_{mn} \frac{MR^2}{2}. \quad (5.22)$$

The moment of inertia is a measure of the rotational inertia of a moving continuous body. For the cylinder it is directly related to the cylinder mass (assumed fixed here) and the radius (which can change). Notably, the moment of inertia about the central vertical axis is not a function of the cylinder height. The reason is that the moment measures the inertia relative to the rotational axis, which is here along the central vertical axis. The angular momentum for the cylinder is thus given by

$$\mathbf{L} = \frac{MR^2}{2} \Omega \hat{\mathbf{z}}. \quad (5.23)$$

---

<sup>1</sup>See [Marion and Thornton \(1988\)](#) or other classical mechanics texts for a discussion of the moment of inertia for a variety of bodies.

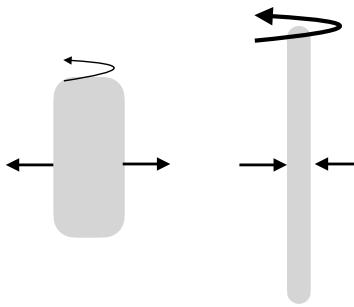


FIGURE 5.1: Illustrating the conservation of angular momentum for a spinning constant mass cylinder of shallow water fluid rotating around its central axis. Fluid particles within the cylinder are assumed to rotate as a rigid-body, meaning there are no strains in the fluid, and yet the cylinder radius and height are allowed to change. The moment of inertia (relative to the central axis) for the left cylinder is larger since more of its mass is distributed away from the central axis than in the right configuration. If the initial spin for the left cylinder is counterclockwise, then the two configurations have identical angular momentum if the right cylinder spins more rapidly than the left, since the moment of inertia for the right cylinder is smaller. This example exemplifies the familiar ice skater experience, whereby the skater's spin rate increases when bringing arms (mass) inward towards the central axis of the body (depicted by the inward arrows on the right panel), whereas the skater's rotation slows when extending arms outward (depicted by the outward arrows on the left panel).

The familiar *ice skater* example occurs when the cylinder radius changes and thus changes the moment of inertia (e.g., the ice skater's arms are brought in toward the central axis of the body or out away from the body). Maintaining constant angular momentum and constant mass means that the angular velocity,  $\Omega$ , increases in magnitude (rotates faster) when the cylinder radius decreases, and vice versa. Explicitly for the cylinder we have  $d\mathbf{L}/dt = 0$  and  $dM/dt = 0$  thus rendering

$$\frac{2}{R} \frac{DR}{Dt} = - \frac{1}{\Omega} \frac{D\Omega}{Dt}. \quad (5.24)$$

We see that reducing the moment of inertia for a constant mass body by bringing its mass distribution towards the central axis (converging mass) leads, through angular momentum conservation, to an increase in rotation speed. The opposite occurs when mass diverges from a region, thus reducing the rotation speed.

Although the angular momentum constraint means that the spin rate changes when changing the moment of inertia, it does not impose a preferred direction. For example, no matter what direction a skater is rotating, decreasing the moment of inertia increases the spin rate in that particular direction. Yet when placing the spinning column on a rotating planet, the planetary rotation breaks the symmetry and thus prescribes the direction for the spin changes. The reason is that planetary rotation contributes to the spin of the column, even if the column has no spin relative to the rotating planetary reference frame. We encounter this additional part of the story in Section 5.3.

### 5.2.3 Material invariance of potential vorticity

Combining angular momentum conservation (5.24) with mass conservation (5.20) leads to the material conservation law

$$\frac{D(\Omega/h)}{Dt} = 0. \quad (5.25)$$

Equation (5.25) means that the potential vorticity is constant for a material fluid column, with potential vorticity for the cylinder given by

$$Q \equiv \Omega/h. \quad (5.26)$$

For example, if the column thickens then the rotational velocity increases in order to maintain  $Q = \Omega/h$  constant. Equivalently, if the column cross-sectional area decreases, the column thickness increases according to volume conservation, which in turn results in an increase in the spin according to angular momentum conservation.

#### 5.2.4 Connecting angular momentum and vorticity

When allowing the shallow water fluid to exhibit motion that is more general than a rigid-body cylinder rotation (i.e., when allowing for strains in the fluid), then the angular rotation rate appearing in the potential vorticity (5.26) is generalized to the absolute vorticity, and we consider this generalization in Section 5.3. Furthermore, as shown in Section 3.6, the vorticity equals to twice the rotation rate,  $2\Omega$ . Hence, the numerator for the potential vorticity of the rigid-body rotating cylinder equals to one-half the vorticity.

#### 5.2.5 Comments and further study

The discussion in this section is motivated by Section 2.4 of [Salmon \(1998\)](#). The rotating cylinder succinctly identifies the two mechanical properties contributing to the potential vorticity conservation law (5.25): a kinematic property (mass conservation) and a dynamic property (angular momentum conservation). For the rotating cylinder, the implications of potential vorticity conservation are well gleaned from the separate mass and angular momentum conservation principles. Hence, potential vorticity conservation lends little novel insight for the cylinder. However, the material invariance of potential vorticity is of fundamental use for studies of rotating and stratified fluids where the flow generally has strains that make vorticity distinct from angular momentum (Section 3.9).

Another important element missing from this discussion is the beta effect ( $\beta$ -effect), which accounts for the changes in planetary vorticity when moving on a rotating spherical planet. We encounter this effect in the following sections.

### 5.3 Shallow water (Rossby) potential vorticity

We now consider the potential vorticity for a single layer of shallow water fluid. The form of the shallow water potential vorticity is sometimes referred to as the Rossby potential vorticity. The derivation here makes use of fluid mechanical equations rather than those from rigid-body mechanics, thus allowing for the added feature of strains in the fluid that distinguish vorticity from angular momentum. We present two derivations: one based on manipulations of the mass and momentum equations, and one based on the small aspect ratio limit of Kelvin's circulation theorem, with Kelvin's theorem more thoroughly studied in Chapter 6.

Figure 5.2 summarizes key elements leading to potential vorticity conservation for a shallow water fluid layer. Namely, as shown in this section, shallow water potential vorticity conservation arises from combining the kinematic constraint of mass conservation (material invariance of  $h A$ ) with either the vorticity equation or Kelvin's circulation theorem for a small aspect ratio fluid.

In the absence of non-conservative processes, we show that shallow water potential vorticity for a material column of fluid remains constant. We refer to this conserved quantity as a material invariant, since it remains invariant (constant) when following material fluid columns.

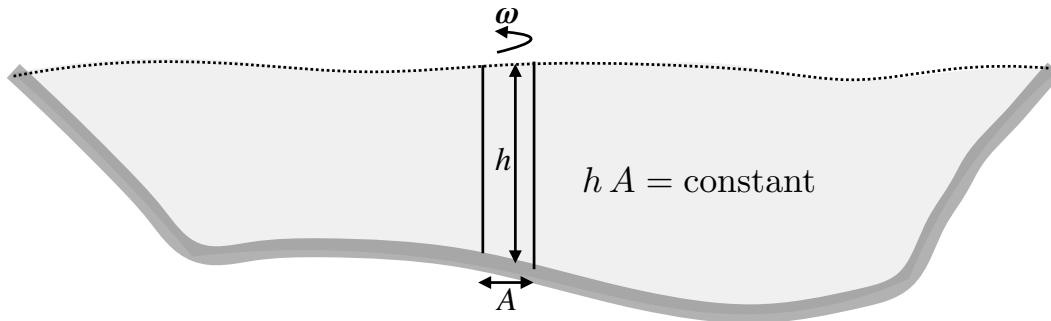


FIGURE 5.2: Illustrating the material invariance of potential vorticity for a layer of shallow water fluid column. Material invariance results from merging mass conservation (material invariance of the column volume,  $h A$ ), to either the vorticity equation or Kelvin's circulation theorem for a small aspect ratio fluid (material invariance of  $\zeta A$ ).

### 5.3.1 Mass conservation plus the vorticity equation

To derive the potential vorticity equation, we here make use of the vorticity equation (5.9) and combine it with mass conservation.

#### Shallow water vorticity and vortex stretching

Mass conservation in the form of the material thickness equation (1.20) leads to the following expression for the divergence of the horizontal velocity

$$\nabla \cdot \mathbf{u} = -\frac{1}{h} \frac{Dh}{Dt}. \quad (5.27)$$

Making use of this result in the vorticity equation (5.9) allows us to eliminate the horizontal convergence

$$\frac{D\zeta_a}{Dt} = -\zeta_a \nabla \cdot \mathbf{u} = \frac{\zeta_a}{h} \frac{Dh}{Dt}. \quad (5.28)$$

This equation says that material changes in shallow water absolute vorticity arise only from material changes in the layer thickness; i.e., absolute vorticity increases in magnitude if the column stretches and decreases if the column compresses. We refer to this process as **vortex stretching**.

We see in Section 6.5.3 that vorticity in continuously stratified fluids is affected by vortex stretching and **vortex tilting**, as well as torques from baroclinicity. In contrast, equation (5.28) says that the material evolution of absolute vorticity for a shallow water fluid is affected only through vortex stretching. This behavior is a result of the depth independence of the horizontal velocity within a shallow water layer and the associated vertical columnar motion of fluid within the layer. On a related note, we see in Section 5.3.3 that vortex tubes in a shallow water layer are nearly vertical, so that we are only concerned with the vertical component of shallow water vorticity. Correspondingly, shallow water vortex tubes never close.

### Material invariance of shallow water potential vorticity

Equation (5.28) can be written as an expression of the material invariance of the shallow water potential vorticity

$$\frac{DQ}{Dt} = 0, \quad (5.29)$$

where

$$Q = \frac{\zeta_a}{h} = \frac{\zeta + f}{h} \quad (5.30)$$

is the shallow water potential vorticity. As defined, shallow water potential vorticity is the ratio of absolute vorticity to the thickness of the fluid layer. The material conservation law (5.29) says that this ratio remains constant for the shallow water layer in the absence of non-conservative processes such as friction.

#### 5.3.2 Motivating the name

Material invariance of the shallow water potential vorticity in equation (5.30) is most practically a statement about how the relative vorticity,  $\zeta$ , changes when changing column thickness or latitude. That is, by maintaining  $Q$  fixed,  $\zeta$  must change when either the column thickness,  $h$ , changes or when the column moves meridionally and thus alters the planetary vorticity,  $f$ . By focusing on relative vorticity we are offered insights into how the fluid motion is constrained and thus a means to predict changes in that motion. In turn, these changes in  $\zeta$  motivate the name “potential vorticity” as we now see.

Potential vorticity measures the ability for a shallow water fluid column to either spin up or spin down (change its relative vorticity) relative a standard configuration. For example, let the standard configuration be defined by an arbitrary standard thickness,  $h_s$ , at the equator (where  $f = 0$ ). Now move an off-equatorial shallow water fluid column with zero relative vorticity to the equator and stretch/compress the column to the standard thickness. Material invariance of the column’s potential vorticity allows us to deduce the column’s relative vorticity at the equator, given information about the initial column thickness and initial Coriolis parameter (see Figure 5.3). Hence, potential vorticity, as an invariant material property, provides the “potential” for a fluid column to manifest a particular value of the relative vorticity when moved and stretched into a standard configuration. In this manner, the use of “potential” in “potential vorticity” is directly analogous to the use of “potential” in potential temperature or gravitational potential energy.

#### 5.3.3 Mass conservation + Kelvin’s circulation theorem

Although we have yet to discuss Kelvin’s theorem (Section 6.2), we here invoke it to illustrate another way to derive the material invariance of shallow water potential vorticity. As we see, this derivation provides a direct analog to the rotating cylinder discussed in Section 5.2.

When applied to an infinitesimal circuit in an inviscid and constant density fluid, Kelvin’s theorem says that

$$\frac{D(\omega_a \cdot \hat{n} \delta S)}{Dt} = 0, \quad (5.31)$$

where  $\omega_a$  is the absolute vorticity

$$\omega_a = \omega + f \hat{z}, \quad (5.32)$$

$\hat{n} \delta S$  is the infinitesimal surface area enclosed by the closed circuit, with  $\hat{n}$  the unit outward normal to the surface. Hence, equation (5.31) says that the projection of the absolute vorticity

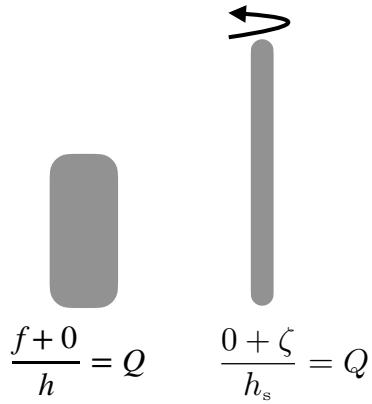


FIGURE 5.3: Left panel: an arbitrary shallow water column with zero relative vorticity and potential vorticity  $Q = f/h$ , with  $f > 0$  assumed for this figure (northern hemisphere). Right panel: the same fluid column moved to the equator (where  $f = 0$ ) and stretched to have the standard thickness,  $h_s > h$ . The relative vorticity of the column at the equator is given by  $\zeta = f(h_s/h)$ , where  $f$  is the Coriolis parameter at the original latitude where  $f > 0$ . Potential vorticity thus provides a means to deduce the relative vorticity that can be realized by moving any particular configuration to a standard location and with a standard thickness. This property motivates the “potential” used in the name.

onto the local normal of an area element, multiplied by that area element, remains materially constant. This identity offers a very strong constraint on the flow.

To make use of equation (5.31) for the shallow water layer, decompose absolute vorticity into

$$\boldsymbol{\omega}_a = \hat{\mathbf{z}}(\zeta + f) + \boldsymbol{\omega}_h, \quad (5.33)$$

where

$$\boldsymbol{\omega}_h = -\hat{\mathbf{z}} \times \nabla w = \hat{\mathbf{x}} \partial_y w - \hat{\mathbf{y}} \partial_x w \quad (5.34)$$

is the horizontal component to the shallow water relative vorticity from equation (1.106d) (recall the expression for  $\boldsymbol{\omega}_h$  follows since the horizontal velocity components have no vertical dependence within a shallow water layer:  $\partial u / \partial z = \partial v / \partial z = 0$ ). Inserting the absolute vorticity (5.33) into Kelvin’s theorem (5.31) leads to

$$\frac{D}{Dt} [(\zeta + f) \delta A + \boldsymbol{\omega}_h \cdot \hat{\mathbf{n}} \delta \mathcal{S}] = 0, \quad (5.35)$$

where the horizontal area element,  $\delta A$ , is the projection of the surface area element onto the vertical direction

$$\delta A = \hat{\mathbf{z}} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (5.36)$$

Shallow water fluid mechanics arises from considering a constant density fluid layer whose flow respects the small aspect ratio limit:  $H/L \ll 1$ , with  $H$  the vertical length scale of the flow, and  $L$  the horizontal length scale of the flow. Under this limit, the second term in equation (5.35) is much smaller than the first. It is further reduced in size since  $\hat{\mathbf{n}}$  is nearly vertical, so that  $\hat{\mathbf{n}} \cdot \hat{\mathbf{x}} \approx 0$  and  $\hat{\mathbf{n}} \cdot \hat{\mathbf{y}} \approx 0$ , in which case we are led to the scaling

$$\frac{|\boldsymbol{\omega}_h \cdot \hat{\mathbf{n}} \delta \mathcal{S}|}{|(\zeta + f) \delta A|} \ll 1. \quad (5.37)$$

This result is consistent with our earlier comment in Section 5.3.1 that shallow water vortex tubes never close. Rather, they are nearly vertical, running from the bottom of a shallow water

layer to the top. With the scaling (5.37), we find

$$\frac{D}{Dt} \left[ \left( \frac{\zeta + f}{h} \right) h \delta A \right] = 0, \quad (5.38)$$

where  $h$  is the layer thickness and  $h \delta A$  is the volume of a fluid column extending through the shallow water layer. Given the incompressible nature of the fluid in a shallow water layer, the column volume is materially constant

$$\frac{D(h \delta A)}{Dt} = 0, \quad (5.39)$$

so that equation (5.38) yields material invariance of shallow water potential vorticity

$$\frac{D}{Dt} \left[ \frac{\zeta + f}{h} \right] = \frac{DQ}{Dt} = 0, \quad (5.40)$$

where  $Q = (\zeta + f)/h$  is the same shallow water potential vorticity derived above in Section 5.3.1.

#### 5.3.4 A fluid column with constant $f \neq 0$

Some of the essential features of shallow water potential vorticity material conservation are depicted in Figure 5.4 for the case of a fluid column with constant  $f \neq 0$ . In the left panel, the column thickness increases (column is stretched). Volume conservation for the column means that the column radius decreases. As the material in the column moves radially inward toward the center, it experiences a Coriolis deflection to the right in the northern hemisphere and to the left in the southern. Both of these deflections renders a cyclonic tendency to the relative vorticity, creating a positive relative vorticity tendency in the northern hemisphere and negative relative vorticity tendency in the southern hemisphere.

An equivalent way to understand the cyclonic tendency is to consider the angular momentum of the fluid column, assuming the column moves coherently as a rigid-body as in Section 5.2.<sup>2</sup> As the radius of the column decreases so too does its moment of inertia. Angular momentum conservation means that the column picks up a tendency that causes its spin to increase. The direction of this spin increase accords with the background  $f$  of the environment. Returning to the skater analog in Figure 5.1, we consider  $f$  to be the initial spin of the skater so that when the moment of inertia decreases the column picks up a spin in the same direction as  $f$ ; i.e., a cyclonic tendency.<sup>3</sup>

In both hemispheres, the Coriolis deflection, or equivalently angular momentum conservation, creates a cyclonic relative vorticity tendency when the column stretches, thus maintaining  $Q = (f + \zeta)/h$  fixed. Again, the cyclonic relative vorticity adds to the magnitude of the planetary vorticity to ensure that the absolute vorticity magnitude increases in accord with the column thickness increase, thus keeping  $Q = \zeta_s/h$  constant. The converse holds when the column is flattened/squashed, whereby the relative vorticity picks up an anti-cyclonic tendency (negative relative vorticity tendency in the northern hemisphere and positive relative vorticity tendency in the southern). Doing so reduces the magnitude of the absolute vorticity in accord

---

<sup>2</sup>Recall that in Section 5.2, the column rotates as a rigid-body, meaning there is no strain in the fluid, and yet the radius and thickness are allowed to change.

<sup>3</sup>The connection between the Coriolis acceleration and angular momentum conservation is considered for particle mechanics in VOLUME 1.

with the reduced column thickness.

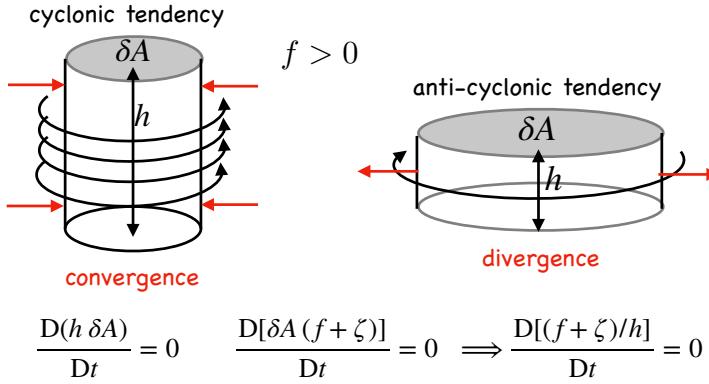


FIGURE 5.4: Material invariance for shallow water potential vorticity results from combining material invariance of the volume of a coherent fluid column with the material invariance of the area weighted absolute vorticity. As the cross-sectional area of the column decreases, as in a converging flow, the thickness of the fluid must increase in order to maintain constant volume. Furthermore, potential vorticity material invariance can be maintained by changing the fluid spin, as measured by the relative vorticity, or by changing the latitude and thus changing its planetary vorticity (the  $\beta$ -effect discussed in Section 6.6). We here depict the case with  $f > 0$  constant (so that  $\beta = 0$ ). Flow converging toward the center of the column picks up a Coriolis acceleration that creates a cyclonic tendency in the relative vorticity, just like the spinning cylinder in Figure 5.1. Equivalently, as the radius of the column decreases so too does its moment of inertia so that the column must pick up a cyclonic tendency to conserve angular momentum. Conversely, as the cross-sectional area increases, the diverging flow creates an anti-cyclonic tendency in the relative vorticity.

### 5.3.5 Material invariance of an arbitrary function of PV

The material invariance of shallow water potential vorticity, equation (5.29), means that any function,  $F(Q)$  is also materially constant. We see this property through the chain rule

$$\frac{DF}{Dt} = \frac{dF}{dQ} \frac{DQ}{Dt} = 0. \quad (5.41)$$

Since  $F$  is arbitrary, there are an infinite number of material invariants corresponding to distinct functions  $F$ . This result holds for all materially invariant scalar properties of the fluid.

### 5.3.6 $N$ -layer potential vorticity

The thickness equation (1.79a) and the vorticity equation (5.12) for an  $N$ -layer shallow water model are given by

$$\frac{D_k h_k}{Dt} = -h_k \nabla \cdot \mathbf{u}_k \quad \text{and} \quad \frac{D_k \zeta_{ak}}{Dt} = -\zeta_{ak} \nabla \cdot \mathbf{u}_k, \quad (5.42)$$

where there is no implied summation over the layer index  $k$ . These forms are isomorphic to the single layer equations so that the potential vorticity of layer- $k$  is given by

$$Q_k = \frac{f + \zeta_k}{h_k}, \quad (5.43)$$

and for a perfect shallow water fluid this layer potential vorticity is materially constant

$$\frac{D_k Q_k}{Dt} = \frac{\partial Q_k}{\partial t} + \mathbf{u}_k \cdot \nabla Q_k = 0, \quad (5.44)$$

where, again, there is no implied summation over  $\kappa$ .

### 5.3.7 Further study

The shallow water potential vorticity (5.30) was introduced by [Rossby \(1940\)](#) and as such it is sometimes referred to as the [Rossby potential vorticity](#). Non-rotating shallow water potential vorticity,  $\zeta/h$ , is illustrated in this [video from Prof. Shapiro at around the 11 minute mark](#). Note that he does not use the term “potential vorticity”, instead invoking mass conservation and angular momentum conservation to describe the motion.

## 5.4 Potential vorticity with non-conservative processes

In this section we consider the role of a non-conservative acceleration,  $\mathbf{F}$ , with this term arising from friction and boundary stresses (Section 1.6.5). Additionally, we allow for the presence of a boundary volume source,  $w^{(\dot{\eta})}$  (as in precipitation minus evaporation), thus changing the volume in the layer. We introduced such processes in Section 1.6, in which case the shallow water equations take on the form

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta + \mathbf{F} \quad \text{and} \quad \frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} = w^{(\dot{\eta})}. \quad (5.45)$$

### 5.4.1 Material time evolution of potential vorticity

In the presence of non-conservative forces, the absolute vorticity equation (5.7) becomes

$$\partial_t \zeta_a + \nabla \cdot (\mathbf{u} \zeta_a) = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) \implies \frac{D\zeta_a}{Dt} + \zeta_a \nabla \cdot \mathbf{u} = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}), \quad (5.46)$$

so that vorticity is now affected by the curl of  $\mathbf{F}$ . As before, we make use of the thickness equation to replace the horizontal divergence according to

$$\nabla \cdot \mathbf{u} = \frac{1}{h} \left[ -\frac{Dh}{Dt} + w^{(\dot{\eta})} \right]. \quad (5.47)$$

The presence of  $w^{(\dot{\eta})}$  modifies the divergence of the horizontal velocity beyond that for a conservative fluid. We are thus led to the potential vorticity equation

$$h \frac{DQ}{Dt} = -Q w^{(\dot{\eta})} + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}). \quad (5.48)$$

Hence, with  $w^{(\dot{\eta})} \neq 0$  and/or  $\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) \neq 0$ , shallow water potential vorticity is no longer materially invariant.

### 5.4.2 The potential vorticity flux

#### Deriving the flux-form equation

We can convert the potential vorticity equation (5.48) into a flux-form conservation equation by making use of the thickness equation

$$h \frac{DQ}{Dt} = h \left[ \frac{\partial Q}{\partial t} + \mathbf{u} \cdot \nabla Q \right] + Q \left[ \frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) - w^{(\dot{\eta})} \right] = \frac{\partial(hQ)}{\partial t} + \nabla \cdot (\mathbf{u} h Q) - Q w^{(\dot{\eta})}, \quad (5.49)$$

thus rendering

$$\partial_t(h Q) + \nabla \cdot (h \mathbf{u} Q) = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}). \quad (5.50)$$

As a final step, make use of the identity

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) = -\nabla \cdot (\hat{\mathbf{z}} \times \mathbf{F}), \quad (5.51)$$

so that the thickness weighted potential vorticity equation (5.50) can be written

$$\partial_t(h Q) = -\nabla \cdot (h Q \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}). \quad (5.52)$$

Note how the volume source term,  $w^{(\dot{\eta})}$ , does not explicitly appear in the flux-form equation (5.52) since the effects from  $w^{(\dot{\eta})}$  are captured by the divergence,  $\nabla \cdot \mathbf{u}$ , as per equation (5.47).

For the shallow water fluid, the thickness weighted potential vorticity equals to the absolute vorticity

$$h Q = \zeta_a. \quad (5.53)$$

Consequently, the flux-form conservation form of the potential vorticity equation (5.52) is identical to equation (5.46) for the absolute vorticity

$$\partial_t(h Q) = -\nabla \cdot (h Q \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}) \iff \partial_t \zeta_a = -\nabla \cdot (\zeta_a \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}). \quad (5.54)$$

### Potential vorticity flux vector

It is remarkable that even with non-conservative forcing, the thickness weighted potential vorticity (equivalently, the absolute vorticity) has its Eulerian evolution determined by the convergence of a flux,

$$\mathbf{J}^Q = h Q \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}. \quad (5.55)$$

Observe that this potential vorticity flux is oriented in the horizontal direction. As detailed in Chapter 8, the potential vorticity flux never crosses the interface between two shallow water layers, even in the presence of non-conservative processes such as mixing and friction. This result is a special case of the [impermeability theorem](#) that holds for continuously stratified fluids (Section 8.2).

### A kinematic expression for the potential vorticity flux vector

The flux convergence evolution for the potential vorticity equation (5.54) is a kinematic result of the definition of potential vorticity. Namely,

$$h Q - f = \zeta = \nabla \cdot (\mathbf{u} \times \hat{\mathbf{z}}) \quad (5.56)$$

so that

$$\partial_t(h Q) = \nabla \cdot (\partial_t \mathbf{u} \times \hat{\mathbf{z}}) \equiv -\nabla \cdot \mathbf{J}^{kin}, \quad (5.57)$$

where we defined the kinematic potential vorticity flux vector

$$\mathbf{J}^{kin} \equiv -\partial_t \mathbf{u} \times \hat{\mathbf{z}}. \quad (5.58)$$

We can relate the kinematic potential vorticity flux vector to  $\mathbf{J}^Q$  through the following.

Take the cross product of the vector invariant velocity equation (5.2) with  $\hat{z}$  to find

$$\partial_t \mathbf{u} \times \hat{z} + \zeta_a \mathbf{u} = -\nabla(g\eta + \mathbf{u} \cdot \mathbf{u}/2) \times \hat{z} + \mathbf{F} \times \hat{z}. \quad (5.59)$$

Now write

$$\zeta_a \mathbf{u} = h Q \mathbf{u} = \mathbf{J}^Q - \hat{z} \times \mathbf{F} \quad (5.60)$$

to have

$$\mathbf{J}^{kin} = \mathbf{J}^Q + \nabla \times [\hat{z}(g\eta + \mathbf{u} \cdot \mathbf{u}/2)]. \quad (5.61)$$

Since the two potential vorticity fluxes differ by a curl; i.e., a rotational term, their divergences are identical

$$\nabla \cdot \mathbf{J}^{kin} = \nabla \cdot \mathbf{J}^Q, \quad (5.62)$$

so that their convergence leads to the same evolution of  $hQ$ . Stated more formally,  $\mathbf{J}^{kin}$  and  $\mathbf{J}^Q$  differ by a gauge, with the gauge function given by  $g\eta + \mathbf{u} \cdot \mathbf{u}/2$ .

## 5.5 Example implications of material PV invariance

The material invariance of shallow water potential vorticity constrains the shallow water motion by stating that  $f, h, \zeta$  cannot change independently of the other. Rather, the combination  $Q = (f + \zeta)/h$  must remain materially invariant (in the absence of non-conservative processes). There are a variety of situations that induce changes in one or two of the terms, with the third term constrained to ensure  $Q$  remains unchanged. We here consider some thought experiments to garner experience with shallow water PV-thinking.

### 5.5.1 Topographic beta effect

Changes in the topography affect the potential vorticity by changing the thickness of a fluid column via (see Figure 1.1)

$$h = H + \Delta\eta - \Delta\eta_b, \quad (5.63)$$

with  $H = \bar{h}$  the area mean layer thickness,  $\Delta\eta = \eta - \bar{\eta}$  the deviation of the free surface from its area mean, and  $\Delta\eta_b = \eta_b - \bar{\eta}_b$  the deviation of the bottom from its area mean. For relative vorticity, we note that spatial changes in the topography act to drive a vertical velocity at the layer bottom as per equation (1.29),

$$w = \frac{D\eta_b}{Dt} \quad \text{at } z = \eta_b, \quad (5.64)$$

which then leads to vortex stretching and hence to a change in relative vorticity. In this subsection we highlight the analog between topographic slopes and planetary beta to thus motivate the term **topographic beta effect**.

To mathematically exhibit the topographic beta effect, consider a fluid column whose vorticity is dominated by planetary vorticity and with bottom topography having a small and linear slope in the meridional direction

$$\Delta\eta_b = \delta y, \quad (5.65)$$

where  $|\delta| \ll 1$  is the topographic slope. Assuming the free surface undulations are small relative

to the resting layer thickness,  $\Delta\eta \ll H$ , we can expand the potential vorticity according to

$$Q = \frac{f + \zeta}{h} \quad (5.66a)$$

$$= \frac{f_o + \beta y + \zeta}{H + \Delta\eta - \delta y} \quad (5.66b)$$

$$\approx \frac{f_o + \beta y + \zeta}{H} [1 - H^{-1}(\Delta\eta - \delta y)] \quad (5.66c)$$

$$\approx \frac{f_o + \beta y + \zeta}{H} - \frac{f_o}{H^2} (\Delta\eta - \delta y). \quad (5.66d)$$

Setting  $DQ/Dt = 0$  and rearranging leads to the material evolution of relative vorticity

$$\frac{D\zeta}{Dt} = -v(\beta + f_o\delta/H) + \frac{f_o w(\eta)}{H}. \quad (5.67)$$

The second term on the right hand side is the vortex stretching associated with vertical motion at the top of the layer, where

$$w(\eta) = \frac{D(\Delta\eta)}{Dt} = \frac{D\eta}{Dt}, \quad (5.68)$$

according to the surface kinematic boundary condition (1.37). The first term on the right hand side is vortex stretching arising from both planetary beta and topographic slopes. It is written in a form revealing the parallels between these two contributions, and it is readily generalized to the following for arbitrary topography

$$\beta^{\text{eff}} = (H - \eta_b) \nabla[f/(H - \eta_b)] \approx \beta \hat{\mathbf{y}} + (f_o/H) \nabla \eta_b. \quad (5.69)$$

One of the more prominent roles for topographic beta is in supporting [topographic Rossby waves](#), which are analogous to the planetary [Rossby waves](#) supported by planetary beta (see [VOLUME 4](#)). We also encounter the topographic beta effect in Section 9.6.1 in our study of the quasi-geostrophic shallow water model.

### 5.5.2 Planetary geostrophic potential vorticity and $f/H$ contours

As introduced in [VOLUME 2](#), [planetary geostrophy \(PG\)](#) is used to study the large-scale laminar ocean circulation where relative vorticity is ignored. Furthermore, as shown in Section 9.4, the inviscid and adiabatic PG system materially preserves the PG potential vorticity,  $Q = f/h$ , so that

$$\frac{D(f/h)}{Dt} = 0. \quad (5.70)$$

Consequently, fluid particles respecting the inviscid planetary geostrophic equations follow contours of constant  $f/h$ . These contours are defined by the [geostrophic streamfunction](#), which provides the pathways for the steady geostrophic flow. In Section 5.7.3 we offer a geometrical interpretation of  $Q = f/h$  material invariance in terms of the [Taylor-Proudman effect](#).

#### **Example $f/H$ contours**

If we assume the free surface undulations are negligible compared to the bottom topography (a useful assumption for planetary geostrophic flow), then shallow water columns follow contours of constant  $f/H$ , where  $z = -H(x, y) = \eta_b(x, y)$  is the vertical position of the bottom topography. In Figure 5.5 we illustrate  $f/H$  contours for a topographic seamount (bump), a topographic

depression (bowl), and a shelf/slope along the western boundary. We see that  $f/H$  contours are diverted equatorward when depth decreases, whereas they are diverted poleward when encountering deeper water. Furthermore, those contours near to either a bump or bowl are closed, so that fluid columns following these contours are trapped around the topographic feature.

For the shelf/slope region in Figure 5.5, the  $f/H$  contours are horizontal where the topography is flat, which for this example is on the shelf and in the open ocean, whereas they are steered toward the equator as they pass from the shelf toward the coast. In this example, the difference between the shelf and deep ocean is only around 200 m, so that this example corresponds to upper ocean columns interacting with the continental slope. Part of the reason for choosing this geometry is that thicker fluid columns from the open ocean are unable to reach the continental shelf. The reason is that thick columns, as they reach the slope, have their  $f/H$  contours reach the equator along the slope before they reach the shelf. That is, the  $f/H$  contours for thick open ocean columns are almost entirely southward when reaching the continental slope, so that they cannot climb to shallower depths on the shelf.

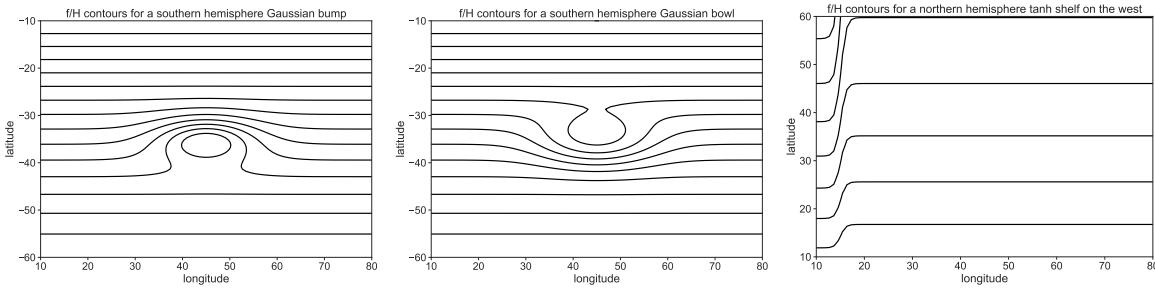


FIGURE 5.5: When a shallow water fluid is governed by the inviscid and adiabatic planetary geostrophic equations and there is no external forcing, then shallow water fluid columns maintain fixed planetary geostrophic potential vorticity,  $D(f/h)/Dt = 0$ . Ignoring free surface undulations relative to changes in the bottom depth means that  $f/H$  remains fixed following the inviscid geostrophic flow, where  $z = -H(x, y) = \eta_b(x, y)$  is the bottom topography. We here illustrate contours of the [geostrophic streamfunction](#) for three topographic features: a Gaussian seamount or bump and a Gaussian bowl, both in the southern hemisphere; and a western boundary continental slope and shelf in the northern hemisphere. Contours of  $f/H$  follow lines of constant latitude when  $H$  is constant,  $f/H$  contours are steered equatorward when moving into a region of shoaling water ( $H$  decreases), and steered poleward when moving into deeper water ( $H$  increases). Furthermore, note that contours near the seamount and bowl can close, in which case the associated geostrophic contours are trapped next to the topographic features.

### How planetary $\beta$ affects $f/H$ contours

The shelf example in Figure 5.5 reveals that for a given change in depth, the latitudinal diversion of an  $f/H$  contour is larger in magnitude for poleward contours relative to equatorward contours. We here show that this property of the  $f/H$  contours arises from planetary  $\beta$ .

For this purpose, consider a particular  $f/H$  contour, on which the latitude and depth change so as to keep  $f/H$  fixed so that

$$\delta(f/H) = 0, \quad (5.71)$$

where  $\delta$  is a differential operator following the  $f/H$  contour. With  $f = 2\Omega \sin \phi$ , the constraint (5.71) relates deviations in latitude,  $\delta\phi$ , to deviations in depth,  $\delta H$ , along an  $f/H$  contour

$$\delta\phi = (\delta H/H) \tan \phi = \frac{f/H}{2\Omega \cos \phi} = \frac{f}{H} \frac{\delta H}{R_e \beta} \implies \frac{\delta\phi}{\delta H} = \frac{f}{H} \frac{1}{R_e \beta}, \quad (5.72)$$

where we introduced planetary  $\beta$  according to

$$\beta = \partial_y f = (2\Omega/R_e) \cos \phi. \quad (5.73)$$

For a particular  $f/H$  contour, we see that  $\delta\phi/\delta H$  is proportional to  $\beta^{-1}$ . Consequently, the magnitude of  $\delta\phi/\delta H$  increases moving poleward, as  $\beta$  decreases toward zero.

### 5.5.3 Spin up of converging flow

Consider the flow shown in Figure 5.6, whereby mass in the shallow water layer converges into a region. Just as described in the potential vorticity derivation Figure 5.4, increasing the column thickness, without substantially altering the planetary vorticity (e.g.,  $f$ -plane), requires  $\partial\zeta/\partial t > 0$  in order to maintain  $Q = (\zeta + f)/h$  materially constant. Following our discussion of the rotating column in Section 5.2, note that convergence of mass reduces the moment of inertia relative to the center of the region. Angular momentum conservation requires the fluid to rotate faster thus picking up a positive relative vorticity. This dynamical process is embedded in the material invariance of potential vorticity. Finally, note that the opposite occurs in a region of diverging fluid, whereby potential vorticity material invariance implies that the relative vorticity has a negative tendency ( $\partial\zeta/\partial t < 0$ ) (see also Figure 5.4).

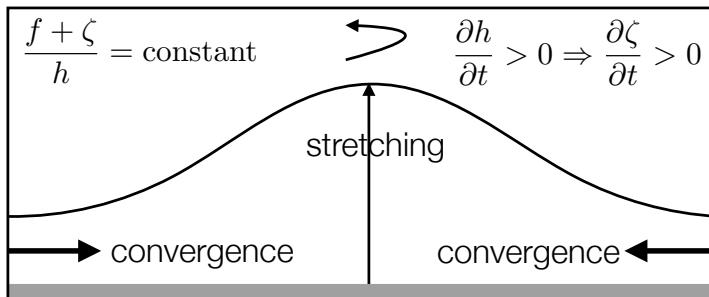


FIGURE 5.6: Illustrating the implications of potential vorticity material invariance for a shallow water fluid on an  $f$ -plane. If mass converges into a region, thus stretching the fluid column, then potential vorticity material invariance implies the relative vorticity increases,  $\partial\zeta/\partial t > 0$ . This result is directly analogous to the rotating cylinder example considered in Figures 5.1 and 5.4. Namely, converging a region of constant mass reduces its moment of inertia so that angular momentum conservation leads to an increase in spin.

### 5.5.4 Further study

Section 4.5 of [Holton and Hakim \(2013\)](#) discusses the case of flow over topography where the full shallow water potential vorticity is materially invariant,  $D(f + \zeta)/Dt = 0$ . In that case there is a dramatic difference between easterly and westerly flows. In the northern hemisphere, westerly winds (eastward flow) deflect over the topography and downstream it undulates as topographic leewaves. A rotating tank offers a useful controlled setting to observe leewaves, such as shown near the 20 minute mark in [this video from Prof. Fultz](#). Easterly winds (westward flow) do not exhibit a wavelike pattern, instead following a trajectory similar to the  $f/H$  contours of planetary geostrophic case, though modified by relative vorticity. In general, the study of flow near topography, either in the shallow water or continuously stratified, introduces a wealth of dynamical behaviors where material invariance of potential vorticity provides an important tool to help unravel mechanisms.

## 5.6 Circulation with non-conservative processes

We follow the discussion in Section 5.4 to study the evolution of circulation in the presence of non-conservative processes such as dia-surface transport and boundary stresses (Section 1.6). For this purpose, consider the velocity circulation around a closed horizontal area,  $\mathcal{S}$  (see Figure 5.7)

$$\mathcal{C} = \int_{\mathcal{S}} \zeta \, d\mathcal{S} = \oint_{\partial\mathcal{S}} \mathbf{u} \cdot \hat{\mathbf{t}} \, d\ell. \quad (5.74)$$

In this equation,  $\hat{\mathbf{t}} \, d\ell$  is the vector line increment around the contour, and  $\hat{\mathbf{t}}$  is the tangent vector orienting the contour integral in a counterclockwise direction. We assume the circulation contour extends vertically through the non-vanishing fluid layer<sup>4</sup> so that the unit outward normal,  $\hat{\mathbf{n}}$ , to the contour is strictly horizontal, as is the tangent vector,  $\hat{\mathbf{t}}$ . We now seek an evolution equation for this circulation by making use of the vorticity equation (5.54).

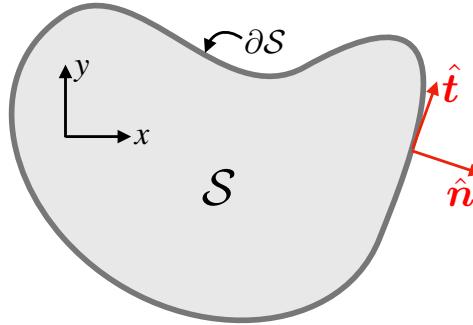


FIGURE 5.7: Illustrating the calculation of circulation around the contour,  $\partial\mathcal{S}$ , of a static horizontal area,  $\mathcal{S}$ , within a layer of shallow water fluid. The circulation theorem (5.78) provides the means to compute the time changes in circulation as a function of the advection of absolute vorticity crossing the contour plus the circulation of friction. Note that the contour has a vertical extent throughout the shallow water layer, so that the unit outward normal,  $\hat{\mathbf{n}}$ , and the unit tangent direction,  $\hat{\mathbf{t}}$ , are both horizontal vectors.

To develop an evolution equation for  $\mathcal{C}$ , integrate the vorticity equation (5.54) over the area  $\mathcal{S}$  to yield

$$\int_{\mathcal{S}} \frac{\partial \zeta}{\partial t} \, d\mathcal{S} = \int_{\mathcal{S}} [-\nabla \cdot (\zeta_a \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F})] \, d\mathcal{S} = - \oint_{\partial\mathcal{S}} [\zeta_a \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}] \cdot \hat{\mathbf{n}} \, d\ell, \quad (5.75)$$

where  $\oint$  symbolizes an integral around the contour. To reach this equation we used the divergence theorem on the right hand side, with  $\hat{\mathbf{n}}$  the horizontal unit outward normal vector along the contour,  $\partial\mathcal{S}$ , and  $d\ell$  is the line increment along the contour. We also set  $\partial f / \partial t = 0$  as part of the time derivative of the absolute vorticity. The non-conservative forcing term can be written

$$(\hat{\mathbf{z}} \times \mathbf{F}) \cdot \hat{\mathbf{n}} = (\hat{\mathbf{n}} \times \hat{\mathbf{z}}) \cdot \mathbf{F} = -\hat{\mathbf{t}} \cdot \mathbf{F}, \quad (5.76)$$

thus leading to

$$\int_{\mathcal{S}} \frac{\partial \zeta}{\partial t} \, d\mathcal{S} = - \oint_{\partial\mathcal{S}} \zeta_a \mathbf{u} \cdot \hat{\mathbf{n}} \, d\ell + \oint_{\partial\mathcal{S}} \mathbf{F} \cdot \hat{\mathbf{t}} \, d\ell, \quad (5.77)$$

where  $\oint$  is the counter-clockwise oriented closed contour integral. We next assume the area  $\mathcal{S}$

<sup>4</sup>The case of a vanishing layer thickness, such as occurs when allowing for sloped side boundaries as per a seashore, is handled by studying the thickness weighted velocity equation and the corresponding column vorticity. We consider this topic in Section 5.8.

is constant in time, so that the Eulerian time derivative can be pulled across the integral to render

$$\frac{d\mathcal{C}}{dt} = - \oint_{\partial S} \zeta_a \mathbf{u} \cdot \hat{\mathbf{n}} d\ell + \oint_{\partial S} \mathbf{F} \cdot \hat{\mathbf{t}} d\ell. \quad (5.78)$$

The first term on the right hand side arises from the horizontal advection of absolute vorticity across the contour. This term is not oriented and so the integral sign has no arrow. The second term arises from the counter-clockwise oriented circulation of any non-conservative accelerations. We refer to equation (5.78) as a *circulation theorem*. It has many uses under specific cases, some of which are described in the remainder of this chapter.

### 5.6.1 Circulation around a closed streamline in steady flow

As a particular example of the circulation theorem (5.78), consider a steady flow in the absence of boundary volume sources ( $w^{(\dot{\eta})} = 0$ ). In this case, the thickness equation (1.19) reduces to  $\nabla \cdot (h \mathbf{u}) = 0$  so that we can introduce a transport streamfunction,  $\Psi$  (with dimensions of  $L^3 T^{-1}$ )

$$h \mathbf{u} = \hat{\mathbf{z}} \times \nabla \Psi. \quad (5.79)$$

In a bounded domain, the streamlines (contours of constant  $\Psi$ ) are closed. Furthermore, the unit outward normal to a closed streamline is perpendicular to the velocity,  $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$ . We thus see that the steady state expression of the circulation theorem (5.78), computed around a closed streamline, leads to the following constraint on the non-conservative forces

$$\oint_{\text{streamline}} \mathbf{F} \cdot \hat{\mathbf{t}} d\ell = \int_{\text{streamline area}} (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{z}} dS = 0, \quad (5.80)$$

where the second equality follows from Stokes' theorem applied over the area bounded by the streamline. Equation (5.80) provides a constraint on the non-conservative forcing that must be satisfied to enable a steady flow. For example, when integrated around a closed streamline, the wind stress forcing must balance dissipation. If the constraint (5.80) is not satisfied, then the flow cannot reach a steady state. Although we may not know explicit details of the streamlines, we can still make use of this constraint if we assume the flow is steady. The analysis in Section 5.7 offers an example application of these ideas for studies of circulation in steady ocean gyres.

### 5.6.2 Circulation from wind stress and Rayleigh drag

A particularly simple form for the non-conservative acceleration is given by

$$\mathbf{F} = -\gamma \mathbf{u} + \boldsymbol{\tau}^{\text{wind}} / (h \rho) \equiv -\gamma \mathbf{u} + \mathbf{F}^{\text{wind}}. \quad (5.81)$$

The first term is referred to as **Rayleigh drag** with  $\gamma > 0$  a constant with dimensions of inverse time.<sup>5</sup> Rayleigh drag damps all flow to rest with  $\gamma^{-1}$  the e-folding time for the damping. The second term in equation (5.81) is the acceleration on the layer from wind stress,

$$\mathbf{F}^{\text{wind}} = \frac{\boldsymbol{\tau}^{\text{wind}}}{h \rho}, \quad (5.82)$$

with this form following from the discussion of boundary stresses in Section 1.6.5. Namely, homogeneity of the shallow water layer renders the contact stress from winds into a body stress

---

<sup>5</sup>See VOLUME 2 for a discussion of Rayleigh drag in the context of **Ekman mechanics**.

applied throughout the layer.

Plugging the acceleration (5.81) into the time dependent circulation theorem (5.78) leads to

$$(d/dt + \gamma) \mathcal{C} = - \int_{\partial S} \zeta_a \mathbf{u} \cdot \hat{\mathbf{n}} d\ell + \oint_{\partial S} \mathbf{F}^{\text{wind}} \cdot \hat{\mathbf{t}} d\ell. \quad (5.83)$$

Specializing to a steady state and choosing the contour as a closed streamline (along which  $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$ ), renders

$$\mathcal{C} = \gamma^{-1} \oint_{\text{streamline}} \mathbf{F}^{\text{wind}} \cdot \hat{\mathbf{t}} d\ell = (\gamma \rho)^{-1} \oint_{\text{streamline}} (\boldsymbol{\tau}^{\text{wind}}/h) \cdot \hat{\mathbf{t}} d\ell. \quad (5.84)$$

This equation says that velocity circulation around a closed streamline is determined by wind stress circulation around that streamline plus knowledge of the Rayleigh drag damping time scale  $\gamma^{-1}$ . This result supports our expectation that the steady circulation around a closed streamline is oriented with the same sense as the applied wind stress.

## 5.7 A primer on steady ocean gyres

Large-scale gyres are a prominent feature of ocean circulation, with the North Atlantic and North Pacific middle-latitude gyres two canonical examples. It is particularly remarkable that ocean gyres are not symmetric in the east-west direction, with a prominent western side where poleward flow is stronger than the more sluggish equatorward flow in the interior. As shown in this section, gyre zonal asymmetry is not a response to the wind forcing, with the asymmetry found even without any zonal variations in the boundary forcing. Instead, it is a manifestation of the beta effect present for flow on a rotating spherical planet (or idealized as the [beta-plane](#) ( $\beta$ -plane) from [VOLUME 2](#)). The role of

$$\beta = \partial_y f > 0 \quad (5.85)$$

in western intensification was first articulated by [Stommel \(1948\)](#), with  $\beta$  encapsulating the leading order role of the earth's sphericity on large-scale flows in the atmosphere and ocean.

We have the basic tools in hand to understand the physical balances leading to western intensification in steady ocean gyres. We follow the traditional approach by focusing on vorticity balance, which offer a more direct path towards understanding western intensification than the momentum or axial angular momentum balances used to explore channel flow in [VOLUME 2](#) as well as [Section 2.7](#). Furthermore, observe that with western intensification fundamentally relying on  $\beta > 0$ , then the arguments given below hold for both hemispheres; i.e., gyres are western intensified in both hemispheres and with either signed wind stress curl. So although we orient the discussion according to a northern hemisphere anti-cyclonic gyre, the arguments hold in general.

### 5.7.1 Steady and large-scale vorticity balance

The steady circulation theorem (5.84) holds regardless the bottom topography or surface height undulations. Again, it says that circulation around a closed streamline is in the same sense as the wind circulation. However, we need more information to see how western intensification emerges as a property of the flow in ocean gyres. For that purpose, consider the steady absolute

vorticity equation (5.54), again in the presence of wind forcing and Rayleigh drag

$$\nabla \cdot (\mathbf{u} \zeta_a) = -\gamma \zeta + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.86)$$

Introducing the shallow water potential vorticity,  $Q = \zeta_a/h$ , allows us to write

$$\nabla \cdot (\mathbf{u} \zeta_a) = \nabla \cdot (h \mathbf{u} Q). \quad (5.87)$$

The steady state thickness equation (1.19) means that  $\nabla \cdot (h \mathbf{u}) = 0$ , so that the vorticity equation (5.86) takes the form

$$h \mathbf{u} \cdot \nabla Q = -\gamma \zeta + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.88)$$

This equation says that in the absence of the Rayleigh friction ( $\gamma = 0$ ) and with a zero wind stress curl, the steady horizontal flow is aligned with potential vorticity contours. This result follows directly from the material invariance of potential vorticity in the absence of non-conservative processes. However, in the presence of Rayleigh drag and/or wind stress curl, the potential vorticity is modified when following the flow so that  $\mathbf{u} \cdot \nabla Q \neq 0$ , in which case the circulation does not follow  $Q$  contours.

## 5.7.2 Planetary geostrophic flow and the Sverdrup balance

For large-scale flow away from lateral boundaries, the flow has an absolute vorticity that is dominated by planetary vorticity so that

$$Q \approx Q^{\text{pg}} = f/h, \quad (5.89)$$

which is the potential vorticity for shallow water planetary geostrophic flow introduced in Section 5.5.2 and studied more thoroughly in VOLUME 2 as well as in Section 9.4. In this flow the only means for changing potential vorticity arise from changes to planetary vorticity (changes to  $f$ ) and changes to layer thickness,  $h$ .

Away from boundaries we also assume the Rayleigh drag term is negligible since the relative vorticity is small. In this case, the potential vorticity equation (5.88) takes the form

$$h \mathbf{u} \cdot \nabla Q^{\text{pg}} = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.90)$$

Expanding the left hand side and introducing the planetary vorticity gradient renders the shallow water Sverdrup balance

$$\beta v = Q^{\text{pg}} \mathbf{u} \cdot \nabla h + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}) \quad \text{shallow water Sverdrup balance.} \quad (5.91)$$

This balance states how horizontal advection of layer thickness (first right hand side term) plus the wind stress curl (second term) balance meridional motion for flow on a rotating sphere (beta effect on left hand side). Gradients in the layer thickness arise from free surface undulations as well as gradients in the bottom topography (see Figure 1.1). The traditional Sverdrup balance arises when we assume the flow takes place over a flat bottom and the free surface undulations are negligible, in which case  $\nabla h = 0$  so that

$$\beta v = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}) \quad \text{traditional Sverdrup balance.} \quad (5.92)$$

This balance is also suited to a stratified ocean in regions where the upper ocean flow does not interact with the bottom, such as in regions far from the continental shelves, or in regions where vertical stratification (i.e., the pycnocline) shields the upper ocean flows from bottom topography.

### 5.7.3 Taylor-Proudman, Sverdrup balance, and $f/h$ invariance

Material invariance of planetary geostrophic potential vorticity,  $Q^{\text{pg}} = f/h$ , has a geometric interpretation in terms of the [Taylor-Proudman effect](#) of vertical stiffening discussed in VOLUME 2. As shown in Figure 5.8, since  $Q^{\text{pg}}$  is materially conserved then so is the stiffness distance,  $h^{\text{stiff}} = h/\sin \phi = 2\Omega/Q^{\text{pg}}$ , where we wrote the Coriolis parameter as  $f = 2\Omega \sin \phi$ . Material invariance of  $h^{\text{stiff}}$  manifests the [Taylor-Proudman effect](#), whereby geostrophically balanced fluid columns are stiffened in the direction of the rotation axis. Here, the rotation axis is the planetary axis through the poles.

Vertical fluid columns respecting the planetary geostrophic balance are stiffened in a direction that parallels the planetary rotation axis. As a result, processes that cause the layer thickness,  $h$ , to squash/compress must occur with equatorward movement, whereas expansion/stretching of layer thickness occurs with poleward movement. Figure 5.8 provides a geometric lens for understanding the shallow water Sverdrup balance (5.91) discussed in Section 5.7.2, with boundary stresses and nonlinear advection leading to changes in layer thickness and so leading to meridional motion. This geometrical view makes it clear that Sverdrup balance requires planetary curvature.

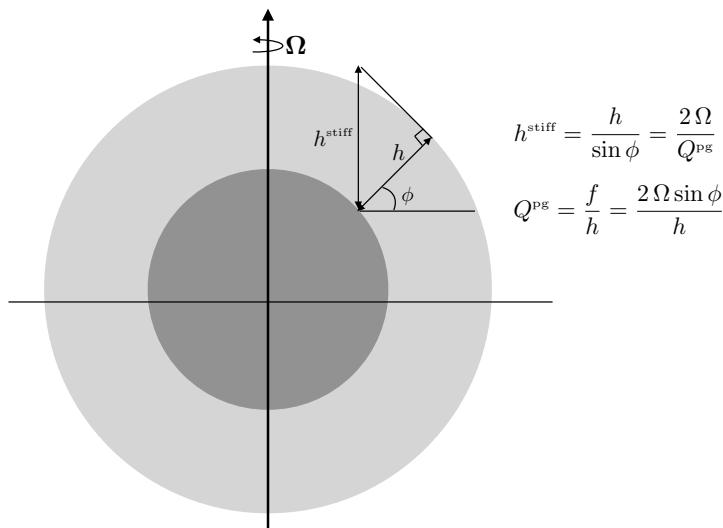


FIGURE 5.8: This figure depicts a homogeneous layer of shallow water fluid (light gray shell; not drawn to scale!) covering a rotating planet, with  $h$  the fluid layer thickness. Material invariance of planetary geostrophic potential vorticity,  $Q^{\text{pg}} = f/h = 2\Omega \sin \phi/h$ , also means that the distance,  $h^{\text{stiff}} = h/\sin \phi = 2\Omega/Q^{\text{pg}}$ , is materially conserved. Material conservation of  $h^{\text{stiff}}$  reflects the vertical stiffening along the rotational axis that occurs for geostrophic flows as per the [Taylor-Proudman effect](#) from VOLUME 2. Processes that change the layer thickness correspond to meridional movement of the fluid column in such a manner to maintain material invariance of  $h^{\text{stiff}}$ . Namely, a process that causes  $h$  to squash/compress leads to equatorward movement whereas expanding/stretching  $h$  leads to poleward movement. This constrained motion extends to multiple shallow water layers, yet with the interior interfaces allowed to expand or contract. This figure is inspired by Figure 1 of [Rhines \(1980\)](#).

### 5.7.4 Sverdrup flow in a closed domain with anti-cyclonic wind stress

Consider a closed northern hemisphere middle latitude  $\beta$ -plane domain driven by an anti-cyclonic wind stress

$$\hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}}) < 0 \quad \text{northern hemisphere anti-cyclonic wind stress.} \quad (5.93)$$

This situation is depicted in Figure 5.9, where we also illustrate a commonly used wind stress profile that is purely zonal and has a co-sinusoidal meridional structure that is symmetric about the central latitude of the domain

$$\mathbf{F}^{\text{wind}} = -\hat{x} A \cos[\pi (y - y_0 + L/2)/L] \quad (5.94a)$$

$$\hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}}) = -(\pi A/L) \sin[\pi (y - y_0 + L/2)/L]. \quad (5.94b)$$

In these equations,  $A > 0$  is the magnitude of the wind stress acceleration applied to the layer, and the domain extends meridionally from  $y_0 - L/2 \leq y \leq y_0 + L/2$  with  $y = y_0$  the central latitude. This wind stress has westerlies on the poleward side of the domain and easterlies (trade winds) on the equatorward side so that  $\hat{z} \cdot (\nabla \times \tau) < 0$  throughout the domain.

The Sverdrup balance (5.92) indicates that an anti-cyclonic wind stress curl drives an equatorward Sverdrup flow. We emphasize that this flow is *not* the result of meridional winds pushing the fluid to the south. Instead, it arises in response to the constraints of vorticity balance with an anti-cyclonic wind stress curl in the presence of the beta effect. Indeed, for the idealized wind stress (5.94a) there is no meridional wind component. Although fluid satisfying Sverdrup balance flows south, all the fluid in the domain cannot be moving to the south. Rather, volume conservation requires a poleward return flow somewhere outside the region of Sverdrup balance.

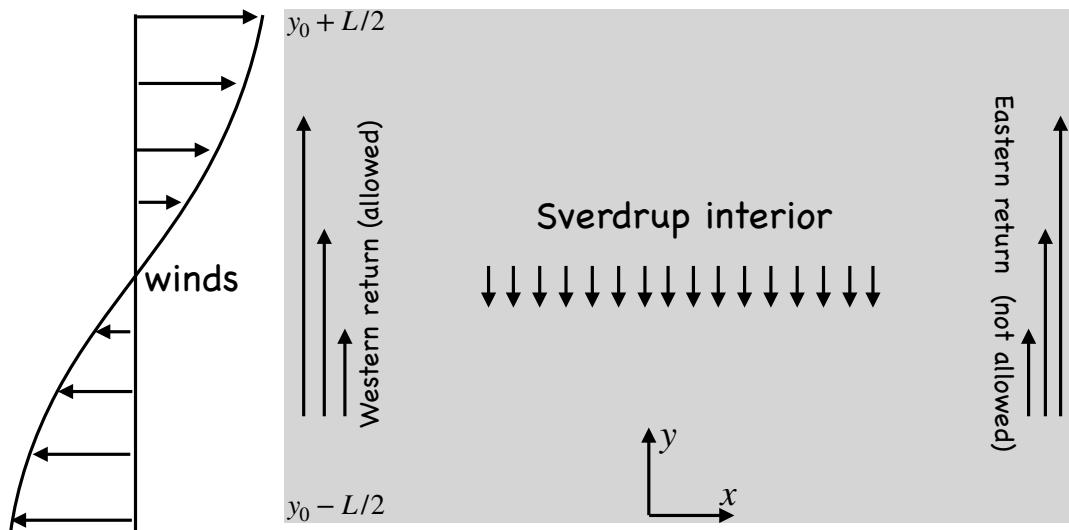


FIGURE 5.9: Illustrating the southward Sverdrup flow in response to an anti-cyclonic wind stress forcing in a bounded northern hemisphere domain. A northward return flow is required to satisfy volume conservation. As seen in Section 5.7.5, a linear vorticity balance between beta, winds, and friction lead to a western boundary return flow and corresponding western intensification. An analogous eastern intensification is not dynamically allowed.

### 5.7.5 Western intensification and the role of beta

Volume conservation is a kinematic constraint that requires a return flow on either the eastern or western side of the domain, outside the region of Sverdrup balance. But what side? We offer the following arguments for the western side, with these arguments representing the basic elements to the Stommel model for western boundary current intensification. In brief, planetary vorticity variation (the beta effect ( $\beta$ -effect)) plus linear bottom friction forces the return flow of a wind-driven gyre to be more intense (and thus to be more narrow) along the western side of the basin.

#### Inertial entry into a boundary layer region

Recall the discussion of Figure 4.4 where we considered how inviscid flow of a two-dimensional non-divergent fluid over a flat bottom region materially preserves absolute vorticity in the presence of a meridional barrier. To materially preserve absolute vorticity (and ignoring free surface undulations relative to the depth of the fluid), the flow can deviate meridionally, either northward or southward, when encountering a western wall. In contrast, such meridional deviation is prohibited for absolute vorticity preserving flow that encounters an eastern wall. So in referring to Figure 5.9, southward flow can make a turn westward towards the western boundary, enter the boundary layer, and move northward within the boundary layer. It cannot do so for the eastern side. The central limitation of this argument concerns the presence of dissipation in the boundary, in which absolute vorticity is no longer materially preserved. Even so, the argument offers a useful first suggestion for the flow favoring the western side.

#### Steady vorticity balance and a role for dissipation

Even if the western side is the preferred region for the return flow, we still need an argument for intensification of that flow. To develop an argument, recall that the return flow region is not in Sverdrup balance. To see what terms can break that balance, consider again the steady vorticity balance (5.88). Continuing to assume a flat bottom and rigid lid surface leads to

$$\beta v = -\gamma \zeta + \hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}}), \quad (5.95)$$

where we dropped the nonlinear advection term  $\mathbf{u} \cdot \nabla \zeta$  (the “inertial” term) since we wish to determine whether a linear balance can give rise to western intensification (we return to this assumption in Section 5.7.9). The Rayleigh drag on the right hand side breaks the Sverdrup balance in regions where the relative vorticity is nontrivial. Since we know there must be a return flow somewhere in the domain, we know there must be a region where dissipation is sufficiently strong to break Sverdrup balance.

What is required for the steady and linear dissipative vorticity balance (5.95) to be maintained in the northward return flow region? To answer this question, expose the signs on the terms in equation (5.95)

$$\underbrace{\beta v}_{\text{positive}} + \underbrace{\gamma \zeta}_{\text{unspecified}} = \underbrace{\hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}})}_{\text{negative}}. \quad (5.96)$$

We have  $\beta v > 0$  since we are concerned with the region of northward return flow, and  $\hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}}) < 0$  by assumption of anti-cyclonic wind stress over the full domain. Hence, for the balance (5.96) to be realized requires  $\zeta < 0$ , with the value large enough to balance both

the winds and the meridional advection of planetary vorticity

$$\underbrace{\gamma \zeta}_{\text{negative}} = \underbrace{-\beta v}_{\text{negative}} + \underbrace{\hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}})}_{\text{negative}}. \quad (5.97)$$

For anti-cyclonic gyre flow, as required by the circulation condition (5.84), the only way to realize  $\zeta < 0$  of sufficient magnitude is to have an intensified flow along the western side of the gyre. In this region,  $\zeta \approx \partial v / \partial x < 0$  can become sufficiently large in magnitude. Furthermore, since the wind stress is applied throughout the domain, all streamlines feel the winds and must pass through the western boundary region where vorticity is enhanced and Rayleigh drag is able to balance the winds and planetary advection. The required boundary current flow is depicted in Figure 5.9.

### The importance of beta

In the absence of beta, there would be no interior region in Sverdrup balance driving southward flow. The flow would thus only be subject to the circulation condition (5.84) whereby linear flow can symmetrically dissipate the wind stress. Hence, the beta effect is the fundamental element that causes poleward flow to intensify along the western side of the gyre in response to the equatorward interior flow.

### Western intensification regardless the sense for the wind stress curl

The steady circulation theorem (5.84) means that the flow circulation is in the same sense as the wind circulation. The arguments offered above for western intensification focused on the anti-cyclonic winds, as per the middle latitude gyres in the northern hemisphere Atlantic and Pacific oceans. What if the winds were cyclonic? In that case, the steady circulation theorem (5.84) means that the circulation is also cyclonic. Even so, the arguments based on the beta effect still result in western intensification. Namely, with cyclonic winds the interior Sverdrup flow is northward, so that the return flow must be southward. For this case, equation (5.97) now takes the form

$$\underbrace{\gamma \zeta}_{\text{positive}} = \underbrace{-\beta v}_{\text{positive}} + \underbrace{\hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}})}_{\text{positive}}. \quad (5.98)$$

Again, we find that vorticity arguments lead to western intensification of the cyclonic gyre.

### The Stommel equation for the streamfunction

By assuming a rigid lid and homogeneous fluid layer, the fluid velocity is horizontally non-divergent and so it can be written in terms of a streamfunction

$$\nabla \cdot (h \mathbf{u}) = h \nabla \cdot \mathbf{u} = 0 \implies \mathbf{u} = \hat{z} \times \nabla \psi = h^{-1} \hat{z} \times \nabla \Psi, \quad (5.99)$$

where  $\Psi$  is the streamfunction for the thickness weighted velocity from equation (5.79). Introducing the streamfunction into the linearized vorticity balance (5.95) leads to

$$(\gamma \nabla^2 + \beta \partial_x) \psi = \hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.100)$$

We encountered this linear partial differential equation in Section 4.8.4 when studying the beta plume. The two problems describe steady gyre circulations in the presence of Rayleigh drag, the beta effect, and wind stress curl. For the beta plume we ignored boundaries and

solved for the free space Green's function shown in Figure 4.7. Here, the western boundary is a fundamental feature of the problem, with further analysis prompting the use of rudimentary boundary layer theory to match the interior Sverdrup solution to the boundary region. That analysis is summarized in Section 19.1.3 of [Vallis \(2017\)](#).

### 5.7.6 A role for bottom pressure torques

Recall the shallow water Sverdrup balance (5.91), here with the addition of Rayleigh drag

$$\beta v = Q^{\text{pg}} \mathbf{u} \cdot \nabla h - \gamma \zeta + \hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.101)$$

For the right hand side, we have thus far considered a flat bottom and ignored free surface undulations, in which case  $\nabla h = 0$ . In this case, the only way to balance meridional motion on a beta plane is to invoke non-conservative processes either from wind stress or Rayleigh drag.

Spatial variations in the bottom topography open up the possibility for an inviscid balance

$$\beta v = Q^{\text{pg}} \mathbf{u} \cdot \nabla h, \quad (5.102)$$

or more generally a balance where frictional vorticity sinks are unimportant except for regions very close to the boundary. For example, consider northward flow in the northern hemisphere along a shallow western continental shelf. Assume the bottom topography only has variations in the zonal direction, with  $\partial_x h > 0$  reflecting deeper water to the east of the shelf. The inviscid vorticity balance (5.102) thus takes the form

$$\beta v = (f/h) u \partial_x h. \quad (5.103)$$

Since  $\beta v > 0$  and  $(f/h) \partial_x h > 0$ , we must have  $u > 0$ . Hence, flow departs from purely northward motion by leaving the shelf and moving into deeper waters. This behavior is also revealed by the western shelf example in Figure 5.5, where the  $f/H$  contours deviate from contours of constant depth when there is a nonzero  $\beta$ . As examined in this section, the bottom pressure torque, arising from bottom topographic form stress, plays the lead role in the corresponding vorticity balance.

#### Curl of the form stresses

In the present analysis, we are only concerned with that portion of  $\mathbf{u} \cdot \nabla h$  directly arising from pressure gradients. We thus consider just the geostrophic flow to arrive at

$$\mathbf{u} \cdot \nabla h = (g/f) \nabla h \cdot (\hat{z} \times \nabla \eta) = (g/f) \hat{z} \cdot (\nabla \eta \times \nabla h). \quad (5.104)$$

It is this term that contains the various pressure torques arising from form stresses. But before unpacking those torques, briefly return to the above shelf example and note that with  $\nabla h = \hat{x} \partial_x h$  then there is a nonzero  $\mathbf{u} \cdot \nabla h$  only with a nonzero  $\partial_y \eta$ , which in turn means there is a nonzero zonal geostrophic flow. As noted above, this zonal flow is a consequence of the sloping bottom topography.

#### Pressure torques

Returning to the general situation, we determine the expressions for the pressure torques contained in the term  $\mathbf{u} \cdot \nabla h$  by making use of the hydrostatic relation  $p_b = p_a + \rho g h$ , and the

layer thickness,  $h = \eta - \eta_b$  (see Figure 1.1), in which case

$$\mathbf{u} \cdot \nabla h = (g/f) \hat{\mathbf{z}} \cdot (\nabla \eta \times \nabla h) \quad (5.105a)$$

$$= 1/(\rho f) \hat{\mathbf{z}} \cdot [\nabla \times (p_a \nabla \eta) + \nabla \eta \times \nabla p_b] \quad (5.105b)$$

$$= 1/(\rho f) \hat{\mathbf{z}} \cdot [\nabla \times (p_a \nabla \eta) - \nabla \times (p_b \nabla \eta_b) + \nabla h \times \nabla p_b] \quad (5.105c)$$

$$= 1/(\rho f) \hat{\mathbf{z}} \cdot [\nabla \times (p_a \nabla \eta) - \nabla \times (p_b \nabla \eta_b) + \nabla p_b \times \nabla p_a / (\rho g)]. \quad (5.105d)$$

The first and second terms are the curls of the pressure form stress applied to the surface (atmospheric) and bottom boundaries (see the form stress chapter in VOLUME 2), where we refer to the atmospheric and bottom pressure torques. The third term is a torque arising from misalignment of the applied pressure and bottom pressure. Inserting equation (5.105d) in the vorticity balance (5.101) leads to the steady balance

$$\beta \rho h v = \hat{\mathbf{z}} \cdot [\nabla \times (p_a \nabla \eta) - \nabla \times (p_b \nabla \eta_b) + \nabla p_b \times \nabla p_a / (\rho g)] + \rho h [-\gamma \zeta + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}})]. \quad (5.106)$$

We consider a few special cases to see how the pressure torques affect the meridional flow in a gyre circulation, with a focus on bottom pressure torque since this term is generally far larger than those torques involving the atmospheric pressure.

### Inviscid balance between meridional flow and bottom pressure torque

In the absence of friction and wind forcing, and with a uniform atmospheric pressure ( $\nabla p_a = 0$ ), then  $\beta$  times the depth integrated meridional mass transport in equation (5.106) is balanced by the bottom pressure torque (i.e., curl of the topographic form stress)

$$\beta \rho h v = -\hat{\mathbf{z}} \cdot (\nabla \times p_b \nabla \eta_b) \quad \text{linear inviscid and unforced.} \quad (5.107)$$

Again consider the example of northward flow along a western continental shelf. In this case,  $\nabla \eta_b = \hat{\mathbf{x}} \partial_x \eta_b < 0$ , so that the balance (5.107) reduces to

$$\beta \rho h v = \partial_y p_b \partial_x \eta_b. \quad (5.108)$$

This balance says that if the bottom depth increases eastward away from the shelf, so that  $\partial_x \eta_b < 0$ , then northward flow corresponds to a northward decrease in the bottom pressure,  $\partial_y p_b < 0$ . That is, the northward flow is directed down the bottom pressure gradient.

### Association rather than causality

The inviscid planetary geostrophic balance (5.107) does not express causality. Rather, it expresses a balance or association that is maintained by steady linear flows in the presence of sloping topography. Hence, it is incorrect to say that bottom pressure torque gives rise to (i.e., causes) inviscid planetary geostrophic flow. Rather, the balance (5.107) says for planetary geostrophic flow, there is no bottom pressure torque without a meridional flow, and conversely there is no meridional flow without bottom pressure torque.

### Vanishing topographic form stress curl for geostrophic *f*-plane motion

From the balance (5.107), we find that a linear inviscid and unforced geostrophic flow on an *f*-plane satisfies

$$\hat{\mathbf{z}} \cdot (\nabla \times p_b \nabla \eta_b) = \hat{\mathbf{z}} \cdot (\nabla p_b \times \nabla \eta_b) = 0 \quad \beta = 0. \quad (5.109)$$

When we can ignore the applied surface pressure,  $p_a = 0$ , then the bottom pressure is given by  $p_b = g \rho h = g \rho (\eta - \eta_b)$ , in which case  $\hat{z} \cdot (\nabla p_b \times \nabla \eta_b) = \hat{z} \cdot (\nabla \eta \times \nabla \eta_b)$ . Evidently, linear inviscid and unforced geostrophic flow on an  $f$ -plane means that  $\hat{z} \cdot (\nabla \eta \times \nabla \eta_b) = 0$ . That is, isolines of surface height align with isolines of bottom topography. Conversely, surface height contours that deviate from bottom topography contours signal the role of friction and/or  $\beta$  acting on the planetary geostrophic flow.

### Local generation of bottom pressure torque

The beta effect provides an inviscid means to balance a misalignment between the surface height and the bottom topography, with misalignment required to generate a nonzero topographic form stress curl, which we have been referring to as the bottom pressure torque. We discussed an analogous misalignment in Section 2.7 when studying the force balances in a steady zonally re-entrant channel with bottom topography. Wind stress forcing and dissipation offer another means to balance  $(\eta, \eta_b)$ -misalignment, as seen merely by rewriting the vorticity balance (5.106) as an expression for the bottom pressure torque

$$\begin{aligned} \hat{z} \cdot (\nabla \times p_b \nabla \eta_b) = \\ -\beta \rho h v + \hat{z} \cdot [\nabla \times (p_a \nabla \eta) + \nabla p_b \times \nabla p_a / (\rho g)] + \rho h [-\gamma \zeta + \hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}})]. \end{aligned} \quad (5.110)$$

In our discussion of western intensification in Section 5.7.5, we ignored the role of bottom pressure torque. However, as seen by this balance, bottom pressure torque plays a role when topography and surface height are misaligned, with that role in some locations more important than friction. See [Becker and Salmon \(1997\)](#), [Hughes \(2000\)](#), [Hughes and de Cueves \(2001\)](#), [Jackson et al. \(2006\)](#) and [Patmore et al. \(2019\)](#) for examples.

### 5.7.7 Properties of area integrated bottom pressure torques

Consider a simply connected fluid domain bounded by an isobath. Stokes' theorem reveals that the integral of the bottom pressure torque,  $\hat{z} \cdot (\nabla \times p_b \nabla \eta_b)$ , vanishes when computed over the area bounded by any isobath

$$\int_S \hat{z} \cdot (\nabla \times p_b \nabla \eta_b) dS = \oint_{\partial S} p_b \nabla \eta_b \cdot \hat{\mathbf{t}} d\ell = 0. \quad (5.111)$$

The integrand for the line integral vanishes pointwise since, by construction, the boundary,  $\partial S$ , is determined by an isobath so its the tangent vector,  $\hat{\mathbf{t}}$ , is orthogonal to  $\nabla \eta_b$ . The identical argument holds for an area bounded by an isobar of bottom pressure since, for any region,

$$\int_S \hat{z} \cdot (\nabla \times p_b \nabla \eta_b) dS = - \int_S \hat{z} \cdot (\nabla \times \eta_b \nabla p_b) dS = - \oint_{\partial S} \eta_b \nabla p_b \cdot \hat{\mathbf{t}} d\ell. \quad (5.112)$$

Now if the region is bounded by an isobar, then  $\nabla p_b \cdot \hat{\mathbf{t}} = 0$  at each point along the region boundary. With a bit more work, we can show that the area integral of the bottom pressure torque vanishes for an annular region bounded by any two isobaths or any two bottom pressure isobars. The proof is presented in the caption to Figure 5.10.

These mathematical identities mean that bottom pressure torque plays no role in the area integrated vorticity budget for regions bounded by isobaths or bottom pressure isobars. In particular, taking a contour that encircles the global ocean reveals that bottom pressure torque does not alter the global area integrated vorticity. Equivalently, topographic form

stresses integrated around a closed contour defined by an isobath or isobar do not alter the circulation around that contour. Evidently, topographic form stress cannot spin-up or spin-down the circulation around a contour determined by isobaths or bottom pressure isobars. Other processes, such as flow nonlinearities, viscous friction, and/or boundary stresses, must play role in determining the circulation since their contour integrals are unconstrained. As studied by [Stewart et al. \(2021\)](#), these very special properties of bottom pressure torque and topographic form stress introduce a variety of nuances when interpreting the area integrated vorticity budget.

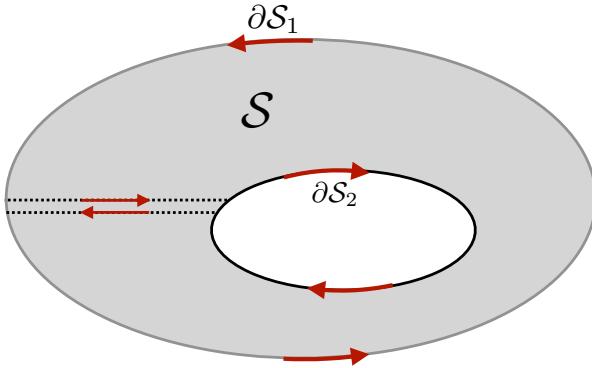


FIGURE 5.10: Stokes' theorem is generally applicable in a simply connected region. However, an annulus region,  $\mathcal{S}$ , is not simply connected since it has a hole. To apply Stokes' theorem over an annulus we imagine snipping the outer contour,  $\partial\mathcal{S}_1$ , and inner contour,  $\partial\mathcal{S}_2$ , and then connecting the two ends of the snipped contours as shown here. In this manner we convert the non-simply connected annulus into a simply connected region, over which we can apply Stokes' theorem in the naive manner. Importantly, we see that the two segments connecting the inner and outer contour are traversed in opposite directions when performing the contour integral. Assuming all functions are smooth, we can take the limit as these two contours get infinitesimally close, in which they cancel identically. This limit also recovers the connected outer and inner contours. If the outer and inner contours are defined by isobaths or bottom pressure isobars, then the bottom pressure torque identically vanishes on the contours. In this manner we have proven that the bottom pressure torque vanishes within the annulus region. We make use of this method in Exercise 6.10 when considering the circulation in an ocean domain with islands.

### 5.7.8 Advection-diffusion of the steady streamfunction<sup>6</sup>

Since the steady state flow satisfies  $\nabla \cdot (h \mathbf{u}) = 0$ , we can introduce a streamfunction

$$h \mathbf{u} = \hat{\mathbf{z}} \times \nabla \Psi, \quad (5.113)$$

in which case the relative vorticity becomes

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \nabla \cdot (h^{-1} \nabla \psi), \quad (5.114)$$

and the steady vorticity equation (5.88) can be written

$$\hat{\mathbf{z}} \cdot (\nabla \Psi \times \nabla Q) = -\gamma \nabla \cdot (h^{-1} \nabla \Psi) + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.115)$$

Following [Welander \(1968\)](#), we interpret equation (5.115) as a steady advection-diffusion equation for  $\Psi$ . Namely, introduce the horizontally non-divergent vector

$$\mathbf{u}^{(Q)} = \hat{\mathbf{z}} \times \nabla Q, \quad (5.116)$$

---

<sup>6</sup>We study the physics of advection and diffusion in VOLUME 1. For Section 5.7.8 we only require a few basic features of this equation.

in which case the streamfunction equation (5.115) becomes

$$\mathbf{u}^{(Q)} \cdot \nabla \Psi = \gamma \nabla \cdot (h^{-1} \nabla \Psi) - \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}}). \quad (5.117)$$

With zero wind stress curl and zero Rayleigh drag, the contours of the steady streamfunction are aligned with  $\mathbf{u}^{(Q)}$ ; that is,  $\Psi$  and  $Q$  contours are parallel. In the presence of Rayleigh drag, the streamfunction deviates from the  $Q$  contours, as it does in the presence of a wind stress curl.

These results represent a mere repackaging of results found earlier. Even so, the advection-diffusion interpretation offers complementary insights into the patterns of the steady streamfunction. Indeed, as we see in Section 5.9, the advective-diffusive interpretation proves very useful when studying patterns of steady sea level in planetary geostrophic gyres.

### 5.7.9 Comments and further study

#### Friction is needed to close the vorticity budget in a flat bottom linear ocean

This video [from SciencePrimer](#) provides a concise summary of the dynamics of ocean gyres and western boundary intensification due to the beta effect. Chapter 19 of [Vallis \(2017\)](#) provides a lucid treatment of ocean gyre dynamics by working through the key features of the [Stommel \(1948\)](#) model, as well as variants such as that from [Munk \(1950\)](#), who considered a viscous closure (most important next to side boundaries) rather than the Rayleigh drag used by Stommel. The Stommel model and its variants are themselves very idealized renditions of the ocean gyres occurring in Nature. Notable further factors become important in studying Nature's gyres, such as topography (briefly discussed in Section 5.7.6), flow nonlinearities and instabilities (recall we dropped the inertial term in Section 5.7.5, thus focusing on linear balances), turbulent boundary layers, and coupled air-sea processes. Each of these processes render the study of western boundary currents one of the most complex and timeless areas of ocean physics.

#### Distinguishing beta-induced western boundary currents from frictional boundary layers

Western intensification of the gyre, implied by  $\beta > 0$ , is distinct from a frictional boundary layer, implied by a [no-slip boundary condition](#). Namely, a frictional western boundary layer, supported by a no-slip solid earth boundary, leads to cyclonic flow within the boundary layer. In contrast, as discussed in this section, planetary  $\beta$ -induced western intensification gives rise to anti-cyclonic flow in the western boundary current. This distinction emphasizes the importance of the side boundary condition, be it no-slip, free-slip, or partial slip. Namely, the free-slip boundary does not support the cyclonic boundary layer of the no-slip. As explored in [Kiss \(2002\)](#) and [Kiss \(2004\)](#), these distinct boundary conditions lead to distinct boundary current properties, with impacts on the dynamics of how the boundary current separates from the side.

#### Topographic form stress reduces the fundamental role of friction

Friction plays a central role in the Stommel and Munk models of western intensification. As described in Chapter 19 of [Vallis \(2017\)](#), there have been attempts to produce an inviscid (and unforced) gyre solution, with the study from [Fofonoff \(1954\)](#) of particular note. We consider elements of the Fofonoff gyre in Exercise 4.11. Additionally, we raised the importance of sloping sides in Section 5.7.6 and in the discussion of axial angular momentum in VOLUME 2. Sloping sides enable bottom topographic form stress and bottom pressure torques to dominate over

bottom turbulent stresses and turbulent torques. In so doing, sloping sides play a leading role in gyre balances, though we did note in the discussion surrounding equation (5.111) that friction remains essential to close the vorticity balance integrated around the gyre. The role of sloping sides for gyre circulations was emphasized by [Hughes \(2000\)](#), [Hughes and de Cueves \(2001\)](#), with [Stewart et al. \(2021\)](#) identifying further nuances related to the integrated properties in Section 5.7.7.

### Flow nonlinearities also reduce the role of friction

Furthermore, we ignored the role of nonlinearities, which generally require the use of numerical models to investigate. Indeed, one of the first numerical simulations was from [Bryan \(1963\)](#), who showed that western boundary currents in a flat bottom gyre generally experience hydrodynamical instabilities when the flows become strong enough. [Becker and Salmon \(1997\)](#) and [Becker \(1999\)](#) further studied the case of a nonlinear gyre circulation in the presence of a sloping side shelf, thus allowing for the role of both nonlinearities and topographic form stress in the vorticity balance. These, and many other, studies allow for inviscid contributions to the vorticity balance, thus alleviating the need for invoking strong frictional effects found in the models of Stommel and Munk.

## 5.8 Column vorticity

Throughout this chapter we have focused on vorticity as defined by the curl of the horizontal velocity,  $\zeta = \hat{z} \cdot (\nabla \times \mathbf{u})$ . Since flow within a shallow water layer moves in coherent and extensible vertical columns (Section 1.2), the shallow water vorticity measures the spin of a shallow water column.

We can also measure the column spin by considering the curl of the thickness weighted velocity

$$\Sigma = \hat{z} \cdot \nabla \times (h \mathbf{u}) = h \zeta + \hat{z} \cdot \nabla h \times \mathbf{u}. \quad (5.118)$$

We observe that  $\Sigma$  provides a measure of the column vorticity, and it has dimensions of a velocity due to the thickness weighting. In addition to the thickness weighted relative vorticity, the column vorticity measures the misalignment between the layer thickness gradient and the velocity, with  $\hat{z} \cdot \nabla h \times \mathbf{u}$  generally nonzero especially for geostrophic flows. Hence, the second term offers an extra measure of the spin for the fluid column beyond the relative vorticity measure.<sup>7</sup> The layer thickness factor ensures that  $\Sigma$  directly probes stresses acting at the layer interfaces, as well as stresses within the fluid layer. Besides its intrinsic interest, this analysis is motivated by studies of the depth integrated flow commonly pursued in ocean circulation studies, with examples studied in Chapter 10. It also provides a venue for studying vorticity in those situations where the shallow water layer thickness vanishes, such as for a shallow water layer along a sloping side boundary.

### 5.8.1 Formulating the column vorticity equation

To formulate the dynamical equation for  $\Sigma$ , we start from the momentum equation (2.29)

$$\frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot (h \mathbf{u} \otimes \mathbf{u}) + f \hat{z} \times (h \mathbf{u}) = -g h \nabla \eta^{\text{eff}} + \mathbf{f}^{\text{nc}} \quad (5.119)$$

---

<sup>7</sup>Exercise 5.8 offers a particular example of a geostrophic flow that is aligned perpendicular to  $\nabla h$ , as found in a geostrophic shallow water front.

where

$$\mathbf{f}^{\text{nc}} = h \mathbf{F} \quad (5.120)$$

is the thickness weighted acceleration (dimension of squared velocity) arising from non-conservative processes, such as from horizontal strains in the presence of viscosity and the boundary transfer of turbulent momentum such as through winds and bottom drag (see Section 1.6). For example, recall the acceleration given by equation (5.81), which is built from Rayleigh drag plus wind stress, in which case

$$\mathbf{f}^{\text{nc}} = -\gamma h \mathbf{u} + \boldsymbol{\tau}^{\text{wind}}/\rho = -\gamma h \mathbf{u} + h \mathbf{F}^{\text{wind}}. \quad (5.121)$$

Taking the vertically projected curl of equation (5.119) leads to

$$\partial_t \Sigma + \hat{\mathbf{z}} \cdot \nabla \times [\nabla \cdot (h \mathbf{u} \otimes \mathbf{u})] + \nabla \cdot (f h \mathbf{u}) = -g \hat{\mathbf{z}} \cdot (\nabla h \times \nabla \eta^{\text{eff}}) + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{f}^{\text{nc}}). \quad (5.122)$$

Let us now examine each of these terms and offer physical interpretations.

### Frictional torques and boundary pressure torques

The term  $\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{f}^{\text{nc}})$  in equation (5.122) provides a torque from boundary stresses and interior viscous stresses. The term  $g \nabla h \times \nabla \eta^{\text{eff}}$  provides a torque whenever the thickness gradients are not aligned with the gradients in the effective surface height. We can write this term in the equivalent form based on the following identities

$$\rho g \nabla h \times \nabla \eta^{\text{eff}} = \rho g \nabla h \times \nabla \eta + \nabla h \times \nabla p_a \quad (5.123a)$$

$$= -\nabla p_a \times \nabla \eta + \nabla p_b \times \nabla \eta + \nabla(p_b - p_a) \times \nabla p_a / (\rho g) \quad (5.123b)$$

$$= -\nabla p_a \times \nabla \eta + \nabla p_b \times \nabla[\eta + p_a / (\rho g)] \quad (5.123c)$$

$$= -\nabla p_a \times \nabla \eta + \nabla p_b \times \nabla[\eta_b + h + p_a / (\rho g)] \quad (5.123d)$$

$$= -\nabla p_a \times \nabla \eta + \nabla p_b \times \nabla \eta_b \quad (5.123e)$$

$$= \nabla \times (-p_a \nabla \eta + p_b \nabla \eta_b), \quad (5.123f)$$

where we made use of the hydrostatic relation  $p_b = p_a + \rho g h$ . We thus see that  $-g \nabla h \times \nabla \eta^{\text{eff}} = \rho^{-1} \nabla \times (p_a \nabla \eta - p_b \nabla \eta_b)$  arises from pressure torques acting on the surface and bottom interfaces of the shallow water layer. These torques spin the column if there is a misalignment between the boundary pressure gradients and the boundary surface slopes. Notably, it is commonly the case that the torque associated with the applied surface pressure is far smaller than that from the bottom pressure, in which case

$$-g \nabla h \times \nabla \eta^{\text{eff}} \approx -\rho^{-1} \nabla \times (p_b \nabla \eta_b) = -\rho^{-1} \nabla p_b \times \nabla \eta_b. \quad (5.124)$$

### Torque from the nonlinear transport

The nonlinear term  $\hat{\mathbf{z}} \cdot \nabla \times [\nabla \cdot (h \mathbf{u} \otimes \mathbf{u})]$  can be written

$$\hat{\mathbf{z}} \cdot \nabla \times [\nabla \cdot (h \mathbf{u} \otimes \mathbf{u})] = \hat{z}_m \epsilon_{mst} \partial_s [\partial_n (h u_n u_t)] \quad (5.125a)$$

$$= \partial_n \partial_s [(h u_n) \epsilon_{mst} \hat{z}_m u_t] \quad (5.125b)$$

$$= \partial_n [\partial_s (h u_n) \epsilon_{mst} \hat{z}_m u_t + h u_n \zeta] \quad (5.125c)$$

$$= -\partial_n [\epsilon_{smt} \hat{z}_m u_t \partial_s (h u_n)] + \nabla \cdot (h \zeta \mathbf{u}), \quad (5.125d)$$

which exposes the divergence of the thickness weighted advective flux of relative vorticity. Further manipulations lead to

$$-\partial_n[\epsilon_{smt}\hat{z}_m u_t \partial_s(h u_n)] = -\partial_n[\epsilon_{smt}\hat{z}_m u_t u_n \partial_s h + \epsilon_{smt}\hat{z}_m u_t h \partial_s u_n] \quad (5.126a)$$

$$= \partial_n[\hat{z}_m(\epsilon_{smt}\partial_s h u_t) u_n - h \epsilon_{smt}\hat{z}_m u_t \partial_s u_n] \quad (5.126b)$$

$$= \nabla \cdot [\hat{\mathbf{z}} \cdot (\nabla h \times \mathbf{u}) \mathbf{u}] - \partial_n[h(\hat{\mathbf{z}} \times \mathbf{u}) \cdot \nabla u_n] \quad (5.126c)$$

$$= \nabla \cdot [\hat{\mathbf{z}} \cdot (\nabla h \times \mathbf{u}) \mathbf{u}] + \nabla \cdot [h(v \partial_x - u \partial_y) \mathbf{u}], \quad (5.126d)$$

so that the nonlinear term takes the form

$$\hat{\mathbf{z}} \cdot \nabla \times [\nabla \cdot (h \mathbf{u} \otimes \mathbf{u})] = \nabla \cdot [\Sigma \mathbf{u} + h(v \partial_x - u \partial_y) \mathbf{u}] = \nabla \cdot [\Sigma \mathbf{u} + h(\mathbf{u}_{\text{clock}} \cdot \nabla) \mathbf{u}]. \quad (5.127)$$

The first term inside the square bracket is the advective flux of  $\Sigma$ . The second term is the thickness weighted transport of  $\mathbf{u}$  by the clockwise rotated horizontal velocity

$$\mathbf{u}_{\text{clock}} = -\hat{\mathbf{z}} \times \mathbf{u}. \quad (5.128)$$

### 5.8.2 Summary of the column vorticity equation

The above manipulations bring the  $\Sigma$  equation (5.122) into the flux-form

$$\partial_t \Sigma = -\nabla \cdot \mathbf{J}^\Sigma + \rho^{-1} \hat{\mathbf{z}} \cdot \nabla \times [-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{\text{nc}}], \quad (5.129)$$

where

$$\mathbf{J}^\Sigma \equiv [h f + \Sigma + h(v \partial_x - u \partial_y)] \mathbf{u} = h \zeta_a \mathbf{u} + [v(\partial_x h + h \partial_x) - u(\partial_y h + h \partial_y)] \mathbf{u} \quad (5.130)$$

is the total flux of  $\Sigma$ . The corresponding material time derivative form of equation (5.129) is given by

$$\frac{D\Sigma}{Dt} + \Sigma \nabla \cdot \mathbf{u} + \nabla \cdot [h f \mathbf{u} + h(v \partial_x - u \partial_y) \mathbf{u}] = \rho^{-1} \hat{\mathbf{z}} \cdot \nabla \times [-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{\text{nc}}]. \quad (5.131)$$

The thickness equation (1.119)

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} + w^{(\dot{\eta})}, \quad (5.132)$$

brings equation (5.131) into the form

$$h \frac{D(\Sigma/h)}{Dt} + \nabla \cdot [h f \mathbf{u} - h \hat{\mathbf{z}} \cdot (\mathbf{u} \times \nabla) \mathbf{u}] = -\frac{\Sigma w^{(\dot{\eta})}}{h} + \rho^{-1} \hat{\mathbf{z}} \cdot \nabla \times [-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{\text{nc}}], \quad (5.133)$$

where we wrote

$$v \partial_x - u \partial_y = -\hat{\mathbf{z}} \cdot (\mathbf{u} \times \nabla). \quad (5.134)$$

### 5.8.3 Steady linear column vorticity balance and the island rule

To garner experience with the column vorticity equation (5.129), consider the steady state and assume  $\mathbf{J}^\Sigma$  is dominated by the planetary vorticity

$$\mathbf{J}^\Sigma \approx h f \mathbf{u}, \quad (5.135)$$

in which case the budget equation (5.129) reduces to the balance

$$\nabla \cdot (h f \mathbf{u}) = \rho^{-1} \hat{\mathbf{z}} \cdot \nabla \times [-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{nc}]. \quad (5.136)$$

Hence, torques due to interface pressures and turbulent stresses are balanced by the divergence of the thickness weighted advective flux of planetary vorticity. Furthermore, since  $\nabla \cdot (h \mathbf{u}) = 0$  in the steady state, equation (5.136) takes the form

$$\rho h \beta v = \hat{\mathbf{z}} \cdot \nabla \times (-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{nc}), \quad (5.137)$$

which connects to the discussion of pressure torques given in Section 5.7.6. It says that for planetary geostrophic flow, meridional transport in the presence of  $\beta \neq 0$  is balanced by the torques arising from atmospheric and bottom form stresses, as well as from non-conservative processes such as wind stress, viscous friction, and bottom drag.

### Formulating an integral balance

The local balance (5.136) holding for planetary geostrophic flow leads to an integral balance through an area integral computed over an arbitrary closed domain

$$\oint_S \nabla \cdot (h f \mathbf{u}) dS = \rho^{-1} \oint_S \hat{\mathbf{z}} \cdot \nabla \times [-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{nc}] dS. \quad (5.138)$$

Making use of Gauss' divergence theorem on the left hand side and Stokes' curl theorem on the right hand side renders

$$\oint_{\partial S} h f \mathbf{u} \cdot \hat{\mathbf{n}} d\ell = \rho^{-1} \oint_{\partial S} [-p_a \nabla \eta + p_b \nabla \eta_b + \rho \mathbf{f}^{nc}] \cdot \hat{\mathbf{t}} d\ell. \quad (5.139)$$

The left hand side is a contour integral with  $\hat{\mathbf{n}}$  the unit outward normal along the contour. The right hand side is also a contour integral, yet with  $\hat{\mathbf{t}}$  the unit tangent along the contour oriented in the counter-clockwise sense. The integral balance (5.139) says that the advective transport of planetary vorticity leaving the closed region (left hand side) is balanced by the oriented contour integral of the pressure and turbulent stresses.

The integral balance (5.139) is a rather remarkable statement that equates the transport leaving a region (left hand side) to pressure and turbulent stresses integrated along the region boundary (right hand side). For example, if the contour used to compute the integrals in equation (5.139) follows a closed streamline, on which  $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$ , then the left hand side vanishes, which in turn means that the right hand side contour integral must also vanish. [Godfrey \(1989\)](#) chose another contour, such as that depicted in Figure 5.11, and made some assumptions about the stresses along the contour. His assumptions allow for an estimate of the transport, with this estimate having proven quite useful for many purposes. We present the arguments next.

### Assumptions about the stresses

A portion of the red contour in Figure 5.11 traverses the eastern boundary on both Island A and Island B. As shown in Section 5.7, friction in the eastern boundary region of an ocean gyre is much less than the western boundary, motivating us to ignore friction (either viscous or bottom drag) along the eastern boundary portions of the contour. Likewise, friction is generally small for the open ocean portion along the two latitudinal lines. Hence, wind stress is the only non-conservative process that affects the right hand side of equation (5.139).

[Godfrey \(1989\)](#) furthermore ignored pressure form stresses and their associated torques. This assumption is reasonable for the atmospheric form stress,  $-p_a \nabla \eta$ , which is generally quite small (it is zero for a rigid lid approximation). However, the bottom form stress,  $p_b \nabla \eta_b$ , and the associated bottom pressure torque, can be larger than the wind stress and wind stress curl, especially near the continental margins and with strong currents. Hence, ignoring bottom pressure torque is an unsatisfying assumption.

### The island rule

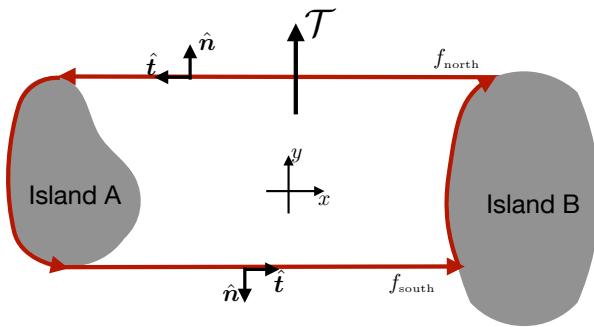


FIGURE 5.11: Illustrating the island rule with two land masses (“islands”). The red contour surrounds all of Island A and traverses along the eastern side of both Island A and Island B. The northern contour extends along a constant latitude line with a corresponding Coriolis parameter  $f_{\text{north}}$ , whereas the southern contour has Coriolis parameter  $f_{\text{south}}$ . Ignoring any precipitation or evaporation crossing the surface, steady state volume conservation means that the same meridional transport,  $\mathcal{T}$ , crosses both the southern and northern contours. Godfrey’s island rule (5.142) provides an estimate for this transport when given the wind stress along the contour.

The islands are material surfaces so that  $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$  along those portions of the contour that are adjacent to the coasts. Integrating along the northern latitude, with  $\hat{\mathbf{n}} = \hat{\mathbf{y}}$ , yields

$$f_{\text{north}} \int_{y_{\text{north}}} h v dx \equiv f_{\text{north}} \mathcal{T}, \quad (5.140)$$

where  $\mathcal{T}$  is the meridional transport. Ignoring any volume transport through the layer surface (e.g.,  $w^{(\dot{\eta})} = 0$ ), the steady transport crossing the northern boundary equals to that crossing the southern boundary, so that

$$f_{\text{south}} \int_{y_{\text{south}}} h v dx \equiv f_{\text{south}} \mathcal{T}. \quad (5.141)$$

Making use of these results in the integral balance (5.139), and following the above assumptions about the boundary stresses, leads to [Godfrey’s island rule](#)

$$\rho \mathcal{T} = \frac{1}{f_{\text{north}} - f_{\text{south}}} \oint_{\partial S} \boldsymbol{\tau}^{\text{wind}} \cdot \hat{\mathbf{t}} d\ell, \quad (5.142)$$

where  $\boldsymbol{\tau}^{\text{wind}}$  is the wind stress, and where the minus sign on  $f_{\text{south}}$  arises since  $\hat{\mathbf{n}} = -\hat{\mathbf{y}}$  along the southern latitude contour. This expression provides an approximation to the flow around Island A, noting that the only nonzero flow normal to the contour is through the two latitudinal segments.

### 5.8.4 Further study

*Godfrey* (1989) applied the island rule (5.142) to estimate transport around Austral-Asia, New Zealand, and Malagasy. Further discussion of Godfrey's island rule can be found in *Tomczak and Godfrey* (1994), *Pedlosky et al.* (1997), and *Klinger and Haine* (2019).

## 5.9 Free surface patterns in steady ocean gyres

We here consider basic features of steady free surface patterns realized by a shallow water fluid. The basic question is how to relate the free surface near the coast to that in the interior of an ocean gyre. We formulate the steady linear case with sloping bottom topography.

### 5.9.1 Formulating the free surface equation

We make use of the steady frictional geostrophic equations, with the velocity equation given by the linearized version of equation (5.119)

$$f \hat{z} \times h \mathbf{u} = -g h \nabla \eta + \mathbf{f}^{\text{nc}}. \quad (5.143)$$

Here,  $\mathbf{f}^{\text{nc}}$  is the thickness weighted acceleration from non-conservative processes, such as the wind stress, viscous friction, and bottom drag, with details provided in Section 5.8.1. Since the flow is steady, the layer thickness equation (1.20) leads to the non-divergence condition

$$\nabla \cdot (h \mathbf{u}) = 0. \quad (5.144)$$

We now formulate a vorticity-like equation. Yet rather than taking the curl of equation (5.143), thus emulating the work from Section 5.8, we first divide equation (5.143) by the Coriolis parameter to write

$$\hat{z} \times h \mathbf{u} = -(g h/f) \nabla \eta + \mathbf{f}^{\text{nc}}/f, \quad (5.145)$$

so that the curl leads to

$$\hat{z} \cdot [\nabla(g h/f) \times \nabla \eta] = \hat{z} \cdot \nabla \times (\mathbf{f}^{\text{nc}}/f), \quad (5.146)$$

where we used the non-divergence condition (5.144). Dividing by the Coriolis parameter limits the analysis to regions bounded away from the equator, which is not a problem for our focus on middle or high latitude planetary geostrophic ocean gyres. We are motivated to pursue this formulation since it supports the use of equation (5.146) next to sloping boundaries, in which the layer thickness vanishes ( $h = 0$ ) at the shoreline edge of the domain.

### 5.9.2 Advection the free surface

To help interpret the conservative portion of the free surface equation (5.146), we introduce the streamfunction<sup>8</sup>

$$\psi^{(\eta)} \equiv -g h/f, \quad (5.147)$$

<sup>8</sup>We follow the sign convention of *Wise et al.* (2018) by introducing  $\psi^{(\eta)} = -g h/f$  in equation (5.147). The alternative choice of  $\psi^{(\eta)} = +g h/f$  is less convenient for interpreting the direction of the velocity induced by topographic slopes. Additionally, in equation (5.155) we introduce Rayleigh drag and interpret the resulting equation as an advective-diffusion balance. If we take  $\psi^{(\eta)} = +g h/f$  then it would be an advective-anti-diffusive balance, which is far less physically satisfying.

so that equation (5.146) takes the form

$$\mathbf{u}^{(\eta)} \cdot \nabla \eta = -\hat{\mathbf{z}} \cdot \nabla \times (\mathbf{f}^{\text{nc}}/f), \quad \text{with} \quad \mathbf{u}^{(\eta)} = \hat{\mathbf{z}} \times \nabla \psi^{(\eta)} = -\hat{\mathbf{z}} \times \nabla(g h/f). \quad (5.148)$$

The left hand side is written as the advection of the free surface by the horizontally non-divergent velocity,  $\mathbf{u}^{(\eta)}$ . That is, the velocity,  $\mathbf{u}^{(\eta)}$ , advects the free surface, whereas the fluid velocity,  $\mathbf{u}$ , advects matter. In the absence of non-conservative forces ( $\mathbf{f}^{\text{nc}} = 0$ ), the free surface is aligned with isolines of  $h/f = 1/Q^{\text{pg}}$ , or equivalently with lines of constant planetary geostrophic potential vorticity,  $Q^{\text{pg}}$  (recall that  $Q^{\text{pg}} = f/h$  as defined by equation (5.89)). For free surface contours to deviate from  $Q^{\text{pg}}$  contours requires non-conservative forces, such as those from wind stresses and frictional dissipation. This discussion is related to that given in Section 5.5.2 where we introduced the notion of  $f/H$  contours for planetary geostrophic flow.

### The advection operator

The layer thickness is the difference between the free surface and bottom topography,  $h = \eta - \eta_b \geq 0$  (see Figure 1.1), in which case the advection operator takes on the form

$$\nabla \psi^{(\eta)} \times \nabla \eta = -(g/f) \nabla(\eta - \eta_b) \times \nabla \eta + (g h/f^2) \beta \hat{\mathbf{y}} \times \nabla \eta \quad (5.149\text{a})$$

$$= (g/f) [\nabla \eta_b + (h \beta/f) \hat{\mathbf{y}}] \times \nabla \eta. \quad (5.149\text{b})$$

This result motivates decomposing  $\mathbf{u}^{(\eta)}$  into two terms

$$\mathbf{u}^{(\text{b})} = (g/f) \hat{\mathbf{z}} \times \nabla \eta_b = (g/f) (-\hat{\mathbf{x}} \partial_y \eta_b + \hat{\mathbf{y}} \partial_x \eta_b) \quad (5.150\text{a})$$

$$\mathbf{u}^{(\beta)} = \hat{\mathbf{z}} \times (g h \beta/f^2) \hat{\mathbf{y}} = -(g h \beta/f^2) \hat{\mathbf{x}}, \quad (5.150\text{b})$$

so that the free surface equation (5.148) takes on the form

$$(\mathbf{u}^{(\text{b})} + \mathbf{u}^{(\beta)}) \cdot \nabla \eta = -\hat{\mathbf{z}} \cdot \nabla \times (\mathbf{f}^{\text{nc}}/f). \quad (5.151)$$

As so defined,  $\mathbf{u}^{(\text{b})}$  arises from gradients in the bottom topography and it is directed with shallow water to the right in the northern hemisphere and to the left in the southern hemisphere, which is the same orientation for the phase velocity in a coastally trapped Kelvin wave (VOLUME 4). The velocity,  $\mathbf{u}^{(\beta)}$ , arises from the planetary vorticity gradient and is always directed to the west, just as for the phase velocity of planetary Rossby waves (VOLUME 4). Notably,  $\mathbf{u}^{(\eta)}$  has a meridional component only so long as there is a zonal topographic slope,  $\partial_x \eta_b \neq 0$ .

### Free surface trajectories in the absence of non-conservative forcing

In the absence of non-conservative forces the steady free surface equation (5.151) is

$$(\mathbf{u}^{(\text{b})} + \mathbf{u}^{(\beta)}) \cdot \nabla \eta = 0 \implies [\hat{\mathbf{z}} \times \nabla \eta_b - (h \beta/f) \hat{\mathbf{x}}] \cdot \nabla \eta = 0. \quad (5.152)$$

For the case of a flat bottom domain, so that  $\mathbf{u}^{(\text{b})} = 0$ , free surface contours are purely zonal ( $\partial_x \eta = 0$ ). Conversely, on an  $f$ -plane, so that  $\mathbf{u}^{(\beta)} = 0$ , free surface contours are aligned with the bottom topography,  $\hat{\mathbf{z}} \cdot (\nabla \eta_b \times \nabla \eta) = 0$ . In the general case, the unforced steady free surface contours are aligned according to the interplay between planetary beta and bottom topography according to equation (5.152).

It is instructive to imagine a free surface ‘‘trajectory’’ as defined by integral curves of  $\mathbf{u}^{(\eta)}$ . Consider one such trajectory that moves from the interior of a northern hemisphere gyre onto

the continental shelf, such as in the example depicted in Figure 5.5. In the interior we assume the bottom topography is nearly flat, so that  $\mathbf{u}^{(b)} \approx 0$  and the corresponding free surface trajectory is along a constant latitude line. As we reach the continental slope, the trajectory becomes southwestward with a trajectory slope determined by the zonal slope of the topography. Finally, as the continental shelf is reached, which is relatively flat, the trajectory returns to a near constant latitude contour.

### 5.9.3 Rayleigh drag and free surface diffusion

The free surface equation (5.151) indicates that any non-conservative acceleration,  $\mathbf{f}^{\text{nc}}$ , causes the free surface to deviate from contours of constant streamfunction,  $\psi^{(\eta)} = -g h/f$ . The dominant contribution to  $\mathbf{f}^{\text{nc}}$  arises from wind stress. Here, we consider the role of a thickness weighted Rayleigh drag assumed to take the form given by equation (5.121)

$$\mathbf{f}^{\text{nc}} = -\gamma h \mathbf{u}_g, \quad (5.153)$$

where  $\gamma$  has dimensions of inverse time and  $\mathbf{u}_g$  is the geostrophic flow given by

$$\mathbf{u}_g = (g/f) \hat{\mathbf{z}} \times \nabla \eta. \quad (5.154)$$

Substituting into the free surface equation (5.151) leads to

$$\underbrace{(\mathbf{u}^{(b)} + \mathbf{u}^{(\beta)}) \cdot \nabla \eta}_{\text{advection}} = \underbrace{\nabla \cdot [(\gamma h g/f^2) \nabla \eta]}_{\text{diffusion}} - \underbrace{\hat{\mathbf{z}} \cdot \nabla \times [\boldsymbol{\tau}^{\text{wind}} / (\rho f)]}_{\text{source from } \nabla \times \boldsymbol{\tau}^{\text{wind}} / f}. \quad (5.155)$$

In the absence of a wind stress, the free surface satisfies a steady advective-diffusive balance.<sup>9</sup> Evidently, the Rayleigh drag, acting as a sink to the geostrophic momentum, also acts to diffuse the free surface and so to cause the free surface contours to deviate or spread out from contours of constant  $g h/f$ . Notice that the diffusion coefficient,

$$\kappa = \gamma h g/f^2 = -(\gamma/f) \psi^{(\eta)}, \quad (5.156)$$

gets larger moving toward the equator. Hence, free surface contours can more readily deviate from contours of constant  $\psi^{(\eta)}$  when moving towards the equator.

### 5.9.4 Further study

Inspiration for this section follows from [Minobe et al. \(2017\)](#), [Wise et al. \(2018\)](#), [Wise et al. \(2020a\)](#) and [Wise et al. \(2020b\)](#). Also note that the advection-diffusion interpretation of the streamfunction equation follows [Welander \(1968\)](#), and was already introduced in Section 5.7.8 and used in Section 4.8 for the study of beta plumes. We formulated the steady linear case following [Wise et al. \(2018\)](#), who considered sloping bottom topography rather than the vertical walls assumed by [Minobe et al. \(2017\)](#). The transient case, which involves coastal boundary waves, is considered by [Hughes et al. \(2019\)](#), [Wise et al. \(2020a\)](#) and [Wise et al. \(2020b\)](#).



<sup>9</sup>We study the advection and diffusion of scalar fields in VOLUME 1 and VOLUME 4.

## 5.10 Exercises

### EXERCISE 5.1: FLOW NEAR A TOPOGRAPHIC BUMP IN A REDUCED GRAVITY MODEL

Elements of this exercise are motivated by Figure 2 from [Adcock and Marshall \(2000\)](#) and Figure 1 from [Marshall et al. \(2012\)](#), where we consider a reduced gravity model with a dynamic lower layer and stagnant upper layer. Place a topographic bump (e.g., seamount or mountain) fully within the lower layer as shown in Figure 5.12.

- Following Exercise 1.5, derive the momentum and thickness equations for a reduced gravity model with a stagnant upper layer. Then derive the potential vorticity equation for the dynamical layer.
- Now assume an  $f$ -plane. If the lower layer potential vorticity is (somehow) horizontally homogenized (e.g., some form of mixing produces a horizontally constant potential vorticity), draw the resulting layer interface  $\eta_{3/2}$ . Assume the relative vorticity is negligible compared to the planetary vorticity so that the flow satisfies the planetary geostrophic scaling introduced in Section 5.5.2 and further pursued in Section 9.4. Also, ignore any non-steady processes; we are only interested here in the steady flow.
- For the case of horizontally homogenized potential vorticity from the previous part, what is the direction for a geostrophically balanced flow: cyclonic or anti-cyclonic? Hint: make use of Exercise 1.5 for the momentum equation of an inverted reduced gravity model.

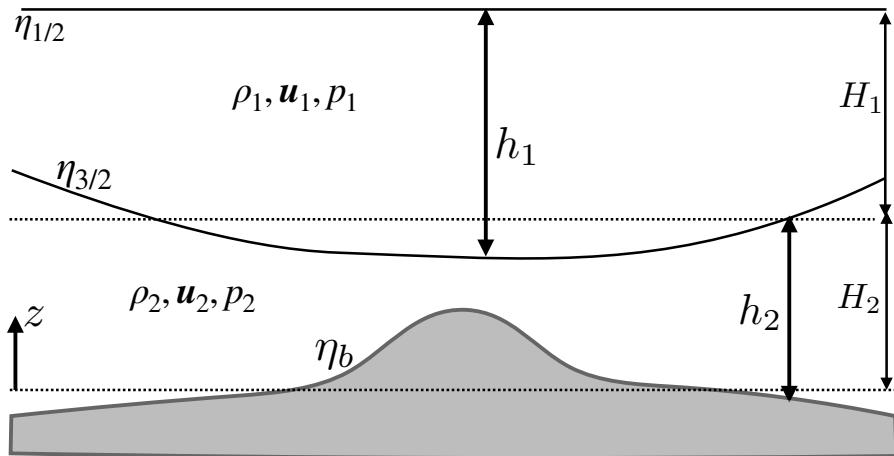


FIGURE 5.12: A reduced gravity model (see Section 1.3) with a stagnant upper layer and a dynamic lower layer as in Exercise 1.5. A seamount sits fully within the lower layer.

### EXERCISE 5.2: POTENTIAL VORTICITY FOR TWO SHALLOW WATER LAYERS

Consider the inviscid Boussinesq two-layer shallow water model as discussed in Section 1.4. Derive the potential vorticity equation for each layer, showing the mathematical steps used in the derivation. Hint: the answer is given in Section 5.3.6.

### EXERCISE 5.3: AVERAGE VORTICITY IN A SHALLOW WATER LAYER

Consider a single layer of shallow water fluid on a rotating plane with rotation rate  $\Omega = \hat{z} \Omega$ . Assume the fluid is contained in an arbitrary horizontal region and that it has a constant total volume given by

$$\mathcal{V} = \int \left[ \int dz \right] d\mathcal{S} = \int h d\mathcal{S} = \int (H + \Delta\eta - \eta_b) d\mathcal{S} = H \mathcal{S}, \quad (5.157)$$

where  $\mathcal{S}$  is the horizontal area of the domain,  $h(x, y, t) = H + \Delta\eta(x, y, t) - \eta_b(x, y)$  is the layer thickness,  $H$  is the resting depth relative to  $z = 0$ ,  $\Delta\eta$  is the sea level deviation from resting, and  $\eta_b$  is the undulation of the bottom topography (see Figure 1.1). Additionally, recall that  $z = 0$  is set according to

$$\int \eta_b d\mathcal{S} = 0, \quad (5.158)$$

and volume conservation ensures that

$$\int \Delta\eta d\mathcal{S} = 0. \quad (5.159)$$

- (a) Determine the volume average of the vorticity  $\hat{\mathbf{z}} \cdot \boldsymbol{\omega}_{\text{rigid}}$  arising from the rigid-body rotation

$$\langle \hat{\mathbf{z}} \cdot \boldsymbol{\omega}_{\text{rigid}} \rangle = \mathcal{V}^{-1} \int \hat{\mathbf{z}} \cdot \boldsymbol{\omega}_{\text{rigid}} dV. \quad (5.160)$$

- (b) Determine the area average of the relative vorticity,

$$\bar{\zeta} = \mathcal{S}^{-1} \int \hat{\mathbf{z}} \cdot \boldsymbol{\omega} d\mathcal{S}, \quad (5.161)$$

in terms of the circulation around the boundary of the domain.

- (c) Determine the volume average of the relative vorticity

$$\langle \zeta \rangle = \mathcal{V}^{-1} \int \hat{\mathbf{z}} \cdot \boldsymbol{\omega} dV. \quad (5.162)$$

Write the expression in terms of the area average vorticity,  $\bar{\zeta}$ , and the correlation,  $\overline{\zeta' h'}$ , where primes are deviations from the area mean.

#### EXERCISE 5.4: APPLICATION OF THE MATERIAL INVARIANCE OF POTENTIAL VORTICITY

In an adiabatic shallow water fluid in a rotating reference frame, show that the potential vorticity satisfies

$$\frac{D}{Dt} \left[ \frac{\zeta + f}{\eta - \eta_b} \right] = 0, \quad (5.163)$$

where  $\eta$  is the height of the free surface and  $\eta_b$  is the height of the bottom topography (see Figure 1.1). For both of the following questions, assume constant volume for the fluid column. Also, assume the column rotates coherently about its axis.

- (a) A cylindrical column of air at  $30^\circ$  latitude with radius 100 km expands horizontally to twice its original radius. If the air is initially at rest, what is the mean tangential velocity at the perimeter after the expansion?
- (b) An air column at  $60^\circ\text{N}$  with zero relative vorticity ( $\zeta = 0$ ) stretches from the surface to the tropopause, which we assume is a rigid lid at 10 km. The air column moves zonally onto a plateau 2.5 km high. What is its relative vorticity? Suppose it then moves southward along the plateau to  $30^\circ\text{N}$ , starting from the relative vorticity it obtained from the plateau. What is its new relative vorticity?

#### EXERCISE 5.5: APPLICATION OF THE MATERIAL INVARIANCE OF POTENTIAL VORTICITY

An air column at  $60^\circ\text{N}$  with  $\zeta = 0$  initially reaches from the surface to a fixed tropopause at 10 km height. If the air column moves across a mountain 2.5 km high at  $45^\circ\text{N}$ , what is its absolute vorticity and relative vorticity as it passes the mountaintop? Hint: Use the material

invariance of shallow water potential vorticity, and assume the top of the column remains at 10 km.

**EXERCISE 5.6: STEADY SHALLOW WATER FLOW IN A ROTATING CHANNEL**

In Figure 5.13 we depict a single layer of shallow water fluid moving within a zonal channel on a northern hemisphere  $f$ -plane with meridional extent  $y = y_0 - L/2$  and  $y = y_0 + L/2$ , where  $y_0$  is the latitude of the  $f$ -plane. The fluid moves in the zonal direction under the influence of an imposed pressure gradient and under a rigid lid at  $z = H$ . The lower boundary that starts at  $z = 0$  and transitions to  $z = d$  at some upstream position. Assume the fluid moves without frictional dissipation so that all solid surfaces are free-slip. As a hint to this exercise, refer to Section 2.4 of *Stern (1975)*.

- Assuming a non-rotating reference frame, determine the velocity in the far downstream where the fluid thickness is  $H - d$ .
- Now assume an  $f$ -plane and determine the meridional profile of the zonal velocity at a distance far downstream.

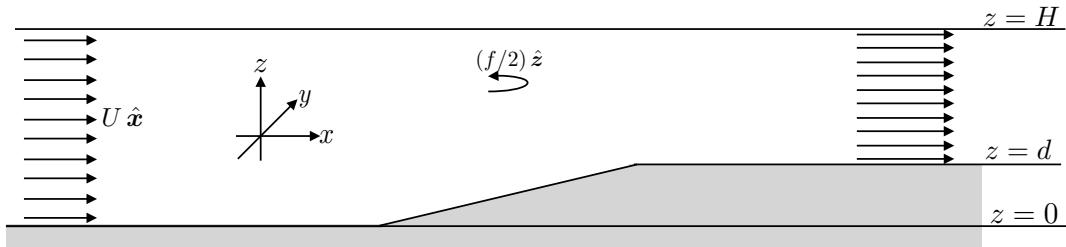


FIGURE 5.13: A single layer of shallow water fluid flows in a free-slip zonal channel on a northern hemisphere  $f$ -plane with vertical walls at  $y = y_0 \pm L/2$  and rigid upper surface at  $z = H$ . An imposed pressure gradient causes the fluid to flow from left to right, with  $\mathbf{u} = U \hat{x}$  the velocity in the far upstream direction given.

**EXERCISE 5.7: PROPERTIES OF THE STEADY STATE SHALLOW WATER FLUID**

Consider a single layer of shallow water fluid in steady state (i.e., all Eulerian time derivatives vanish).

- Show that there exists a streamfunction for the steady state thickness weighted horizontal flow

$$h \mathbf{u} = \hat{z} \times \nabla \Psi. \quad (5.164)$$

- What are the physical dimensions of  $\Psi$ ?
- Show that the shallow water potential vorticity is a constant along the steady state streamlines of the thickness weighted flow

$$Q = Q(\Psi). \quad (5.165)$$

- Show that the Bernoulli function,

$$B = g \eta + \mathbf{u} \cdot \mathbf{u}/2 \quad (5.166)$$

is also a constant along the same streamlines; i.e.,

$$B = B(\Psi). \quad (5.167)$$

- What is the functional relation between the Bernoulli function and the potential vorticity?

Hint: make use of the physical dimensions for  $Q$ ,  $B$ , and  $\Psi$  to be sure that your solution is dimensionally consistent.

#### EXERCISE 5.8: ZONALLY SYMMETRIC SHALLOW WATER FRONT

Consider a single layer of shallow water fluid on a  $\beta$ -plane ( $f = f_0 + \beta y$ ) with a flat bottom. Assume all fields possess zonal symmetry as in the zonal front shown in Figure 5.14. Since the zonal pressure gradient vanishes under the assumption of zonal symmetry, the geostrophic portion of the meridional velocity vanishes. However, there is generally a non-zero ageostrophic component to this velocity, and we retain that possibility throughout this exercise.

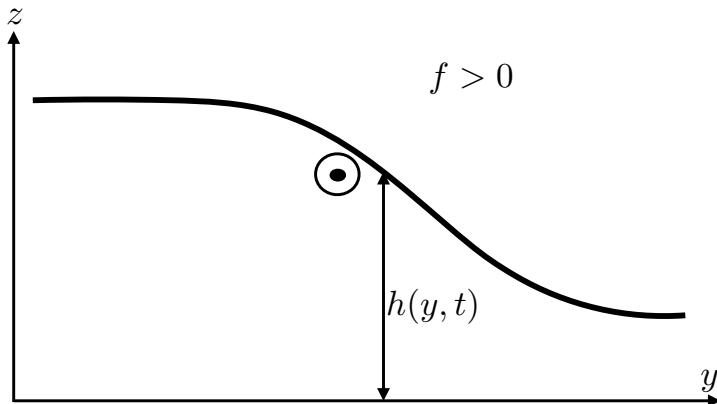


FIGURE 5.14: Schematic of a zonally symmetric front in a shallow water layer in the northern hemisphere ( $f > 0$ ), here used for Exercise 5.8. The thickness decreases to the north. If the zonal flow is in geostrophic balance, then the northward pressure gradient is in geostrophic balance with a southward Coriolis acceleration arising from an eastward (out of the page) geostrophic current (see also Figure 2.1). The geostrophic component of the meridional flow vanishes due to zonal symmetry, but the full meridional flow need not vanish if there are ageostrophic processes. We allow for the possibility of a nonzero meridional velocity throughout this exercise.

- Write the potential vorticity,  $Q$ , assuming the zonal flow is in geostrophic balance. Write  $Q$  terms of meridional derivatives of the layer thickness.
- From the shallow water equations, explicitly show that the potential vorticity is materially constant (i.e., it is a Lagrangian invariant). To do so, work through the usual shallow water potential vorticity material conservation derivation yet make use of the zonally symmetric equations of motion. Allow for a nonzero ageostrophic meridional flow. Show all relevant steps.
- Show that the potential vorticity can be written as  $Q = -(\partial_y M)/h$ , where  $h$  is the layer thickness. What is the expression for  $M$ ? Hint: recall our discussion of **potential momentum** in VOLUME 1.
- Potential vorticity is not the only material constant for this system. Due to the zonal symmetry, **Noether's theorem** indicates there is another. Show that  $M$  is materially constant. Again, continue to allow for a nonzero meridional velocity component.

#### EXERCISE 5.9: RAYLEIGH DRAG AND GALILEAN INVARIANCE

In VOLUME 2, we discuss **Galilean covariance** for the fluid mechanical equations. Is the **Rayleigh drag** used in equation (5.81) Galilean invariant? Why? If not, then should that be of concern for its use in studying flow in a closed and bounded domain?

#### EXERCISE 5.10: GYRES IN THE PRESENCE OF CYCLONIC WINDS

Consider the ocean gyre discussion in Section 5.7. Rather than anti-cyclonic winds, now apply a cyclonic wind stress to the domain. As per the circulation theorem (5.84), the gyre flow will have a cyclonic sense. Will the resulting gyre exhibit eastern intensification or western? Appeal

to whatever arguments you wish.

**EXERCISE 5.11: SHALLOW WATER EQUATIONS WITH DIVERGENCE-DAMPING**

When breaking the continuous symmetry of the equations of motion, a discretized numerical simulation admits unphysical flow features sometimes referred to as [computational modes](#). Some computational modes can evolve in time with energy accumulating at high wave numbers, in which case the numerical simulation produces unphysical grid noise and becomes of little physical use. To suppress grid noise, numerical models commonly introduce numerical dissipation, even if the continuous equations have zero dissipation. The formulation of numerical dissipation is largely an art guided by the dual needs of suppressing grid noise without otherwise damaging physical properties of the simulated flow. We here consider physical properties of a specific form of numerical dissipation known as [divergence damping](#). We work within the framework of the continuous equations so to develop generic physical properties of the divergence-damping operator. No knowledge of numerical methods is required to solve this problem.

Divergence-damping is motivated by the desire to leave the vorticity equation untouched while damping divergent motion that can arise in numerical simulations. This motivation is based on noting that much of the large-scale circulation in a rotating fluid has a nontrivial absolute vorticity yet a relatively small horizontal divergence. For example, geostrophic flow on an  $f$ -plane has vorticity dominated by planetary vorticity  $f$ , while it has zero horizontal divergence (see our discussion of geostrophic in [VOLUME 2](#), or the two-dimensional barotropic vorticity equation in [Section 4.2](#)). The divergence-damping operator is thus designed to reduce the magnitude of the horizontal divergence while leaving the vorticity untouched.

We here examine the impacts of divergence-damping on mechanical energy and angular momentum. For this purpose, consider a single layer of shallow water fluid with divergence-damping. This system is described by the momentum and thickness equations

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla(g\eta + \alpha\Gamma) \quad (5.168a)$$

$$\frac{Dh}{Dt} = -h\nabla \cdot \mathbf{u}. \quad (5.168b)$$

The parameter  $\alpha > 0$  is a constant and the field  $\Gamma$  is given by the Laplacian of the horizontal flow divergence

$$\Gamma = \nabla^2 \mathcal{D}, \quad (5.169)$$

where

$$\mathcal{D} = \nabla \cdot \mathbf{u}. \quad (5.170)$$

The divergence has physical dimensions of inverse time ( $T^{-1}$ ), so that its Laplacian,  $\Gamma$ , has dimensions of  $L^{-2} T^{-1}$ , and the coefficient  $\alpha$  has dimensions  $L^4 T^{-1}$ .

Divergence damping leads to a modification to the horizontal pressure gradient. We may think of this modification as arising from the horizontal gradient of a modified free surface height

$$\tilde{\eta} = \eta + \frac{\alpha\Gamma}{g}. \quad (5.171)$$

Notably, mass conservation remains the same since the thickness equation is unchanged. Hence, momentum evolution is modified by changing the pressure gradient, yet the thickness equation remains the same.

- (a) Show that the vorticity equation (5.9) remains unchanged in the presence of divergence-damping.

- (b) Show that the potential vorticity equation (5.29) remains unchanged in the presence of divergence-damping.  
 (c) Show that the horizontal divergence evolves according to

$$\frac{\partial \mathcal{D}}{\partial t} = \left[ \frac{\partial \mathcal{D}}{\partial t} \right]_{\alpha=0} - \alpha \nabla^2 \Gamma. \quad (5.172)$$

- (d) Show that the evolution of gravitational potential energy per horizontal area

$$\mathcal{P} = g \rho \int_{\eta_b}^{\eta} z \, dz \quad (5.173)$$

remains unchanged from that determined in Section 2.5.1.

- (e) Show that the kinetic energy per horizontal area evolves according to

$$\frac{\partial \mathcal{K}}{\partial t} + \nabla \cdot (\mathbf{u} \mathcal{K}) = -h \rho g \mathbf{u} \cdot \nabla \tilde{\eta}, \quad (5.174)$$

where

$$\mathcal{K} = \frac{1}{2} \int_{\eta_b}^{\eta} \rho \mathbf{u}^2 \, dz = \rho h \mathbf{u}^2 / 2, \quad (5.175)$$

is the horizontal kinetic energy per area (Section 2.5.2).

- (f) Determine the evolution equation for global integrated kinetic energy

$$\frac{\partial}{\partial t} \left[ \int \mathcal{K} \, dA \right] = \frac{\partial}{\partial t} \left[ \int \int_{\eta_b}^{\eta} (\rho \mathbf{u} \cdot \mathbf{u} / 2) \, dz \, dA \right]. \quad (5.176)$$

Hint: drop all lateral boundary terms by assuming either solid lateral walls or periodicity.

- (g) Consider a single shallow water layer in a rotating tank as in Section 2.8. Show that the material evolution of angular momentum relative to the vertical rotational axis is given by

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g \frac{\partial \eta}{\partial \phi} + \mathcal{T}. \quad (5.177)$$

What is the mathematical form for  $\mathcal{T}$ ? Hint: check your answer with the next part of this exercise.

- (h) Show that the domain integrated angular momentum satisfies the equation

$$\frac{\partial}{\partial t} \int L^z = \alpha \rho \int \Gamma \frac{\partial \eta}{\partial \phi} \, dA. \quad (5.178)$$

where we assume the bottom topography is flat so that  $h = \eta$ .

- (i) The linearized thickness equation for a flat bottom is given by

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0, \quad (5.179)$$

where  $H$  is the thickness of the resting fluid layer. Show that the time change for the global integrated angular momentum is given by

$$\frac{\partial}{\partial t} \int L^z = -\frac{\alpha \rho}{H} \int \left[ \frac{\partial}{\partial t} \nabla^2 \eta \right] \frac{\partial \eta}{\partial \phi} \, dA. \quad (5.180)$$



## Chapter 6

# VORTICITY AND CIRCULATION MECHANICS

In this chapter we study the kinematics and dynamics of vorticity and circulation, extending our introductory study in earlier chapters in this part of the book. An understanding of vorticity mechanics offers many insights into the nature of fluid flow and how that flow is constrained. In making use of vorticity for geophysical fluids, it can be useful to move beyond the vorticity of a fluid element as defined by the curl of the velocity field. For example, as a means to summarize facets of the vorticity contained within a three-dimensional fluid, we study the depth integral of the vorticity equation in Section 6.8 for a hydrostatic and Boussinesq fluid and in Section 10.4 for a planetary geostrophic flow. This analysis is particularly useful in understanding facets of ocean circulation. Relatedly, there are occasions to study vorticity of the depth integrated flow (i.e., vorticity of the transport) or depth averaged flow (i.e., vorticity of the barotropic flow), and we do so in Section 6.9. Such studies emphasize the importance of boundary forces and their curls (“torques”) for the vorticity of a fluid column. Details of the terms affecting such vorticities depend on the form of the vorticity, with somewhat complementary features emphasized.

### CHAPTER GUIDE

This chapter assumes an understanding of vorticity developed in earlier chapters in this part of the book. We also make use of fluid kinematics and fluid dynamics from VOLUME 2. As for the shallow water vorticity discussed in Chapter 5, we here make use of vector calculus identities from VOLUME 1. The concepts and methods developed in this chapter are fundamental to the notions of vorticity and circulation, with elements encountered in the remainder of this part of the book as well as in the study of balanced models in Part III.

Throughout this chapter, when considering spherical geometry, we make use of the [planetary Cartesian coordinates](#) from VOLUME 2. Since the sphere is assumed to be embedded in Euclidean space, we can naively perform integrals of vectors over the Euclidean space.

<b>6.1</b>	<b>Loose threads</b>	<b>211</b>
<b>6.2</b>	<b>Kelvin's circulation theorem</b>	<b>211</b>
6.2.1	Formulation	211
6.2.2	Two processes affecting circulation	212
6.2.3	Barotropic flow	213
6.2.4	Pressure contribution to circulation in an ideal gas	214

---

6.2.5	Circulation around a loop with constant entropy and concentration	214
6.2.6	Circulation around a loop with constant $S$ and $\Theta$ . . . . .	215
6.2.7	Comments and further reading . . . . .	216
<b>6.3</b>	<b>Vorticity dynamics</b> . . . . .	<b>216</b>
6.3.1	Vector-invariant velocity equation . . . . .	216
6.3.2	Basic form of the vorticity equation . . . . .	217
6.3.3	Massaged form of the vorticity equation . . . . .	217
6.3.4	Evolution of Cartesian vorticity components . . . . .	218
6.3.5	Evolution of the normal component of absolute vorticity . . . . .	220
6.3.6	Vorticity, angular momentum, and torques . . . . .	221
<b>6.4</b>	<b>Mechanics of baroclinicity</b> . . . . .	<b>221</b>
6.4.1	Curl of the pressure gradient body force . . . . .	221
6.4.2	Kelvin's circulation theorem and contact pressure forces . . . . .	222
6.4.3	Bottom pressure contributions at the solid-earth boundary . . . . .	224
6.4.4	Further study . . . . .	224
<b>6.5</b>	<b>Vortex lines and material lines</b> . . . . .	<b>224</b>
6.5.1	Vortex lines evolve through the strain rate tensor . . . . .	225
6.5.2	Frozen-in nature of vorticity . . . . .	226
6.5.3	Stretching and tilting of vortex tubes . . . . .	226
6.5.4	Shallow water vorticity revisited . . . . .	229
6.5.5	Concerning three-dimensional turbulence . . . . .	230
<b>6.6</b>	<b>Circulation viewed in a rotating reference frame</b> . . . . .	<b>230</b>
6.6.1	Material evolution of absolute circulation . . . . .	231
6.6.2	The beta effect . . . . .	233
6.6.3	The case of two-dimensional non-divergent flow . . . . .	234
6.6.4	Planetary circulation, planetary vorticity, and the Coriolis acceleration . . . . .	235
6.6.5	Further study . . . . .	237
<b>6.7</b>	<b>Vorticity budget for a primitive equation Boussinesq ocean</b> . . . . .	<b>237</b>
6.7.1	Deriving the vorticity equation . . . . .	238
6.7.2	Boussinesq baroclinicity . . . . .	239
6.7.3	Vertical vorticity equation . . . . .	241
<b>6.8</b>	<b>Evolution of depth integrated vertical vorticity</b> . . . . .	<b>242</b>
6.8.1	Comments on the role of baroclinicity . . . . .	242
6.8.2	Leibniz rule expressions . . . . .	243
6.8.3	Bottom boundary contribution . . . . .	243
6.8.4	Surface boundary contribution . . . . .	245
6.8.5	Comments . . . . .	246
<b>6.9</b>	<b>Vorticity for depth integrated hydrostatic flow</b> . . . . .	<b>246</b>
6.9.1	Comparing the two vorticities . . . . .	247
6.9.2	Evolution of vorticity for the depth integrated horizontal flow . .	247
6.9.3	Boundary pressure torques . . . . .	248
6.9.4	Steady state vorticity budget . . . . .	250
6.9.5	Integral balances satisfied by steady flows . . . . .	252
6.9.6	Formulation based on the vector-invariant velocity equation . . . .	253
6.9.7	Vorticity of the depth averaged flow . . . . .	254
6.9.8	Comments and further study . . . . .	254
<b>6.10</b>	<b>Exercises</b> . . . . .	<b>255</b>

---

## 6.1 Loose threads

- Rewrite Section 6.8 using the full vertical vorticity equation derived in Section 6.3.4. The only difference is the presence of a component of baroclinicity,  $-\hat{z} \times \rho^{-1} \nabla p$ .
- Summarize the various forms of the vorticity budget as per *McWilliams et al. (2024)* Section 2.
- Consider vertical velocity as per *Fraser et al. (2024)*.

## 6.2 Kelvin's circulation theorem

In this section we study the evolution of circulation around a closed loop that follows the flow, or equivalently (through Stokes' theorem) with the change in vorticity penetrating the area enclosed by the loop

$$\frac{d\mathcal{C}}{dt} = \frac{d}{dt} \oint_{\partial\mathcal{S}(\mathbf{v})} \mathbf{v} \cdot d\mathbf{x} = \frac{d}{dt} \int_{\mathcal{S}(\mathbf{v})} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (6.1)$$

where  $\mathcal{S}(\mathbf{v})$  designates a surface whose points all move with the fluid flow. We here consider the case of a non-rotating reference frame, with the straightforward extension to rotating reference frames in Section 6.6. Kelvin's theorem refers to the special case of a perfect barotropic fluid, whereas the treatment here considers how circulation evolves for more general flows.

We emphasize that our concern is with circulation computed around closed loops that follow the flow. For single-component fluids, such loops are material (i.e., the same fluid particles are fixed to the loop). Correspondingly, we can make use of Kelvin's circulation theorem for any material loop, including loops next to static material boundaries such as that considered in Exercise 6.11. For multi-component fluids there is no perfectly material loop in the presence of diffusion, and yet we can still make use of Kelvin's theorem for loops that follow the **barycentric velocity**,  $\mathbf{v}$ .

### 6.2.1 Formulation

Since the circulation is computed for a closed circuit following the flow, the time derivative in equation (6.1) moves inside the integral as a material/Lagrangian time derivative<sup>1</sup>

$$\frac{d\mathcal{C}}{dt} = \frac{d}{dt} \oint_{\partial\mathcal{S}(\mathbf{v})} \mathbf{v} \cdot d\mathbf{x} = \oint_{\partial\mathcal{S}(\mathbf{v})} \frac{D(\mathbf{v} \cdot d\mathbf{x})}{Dt}. \quad (6.2)$$

The material evolution of  $\mathbf{v}$  is determined by Newton's law of motion, which for a non-rotating reference frame is given by (see VOLUME 2)

$$\frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F}. \quad (6.3)$$

In this equation,  $p$  is the pressure,  $\rho$  is the mass density,  $\Phi$  is the geopotential (and/or the potential for any conservative force), and  $\mathbf{F}$  is the acceleration from any non-conservative forces such as from viscous stresses and/or boundary stresses.

---

<sup>1</sup>In VOLUME 2, we studied how to take derivatives of flow-following integrals.

The material time derivative of the differential line element moving around the circuit equals to the differential of the velocity on the circuit

$$\frac{D(dx)}{Dt} = dv. \quad (6.4)$$

This result follows since all points along the circuit follow the flow, by construction. Consequently, evolution of circulation following a loop becomes

$$\frac{d\mathcal{C}}{dt} = \oint_{\partial S(v)} [(-\rho^{-1} \nabla p - \nabla \Phi + \mathbf{F}) \cdot dx + v \cdot dv] \quad (6.5a)$$

$$= \oint_{\partial S(v)} [(-\rho^{-1} \nabla p + \mathbf{F}) \cdot dx + d(-\Phi + v \cdot v/2)] \quad (6.5b)$$

$$= \oint_{\partial S(v)} (-\rho^{-1} \nabla p + \mathbf{F}) \cdot dx \quad (6.5c)$$

$$= \int_{S(v)} [-\nabla \times (\rho^{-1} \nabla p) + \nabla \times \mathbf{F}] \cdot \hat{n} dS \quad (6.5d)$$

$$= \int_{S(v)} (\mathbf{B} + \nabla \times \mathbf{F}) \cdot \hat{n} dS. \quad (6.5e)$$

We noted that when integrating over space at a particular time,  $\nabla \Phi \cdot dx = d\Phi$  is an exact spatial differential, and so is  $dv^2$ . Hence, they both have a zero line integral around a closed circuit in space<sup>2</sup>

$$\oint_{\partial S(v)} d\Phi = 0 \quad \text{and} \quad \oint_{\partial S(v)} dv^2 = 0. \quad (6.6)$$

### 6.2.2 Two processes affecting circulation

Equation (6.5e) says that the circulation around a flow-following loop is affected by two processes, whose form depends on whether considering their line integral or surface integral expressions. The contribution from non-conservative forces take the form

$$\oint_{\partial S(v)} \mathbf{F} \cdot dx = \int_{S(v)} (\nabla \times \mathbf{F}) \cdot \hat{n} dS. \quad (6.7)$$

The line integral form expresses the mechanical work per unit mass (acceleration times distance) done by the non-conservative forces around the closed loop.<sup>3</sup> If the force is associated with friction, then friction acts to dissipate kinetic energy, with friction also leading, in general, to a reduction of the circulation magnitude. The surface integral form expresses the curl of friction as integrated over the surface.

The pressure gradient acceleration appears in equation (6.5e), and it provides a reversible mechanical process affecting circulation

$$\oint_{\partial S(v)} (-\rho^{-1} \nabla p) \cdot dx = \int_{S(v)} \rho^{-2} (\nabla \rho \times \nabla p) \cdot \hat{n} dS = \int_{S(v)} \mathbf{B} \cdot \hat{n} dS. \quad (6.8)$$

As for friction, the line integral form expresses the mechanical work per unit mass (acceleration times distance) done by the pressure gradient acceleration as integrated around the circuit.

<sup>2</sup>See VOLUME 1 for more on exact differentials.

<sup>3</sup>We introduced the notions of mechanical work in VOLUME 2.

The vector  $\mathbf{B}$  is referred to as the baroclinicity

$$\mathbf{B} = -\nabla \times (\rho^{-1} \nabla p) = \rho^{-2} \nabla \rho \times \nabla p, \quad (6.9)$$

and it has physical dimensions of inverse squared time,  $T^{-2}$ . Its appearance in the circulation theorem arises from the non-alignment of density and pressure isolines.

Work done by pressure around an arbitrary loop does not generally vanish, nor does it have a specific sign. However, there are a variety of special loops around which the pressure work does vanish. For example, the pressure work vanishes for closed contours on surfaces of constant density or constant pressure. However, such contours are generally not flow-following, and so their circulation is not considered as part of Kelvin's theorem. In the remainder of this section, we consider some further cases where baroclinicity vanishes.

### 6.2.3 Barotropic flow

The solenoidal/baroclinicity vector vanishes for a constant density fluid, in which  $\nabla \rho = 0$  such as for a single layer of shallow water fluid. More generally, the baroclinicity vector vanishes for barotropic flow, in which

$$p = p(\rho) \implies \text{barotropic flow.} \quad (6.10)$$

Kelvin's circulation theorem then follows, which states that for inviscid barotropic flow the circulation around any closed flow-following circuit remains constant

$$\frac{d\mathcal{C}}{dt} = \frac{d}{dt} \oint_{\partial S(v)} \mathbf{v} \cdot d\mathbf{x} = \frac{d}{dt} \int_{S(v)} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0 \quad \iff \text{inviscid barotropic flow.} \quad (6.11)$$

That is, the circulation around any closed flow-following circuit in a perfect barotropic fluid remains materially constant. This remarkable result greatly constrains the flow and thus provides a wealth of insights into the nature of the flow field.

Another way to recognize that baroclinicity vanishes for a barotropic flow is to note that the curl of the pressure gradient acceleration vanishes

$$\nabla \times (\rho^{-1} \nabla p) = -\rho^{-2} \nabla \rho \times \nabla p = -\rho^{-2} (\partial p / \partial \rho) \nabla \rho \times \nabla \rho = 0. \quad (6.12)$$

Hence, for a barotropic fluid there is a scalar potential whereby

$$\nabla \Phi_p = \rho^{-1} \nabla p. \quad (6.13)$$

This identity means that the pressure gradient acceleration for a barotropic flow is an exact spatial differential

$$\rho^{-1} \nabla p \cdot d\mathbf{x} = \rho^{-1} dp \equiv d\Phi_p, \quad (6.14)$$

with integration leading to

$$\Phi_p = \int_{p_0}^p \frac{dp'}{\rho(p')}, \quad (6.15)$$

where  $p_0$  is an arbitrary reference pressure. Since the closed loop integral of an exact differential vanishes, we again see that the pressure gradient acceleration has no impact on circulation around flow-following loops in a barotropic flow. Stated alternatively, with  $\nabla \Phi_p = \rho^{-1} \nabla p$ , then

the Euler equation takes on the form

$$\frac{D\mathbf{v}}{Dt} = -\nabla(\Phi_p + \Phi), \quad (6.16)$$

which renders a materially constant circulation.

#### 6.2.4 Pressure contribution to circulation in an ideal gas

Building on the notions from a barotropic flow in Section 6.2.3, we here determine a class of contours for an ideal gas where baroclinicity vanishes even if the flow is not barotropic. To start, we make use of an exercise in VOLUME 2 in which we derived the following equation for an ideal gas

$$\rho^{-1} \nabla p = \theta \nabla \Pi, \quad (6.17)$$

where  $\theta$  is the [potential temperature](#) and  $\Pi$  is the [Exner function](#). This equation says that the pressure gradient acceleration, for an ideal gas, is equal to the potential temperature times the gradient of the Exner function. We can thus write the pressure gradient acceleration contribution in equation (6.5c) as

$$-\oint_{\partial S(\mathbf{v})} \rho^{-1} \nabla p \cdot d\mathbf{x} = -\oint_{\partial S(\mathbf{v})} \theta \nabla \Pi \cdot d\mathbf{x} = -\oint_{\partial S(\mathbf{v})} \theta d\Pi = \oint_{\partial S(\mathbf{v})} \Pi d\theta, \quad (6.18)$$

where the final step noted that the closed loop integral of an exact differential vanishes

$$\oint_{\partial S(\mathbf{v})} d(\theta \Pi) = 0. \quad (6.19)$$

Hence, the contribution from baroclinicity (i.e., pressure gradient acceleration) vanishes for closed contours drawn either on a constant  $\Pi$  surface or a constant  $\theta$  surface. For an ideal gas, changes in  $\theta$  are directly related to changes in specific entropy<sup>4</sup> Hence, for a perfect fluid flow of an ideal gas, where specific entropy is materially invariant, so too is potential temperature:  $D\theta/Dt = 0$ . It follows that a contour drawn on a potential temperature surface remains a flow-following contour for a perfect fluid. We have thus deduced that isentropic flow of an ideal gas has a flow-following circulation that is unaffected by baroclinicity.

#### 6.2.5 Circulation around a loop with constant entropy and concentration

We here make use of an expression from VOLUME 2 for the pressure gradient acceleration in terms of thermodynamic functions

$$-\rho^{-1} \nabla p = -\nabla \mathcal{H} + T \nabla \mathcal{S} + \mu \nabla C, \quad (6.20)$$

where  $\mathcal{H}$  is the specific enthalpy,  $\mathcal{S}$  is the specific entropy,  $T$  is the thermodynamic (*in situ*) temperature,  $C$  is the material tracer concentration, and  $\mu$  is the chemical potential for a binary fluid such as commonly assumed for the ocean (freshwater and salt) and atmosphere (dry air and water vapor). Equation (6.20) holds for a compressible fluid, in which the thermodynamic pressure and [mechanical pressure](#) are the same.<sup>5</sup>

<sup>4</sup>We pursue these ideas in the chapter on energetics in VOLUME 2, in particular within one of its exercises.

<sup>5</sup>We discuss this equivalence in VOLUME 2.

The identity (6.20) brings the pressure gradient contribution to Kelvin's circulation theorem (6.5e) into

$$-\oint_{\partial S(v)} \rho^{-1} \nabla p \cdot d\mathbf{x} = \oint_{\partial S(v)} (T \nabla S + \mu \nabla C) \cdot d\mathbf{x}, \quad (6.21)$$

where we set

$$\oint_{\partial S(v)} \nabla \mathcal{H} \cdot d\mathbf{x} = 0, \quad (6.22)$$

which holds since  $\mathcal{H}$  is a state function so that for any time instance,

$$\nabla \mathcal{H} \cdot d\mathbf{x} = d\mathcal{H} \quad (6.23)$$

is an exact spatial differential. The decomposition (6.21) reveals that the pressure contribution to circulation vanishes when computing circulation for an isentropic and constant concentration loop<sup>6</sup>

$$\oint_{\partial S(v)} \rho^{-1} \nabla p \cdot d\mathbf{x} = 0 \quad \text{if } dS = 0 \text{ and } dC = 0. \quad (6.24)$$

Such loops follow the flow in those cases where specific entropy and matter concentration are materially invariant

$$\frac{DS}{Dt} = 0 \quad \text{and} \quad \frac{DC}{Dt} = 0. \quad (6.25)$$

It follows that for the special case of a homogeneous fluid ( $C = \text{constant}$ ) undergoing isentropic quasi-static changes, pressure plays no role in the circulation computed around an isentropic loop.

### 6.2.6 Circulation around a loop with constant $S$ and $\Theta$

Rather than invoking the gradient form of the fundamental thermodynamic relation (VOLUME 2), consider an ocean application where equation the *in situ* density,  $\rho$ , takes on the functional form

$$\rho = \rho(S, \Theta, p), \quad (6.26)$$

with  $S$  the salinity and  $\Theta$  the Conservative Temperature. It follows that along a contour that maintains fixed  $S$  and  $\Theta$ , the pressure gradient acceleration is a function just of the pressure, in which case we write

$$\rho^{-1} \nabla p \cdot d\mathbf{x} = \rho^{-1}(S_{\text{const}}, \Theta_{\text{const}}, p) dp \equiv d\Psi_p \quad \text{if } S = S_{\text{const}} \text{ and } \Theta = \Theta_{\text{const}}, \quad (6.27)$$

where we followed the barotropic case of equation (6.15) to write

$$\Psi_p(S_{\text{const}}, \Theta_{\text{const}}, p) = \int_{p_0}^p \frac{dp'}{\rho(S_{\text{const}}, \Theta_{\text{const}}, p')}, \quad (6.28)$$

with  $p_0$  an arbitrary reference pressure. As in Section 6.2.5, we conclude that pressure plays no role in affecting circulation around loops with fixed  $S$  and  $\Theta$

$$\oint_{\partial S(v)} \rho^{-1} \nabla p \cdot d\mathbf{x} = \oint_{\partial S(v)} d\Psi_p(S, \Theta, p) = 0 \quad \text{if } S = S_{\text{const}} \text{ and } \Theta = \Theta_{\text{const}}, \quad (6.29)$$

---

<sup>6</sup>For the ocean, an isentropic and constant salt concentration process maintains a constant Conservative Temperature.

which follows since  $d\Psi_p$  is an exact spatial differential. Such closed loop contours follow the flow if  $S$  and  $\Theta$  are materially invariant

$$\frac{DS}{Dt} = 0 \quad \text{and} \quad \frac{D\Theta}{Dt} = 0, \quad (6.30)$$

which is the case in the absence of mixing and/or sources of  $S$  and  $\Theta$ .

### 6.2.7 Comments and further reading

There is no guarantee that the closed flow-following contours discussed in Sections 6.2.5 and 6.2.6 exist in any particular flow. Rather, all we showed is that if such closed contours exist, and if the flow maintains materially invariant specific entropy and concentration (Section 6.2.5) or salinity and Conservative Temperature (Section 6.2.6), then the circulation around such loops is unaffected by the pressure gradient acceleration.

Our presentation of Kelvin's circulation theorem anticipates analogous considerations encountered with potential vorticity in Chapter 7, with portions of the discussion motivated by our study of [Kooloth et al. \(2022\)](#) given in Section 7.7.

## 6.3 Vorticity dynamics

We now study the time evolution of vorticity and the processes leading to this evolution. We considered this question in Chapter 5 when focused on shallow water vorticity, and in Chapter 4 for the horizontally non-divergent barotropic flow. Here, we consider the general case of a stratified flow with non-conservative forces. As for Kelvin's theorem, we make use of Newton's law of motion, written here in the form for a rotating fluid (see VOLUME 2)

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\rho^{-1} \nabla p - \nabla \Phi + \mathbf{F}, \quad (6.31)$$

where  $\boldsymbol{\Omega}$  is the angular velocity of the rotating reference frame.

### 6.3.1 Vector-invariant velocity equation

As for the shallow water fluid in Section 5.1, we find it useful to convert the advective-form momentum equation to vector-invariant velocity equation. For this purpose, make use of a vector identity from VOLUME 1

$$\boldsymbol{\omega} \times \mathbf{v} = -(1/2) \nabla(\mathbf{v} \cdot \mathbf{v}) + (\mathbf{v} \cdot \nabla) \mathbf{v} \quad (6.32)$$

to eliminate velocity self-advection in favor of vorticity and kinetic energy

$$\partial_t \mathbf{v} + \boldsymbol{\omega}_a \times \mathbf{v} = -\rho^{-1} \nabla p - \nabla(\mathbf{v}^2/2 + \Phi) + \mathbf{F}. \quad (6.33)$$

We here introduced the absolute vorticity

$$\boldsymbol{\omega}_a = \nabla \times (\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) = \boldsymbol{\omega} + 2\boldsymbol{\Omega}, \quad (6.34)$$

which is the curl of the absolute (inertial frame) velocity, and which equals to the sum of the relative vorticity plus the planetary vorticity (see Section 3.6.1).

### 6.3.2 Basic form of the vorticity equation

Taking the curl of the vector-invariant momentum equation (6.33) removes the mechanical energy per mass,  $\mathbf{v}^2/2 + \Phi$ , thus leaving

$$\partial_t \boldsymbol{\omega} + \nabla \times (\boldsymbol{\omega}_a \times \mathbf{v}) = \rho^{-2} (\nabla \rho \times \nabla p) + \nabla \times \mathbf{F}. \quad (6.35)$$

For geophysical fluid mechanics, we generally assume that  $\boldsymbol{\Omega}$  has zero time tendency so that

$$\partial_t \boldsymbol{\omega}_a = \partial_t (\boldsymbol{\omega} + 2 \boldsymbol{\Omega}) = \partial_t \boldsymbol{\omega}, \quad (6.36)$$

in which case equation (6.35) can be written as an equation for absolute vorticity

$$\partial_t \boldsymbol{\omega}_a + \nabla \times (\boldsymbol{\omega}_a \times \mathbf{v}) = \mathbf{B} + \nabla \times \mathbf{F}, \quad (6.37)$$

where  $\mathbf{B}$  is the baroclinicity vector given by equation (6.9).

### 6.3.3 Massaged form of the vorticity equation

Physical interpretation of the term  $\nabla \times (\boldsymbol{\omega}_a \times \mathbf{v})$  appearing in the prognostic equation (6.37) can be made more transparent by using the following vector identity

$$\nabla \times (\boldsymbol{\omega}_a \times \mathbf{v}) = (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \boldsymbol{\omega}_a \nabla \cdot \mathbf{v} - \mathbf{v} \nabla \cdot \boldsymbol{\omega}_a \quad (6.38a)$$

$$= (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} - \frac{\boldsymbol{\omega}_a}{\rho} \frac{D\rho}{Dt}. \quad (6.38b)$$

The second equality required the continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}, \quad (6.39)$$

and the non-divergent nature of the absolute vorticity

$$\nabla \cdot \boldsymbol{\omega}_a = \nabla \cdot (\nabla \times \mathbf{v} + 2 \boldsymbol{\Omega}) = 0. \quad (6.40)$$

Equation (6.37) thus takes the form

$$\frac{D\boldsymbol{\omega}_a}{Dt} - \frac{\boldsymbol{\omega}_a}{\rho} \frac{D\rho}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \frac{1}{\rho^2} (\nabla \rho \times \nabla p) + \nabla \times \mathbf{F}, \quad (6.41)$$

which can be written

$$\rho \frac{D(\boldsymbol{\omega}_a/\rho)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \mathbf{B} + \nabla \times \mathbf{F}. \quad (6.42)$$

Each term on the right hand side of the material evolution equation (6.42) represents a distinct physical process that affects  $\boldsymbol{\omega}_a/\rho$  of a fluid element. The first term,  $(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}$ , embodies stretching and twisting and is explored in Section 6.5 in the simplified context of a barotropic fluid. The second term arises from baroclinicity as introduced in equation (6.9) and given a mechanical interpretation in Section 6.4. The third term arises from the curl of the non-conservative forces (e.g., friction). Such forces contribute especially in boundary layer regions where friction curls are relatively large in magnitude.

### 6.3.4 Evolution of Cartesian vorticity components

Terms appearing on the right hand side of the vorticity equation (6.42) provide sources for the vorticity of a fluid element. Here we derive flux-form conservation equations that are separately satisfied by each of the vorticity components. We find that each vorticity component has an Eulerian time derivative determined by the convergence of a corresponding flux.

#### Vertical component to the absolute vorticity

Consider the material evolution equation for the vertical component of the absolute vorticity

$$\rho \frac{D(\zeta_a/\rho)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) w + \hat{\mathbf{z}} \cdot (\mathbf{B} + \nabla \times \mathbf{F}) \quad \text{with } \zeta_a = \hat{\mathbf{z}} \cdot \boldsymbol{\omega}_a. \quad (6.43)$$

Making use of the identities<sup>7</sup>

$$\rho \frac{D(\zeta_a/\rho)}{Dt} = \frac{\partial \zeta_a}{\partial t} + \nabla \cdot (\mathbf{v} \zeta_a) \quad (6.44a)$$

$$(\boldsymbol{\omega}_a \cdot \nabla) w = \nabla \cdot (\boldsymbol{\omega}_a w) \quad (6.44b)$$

$$\hat{\mathbf{z}} \cdot \mathbf{B} = -\hat{\mathbf{z}} \cdot [\nabla \times (\rho^{-1} \nabla p)] = \nabla \cdot (\hat{\mathbf{z}} \times \rho^{-1} \nabla p) \quad (6.44c)$$

$$\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) = -\nabla \cdot (\hat{\mathbf{z}} \times \mathbf{F}), \quad (6.44d)$$

brings equation (6.43) into the flux-form

$$\partial_t \zeta_a = -\nabla \cdot \mathbf{J}^{\zeta_a} \quad \text{with } \mathbf{J}^{\zeta_a} = \mathbf{v} \zeta_a - \boldsymbol{\omega}_a w - \hat{\mathbf{z}} \times \rho^{-1} \nabla p + \hat{\mathbf{z}} \times \mathbf{F}. \quad (6.45)$$

This budget equation says that  $\zeta_a$  evolves at a point according to the convergence of a vorticity flux,  $\mathbf{J}^{\zeta_a}$ . The vorticity flux is comprised of the following terms:

- advective flux of vertical vorticity:  $\mathbf{v} \zeta_a$ ,
- absolute vorticity transported vertically:  $-\boldsymbol{\omega}_a w$ ,
- pressure gradient acceleration rotated clockwise by  $\pi/2$  radians around the vertical axis:  $-\hat{\mathbf{z}} \times \rho^{-1} \nabla p$ ,
- $\pi/2$  counter-clockwise rotated friction acceleration,  $\hat{\mathbf{z}} \times \mathbf{F}$ .

Note that there is no vertical component to the vorticity flux:

$$\hat{\mathbf{z}} \cdot \mathbf{J}^{\zeta_a} = \hat{\mathbf{z}} \cdot [\mathbf{v} \zeta_a - \boldsymbol{\omega}_a w - \hat{\mathbf{z}} \times \rho^{-1} \nabla p + \hat{\mathbf{z}} \times \mathbf{F}] = 0, \quad (6.46)$$

so that  $\zeta_a$  is only affected by the convergence of a purely horizontal flux. We offer a schematic of this property in Figure 6.1.

#### Flux for the other Cartesian directions

Mathematically, there is nothing special about the vertical vorticity component. Hence, we readily find that the horizontal vorticity components also satisfy their own respective flux-form

<sup>7</sup>Exercise 6.1 asks for a proof of the identity (6.44a), which follows from use of the mass continuity equation in the form  $D\rho/Dt = -\rho \nabla \cdot \mathbf{v}$ . For equation (6.44d) we can use  $\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) = \nabla z \cdot (\nabla \times \mathbf{F}) = \nabla \cdot [z \nabla \times \mathbf{F}] = \nabla \cdot [\nabla \times (z \mathbf{F}) - \nabla z \times \mathbf{F}] = -\nabla \cdot (\hat{\mathbf{z}} \times \mathbf{F})$ .

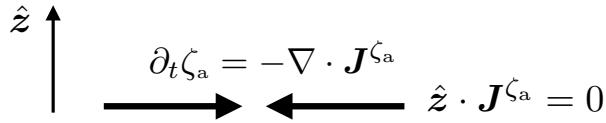


FIGURE 6.1: Tendency for the vertical component to the absolute vorticity arises from convergence of the vorticity flux,  $\mathbf{J}^{\zeta_a}$ , given by equation (6.45). This vorticity flux is strictly horizontal,  $\hat{z} \cdot \mathbf{J}^{\zeta_a} = 0$ . This result generalizes for any arbitrary Cartesian component of the absolute vorticity, whereby the corresponding vorticity flux is orthogonal to its vorticity component. Generalization of this result leads to the impermeability theorem of potential vorticity studied in Section 8.2.

conservation equations, thus leading to the general result

$$\partial_t(\hat{e} \cdot \boldsymbol{\omega}_a) = -\nabla \cdot [\mathbf{v}(\hat{e} \cdot \boldsymbol{\omega}_a) - \boldsymbol{\omega}_a(\hat{e} \cdot \mathbf{v}) - \hat{e} \times \rho^{-1} \nabla p + \hat{e} \times \mathbf{F}], \quad (6.47)$$

where  $\hat{e}$  is any one of the Cartesian unit vectors  $\hat{x}, \hat{y}, \hat{z}$ . Furthermore, we readily see that the vorticity flux satisfies

$$\hat{e} \cdot [\mathbf{v}(\hat{e} \cdot \boldsymbol{\omega}_a) - \boldsymbol{\omega}_a(\hat{e} \cdot \mathbf{v}) - \hat{e} \times \rho^{-1} \nabla p + \hat{e} \times \mathbf{F}] = 0, \quad (6.48)$$

so that the time tendency for  $\hat{e} \cdot \boldsymbol{\omega}_a$  is affected by a flux in the directions orthogonal to  $\hat{e}$ . This property of the vorticity flux is generalized via the impermeability theorem of potential vorticity studied in Section 8.2, with particular connection to the present discussion given in Section 8.2.2.

### The kinematic vorticity flux

Following our study of vorticity for the shallow water fluid in Section 5.4.2, we here show that there is a kinematic reason that each Cartesian component of vorticity has a time tendency given by the convergence of a flux. For this purpose, recall the identity (3.2), in which the vertical component to the relative vorticity is written as the divergence of the velocity rotated by  $\pi/2$  in the clockwise direction around the vertical axis

$$\zeta = \nabla \cdot (\mathbf{v} \times \hat{z}). \quad (6.49)$$

It follows, quite trivially, that

$$\partial_t \zeta_a = \nabla \cdot (\partial_t \mathbf{v} \times \hat{z}) \equiv -\nabla \cdot \mathbf{J}^{\text{kin}}. \quad (6.50)$$

The kinematic flux,

$$\mathbf{J}^{\text{kin}} = -\partial_t \mathbf{v} \times \hat{z}, \quad (6.51)$$

differs from  $\mathbf{J}^{\zeta_a}$  by a total curl, so that their convergences are identical. We prove this assertion by returning to the vector-invariant velocity equation (6.33), which leads to the identity

$$-\partial_t \mathbf{v} \times \hat{z} = (\boldsymbol{\omega}_a \times \mathbf{v}) \times \hat{z} + \nabla \times [\hat{z}(\Phi + \mathbf{v} \cdot \mathbf{v}/2)] - \hat{z} \times \rho^{-1} \nabla p + \hat{z} \times \mathbf{F}, \quad (6.52)$$

which then renders

$$\mathbf{J}^{\text{kin}} = \mathbf{J}^{\zeta_a} + \nabla \times [\hat{z}(\Phi + \mathbf{v} \cdot \mathbf{v}/2)]. \quad (6.53)$$

The rotational term equals to the mechanical energy per mass, which was also found for the shallow water case given by equation (5.61).

### 6.3.5 Evolution of the normal component of absolute vorticity

As a further examination of vorticity components, we here consider the material evolution of vorticity projected onto the unit normal vector,  $\hat{\mathbf{n}}$ , for an infinitesimal material area,  $\delta\mathcal{S}$ . This discussion leads to an infinitesimal version of Kelvin's circulation theorem, thus explicitly linking the evolution equations for vorticity and circulation. In the process we invoke some results from fluid kinematics derived in VOLUME 2.

The unit normal vector to a material surface evolves according to

$$\frac{D\hat{\mathbf{n}}_m}{Dt} = -\hat{\mathbf{n}} \cdot \partial_m^{\text{surf}} \mathbf{v}, \quad (6.54)$$

where the surface derivative, and corresponding surface divergence, are given by

$$\partial_m^{\text{surf}} = \partial_m - \hat{\mathbf{n}}_m (\hat{\mathbf{n}} \cdot \nabla) \quad \text{and} \quad \nabla^{\text{surf}} = \nabla - \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \nabla). \quad (6.55)$$

Making use of the vorticity equation in the form (6.41), along with the continuity equation (6.39), leads to

$$\hat{\mathbf{n}} \cdot \frac{D\omega_a}{Dt} = -(\hat{\mathbf{n}} \cdot \omega_a) \nabla \cdot \mathbf{v} + \hat{n}_j (\omega_a \cdot \nabla) v_j + \hat{\mathbf{n}} \cdot (\mathbf{B} + \nabla \times \mathbf{F}). \quad (6.56)$$

Likewise, taking the dot product of  $\omega_a$  with the evolution equation (6.54) yields

$$\omega_a \cdot \frac{D\hat{\mathbf{n}}}{Dt} = \hat{n}_j [-(\omega_a \cdot \nabla) + (\hat{\mathbf{n}} \cdot \omega_a) (\hat{\mathbf{n}} \cdot \nabla)] v_j. \quad (6.57)$$

Adding these two equations renders the material time evolution

$$\frac{D(\omega_a \cdot \hat{\mathbf{n}})}{Dt} = -(\omega_a \cdot \hat{\mathbf{n}}) \nabla^{\text{surf}} \cdot \mathbf{v} + \hat{\mathbf{n}} \cdot (\mathbf{B} + \nabla \times \mathbf{F}). \quad (6.58)$$

The surface divergence of the velocity measures evolution of the material surface area

$$\frac{1}{\delta S} \frac{D\delta\mathcal{S}}{Dt} = \nabla^{\text{surf}} \cdot \mathbf{v}, \quad (6.59)$$

so that the material evolution equation for the  $\hat{\mathbf{n}}$  component of absolute vorticity is

$$\frac{D(\omega_a \cdot \hat{\mathbf{n}})}{Dt} = -\frac{(\omega_a \cdot \hat{\mathbf{n}})}{\delta S} \frac{D\delta\mathcal{S}}{Dt} + \hat{\mathbf{n}} \cdot (\mathbf{B} + \nabla \times \mathbf{F}). \quad (6.60)$$

The area term arises from the familiar “ice-skater” effect that reflects angular momentum conservation for the column (Section 5.2.2), whereas the other terms are the projection of the baroclinicity and friction curl onto  $\hat{\mathbf{n}}$ . Bringing the area term onto the left hand side then renders the infinitesimal form of Kelvin's circulation theorem (6.5e)

$$\frac{D}{Dt} (\omega_a \cdot \hat{\mathbf{n}} dS) = (\mathbf{B} + \nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad (6.61)$$

Note the presence of the absolute vorticity,  $\omega_a$ , in equation (6.60) rather than the relative vorticity considered in Section 6.2. We return in Section 6.6 to the question of circulation arising from planetary rotation, where we derive the finite version of the circulation theorem (6.60).

### 6.3.6 Vorticity, angular momentum, and torques

Both vorticity and angular momentum offer measures of the rotational motion of a fluid flow. However, there are key distinctions as detailed in Section 3.9. Perhaps the most fundamental distinction is that vorticity measures the rotation or spin without reference to an origin, whereas angular momentum is computed relative to a subjectively chosen origin. Vorticity is thus an intrinsic property of the fluid flow, whereas angular momentum depends on the chosen origin and is affected by fluid strains. Consequently, there is a direct connection between angular momentum and vorticity only for the special case of flow exhibiting rigid-body motion.

Angular momentum of motion relative to an origin changes in the presence of torques computed about the chosen origin, with the torque equal to the cross product of the position vector of a point and the force vector acting at that point. In contrast, vorticity at a point is affected by the curl of the force per mass acting at the point. Furthermore, angular momentum is a property of any mechanical system, including point particles and rigid bodies, whereas vorticity is a property only of a continuous media where we can compute spatial derivatives of the velocity field.

When the curl of a force per mass is applied to a fluid and thus changes its vorticity, we commonly use the term “torque” in reference to this force curl. For example, in Section 6.4 we explore baroclinicity, which is the key mechanism for how inviscid torques from pressure modify vorticity. In that discussion, we see that baroclinicity provides a vorticity source when the pressure force acting on a fluid element does not pass through the center of mass of that element. When there is baroclinicity, the pressure force spins the fluid element thus affecting vorticity. Analogous inviscid and viscous force curls act on boundaries, such as when a fluid interacts with the solid earth. It is within this context that we use the term “torque” when referring to a vorticity source. Correspondingly, the torques providing a vorticity source have the dimension of force per mass per length, whereas the torques altering angular momentum have the dimension of force times length.<sup>8</sup>

## 6.4 Mechanics of baroclinicity

Baroclinicity is present in most geophysical flows, thus affecting the material evolution of circulation and vorticity. Flow with a nonzero baroclinicity vector is generally referred to as **baroclinic flow**, whereas a **barotropic flow** has zero baroclinicity. We illustrate the basic distinction between barotropic and baroclinic fluids in Figure 6.2. We observe that a baroclinic fluid is associated with fluid motion that is a function of the vertical direction. In contrast, a barotropic fluid in the special case of pressure and density surfaces aligned with geopotentials, supports no motion. We further develop these points as we explore the mechanics of baroclinicity.

### 6.4.1 Curl of the pressure gradient body force

Baroclinicity is the curl of the pressure gradient body force

$$\mathbf{B} = \nabla \times \mathbf{F}_{\text{press}} = \nabla \times (-\rho^{-1} \nabla p) = -\nabla \rho^{-1} \times \nabla p = \frac{\nabla \rho \times \nabla p}{\rho^2}. \quad (6.62)$$

As discussed in Section 6.3.6, the curl of a force provides a torque that spins the fluid, thus rendering a vorticity source. Geometrically, baroclinicity arises when there is nonzero change in

---

<sup>8</sup>See Section 2 of [Hughes \(2000\)](#) for a similar perspective on usage of the term “torque” for processes affecting vorticity.

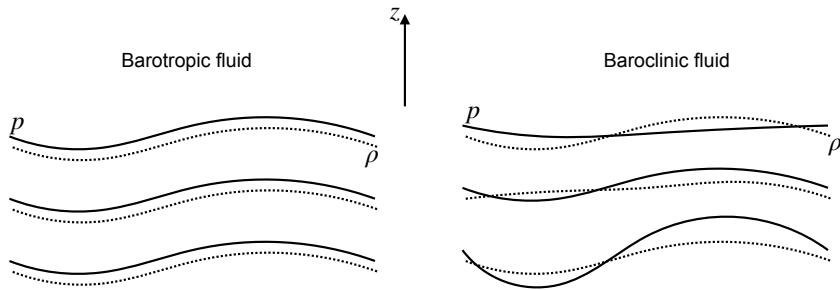


FIGURE 6.2: Left panel: a barotropic fluid, whereby density is a function just of pressure,  $\rho = \rho(p)$ , so that density surfaces (dashed lines) and pressure surfaces (solid lines) are parallel. Horizontal density and pressure surfaces in a barotropic fluid support no motion. Right panel: a baroclinic fluid, whereby density and pressure surfaces generally differ so that density is a function of more than just the pressure. A baroclinic fluid is associated with fluid motion that is a function of the vertical direction.

pressure along contours of constant density, or conversely changes in density along contours of constant pressure. It can be useful to introduce the notion of a **solenoid**, which is a tube region in the fluid that is perpendicular to both  $\nabla\rho$  and  $\nabla p$ . There are no solenoids for barotropic flows, whereby  $p = p(\rho)$  (see equation (6.10)). For baroclinic flow, solenoids are associated with a torque that affects vorticity.

To further understand the mechanical interpretation of solenoids in terms of a torque, consider the cross product

$$\rho \mathbf{B} = \mathbf{F}^{\text{press}} \times \nabla\rho = (-\rho^{-1} \nabla p) \times \nabla\rho. \quad (6.63)$$

The first term on the right hand side is the pressure gradient acceleration that is oriented down the pressure gradient. Now consider a tiny fluid element such as shown in Figure 6.3. By construction, the pressure force acts at the geometric center of the element. However, the nonzero density gradient means that the center of mass for the fluid element is not at the geometric center. In this case, the pressure gradient force does not pass through the center of mass, so that it imparts a torque to the fluid element. This torque then modifies the vorticity and hence the circulation around the boundary of the element. As an example, consider a horizontal pressure gradient acting in a fluid that is vertically stratified so  $\partial_z \rho \neq 0$ . There is a non-zero baroclinicity pointing in the horizontal direction (perpendicular to the horizontal pressure gradient and vertical density gradient), with this horizontal baroclinicity vector providing a source for horizontal vorticity. It is only if the pressure gradient force is aligned with the density gradient (barotropic flow), or if the density is spatially uniform (e.g., constant density homogeneous fluid), that we find the pressure gradient force passing through the center of mass and thus inducing no vorticity.

#### 6.4.2 Kelvin's circulation theorem and contact pressure forces

We are afforded another means to understand baroclinicity by returning to the formulation of Kelvin's circulation theorem in Section 6.2.1. Focusing just on the baroclinicity contribution in equation (6.5e) we have

$$\left[ \frac{dC}{dt} \right]_{\text{baroclinicity}} = \oint_{\partial S(v)} -\rho^{-1} \nabla p \cdot d\mathbf{x} = \oint_{\partial S(v)} \rho^{-2} (-p \nabla \rho) \cdot d\mathbf{x}, \quad (6.64)$$

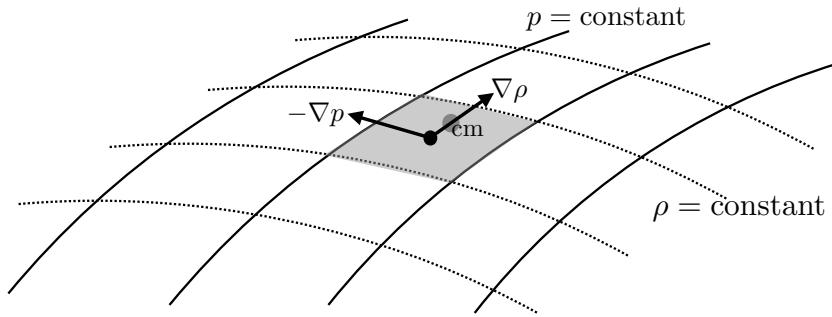


FIGURE 6.3: A mechanical interpretation of the baroclinicity vector. We consider a tiny fluid element bounded by surfaces of constant pressure and density. By construction, the pressure gradient force acts at the geometric center of the element, whereas the center of mass for the element is off-center due to the density gradient across the element. The pressure gradient force thus provides a torque for the fluid element, with the moment-arm for the torque determined by the distance between the geometric center and the center of mass. This torque modifies the vorticity of the fluid element, and in turn modifies the circulation computed around the element's boundary. As depicted here, the baroclinicity vector points into the page (right hand rule for  $(-\rho^{-1} \nabla p) \times \nabla \rho$ ), so that this baroclinicity spins-up a clockwise circulation around the element, or equivalently a clockwise vorticity. This figure is adapted from Figure 14.9 of [Thorne and Blandford \(2017\)](#).

which follows since, at any particular instance,

$$\oint_{\partial S(v)} \nabla(p/\rho) \cdot d\mathbf{x} = \oint_{\partial S(v)} d(p/\rho) = 0. \quad (6.65)$$

The term  $-p \nabla \rho$  in equation (6.64) is proportional to the compressive contact force from pressure force acting normal to a constant density surface. Consequently, if the material surface on which we are computing circulation happens to be parallel to a constant density surface, then pressure cannot generate any circulation around that material circuit. The left circuit in Figure 6.4 illustrates this situation. For the more general case where a material surface crosses constant density surfaces, pressure modifies circulation computed around such circuits (right circuit in Figure 6.4).

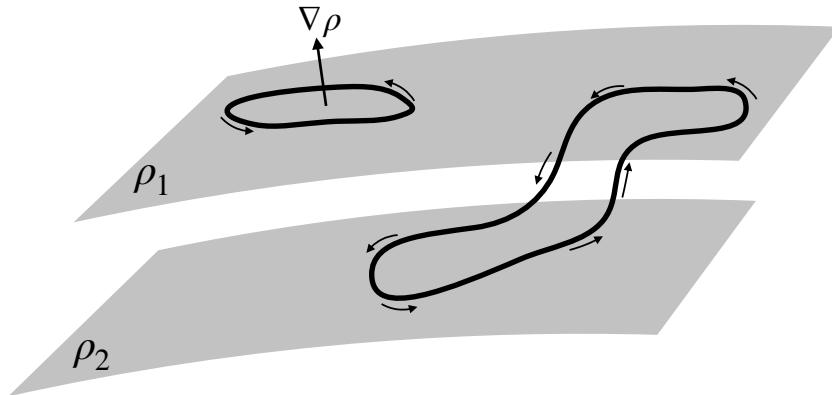


FIGURE 6.4: The material circuit on the left is assumed to be within a constant density surface. In this case,  $\nabla \rho \cdot d\mathbf{x} = 0$  so that pressure cannot modify the circulation around this circuit. However, if a material circuit is not contained fully within constant density surface, such as depicted on the right, then pressure is able to modify the circulation computed around the circuit.

### 6.4.3 Bottom pressure contributions at the solid-earth boundary

As an application of the above ideas, consider a fluid region that intersects the solid-earth boundary. The solid-earth boundary is material so that we can apply Kelvin's circulation theorem to a circuit on the boundary. Consider the situation in Figure 6.5, which shows a vertical slice next to a sloping bottom with constant density surfaces intersecting the bottom. As in our considerations in Section 6.4.2, any material circuit that sits within the bottom boundary crosses density surfaces, in which case circulation is affected by the bottom pressure. Indeed, even if the bottom is flat, so long as density is not constant along the bottom, then a material circuit within the bottom has circulation modified by bottom pressure.

To develop the mathematics of the above ideas, write the differential line element within the bottom circuit as

$$d\mathbf{x} = \hat{\mathbf{x}} dx + \hat{\mathbf{y}} dy + \hat{\mathbf{z}} dz = (\hat{\mathbf{x}} + \hat{\mathbf{z}} \partial_x \eta_b) dx + (\hat{\mathbf{y}} + \hat{\mathbf{z}} \partial_y \eta_b) dy. \quad (6.66)$$

To reach this result we set  $z = \eta_b(x, y)$  since the circuit is along the bottom boundary, which in turn means that<sup>9</sup>

$$dz = d\eta_b = \nabla \eta_b \cdot (\hat{\mathbf{x}} dx + \hat{\mathbf{y}} dy) = \nabla \eta_b \cdot d\mathbf{x}_{\text{horz}}. \quad (6.67)$$

Consequently, the projection of the density gradient onto the circuit is given by

$$\nabla \rho \cdot d\mathbf{x} = (\nabla_h \rho + \partial_z \rho \nabla \eta_b) \cdot d\mathbf{x}_{\text{horz}}. \quad (6.68)$$

Making use of this result in Kelvin's circulation theorem and focusing on the pressure contribution, as in equation (6.64), leads to

$$\left[ \frac{d\mathcal{C}}{dt} \right]_{\text{bottom}} = - \oint_{\partial S_{\text{bottom}}} p \rho^{-2} \nabla \rho \cdot d\mathbf{x} = - \oint_{\partial S_{\text{bottom}}} \frac{p_b}{\rho^2} (\nabla_h \rho + \partial_z \rho \nabla \eta_b) \cdot d\mathbf{x}_{\text{horz}}. \quad (6.69)$$

There are two contributions to the circulation changes revealed by equation (6.69). The first arises from the sloped density surfaces next to the bottom, and the second arises from the sloped bottom multiplied by the vertical density gradient. These two contributions are weighted by the bottom pressure,  $p_b$ , which is normalized by the squared density. Circulation modifications are enhanced by increased horizontal density gradients next to the bottom, as well as increased topographic slopes. For the special case of flat topography and flat density there are no bottom pressure-induced changes to the circulation around a bottom material circuit.

### 6.4.4 Further study

This video from Prof. Shapiro provides a lucid discussion of baroclinicity and its role in affecting vorticity and circulation.

## 6.5 Vortex lines and material lines

We here study the physics of the source term

$$(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} = f \partial_z \mathbf{v} + (\boldsymbol{\omega} \cdot \nabla) \mathbf{v} \quad (6.70)$$

---

<sup>9</sup>Since  $\eta_b = \eta_b(x, y)$ , its gradient is horizontal:  $\nabla \eta_b = \hat{\mathbf{x}} \partial_x \eta_b + \hat{\mathbf{y}} \partial_y \eta_b$ .

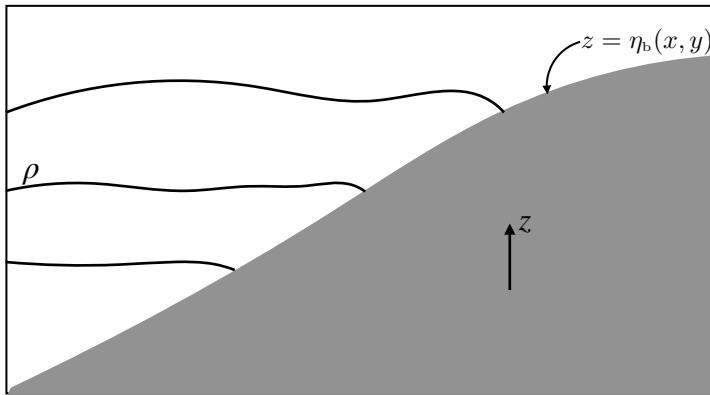


FIGURE 6.5: Constant density surfaces intersecting a sloped solid-earth boundary. Any circuit that sits along the boundary is material since the bottom is material. For circuits that cross density surfaces, the bottom pressure acts to modify circulation computed for this circuit.

appearing in the vorticity equation (6.42). The contribution from  $\hat{z} f \partial_z w$  to the first term is further explored when studying the equations of [planetary geostrophy](#) in [VOLUME 2](#) and Section 10.3, given its importance for large-scale meridional motion on a spherical planet. The second term,  $(\boldsymbol{\omega} \cdot \nabla) \mathbf{v}$ , is the focus of this section.

### 6.5.1 Vortex lines evolve through the strain rate tensor

To help unpack the physics of the source,  $(\boldsymbol{\omega} \cdot \nabla) \mathbf{v}$ , write it in the following form found by exposing Cartesian tensor labels

$$\omega_m \partial_m v_n = (\omega_m/2) [(\partial_m v_n + \partial_n v_m) + (\partial_m v_n - \partial_n v_m)] \quad (6.71a)$$

$$= \omega_m \mathbb{S}_{mn} - \omega_m \mathbb{R}_{mn}, \quad (6.71b)$$

where  $\mathbb{S}_{mn} = (1/2)(\partial_n v_m + \partial_m v_n)$  are components to the [strain rate tensor](#) and  $\mathbb{R}_{mn} = (1/2)(\partial_n v_m - \partial_m v_n)$  are components to the [rotation tensor](#). These tensors were introduced in [VOLUME 2](#) when studying the kinematics of line elements. The rotation tensor is related to the vorticity, thus allowing for the identity

$$2\omega_m \mathbb{R}_{mn} = -\omega_m \epsilon_{mnp} \omega_p = \epsilon_{mnp} \omega_m \omega_p = (\boldsymbol{\omega} \times \boldsymbol{\omega})_n = 0. \quad (6.72)$$

Recalling that the rotation tensor generates rotations about the axis defined by vorticity, we can understand why  $\boldsymbol{\omega} \cdot \mathbb{R} = 0$ . Namely, there is no rotation generated when a vector is rotated about its own axis. We are thus left just with

$$(\boldsymbol{\omega} \cdot \nabla) \mathbf{v} = \boldsymbol{\omega} \cdot \mathbb{S}. \quad (6.73)$$

That is, the source,  $(\boldsymbol{\omega} \cdot \nabla) \mathbf{v}$ , appearing in the vorticity equation is determined by the projection of the vorticity onto the strain rate tensor. This result highlights the fundamental role of flow strains in affecting vorticity.

### 6.5.2 Frozen-in nature of vorticity

Consider an inviscid barotropic fluid in the absence of planetary rotation and whose flow is non-divergent, in which case the vorticity equation (6.42) reduces to

$$\frac{D\omega}{Dt} = (\boldsymbol{\omega} \cdot \nabla) \mathbf{v}, \quad (6.74)$$

and recall the evolution equation for a material line element derived in VOLUME 2

$$\frac{D(\delta\mathbf{x})}{Dt} = (\delta\mathbf{x} \cdot \nabla) \mathbf{v}. \quad (6.75)$$

Now recall from Section 3.7.1 that a vortex line is a line drawn through the fluid that is everywhere parallel to the vorticity. Such a line connects material fluid particles, so that a vortex line constitutes a particular case of a material line. At some initial time,  $t = 0$ , let the vorticity on an infinitesimal vortex line be related to the material line element according to

$$\delta\mathbf{x}(0) = \Gamma \boldsymbol{\omega}(\mathbf{x}, 0), \quad (6.76)$$

where  $\Gamma$  has dimensions  $L T$  and is determined by the initial vorticity and initial line element. Since the vorticity equation (6.74) has precisely the same mathematical form as the material line element equation (6.75), the difference vector

$$\mathbf{A} \equiv \delta\mathbf{x} - \Gamma \boldsymbol{\omega}, \quad (6.77)$$

evolves according to

$$\frac{D\mathbf{A}}{Dt} = (\mathbf{A} \cdot \nabla) \mathbf{v}. \quad (6.78)$$

But since  $\mathbf{A}$  vanishes at  $t = 0$ , we conclude that it vanishes for all time

$$\frac{D\mathbf{A}}{Dt} = 0. \quad (6.79)$$

Consequently, the relation (6.76) holds for all time with  $\Gamma$  a constant. That is, the vortex line and its corresponding line element remain parallel as they both evolve according to their projection onto the strain rate tensor. We thus say that vorticity is a *frozen-in* property as illustrated by Figure 6.6. Although we established this property only for the case of an inviscid, barotropic fluid with non-divergent flow, it offers insight into the more general situation occurring in real fluids.

### 6.5.3 Stretching and tilting of vortex tubes

Vorticity responds when vortex lines or tubes are stretched or bent by the strain rate tensor. To help understand the response, consider again the perfect fluid barotropic vorticity equation with a non-divergent flow (equation (6.74)) and focus on the material evolution of the vertical vorticity component

$$\frac{D\omega_z}{Dt} = \omega_x \frac{\partial w}{\partial x} + \omega_y \frac{\partial w}{\partial y} + \omega_z \frac{\partial w}{\partial z} = \boldsymbol{\omega} \cdot \nabla w = \nabla \cdot (w \boldsymbol{\omega}). \quad (6.80)$$

The following discussion closely emulates that given for the kinematics of a material line element in VOLUME 2.

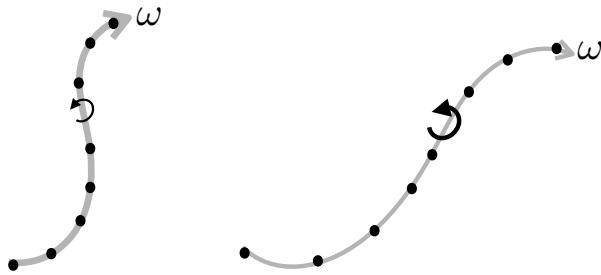


FIGURE 6.6: For the non-divergent flow of a perfect barotropic fluid, vortex lines are also material lines. This property means that for an arbitrary vortex line drawn in the fluid, the fluid particles that are initially on the vortex line remain on the line as it moves through the fluid according to the strain rate tensor. We here show two instances of the same vortex line along with sample test fluid particles. The left configuration stretches into the right configuration, with the vorticity increasing as the vortex line stretches according to the discussion in Section 6.5.3. This property of a vortex line is known as its *frozen-in nature*. The frozen-in nature of vortex lines strictly holds only for perfect barotropic fluid exhibiting non-divergent flow, yet it offers useful insights into the vortex dynamics of more general fluids.

### Stretching

Consider the vortex tube to be initially aligned with the  $\hat{z}$ -axis, so that  $\omega_x = \omega_y = 0$ , in which case there is only a single term impacting the material evolution of vertical vorticity<sup>10</sup>

$$\frac{D\omega_z}{Dt} = \omega_z \frac{\partial w}{\partial z}. \quad (6.81)$$

Since the flow is non-divergent, the volume of an infinitesimal portion of the vortex tube is materially constant

$$\frac{D(\delta V)}{Dt} = 0, \quad (6.82)$$

which means that the vertical extent,  $\delta z$ , and cross-sectional area,  $\delta A$ , are constrained

$$\frac{1}{\delta z} \frac{D(\delta z)}{Dt} + \frac{1}{\delta A} \frac{D(\delta A)}{Dt} = 0. \quad (6.83)$$

As the tube stretches vertically, its horizontal area reduces, and vice versa. Making use of the expression for the evolution of a material line segment (equation (6.75)) allows us to write

$$\frac{1}{\delta z} \frac{D(\delta z)}{Dt} = \frac{\partial w}{\partial z}, \quad (6.84)$$

so that the vorticity equation (6.81) becomes

$$\frac{D\omega_z}{Dt} = \omega_z \frac{\partial w}{\partial z} = \omega_z \left[ \frac{1}{\delta z} \frac{D(\delta z)}{Dt} \right] = -\omega_z \left[ \frac{1}{\delta A} \frac{D(\delta A)}{Dt} \right]. \quad (6.85)$$

Rearrangement leads to

$$\frac{D(\omega_z \delta A)}{Dt} = 0, \quad (6.86)$$

which is an expression of Kelvin's circulation theorem (equation (6.11)) for a horizontal cross-section of the vortex tube.

<sup>10</sup>Be mindful to distinguish the symbols for the vertical component of vorticity,  $\omega_z$ , and the vertical component of velocity,  $w$ .

The above manipulations suggest the following interpretation for the vortex stretching term,  $\omega_z (\partial w / \partial z)$ , appearing in the vertical vorticity equation (6.80) and illustrated in Figure 6.7. Namely, as the vortex tube is stretched and its cross-sectional area is compressed, the vorticity magnitude increases so to maintain a constant circulation around the tube, as per Kelvin's theorem (or equivalently as per Helmholtz's first theorem discussed in Section 3.7.3). Stretching a vortex tube increases the magnitude of the vorticity in the direction of the stretching whereas compressing a tube reduces the vorticity magnitude. This result accords with our understanding of angular momentum conservation as discussed for the rotating cylinder in Section 5.2.2 and depicted by Figure 5.1.

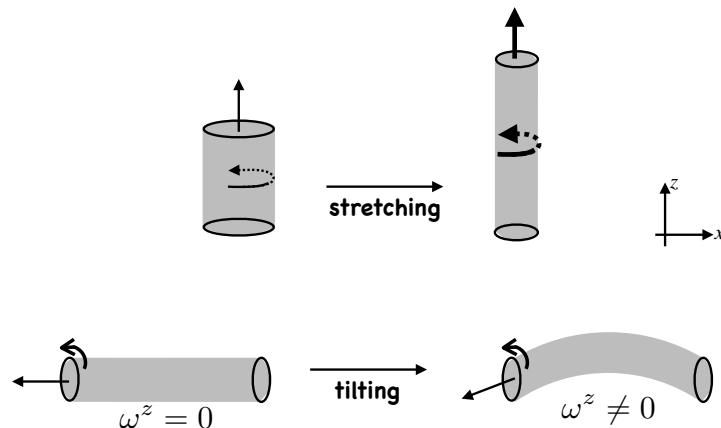


FIGURE 6.7: Illustrating how stretching and tilting of a vortex tube impacts on the vorticity. Top panels: As the cross-sectional area of the vortex tube shrinks, and the vertical extent of the tube stretches, the magnitude of the vorticity along the axis of the tube increases. This result accords with our understanding of angular momentum conservation as discussed for the rotating cylinder in Section 5.2.2 and depicted by Figure 5.1, as well as with Helmholtz's first theorem in Section 3.7.3 and Figure 3.7. Lower panels: The initial vortex tube is assumed to be aligned parallel to the  $x$ -axis, so that it has zero projection in the vertical direction. A horizontal shear of the vertical velocity ( $\partial w / \partial x \neq 0$ ) deforms the vortex tube. Upon deforming (or tilting), the tube picks up a nonzero projection in the vertical, which means that it now has a nonzero vertical component to vorticity.

### Tilting

Now consider an initially horizontal vortex tube as in the lower left panel of Figure 6.7 so that  $\omega_z = 0$ . Furthermore, to focus on just one of the two horizontal directions we set  $\omega_y = 0$  so that equation (6.80) for the vertical vorticity becomes

$$\frac{D\omega_z}{Dt} = \omega_x \frac{\partial w}{\partial x}. \quad (6.87)$$

If there is no horizontal shear in the vertical velocity ( $\partial w / \partial x = 0$ ), then the vortex tube remains horizontal. However, in the presence of  $\partial w / \partial x \neq 0$ , the vorticity vector picks up a nonzero vertical projection. To help visualize this process, recall the frozen-in nature of vortex lines, and consider the evolution of an infinitesimal line segment on the vortex tube. With the vortex tube initially aligned parallel to the  $x$ -axis, the evolution of a material line segment (equation (6.75)) is given by

$$\frac{D(\delta x)}{Dt} = \delta x \frac{\partial v}{\partial x}. \quad (6.88)$$

The initially horizontal line segment thus picks up a projection in the vertical so long as  $\partial w / \partial x \neq 0$ . Correspondingly, the vorticity picks up a vertical component. We can think of this process as a tilting or deforming of the initially horizontal vortex tube, with the tilted tube having a nonzero vertical projection.

#### 6.5.4 Shallow water vorticity revisited

We here revisit our discussion of the shallow water vorticity from Chapter 5 in light of the vorticity equation (6.42). Notably, an inviscid shallow water fluid has zero baroclinicity, so that only stretching and tilting affect shallow water vorticity.

##### Vortex tubes never close in a shallow water layer

The absolute vorticity vector in a shallow water layer is given by equation (1.106d)

$$\boldsymbol{\omega}_a = \nabla \times \mathbf{v} + \hat{\mathbf{z}} f = \boldsymbol{\omega}_h + \hat{\mathbf{z}} (\zeta + f) = \hat{\mathbf{x}} \partial_y w - \hat{\mathbf{y}} \partial_x w + \hat{\mathbf{z}} \zeta_a = -\hat{\mathbf{z}} \times \nabla w + \hat{\mathbf{z}} \zeta_a, \quad (6.89)$$

where we set  $\partial_z u = \partial_z v = 0$  for the horizontal velocity within a shallow water layer. Since the shallow water fluid is hydrostatic, the horizontal vorticity component is much smaller in magnitude than the vertical component,

$$|\partial_x w, \partial_y w| \ll |\zeta|. \quad (6.90)$$

Vortex tubes in a shallow water fluid do not close, since to close requires breaking this inequality. Hence, shallow water vortex tubes reach from the bottom of the layer to the top, with only a slight tilt relative to the vertical.

##### Material time evolution of shallow water vorticity

To determine how shallow water vorticity evolves, we make use of the stretching and tilting term in the form of equation (6.73) so that

$$\frac{D\boldsymbol{\omega}_a}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} \implies \frac{D\omega_{an}}{Dt} = \omega_{am} \mathbb{S}_{mn}. \quad (6.91)$$

The strain rate tensor for the shallow water fluid is

$$\mathbb{S} = \frac{1}{2} \begin{bmatrix} 2\partial_x u & \partial_y u + \partial_x v & \partial_z u + \partial_x w \\ \partial_x v + \partial_y u & 2\partial_y v & \partial_z v + \partial_y w \\ \partial_x w + \partial_z u & \partial_y w + \partial_z v & 2\partial_z w \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2\partial_x u & \partial_y u + \partial_x v & \partial_x w \\ \partial_x v + \partial_y u & 2\partial_y v & \partial_y w \\ \partial_x w & \partial_y w & 2\partial_z w \end{bmatrix}, \quad (6.92)$$

so that material time evolution of the vertical vorticity component, is given by

$$\frac{D(\zeta + f)}{Dt} = \omega_1 \mathbb{S}_{13} + \omega_2 \mathbb{S}_{23} + \hat{\mathbf{z}} \cdot \boldsymbol{\omega}_a \mathbb{S}_{33} \quad (6.93a)$$

$$= (1/2) (\omega_1 \partial_x w + \omega_2 \partial_y w) + \omega_{a3} \partial_z w \quad (6.93b)$$

$$= (\zeta + f) \partial_z w \quad (6.93c)$$

$$= -(\zeta + f) \nabla \cdot \mathbf{u}, \quad (6.93d)$$

which agrees with the shallow water vorticity equation (5.9). For the zonal vorticity component we have

$$\frac{D\omega_1}{Dt} = \omega_1 S_{11} + \omega_2 S_{21} + \hat{z} \cdot \boldsymbol{\omega}_a S_{31} \quad (6.94a)$$

$$= \omega_1 \partial_x u + \omega_2 (\partial_x v + \partial_y u)/2 + \omega_3 \partial_x w, \quad (6.94b)$$

and a similar expression for the meridional component.

### Flux-form evolution of Cartesian vorticity components

In Section 6.3.4, we showed how each of the three Cartesian vorticity components evolves according to a flux-form equation. For the vertical component to the vorticity, the vorticity flux,  $\mathbf{J}^{\zeta_a}$ , is given by equation (6.45) and it takes on the following form for a shallow water layer

$$\mathbf{J}^{\zeta_a} = \mathbf{v} \zeta_a - w \boldsymbol{\omega}_a = \mathbf{u} \zeta_a - w \boldsymbol{\omega}_h = \mathbf{u} \zeta_a + w \hat{z} \times \nabla w = \mathbf{u} \zeta_a + \hat{z} \times \nabla w^2/2. \quad (6.95)$$

The term

$$\hat{z} \times \nabla w^2/2 = -\nabla \times \hat{z} w^2/2 \quad (6.96)$$

has a zero divergence and so has no contribution to the convergence of the vorticity flux. Hence, the vorticity flux,  $\mathbf{J}^{\zeta_a}$ , derived here differs by a gauge from the purely advective flux,  $\mathbf{u} \zeta_a$ , considered in the shallow water vorticity equation (5.7).

For the zonal component to the vorticity, the vorticity flux,  $\mathbf{J}^{\omega_1}$ , is given by equation (6.47) and it takes on the following form for a shallow water layer

$$\mathbf{J}^{\omega_1} = \mathbf{v} \omega_{a1} - u \boldsymbol{\omega}_a = (\hat{y} v + \hat{z} w) \omega_1 - \hat{y} \omega_2 - \hat{z} (\zeta + f) = \hat{y} (v \omega_1 - u \omega_2) + \hat{z} [w \omega_1 - u (\zeta + f)], \quad (6.97)$$

with a similar form for the flux of the meridional vorticity component,  $\mathbf{J}^{\omega_2}$ .

#### 6.5.5 Concerning three-dimensional turbulence

As a vortex tube is stretched in the presence of straining motion, it spins faster as its radius decreases. Hence, its kinetic energy moves from larger to smaller spatial scales. This process of downscale energy cascade (i.e., the movement of kinetic energy from large to small scales) is a fundamental property of three dimensional turbulence, and vortex stretching is the dominant mechanism for the cascade. In contrast, two dimensional turbulence, which occurs in horizontal non-divergent flows, does not support vortex stretching and consequently does not support the downscale energy cascade. Instead, two dimensional turbulence supports an inverse cascade whereby there is a net flow of energy to larger scales, with that flow related to the material conservation of vorticity in two dimensional non-divergent flows (see Chapter 4). [Vallis \(2017\)](#) provides a lucid discussion of energy cascades in both two and three dimensional turbulence.

## 6.6 Circulation viewed in a rotating reference frame

We here tie up a loose end by studying circulation and vorticity for fluids in a rotating reference frame, such as those on a rotating planet. It turns out that incorporating rotation is straightforward, and yet the implications are quite profound for the motion of geophysical fluids. In this section we are careful to make use of [planetary Cartesian coordinates](#), whereby the origin of the coordinate system is at the center of the planet.

Start by recalling the expression from VOLUME 2 for the inertial or **absolute velocity** (i.e., velocity measured in an inertial frame)

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}, \quad (6.98)$$

where  $\mathbf{v}$  is the velocity measured in the rotating frame (relative velocity), and  $\mathbf{x}$  is the position vector relative to the origin (e.g., center of earth). The absolute circulation around an arbitrary circuit (a circuit that is not necessarily material) is thus given by

$$\mathcal{C}_a = \oint_{\partial S} (\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{x} = \mathcal{C} + \mathcal{C}_{\text{planet}}, \quad (6.99)$$

where the circulation measured in the rotating reference frame is

$$\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} \quad (6.100)$$

and the circulation associated with the rotating planet is

$$\mathcal{C}_{\text{planet}} = \oint_{\partial S} (\boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{x}. \quad (6.101)$$

A fluid element at rest in the rotating reference frame still has a nonzero absolute circulation as given by the planetary circulation. Making use of Stokes' theorem leads to the equivalent forms for the circulations

$$\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS \quad \text{relative circulation} \quad (6.102a)$$

$$\mathcal{C}_{\text{planet}} = \oint_{\partial S} (\boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{x} = \int_S \boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{n}} dS \quad \text{planetary circulation} \quad (6.102b)$$

$$\mathcal{C}_a = \oint_{\partial S} \mathbf{v}_a \cdot d\mathbf{x} = \int_S \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} dS \quad \text{absolute circulation,} \quad (6.102c)$$

where

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} \quad \text{relative vorticity} \quad (6.103a)$$

$$\boldsymbol{\omega}_{\text{planet}} = \nabla \times (\boldsymbol{\Omega} \times \mathbf{x}) = 2\boldsymbol{\Omega} \quad \text{planetary vorticity} \quad (6.103b)$$

$$\boldsymbol{\omega}_a = \nabla \times (\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) = \boldsymbol{\omega} + \boldsymbol{\omega}_{\text{planet}} \quad \text{absolute vorticity.} \quad (6.103c)$$

Thus far we have merely substituted in the expression (6.98) for the inertial velocity and then decomposed the vorticity and circulation into its relative and planetary components. Next we consider how circulation evolves, in which case we see how the relative and planetary circulations interact.

### 6.6.1 Material evolution of absolute circulation

Consider how the absolute circulation evolves for a material circuit that moves with the fluid

$$\frac{d\mathcal{C}_a}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} \mathbf{v}_a \cdot d\mathbf{x} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} (\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{x}. \quad (6.104)$$

We measure fluid motion in the rotating frame so that the material time derivative is computed with the velocity,  $\mathbf{v}$ , rather than the absolute velocity,  $\mathbf{v}_a$ . Following the derivation of Kelvin's circulation theorem in a non-rotating reference frame from Section 6.2 leads to

$$\frac{d\mathcal{C}_a}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} (\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{x} \quad (6.105a)$$

$$= \oint_{\partial S(\mathbf{v})} \left[ \frac{D\mathbf{v}}{Dt} + \boldsymbol{\Omega} \times \frac{D\mathbf{x}}{Dt} \right] \cdot d\mathbf{x} + \oint_{\partial S(\mathbf{v})} (\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{v} \quad (6.105b)$$

$$= \oint_{\partial S(\mathbf{v})} \left[ \frac{D\mathbf{v}}{Dt} + \boldsymbol{\Omega} \times \mathbf{v} \right] \cdot d\mathbf{x} + \oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{v} \quad (6.105c)$$

$$= \oint_{\partial S(\mathbf{v})} \left[ \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \times \mathbf{v} \right] \cdot d\mathbf{x}. \quad (6.105d)$$

To reach this result we set

$$\mathbf{v} = \frac{D\mathbf{x}}{Dt}, \quad (6.106)$$

for the velocity of a fluid particle on the circuit. We also used the identity

$$\oint_{\partial S(\mathbf{v})} \mathbf{v} \cdot d\mathbf{v} = \frac{1}{2} \oint_{\partial S(\mathbf{v})} d(\mathbf{v} \cdot \mathbf{v}) = 0 \quad (6.107)$$

as well as

$$\oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \times \mathbf{x}) \cdot d\mathbf{v} = \oint_{\partial S(\mathbf{v})} d[(\boldsymbol{\Omega} \times \mathbf{x}) \cdot \mathbf{v}] - \oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \times d\mathbf{x}) \cdot \mathbf{v} = \oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \times \mathbf{v}) \cdot d\mathbf{x}, \quad (6.108)$$

where we set

$$\oint_{\partial S(\mathbf{v})} d[(\boldsymbol{\Omega} \times \mathbf{x}) \cdot \mathbf{v}] = 0 \quad (6.109)$$

since, as for equation (6.107), the closed loop integral of an exact spatial differential vanishes. We also noted that  $\boldsymbol{\Omega}$  is a constant vector so that  $d\boldsymbol{\Omega} = 0$ . Now insert the momentum equation (6.31) into equation (6.105d) to yield

$$\frac{d\mathcal{C}_a}{dt} = \oint_{\partial S(\mathbf{v})} \left[ \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \times \mathbf{v} \right] \cdot d\mathbf{x}. \quad (6.110a)$$

$$= \oint_{\partial S(\mathbf{v})} \left[ -\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F} \right] \cdot d\mathbf{x}. \quad (6.110b)$$

$$= \oint_{\partial S(\mathbf{v})} \left[ -\frac{dp}{\rho} + \mathbf{F} \cdot d\mathbf{x} \right]. \quad (6.110c)$$

Making use of Stokes' theorem leads to the evolution of absolute circulation around a material loop

$$\frac{d\mathcal{C}_a}{dt} = \oint_{\partial S(\mathbf{v})} \left[ -\frac{dp}{\rho} + \mathbf{F} \cdot d\mathbf{x} \right] = \int_{S(\mathbf{v})} (\mathbf{B} + \nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS, \quad (6.111)$$

where  $\mathbf{B} = \rho^{-2} \nabla \rho \times \nabla p$  is the baroclinicity vector from equation (6.62).

The circulation theorem (6.111) is the same as obtained for Kelvin's circulation theorem in a non-rotating reference frame as discussed in Section 6.2 (see equation (6.5e)). As such, we find that time changes to the absolute circulation are affected by the work applied by pressure and friction when integrated around the material circuit. Evidently, the formalism confirms

that absolute circulation is a frame invariant property of the fluid, in which its evolution is unchanged when moving to a non-inertial rotating frame.

### 6.6.2 The beta effect

As given by equation (6.99), the absolute circulation around an arbitrary circuit equals to the circulation of fluid measured in the rotating frame (relative circulation) plus circulation of the rotating frame itself (planetary circulation)

$$\mathcal{C}_a = \mathcal{C} + \mathcal{C}_{\text{planet}} = \mathcal{C} + 2 \int_S \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS \iff \frac{d\mathcal{C}_a}{dt} = \frac{d\mathcal{C}}{dt} + \frac{d\mathcal{C}_{\text{planet}}}{dt}. \quad (6.112)$$

We can determine the processes that affect the absolute circulation around a material loop by using the circulation theorem (6.111)

$$\frac{d\mathcal{C}}{dt} = -\frac{d\mathcal{C}_{\text{planet}}}{dt} + \frac{d\mathcal{C}_a}{dt} \quad (6.113a)$$

$$= -2 \frac{d}{dt} \left[ \int_{S(v)} \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS \right] + \int_{S(v)} (\mathbf{B} + \nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad (6.113b)$$

We generally assume that the planetary rotation is a constant in time and points through the north pole of the sphere<sup>11</sup>  $\boldsymbol{\Omega} = \Omega \hat{\mathbf{Z}}$ , so that

$$\int_{S(v)} \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS = \Omega \int_{S(v)} \hat{\mathbf{Z}} \cdot \hat{\mathbf{n}} dS = \Omega A_{\perp}. \quad (6.114)$$

The area,  $A_{\perp}$ , is the projection of the spherical area enclosed by the circuit onto the horizontal equatorial plane, with Figure 6.8 illustrating the geometry. This result has profound impact on large scale geophysical fluid motion, whereby relative circulation around a material circuit in the rotating frame changes according to

$$\frac{d\mathcal{C}}{dt} = \underbrace{-2\Omega \frac{dA_{\perp}}{dt}}_{\text{beta effect}} + \underbrace{\int_S (\mathbf{B} + \nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS}_{\text{baroclinicity plus friction curl}}(v). \quad (6.115)$$

Equation (6.115) is sometimes referred to as the [Bjerknes circulation theorem](#) (see [Holton and Hakim \(2013\)](#) equation (4.5)). The second term, comprised of baroclinicity and friction, also appears in case of a non-rotating reference frame that we studied in Sections 6.2 and 6.4.

The first term in the circulation theorem (6.115) is fundamentally new. It is nonzero in the presence of both rotation and curvature of the sphere. The spherical effect arises from latitudinal movement of a material circuit, with the area,  $A_{\perp}$ , changing under such motion. When the circuit moves poleward, the projected area,  $A_{\perp}$ , increases whereas it decreases to zero as it moves equatorward. The material change in  $A_{\perp}$ , when multiplied by the planetary vorticity, modifies the relative circulation around the material circuit. We refer to [planetary induction](#) as the process whereby relative circulation is modified by latitudinal motion of a material circuit on a rotating sphere. Or more commonly, planetary induction is referred to as the [beta effect](#) ( $\beta$ -effect), given its connection to the latitudinal gradient of the Coriolis parameter,  $\beta = \partial_y f$ . Notably, longitudinal motion of the circuit has no impact on  $A_{\perp}$ , so that

<sup>11</sup>Note that in equation (6.114), the vertical Cartesian direction through the north pole is written  $\hat{\mathbf{Z}}$ , to avoid confusion with the local vertical direction  $\hat{\mathbf{z}}$  determined by the geopotential.

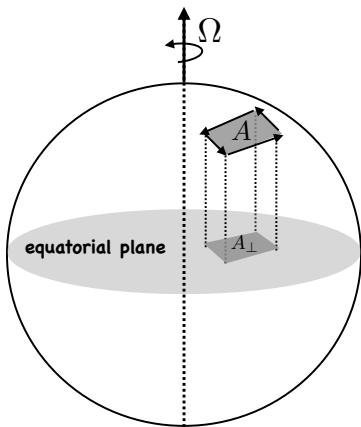


FIGURE 6.8: Geometry of the beta effect on a rotating sphere. According to the Bjerknes circulation theorem (6.115), the circulation for a material loop on the surface of a rotating sphere is affected by baroclinicity and friction, as for a non-rotating sphere, as well as latitudinal motion of the loop. The latitudinal motion alters the area of the loop as projected onto the equatorial plane, with the projected area increasing as the loop moves poleward. When multiplied by the magnitude of the planetary vorticity,  $2\Omega$ , the area contribution is termed planetary induction (i.e., relative circulation is induced by latitudinal motion), or more commonly it is called the beta effect ( $\beta$ -effect). The beta effect requires both rotation ( $2\Omega$ ) and curvature of the sphere ( $\partial_y f = \beta$ ); it is therefore absent on the  $f$ -plane.

longitudinal motion imparts no planetary induction of relative circulation.

In theories of large-scale laminar planetary flows, the baroclinicity and friction terms are typically sub-dominant. For these flows, the material evolution of relative circulation is dominated by the beta effect. Flow satisfying the assumptions of planetary geostrophy is the canonical example of such flow, as studied in VOLUME 2 as well as Chapters 9 and 10. In such flows, forces that lead to meridional motion also give rise to changes in the relative circulation. Conversely, forces that change the circulation around a material loop affect meridional motion of the loop.

### 6.6.3 The case of two-dimensional non-divergent flow

To garner further insight into the nature of the beta effect, consider a perfect (i.e., inviscid) two-dimensional and non-divergent flow (zero vertical velocity) on a rotating sphere. In this case there is only a vertical component to vorticity and baroclinicity vanishes. Hence, relative vorticity is affected only via the beta effect. In addition, the fluid flow materially preserves the area of any material region. This two-dimensional non-divergent barotropic flow is discussed in more detail in Chapter 4. We here use it as an example to expose essential features of the beta effect (see also Section 4.3.2).

In the rotating frame, circulation around an infinitesimal closed material loop is

$$\mathcal{C} = A \zeta, \quad (6.116)$$

where  $\zeta$  is the relative vorticity and  $A$  is the area enclosed by the loop. Because the fluid flow is non-divergent, the loop area  $A$  remains constant even as the loop becomes contorted.<sup>12</sup> This area preservation property simplifies the evolution equation for the circulation, which is given

---

<sup>12</sup>This kinematic result is derived in VOLUME 2, as part of our study of kinematics of non-divergent flows.

by

$$\frac{D\mathcal{C}}{Dt} = \frac{D(A\zeta)}{Dt} = A \frac{D\zeta}{Dt}. \quad (6.117)$$

Equating this result to the circulation change implied by Bjerknes' circulation theorem (6.115) renders

$$\frac{D\mathcal{C}}{Dt} = A \frac{D\zeta}{Dt} = -2\Omega \frac{DA_{\perp}}{Dt}. \quad (6.118)$$

Let the material circuit be at a latitude,  $\phi$ , so that the projection of the loop area onto the equatorial plane is (see Figure 6.8)

$$A_{\perp} = A \sin \phi. \quad (6.119)$$

Hence, material evolution of the circulation is

$$\frac{D\mathcal{C}}{Dt} = A \frac{D\zeta}{Dt} \quad (6.120a)$$

$$= -2\Omega \frac{DA_{\perp}}{Dt} \quad (6.120b)$$

$$= -2A\Omega \frac{D \sin \phi}{Dt} \quad (6.120c)$$

$$= -2A\Omega \cos \phi \frac{D\phi}{Dt} \quad (6.120d)$$

$$= -A \left[ \frac{2\Omega \cos \phi}{R} \right] \left[ R \frac{D\phi}{Dt} \right] \quad (6.120e)$$

$$= -A\beta v, \quad (6.120f)$$

where we introduced the meridional velocity component

$$v = R \frac{D\phi}{Dt} \quad (6.121)$$

and the meridional derivative of the planetary vorticity

$$\beta = \frac{df}{dy} = \frac{1}{R} \frac{d}{d\phi} (2\Omega \sin \phi) = \frac{2\Omega \cos \phi}{R}. \quad (6.122)$$

The result (6.120f)

$$\frac{1}{A} \frac{D\mathcal{C}}{Dt} = \frac{D\zeta}{Dt} = -\beta v, \quad (6.123)$$

shows how meridional motion on a rotating sphere induces relative circulation, and thus relative vorticity. It furthermore motivates the name **beta effect** ( $\beta$ -effect) for the planetary induction.

#### 6.6.4 Planetary circulation, planetary vorticity, and the Coriolis acceleration

From equation (6.103b) we know that the planetary vorticity is given by the constant vector

$$\boldsymbol{\omega}_{\text{planet}} = \nabla \times (\boldsymbol{\Omega} \times \mathbf{x}) = 2\boldsymbol{\Omega}. \quad (6.124)$$

As a constant vector, it is the same everywhere in space. However, its impact on the fluid circulation and vorticity depends on what latitude the rotation vector is sampled. We here focus on the radial component of the planetary vorticity by measuring the circulation per area for fixed radius circuits, with reference to Figure 6.9.

### Planetary circulation centered on the pole and on the equator

Equation (6.114) leads to the planetary circulation

$$\mathcal{C}_{\text{planet}} = 2\Omega A_{\perp}, \quad (6.125)$$

for an arbitrary constant radius circuit. Stokes' theorem then says that the planetary vorticity, as projected onto the local radial direction (outward normal to the surface), has value

$$\boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{n}} \approx \mathcal{C}_{\text{planet}} / \mathcal{S} = 2\Omega A_{\perp} / \mathcal{S}, \quad (6.126)$$

where  $\mathcal{S} = \int_S d\mathcal{S}$  is the area enclosed by the circuit. In the limit that the circuit becomes infinitesimal, then  $A_{\perp} \rightarrow \mathcal{S}$  when the circuit is centered on the north pole, whereas  $A_{\perp} \rightarrow 0$  for an equatorially centered circuit. Correspondingly, the planetary vorticity, when projected into the radial direction, is  $2\Omega$  at the north pole (and  $-2\Omega$  at the south pole), whereas it vanishes at the equator. We emphasize that it is not the planetary vorticity that vanishes at the equator, which is obvious since  $\boldsymbol{\omega}_{\text{planet}} = 2\Omega$  is a constant vector. Rather, it is the radial projection,  $\boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{r}}$ , that vanishes at the equator.

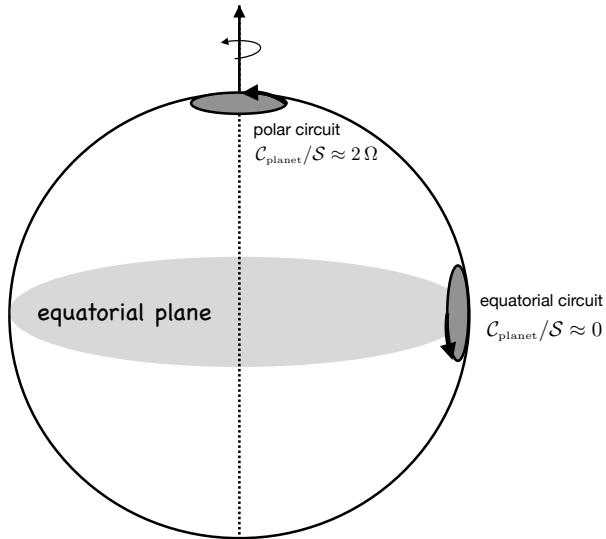


FIGURE 6.9: Area mean of the planetary vorticity as computed around two closed circular loops with area  $\mathcal{S}$ . One circuit is centered on the north pole, in which case the area mean planetary circulation,  $\mathcal{C}_{\text{planet}}/\mathcal{S} \approx 2\Omega$ . The other circuit is centered on the equator so that  $\mathcal{C}_{\text{planet}}/\mathcal{S} \approx 0$ . Stokes' theorem says that the planetary vorticity, as projected onto the local radial direction (outward normal), has value  $\boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{r}} \approx \mathcal{C}_{\text{planet}}/\mathcal{S}$ . Hence, the planetary vorticity equals  $2\Omega$  at the north pole (and  $-2\Omega$  at the south pole), whereas it vanishes at the equator.

### Comments on the Coriolis acceleration

Although  $\boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{r}} = 0$  at the equator, the Coriolis acceleration does not generally vanish there. That is, recall the discussion from VOLUME 2, where the Coriolis acceleration in spherical coordinates is written

$$\mathbf{A}_{\text{coriolis}} = -2\Omega \left[ \hat{\lambda} (w \cos \phi - v \sin \phi) + \hat{\phi} u \sin \phi - \hat{r} u \cos \phi \right], \quad (6.127)$$

which, at  $\phi = 0$ , is

$$\mathbf{A}_{\text{coriolis}}(\phi = 0) = -2\Omega (w \hat{\lambda} - u \hat{r}). \quad (6.128)$$

Evidently, a nonzero Coriolis acceleration at the equator arises since it depends on  $\Omega = \omega_{\text{planet}}/2$  rather than just its radial projection,  $\omega_{\text{planet}} \cdot \hat{\mathbf{r}}$ . For large-scale flows, we commonly ignore  $\mathbf{A}_{\text{coriolis}}(\phi = 0)$  since its radial term is tiny relative to the gravitational acceleration, and the longitudinal term is small for large-scale flows where the vertical velocity is typically small.<sup>13</sup> Indeed, these points were made in VOLUME 2, whereby the Coriolis acceleration for large-scale planetary flows is approximated by

$$\mathbf{A}_{\text{Coriolis}}^{\text{large-scale}} \equiv -2\Omega \sin \phi (-\hat{\lambda} v + \hat{\phi} u) \equiv -f \hat{\mathbf{r}} \times \mathbf{v}. \quad (6.129)$$

This approximate Coriolis acceleration does vanish at the equator, and it is the form resulting from the [Traditional Approximation](#) used for the hydrostatic primitive equations (see VOLUME 2).

### 6.6.5 Further study

The beta effect and its role in vorticity is nicely summarized in [this video from Science Primer](#) in the context of Rossby waves.

## 6.7 Vorticity budget for a primitive equation Boussinesq ocean

In this section we develop the vorticity budget for a hydrostatic primitive equation [Boussinesq ocean](#) in the presence of diabatic sources and frictional forcing. This system is of particular importance for ocean circulation models. The governing primitive equations, as derived in VOLUME 2, are given by

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{v} = -\nabla_h \varphi + \mathbf{F} \quad (6.130a)$$

$$\partial_z \varphi = b \quad (6.130b)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (6.130c)$$

$$\frac{Db}{Dt} = \dot{b}, \quad (6.130d)$$

with the non-divergent velocity field written

$$\mathbf{v} = (\mathbf{u}, w) = \mathbf{u} + w \hat{\mathbf{z}}. \quad (6.131)$$

The perturbation pressure is given by

$$\rho_0 \varphi = \delta p = p - p_0, \quad (6.132)$$

with the reference pressure,  $p_0 = p_0(z)$ , in hydrostatic balance with the constant reference density

$$\frac{dp_0}{dz} = -g \rho_0, \quad (6.133)$$

and  $p$  the hydrostatic pressure satisfying the local hydrostatic balance

$$\partial_z p = -g \rho. \quad (6.134)$$

---

<sup>13</sup> [Stewart and Dellar \(2011\)](#) argue for the importance of the full expression of the Coriolis acceleration (6.127) for the dynamics of cross-equatorial abyssal ocean flows.

The globally referenced Archimedean buoyancy is given by

$$b = -g(\rho - \rho_0)/\rho_0, \quad (6.135)$$

with this field discussed in VOLUME 2. As mentioned at the end of Section 6.6.4, we here assume the Coriolis acceleration of the form relevant to the Traditional Approximation, in which we are only concerned with the local vertical component of planetary rotation so that

$$f \hat{z} \times \mathbf{v} = f \hat{z} \times \mathbf{u}. \quad (6.136)$$

Finally, the friction acceleration vector is horizontal

$$\mathbf{F} = (F^x, F^y, 0) \quad (6.137)$$

and the gradient operator is decomposed into its horizontal plus vertical contribution.

$$\nabla = \nabla_h + \hat{z} \partial_z. \quad (6.138)$$

### 6.7.1 Deriving the vorticity equation

#### Vector invariant velocity equation

To derive the vorticity equation, it is useful to combine the horizontal momentum equation with the hydrostatic balance, in which case

$$\frac{D\mathbf{u}}{Dt} + f \hat{z} \times \mathbf{v} = -\nabla \varphi + b \hat{z} + \mathbf{F}. \quad (6.139)$$

As for the non-hydrostatic case (Section 6.3.1), we rewrite the self-advection operator,  $(\mathbf{v} \cdot \nabla) \mathbf{u}$ , before taking the curl. In turn, we introduce the hydrostatic relative vorticity given by the curl of the horizontal velocity

$$\boldsymbol{\omega}^{hy} = \nabla \times \mathbf{u} = \hat{z} \times \partial_z \mathbf{u} + \hat{z} \zeta = -\hat{x} \partial_z v + \hat{y} \partial_z u + \hat{z} \zeta, \quad (6.140)$$

where

$$\zeta = \partial_x v - \partial_y u \quad (6.141)$$

is the vertical component to the relative vorticity, and the hydrostatic vorticity is non-divergent

$$\nabla \cdot \boldsymbol{\omega}^{hy} = 0. \quad (6.142)$$

It is then straightforward to show that

$$\boldsymbol{\omega}^{hy} \times \mathbf{v} = \hat{x} (w \partial_z u - v \partial_x v + v \partial_y u) + \hat{y} (w \partial_z v - u \partial_y u + u \partial_x v) - \hat{z} \partial_z (u^2 + v^2)/2 \quad (6.143a)$$

$$= w \partial_z \mathbf{u} + \zeta (-v \hat{x} + u \hat{y}) - \hat{z} \partial_z (u^2 + v^2)/2, \quad (6.143b)$$

in which case

$$\nabla (\mathbf{u}^2/2) + \boldsymbol{\omega}^{hy} \times \mathbf{v} = \nabla (u^2 + v^2)/2 - \hat{z} \partial_z (u^2 + v^2)/2 + w \partial_z \mathbf{u} + \zeta (-v \hat{x} + u \hat{y}) \quad (6.144a)$$

$$= (u \partial_x + v \partial_y + w \partial_z) \mathbf{u} \quad (6.144b)$$

$$= (\mathbf{v} \cdot \nabla) \mathbf{u}. \quad (6.144c)$$

The material time derivative of the horizontal velocity can thus be written

$$\frac{D\mathbf{u}}{Dt} = \partial_t \mathbf{u} + (\mathbf{v} \cdot \nabla) \mathbf{u} = \partial_t \mathbf{u} + \boldsymbol{\omega}^{\text{hy}} \times \mathbf{v} + \nabla(\mathbf{u}^2/2), \quad (6.145)$$

which then leads to the vector invariant horizontal velocity equation

$$\partial_t \mathbf{u} + (f \hat{\mathbf{z}} + \boldsymbol{\omega}^{\text{hy}}) \times \mathbf{v} = -\nabla(\varphi + \mathbf{u}^2/2) + b \hat{\mathbf{z}} + \mathbf{F}, \quad (6.146)$$

which can be written in the equivalent form<sup>14</sup>

$$(\partial_t + w \partial_z) \mathbf{u} + (f + \zeta) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h(\varphi + \mathbf{u}^2/2) - (\partial_z \varphi - b) \hat{\mathbf{z}} + \mathbf{F}. \quad (6.147)$$

### Curl of the velocity equation to render the vorticity equation

Now take the curl of the vector invariant velocity equation (6.146), and make use of the identity

$$\nabla \times (\boldsymbol{\omega}_a^{\text{hy}} \times \mathbf{v}) = (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a^{\text{hy}} - (\boldsymbol{\omega}_a^{\text{hy}} \cdot \nabla) \mathbf{v}, \quad (6.148)$$

where we introduced the absolute vorticity for a hydrostatic fluid

$$\boldsymbol{\omega}_a^{\text{hy}} = f \hat{\mathbf{z}} + \boldsymbol{\omega}^{\text{hy}}. \quad (6.149)$$

The result is the vorticity equation

$$\partial_t \boldsymbol{\omega}^{\text{hy}} + (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a^{\text{hy}} = (\boldsymbol{\omega}_a^{\text{hy}} \cdot \nabla) \mathbf{v} + \nabla \times \hat{\mathbf{z}} b + \nabla \times \mathbf{F}. \quad (6.150)$$

Since the Coriolis parameter is time independent, we can add it to the time derivative to yield an equation for absolute vorticity

$$\frac{D\boldsymbol{\omega}_a^{\text{hy}}}{Dt} = \underbrace{(\boldsymbol{\omega}_a^{\text{hy}} \cdot \nabla) \mathbf{v}}_{\text{stretching + tilting}} + \underbrace{\nabla \times \hat{\mathbf{z}} b}_{\text{baroclinicity}} + \underbrace{\nabla \times \mathbf{F}}_{\text{friction curl}}. \quad (6.151)$$

It is notable that the pressure gradient is eliminated from the Boussinesq vorticity equation. Even so, the vorticity is affected by baroclinicity as manifested through horizontal gradients in the buoyancy field, which we discuss next.

### 6.7.2 Boussinesq baroclinicity

Let us compare the Boussinesq vorticity equation (6.151) to the vorticity equation (6.42) for a non-hydrostatic and non-Boussinesq fluid. One difference concerns the form of the vorticity, which differs due to the use of only the horizontal velocity,  $\boldsymbol{\omega}^{\text{hy}} = \nabla \times \mathbf{u}$ , for the hydrostatic fluid whereas the full velocity is used for the non-hydrostatic case,  $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ . Even so, both vorticity equations have a vorticity source due to stretching and tilting, and both have a source due to the curl of friction. The key difference arises in the form of the baroclinicity vector. Namely, the Boussinesq baroclinicity does not involve the Boussinesq pressure gradient acceleration

<sup>14</sup>As discussed in [Griffies et al. \(2020\)](#), the form (6.147) is commonly used for Boussinesq and hydrostatic ocean models.

since it has zero curl. Instead, Boussinesq baroclinicity is given by<sup>15</sup>

$$\mathbf{B}_{\text{bouss}} = \nabla \times \hat{\mathbf{z}} b = \nabla b \times \hat{\mathbf{z}}. \quad (6.152)$$

Boussinesq baroclinicity has a somewhat simpler form than baroclinicity in a compressible fluid, as given by equation (6.62)

$$\mathbf{B} = (\nabla \rho \times \nabla p) / \rho^2 = -\nabla \times (\rho^{-1} \nabla p). \quad (6.153)$$

Again, the fundamental difference arises since the Boussinesq pressure gradient acceleration is annihilated when taking the curl of the velocity equation to produce the Boussinesq vorticity equation. So rather than arise from the misalignment of pressure and density isolines, Boussinesq baroclinicity arises from the misalignment of the gravity field and density gradients.

One practical feature of the Boussinesq baroclinicity (6.152) is that we can readily deduce the presence of baroclinicity (either for the non-hydrostatic or hydrostatic Boussinesq ocean) merely by noting whether there is a slope to the buoyancy surfaces relative to the horizontal (e.g., Figure 6.10). That is, a sloping buoyancy surface provides a vorticity source for the Boussinesq ocean.

### Boussinesq baroclinicity only affects a horizontal vorticity tendency

Given that the Boussinesq baroclinicity (6.152) does not involve the pressure gradient acceleration, we must modify the physical interpretation offered in Section 6.4. In particular, the curl of the non-Boussinesq pressure acceleration has components in all three directions so that the non-Boussinesq baroclinicity affects a source for each of the three vorticity components. In contrast, the Boussinesq baroclinicity is the curl of the Archimedean buoyant acceleration and this acceleration acts only in the vertical. Consequently, the Boussinesq baroclinicity has no direct affect on the vertical component to absolute vorticity

$$\hat{\mathbf{z}} \cdot \mathbf{B}_{\text{bouss}} = \hat{\mathbf{z}} \cdot (\nabla \times \hat{\mathbf{z}} b) = 0. \quad (6.154)$$

Rather, Boussinesq baroclinicity only acts directly as a source for horizontal vorticity. Thus, Boussinesq baroclinicity can only indirectly affect vertical vorticity through the effects of baroclinicity on vertical velocity and the corresponding vertical stretching.

### Comments on shallow water vorticity

In Chapter 5, we studied vorticity in the shallow water fluid. As noted in Section 5.3.3, we are only concerned with the vertical component to vorticity in the shallow water fluid, since the horizontal components are tiny by comparison. Furthermore, equation (5.28) says that the vertical component to shallow water absolute vorticity is materially altered only through material changes to the layer thickness. There is no impact from baroclinicity on the shallow water vorticity. The absence of baroclinicity follows trivially from the absence of any horizontal buoyancy gradients within the shallow water layer. In this manner, the shallow water layer is barotropic.

---

<sup>15</sup>We see in Exercise 7.2 that the baroclinicity vector (6.152) also applies for the non-hydrostatic Boussinesq ocean.

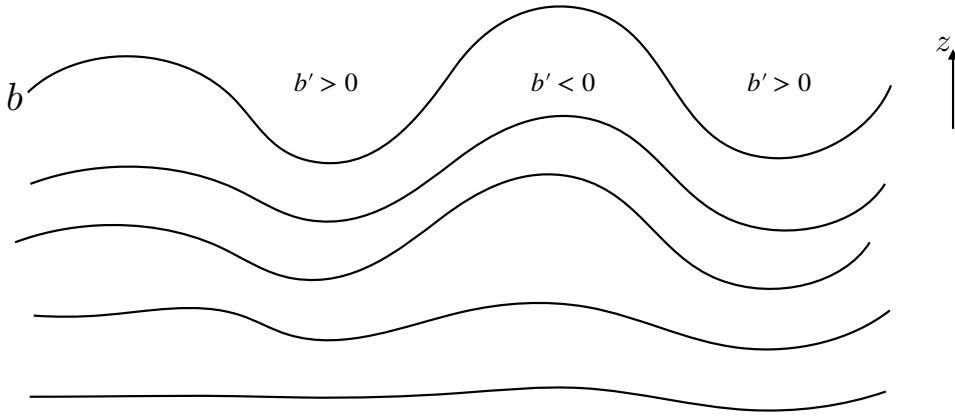


FIGURE 6.10: Baroclinicity in a Boussinesq ocean is manifest by nonzero horizontal gradients in the buoyancy field. Here we depict a region of relatively strong baroclinicity above a region of weaker baroclinicity. A sloping buoyancy surface is therefore synonymous with a nontrivial baroclinic structure. We label anomalously positive ( $b' > 0$ ) and negative buoyancy ( $b' < 0$ ), where the prime denotes anomalies relative to a horizontal average. Furthermore, as per equation (6.154), baroclinicity in a Boussinesq ocean only acts as a source for horizontal vorticity.

### 6.7.3 Vertical vorticity equation

Following the discussion in Section 6.3.4 for the unapproximated vorticity, we here examine the vertical component of the hydrostatic and Boussinesq vorticity equation (6.151)

$$\frac{D\zeta_a}{Dt} = (\boldsymbol{\omega}_a^{hy} \cdot \nabla) w + \hat{z} \cdot (\nabla \times \mathbf{F}), \quad (6.155)$$

with the absence of baroclinicity noted above in Section 6.7.2. The stretching, tilting, and friction curl appearing on the right hand side provide vorticity sources that affect the left hand side's material time evolution. We see this evolution more fully by expanding the terms to render

$$\partial_t \zeta_a + (\mathbf{v} \cdot \nabla) \zeta + \beta v = \hat{z} \cdot (\partial_z \mathbf{u} \times \nabla_h w) + (\zeta + f) \partial_z w + \hat{z} \cdot (\nabla \times \mathbf{F}). \quad (6.156)$$

#### Planetary geostrophic limit

The linearized, inviscid, and steady version of the vorticity equation (6.156) leads to the linear vorticity balance

$$\beta v = f \partial_z w. \quad (6.157)$$

This relation comprises the inviscid vorticity equation for [planetary geostrophy](#). It represents a kinematic balance since no forces are exposed here to explicitly cause motion, though such forces do appear in the momentum equation. Reading the balance from right to left indicates that any process generating vorticity via vortex stretching must be balanced by meridional motion. That is, the fluid responds to vortex stretching by moving meridionally through the planet's vorticity field. Since the vorticity of a planetary geostrophic fluid is solely determined by planetary vorticity, meridional movement is the only means for the fluid to balance vortex sources. Conversely, reading the equality from left to right reveals that any meridional motion itself must be balanced by vortex stretching.

### Vorticity flux vector

We can write the vorticity equation (6.156) in an alternative form by making use of  $\nabla \cdot \mathbf{v} = \nabla \cdot \boldsymbol{\omega}_a^{hy} = 0$  to yield

$$\partial_t \zeta_a = -\nabla \cdot (\mathbf{v} \zeta_a - w \boldsymbol{\omega}_a^{hy}) + \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}). \quad (6.158)$$

Furthermore, we can use equation (6.44d) for the friction curl,  $\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}) = -\nabla \cdot (\hat{\mathbf{z}} \times \mathbf{F})$ . Hence, the vertical component of the Boussinesq vorticity evolves according to the convergence of the vorticity flux

$$\partial_t \zeta_a = -\nabla \cdot \mathbf{J}^{\zeta_a} \quad \text{with} \quad \mathbf{J}^{\zeta_a} = \mathbf{v} \zeta_a - w \boldsymbol{\omega}_a^{hy} + \hat{\mathbf{z}} \times \mathbf{F}, \quad (6.159)$$

which can be compared to the vorticity flux (6.45) for the compressible nonhydrostatic fluid. Again, the main difference arises from the absence of a baroclinicity contribution for the hydrostatic Boussinesq ocean.

The identity  $\boldsymbol{\omega}_a^{hy} = \hat{\mathbf{z}} \times \partial_z \mathbf{u} + \hat{\mathbf{z}} \zeta_a$  allows us to write

$$\mathbf{v} \zeta_a - w \boldsymbol{\omega}_a^{hy} = \mathbf{u} \zeta_a - w \hat{\mathbf{z}} \times \partial_z \mathbf{u}, \quad (6.160)$$

which is a horizontal vector. Furthermore, note that  $\hat{\mathbf{z}} \times \mathbf{F}$  is a horizontal vector, which then means that there is no vertical contribution to the vorticity flux vector,  $\mathbf{J}^{\zeta_a} \cdot \hat{\mathbf{z}} = 0$ . We previously encountered this property in Section 6.3.4 when discussing the vorticity flux for the non-Boussinesq fluid, with Figure 6.1 providing a schematic.

## 6.8 Evolution of depth integrated vertical vorticity

In this section we study the depth integral of the vertical vorticity equation (6.45)

$$\partial_t \zeta_a = -\nabla \cdot \mathbf{J}^{\zeta_a} \quad \text{with} \quad \mathbf{J}^{\zeta_a} = \mathbf{v} \zeta_a - \boldsymbol{\omega}_a w - \hat{\mathbf{z}} \times \rho^{-1} \nabla p + \hat{\mathbf{z}} \times \mathbf{F}. \quad (6.161)$$

We perform the depth integral over the full depth of the ocean from its bottom at  $z = \eta_b(x, y)$  to the ocean surface at  $z = \eta(x, y, t)$  (see Figure 1.1). Studies of the depth integrated vorticity equation allow us to focus on the two dimensional budgets with particular attention to how boundary torques alter the budget. This section anticipates analysis of the depth integrated planetary geostrophic vorticity equation in Section 10.4, with that analysis of use for understanding the role of topography in forcing the large-scale ocean circulation. We also consider the vorticity of the depth integrated flow in Section 6.9, which is commonly considered in numerical applications. Notably, the discussion in the present section does not assume pressure is approximately hydrostatic, so that all results hold for the general case of a non-hydrostatic flow.

### 6.8.1 Comments on the role of baroclinicity

Results of the analysis in this section can be readily specialized to the Boussinesq and hydrostatic case given by equation (6.159). The key distinction, as noted in Section 6.7.2, is that Boussinesq baroclinicity does not directly affect changes to the vertical component of the Boussinesq vorticity (whether hydrostatic or non-hydrostatic; see Exercise 7.2). This property of the Boussinesq baroclinicity means that the boundary pressure torques discussed in Sections 6.8.3 and 6.8.4 play no direct role in the Boussinesq vorticity equation. However, these boundary pressure torques play a direct role in vertical motion next to the boundaries, especially next

to the bottom, with such motion affecting a source to vorticity through stretching. We have much to say in Section 10.4 concerning how boundary pressure torques affect vertical motion for vorticity for the planetary geostrophic fluid. Additionally, as seen in Section 6.9, boundary pressure torques do play a direct role in affecting vorticity of the depth integrated flow in both the Boussinesq and non-Boussinesq fluids.

This discussion exemplifies the sometimes subtle differences between vorticity sources depending on the precise nature of the vorticity, whether it be vorticity for a fluid element as discussed in this section, vorticity of the depth integrated flow in Section 6.9, or vorticity of the depth averaged flow in Section 6.9.7. When studying flavors of vorticity, it is important to be clear on details of their evolution equations since the details color the physical interpretations.

### 6.8.2 Leibniz rule expressions

The necessary manipulations are typical for the analysis of depth integrated budgets, such as considered for the depth integrated momentum equation and depth integrated angular momentum in VOLUME 2. For vorticity we are interested in manipulating following equation

$$\int_{\eta_b}^{\eta} \frac{\partial \zeta_a}{\partial t} dz = - \int_{\eta_b}^{\eta} \nabla \cdot \mathbf{J}^{\zeta_a} dz, \quad (6.162)$$

where  $\mathbf{J}^{\zeta_a}$  is the vorticity flux given by equation (6.161). We make use of Leibniz's rule (VOLUME 2) to move the time and space derivatives from inside the integrals to outside<sup>16</sup>

$$\int_{\eta_b}^{\eta} \frac{\partial \zeta_a}{\partial t} dz = -[\zeta_a \partial_t \eta]_{z=\eta} + \frac{\partial}{\partial t} \int_{\eta_b}^{\eta} \zeta_a dz \quad (6.163)$$

$$- \int_{\eta_b}^{\eta} \nabla_h \cdot \mathbf{J}^{\zeta_a} dz = [\nabla_h \eta \cdot \mathbf{J}^{\zeta_a}]_{z=\eta} - [\nabla_h \eta_b \cdot \mathbf{J}^{\zeta_a}]_{z=\eta_b} - \nabla_h \cdot \int_{\eta_b}^{\eta} \mathbf{J}^{\zeta_a} dz \quad (6.164)$$

$$- \int_{\eta_b}^{\eta} \frac{\partial (\hat{\mathbf{z}} \cdot \mathbf{J}^{\zeta_a})}{\partial z} dz = -[\hat{\mathbf{z}} \cdot \mathbf{J}^{\zeta_a}]_{z=\eta} + [\hat{\mathbf{z}} \cdot \mathbf{J}^{\zeta_a}]_{z=\eta_b}. \quad (6.165)$$

These results then lead to

$$\frac{\partial}{\partial t} \int_{\eta_b}^{\eta} \zeta_a dz = \left[ \zeta_a \partial_t \eta - \nabla(z - \eta) \cdot \mathbf{J}^{\zeta_a} \right]_{z=\eta} + \left[ \nabla(z - \eta_b) \cdot \mathbf{J}^{\zeta_a} \right]_{z=\eta_b} - \nabla_h \cdot \int_{\eta_b}^{\eta} \mathbf{J}^{\zeta_a} dz. \quad (6.166)$$

The time tendency for the depth integral of the vertical component of absolute vorticity, for a fluid column at a fixed horizontal position (left hand side) is determined by a suite of boundary contributions due to baroclinicity, vortex stretching and friction, plus the convergence of the depth integrated vorticity flux (final term on right hand side). We next massage the boundary contributions to expose their associated physical processes.

### 6.8.3 Bottom boundary contribution

The bottom boundary contribution to the vorticity equation (6.166) takes on the form

$$\nabla(z - \eta_b) \cdot \mathbf{J}^{\zeta_a} = \nabla(z - \eta_b) \cdot [\mathbf{v} \zeta_a - \boldsymbol{\omega}_a w - \hat{\mathbf{z}} \times \rho^{-1} \nabla p + \hat{\mathbf{z}} \times \mathbf{F}]_{z=\eta_b} \quad (6.167a)$$

$$= -|\nabla(z - \eta_b)| \hat{\mathbf{n}} \cdot [-\boldsymbol{\omega}_a w - \hat{\mathbf{z}} \times \rho^{-1} \nabla p + \hat{\mathbf{z}} \times \mathbf{F}]_{z=\eta_b}, \quad (6.167b)$$

<sup>16</sup> Recall that since  $\eta(x, y, t)$  is a spatial function just of the horizontal position, there is no difference between  $\nabla \eta$  and  $\nabla_h \eta$ . The same point holds for  $\eta_b(x, y)$  as well.

where we made use of the no-normal flow bottom kinematic boundary condition,  $\hat{\mathbf{n}} \cdot \mathbf{v} = 0$ , and where

$$\hat{\mathbf{n}} = - \left[ \frac{\nabla(z - \eta_b)}{|\nabla(z - \eta_b)|} \right] = - \left[ \frac{\hat{\mathbf{z}} - \nabla_h \eta_b}{\sqrt{1 + \nabla_h \eta_b \cdot \nabla_h \eta_b}} \right] \quad (6.168)$$

is the outward unit normal at the bottom.

### Vortex stretching by vertical flow along a sloping bottom

The first term in the bottom boundary flux (6.167b) provides an inviscid vertical transport of the normal component of the absolute vorticity at the boundary. This term contributes through the action of vertical motion next to a sloping bottom, thus providing a vertical transfer of the vorticity component that is perpendicular to the bottom. This motion provides a form of vortex stretching that vanishes for a flat bottom, in which case  $w(\eta_b) = 0$ . It also vanishes for flow that parallels the bottom, whereby  $\mathbf{u} \cdot \nabla \eta_b = 0$  so that  $w(\eta_b) = 0$  according to the bottom kinematic boundary condition.

### Bottom pressure torques

The second term in the bottom boundary flux (6.167b) arises from baroclinicity next to the bottom, in which case we consider the following term

$$\hat{\mathbf{n}} \cdot [\hat{\mathbf{z}} \times (\rho^{-1} \nabla p)]_{z=\eta_b} = [\rho^{-1} \nabla p]_{z=\eta_b} \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{z}}) \equiv [\rho^{-1} \nabla p]_{z=\eta_b} \cdot \mathbf{t}, \quad (6.169)$$

where we introduced the tangent direction

$$\mathbf{t} = \hat{\mathbf{n}} \times \hat{\mathbf{z}} = \left[ \frac{\nabla \eta_b \times \hat{\mathbf{z}}}{|\nabla(z - \eta_b)|} \right]. \quad (6.170)$$

The vector  $\mathbf{t}$  is horizontal and it points along isolines of constant topography in a direction with land to the left pointing in the direction of  $\mathbf{t}$ , as depicted in Figure 6.11. Since  $\hat{\mathbf{z}}$  and  $\hat{\mathbf{n}}$  are not orthogonal,  $\mathbf{t}$  is not normalized so that it is not adorned with a hat.

Let us decompose the pressure gradient at the bottom according to

$$\nabla p = \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \nabla p) + \hat{\mathbf{t}} (\hat{\mathbf{t}} \cdot \nabla p) \quad z = \eta_b, \quad (6.171)$$

where  $\hat{\mathbf{t}} = \mathbf{t}/|\mathbf{t}|$  is the normalized horizontal tangent vector. Evidently, for the boundary condition (6.169) we only need the  $\hat{\mathbf{t}} (\hat{\mathbf{t}} \cdot \nabla p)$  term. But that term is simply the gradient of the bottom pressure

$$\hat{\mathbf{t}} (\hat{\mathbf{t}} \cdot \nabla p) = \nabla p_b. \quad (6.172)$$

We can provide a bit more thorough derivation of this result by using the methods of generalized vertical coordinates from VOLUME 4, whereby

$$\nabla p = (\hat{\mathbf{z}} - \nabla_\sigma z) \partial_z p + \nabla_\sigma p, \quad (6.173)$$

where  $\sigma$  is an arbitrary generalized vertical coordinate. Evaluating the pressure gradient (6.173) at the ocean bottom ( $z = \eta_b$ ), and letting  $\sigma$  align with the bottom, leads to

$$[\nabla p]_{z=\eta_b} = (\hat{\mathbf{z}} - \nabla \eta_b) \partial_z p + \nabla p_b. \quad (6.174)$$

As before, we conclude that  $\nabla p_b$  is the horizontal component of  $[\nabla p]_{z=\eta_b}$  in the direction tangent

to the bottom.

We thus find the contribution from baroclinicity at the ocean bottom takes the form

$$\nabla(z - \eta_b) \cdot \mathbf{J}_\zeta^{\text{baroclinicity}} = \rho^{-1} \nabla p_b \cdot (\nabla \eta_b \times \hat{z}) = \rho^{-1} \hat{z} \cdot [\nabla p_b \times \nabla \eta_b] = \rho^{-1} \hat{z} \cdot [\nabla \times (p_b \nabla \eta_b)]. \quad (6.175)$$

Evidently, the contribution from baroclinicity next to the bottom arises from the bottom pressure torque due to bottom pressure isolines that are not parallel to bottom topography isolines. We have more to say concerning boundary pressure torques in Section 6.9.3 as they also affect vorticity of the depth integrated flow.

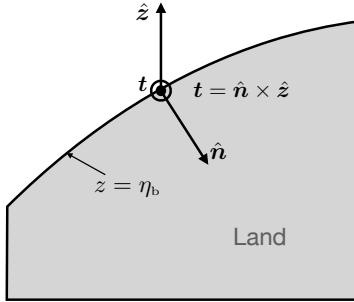


FIGURE 6.11: Orientation of the unit vectors next to the bottom of the fluid. The vertical unit vector,  $\hat{z}$ , points vertically upward and the outward unit vector,  $\hat{n}$ , points downward into the rock (shaded gray). The along-topography horizontal vector,  $\mathbf{t} = \hat{n} \times \hat{z}$ , points along lines of constant topography with land to the left when facing in the direction of  $\mathbf{t}$ ; in this figure it points out from the page. The vector  $\mathbf{t}$  is not necessarily a unit vector since it is not orthogonal to  $\hat{n}$ .

### Torques from bottom friction

The third term in the boundary flux (6.167b) is the contribution from friction along the bottom

$$-\hat{n} \cdot (\hat{z} \times \mathbf{F}) = -\mathbf{F} \cdot (\hat{n} \times \hat{z}) \equiv -\mathbf{F} \cdot \mathbf{t}. \quad (6.176)$$

Hence, contributions to the vertical vorticity evolution arise from the component of friction that projects onto the direction that parallels isobaths. To further our understanding of this result, consider a bottom friction written as a Rayleigh drag so that  $\mathbf{F} = -\gamma \mathbf{u}$  and

$$-\mathbf{F} \cdot \mathbf{t} = \gamma \mathbf{u} \cdot \mathbf{t}, \quad (6.177)$$

with  $\gamma$  an inverse time scale. If the flow is oriented with shallow water to the right; e.g., into the page in Figure 6.11, then  $-\mathbf{F} \cdot \mathbf{t} < 0$ , thus contributing a negative vorticity tendency. In general, the bottom friction acts to damp the depth integrated vorticity, which is expected since bottom friction does not spontaneously spin-up the flow.

### 6.8.4 Surface boundary contribution

The surface boundary contribution to the vorticity equation (6.166) takes on a similar form to the bottom, with the new feature that the free surface is both moving and permeable. This

boundary term is given by

$$\zeta_a \partial_t \eta - \nabla(z - \eta) \cdot \mathbf{J}^{\zeta_a} = \zeta_a \partial_t \eta - \nabla(z - \eta) \cdot [\mathbf{v} \zeta_a - w \omega_a - \hat{\mathbf{z}} \times \rho^{-1} \nabla p + \hat{\mathbf{z}} \times \mathbf{F}]_{z=\eta} \quad (6.178a)$$

$$= \zeta_a [\partial_t \eta + \mathbf{u} \cdot \nabla_h \eta - w]_{z=\eta} + \nabla(z - \eta) \cdot [w \omega_a + \hat{\mathbf{z}} \times \rho^{-1} \nabla p - \hat{\mathbf{z}} \times \mathbf{F}]_{z=\eta} \quad (6.178b)$$

$$= \zeta_a Q_m / \rho + |\nabla(z - \eta)| \hat{\mathbf{n}} \cdot [w \omega_a + \hat{\mathbf{z}} \times \rho^{-1} \nabla p - \hat{\mathbf{z}} \times \mathbf{F}]_{z=\eta}, \quad (6.178c)$$

where we made use of the [kinematic boundary condition](#) from [VOLUME 2](#) to introduce the surface mass flux  $Q_m$ , and where

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\nabla_h \eta + \hat{\mathbf{z}}}{\sqrt{1 + \nabla_h \eta \cdot \nabla_h \eta}} \quad (6.179)$$

is the outward unit normal at the surface. The first term in the surface boundary flux (6.178c) provides transport of boundary vorticity due to the transfer of mass across the boundary. The second term provides an inviscid vertical transport of absolute vorticity at the surface boundary, thus acting as a vortex stretching contribution. The third term provides a torque due to misalignments between the applied pressure isobars and the free surface isolines

$$|\nabla(z - \eta)| \hat{\mathbf{n}} \cdot (\hat{\mathbf{z}} \times \rho^{-1} \nabla p_a) = \rho^{-1} \nabla p_a \cdot (\nabla \eta \times \hat{\mathbf{z}}) = \rho^{-1} \hat{\mathbf{z}} \cdot (\nabla p_a \times \nabla \eta) = \rho^{-1} \hat{\mathbf{z}} \cdot [\nabla \times (p_a \nabla \eta)], \quad (6.180)$$

where the density,  $\rho$ , is evaluated at the ocean surface. Note that we made use of the same kinematics as for the bottom, thus allowing us to write

$$\hat{\mathbf{n}} \cdot [\hat{\mathbf{z}} \times \rho^{-1} \nabla p]_{z=\eta} = \hat{\mathbf{n}} \cdot (\hat{\mathbf{z}} \times \rho^{-1} \nabla p_a), \quad (6.181)$$

which is directly analogous to the bottom pressure equation (6.173). The fourth term in equation (6.178c) provides the corresponding contribution from the friction along the upper surface, with friction acting to reduce the magnitude of the surface boundary vorticity. For a rigid lid surface,  $w(0) = 0$ ,  $\eta = 0$ , and  $Q_m = 0$  so that the only surface boundary contribution arises from friction.

### 6.8.5 Comments

The depth integrated vorticity budget as derived in this section is perhaps the most physically straightforward of the suite of depth integrated vorticity budgets. However, in the practice of ocean modeling, this budget is generally not used since it requires an online coding of the vorticity equation and then its depth integral. As ocean models generally time step the velocity rather than the vorticity, it is common to form a vorticity budget based on the depth integrated flow or the depth averaged flow. We develop these budgets in the following section.

## 6.9 Vorticity for depth integrated hydrostatic flow

In this section we develop dynamical equations for vorticity of the depth integrated flow in a hydrostatic primitive equation fluid. A compelling application of these ideas comes from the study of large-scale ocean circulation. The leading order impacts from bottom pressure torques has emerged from research during recent decades, thus pointing to the fundamental role of bottom topography and flows next to sloping bottom (rather than vertical sidewalls) in affecting the ocean circulation. This recognition contrasts to traditional theories whereby the wind stress curl balances meridional motion through the beta effect. In particular, numerical

model studies reveal that wind stress curl is sub-dominant in any region with nontrivial bottom velocities.<sup>17</sup> In this section we introduce the basics and provide more discussion in Sections 10.6 and 10.7 when studying vorticity dynamics for the planetary geostrophic system.

### 6.9.1 Comparing the two vorticities

In Section 6.8 we derived the evolution equation for the depth integral of the vertical component to the absolute vorticity,

$$\int_{\eta_b}^{\eta} \zeta_a dz = \int_{\eta_b}^{\eta} (f + \hat{z} \cdot \nabla \times \mathbf{u}) dz. \quad (6.182)$$

In this section we study the evolution equation for the relative vorticity in the depth integrated horizontal flow

$$\hat{z} \cdot \nabla \times \mathbf{U}^\rho = \hat{z} \cdot \nabla \times \int_{\eta_b}^{\eta} \mathbf{u} \rho dz, \quad (6.183)$$

where we introduced the depth integrated horizontal mass flux

$$\mathbf{U}^\rho = \int_{\eta_b}^{\eta} \rho \mathbf{u} dz. \quad (6.184)$$

For a Boussinesq ocean we set the density to a constant, in which case the difference between the two relative vorticities is<sup>18</sup>

$$\hat{z} \cdot \nabla \times \left[ \int_{\eta_b}^{\eta} \mathbf{u} dz \right] - \int_{\eta_b}^{\eta} \hat{z} \cdot \nabla \times \mathbf{u} dz = \hat{z} \cdot [\nabla \eta \times \mathbf{u}(\eta) - \nabla \eta_b \times \mathbf{u}(\eta_b)]. \quad (6.185)$$

Flows along boundaries generally have a nontrivial projection in the direction parallel to boundary isosurfaces, in which case the cross products are nonzero thus leading to differences in the two relative vorticities.

### 6.9.2 Evolution of vorticity for the depth integrated horizontal flow

In VOLUME 2, we derived the following depth integrated horizontal momentum equation for a hydrostatic fluid

$$(\partial_t + f \hat{z} \times) \mathbf{U}^\rho = \mathbf{u}(\eta) Q_m - \eta \nabla_h p_a + \eta_b \nabla_h p_b - \nabla_h \mathcal{P} + \mathbf{D} + \nabla_h \cdot \left[ \int_{\eta_b}^{\eta} \mathbb{T}_{\text{hor}}^{\text{kinetic}} dz \right]. \quad (6.186)$$

We here introduced the potential energy per horizontal area of the fluid column, the depth integrated horizontal friction, and the divergence of the horizontal kinetic stress tensor

$$\mathcal{P} = \int_{\eta_b}^{\eta} g \rho z dz \quad \text{and} \quad \mathbf{D} = \int_{\eta_b}^{\eta} \rho \mathbf{F}_{\text{horz}} dz \quad \text{and} \quad \mathbb{T}_{\text{hor}}^{\text{kinetic}} = -\rho \mathbf{u} \otimes \mathbf{u}. \quad (6.187)$$

<sup>17</sup>The natural ocean has no distinction between side and bottom. Rather, as discussed in Figure 2.7, the ocean has a sloping bottom that reaches to the surface along its boundary at the “beach.” [Hallberg and Rhines \(1996\)](#), [Hughes and de Cuevas \(2001\)](#), and many subsequent studies emphasize that theoretical and numerical models using vertical sides and a flat bottom exhibit somewhat unnatural dynamical balances, whereas models with sloping bottoms better capture effects from topography consistent with the theory presented in this chapter.

<sup>18</sup>As noted in footnote 16, we can reduce notational clutter by writing  $\nabla$  rather than  $\nabla_h$  when operating on functions that are independent of  $z$ , such as  $\eta$ ,  $p_a$ ,  $\eta_b$  and  $p_b$ . Since these fields are independent of  $z$ , then  $\nabla_h \eta = \nabla \eta$ , and likewise for  $p_a$ ,  $\eta_b$  and  $p_b$ . We sometimes make use of this notation, though write  $\nabla_h$  where it can help to reduce ambiguity.

All terms on the right hand side of equation (6.186) provide a stress that drives changes in  $\mathbf{U}^\rho$ . Hence, the curl of these terms provides a torque, in which case we have

$$\hat{\mathbf{z}} \cdot \partial_t (\nabla \times \mathbf{U}^\rho) = -\nabla \cdot (f \mathbf{U}^\rho) + \hat{\mathbf{z}} \cdot \nabla \times \left[ \mathbf{u}(\eta) Q_m - \eta \nabla_h p_a + \eta_b \nabla_h p_b + \Delta \boldsymbol{\tau} + \mathbf{D} + \nabla_h \cdot \left( \int_{\eta_b}^{\eta} \mathbb{T}_{\text{hor}}^{\text{kinetic}} dz \right) \right], \quad (6.188)$$

where we used the vector identity

$$\hat{\mathbf{z}} \cdot \nabla \times (f \hat{\mathbf{z}} \times \mathbf{U}^\rho) = \nabla \cdot (f \mathbf{U}^\rho), \quad (6.189)$$

as well as  $\nabla \times \nabla_h \mathcal{P} = 0$ . We now discuss the various physical processes appearing in the vorticity equation (6.188).

### Beta effect

The first term on the right hand side of the vorticity equation (6.188) arises from the convergence of mass within a fluid column due to depth integrated horizontal flow. We can further decompose the effects from this term by performing the product rule

$$-\nabla \cdot (f \mathbf{U}^\rho) = -f \nabla \cdot \mathbf{U}^\rho - \beta V^\rho. \quad (6.190)$$

The contribution from  $\beta V^\rho$  arises from the beta effect as discussed in Section 6.6.2. For the first term, the weighting by the Coriolis parameter means that mass convergence at higher latitudes has more impact on vorticity changes than at lower latitudes. We can understand this weighting by noting that vertical fluid columns are more aligned with the planetary rotation at the high latitudes. Hence, when the mass of vertical columns converges at the higher latitudes, there is more impact on changes to the vorticity of the depth integrated flow.

### Mass transfer, turbulent momentum transfer, and nonlinear effects

The term  $\nabla \times [\mathbf{u}(\eta) Q_m]$  appearing in equation (6.191) accounts for vorticity crossing the ocean surface as affected by the mass flux. The term  $\nabla \times \Delta \boldsymbol{\tau}$  is the torque from turbulent stresses at the ocean surface and bottom, and  $\hat{\mathbf{z}} \cdot \nabla \times \mathbf{D}$  is the torque from horizontal frictional stresses in the fluid interior. The final term arises from the nonlinear kinetic stresses,  $\mathbb{T}_{\text{hor}}^{\text{kinetic}}$ , that account for curls in the self-advection operator.

#### 6.9.3 Boundary pressure torques

The pressure terms in equation (6.188)

$$\hat{\mathbf{z}} \cdot \nabla \times (-\eta \nabla p_a + \eta_b \nabla p_b) = \hat{\mathbf{z}} \cdot \nabla \times (p_a \nabla \eta - p_b \nabla \eta_b) \quad (6.191)$$

arise from curls of the pressure form stress (VOLUME 2) at the ocean surface and bottom, and these contributions are referred to as pressure torques. When the fluid is a column of ocean water, then the surface pressure contribution is the *atmospheric pressure torque* and the bottom pressure term is the *bottom pressure torque*.

### Geometry of boundary pressure torques

Geometrically, there is a nonzero atmospheric pressure torque when the applied pressure,  $p_a$ , has a gradient when moving along contours of constant free surface. Likewise, there is a nonzero bottom pressure torque when bottom pressure,  $p_b$ , changes along contours of constant bottom topography. Mathematically, we reveal these properties through use of Exercise 6.14. For example, the bottom pressure torque along an isobath (contour of constant  $\eta_b$ ) can be written

$$\hat{z} \cdot \nabla \eta_b \times \nabla p_b = -(\hat{n} \cdot \nabla \eta_b)(\hat{t} \cdot \nabla p_b), \quad (6.192)$$

where  $\hat{t}$  is a unit tangent vector directed along the isobath, and  $\hat{n}$  is a unit vector that points to the left of  $\hat{t}$  (see Figure 6.12). Both  $\hat{t}$  and  $\hat{n}$  are horizontal vectors.<sup>19</sup> Hence,  $\hat{n} \cdot \nabla \eta_b$  measures the slope of the bottom topography in the direction normal to an isobath, and  $\hat{t} \cdot \nabla p_b$  measures the change of the bottom pressure along the isobath. There is a nonzero bottom pressure torque along an isobath so long as there is a slope to the bottom pressure along the isobath, and there is a change in bottom pressure moving along the isobath.

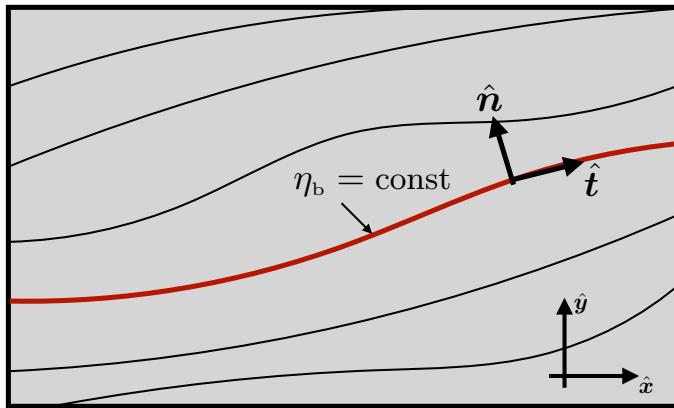


FIGURE 6.12: Geometry depicting a contour along a particular line of constant topography (i.e., an isobath),  $\eta_b(x, y)$ . The along-contour direction is  $\hat{t} = d\mathbf{x}/ds$ , with  $s$  the arc length along the contour. The unit direction pointing to the left of  $\hat{t}$  is written  $\hat{n}$ , with  $\hat{n} \cdot \hat{t} = 0$  and  $\hat{t} \times \hat{n} = \hat{z}$ . Both  $\hat{n}$  and  $\hat{t}$  are horizontal unit vectors. There is a nonzero bottom pressure torque if bottom pressure changes when following an isobath.

### Geostrophic velocity associated with the bottom pressure torque

To further our understanding of the pressure torques in equation (6.191), focus on the bottom pressure and introduce a geostrophic velocity<sup>20</sup>

$$\rho_0 f \mathbf{u}_g = \hat{z} \times (\nabla_h p)_{z=\eta_b} = \hat{z} \times [\nabla p_b + g \rho(\eta_b) \nabla \eta_b], \quad (6.193)$$

where the second equality made use of a result from our study of hydrostatic pressure in VOLUME 2, in which we express the horizontal pressure gradient at the bottom,  $(\nabla_h p)_{z=\eta_b}$ , in

<sup>19</sup>It is important to note that  $\hat{n}$  is not the outward normal direction to the bottom, contrary to its usage in Section 6.8. Here,  $\hat{n}$  it is the horizontal direction within the bottom surface that is normal to contours of constant topography.

<sup>20</sup>The geostrophic velocity is a balance between the Coriolis acceleration and the *horizontal* pressure gradient. We thus need to decompose the horizontal pressure gradient into the bottom pressure gradient and the gradient of the bottom slope, as per our discussion in VOLUME 2 of hydrostatic pressure gradients.

terms of the gradient of bottom pressure and gradient of bottom topography. Hence,

$$\nabla p_b = -\rho_0 f \hat{z} \times \mathbf{u}_g - g \rho(\eta_b) \nabla \eta_b, \quad (6.194)$$

so that the bottom pressure torque takes the form

$$\hat{z} \cdot \nabla \times (\eta_b \nabla p_b) = \hat{z} \cdot (\nabla \eta_b \times \nabla p_b) = -\rho_0 f \mathbf{u}_g \cdot \nabla \eta_b. \quad (6.195)$$

This equation is merely a replacement of the bottom pressure gradient with a corresponding geostrophic velocity. However, if this geostrophic velocity is assumed to satisfy the bottom no normal flow [kinematic boundary condition](#), then we reach the equality<sup>21</sup>

$$\hat{z} \cdot \nabla \eta_b \times \nabla p_b = -\rho_0 f w_g, \quad (6.196)$$

which links the bottom pressure torque to vertical vortex stretching by the vertical component to the geostrophic velocity. Evidently, if the bottom geostrophic velocity is negative ( $w_g < 0$ ), then that induces vortex stretching and a corresponding positive tendency for vorticity of the depth integrated flow. The opposite holds with  $w_g > 0$ , in which vortex squashing induces a negative tendency for vorticity of the depth integrated flow.

The equality (6.196) is sometimes used to infer the bottom pressure torque by diagnosing the bottom vertical velocity,  $w(\eta_b)$  ([Spence et al., 2012](#)). Although this diagnostic is suitable for some studies, there are important caveats. Namely, the bottom vertical velocity is generally affected by bottom frictional effects and thus can have a nontrivial Ekman component.<sup>22</sup> Nonlinear effects can also be important especially when considering motions with sizable Rossby numbers. Neither the Ekman component nor nonlinear terms are directly related to the bottom pressure torque. We thus expect  $\rho_0 f w(\eta_b)$  to be distinct from  $\hat{z} \cdot \nabla \eta_b \times \nabla p_b$  in regions of sizable deep flows where bottom friction and/or nonlinear effects are of leading order importance.<sup>23</sup> The studies from [Gula et al. \(2015\)](#) and [LeCorre et al. \(2020\)](#) illustrate these points from numerical simulations of the subpolar North Atlantic circulation.

#### 6.9.4 Steady state vorticity budget

The steady state form of the vorticity budget (6.188) leads to the balance

$$\begin{aligned} \beta V^\rho = -f \nabla \cdot \mathbf{U}^\rho \\ + \hat{z} \cdot \nabla \times \left[ \mathbf{u}(\eta) Q_m - \eta \nabla_h p_a + \eta_b \nabla_h p_b + \Delta \boldsymbol{\tau} + \mathbf{D} + \nabla_h \cdot \left( \int_{\eta_b}^{\eta} \mathbb{T}_{\text{hor}}^{\text{kinetic}} dz \right) \right]. \end{aligned} \quad (6.197)$$

Writing the balance in this manner reveals how the beta affect affords a steady meridional mass transport as a balance with the variety of terms on the right hand side.

<sup>21</sup>The velocity,  $\mathbf{v}$ , satisfies the no normal flow kinematic boundary condition, in which  $\mathbf{v}(\eta_b) \cdot \hat{n} = 0$ . Decomposing the velocity into its geostrophic and ageostrophic components,  $\mathbf{v} = \mathbf{v}_g + \mathbf{v}_a$ , does not generally imply that  $\mathbf{v}_g$  and  $\mathbf{v}_a$  separately satisfy the kinematic boundary condition. Rather, we must make that assumption in order to reach the equality (6.196).

<sup>22</sup>Recall our discussion of [Ekman mechanics](#) in [VOLUME 2](#).

<sup>23</sup>In addition to the Ekman and nonlinear effects noted here, diagnosing  $w(\eta_b)$  in a numerical model can be fraught with difficulties related to the discrete grid stencil given that grids can be quite coarse in the deep ocean with many ocean model configurations.

### Specializing the budget to expose a variety of balanced flow regimes

Let us further specialize to the case appropriate for many studies of the large-scale circulation, whereby we make the following assumptions.

- Uniform mass atmosphere so that  $p_a$  is a constant.
- The frictional stresses from horizontal strains within the fluid interior,  $\mathbf{D}$ , can be neglected.
- Zero boundary mass transport so that  $Q_m = 0$  and, correspondingly, the steady depth integrated mass budget (VOLUME 2) means that  $\nabla \cdot \mathbf{U}^\rho = 0$  when  $Q_m = 0$ .

These simplifications bring the balance (6.197) to the form

$$\beta V^\rho = \hat{\mathbf{z}} \cdot \nabla \times \left[ \eta_b \nabla_h p_b + \Delta \boldsymbol{\tau} + \nabla_h \cdot \left( \int_{\eta_b}^{\eta} \mathbb{T}_{\text{hor}}^{\text{kinetic}} dz \right) \right] \quad (6.198a)$$

$$\text{MERIDIONAL TRANSPORT} = \text{BOTTOM PRESS TORQUE} + \text{BOUNDARY STRESS} + \text{NONLINEAR}. \quad (6.198b)$$

This steady balance reveals distinct flow regimes depending on which of the terms dominate, and as such it serves as a useful framework for analysis.

#### Topographic nonlinear balance

[Jackson et al. \(2006\)](#), [Patmore et al. \(2019\)](#), and [LeCorre et al. \(2020\)](#) emphasize the importance of the nonlinear term in equation (6.198b) when flows are especially strong. The bottom pressure torque contribution is particularly strong where flows are strong near the bottom. Such nonlinear flow regimes generally have variations over length scales much smaller than that of the wind stress. Hence, if the horizontal friction is small, as it is even for strong flows not directly adjacent to solid boundaries, and the bottom frictional drag is small, then the vorticity balance (6.198a) in the nonlinear inviscid regime takes on the form

$$\beta V^\rho = \hat{\mathbf{z}} \cdot \nabla \times \left[ \eta_b \nabla_h p_b + \nabla_h \cdot \left( \int_{\eta_b}^{\eta} \mathbb{T}_{\text{hor}}^{\text{kinetic}} dz \right) \right] \quad \text{topographic nonlinear balance.} \quad (6.199)$$

Observe that the nonlinear term and bottom pressure torque have derivatives whereas there are none on the  $\beta V^\rho$  term. These derivatives make the right hand side terms have variations at smaller scales than  $\beta V^\rho$ . We infer that the smaller scales present in the bottom pressure torque and the nonlinear term nearly balance, and with any residual leading to the broader scale meridional transport. Figure 6 in [LeCorre et al. \(2020\)](#) provides a striking example of this balance in a numerical simulation of the North Atlantic subpolar gyre.

#### Linear regime of planetary geostrophy

For the linear regime of planetary geostrophic flow (Chapter 10), the nonlinear term from the kinetic stress is small, so that the balance is between meridional transport, bottom pressure torque, and curl of turbulent boundary stresses. The [Sverdrup balance](#) is one particular example of a planetary geostrophic balance, with Sverdrup balance ignoring the bottom pressure torque and bottom turbulent stresses, and thus focuses just on the balance between meridional transport with the turbulent surface stresses largely arising from winds

$$\beta V^\rho = \hat{\mathbf{z}} \cdot \nabla \times \boldsymbol{\tau}^\eta \quad \text{Sverdrup balance.} \quad (6.200)$$

However, as emphasized by [Hallberg and Rhines \(1996\)](#) and [Hughes and de Cueves \(2001\)](#), as well as more recent studies, contributions from bottom pressure torques are of leading order importance in the presence of flow next to sloping side boundaries, thus making the traditional Sverdrup balance mostly relevant in the open ocean away from boundaries. A more general balance is known as [topographic topographic Sverdrup balance](#)

$$\beta V^\rho = \hat{z} \cdot \nabla \times (\eta_b \nabla_h p_b + \boldsymbol{\tau}^\eta) \quad \text{topographic Sverdrup balance.} \quad (6.201)$$

We further study these balances of planetary geostrophy in Sections 10.4, 10.6, and 10.7.

### 6.9.5 Integral balances satisfied by steady flows

Reconsider the steady vorticity balance (6.197), here written in the form

$$\nabla \cdot (f \mathbf{U}^\rho) = \hat{z} \cdot \nabla \times \mathbf{M} \quad (6.202)$$

where we introduced the stress vector

$$\mathbf{M} = \mathbf{u}(\eta) Q_m - \eta \nabla_h p_a + \eta_b \nabla_h p_b + \Delta \boldsymbol{\tau} + \mathbf{D} + \nabla_h \cdot \left( \int_{\eta_b}^{\eta} T_{\text{hor}}^{\text{kinetic}} dz \right). \quad (6.203)$$

Now integrate equation (6.202) over an area,  $\mathcal{S}$ , with Gauss's divergence theorem on the left hand side leading to

$$\int_{\mathcal{S}} \nabla \cdot (f \mathbf{U}^\rho) d\mathcal{S} = \oint_{\partial\mathcal{S}} f \mathbf{U}^\rho \cdot \hat{\mathbf{n}} ds, \quad (6.204)$$

where  $\hat{\mathbf{n}}$  is the horizontal outward unit normal on the boundary,  $\partial\mathcal{S}$ , and  $ds$  is the arc-length increment along the boundary. This term is the mass transport crossing the boundary as weighted by the Coriolis parameter.

Use of Stokes' curl theorem on the right hand side of equation (6.202) leads to

$$\int_{\mathcal{S}} \hat{z} \cdot \nabla \times \mathbf{M} d\mathcal{S} = \oint_{\partial\mathcal{S}} \mathbf{M} \cdot \hat{\mathbf{t}} ds, \quad (6.205)$$

where  $\hat{\mathbf{t}}$  is the horizontal unit tangent vector along the boundary, and the integral is oriented in the counter-clockwise direction. To help interpret the closed loop integral in equation (6.205), consider just the contribution from bottom pressure

$$\oint_{\partial\mathcal{S}} \mathbf{M}_{\text{bottom press}} \cdot \hat{\mathbf{t}} ds = \oint_{\partial\mathcal{S}} \eta_b \nabla p_b \cdot \hat{\mathbf{t}} ds = - \oint_{\partial\mathcal{S}} p_b \nabla \eta_b \cdot \hat{\mathbf{t}} ds, \quad (6.206)$$

which is the work done by bottom topographic form stress around the closed contour. The other terms in equation (6.202) have interpretations as the work arising from integrating stresses from mass transport through the surface, atmospheric form stress, turbulent boundary stresses, interior frictional stresses, and nonlinear kinetic stress. Observe that the integral of bottom pressure torque in equation (6.206) vanishes if the closed contour follows either an isobath or a bottom pressure isobar. The vanishing of this integral means that bottom pressure torques have zero net circulation around isobaths or bottom isobars. An analogous property is satisfied by the atmospheric pressure torque when integrated around closed contours of constant atmospheric pressure,  $p_a$ , or constant surface height,  $\eta$ .

Bringing the above results together renders the general balance around the boundary of an

arbitrary closed region

$$\oint_{\partial S} f \mathbf{U}^\rho \cdot \hat{\mathbf{n}} ds = \oint_{\partial S} \mathbf{M} \cdot \hat{\mathbf{t}} ds. \quad (6.207)$$

We thus see that transport across the closed boundary, as weighted by the Coriolis parameter, arises from a nonzero net work around the boundary by the variety of stresses comprising  $\mathbf{M}$ . We are afforded a key simplification if  $\mathbf{U}^\rho \cdot \hat{\mathbf{n}} = 0$  at each point along the boundary. For example, if  $\nabla \cdot \mathbf{U}^\rho = 0$ , which generally also requires  $Q_m = 0$ , then contours along which  $\mathbf{U}^\rho \cdot \hat{\mathbf{n}} = 0$  correspond to closed streamlines of the steady  $\mathbf{U}^\rho$ . Hence, we find the following balance holds around any closed streamline

$$\oint_{\partial S} \mathbf{M} \cdot \hat{\mathbf{t}} ds = 0. \quad \text{for } \partial S \text{ a closed streamline of } \mathbf{U}^\rho. \quad (6.208)$$

Closed streamlines do not always exist. But when they do, such as for steady ocean gyre circulations, this balance holds. In Section 10.6.3, we consider the planetary geostrophic version of this balance.

### 6.9.6 Formulation based on the vector-invariant velocity equation

In formulating the budget equation (6.188) for vorticity of the depth integrated flow, we started with the depth integrated momentum in from VOLUME 2. However, many numerical models are formulated using the vector invariant form of the horizontal velocity equation (6.147), here written in the equivalent form for a Boussinesq ocean

$$\underbrace{\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} + \nabla_h p / \rho_0 - \mathbf{F}}_{\text{linear terms plus friction}} = \underbrace{-\zeta \hat{\mathbf{z}} \times \mathbf{u} - w \partial_z \mathbf{u} - \nabla_h \mathbf{u}^2 / 2}_{\text{Magnus + vertical advection + kinetic energy}}. \quad (6.209)$$

The three nonlinear terms on the right hand side arise from expanding the nonlinear self-advection term,  $(\mathbf{v} \cdot \nabla) \mathbf{u}$ , following the manipulations in Section 6.7.1. Much of the formulation to follow emulates that considered thus far, with the exception of the nonlinear terms and elements of the boundary contributions.

We take the vertical integral of equation (6.209)

$$\int_{\eta_b}^{\eta} [\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} + \nabla_h p / \rho_0 - \mathbf{F}] dz = - \int_{\eta_b}^{\eta} (\zeta \hat{\mathbf{z}} \times \mathbf{u} + w \partial_z \mathbf{u} + \nabla_h \mathbf{u}^2 / 2) dz, \quad (6.210)$$

and then the curl

$$\nabla \times \int_{\eta_b}^{\eta} (\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} + \nabla_h p / \rho_0 - \mathbf{F}) dz = -\nabla \times \int_{\eta_b}^{\eta} (\zeta \hat{\mathbf{z}} \times \mathbf{u} + w \partial_z \mathbf{u} + \nabla_h \mathbf{u}^2 / 2) dz. \quad (6.211)$$

Making use of the following identities

$$\hat{\mathbf{z}} \cdot \left[ \nabla \times \int_{\eta_b}^{\eta} \partial_t \mathbf{u} dz \right] = \hat{\mathbf{z}} \cdot \partial_t (\nabla \times \mathbf{U}) - \hat{\mathbf{z}} \cdot \nabla \times [\mathbf{u}(\eta) \partial_t \eta] \quad (6.212a)$$

$$\hat{\mathbf{z}} \cdot \left[ \nabla \times \int_{\eta_b}^{\eta} f \hat{\mathbf{z}} \times \mathbf{u} dz \right] = \nabla \cdot (f \mathbf{U}) \quad (6.212b)$$

$$\hat{\mathbf{z}} \cdot \left[ \nabla \times \int_{\eta_b}^{\eta} \nabla_h p dz \right] = \hat{\mathbf{z}} \cdot \nabla \times (\eta \nabla p_a - \eta_b \nabla p_b) \quad (6.212c)$$

$$\hat{z} \cdot \left[ \nabla \times \int_{\eta_b}^{\eta} \zeta \hat{z} \times \mathbf{u} dz \right] = \nabla \cdot \left[ \int_{\eta_b}^{\eta} \zeta \mathbf{u} dz \right] \quad (6.212d)$$

$$\hat{z} \cdot \left[ \nabla \times \int_{\eta_b}^{\eta} \rho_o \mathbf{F} dz \right] = \hat{z} \cdot \nabla \times (\Delta \boldsymbol{\tau} + \mathbf{D}), \quad (6.212e)$$

leads to

$$\begin{aligned} \hat{z} \cdot \partial_t (\nabla \times \mathbf{U}) &= -\nabla \cdot \left[ f \mathbf{U} + \int_{\eta_b}^{\eta} \zeta \mathbf{u} dz \right] \\ &+ \hat{z} \cdot \nabla \times \left[ \mathbf{u}(\eta) \partial_t \eta - \eta \nabla p_a + \eta_b \nabla p_b + (\Delta \boldsymbol{\tau} + \mathbf{D})/\rho_o - \int_{\eta_b}^{\eta} (w \partial_z \mathbf{u} + \nabla_h \mathbf{u}^2/2) dz \right]. \end{aligned} \quad (6.213)$$

The left hand side is the time tendency of the vorticity of the depth integrated horizontal flow, with this time tendency driven by the various linear and nonlinear terms on the right hand side. This evolution equation should be compared to equation (6.188) as derived from the advective form of the momentum equation. Likewise, we derive a steady state balance by setting the time tendencies to zero to yield

$$\begin{aligned} \beta V &= -f \nabla \cdot \mathbf{U} - \nabla \cdot \left[ \int_{\eta_b}^{\eta} \zeta \mathbf{u} dz \right] \\ &+ \hat{z} \cdot \nabla \times \left[ \mathbf{u}(\eta) \partial_t \eta - \eta \nabla p_a + \eta_b \nabla p_b + (\Delta \boldsymbol{\tau} + \mathbf{D})/\rho_o - \int_{\eta_b}^{\eta} (w \partial_z \mathbf{u} + \nabla_h \mathbf{u}^2/2) dz \right], \end{aligned} \quad (6.214)$$

which should be compared to equation (6.197).

### 6.9.7 Vorticity of the depth averaged flow

The vorticity of the depth averaged flow is given by  $\hat{z} \cdot \nabla \times \bar{\mathbf{u}}$ , where  $\bar{\mathbf{u}}$  takes on the following form for a Boussinesq ocean

$$\bar{\mathbf{u}} = \frac{\int_{\eta_b}^{\eta} \mathbf{u} dz}{\eta - \eta_b} = \frac{\mathbf{U}}{\eta - \eta_b}. \quad (6.215)$$

The difference is given by

$$\nabla \times \mathbf{U} - (\eta - \eta_b) \nabla \times \bar{\mathbf{u}} = \nabla(\eta - \eta_b) \times \bar{\mathbf{u}}, \quad (6.216)$$

so that the two vorticities are the same in the special case of a depth averaged flow that is parallel to  $\nabla(\eta - \eta_b)$ . Quite trivially,  $\nabla(\eta - \eta_b) \times \bar{\mathbf{u}} = 0$  occurs for a rigid lid and flat bottom ocean, in which  $\nabla \eta = \nabla \eta_b = 0$ . More generally,  $\nabla(\eta - \eta_b) \times \bar{\mathbf{u}} \neq 0$ , particularly in the presence of topography. We further study the budgets for these two vorticities, for planetary geostrophic flow, in Sections 10.6 and 10.7.

### 6.9.8 Comments and further study

The diagnostic budgets derived in this section have appeared in many studies of ocean vorticity. When diagnosing the budget terms in a numerical model, the choice for how to mathematically formulate the diagnostic balances is largely driven by physical transparency as well as by numerical precision. Concerning numerical precision, it is useful to note that vorticity, as the derivative of velocity, has more power at the high spatial wave numbers than does velocity. In a numerical model, such power can manifest as grid scale noise. It is thus of use to perform much

of the calculation online to enable the most accurate available diagnostic. Even so, further spatial smoothing is generally required, especially in realistic models, to extract physically interpretable signals.



## 6.10 Exercises

### EXERCISE 6.1: FILLING IN DETAILS TO A DERIVATION

Fill in the mathematical details to prove the identity (6.44a)

$$\rho \frac{D(\zeta_a/\rho)}{Dt} = \frac{\partial \zeta_a}{\partial t} + \nabla \cdot (\mathbf{v} \zeta_a) \quad (6.217)$$

Hint: make use of mass continuity in the form of

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}, \quad (6.218)$$

### EXERCISE 6.2: RELATING THE INTEGRAL OF DIVERGENCE AND VORTICITY

For some purposes, it is useful to consider evolution of the flow divergence as well as the vorticity. In this chapter we focused on the vorticity. Here are to prove, using Cartesian tensors, that

$$\int_{\mathcal{R}} (\mathbf{v} \nabla \cdot \mathbf{v} + \boldsymbol{\omega} \times \mathbf{v}) dV = \oint_{\partial\mathcal{R}} [\mathbf{v}(\mathbf{v} \cdot \hat{\mathbf{n}}) - \hat{\mathbf{n}} \mathcal{K}] d\mathcal{S}, \quad (6.219)$$

where  $\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2$  is the kinetic energy per mass. Hint: make use of the vector identity (6.32) as well as the scalar form of Gauss's divergence theorem from VOLUME 1, as given by

$$\oint_{\partial\mathcal{R}} \phi \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{R}} \nabla \phi dV. \quad (6.220)$$

### EXERCISE 6.3: STRAIN AND ROTATION FOR STRETCHING AND TILTING

In this exercise we write the  $3 \times 3$  strain rate tensor,  $\mathbf{S}$ , and rotation tensor,  $\mathbf{R}$ , for the examples of vortex stretching and vortex tilting considered in Section 6.5.3. Hint: there is no unique answer for the strain rate tensors, so offer a simple example that renders the desired behavior of a vortex line.

- (a) Write a strain rate tensor corresponding to vortex stretching as per equation (6.81) along with  $\omega^x = \omega^y = 0$ , and write the corresponding vorticity source term  $\boldsymbol{\omega} \cdot \mathbf{S}$ .
- (b) Write the rotation tensor for vortex stretching as per equation (6.81) and verify that  $\omega_m \cdot \mathbb{R}_{mn} = 0$ .
- (c) Write a strain rate tensor corresponding to vortex tilting as per equation (6.87) along with  $\omega^y = \omega^z = 0$ .
- (d) Write the rotation tensor for vortex tilting as per equation (6.87) and verify that  $\omega_m \mathbb{R}_{mn} = 0$ .

### EXERCISE 6.4: FRICTION IN THE VORTICITY EQUATION

Assume a viscous friction operator of the form

$$\mathbf{F} = \nu \nabla^2 \mathbf{v}, \quad (6.221)$$

with  $\nu$  a constant molecular kinematic viscosity. Assuming Cartesian coordinates, write the vorticity equation (6.42) with this term included.

**EXERCISE 6.5: FRICTION FOR NON-DIVERGENT FLOWS**

Consider a non-divergent flow with a Laplacian frictional acceleration

$$\mathbf{F} = \nu \nabla^2 \mathbf{v} \quad \text{with} \quad \nabla \cdot \mathbf{v} = 0, \quad (6.222)$$

with  $\nu$  a constant molecular kinematic viscosity. Write this expression in terms of the vorticity.  
Hint: check that  $\nabla \times \mathbf{F}$  equals to the friction appearing in the vorticity equation derived in exercise 6.4. Further hint: the derivation is given in VOLUME 2.

**EXERCISE 6.6: VORTICITY FOR STEADY NON-DIVERGENT  $y$ - $z$  CIRCULATION**

This exercise is based on exercise (1) in Section 1.1 of [Pratt and Whitehead \(2008\)](#). Consider inviscid, constant density, and non-divergent flow in the  $y$ - $z$  (meridional-vertical) plane and in a non-rotating reference frame

$$\rho (\partial_t + v \partial_y + w \partial_z) v = -\partial_y p \quad (6.223a)$$

$$\rho (\partial_t + v \partial_y + w \partial_z) w = -\partial_z p - \rho g \quad (6.223b)$$

$$\partial_y v + \partial_z w = 0. \quad (6.223c)$$

- (a) Show that the zonal component of the relative vorticity is materially constant following the  $y$ - $z$  flow

$$(\partial_t + v \partial_y + w \partial_z) \omega^x = 0 \quad \text{with} \quad \omega^x = \hat{\mathbf{x}} \cdot (\nabla \times \mathbf{v}) = \partial_y w - \partial_z v. \quad (6.224)$$

- (b) Introduce the  $y$ - $z$  (meridional-vertical) overturning streamfunction

$$(v, w) = \hat{\mathbf{x}} \times \nabla \psi = -\hat{\mathbf{y}} \partial_z \psi + \hat{\mathbf{z}} \partial_y \psi, \quad (6.225)$$

so that the vorticity is the Laplacian of the streamfunction

$$\omega^x = (\partial_{yy} + \partial_{zz}) \psi. \quad (6.226)$$

Write the vorticity equation (6.224) in terms of the streamfunction. Check your answer by showing that the steady vorticity equation implies that the Jacobian of the streamfunction with the vorticity vanishes

$$J(\psi, \omega^x) = \partial_y \psi \partial_z \omega^x - \partial_z \psi \partial_y \omega^x = 0. \quad (6.227)$$

- (c) Following from the previous part, show that for steady flow that the vorticity is a function just of the streamfunction,

$$\omega^x = F(\psi), \quad (6.228)$$

where the function,  $F$ , is determined by the value of the vorticity along the streamlines.  
Hint: we already know that  $\omega^x = (\partial_{yy} + \partial_{zz}) \psi$ , even for time dependent flow, which follows from the non-divergent nature of the  $y$ - $z$  overturning circulation. What equation (6.228) says is that for steady flow, the vorticity is a function just of the streamfunction. Consequently, if we specify the vorticity at any point along a streamline, then we know the vorticity everywhere along the streamline since it remains constant. Furthermore, it

means that the streamfunction satisfies the elliptic problem

$$(\partial_{yy} + \partial_{zz}) \psi = F(\psi). \quad (6.229)$$

**EXERCISE 6.7: BAROCLINICITY WITH  $\rho_o(z)$**

Recall the discussion of the Boussinesq ocean's momentum equation in VOLUME 2. We started the derivation by assuming a depth-dependent reference density,  $\rho_o = \rho_o(z)$ , and then stated that the form of the baroclinicity vector appearing in the Boussinesq vorticity equation is greatly simplified by setting  $\rho_o$  to a global constant, and thus dropping the  $z$  dependence. Derive the following baroclinicity vector

$$\mathbf{B} = \nabla \left[ b - \frac{\delta p}{\rho_o^2} \frac{d\rho_o}{dz} \right] \times \hat{z}, \quad (6.230)$$

so that  $\mathbf{B} = \nabla b \times \hat{z}$  when  $\rho_o$  is assumed to be a global constant. Hint: write the vector-invariant form of the Boussinesq momentum equation with  $\rho_o(z)$ , and then take the curl.

**EXERCISE 6.8: GENERATION OF VORTICITY BY BAROCLINICITY**

Consider a body of water with a flat bottom and rigid sides. Let the top surface be at  $z = 0$  and bottom at  $z = -H$ , and assume zero pressure applied at the top surface. Let the density have a horizontal structure given by

$$\rho(x) = \rho_o (1 - \gamma x) \quad (6.231)$$

where  $\rho_o$  and  $\gamma$  are positive constants (with dimensions of density and inverse length, respectively). We furthermore assume that  $\gamma|x| \ll 1$  so that the density is strictly positive. Note that a study of Figure 6.3 helps with this exercise.

As posed, the fluid is not in mechanical equilibrium since there is a horizontal density gradient. Hence, the fluid will adjust as a result of the nonzero horizontal pressure gradient force. Our aim here is to compute the baroclinicity contained in the fluid to garner a sense for the initial adjustment of vorticity.

- (a) Compute the density gradient  $\nabla \rho$  and draw a schematic.
- (b) Compute the pressure gradient,  $\nabla p$ , assuming approximate hydrostatic balance so that  $\partial p / \partial z = -\rho g$ . Draw a schematic at  $x = 0$ .
- (c) Compute the baroclinicity/solenoidal vector  $\mathbf{B} = \rho^{-2} (\nabla \rho \times \nabla p)$ . Draw a schematic.
- (d) Describe the vorticity induced by the baroclinicity vector.

**EXERCISE 6.9: GENERATION OF CIRCULATION BY BAROCLINICITY IN AN IDEAL GAS**

In this exercise we examine the baroclinicity vector for a simple ideal gas, which is described by the equation of state (VOLUME 2)

$$\rho = \frac{p M_{\text{mole}}}{T R^g} \equiv \frac{p}{T R^M}, \quad (6.232)$$

where  $R^g$  is the universal gas constant and  $R^M$  is the specific gas constant. We also assume the atmosphere is in approximate hydrostatic balance, and ignore rotation (relatively small lateral region of the atmosphere). For further hints to this exercise, see Section 4.1 of [Holton and Hakim \(2013\)](#) or Section 2.4.3 of [Markowski and Richardson \(2010\)](#), where they discuss circulation generated by differences in land-sea temperatures, thus leading to a sea breeze.

- (a) Express the baroclinicity vector,  $\mathbf{B}$ , in terms of pressure and temperature gradients.
- (b) Express the baroclinicity vector in terms of pressure and potential temperature gradients.

Hint: see VOLUME 2 for the potential temperature in an ideal gas.

- (c) Consider an ideal gas atmosphere straddling the ocean and flat land as in Figure 6.13. Let the daytime air be relatively cool over the ocean and relatively warm over the land. Furthermore, assume the sea level pressure is the same value over land and ocean. Ignoring rotation, draw isolines of constant temperature and constant pressure. Assume the horizontal temperature gradient is constant with height. Here are some hints.
- Temperature decreases from land to ocean and decreases when ascending into the atmosphere.
  - Pressure is assumed to be horizontally constant at sea level and it decreases upward. Use the ideal gas law to determine the sense for the horizontal pressure gradient as one ascends. Consult the discussion in VOLUME 2 for geopotentials in an ideal gas atmosphere.
  - We are only concerned with a qualitative sense for the isolines in the lower atmosphere and over a horizontal region small enough that rotation can be ignored.
- (d) Describe the sense for the circulation induced by the baroclinicity. Does circulation correspond to your experience at a sunny beach day as the air warms over the land faster than over the adjacent ocean? What force causes air to rise and to fall?

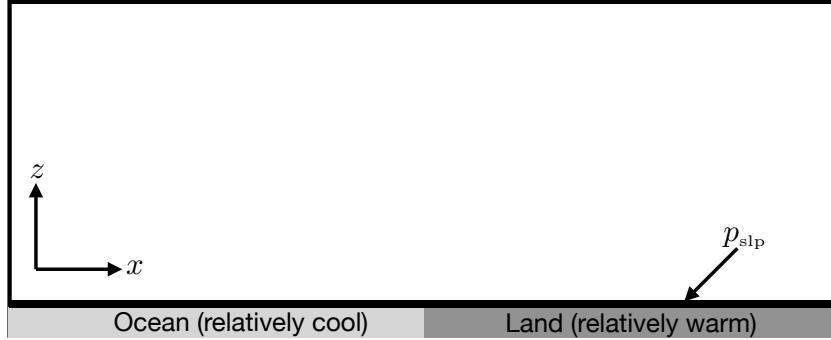


FIGURE 6.13: Setup for the sea breeze Exercise 6.9. We here depict a vertical-zonal crosssection of the atmosphere where the lower boundary straddles the ocean and land. The atmosphere over the ocean is assumed to be cooler than the atmosphere over the land, as typically occurs on a sunny afternoon with solar radiation warming land faster than the ocean.

#### EXERCISE 6.10: CIRCULATION WITH ISLANDS

Our discussion of Stokes' theorem has been thus far restricted to a simply connected domain, in which

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS. \quad (6.233)$$

For a simply connected domain, the closed contour can be shrunk to a point without leaving the domain.

A more general topology consists of a region with holes, whereby closed contours cannot in general be shrunk to a point without leaving the region. In an oceanographic context, the “holes” are islands or continents and the circulation is that for the depth integrated flow. Figure 6.14 shows a region of the ocean containing three arbitrarily shaped impenetrable islands, with the three islands surrounded by a contour. The contour cannot be shrunk to a point without crossing over the islands, thus making this region of the ocean multiply-connected. The presence of islands thus adds a level of complexity to the World Ocean that is absent an AquaPlanet or the global atmosphere.

Derive the following expression for the circulation in multiply-connected regions

$$\oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{x} = \sum_{n=1}^N \left( \oint_{\partial\mathcal{S}_n} \mathbf{v} \cdot d\mathbf{x} \right) + \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (6.234)$$

where  $N$  is the number of islands,  $\mathcal{S}_n$  is the contour surrounding each island, and  $\mathcal{S}$  is the region of water that excludes the islands. In words, this result says that the circulation around a region equals to the circulation around the islands within the region, plus the normal component of the vorticity integrated over the area within the fluid region. Removing the islands allows the island contours to be shrunk to zero size, in which case we recover the simply connected result (6.233). As part of your solution, make use of the contour integral method detailed in Figure 5.10.

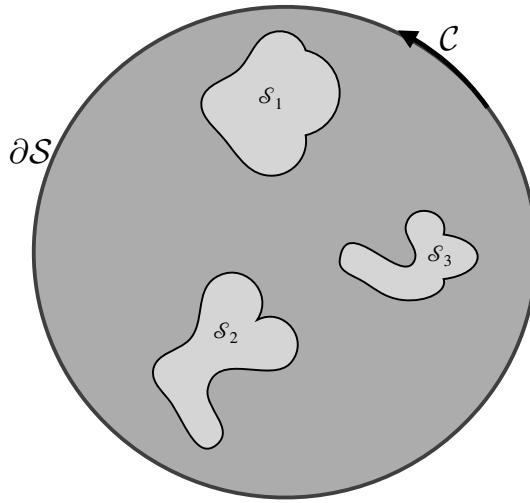


FIGURE 6.14: A region of the ocean consisting of three islands,  $\mathcal{S}_1$ ,  $\mathcal{S}_2$ , and  $\mathcal{S}_3$ , each with boundaries  $\partial\mathcal{S}_n$  and with the closed contour,  $\partial\mathcal{S}$ , drawn around the three islands. The contour  $\partial\mathcal{S}$  cannot be shrunk to a point without crossing over the islands, thus indicating that the domain is multiply connected. Exercise 6.10 is concerned with deriving an expression for the circulation of the depth-integrated flow as defined along the closed contour,  $\partial\mathcal{S}$ . In that derivation we can make use of the contour integral method detailed in Figure 5.10.

#### EXERCISE 6.11: EVOLUTION OF CIRCULATION AROUND ISLANDS

The momentum equation for a homogeneous layer of inviscid shallow water fluid on a tangent plane is given by

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta. \quad (6.235)$$

In this equation,  $\mathbf{u} = (u, v)$  is the horizontal velocity,  $f$  is the Coriolis parameter (need not be constant),  $g$  is the effective gravitational acceleration, and  $\eta$  is the deviation of the free surface from its horizontal resting position. All spatial derivatives are horizontal, so that

$$\mathbf{u} \cdot \nabla = u \partial_x + v \partial_y. \quad (6.236)$$

Use of a vector identity allows us to write

$$\partial_t \mathbf{u} + (f + \zeta) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla (\mathbf{u}^2/2 + g \eta), \quad (6.237)$$

where

$$\zeta = \hat{z} \cdot (\nabla \times \mathbf{u}) \quad (6.238)$$

is the vorticity of the shallow water fluid.

Consider an island, such as one shown in Figure 6.14. Each island is static and impenetrable to fluid flow, which means that

$$\mathbf{u} \cdot \hat{\mathbf{n}} = 0 \quad (6.239)$$

where  $\hat{\mathbf{n}}$  is the outward normal on an island boundary. For simplicity, assume this island outward normal is horizontal; i.e., the island is bounded by a vertical side. This no-normal flow constraint means that the velocity just next to an island is parallel to the island<sup>24</sup>

$$\mathbf{u} \times d\mathbf{x} = 0. \quad (6.240)$$

Equivalently, the island represents a solid material boundary across which no flow passes.

Show that the inviscid shallow-water circulation around an island remains constant in time

$$\frac{d}{dt} \oint_I \mathbf{u} \cdot d\mathbf{x} = 0. \quad (6.241)$$

Recall that Kelvin's circulation theorem is formulated for a material circuit in an inviscid fluid, with the circuit moving with the flow. This exercise shows that the circulation theorem also holds for a material circuit enclosing a static solid boundary.

#### EXERCISE 6.12: HELICITY FOR A PERFECT BAROTROPIC FLUID IN A GRAVITY FIELD AND NON-ROTATING REFERENCE FRAME

Consider a closed material volume,  $\mathcal{R}$ , of a perfect single-constituent barotropic fluid ( $\rho = \rho(p)$ ) in a gravity field ( $\mathbf{g} = -\nabla\Phi$ ) and in a non-rotating reference frame ( $\boldsymbol{\Omega} = 0$ ). Let this material volume have a boundary that is always tangent to the fluid vorticity,  $\boldsymbol{\omega}$ . Hence, the outward normal to the region boundary is orthogonal to the vorticity,

$$\hat{\mathbf{n}} \cdot \boldsymbol{\omega} = 0. \quad (6.242)$$

Such volumes define closed vortex tubes, such as a smoke ring or linked smoke rings. The helicity of the fluid within the vortex tube volume is defined as the integration of the helicity density,  $\mathbf{v} \cdot \boldsymbol{\omega}$ , over the closed volume

$$\mathbb{H} = \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \cdot \boldsymbol{\omega} dV, \quad (6.243)$$

where the volume  $\mathcal{R}(\mathbf{v})$  is material. In Cartesian coordinates, the helicity density takes the form

$$\mathbf{v} \cdot \boldsymbol{\omega} = u(\partial_y w - \partial_z v) + v(\partial_z u - \partial_x w) + w(\partial_x v - \partial_y u). \quad (6.244)$$

Although the helicity density vanishes for some common examples, such as for a fluid in rigid-body rotation, it need not vanish in general.

- (a) Show that the helicity density identically vanishes for a two-dimensional non-divergent flow.

---

<sup>24</sup>This boundary condition is valid only for inviscid fluids such as that considered here. For a real fluid with nonzero viscosity, all components of the velocity vector vanish at solid boundaries due to the no-slip boundary condition.

(b) Show that helicity is materially constant following the material volume

$$\frac{d\mathbb{H}}{dt} = 0. \quad (6.245)$$

(c) Discuss why helicity is not defined for a shallow water fluid.

Use the following hints.

- Make use of  $\Phi_p$  that satisfies equation (6.13).
- The shallow water fluid model is based on the small aspect ratio limit, in which the fluid depth is much smaller than its lateral extent. In this limit, the vertical component of vorticity dominates over the horizontal. See further discussion in Section 6.5.4.

#### EXERCISE 6.13: DISCRETE CALCULATION OF BOTTOM PRESSURE TORQUE

In many diagnostic studies with numerical models it is of interest to compute pressure torques affecting vorticity. One particularly common diagnostic concerns the bottom pressure torque arising in equation (6.191). Derive a discrete expression for the area averaged bottom pressure torque

$$\overline{\text{BPT}} = A^{-1} \int_S \hat{z} \cdot \nabla \times (\eta_b \nabla p_b) dS = A^{-1} \oint_{\partial S} \eta_b \nabla p_b \cdot \hat{t} d\ell = -A^{-1} \oint_{\partial S} p_b \nabla \eta_b \cdot \hat{t} d\ell, \quad (6.246)$$

over the shaded region depicted in Figure 6.15, with  $A = \int_S dS$  the horizontal area of this region. Hint: this exercise shares much with the area averaged vorticity in Exercise 3.8, although the final result is distinct. Note: given that the bottom pressure torque is generally the small difference between large numbers, it is very useful to perform the diagnostic calculation online so that full computational precision can be maintained.

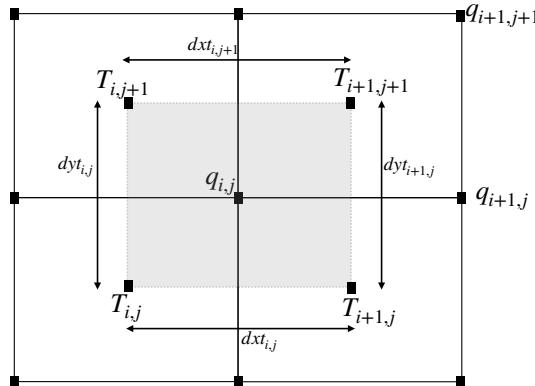


FIGURE 6.15: Discrete grid layout for variables needed to compute the bottom pressure torque as averaged over the shaded vorticity region. The bottom topography and bottom pressure are both known at the tracer points.

#### EXERCISE 6.14: DYNAMICAL PORTION OF THE TOPOGRAPHIC FORM STRESS

We discussed bottom topographic form stress in VOLUME 2, with its curl leading to the bottom pressure torque in equation (6.191). The dominant portion of the bottom topographic form stress acting on the ocean has little to do with fluid motion. Rather, it merely holds the ocean fluid within the basin, much as water is held within a drinking container through pressure imparted by the container sides.

- (a) To help isolate the dynamically relevant portion of the bottom bottom pressure, show that we can write the horizontal gradient of the bottom pressure for a hydrostatic fluid according to

$$\nabla_h p_b = g [\rho(\eta) \nabla_h \eta - \rho(\eta_b) \nabla_h \eta_b] + g \int_{\eta_b}^{\eta} \nabla_h \rho dz. \quad (6.247)$$

where we ignore the applied surface pressure,  $p_a$ , for simplicity.

- (b) Hence, show that the bottom pressure torque takes the form

$$\hat{z} \cdot \nabla_h \eta_b \times \nabla_h p_b = g \hat{z} \cdot \nabla_h \eta_b \times \left[ \rho(\eta) \nabla_h \eta + \int_{\eta_b}^{\eta} \nabla_h \rho dz \right] \equiv \hat{z} \cdot \nabla \eta_b \times \nabla p_b^*, \quad (6.248)$$

where

$$\nabla_h p_b^* = g \rho(\eta) \nabla_h \eta + g \int_{\eta_b}^{\eta} \nabla_h \rho dz. \quad (6.249)$$

Note how  $\nabla_h p_b^*$  has no contribution from the potentially very large term,  $-g \rho(\eta_b) \nabla_h \eta_b$ , arising from gradients in the bottom topography.

- (c) [Molemaker et al. \(2015\)](#) and [Gula et al. \(2015\)](#) assume knowledge of the bottom pressure torque along contours of constant topography. Given that knowledge they then make use of the following diagnostic expression for  $p_b^*$

$$p_b^*(s) - p_b^*(s_0) = - \int_{s_0}^s \frac{\hat{z} \cdot (\nabla \eta_b \times \nabla p_b)}{\hat{n} \cdot \nabla \eta_b} ds, \quad (6.250)$$

with  $p_b^*(s_0)$  the value at the arbitrary starting point for the contour. Derive equation (6.250), with the following information of possible use.

- As depicted in Figure 6.12,  $s$  is the arc length along the chosen contour of constant  $\eta_b$ , with  $s$  increasing in the tangent direction,  $\hat{t}$ . Likewise,  $\hat{n}$  is a unit vector pointing to the left of the contour so that  $\hat{n} \cdot \hat{t} = 0$  and  $\hat{t} \times \hat{n} = \hat{z}$ .
- Along any contour of constant  $\eta_b(x, y)$  we have

$$0 = d\eta_b = \nabla \eta_b \cdot d\mathbf{x} = \nabla \eta_b \cdot \frac{d\mathbf{x}}{ds} ds = \nabla \eta_b \cdot \hat{t} ds. \quad (6.251)$$

- The main mathematics of this exercise are contained in an exercise in VOLUME 1.

Equation (6.250) provides a means to compute the anomalous  $p_b^*$  [anomalous relative to  $p_b^*(s_0)$ ] along a constant topography contour. Mapping  $p_b^*(s) - p_b^*(s_0)$  for a suite of contours then provides the means to determine the dynamically relevant portion of the bottom pressure and then, when multiplying by the bottom slope, compute the dynamically relevant portion of the form stress.



## Chapter 7

# POTENTIAL VORTICITY MECHANICS

Potential vorticity (PV) is a dynamical tracer of immense importance to the study of geophysical fluid mechanics. One application of potential vorticity concerns its direct connection to the flow field in certain balanced models (e.g., geostrophically balanced models), with [Hoskins \(1991\)](#) providing an insightful starting point for this perspective. Potential vorticity is also useful as a tracer whose structure signals a variety of dynamical interactions, particularly with boundaries, and that can be directly tied to flow stability properties. In this chapter we establish fundamental properties of potential vorticity and its time evolution. The potential vorticity we consider here is sometimes referred to as Ertel potential vorticity ([Ertel, 1942](#)), which is the most basic of the many potential vorticities encountered in geophysical fluids mechanics.

The barotropic fluid forms a pedagogically useful starting point for our study. However, realistic geophysical flows are baroclinic, and it is the baroclinic fluid where “PV thinking” is arguably the most useful and powerful. The general method exploited for the construction of potential vorticity is to choose a scalar field to strategically orient the absolute vorticity. If the scalar is a material invariant, and it annihilates the baroclinicity vector, then the corresponding potential vorticity is a material invariant in the absence of irreversible processes. For a barotropic fluid, the choice of scalar field is rather arbitrary, with preference given to one that is materially invariant. For a baroclinic fluid we are more restricted since the scalar must orient vorticity in a direction that annihilates the torque from baroclinicity and, ideally, be itself materially invariant in the absence of irreversible processes. Even in the presence of irreversible processes, potential vorticity remains an important flow property that constrains the motion and provides insights into the mechanics of that motion.

### CHAPTER GUIDE

This chapter requires an understanding of vorticity from Chapter 6 as well as skills with vector calculus identities from VOLUME 1. The concepts and methods developed in this chapter are fundamental to the notions of potential vorticity, and are essential for the budget equations developed in Chapter 8.

7.1	Loose threads . . . . .	264
7.2	Potential vorticity in perfect fluids . . . . .	264
7.2.1	Perfect barotropic fluid . . . . .	264
7.2.2	Cylinder between two constant $\chi$ surfaces . . . . .	265
7.2.3	Material invariance . . . . .	266

7.2.4	Perfect baroclinic fluid . . . . .	267
7.2.5	A variety of materially invariant potential vorticities . . . . .	268
7.2.6	Some remarks about potential vorticity . . . . .	268
7.3	Potential vorticity and seawater . . . . .	<b>271</b>
7.3.1	Baroclinicity vector . . . . .	271
7.3.2	Potential vorticity based on potential density, $\varrho$ . . . . .	271
7.3.3	An example EOS admitting a materially invariant PV . . . . .	272
7.3.4	Further reading . . . . .	273
7.4	Potential vorticity evolution in real fluids . . . . .	<b>273</b>
7.5	Flux-form potential vorticity budget . . . . .	<b>274</b>
7.5.1	Deriving the flux-form potential vorticity budget . . . . .	275
7.5.2	PV-substance and the potential vorticity flux . . . . .	275
7.5.3	Gauge freedom in $J^Q$ and the kinematic flux . . . . .	276
7.5.4	Gauge freedom resulting from $\nabla \cdot \omega_a = 0$ . . . . .	278
7.6	A hydrostatic primitive equation and Boussinesq ocean . . . . .	<b>279</b>
7.6.1	Potential vorticity . . . . .	279
7.6.2	Potential vorticity flux vector . . . . .	280
7.6.3	Kinematic derivation of the potential vorticity flux . . . . .	281
7.6.4	A potential vorticity flux vector suited to steady flows . . . . .	282
7.7	Potential vorticity over finite regions . . . . .	<b>284</b>
7.7.1	Differential relations . . . . .	284
7.7.2	Integral relations . . . . .	285
7.8	Exercises . . . . .	<b>286</b>

---

## 7.1 Loose threads

- Exercises

## 7.2 Potential vorticity in perfect fluids

In this section we derive the material invariance of potential vorticity (PV) for a perfect homogeneous fluid.<sup>1</sup> We make use of Kelvin's circulation theorem for an infinitesimal closed loop, in which case the primary object of interest is a particular component of the absolute vorticity.

### 7.2.1 Perfect barotropic fluid

Consider a perfect barotropic fluid. As for the shallow water discussion in Section 5.3.3, we can apply Kelvin's circulation theorem (Section 6.2.3) to an infinitesimal material circuit within the fluid (Figure 7.1) to render the material invariance

$$\frac{D}{Dt}(\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \delta S) = 0, \quad (7.1)$$

<sup>1</sup>A perfect homogeneous fluid has zero viscosity (inviscid) and a single material component. There can be no mixing of matter in this fluid since every fluid element has the same homogeneous concentration. So without viscosity, the homogeneous fluid is perfect.

with  $\delta\mathcal{S}$  the area enclosed by the circuit. The conservation of potential vorticity is built from specializing this result. For that purpose, introduce a materially invariant field

$$\frac{D\chi}{Dt} = 0. \quad (7.2)$$

In most applications,  $\chi$  is a scalar field such as tracer concentration, globally referenced Archimedean buoyancy, Conservative Temperature, or specific entropy. However, in equation (7.21) we consider the non-standard case of  $\chi = z$ , which is relevant for two-dimensional non-divergent barotropic fluids. The one key assumption we make is that  $\chi$  is a smooth field that is not a spatial constant, so that  $|\nabla\chi| \neq 0$ .

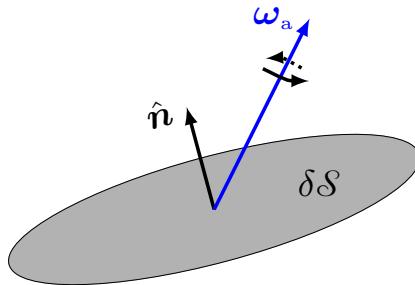


FIGURE 7.1: The projection of the absolute vorticity,  $\omega_a$ , onto the normal direction of an infinitesimal moving surface,  $\hat{\mathbf{n}}\delta\mathcal{S}$ .

### 7.2.2 Cylinder between two constant $\chi$ surfaces

We make use of isosurfaces of  $\chi$  to orient a material circuit used to compute circulation and hence to orient the vorticity. In particular, referring to Figure 7.2, let the circuit bound a small cylinder whose ends sit on isosurfaces with concentrations  $\chi - \delta\chi/2$  and  $\chi + \delta\chi/2$ . We can consider this cylinder to be a portion of a vortex tube that is bounded by the two  $\chi$  isosurfaces. The cylinder's volume is given by

$$\delta V = \delta\mathcal{S} \delta h, \quad (7.3)$$

where  $\delta h$  is the distance between the  $\chi$  isosurfaces. The unit normal direction orienting the area,  $\delta\mathcal{S}$ , is given by

$$\hat{\mathbf{n}} = \nabla\chi/|\nabla\chi|. \quad (7.4)$$

It is the need to define  $\hat{\mathbf{n}}$  that requires us to assume  $|\nabla\chi| \neq 0$ . The distance,  $\delta h$ , between the two isosurfaces is related to the  $\chi$  increment,  $\delta\chi$ , through

$$\delta\chi = \nabla\chi \cdot \delta\mathbf{x} = |\nabla\chi| \hat{\mathbf{n}} \cdot \delta\mathbf{x} = |\nabla\chi| \delta h. \quad (7.5)$$

This result takes on the equivalent form

$$\delta\chi = |\nabla\chi| \delta h = (\hat{\mathbf{n}} \cdot \nabla\chi) \delta h, \quad (7.6)$$

so that the distance (or thickness) between the two isosurfaces is

$$\delta h = \delta\chi/|\nabla\chi|. \quad (7.7)$$

As seen in Figure 7.2, the spatial separation between the two isosurfaces is relatively small in regions of strong scalar gradients (large  $|\nabla\chi|$ ), whereas the separation is relatively large in regions of small  $|\nabla\chi|$ .

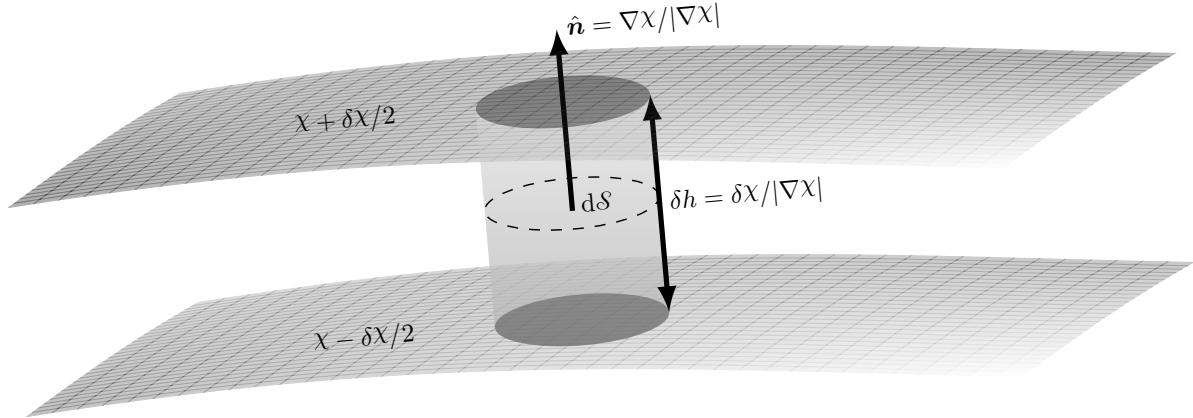


FIGURE 7.2: Illustrating the geometry of a cylindrical region of fluid between two iso-surfaces of a field  $\chi$ , separated by the infinitesimal amount  $\delta\chi$ . The volume of the cylinder is  $\delta V = \delta\mathcal{S}\delta h$ , with  $\delta h$  the thickness and  $\delta\mathcal{S}$  the area. By convention, the unit normal vector,  $\hat{\mathbf{n}} = \nabla\chi/|\nabla\chi|$ , points towards larger values of  $\chi$ . It is here depicted at the center of the cylinder, which differs by an infinitesimal amount from the normal computed on either  $\chi + \delta\chi/2$  or  $\chi - \delta\chi/2$ . If  $\chi$  is a material invariant so that  $D\chi/Dt = 0$ , then so too is its infinitesimal increment,  $D(\delta\chi)/Dt = 0$ . As per equation (7.7), the geometric thickness between the isosurfaces is related to the field increment by  $\delta h = \delta\chi/|\nabla\chi|$ , so that the larger the magnitude of the gradient in the scalar field, the smaller the layer thickness. For a baroclinic fluid, material invariance of potential vorticity in a perfect fluid holds if we can find a field such that  $\hat{\mathbf{n}} \cdot \mathbf{B} = 0$ , with baroclinicity  $\mathbf{B} = (\nabla\rho \times \nabla p)/\rho^2$ . The cylindrical tube acts as a vortex tube for that component of absolute vorticity,  $\omega_a \cdot \hat{\mathbf{n}}$ , that is parallel to the tube.

### 7.2.3 Material invariance

We now have the necessary pieces to write the normal projection of the absolute vorticity according to the following

$$\omega_a \cdot \hat{\mathbf{n}} \delta\mathcal{S} = \frac{\omega_a \cdot \nabla\chi}{|\nabla\chi|} \delta\mathcal{S} \quad \text{equation (7.4)}$$

$$= \frac{\omega_a \cdot \nabla\chi}{|\nabla\chi|} \frac{\delta V}{\delta h} \quad \text{equation (7.3)}$$

$$= (\omega_a \cdot \nabla\chi) \frac{\delta V}{\delta\chi} \quad \text{equation (7.7)}$$

$$= \frac{\omega_a \cdot \nabla\chi}{\rho} \frac{\rho \delta V}{\delta\chi} \quad \text{multiply by } \rho/\rho. \quad (7.8d)$$

Mass is materially invariant so that

$$\frac{D(\rho \delta V)}{Dt} = 0. \quad (7.9)$$

Likewise, by assumption  $\chi$  is materially invariant so that the increment between two  $\chi$  isosurfaces is materially invariant

$$\frac{D(\delta\chi)}{Dt} = 0. \quad (7.10)$$

Bringing these elements into Kelvin's circulation theorem (7.1) leads us to conclude that the potential vorticity,  $Q$ , is also materially invariant

$$Q \equiv \frac{\boldsymbol{\omega}_a \cdot \nabla \chi}{\rho} = \frac{\nabla \cdot (\boldsymbol{\omega}_a \chi)}{\rho} \quad \text{with} \quad \frac{DQ}{Dt} = 0. \quad (7.11)$$

This expression for the potential vorticity is the most general form and it is often referred to as the Ertel potential vorticity ([Ertel, 1942](#)). The first expression shows the numerator as the projection of the absolute vorticity into the direction normal to  $\chi$  isosurfaces. Conversely, it is a measure of the  $\chi$  stratification in the direction of the absolute vorticity vector. The second expression follows since the absolute vorticity has zero divergence so that the numerator is a total divergence. This divergence form of the potential vorticity numerator has important implications for the potential vorticity budgets studied in Section 7.5 and throughout Chapter 8.

#### 7.2.4 Perfect baroclinic fluid

Consider the case of a perfect baroclinic fluid, in which Kelvin's circulation theorem for an infinitesimal circuit takes the form

$$\frac{D}{Dt}(\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \delta S) = \mathbf{B} \cdot \hat{\mathbf{n}} \delta S. \quad (7.12)$$

The source on the right hand side involves the baroclinicity vector,  $\mathbf{B}$ , discussed in Sections 6.2 and 6.4, which is the curl of the pressure gradient acceleration

$$\mathbf{B} = \nabla \times (-\rho^{-1} \nabla p) = \rho^{-2} \nabla \rho \times \nabla p. \quad (7.13)$$

Now assume there exists a materially invariant field,  $D\chi/Dt = 0$ , that also annihilates the baroclinicity vector as in Figure 7.3, so that

$$\mathbf{B} \cdot \hat{\mathbf{n}} = \frac{\mathbf{B} \cdot \nabla \chi}{|\nabla \chi|} = 0. \quad (7.14)$$

In that case, the derivation detailed earlier for the barotropic fluid follows directly for the baroclinic case, in which case we conclude that potential vorticity remains materially invariant

$$\frac{DQ}{Dt} = 0 \quad \text{where} \quad Q = \frac{\boldsymbol{\omega}_a \cdot \nabla \chi}{\rho}. \quad (7.15)$$

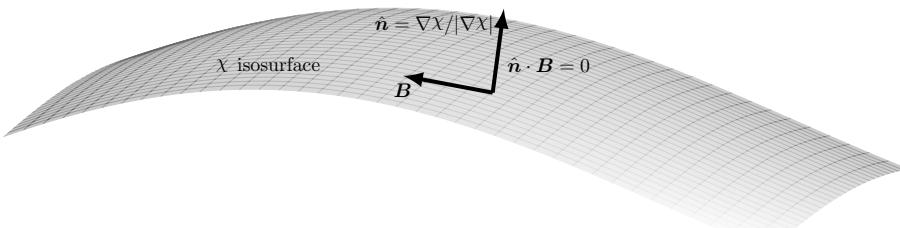


FIGURE 7.3: Material invariance of potential vorticity is ensured for perfect fluids that admit a materially invariant field that also annihilates the baroclinicity vector. Mathematically, this property means that  $\hat{\mathbf{n}} \cdot \mathbf{B} = 0$  where  $\hat{\mathbf{n}} = \nabla \chi / |\nabla \chi|$  is the unit normal direction for the surface. In this figure we depict the baroclinicity vector,  $\mathbf{B}$ , that is aligned with  $\chi$  isosurfaces.

The existence of a materially invariant potential vorticity for perfect baroclinic fluids depends on the existence of a materially invariant scalar field that annihilates the baroclinicity vector. Buoyancy is the most common choice for this field in geophysical fluid applications, with buoyancy typically measured by specific entropy or potential temperature in the atmosphere and potential density in the ocean. We have more to say on the chosen field in the remainder of this chapter as well as in Section 8.3.

### 7.2.5 A variety of materially invariant potential vorticities

The material invariant statement

$$\frac{DQ}{Dt} = \frac{D(\rho^{-1} \boldsymbol{\omega}_a \cdot \nabla \chi)}{Dt} = 0 \quad (7.16)$$

generates a number of further materially invariant fields. First, consider any function of  $Q$ , whereby

$$\frac{D\mathcal{F}(Q)}{Dt} = \frac{d\mathcal{F}}{dQ} \frac{DQ}{Dt} = 0, \quad (7.17)$$

which follows since  $DQ/Dt = 0$ . Among the infinite number of possible functions,  $\mathcal{F}(Q)$ , the most commonly considered is  $\mathcal{F}(Q) = Q^2$ , whose global integral is referred to as the **potential enstrophy**.

Second, consider the *iterated potential vorticity* defined according to

$$Q^{(1)} = \rho^{-1} \boldsymbol{\omega}_a \cdot \nabla \chi \quad \text{and} \quad Q^{(n)} = \boldsymbol{\omega}_a \cdot \nabla Q^{(n-1)} \quad \text{for } n = 2, 3, 4\dots \quad (7.18)$$

As defined,  $Q^{(1)}$  is the familiar Ertel potential vorticity, whereas higher iterations replace the field,  $\chi$ , with  $Q^{(n-1)}$ . Since  $Q^{(n-1)}$  is materially invariant, so too is  $Q^{(n)}$ . Consider the example

$$Q^{(2)} = \boldsymbol{\omega}_a \cdot \nabla Q^{(1)} = \boldsymbol{\omega}_a \cdot \nabla (\rho^{-1} \boldsymbol{\omega}_a \cdot \nabla \chi) = \nabla \cdot [\boldsymbol{\omega}_a \rho^{-1} (\boldsymbol{\omega}_a \cdot \nabla \chi)], \quad (7.19)$$

which reveals that there are  $n$  powers of absolute vorticity for  $Q^{(n)}$ .

### 7.2.6 Some remarks about potential vorticity

#### Perfect fluid PV material invariance $\leftrightarrow$ Kelvin's circulation theorem

Kelvin's circulation theorem from Section 6.2.3 is at the heart of the derivations presented in this section, with the theorem applied to a strategically chosen infinitesimal loop. Because the loop is tiny, we use Stokes' theorem to convert the line integral expression of Kelvin's theorem into a statement about the material evolution of absolute vorticity projected onto the normal direction of the loop, and multiplied by the loop area. We further specialize the theorem to a cylindrical region between two isosurfaces of a materially invariant field. For the perfect barotropic fluid, we require the mass of the cylinder to be materially invariant, as well as the scalar field. In this case there is a potential vorticity that is also materially invariant. For a baroclinic fluid, material invariance of potential vorticity requires a field that is both materially invariant and that annihilates the baroclinicity vector. We have more to say regarding the availability of such fields in the remainder of this chapter.

### Motivating the adjective “potential”

In Section 5.3.2 we motivated the advective “potential” for the shallow water potential vorticity. We do so here for Ertel’s potential vorticity. For that purpose, write potential vorticity in the form

$$Q = \frac{\omega_a \cdot \hat{n}}{\rho} |\nabla \chi| \quad \text{with} \quad \hat{n} = \frac{\nabla \chi}{|\nabla \chi|}. \quad (7.20)$$

In cases where  $\rho$  is roughly a constant (e.g., Boussinesq ocean), and when  $Q$  is materially invariant, the component of the absolute vorticity increases in the direction parallel to  $\nabla \chi$  when the fluid parcel moves into a region where  $|\nabla \chi|$  decreases. Hence, there is a “release” of absolute vorticity aligned with  $\nabla \chi$  in regions where  $\chi$  isosurfaces are spread. We conceive of this increased vorticity as through the stretching of the cylindrical tube extending between  $\chi$  isosurfaces along the  $\hat{n}$  direction, with this tube acting as a vortex tube for that component of vorticity parallel to the tube. In contrast, when the  $\chi$  isosurfaces are tightly packed, then the vortex tube is squashed and so too is the vorticity. We thus conceive of potential vorticity as the “potential” for releasing absolute vorticity that is oriented in the direction parallel to  $\nabla \chi$ .

### Potential vorticity as a dynamical tracer

We refer to potential vorticity as a [dynamical tracer](#) since it depends directly on the velocity field through the vorticity, and since it helps to trace dynamical pathways in the flow. It also depends on the scalar field,  $\chi$ , which is commonly taken as a thermodynamic tracer such as the potential temperature,  $\theta$ . In these cases, potential vorticity embodies both dynamical and thermodynamical information. In contrast, material tracers such as salinity, and thermodynamic tracers such as  $\theta$ , are properties of the fluid whose distribution is affected by the velocity but whose local measurement does not require knowledge of the flow.

### Entropic potential vorticity as the grand unifier

Entropic potential vorticity provides a connection between vorticity (mechanics) and stratification (thermodynamics). By connecting these two basic facets of geophysical fluid flows, the study of potential vorticity and its conservation properties provides a powerful and unique lens to help rationalize the huge variety of geophysical flow regimes, and to predict their response to changes in forcing. It is for this reason that potential vorticity is sometimes considered the grand unifying concept in geophysical fluid mechanics.

### Potential vorticity as a diagnostic tracer

Suppose we have an initial flow field in which  $\rho Q = \omega_a \cdot \nabla \chi = 0$ , which means that the absolute vorticity is aligned with surfaces of constant  $\chi$ . In an inviscid and adiabatic fluid, material conservation of  $Q$  means that  $\omega_a$  remains within constant  $\chi$  surfaces for all time. We infer from this particular example that if we know the evolution of  $\chi$ , then we know the evolution of vortex lines defined by  $\omega_a$ , which in turn allows for the inference of a number of further flow properties. This particular example offers a hint at the multiple applications of “PV thinking” to understand and predict fluid motion, with [Hoskins \(1991\)](#) providing an elegant survey of such thinking.

### Potential vorticity versus momentum

Momentum is affected by pressure, and pressure fluctuations propagate at the speed of sound (see discussion of acoustic waves in VOLUME 4) for compressible flows (including the Boussinesq ocean as discussed in VOLUME 2), whereas they move with infinite speed for constant density (incompressible) fluids. In contrast, potential vorticity, for those cases where we can remove the effects of the baroclinicity vector, does not directly feel the impacts from pressure fluctuations. Hence, potential vorticity, much like vorticity in a barotropic fluid, evolves much slower and more locally than momentum. This dynamical difference offers a key reason that potential vorticity offers added insights into fluid flows beyond that afforded by momentum.

### There are numerous forms for potential vorticity

The expression (7.11) is, on first glance, quite distinct from the shallow water (Rossby) potential vorticity,  $Q = (\zeta + f)/h$ , studied in Chapter 5 (see equation (5.30)). However, as shown in Section 14.3, they are closely related for the special case of entropic potential vorticity in a Boussinesq ocean when formulated using isopycnal/isentropic coordinates.

Even so, there are a variety of other forms for potential vorticity, with the forms (and physical dimensions) depending on the dynamical and thermodynamical properties. We encounter some further forms of potential vorticity in the remainder of this chapter, as well as in the oceanic potential vorticity discussions of Chapter 14 and in our study of balanced models in Part III. The review paper by Müller (1995) offers a lucid presentation of potential vorticity and its many forms encountered in ocean physics.

### Potential vorticity for horizontally non-divergent barotropic flow

A non-standard, but relevant, choice for the function  $\chi$  used to define potential vorticity is given by the vertical coordinate

$$\chi = z, \quad (7.21)$$

in which case

$$\rho Q = \omega_a \cdot \hat{z} = \zeta_a. \quad (7.22)$$

In Chapter 4, we consider the two-dimensional non-divergent barotropic model, in which case  $\rho$  is a constant and

$$w = \frac{Dz}{Dt} = 0. \quad (7.23)$$

The corresponding materially invariant potential vorticity is the vertical component of the absolute vorticity divided by the constant density. Ignoring the constant density factor leads us to identify the absolute vorticity as the Ertel potential vorticity for this flow

$$q = \zeta_a = \zeta + f. \quad (7.24)$$

As discussed in Section 4.3.3, the horizontally non-divergent barotropic flow also maintains material constancy (in the absence of non-conservative processes) of the Rossby potential vorticity,  $(\zeta + f)/h$ , with this property holding since  $Dh/Dt = 0$  for this flow.

## 7.3 Potential vorticity and seawater

As seen in Section 7.2, material invariance of potential vorticity for a perfect fluid requires a materially invariant scalar field to annihilate the baroclinicity vector. There is no such scalar field for the ocean with a realistic seawater equation of state. Nonetheless, there are important approximate cases that allow for material potential vorticity invariance, and we explore such cases in this section.

### 7.3.1 Baroclinicity vector

Recall the baroclinicity vector given by (Sections 6.2 and 6.4)

$$\mathbf{B} = \frac{\nabla\rho \times \nabla p}{\rho^2}. \quad (7.25)$$

If we take the *in situ* density as the scalar field to define potential vorticity, then  $\mathbf{B} \cdot \nabla\rho = 0$ . However, *in situ* density is not a conserved scalar in the ocean due to pressure effects. Namely, with *in situ* density having the function dependence  $\rho = \rho(S, \Theta, p)$  (see VOLUME 2), its material time derivative is

$$\frac{D\rho}{Dt} = \frac{\partial\rho}{\partial S} \frac{DS}{Dt} + \frac{\partial\rho}{\partial\Theta} \frac{D\Theta}{Dt} + \frac{\partial\rho}{\partial p} \frac{Dp}{Dt}. \quad (7.26)$$

Even when salinity and Conservative Temperature are materially constant,  $DS/Dt = 0$  and  $D\Theta/Dt = 0$ , the *in situ* density has a nonzero material time derivative due to material pressure changes,  $Dp/Dt \neq 0$ . Material changes in the pressure of a fluid element arise even in the absence of irreversible processes such as mixing. In general, such mechanical changes arise due to the gradients in the pressure field that the fluid element feels. Given that pressure affects *in situ* density, with such effects occurring even in a perfect fluid, we conclude that *in situ* density is not an appropriate scalar for developing a materially invariant potential vorticity. For the same reason, we do not consider pressure as a suitable scalar field.

### 7.3.2 Potential vorticity based on potential density, $\varrho$

Potential density is commonly used in ocean physics (see VOLUME 2), where potential density is the *in situ* density referenced to a chosen pressure, and so it is written as<sup>2</sup>

$$\varrho(S, \Theta) = \rho(S, \Theta, p = p_{\text{ref}}). \quad (7.27)$$

The material time derivative of potential density is

$$\frac{D\varrho}{Dt} = \frac{\partial\varrho}{\partial S} \frac{DS}{Dt} + \frac{\partial\varrho}{\partial\Theta} \frac{D\Theta}{Dt}, \quad (7.28)$$

which vanishes in the absence of material changes to salinity and Conservative Temperature, which generally arise from irreversible processes such as mixing. When using potential density as the scalar field for potential vorticity, the baroclinicity vector projects onto the diapycnal

---

<sup>2</sup>Ocean physicists often choose the reference pressure as the standard atmospheric sea level pressure. However, that is not required for the following formalism to hold, with any reference pressure suitable.

direction according to

$$\rho^2 \mathbf{B} \cdot \nabla \varrho = (\nabla \rho \times \nabla p) \cdot \nabla \varrho \quad (7.29a)$$

$$= (\nabla \varrho \times \nabla \rho) \cdot \nabla p \quad (7.29b)$$

$$= [(\varrho_S \nabla S + \varrho_\Theta \nabla \Theta) \times (\rho_S \nabla S + \rho_\Theta \nabla \Theta + \rho_p \nabla p)] \cdot \nabla p \quad (7.29c)$$

$$= [(\varrho_S \nabla S + \varrho_\Theta \nabla \Theta) \times (\rho_S \nabla S + \rho_\Theta \nabla \Theta)] \cdot \nabla p \quad (7.29d)$$

$$= [\varrho_S \nabla S \times \rho_\Theta \nabla \Theta + \varrho_\Theta \nabla \Theta \times \rho_S \nabla S] \cdot \nabla p \quad (7.29e)$$

$$= (\varrho_S \rho_\Theta - \varrho_\Theta \rho_S) (\nabla S \times \nabla \Theta) \cdot \nabla p, \quad (7.29f)$$

where we used the shorthand notation for partial derivatives

$$\rho_S = \frac{\partial \rho}{\partial S} \quad \text{and} \quad \varrho_S = \frac{\partial \varrho}{\partial S} \quad (7.30a)$$

$$\rho_\Theta = \frac{\partial \rho}{\partial \Theta} \quad \text{and} \quad \varrho_\Theta = \frac{\partial \varrho}{\partial \Theta}. \quad (7.30b)$$

Note that the triple product,  $(\nabla S \times \nabla \Theta) \cdot \nabla p$ , also appears in the discussion of [neutral helicity](#) in [VOLUME 2](#). Equation (7.29f) allows us to identify cases where the baroclinicity vector is annihilated,  $\mathbf{B} \cdot \nabla \varrho = 0$ , thus yielding a materially invariant potential vorticity in the absence of irreversible processes.

- **UNIFORM SALINITY OR UNIFORM CONSERVATIVE TEMPERATURE:** If salinity or Conservative Temperature are spatially uniform, then  $\mathbf{B} \cdot \nabla \varrho = 0$ .
- **ADDITIVE PRESSURE DEPENDENCE TO THE *in situ* DENSITY:** There is a materially invariant potential vorticity with a vanishing thermodynamic pre-factor in equation (7.29f),  $\varrho_S \rho_\Theta - \varrho_\Theta \rho_S$ . This term does not generally vanish since the ocean has a pressure dependent equation of state, and this pressure dependence generally means that  $\mathbf{B} \cdot \nabla \varrho \neq 0$ . Even so, we can annihilate the baroclinicity vector if the *in situ* density has a pressure dependence that is additive, in which case we can write

$$\rho(S, \Theta, p) = \varrho(S, \Theta) + F(p) - F(p_{\text{ref}}) \implies \varrho_S \rho_\Theta - \varrho_\Theta \rho_S = 0. \quad (7.31)$$

Notably, we did not assume a linear equation of state; only that it has the special functional form in equation (7.31). For some cases, we may assume  $F$  to be a constant, in which case there is no pressure dependence so that *in situ* density is the same as potential density.

### 7.3.3 An example EOS admitting a materially invariant PV

An explicit realization of the equation of state (7.31) can be found by taking a Taylor series expansion of the *in situ* density around the reference pressure, and evaluating the derivatives in the expansion in terms of a chosen reference pressure, reference salinity, and reference Conservative Temperature

$$\rho(S, \Theta, p) \approx \varrho(S, \Theta) + (p - p_{\text{ref}}) \underbrace{\left[ \frac{\partial \rho}{\partial p} \right]_{S=S_{\text{ref}}, \Theta=\Theta_{\text{ref}}, p=p_{\text{ref}}} + H.O.T.}_{F(p) - F(p_{\text{ref}})} \quad (7.32)$$

where

$$\varrho(S, \Theta) = \rho(S, \Theta, p_{\text{ref}}) \quad (7.33)$$

is the potential density referenced to  $p = p_{\text{ref}}$ , and where *H.O.T.* symbolizes higher order terms. This approach ignores the salinity and Conservative Temperature dependence of terms in the Taylor series expansion. Ignoring this dependence is a rather good approximation for many purposes since the ocean sound speed is not far from a constant

$$c_s^{-2} = \frac{\partial \rho}{\partial p} \approx \text{constant}. \quad (7.34)$$

In this case, the equation of state takes the form

$$\rho(S, \Theta, p) \approx \varrho(S, \Theta) + \frac{p - p_{\text{ref}}}{c_s^2}, \quad (7.35)$$

### 7.3.4 Further reading

The presentation given here follows that given in Section 4.5.4 of [Vallis \(2017\)](#). [Straub \(1999\)](#) focuses on the source of potential vorticity arising from a nonzero thermobaricity parameter,  $\mathcal{T} = \partial_p(\alpha/\beta)$  (see Section 19.3.3). In Section 7.7 we reconsider the notions presented here by suggesting the relevance of an alternative potential vorticity field that is attached to a finite sized region rather than to a fluid particle.

## 7.4 Potential vorticity evolution in real fluids

Thus far we have considered perfect fluids, with the use of Kelvin's circulation theorem a suitable framework to derive the material invariance of potential vorticity. In this section we consider a real fluid that contains non-conservative processes. Potential vorticity is no longer materially invariant when exposed to non-conservative processes such as mixing, friction, and diabatic sources.

To develop the potential vorticity budget in the presence of non-conservative processes, we pursue an algebraic approach that starts from the vorticity equation (6.42)

$$\rho \frac{D(\boldsymbol{\omega}_a/\rho)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \mathbf{B} + \nabla \times \mathbf{F}, \quad (7.36)$$

where  $\mathbf{F}$  is the acceleration from non-conservative forces and  $\mathbf{B}$  the baroclinicity vector. Furthermore, we introduce a scalar field that generally has a nonzero material evolution

$$\frac{D\chi}{Dt} = \dot{\chi}, \quad (7.37)$$

with  $\dot{\chi}$  arising from diffusion, sources, boundary fluxes, or other processes that lead to material evolution of  $\chi$ .

As part of the manipulations in this section, we make use of the identity

$$(\boldsymbol{\omega}_a \cdot \nabla) \frac{D\chi}{Dt} = \boldsymbol{\omega}_a \cdot \frac{D(\nabla\chi)}{Dt} + [(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}] \cdot \nabla\chi, \quad (7.38)$$

which is readily proven by expanding terms and assuming Cartesian coordinates. Rearrangement,

and use of the scalar equation (7.37), leads to

$$\boldsymbol{\omega}_a \cdot \frac{D(\nabla\chi)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{\chi} - [(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}] \cdot \nabla\chi. \quad (7.39)$$

Now project the vorticity equation (7.36) onto the direction normal to the  $\chi$  isosurfaces

$$\rho \nabla\chi \cdot \frac{D(\boldsymbol{\omega}_a/\rho)}{Dt} = \nabla\chi \cdot [(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}] + \nabla\chi \cdot (\mathbf{B} + \nabla \times \mathbf{F}). \quad (7.40)$$

The sum of equations (7.39) and (7.40) leads to

$$\rho \frac{D(\nabla\chi \cdot \boldsymbol{\omega}_a/\rho)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{\chi} + \nabla\chi \cdot (\mathbf{B} + \nabla \times \mathbf{F}). \quad (7.41)$$

This equation is general so that it applies to any scalar field.

To simplify the source terms on the right hand side of equation (7.41), follow the discussion from Section 7.2.4 by assuming that  $\chi$  annihilates the baroclinicity vector.<sup>3</sup> This scalar field is typically given by potential temperature, specific entropy, buoyancy, or potential density. We thus have

$$\nabla\chi \cdot \mathbf{B} = 0, \quad (7.42)$$

which in turn leads to the potential vorticity equation in the presence of irreversible processes such as friction and mixing

$$\rho \frac{DQ}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{\chi} + \nabla\chi \cdot (\nabla \times \mathbf{F}), \quad (7.43)$$

where the potential vorticity is again given by

$$Q = \rho^{-1} \boldsymbol{\omega}_a \cdot \nabla\chi. \quad (7.44)$$

If  $\chi$  is a thermodynamic scalar such as potential entropy, then the material evolution of potential vorticity is affected by diabatic processes (heating and cooling) as well as friction. Hence, the potential vorticity of a fluid element can be either generated or destroyed depending on details of these irreversible process. Such processes are often localized to areas of mixing as well as to boundaries where strong mechanical and/or buoyant processes are active. The study of how potential vorticity is materially modified by irreversible processes forms an important area of research in potential vorticity dynamics. We have more to say on this notion when studying finite volume budgets of potential vorticity in Chapter 8.

## 7.5 Flux-form potential vorticity budget

The material invariance of potential vorticity is an example of a material or Lagrangian conservation property of perfect fluids, with the material conservation statement  $\rho DQ/Dt = 0$  having its flux-form expression

$$\partial_t(\rho Q) + \nabla \cdot (\rho \mathbf{v} Q) = 0 \quad \text{perfect fluid}, \quad (7.45)$$

---

<sup>3</sup>In Section 8.3 we study what happens when no such scalar exists.

which is derived through use of mass conservation in the form

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}. \quad (7.46)$$

Following the formalism established for material tracers in VOLUME 2, the flux-form local conservation law (7.45) leads to conservation properties over finite regions, which we refer to as **global conservation** laws. In this section we examine the flux-form budget in the presence of non-conservative processes. In particular, we show that the Eulerian evolution of potential vorticity continues to be determined by the convergence of a flux, thus allowing for natural extensions to global conservation laws and budget analyses. These properties were observed earlier in our study of shallow water potential vorticity in Section 5.4.2 and absolute vorticity in Section 6.3.4. These common features of the various vorticity budgets relate to the ability to write components of the absolute vorticity (including the potential vorticity) as the divergence of a vector.

### 7.5.1 Deriving the flux-form potential vorticity budget

To transform the material evolution equation (7.43) into a flux-form equation we make use of the following identities

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad \text{relating material and Eulerian time changes} \quad (7.47a)$$

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v} \quad \text{mass conservation} \quad (7.47b)$$

$$(\boldsymbol{\omega}_a \cdot \nabla) \dot{\chi} = \nabla \cdot (\boldsymbol{\omega}_a \dot{\chi}) \quad \text{absolute vorticity is non-divergent: } \nabla \cdot \boldsymbol{\omega}_a = 0 \quad (7.47c)$$

$$\nabla \chi \cdot (\nabla \times \mathbf{F}) = \nabla \cdot (\mathbf{F} \times \nabla \chi) \quad \text{divergence of curl vanishes.} \quad (7.47d)$$

The identity (7.47d) follows from

$$\nabla \chi \cdot (\nabla \times \mathbf{F}) = \nabla \cdot (\chi \nabla \times \mathbf{F}) = \nabla \cdot [\nabla \times (\chi \mathbf{F}) - \nabla \chi \times \mathbf{F}] = \nabla \cdot (\mathbf{F} \times \nabla \chi), \quad (7.48)$$

where a vanishing divergence of a curl is needed to reach the first and third equalities. These identities then lead to the material evolution equation

$$\rho \frac{DQ}{Dt} = \nabla \cdot (\boldsymbol{\omega}_a \dot{\chi} + \mathbf{F} \times \nabla \chi). \quad (7.49)$$

Now converting the material time derivative into its Eulerian expression, and making use of mass conservation, renders the flux-form potential vorticity budget equation

$$\partial_t(\rho Q) + \nabla \cdot [\rho Q \mathbf{v} - \boldsymbol{\omega}_a \dot{\chi} - \mathbf{F} \times \nabla \chi] = 0. \quad (7.50)$$

### 7.5.2 PV-substance and the potential vorticity flux

The budget equation (7.50) says that the density-weighted potential vorticity,

$$\rho Q = \boldsymbol{\omega}_a \cdot \nabla \chi, \quad (7.51)$$

has a local time tendency determined by the convergence of the potential vorticity flux vector

$$\partial_t(\rho Q) = -\nabla \cdot \mathbf{J}^Q \quad \text{with} \quad \mathbf{J}^Q = \rho Q \mathbf{v} - \boldsymbol{\omega}_a \dot{\chi} + \nabla \chi \times \mathbf{F}. \quad (7.52)$$

The potential vorticity flux vector,  $\mathbf{J}^Q$ , compares to that found for shallow water potential vorticity given by equation (5.55). The budget (7.50) follows a form similar to material tracers detailed in VOLUME 2, though here with some particularly specific terms in the potential vorticity flux vector,  $\mathbf{J}^Q$ . The correspondence suggests that one consider equation (7.50) as the local budget for **potential vorticity substance** (PV-substance), with  $Q$  the concentration of PV-substance and  $\mathbf{J}^Q$  its flux. This interpretation is pursued further in Chapter 8 when exposing the rather novel properties of budgets for PV-substance when integrated over regions bounded by isentropes.

The first term in the PV-substance flux vector (7.52) arises from the advection of PV-substance; the second contribution arises from processes leading to material evolution of  $\chi$ ; and the third from any non-conservative acceleration,  $\mathbf{F}$ , that is not parallel to  $\nabla\chi$ . Note that the form of the non-conservative contribution,  $\nabla\chi \times \mathbf{F}$ , suggests that we think of  $\mathbf{F}$  as contributing to a torque that rotates the  $\chi$  isosurfaces as it modifies the PV-substance (see Figure 7.4).

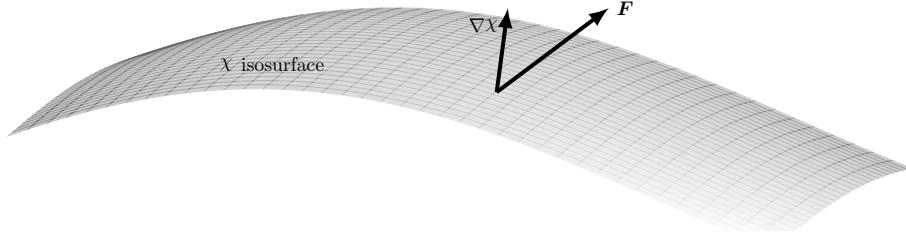


FIGURE 7.4: The contribution from any non-conservative acceleration (e.g., friction) to the potential vorticity flux is given by  $\mathbf{J}_{\text{friction}} = \nabla\chi \times \mathbf{F}$ . This cross product is nonzero only when  $\mathbf{F}$  is not fully aligned with  $\nabla\chi$ , so that non-conservative forces create potential vorticity by rotating  $\chi$  isosurfaces. Hence, if  $\mathbf{F}$  is aligned with  $\nabla\chi$ , or when there is no spatial structure to  $\nabla\chi \times \mathbf{F}$  (i.e., zero divergence), then non-conservative accelerations do not contribute to potential vorticity evolution. This interpretation is analogous to that given to the effects from baroclinicity on vorticity given in Section 6.4.

### 7.5.3 Gauge freedom in $\mathbf{J}^Q$ and the kinematic flux

As seen by the potential vorticity equation (7.52), the time tendency for PV-substance,  $\partial_t(\rho Q)$ , is unchanged by adding the curl of a vector to the flux,  $\mathbf{J}^Q$ . This ambiguity manifests a **gauge symmetry**. We here exhibit the gauge symmetry associated with  $\mathbf{J}^Q$ , and then in Section 7.5.4 expose yet another gauge symmetry associated with the potential vorticity itself.<sup>4</sup>

#### Kinematic potential vorticity flux

Consider the identity

$$\rho Q = \nabla\chi \cdot \boldsymbol{\omega}_a = \nabla \cdot (\chi \boldsymbol{\omega}_a), \quad (7.53)$$

which means that

$$\partial_t(\rho Q) = \partial_t[\nabla \cdot (\chi \boldsymbol{\omega}_a)] = \nabla \cdot [\partial_t(\chi \boldsymbol{\omega}_a)] = -\nabla \cdot \mathbf{J}^{\text{kin}}, \quad (7.54)$$

<sup>4</sup>We offer more discussion of gauge symmetry in VOLUME 2, as part of the discussion of non-divergent flow kinematics. We further the use of gauge freedom for the study of potential vorticity in Section 7.6 as well as throughout Chapter 8.

where the time derivative commutes with the divergence operator. The final equality introduced the kinematic potential vorticity flux

$$\mathbf{J}^{\text{kin}} \equiv -\partial_t(\chi \boldsymbol{\omega}_{\text{a}}), \quad (7.55)$$

which compares to the shallow water version given by equation (5.58). We conclude that the two potential vorticity fluxes,  $\mathbf{J}^{\text{kin}}$  and  $\mathbf{J}^Q$ , have equal divergence

$$\nabla \cdot \mathbf{J}^{\text{kin}} = \nabla \cdot \mathbf{J}^Q, \quad (7.56)$$

and so they differ at most by the curl of a vector

$$\mathbf{J}^{\text{kin}} = \mathbf{J}^Q + \nabla \times \mathbf{A}, \quad (7.57)$$

with  $\mathbf{A}$  referred to as a **gauge function**. There are a variety of potential vorticity fluxes encountered in the study of potential vorticity, with each flux differing by the curl of a gauge function.

### Relating the potential vorticity fluxes $\mathbf{J}^{\text{kin}}$ and $\mathbf{J}^Q$

We determine the relation between  $\mathbf{J}^{\text{kin}}$  and  $\mathbf{J}^Q$  through the following steps starting from

$$\mathbf{J}^{\text{kin}} = -\partial_t(\chi \boldsymbol{\omega}_{\text{a}}) = -\partial_t \chi \boldsymbol{\omega}_{\text{a}} - \chi (\nabla \times \partial_t \mathbf{v}). \quad (7.58)$$

To reach this result required the identity

$$\partial_t \boldsymbol{\omega}_{\text{a}} = \nabla \times \partial_t(\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) = \nabla \times \partial_t \mathbf{v}, \quad (7.59)$$

which follows since  $\boldsymbol{\Omega}$  is time independent, and  $\partial_t \mathbf{x} = 0$  since the Eulerian time derivative is computed at a fixed space point. We next use the identity

$$-\chi (\nabla \times \partial_t \mathbf{v}) = -\nabla \times (\chi \partial_t \mathbf{v}) + \nabla \chi \times \partial_t \mathbf{v}, \quad (7.60)$$

so that the kinematic flux from equation (7.58) becomes

$$\mathbf{J}^{\text{kin}} = -\partial_t \chi \boldsymbol{\omega}_{\text{a}} + \nabla \chi \times \partial_t \mathbf{v} - \nabla \times (\chi \partial_t \mathbf{v}). \quad (7.61)$$

Next recall the vector-invariant form of the velocity equation (6.33), here written in the form

$$\partial_t \mathbf{v} + \boldsymbol{\omega}_{\text{a}} \times \mathbf{v} = -\rho^{-1} \nabla p - \nabla \mathcal{M} + \mathbf{F}, \quad (7.62)$$

where we introduced the mechanical energy per mass

$$\mathcal{M} = \mathbf{v} \cdot \mathbf{v}/2 + \Phi. \quad (7.63)$$

Equation (7.62) then leads to the cross product

$$\nabla \chi \times \partial_t \mathbf{v} = \rho Q \mathbf{v} - (\dot{\chi} - \partial_t \chi) \boldsymbol{\omega}_{\text{a}} - \nabla \chi \times \rho^{-1} \nabla p - \nabla \times (\chi \nabla \mathcal{M}) + \nabla \chi \times \mathbf{F}, \quad (7.64)$$

where we used the identities

$$(\boldsymbol{\omega}_{\text{a}} \times \mathbf{v}) \times \nabla \chi = (\nabla \chi \cdot \boldsymbol{\omega}_{\text{a}}) \mathbf{v} - (\nabla \chi \cdot \mathbf{v}) \boldsymbol{\omega}_{\text{a}} = \rho Q \mathbf{v} - (\dot{\chi} - \partial_t \chi) \boldsymbol{\omega}_{\text{a}}. \quad (7.65)$$

Making use of equation (7.64) brings the kinematic flux (7.61) into the form

$$\mathbf{J}^{\text{kin}} = \rho Q \mathbf{v} - \dot{\chi} \boldsymbol{\omega}_{\text{a}} - \nabla \chi \times \rho^{-1} \nabla p + \nabla \chi \times \mathbf{F} - \nabla \times (\chi \nabla \mathcal{M} + \chi \partial_t \mathbf{v}). \quad (7.66)$$

Comparing to equation (7.52) for  $\mathbf{J}^Q$  leads to the relation

$$\mathbf{J}^{\text{kin}} = \mathbf{J}^Q - \chi \mathbf{B} + \nabla \times [\chi (\partial_t \mathbf{v} + \rho^{-1} \nabla p + \nabla \mathcal{M})], \quad (7.67)$$

where we introduced the baroclinicity vector,  $\mathbf{B} = \nabla \times (-\rho^{-1} \nabla p)$ . Although the term  $-\chi \mathbf{B}$  does not appear in the form of a curl, it does have a zero divergence

$$\nabla \cdot (\chi \mathbf{B}) = \nabla \chi \cdot \mathbf{B} + \chi \nabla \cdot \mathbf{B} = 0 + 0, \quad (7.68)$$

where we set  $\nabla \chi \cdot \mathbf{B} = 0$ , and noted that the baroclinicity has zero divergence,  $\nabla \cdot \mathbf{B} = \nabla \cdot [\nabla \times (-\rho^{-1} \nabla p)] = 0$ . As a result, we have shown that  $\nabla \cdot \mathbf{J}^{\text{kin}} = \nabla \cdot \mathbf{J}^Q$ , which means that each flux has the same affect on the time evolution of  $\rho Q$ .

#### 7.5.4 Gauge freedom resulting from $\nabla \cdot \boldsymbol{\omega}_{\text{a}} = 0$

The divergence form of PV-substance given by equation (7.53) reveals that  $\rho Q$  remains unchanged if we add a total curl to the argument of the divergence operator. In particular, consider the identities<sup>5</sup>

$$\chi \boldsymbol{\omega}_{\text{a}} = \chi (\nabla \times \mathbf{v} + 2 \boldsymbol{\Omega}) \quad \text{since } \boldsymbol{\omega}_{\text{a}} = \boldsymbol{\omega} + 2 \boldsymbol{\Omega} \quad (7.69a)$$

$$= \nabla \times (\chi \mathbf{v}) - \nabla \chi \times \mathbf{v} + \chi 2 \boldsymbol{\Omega} \quad \text{move gradient operator.} \quad (7.69b)$$

We have thus moved the curl operation acting on the velocity field (to compute relative vorticity) onto a gradient of the scalar field, with a total curl making up the difference. Since  $\nabla \cdot [\nabla \times (\chi \mathbf{v})] = 0$ , we are led to the equivalent expressions for PV-substance

$$\rho Q = \nabla \chi \cdot \boldsymbol{\omega}_{\text{a}} = \nabla \cdot (\chi 2 \boldsymbol{\Omega} + \chi \boldsymbol{\omega}) = \nabla \cdot (2 \boldsymbol{\Omega} \chi - \nabla \chi \times \mathbf{v}) = \nabla \chi \cdot (2 \boldsymbol{\Omega}) - \nabla \cdot (\nabla \chi \times \mathbf{v}). \quad (7.70)$$

The result of these manipulations is an expression for the Ertel potential vorticity that does not involve the relative vorticity, but is instead written as the convergence of a vector

$$\rho Q = -\nabla \cdot (-2 \boldsymbol{\Omega} \chi + \nabla \chi \times \mathbf{v}). \quad (7.71)$$

This form involves the component of the velocity that is parallel to  $\chi$  isosurfaces since

$$(\nabla \chi \times \mathbf{v}) \cdot \nabla \chi = 0. \quad (7.72)$$

Besides offering a curious expression for potential vorticity that does not require the relative vorticity, we show in Section 8.4 how the formulation (7.71) can be especially useful for developing budgets of integrated PV-substance.

In the manipulations (7.69a)-(7.69b), we did not touch the planetary vorticity term. We certainly could do so, in which case

$$\chi 2 \boldsymbol{\Omega} = \chi \nabla \times (\boldsymbol{\Omega} \times \mathbf{x}) = \nabla \times (\chi \boldsymbol{\Omega} \times \mathbf{x}) - \nabla \chi \times (\boldsymbol{\Omega} \times \mathbf{x}). \quad (7.73)$$

---

<sup>5</sup>We encounter similar mathematical manipulations in Section 16.4 when studying the connection between advection and skew diffusion in the tracer equation.

The term  $\nabla \times (\chi \boldsymbol{\Omega} \times \mathbf{x})$  drops out when taking the divergence. However, the term  $\boldsymbol{\Omega} \times \mathbf{x}$  requires us to evaluate the position vector,  $\mathbf{x}$ , for each point in the fluid, and doing so is not generally convenient. For this reason, we do not modify the planetary vorticity contribution to the potential vorticity, preferring to keep the form  $\chi 2\boldsymbol{\Omega}$  in equation (7.71).

## 7.6 A hydrostatic primitive equation and Boussinesq ocean

Building on the vorticity budget in Section 6.7, we here develop the PV-substance budget for a hydrostatic primitive equation Boussinesq ocean in the presence of diabatic processes and non-conservative forces such as friction. For that purpose, recall the vorticity equation for a hydrostatic and Boussinesq ocean (6.151)

$$\frac{D\omega_a^{hy}}{Dt} = \underbrace{(\omega_a^{hy} \cdot \nabla) \mathbf{v}}_{\text{stretching + tilting}} + \underbrace{\nabla \times \hat{\mathbf{z}} b}_{\text{baroclinicity}} + \underbrace{\nabla \times \mathbf{F}}_{\text{friction curl}}, \quad (7.74)$$

where  $b$  is the Archimedean buoyancy used in our discussion of the Boussinesq ocean in VOLUME 2, and

$$\omega_a^{hy} = \omega^{hy} + f \hat{\mathbf{z}} = \nabla \times \mathbf{u} + f \hat{\mathbf{z}} \quad (7.75)$$

is the absolute vorticity for the hydrostatic fluid with  $\mathbf{u}$  the horizontal velocity (see equations (6.140) and (6.149)).

### 7.6.1 Potential vorticity

Baroclinicity is eliminated from the vorticity equation (7.74) by projecting the absolute vorticity onto the direction normal to buoyancy surfaces

$$\nabla b \cdot \frac{D\omega_a^{hy}}{Dt} = \nabla b \cdot [(\omega_a^{hy} \cdot \nabla) \mathbf{v}] + \nabla b \cdot (\nabla \times \mathbf{F}), \quad (7.76)$$

where we used

$$\nabla b \cdot (\nabla \times \hat{\mathbf{z}} b) = \nabla b \cdot (\nabla b \times \hat{\mathbf{z}}) = 0. \quad (7.77)$$

We next make use of the identity

$$\frac{D(\partial b / \partial x^i)}{Dt} = \frac{\partial}{\partial x^i} \left[ \frac{Db}{Dt} \right] - \nabla b \cdot \frac{\partial \mathbf{v}}{\partial x^i} = \frac{\partial \dot{b}}{\partial x^i} - \nabla b \cdot \frac{\partial \mathbf{v}}{\partial x^i}, \quad (7.78)$$

so that

$$\omega_a^{hy} \cdot \left[ \frac{D \nabla b}{Dt} \right] = \omega_a^{hy} \cdot \nabla \dot{b} - \nabla b \cdot [(\omega_a^{hy} \cdot \nabla) \mathbf{v}]. \quad (7.79)$$

Making use of this result in equation (7.76) renders

$$\nabla b \cdot \frac{D\omega_a^{hy}}{Dt} + \omega_a^{hy} \cdot \frac{D \nabla b}{Dt} = \omega_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \times \mathbf{F}), \quad (7.80)$$

which leads to

$$\frac{DQ}{Dt} = (\omega_a^{hy} \cdot \nabla) \dot{b} + \nabla b \cdot (\nabla \times \mathbf{F}) \quad (7.81)$$

where

$$Q = \omega_a^{hy} \cdot \nabla b = \omega^{hy} \cdot \nabla b + f \partial_z b \quad (7.82)$$

is the potential vorticity for a rotating hydrostatic Boussinesq ocean. Potential vorticity is materially invariant for the inviscid and adiabatic case, in which  $\mathbf{F} = 0$  and  $\dot{b} = 0$ .

It is sometimes useful to split the hydrostatic vorticity into its vertical and horizontal terms as per equation (6.140). In this way, potential vorticity takes on the form

$$Q = \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + (\zeta + f) \frac{\partial b}{\partial z} = \hat{\mathbf{z}} \cdot \left[ \frac{\partial \mathbf{u}}{\partial z} \times \nabla b \right] + (\zeta + f) \frac{\partial b}{\partial z}. \quad (7.83)$$

This expression plays an important role in characterizing flows with order unity Rossby number, where the term  $\hat{\mathbf{z}} \cdot (\partial_z \mathbf{u} \times \nabla b)$  can become comparable to  $(\zeta + f) \partial_z b$ , particularly in regions of strong horizontal buoyancy fronts such as those studied by [Thomas et al. \(2008\)](#) and [Thomas et al. \(2013\)](#).

### 7.6.2 Potential vorticity flux vector

The material form of the potential vorticity equation (7.81) is converted into its flux-form via

$$\partial_t Q + \nabla \cdot (\mathbf{v} Q) = \omega_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \times \mathbf{F}) \quad (7.84a)$$

$$= \nabla \cdot [\dot{b} \omega_a^{hy} + b (\nabla \times \mathbf{F})] \quad (7.84b)$$

$$= \nabla \cdot [\dot{b} \omega_a^{hy} + \nabla \times (b \mathbf{F}) - \nabla b \times \mathbf{F}] \quad (7.84c)$$

$$= \nabla \cdot (\dot{b} \omega_a^{hy} - \nabla b \times \mathbf{F}), \quad (7.84d)$$

where we used

$$\nabla \cdot \mathbf{v} = 0 \quad \text{Boussinesq flow is non-divergent} \quad (7.85a)$$

$$\nabla \cdot \omega_a^{hy} = 0 \quad \text{vorticity always has zero divergence} \quad (7.85b)$$

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0 \quad \text{divergence of curl vanishes} \quad (7.85c)$$

$$\nabla \cdot [\nabla \times (b \mathbf{F})] = 0 \quad \text{divergence of curl vanishes.} \quad (7.85d)$$

The conservation equation (7.84d) allows us to identify a potential vorticity flux vector for the hydrostatic Boussinesq ocean

$$\mathbf{J}^Q = \mathbf{v} Q - \dot{b} \omega_a^{hy} + \nabla b \times \mathbf{F}, \quad (7.86)$$

so that the potential vorticity equation takes the form

$$\partial_t Q + \nabla \cdot \mathbf{J}^Q = 0. \quad (7.87)$$

The potential vorticity flux (7.86) compares directly to the potential vorticity flux (7.52) suited to the non-hydrostatic and non-Boussinesq fluid. As such, it is comprised of an advective term

$$\mathbf{J}_{\text{advective}} = \mathbf{v} Q, \quad (7.88)$$

and non-advective terms arising from diabatic processes and non-conservative accelerations (e.g., friction)

$$\mathbf{J}_{\text{non-advective}} = -\dot{b} \omega_a^{hy} + \nabla b \times \mathbf{F}. \quad (7.89)$$

### 7.6.3 Kinematic derivation of the potential vorticity flux

Following the discussion in Section 7.5.3, we consider a kinematic derivation of the potential vorticity flux vector for the hydrostatic and Boussinesq ocean. For that purpose, write the hydrostatic Boussinesq potential vorticity (7.82) in the form

$$Q = \omega_a^{hy} \cdot \nabla b = \nabla \cdot (b \omega_a^{hy}), \quad (7.90)$$

which follows since  $\nabla \cdot \omega_a^{hy} = 0$ . Taking the Eulerian time derivative then leads to

$$\partial_t Q = -\nabla \cdot \mathbf{J}^{\text{kin}} \quad (7.91)$$

where we defined the kinematic potential vorticity flux

$$\mathbf{J}^{\text{kin}} = -\partial_t(b \omega_a^{hy}). \quad (7.92)$$

The kinematic potential vorticity flux (7.92) is directly analogous to the potential vorticity flux (7.55) appearing in the non-hydrostatic and non-Boussinesq fluid.

#### Manifesting impermeability

The potential vorticity flux (7.92) manifests the impermeability property of Chapter 8 since

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{kin}} / Q = -\frac{1}{|\nabla b|} \frac{\partial b}{\partial t} = \hat{\mathbf{n}} \cdot \mathbf{v}_{b\perp}, \quad (7.93)$$

where  $\hat{\mathbf{n}} = \nabla b / |\nabla b|$  is the outward unit normal for a buoyancy surface, and with  $\mathbf{v}_{b\perp}$  the velocity of a point on the isopycnal that satisfies

$$(\partial_t + \mathbf{v}_{b\perp} \cdot \nabla) b = 0. \quad (7.94)$$

We return in Chapter 8 to help clarify this particular point about impermeability.

#### Relating $\mathbf{J}^Q$ and $\mathbf{J}^{\text{kin}}$

We relate the two potential vorticity fluxes,  $\mathbf{J}^Q$  and  $\mathbf{J}^{\text{kin}}$ , just like in Section 7.5.3 for the non-hydrostatic and non-Boussinesq fluid. Although the steps are exactly analogous, it is good practice to work through the details. For that purpose, start with the kinematic flux to write

$$\mathbf{J}^{\text{kin}} = -\partial_t(b \omega_a^{hy}) = -\partial_t b \omega_a^{hy} - b (\nabla \times \partial_t \mathbf{u}), \quad (7.95)$$

where we used

$$\partial_t \omega_a^{hy} = \partial_t \omega^{hy} = \nabla \times \partial_t \mathbf{u}. \quad (7.96)$$

Next use the identity

$$-b (\nabla \times \partial_t \mathbf{u}) = -\nabla \times (b \partial_t \mathbf{u}) + \nabla b \times \partial_t \mathbf{u}, \quad (7.97)$$

so that the kinematic flux from equation (7.95) becomes

$$\mathbf{J}^{\text{kin}} = -\partial_t b \omega_a^{hy} + \nabla b \times \partial_t \mathbf{u} - \nabla \times (b \partial_t \mathbf{u}). \quad (7.98)$$

Next recall the vector-invariant form of the hydrostatic and Boussinesq velocity equation (6.146)

$$\partial_t \mathbf{u} + \boldsymbol{\omega}_{\text{a}}^{\text{hy}} \times \mathbf{v} = -\nabla(\varphi + \mathbf{u}^2/2) + b \hat{\mathbf{z}} + \mathbf{F}, \quad (7.99)$$

This equation then leads to the cross product

$$\nabla b \times \partial_t \mathbf{u} = Q \mathbf{v} - (\dot{b} - \partial_t b) \boldsymbol{\omega}_{\text{a}}^{\text{hy}} - \nabla \times [b \nabla(\varphi + \mathbf{u} \cdot \mathbf{u}/2) - (b^2/2) \hat{\mathbf{z}}] + \nabla b \times \mathbf{F}, \quad (7.100)$$

where we used the identities

$$(\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \times \mathbf{v}) \times \nabla b = (\nabla b \cdot \boldsymbol{\omega}_{\text{a}}^{\text{hy}}) \mathbf{v} - (\nabla b \cdot \mathbf{v}) \boldsymbol{\omega}_{\text{a}}^{\text{hy}} = Q \mathbf{v} - (\dot{b} - \partial_t b) \boldsymbol{\omega}_{\text{a}}^{\text{hy}}. \quad (7.101)$$

Bringing everything together leads to the kinematic potential vorticity flux

$$\mathbf{J}^{\text{kin}} = -\partial_t b \boldsymbol{\omega}_{\text{a}}^{\text{hy}} + \nabla b \times \partial_t \mathbf{u} - \nabla \times (b \partial_t \mathbf{u}) \quad (7.102\text{a})$$

$$= Q \mathbf{v} - \dot{b} \boldsymbol{\omega}_{\text{a}}^{\text{hy}} + \nabla b \times \mathbf{F} + \nabla \times [-b \nabla(\varphi + \mathbf{u} \cdot \mathbf{u}/2) + b^2 \hat{\mathbf{z}} - b \partial_t \mathbf{u}] \quad (7.102\text{b})$$

$$= \mathbf{J}^{\text{Q}} + \nabla \times [-b \nabla(\varphi + \mathbf{u} \cdot \mathbf{u}/2) + b^2 \hat{\mathbf{z}} - b \partial_t \mathbf{u}]. \quad (7.102\text{c})$$

We have thus verified that  $\mathbf{J}^{\text{kin}}$  and  $\mathbf{J}^{\text{Q}}$  differ by the curl of a vector, so that their divergences are indeed equal.

#### 7.6.4 A potential vorticity flux vector suited to steady flows

*Schär (1993)* provided a generalization of Bernoulli's theorem for understanding steady geophysical flows, with *Marshall (2000)*, *Marshall et al. (2001)*, and *Polton and Marshall (2007)* applying this theorem to oceanic contexts within a hydrostatic and Boussinesq ocean.<sup>6</sup> We here derive their potential vorticity flux vector for the hydrostatic and Boussinesq ocean. The manipulations share similarities with those in Section 7.6.3, yet we present the details for further developing an appreciation of the various manipulations, which are a central facet of working with potential vorticity. The presentation here is a warm-up to the non-Boussinesq and non-hydrostatic case studied in Section 8.3.1, where we also provide example uses for the formulation.

##### Momentum equation

We start by exposing the Boussinesq form of the Bernoulli potential within the vector-invariant velocity equation. For this purpose, return to the horizontal momentum equation (6.146), and expand the expressions for perturbation pressure and the buoyancy

$$\partial_t \mathbf{u} + (\mathbf{f} + \boldsymbol{\omega}_{\text{a}}^{\text{hy}}) \times \mathbf{v} = -\nabla(\varphi + |\mathbf{u}|^2/2) + \hat{\mathbf{z}} b + \mathbf{F} \quad (7.103\text{a})$$

$$= -\nabla(|\mathbf{u}|^2/2) - \nabla(p - p_0)/\rho_{\text{o}} - \hat{\mathbf{z}} g (\rho - \rho_{\text{o}})/\rho_{\text{o}} + \mathbf{F} \quad (7.103\text{b})$$

$$= -\nabla(|\mathbf{u}|^2/2) - \nabla(p/\rho_{\text{o}}) - \hat{\mathbf{z}} g \rho/\rho_{\text{o}} + \mathbf{F} \quad (7.103\text{c})$$

$$= -\nabla(|\mathbf{u}|^2/2 + p/\rho_{\text{o}}) - \hat{\mathbf{z}} \left[ \frac{g \rho - g \rho_{\text{o}} + g \rho_{\text{o}}}{\rho_{\text{o}}} \right] + \mathbf{F} \quad (7.103\text{d})$$

$$= -\nabla(|\mathbf{u}|^2/2 + p/\rho_{\text{o}} + g z) - \hat{\mathbf{z}} g (\rho - \rho_{\text{o}})/\rho_{\text{o}} + \mathbf{F} \quad (7.103\text{e})$$

$$= -\nabla \mathcal{B} + \hat{\mathbf{z}} b + \mathbf{F}, \quad (7.103\text{f})$$

<sup>6</sup>We studied the Bernoulli potential and Bernoulli's theorem in VOLUME 2.

where we introduced the Bernoulli potential for a hydrostatic and Boussinesq fluid<sup>7</sup>

$$\mathcal{B} = |\mathbf{u}|^2/2 + p/\rho_0 + g z. \quad (7.104)$$

### Potential vorticity flux

The flux-form potential vorticity conservation statement remains as given by equation (7.84d), and the PV-substance flux is given by equation (7.86). However, we can make use of the gauge invariance of the potential vorticity flux to write it in a manner conducive to analyzing steady state conditions. For this purpose, operate with  $\nabla b \times$  on the velocity equation (7.103f) to have

$$\nabla b \times \partial_t \mathbf{u} + \nabla b \times (\boldsymbol{\omega}_a^{hy} \times \mathbf{v}) = -\nabla b \times \nabla \mathcal{B} + \nabla b \times \hat{z} b + \nabla b \times \mathbf{F}. \quad (7.105)$$

Now make use of the identity

$$\nabla b \times (\boldsymbol{\omega}_a^{hy} \times \mathbf{v}) = (\nabla b \cdot \mathbf{v}) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \mathbf{v} \quad (7.106)$$

in equation (7.105) to yield

$$(\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \mathbf{v} = -\nabla b \times \partial_t \mathbf{u} - \nabla b \times \nabla \mathcal{B} + \nabla b \times \hat{z} b + \nabla b \times \mathbf{F}. \quad (7.107)$$

We next make use of this identity for the purpose of manipulating the potential vorticity flux given by equation (7.86)

$$\mathbf{J}^Q = \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \times \mathbf{F} \quad (7.108a)$$

$$= \mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) - [\partial_t b + \mathbf{v} \cdot \nabla b] \boldsymbol{\omega}_a^{hy} + \nabla b \times \mathbf{F} \quad (7.108b)$$

$$= [\mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) - (\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a^{hy}] - \partial_t b \boldsymbol{\omega}_a^{hy} + \nabla b \times \mathbf{F} \quad (7.108c)$$

$$= [\nabla b \times \partial_t \mathbf{u} + \nabla b \times \nabla \mathcal{B} - \nabla b \times \hat{z} b - \nabla b \times \mathbf{F}] - \partial_t b \boldsymbol{\omega}_a^{hy} + \nabla b \times \mathbf{F} \quad (7.108d)$$

$$= \nabla b \times \partial_t \mathbf{u} - \partial_t b \boldsymbol{\omega}_a^{hy} + \nabla b \times \nabla \mathcal{B} - \nabla b \times \hat{z} b \quad (7.108e)$$

$$= \nabla b \times [\partial_t \mathbf{u} + \nabla \mathcal{B}] - \partial_t b \boldsymbol{\omega}_a^{hy} - \nabla \times (\hat{z} b^2/2). \quad (7.108f)$$

Dropping the rotational (gauge) term,  $-\nabla \times (\hat{z} b^2/2)$ , leads to the flux vector

$$\mathbf{J}^{ss} = \nabla b \times (\partial_t \mathbf{u} + \nabla \mathcal{B}) - \partial_t b \boldsymbol{\omega}_a^{hy} \quad (7.109)$$

We have now reached our goal whereby the steady state version of the potential vorticity flux (7.109) takes the rather elegant form

$$\mathbf{J}^{ss} = \nabla b \times \nabla \mathcal{B} = \nabla \times b \nabla \mathcal{B} = -\nabla \times \mathcal{B} \nabla b \quad \text{steady state.} \quad (7.110)$$

Hence, the steady state potential vorticity flux is aligned with the intersection of surfaces of constant buoyancy and Bernoulli potential

$$\nabla b \cdot \mathbf{J}^{ss} = 0 \quad \text{and} \quad \nabla \mathcal{B} \cdot \mathbf{J}^{ss} = 0 \quad \text{steady state.} \quad (7.111)$$

Recalling our discussion of a vector **streamfunction** in VOLUME 2, where here see that the buoyancy and Bernoulli potential serve as the two scalar functions that build the vector streamfunction for the steady state potential vorticity flux. Mapping surfaces of constant

<sup>7</sup>See VOLUME 2 for the non-Boussinesq Bernoulli potential.

buoyancy and constant Bernoulli potential, and determining their intersections, then determines the pathways for potential vorticity flux operating in a steady state. Figure 7.5 offers a schematic based on the analogous situation for a velocity streamfunction in a non-divergent flow, as studied in VOLUME 2. This result is the Boussinesq/hydrostatic form of the more general non-Boussinesq/non-hydrostatic case derived in Section 8.3.3.

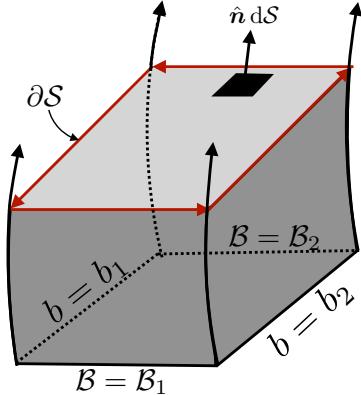


FIGURE 7.5: Isosurfaces of constant buoyancy,  $b$ , and Bernoulli potential,  $\mathcal{B}$ , serving as the two scalar functions building the vector streamfunction for potential vorticity flux,  $\mathbf{J}_{ss}^Q$ , in the steady state (see equation (7.111)). Streamlines are defined by the intersections of the  $b$  and  $\mathcal{B}$  isosurfaces, as shown by four streamlines along the corners of this particular volume. The transport of PV-substance through the surface,  $S$ , is determined by the line integral,  $\oint_{\partial S} \mathcal{B} db = -\oint_{\partial S} b d\mathcal{B} = (\mathcal{B}_1 - \mathcal{B}_2)(b_2 - b_1)$ , around the boundary circuit.

## 7.7 Potential vorticity over finite regions

In Section 7.3, we detailed why there is no materially invariant potential vorticity for a realistic equation of state for seawater, in which  $\rho = \rho(S, \Theta, p)$ . A similar limitation holds for the atmosphere, where the specific entropy is a function of pressure, density, and moisture concentration. We here build from our discussion of Kelvin's circulation theorem in Section 6.2 to consider a finite volume [potential vorticity pancake](#). Rather than being a property carried by each fluid particle, the pancake potential vorticity is carried by a finite fluid region and it is materially invariant for perfect fluid flows, even for a realistic equation of state. The discussion here follows that given by [Kooloth et al. \(2022\)](#).

### 7.7.1 Differential relations

The starting point is the potential vorticity equation written as in equation (7.41)

$$\rho \frac{DQ}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{\chi} + \nabla \chi \cdot [\nabla \times (-\rho^{-1} \nabla p)] + \nabla \chi \cdot (\nabla \times \mathbf{F}) \quad \text{with} \quad Q = \rho^{-1} \nabla \chi \cdot \boldsymbol{\omega}_a. \quad (7.112)$$

The key step in the derivation is to replace the pressure gradient acceleration with its equivalent in terms of specific enthalpy and its partial derivatives.<sup>8</sup> For this purpose, write the specific enthalpy as a function of Conservative Temperature, salinity, and pressure

$$\mathcal{H} = \mathcal{H}(\Theta, S, p), \quad (7.113)$$

<sup>8</sup>We made use of a similar approach when studying circulation in Section 6.2.5.

which leads to the spatial gradient<sup>9</sup>

$$\nabla \mathcal{H} = \rho^{-1} \nabla p + \left[ \frac{\partial \mathcal{H}}{\partial \Theta} \right]_{p,S} \nabla \Theta + \left[ \frac{\partial \mathcal{H}}{\partial S} \right]_{\Theta,p} \nabla S, \quad (7.114)$$

which then leads to the expression for the pressure gradient acceleration

$$-\rho^{-1} \nabla p = -\nabla \mathcal{H} + \mathcal{H}_\Theta \nabla \Theta + \mathcal{H}_S \nabla S, \quad (7.115)$$

where we introduced the shorthand notation

$$\mathcal{H}_\Theta = \left[ \frac{\partial \mathcal{H}}{\partial \Theta} \right]_{p,S} \quad \text{and} \quad \mathcal{H}_S = \left[ \frac{\partial \mathcal{H}}{\partial S} \right]_{\Theta,p}. \quad (7.116)$$

Use of the identity (7.115) within the potential vorticity equation (7.112) leads to

$$\rho \frac{DQ}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{\chi} + \nabla \chi \cdot \nabla \times [\mathcal{H}_\Theta \nabla \Theta + \mathcal{H}_S \nabla S + \mathbf{F}], \quad (7.117)$$

where we set  $\nabla \times \nabla \mathcal{H} = 0$ . Now make use of the following vector identity, with  $\mathbf{C}$  an arbitrary vector,

$$\nabla \chi \cdot (\nabla \times \mathbf{C}) = \nabla \cdot (\chi \nabla \times \mathbf{C}) = \nabla \cdot [\nabla \times (\chi \mathbf{C}) - \nabla \chi \times \mathbf{C}] = -\nabla \cdot (\nabla \chi \times \mathbf{C}), \quad (7.118)$$

to bring the potential vorticity equation (7.117) to

$$\rho \frac{DQ}{Dt} = \nabla \cdot [\boldsymbol{\omega}_a \dot{\chi} - \nabla \chi \times (\mathcal{H}_\Theta \nabla \Theta + \mathcal{H}_S \nabla S + \mathbf{F})]. \quad (7.119)$$

## 7.7.2 Integral relations

Equation (7.119) is quite general. To reduce by one the terms on the right hand side, we can either choose  $\chi = \Theta$  or  $\chi = S$ . Let us choose  $\chi = \Theta$ , in which case

$$\rho \frac{DQ^{(\Theta)}}{Dt} = \nabla \cdot [\mathcal{H}_S \nabla S \times \nabla \Theta + \mathbf{F} \times \nabla \Theta + \boldsymbol{\omega}_a \dot{\Theta}] \quad \text{with} \quad Q^{(\Theta)} = \rho^{-1} \nabla \Theta \cdot \boldsymbol{\omega}_a. \quad (7.120)$$

We still have the reversible term,  $\mathcal{H}_S \nabla S \times \nabla \Theta$ , contributing to the material time evolution of  $Q^{(\Theta)}$ , and this term vanishes only for particularly idealized forms of the fluid thermodynamics. Rather than pursue material invariance for potential vorticity along a fluid particle, consider an integration of  $Q^{(\Theta)}$  over a finite region that moves with the flow, and take its time derivative

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \rho Q^{(\Theta)} dV = \int_{\mathcal{R}(\mathbf{v})} \rho \frac{DQ^{(\Theta)}}{Dt} dV \quad (7.121a)$$

$$= \int_{\mathcal{R}(\mathbf{v})} \nabla \cdot [\mathcal{H}_S \nabla S \times \nabla \Theta + \mathbf{F} \times \nabla \Theta + \boldsymbol{\omega}_a \dot{\Theta}] dV \quad (7.121b)$$

$$= \oint_{\partial \mathcal{R}(\mathbf{v})} [\mathcal{H}_S \nabla S \times \nabla \Theta + \mathbf{F} \times \nabla \Theta + \boldsymbol{\omega}_a \dot{\Theta}] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (7.121c)$$

The first and second terms on the right hand side vanish if we choose a volume whose boundary has an outward unit normal vector,  $\hat{\mathbf{n}}$ , parallel to  $\nabla \Theta$  at every point around the closed region.

---

<sup>9</sup>We performed a similar manipulation in our study of energetics in VOLUME 2.

This sort of  $\Theta$ -bubble is uncommon in a stably stratified fluid, though it may occur over relatively small scales in turbulent flows.

For another arrangement, consider the case of salinity with closed contours on Conservative Temperature surfaces. We thus take a pancake/disk region with  $\hat{\mathbf{n}}$  parallel to  $\nabla\Theta$  on its top and bottom and  $\hat{\mathbf{n}}$  parallel to  $\nabla S$  on the sides.<sup>10</sup> For this region, the  $\nabla S \times \nabla\Theta$  contribution vanishes along all the boundaries, and the contribution from  $\mathbf{F}$  only appears on the sides where  $\hat{\mathbf{n}}$  is parallel to  $\nabla S$ . We can take a complementary approach by setting  $\chi = S$ , in which case the previous discussion holds yet with  $\Theta$  and  $S$  interchanged. Either of these pancake regions preserves the materially integrated potential vorticity for inviscid flows with  $S$  and  $\Theta$  both materially invariant.



## 7.8 Exercises

### EXERCISE 7.1: POTENTIAL VORTICITY FOR A PERFECT NON-HYDROSTATIC BOUSSINESQ OCEAN

Consider a perfect tangent plane Boussinesq ocean whose governing equations are given by

$$\frac{D\mathbf{v}}{Dt} + f(\hat{\mathbf{z}} \times \mathbf{v}) = -\nabla\varphi + b\hat{\mathbf{z}} \quad (7.122)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (7.123)$$

$$\frac{Db}{Dt} = 0 \quad (7.124)$$

$$b = -\frac{g(\rho - \rho_0)}{\rho_0} = g\alpha\theta, \quad (7.125)$$

where  $\theta$  is the potential temperature and  $\alpha > 0$  is a constant thermal expansion coefficient, and we assumed a linear equation of state for density whereby

$$\rho = \rho_0(1 - \alpha\theta). \quad (7.126)$$

Further details are provided in our study of the [Boussinesq ocean](#) in [VOLUME 2](#). Some of this exercise follows the hydrostatic Boussinesq discussion in Section 7.6, though they differ in important places, so be careful!

- (a) Derive the equation for the material time evolution of potential vorticity in this fluid system. Show all steps in the derivation.
- (b) The baroclinicity vector appearing in the Boussinesq vorticity equation is  $\mathbf{B} = \nabla \times \hat{\mathbf{z}} b = \nabla b \times \hat{\mathbf{z}}$  (see Section 6.7.1 to check the derivation in part (a) above). Show that this vector results upon making the Boussinesq approximation to the non-Boussinesq expression  $\mathbf{B} = (\nabla\rho \times \nabla p)/\rho^2$ . Hint: drop the  $\delta\rho$  and  $\delta p$  product given that it is a higher order quantity.
- (c) Show that the vertical portion of  $Q^{\text{bouss}}$  can be written

$$Q_{\text{vert}}^{\text{bouss}} = (\zeta + f)N^2 \quad (7.127)$$

where  $\zeta = \partial_x v - \partial_y u$  is the vertical component to the relative vorticity and  $N^2 = \partial b/\partial z$

---

<sup>10</sup>The geometry is similar to that shown in Figure 7.5 when discussing the steady state potential vorticity flux.

is the squared buoyancy frequency. Hint: there is no trick.

- (d) If flow maintains hydrostatic and geostrophic balance, show that the horizontal portion of  $Q^{\text{bouss}}$  can be written

$$Q_{\text{horz}}^{\text{bouss}} = \boldsymbol{\omega} \cdot \nabla_h b \approx -f^{-1} |\nabla_h b|^2. \quad (7.128)$$

Hint: recall that for hydrostatic and geostrophic flow, the vertical velocity is much smaller than horizontal.

#### EXERCISE 7.2: POTENTIAL VORTICITY FOR DIABATIC AND FRICTIONAL NON-HYDROSTATIC BOUSSINESQ OCEAN

Reconsider Exercise 7.1 in the presence of irreversible forces (e.g., friction) and buoyancy sources so that the governing tangent plane equations are

$$\frac{D\mathbf{v}}{Dt} + f(\hat{\mathbf{z}} \times \mathbf{v}) = -\nabla\varphi + b\hat{\mathbf{z}} + \mathbf{F} \quad (7.129)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (7.130)$$

$$\frac{Db}{Dt} = \dot{b} \quad (7.131)$$

$$b = -\frac{g(\rho - \rho_0)}{\rho_0} = g\alpha\theta. \quad (7.132)$$

In these equations we assumed a linear equation of state for density whereby

$$\rho = \rho_0(1 - \alpha\theta), \quad (7.133)$$

with  $\alpha > 0$  a constant thermal expansion coefficient. Hence, material time changes to the buoyancy are given by

$$\frac{Db}{Dt} = \dot{b} = g\alpha\dot{\theta}, \quad (7.134)$$

where  $\dot{\theta}$  is a diabatic heating source/sink. We also included  $\mathbf{F}$  to the velocity equation (7.129), which represents a non-conservative acceleration such as from friction or boundary stresses.

- (a) Derive the equation for the material time evolution of potential vorticity in this fluid system, including the irreversible contributions from non-conservative accelerations and from heating.
- (b) Derive an equation for the potential vorticity time tendency (i.e., Eulerian time derivative), written in the form

$$\partial_t Q = -\nabla \cdot \mathbf{J}^Q. \quad (7.135)$$

What is the potential vorticity flux  $\mathbf{J}^Q$ ? Note that your answer is unique up to the curl of an arbitrary vector (gauge symmetry). Also note that for a Boussinesq flow we drop the constant reference density in the definition of  $\mathbf{J}^Q$ .

- (c) A common diabatic process is written in the form of a damping source

$$\dot{b} = -\mu(b - b^*), \quad (7.136)$$

where  $\mu$  is a constant Newtonian damping coefficient (dimensions of inverse time), and  $b^*$  is a specified buoyancy profile. This form of a buoyancy source acts to damp the buoyancy towards a specified profile  $b^*$ . Show that Newtonian damping of buoyancy corresponds to potential vorticity damping towards  $Q^* = \boldsymbol{\omega}_a \cdot \nabla b^*$ .

- (d) A form for the friction operator is given by Rayleigh drag

$$\mathbf{F} = -\gamma \mathbf{v}, \quad (7.137)$$

with  $\gamma$  a constant Rayleigh damping coefficient with dimension of inverse time. Show that Rayleigh drag in the momentum equation, which acts to damp velocity towards zero, corresponds to a damping of potential vorticity towards its planetary geostrophic form,  $Q^{\text{pg}} = f N^2$ , where  $N^2 = \partial b / \partial z$  is the squared buoyancy frequency.

- (e) Discuss the balance needed between forcing terms in  $\mathbf{J}^Q$  to arrive at a steady state (i.e., zero Eulerian time tendency). Continue to assume the friction is in the form of Rayleigh drag and heating is in the form of Newtonian damping.



## POTENTIAL VORTICITY BUDGETS

In Chapter 7 we studied properties of the Ertel potential vorticity,  $Q = \rho^{-1} \boldsymbol{\omega}_a \cdot \nabla \chi$ , which is the specific volume times the projection of the absolute vorticity,  $\boldsymbol{\omega}_a$ , onto the gradient of a scalar field,  $\nabla \chi$ . For a perfect fluid (i.e., no friction, no diabatic heating, nor diffusive mixing) with  $\nabla \chi$  orthogonal to the baroclinicity,  $\mathbf{B} = \rho^{-2} \nabla \rho \times \nabla p$ , and if  $\chi$  is itself materially invariant ( $D\chi/Dt = 0$ ), then  $Q$  is a material invariant:  $DQ/Dt = 0$ , meaning that  $Q$  remains constant when following a fluid particle. In geophysical fluid mechanics, it is common to define potential vorticity with  $\chi$  equal to the specific entropy (for the atmosphere) or Archimedean buoyancy (for the ocean). Specific entropy and buoyancy are thermodynamic tracers that are a function of the fluid state (e.g., temperature, pressure, salinity), whereas potential vorticity is a function of the flow (via the absolute vorticity). It is remarkable that a materially invariant property of the fluid can, if it annihilates baroclinicity, generate a materially invariant property of the flow.

In this chapter we develop finite volume budgets for potential vorticity substance,  $\int_{\mathcal{R}} Q \rho dV$ , with such budgets fundamentally affected by the impermeability theorem respected by the potential vorticity flux,  $\mathbf{J}^Q$ . Impermeability says that there is identically zero potential vorticity flux crossing  $\chi$ -isosurfaces, with this property holding even when there are mass and thermal fluxes crossing  $\chi$ -isosurfaces. Such generality signals the kinematic nature of impermeability, which ultimately follows from the non-divergent nature of vorticity. It offers further insight into why, as developed in Section 8.4, the volume integrated potential vorticity substance,  $\int_{\mathcal{R}} Q \rho dV$ , changes only when  $\chi$  surfaces intersect a boundary.

Impermeability and material invariance are two aspects of potential vorticity that are easily confused. Here we summarize their basic elements that are explored within this chapter.

- The potential vorticity flux,  $\mathbf{J}^Q$ , does not cross surfaces of constant  $\chi$ , and this property is referred to as impermeability or the impermeability theorem. Impermeability holds for all flows, even in the presence of irreversible processes and with  $D\chi/Dt \neq 0$ . Hence, any closed volume,  $\mathcal{R}_\chi$ , moving with velocity,  $\mathbf{v}_Q = \mathbf{J}^Q/(\rho Q)$ , maintains a constant  $\int_{\mathcal{R}_\chi} \rho Q dV$ , so that the addition or removal of mass to the region only dilutes or concentrates the potential vorticity substance. This finite domain conservation of potential vorticity substance is a kinematic property that follows from the non-divergent nature of absolute vorticity, or, equivalently, since  $\rho Q = \boldsymbol{\omega}_a \cdot \nabla \chi = \nabla \cdot (\boldsymbol{\omega}_a \chi)$  equals to a divergence.
- Material invariance of potential vorticity,  $DQ/Dt = 0$ , holds for a perfect fluid and for  $Q$  that is defined according to a materially constant scalar field,  $\chi$ , that annihilates baroclinicity via  $\rho^{-2} (\nabla \rho \times \nabla p) \cdot \nabla \chi = 0$ . Material conservation of potential vorticity is a local property holding for each fluid particle, and as such it is a far more restrictive property than impermeability and the associated finite volume conservation of potential vorticity substance.

## CHAPTER GUIDE

The goal of this chapter is to fill in the conceptual and technical details needed to understand the above two bullet points about finite volume and local conservation properties of potential vorticity. We build from the potential vorticity mechanics introduced in Chapter 7, and make use of vector calculus from VOLUME 1. This chapter is an essential read for those interested in potential vorticity theory and potential vorticity budgets. The primary references for this chapter are *Haynes and McIntyre* (1987), *Haynes and McIntyre* (1990), and *Müller* (1995).

<b>8.1</b>	<b>Loose threads</b>	<b>290</b>
<b>8.2</b>	<b>Variations on the impermeability theorem</b>	<b>291</b>
8.2.1	Impermeability for the Haynes-McIntyre potential vorticity flux	291
8.2.2	A kinematic derivation of impermeability	292
8.2.3	Comments	294
<b>8.3</b>	<b>Impermeability theorem for seawater</b>	<b>294</b>
8.3.1	Ocean potential vorticity in terms of potential density	294
8.3.2	A modified PV-substance flux	295
8.3.3	Integral constraints for steady state	296
8.3.4	Further study	297
<b>8.4</b>	<b>Integrated potential vorticity substance</b>	<b>297</b>
8.4.1	The primary role of boundaries	298
8.4.2	Region bounded by a single buoyancy surface	299
8.4.3	Region bounded by two buoyancy surfaces	300
8.4.4	Region bounded by land and a buoyancy surface	300
8.4.5	A layer outcropping at the ocean surface	302
8.4.6	Further study	302
<b>8.5</b>	<b>Boundary fluxes of PV-substance</b>	<b>302</b>
8.5.1	Layer integrated budget	304
8.5.2	Impermeability across interior layer interfaces	304
8.5.3	Potential vorticity flux at a land-sea boundary	305
8.5.4	Potential vorticity flux at the air-sea boundary	306
8.5.5	Thought experiments	307
8.5.6	Is there a preferred form of the PV-substance flux?	308
8.5.7	Further study	309

## 8.1 Loose threads

- Make a table with PV and PV fluxes. They can get rather confusing.
- *Callies and Ferrari* (2018) and bottom mixing induced circulation. Connect the thermal wind next to the bottom, inducing flow counter to Kelvin waves, to the potential vorticity boundary fluxes in Section 8.5.3. Discuss the spin up experiment in Section 4 of *Callies and Ferrari* (2018) from a potential vorticity perspective. Note the role of bottom friction in enabling the bottom buoyancy mixing to impart potential vorticity to the flow. without friction then there would be no way to introduce potential vorticity to the flow. That then couples the buoyancy mixing to the friction. *Callies and Ferrari* (2018) also note that to satisfy  $\hat{\mathbf{n}} \cdot \nabla b = 0$  requires friction. I do not understand that fully, but must be

related to this potential vorticity argument.

- Exercises

## 8.2 Variations on the impermeability theorem

In this section we derive the impermeability theorem satisfied by the potential vorticity flux vector. We illustrate the theorem for a variety of potential vorticity flux vectors that differ by a gauge function. So although these fluxes have identical divergences, the physical content of the fluxes is distinct. Consequently, one may choose to use a particular flux depending on the context of their use. We return to this point in Section 8.5.6.

### 8.2.1 Impermeability for the Haynes-McIntyre potential vorticity flux

We start with the flux-form evolution equation for PV-substance given by equation (7.50)

$$\partial_t(\rho Q) + \nabla \cdot \mathbf{J}^Q = 0 \quad \text{with} \quad \mathbf{J}^Q = \rho \mathbf{v} Q - \dot{\chi} \boldsymbol{\omega}_a + \nabla \chi \times \mathbf{F}, \quad (8.1)$$

with  $\dot{\chi} = D\chi/Dt$ , and with the PV-substance flux vector,  $\mathbf{J}^Q$ , given in the form examined in [Haynes and McIntyre \(1987\)](#).<sup>1</sup> Following the derivation in [Haynes and McIntyre \(1987\)](#), we decompose the velocity into two components, one oriented parallel to constant  $\chi$  surfaces and one oriented perpendicular

$$\mathbf{v}_{\parallel} = \mathbf{v} - \hat{\mathbf{n}} (\mathbf{v} \cdot \hat{\mathbf{n}}) \quad \text{and} \quad \mathbf{v}_{\perp} = -\frac{\hat{\mathbf{n}} \partial \chi / \partial t}{|\nabla \chi|} \quad \Rightarrow \mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} + \frac{\hat{\mathbf{n}} \dot{\chi}}{|\nabla \chi|} \quad (8.2)$$

where

$$\hat{\mathbf{n}} = \nabla \chi / |\nabla \chi| \quad (8.3)$$

is the unit normal vector for  $\chi$ -isosurfaces. By construction, the velocity  $\mathbf{v}_{\perp}$  satisfies

$$(\partial_t + \mathbf{v}_{\perp} \cdot \nabla) \chi = 0. \quad (8.4)$$

Hence, according to the kinematics detailed in our study of the [kinematic boundary condition](#) in [VOLUME 2](#),

$$\mathbf{v}_{\perp} \cdot \hat{\mathbf{n}} = \mathbf{v}_{\chi} \cdot \hat{\mathbf{n}}, \quad (8.5)$$

where  $\mathbf{v}_{\chi}$  is the velocity of a point fixed on a constant  $\chi$  surface. That is,  $\mathbf{v}_{\perp}$  provides a measure of the velocity for a point following a constant  $\chi$  surface, even as that surface moves through the fluid. We make use of this key identity below.

With the velocity decomposition (8.2), the PV-substance flux vector takes the form

$$\mathbf{J}^Q = \rho \mathbf{v} Q - \dot{\chi} \boldsymbol{\omega}_a + \nabla \chi \times \mathbf{F} \quad (8.6a)$$

$$= \left[ \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} + \frac{\dot{\chi} \nabla \chi}{|\nabla \chi|^2} \right] \rho Q - \dot{\chi} \boldsymbol{\omega}_a + \nabla \chi \times \mathbf{F} \quad (8.6b)$$

$$= (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) \rho Q - \dot{\chi} [\boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}}] + \nabla \chi \times \mathbf{F} \quad (8.6c)$$

$$= (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) \rho Q - \dot{\chi} (\boldsymbol{\omega}_a)_{\parallel} + \nabla \chi \times \mathbf{F} \quad (8.6d)$$

<sup>1</sup>Following [Haynes and McIntyre \(1987\)](#), we set the gauge function,  $\mathbf{A}$ , to zero in equation (8.1).

$$= \underbrace{\mathbf{v}_\perp \rho Q}_{\mathbf{J}_\perp} + \underbrace{\left[ \rho Q \mathbf{v}_\parallel - \dot{\chi} (\boldsymbol{\omega}_a)_\parallel \right]}_{\mathbf{J}_\parallel} + \nabla \chi \times \mathbf{F} \quad (8.6e)$$

$$\equiv \mathbf{J}_\perp + \mathbf{J}_\parallel, \quad (8.6f)$$

where

$$(\boldsymbol{\omega}_a)_\parallel = \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \boldsymbol{\omega}_a - \left[ \frac{\boldsymbol{\omega}_a \cdot \nabla \chi}{|\nabla \chi|^2} \right] \nabla \chi = \boldsymbol{\omega}_a - \frac{\rho Q}{|\nabla \chi|} \hat{\mathbf{n}}. \quad (8.7)$$

The above results motivate us to write the PV-substance budget equation (8.1) in the form

$$\partial_t (\rho Q) + \nabla \cdot (\mathbf{v}_Q \rho Q) = 0, \quad (8.8)$$

where

$$\mathbf{v}_Q \equiv \frac{\mathbf{J}^Q}{\rho Q} \quad \text{definition of } \mathbf{v}_Q \quad (8.9a)$$

$$= \mathbf{v} + \frac{-\dot{\chi} \boldsymbol{\omega}_a + \nabla \chi \times \mathbf{F}}{\rho Q} \quad \text{equation (8.1)} \quad (8.9b)$$

$$= \mathbf{v}_\perp + \mathbf{v}_\parallel + \frac{-\dot{\chi} (\boldsymbol{\omega}_a)_\parallel + \nabla \chi \times \mathbf{F}}{\rho Q} \quad \text{equation (8.6e),} \quad (8.9c)$$

so that  $\mathbf{v}_Q$  is the velocity that advects the PV-substance through the fluid. A direct calculation shows that  $\mathbf{v}_Q$  satisfies the following property

$$\mathbf{v}_Q \cdot \hat{\mathbf{n}} = \mathbf{v}_\perp \cdot \hat{\mathbf{n}} = \mathbf{v}_\chi \cdot \hat{\mathbf{n}}, \quad (8.10)$$

where the final equality made use of the identity (8.5). As a result,  $\mathbf{v}_Q$  has a normal component that is identical to that of the velocity of a point fixed on the  $\chi$  surface

$$(\partial_t + \mathbf{v}_Q \cdot \nabla) \chi = 0. \quad (8.11)$$

We depict this result in Figure 8.1, whereby the PV-substance flux never crosses the  $\chi$ -isosurface, even as the surface moves and even in the presence of processes that allow for matter and thermal properties to cross the surface. This result holds since the  $\chi$ -isosurface moves in a way to precisely track the PV-substance flux. In general,  $\chi$  surfaces are permeable to matter and thermal properties but, as we have just shown, are impermeable to PV-substance. This is a rather remarkable kinematic result that has important implications for budgets of PV-substance within regions bounded by constant  $\chi$  surfaces.

### 8.2.2 A kinematic derivation of impermeability

The derivation of impermeability in Section 8.2.1 follows that given by [Haynes and McIntyre \(1987\)](#). We now complement that derivation by an alternative that emphasizes the kinematic origins of impermeability. For that purpose, we make no use of the vorticity equation or the equation for  $\chi$ . Instead, we merely use the definition of potential vorticity and the non-divergence property of absolute vorticity. This derivation follows our discussion of gauge freedom introduced in Section 7.5.3, as well as the discussion of potential vorticity for a hydrostatic Boussinesq ocean in Section 7.6.

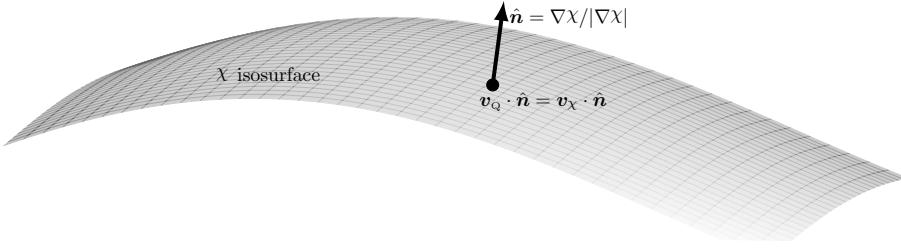


FIGURE 8.1: The flux,  $\mathbf{J}^Q$ , of PV-substance,  $\rho Q = \nabla \cdot (\chi \boldsymbol{\omega}_a)$ , does not penetrate a surface of constant  $\chi$ . This kinematic result follows since the effective velocity of PV-substance,  $\mathbf{v}_Q = \mathbf{J}^Q/(\rho Q)$ , has the same normal component as a point fixed on a  $\chi$  surface,  $\mathbf{v}_Q \cdot \hat{\mathbf{n}} = \mathbf{v}_\chi \cdot \hat{\mathbf{n}}$ . Consequently, the  $\chi$  surface moves in a manner so that no flux of PV-substance crosses the surface, even in the presence of irreversible processes. This result is known as the impermeability theorem since  $\chi$  surfaces are impermeable to the flux of PV-substance, even though they are permeable to matter and thermal properties.

The key identity we need was already given by equation (7.11)

$$\rho Q = \boldsymbol{\omega}_a \cdot \nabla \chi = \nabla \cdot (\boldsymbol{\omega}_a \chi), \quad (8.12)$$

thus revealing that  $\rho Q$  is a pure divergence. Taking the Eulerian time derivative then leads to

$$\partial_t(\rho Q) = -\nabla \cdot \mathbf{J}^{\text{kin}}, \quad (8.13)$$

where

$$\mathbf{J}^{\text{kin}} = -\partial_t(\boldsymbol{\omega}_a \chi) \quad (8.14)$$

is the kinematic form of the PV-substance flux. By construction, this flux vanishes in the steady state

$$\mathbf{J}^{\text{kin}} = 0 \quad \text{in steady state,} \quad (8.15)$$

which certainly contrasts to the steady state Haynes-McIntyre flux given by equation (8.1).

Introducing the velocity seen from an inertial reference frame (also called the absolute velocity in VOLUME 1 and VOLUME 2)

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x} \quad (8.16)$$

leads to

$$-(\partial_t \boldsymbol{\omega}_a) \chi = -\partial_t(\nabla \times \mathbf{v}_a) \chi = -(\nabla \times \partial_t \mathbf{v}_a) \chi = -\nabla \times (\partial_t \mathbf{v}_a \chi) + \nabla \chi \times \partial_t \mathbf{v}_a. \quad (8.17)$$

Dropping the total curl (which amounts to choosing a gauge function) yields the modified kinematic expression for the flux of PV-substance

$$\tilde{\mathbf{J}}^{\text{kin}} = -\partial_t \mathbf{v}_a \times \nabla \chi - \boldsymbol{\omega}_a \partial_t \chi. \quad (8.18)$$

It follows that

$$\tilde{\mathbf{v}}^{\text{kin}} \cdot \hat{\mathbf{n}} = \frac{\tilde{\mathbf{J}}^{\text{kin}} \cdot \hat{\mathbf{n}}}{\rho Q} = -\frac{\omega_a \cdot \nabla \chi}{\rho Q} \frac{\partial \chi}{\partial t} \frac{1}{|\nabla \chi|} = -\frac{\partial \chi}{\partial t} \frac{1}{|\nabla \chi|} = \mathbf{v}_\chi \cdot \hat{\mathbf{n}}, \quad (8.19)$$

which is the same result (8.10) as found for the Haynes-McIntyre flux. This result allows us to conclude that  $\tilde{\mathbf{J}}^{\text{kin}}$  satisfies the impermeability theorem. We again emphasize that there has been no use of the dynamical equations for vorticity or for  $\chi$ . Instead, this expression of impermeability only used the definition of potential vorticity, along with the non-divergent nature of vorticity,  $\nabla \cdot \omega_a = 0$ .

Throughout this discussion, we assumed  $\chi$  to be an arbitrary smooth scalar field. Hence, any scalar field used to project out a component of the absolute vorticity has its iso-surfaces impenetrable to the flux of the corresponding component of absolute vorticity. This result trivializes the impermeability theorem from a mathematical perspective. In Section 6.3.4 we somewhat anticipated this result when studying the Cartesian components of the absolute vorticity (see also Section 5 of [Haynes and McIntyre \(1987\)](#)). This trivial mathematical result does not reduce the importance of the entropic potential vorticity impermeability theorem for studying stratified fluid flows. The importance holds since this particular potential vorticity has direct connection to dynamics and thermodynamics.

### 8.2.3 Comments

The impermeability theorem was introduced by [Haynes and McIntyre \(1987\)](#). Their paper was met by some confusion thus prompting them to write [Haynes and McIntyre \(1990\)](#). Besides exposing the purely kinematic aspects of impermeability, the presentation in this section reveals that there are multiple potential vorticity flux vectors that satisfy impermeability, with the vectors differing by a gauge transformation. Which flux vector is preferred depends on the application, with [Bretherton and Schär \(1993\)](#), [Davies-Jones \(2003\)](#), and [Marshall et al. \(2001\)](#) proposing criteria favoring one form over another. We pursue such considerations in Section 8.3.

## 8.3 Impermeability theorem for seawater

As seen from Section 8.2.2, impermeability holds for any component of vorticity and the corresponding scalar isosurface. In contrast, material invariance of potential vorticity requires a materially conserved scalar to annihilate the baroclinicity vector (e.g., Section 7.2.4). Consequently, material invariance is much tougher to satisfy than impermeability. Indeed, as shown in Section 7.3, there is no materially invariant potential vorticity for an ocean with a realistic nonlinear equation of state (EOS). Hence, there is no materially invariant potential vorticity for the ocean even in the absence of irreversible processes. Nevertheless, one can define an ocean potential vorticity according to any scalar field, such as [potential density](#), and still make use of the impermeability theorem when performing a potential vorticity budget. We here expose the details.

### 8.3.1 Ocean potential vorticity in terms of potential density

Following [Marshall et al. \(2001\)](#), we introduce an ocean potential vorticity field according to

$$Q^{\text{ocn}} = \rho^{-1} \nabla b \cdot \omega_a, \quad (8.20)$$

where the **Archimedean buoyancy** field,  $b$ , is approximated by a chosen potential density. As shown in Section 7.3, a globally defined buoyancy does not annihilate the baroclinicity vector for a realistic seawater equation of state

$$\mathbf{B} \cdot \nabla b = [-\nabla(1/\rho) \times \nabla p] \cdot \nabla b \neq 0. \quad (8.21)$$

Consequently,  $DQ^{\text{ocn}}/Dt \neq 0$  even in the absence of irreversible processes. Nonetheless, the Eulerian budget for PV-substance satisfies

$$\partial_t(\rho Q^{\text{ocn}}) = -\nabla \cdot \tilde{\mathbf{J}}^{\text{Q-ocn}}, \quad (8.22)$$

and  $\tilde{\mathbf{J}}^{\text{Q-ocn}}$  satisfies the impermeability theorem for  $b$ -surfaces. A flux-form budget equation greatly facilitates the study of budgets for PV-substance even within an ocean with a realistic equation of state. Derivation of the flux-form equation (8.22) follows from the discussion in Section 8.2.2, where we know that the kinematic flux

$$\tilde{\mathbf{J}}^{\text{Q-ocn}} = -\partial_t \mathbf{v}_a \times \nabla b - \boldsymbol{\omega}_a \partial_t b = -\partial_t \mathbf{v} \times \nabla b - \boldsymbol{\omega}_a \partial_t b \quad (8.23)$$

satisfies the impermeability theorem for  $b$ -surfaces and whose convergence drives the time tendency for the PV-substance. Note that the second equality in equation (8.23) follows since

$$\partial_t \mathbf{v}_a = \partial_t(\mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}) = \partial_t \mathbf{v}, \quad (8.24)$$

given that the Eulerian time derivative is computed at a fixed position,  $\mathbf{x}$ , and the planetary rotation is assumed constant,  $\partial_t \boldsymbol{\Omega} = 0$ .

### 8.3.2 A modified PV-substance flux

The kinematic flux (8.23) vanishes in the steady state. We here motivate a gauge transformed flux that leads to the same flux divergence yet that renders a nonzero steady state flux. For this purpose, make use of the vector-invariant velocity equation (equation (6.33))

$$\partial_t \mathbf{v} + \boldsymbol{\omega}_a \times \mathbf{v} = -\rho^{-1} \nabla p - \nabla \mathcal{M} + \mathbf{F}, \quad (8.25)$$

where

$$\mathcal{M} = \mathbf{v} \cdot \mathbf{v}/2 + \Phi \quad (8.26)$$

is the mechanical energy per mass of a fluid element. Bringing the pressure term inside of the gradient operator leads to

$$\partial_t \mathbf{v} + \boldsymbol{\omega}_a \times \mathbf{v} = p \nabla(1/\rho) - \nabla(\mathcal{M} + p/\rho) + \mathbf{F}. \quad (8.27)$$

Following our treatment of the hydrostatic Boussinesq ocean in Section 7.6.4, we introduce the Bernoulli function<sup>2</sup>

$$B = \mathcal{M} + p/\rho. \quad (8.28)$$

<sup>2</sup>The **Bernoulli potential**,  $\mathcal{B}$ , arises from an analysis of the total energy budget VOLUME 2, where we find that the Bernoulli potential in a compressible (non-Boussinesq) fluid,  $\mathcal{B} = \mathcal{M} + p/\rho + \mathcal{I}$ , also includes the internal energy per mass,  $\mathcal{I}$ . However, the internal energy is missing from equation (8.28), thus motivating our use of the terminology “a Bernoulli function” rather than “the Bernoulli potential”.

The vector-invariant velocity equation (8.27) thus leads to the cross-product

$$\partial_t \mathbf{v} \times \nabla b = -(\boldsymbol{\omega}_a \times \mathbf{v}) \times \nabla b + [p \nabla(1/\rho) - \nabla(\mathcal{M} + p/\rho) + \mathbf{F}] \times \nabla b \quad (8.29a)$$

$$= -(\nabla b \cdot \boldsymbol{\omega}_a) \mathbf{v} + (\nabla b \cdot \mathbf{v}) \boldsymbol{\omega}_a + [p \nabla(1/\rho) - \nabla(\mathcal{M} + p/\rho) + \mathbf{F}] \times \nabla b \quad (8.29b)$$

$$= -\mathbf{v} \rho Q^{ocn} + (\dot{b} - \partial_t b) \boldsymbol{\omega}_a + [p \nabla(1/\rho) - \nabla(\mathcal{M} + p/\rho) + \mathbf{F}] \times \nabla b. \quad (8.29c)$$

Use of this result leads to the flux (8.23)

$$\tilde{\mathbf{J}}^{Q-ocn} = -\partial_t \mathbf{v} \times \nabla b - \boldsymbol{\omega}_a \partial_t b \quad (8.30a)$$

$$= \mathbf{v} \rho Q^{ocn} - \dot{b} \boldsymbol{\omega}_a - \mathbf{F} \times \nabla b + [\nabla(\mathcal{M} + p/\rho) - p \nabla(1/\rho)] \times \nabla b \quad (8.30b)$$

$$= \mathbf{J}^Q + [\nabla(\mathcal{M} + p/\rho) - p \nabla(1/\rho)] \times \nabla b, \quad (8.30c)$$

where  $\mathbf{J}^Q$  is the Haynes-McIntyre form of the PV-substance flux given by equation (8.1). The term

$$\nabla(\mathcal{M} + p/\rho) \times \nabla b = \nabla \times [(\mathcal{M} + p/\rho) \nabla b] \quad (8.31)$$

is a total curl and as such it can be moved around without altering the evolution of PV-substance. Furthermore, since it is parallel to buoyancy isosurfaces it does not alter the impermeability properties of the PV-substance flux.

*Marshall et al. (2001)* focused attention on the flux

$$\mathbf{J}^{marshall} = \tilde{\mathbf{J}}^{Q-ocn} - \nabla(\mathcal{M} + p/\rho) \times \nabla b = -[\partial_t \mathbf{v} + \nabla(\mathcal{M} + p/\rho)] \times \nabla b - \boldsymbol{\omega}_a \partial_t b. \quad (8.32)$$

Since  $\mathbf{J}^{marshall}$  differs from  $\tilde{\mathbf{J}}^{Q-ocn}$  by a total curl, their divergences are equal. Diagnostically desirable features of  $\mathbf{J}^{marshall}$  include the following:

- $\nabla b \cdot \mathbf{J}^{marshall}/(\rho Q) = \partial_t b$ , thus satisfying the impermeability theorem.
- $\mathbf{J}^{marshall}$  has no explicit reference to irreversible processes. Consequently, in some cases it can be simpler to diagnose than the Haynes-McIntyre flux,  $\mathbf{J}^Q$ .
- In a steady state, the flux is given by

$$\mathbf{J}^{marshall} = \nabla b \times \nabla(\mathcal{M} + p/\rho) = \nabla \times [b \nabla(\mathcal{M} + p/\rho)]. \quad (8.33)$$

Consequently,  $\mathcal{M} + p/\rho$  provides a streamfunction for the steady state flux on buoyancy surfaces. As emphasized by *Schär (1993)*, this result holds even when there are irreversible processes, thus providing useful diagnostics even in the presence of dissipation.

### 8.3.3 Integral constraints for steady state

The steady state PV-substance flux in the form (8.33) can be used to develop some integral constraints on the steady flow. For this purpose consider the steady form of  $\mathbf{J}^{marshall}$  and integrate over an arbitrary simply connected area making use of Stokes' theorem

$$\int_S \nabla \times [b \nabla B] \cdot \hat{\mathbf{n}} dS = \oint_{\partial S} b \nabla B \cdot d\mathbf{r} = \oint_{\partial S} b dB = - \oint_{\partial S} B db. \quad (8.34a)$$

The first equality made use of Stokes' theorem; the second make use of the identity for exact differentials

$$\nabla B \cdot d\mathbf{r} = dB; \quad (8.35)$$

and the final equality made use of

$$b \, dB = d(b \, B) - B \, db \quad (8.36)$$

and noted that the closed loop integral of an exact differential vanishes, so that

$$\oint_{\partial S} d(B \, b) = 0. \quad (8.37)$$

If we can find a closed contour where either  $B$  is a constant ( $dB = 0$ ), or the buoyancy is a constant ( $db = 0$ ), then we have the steady state constraint

$$\int_S \mathbf{J}^{\text{marshall}} \cdot \hat{\mathbf{n}} \, dS = 0 \quad \text{area enclosed by contour with } \mathcal{M} + p/\rho \text{ constant or } b \text{ constant.} \quad (8.38)$$

In regions where there are such closed contours, this constraint offers useful insight into the steady state balances. [Marshall \(2000\)](#) and [Polton and Marshall \(2007\)](#) make particular use of closed  $B$  contours on constant depth surfaces (so that  $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ ) in a Boussinesq ocean.

### 8.3.4 Further study

[Marshall et al. \(2001\)](#) builds from the generalized Bernoulli theorem of [Schär \(1993\)](#) and [Bretherton and Schär \(1993\)](#). We also consider these topics for a hydrostatic Boussinesq ocean in Section 7.6.4.

## 8.4 Integrated potential vorticity substance

In this section we derive some properties of integrated potential vorticity, with these properties merely the result of how potential vorticity is defined. We write potential vorticity using a global buoyancy field,  $b$ , as in our discussion of ocean potential vorticity in Section 8.3

$$\rho Q = \boldsymbol{\omega}_a \cdot \nabla b = \nabla \cdot (\boldsymbol{\omega}_a b) = \nabla \cdot [(\boldsymbol{\omega} + 2\boldsymbol{\Omega}) b]. \quad (8.39)$$

The following points are central to the results derived in this section, and they all follow from the non-divergent nature of the absolute vorticity.

- The divergence form given in the second and third equalities of equation (8.39) is the starting point for the derivations in this section. Indeed, as emphasized by [Morel et al. \(2019\)](#), the divergence form is appropriate for deriving discrete approximations since in this case the discrete potential vorticity also satisfies the properties developed in this section.
- As emphasized in Section 8.2, the properties in this section hold for any smooth scalar field that is used to define the potential vorticity.
- The properties in this section hold even when there is no materially invariant potential vorticity since we only make use of the non-divergent nature of the absolute vorticity.

### 8.4.1 The primary role of boundaries

#### Integral in terms of boundary vorticity and boundary buoyancy

We consider  $Q$  to be an intensive fluid property measuring the amount of PV-substance per unit mass (i.e., the concentration of PV-substance), and correspondingly with  $\rho Q$  the amount of PV-substance per volume.<sup>3</sup> With this interpretation, the amount of PV-substance within an arbitrary finite region is determined by the volume integral of  $\rho Q$

$$\mathcal{I} = \int_{\mathcal{R}} Q \rho dV = \int_{\mathcal{R}} \nabla \cdot (\omega_a b) dV = \oint_{\partial\mathcal{R}} b \omega_a \cdot \hat{n} dS, \quad (8.40)$$

where the final equality used the [divergence theorem](#). Hence, the volume integrated PV-substance in a region is determined solely by values of the absolute vorticity and buoyancy on the region boundary. This property is strikingly distinct from material tracers. In practice it can be useful to decompose the absolute vorticity into the relative vorticity plus planetary vorticity

$$\mathcal{I} = \int_{\mathcal{R}} Q \rho dV = \oint_{\partial\mathcal{R}} b \omega_a \cdot \hat{n} dS = \oint_{\partial\mathcal{R}} b \omega \cdot \hat{n} dS + \oint_{\partial\mathcal{R}} b 2\Omega \cdot \hat{n} dS. \quad (8.41)$$

#### Integral in terms of boundary velocity and boundary buoyancy gradient

We follow [Morel et al. \(2019\)](#) by deriving an alternative expression for  $\mathcal{I}$  in equation (8.41), with this alternative expression more convenient in some cases. For this purpose we write

$$b \omega = b \nabla \times \mathbf{v} = \nabla \times (b \mathbf{v}) - \nabla b \times \mathbf{v}, \quad (8.42)$$

and use the divergence theorem to eliminate the total curl term

$$\oint_{\partial\mathcal{R}} \nabla \times (b \mathbf{v}) \cdot \hat{n} dS = \int_{\mathcal{R}} \nabla \cdot [\nabla \times (b \mathbf{v})] dV = 0. \quad (8.43)$$

We are thus led to the equivalent expressions for the integrated potential vorticity substance

$$\mathcal{I} = \int_{\mathcal{R}} Q \rho dV \quad (8.44a)$$

$$= \oint_{\partial\mathcal{R}} b \omega \cdot \hat{n} dS + 2 \oint_{\partial\mathcal{R}} b \Omega \cdot \hat{n} dS \quad (8.44b)$$

$$= - \oint_{\partial\mathcal{R}} (\nabla b \times \mathbf{v}) \cdot \hat{n} dS + 2 \oint_{\partial\mathcal{R}} b \Omega \cdot \hat{n} dS. \quad (8.44c)$$

In effect, the alternative expressions in the above equations move the derivative operator between the boundary velocity (for computing the relative vorticity) and the boundary buoyancy. One formulation may be more convenient than the other, depending on the boundary conditions. We emphasize that once a particular formulation is chosen, it is necessary to use that formulation for all of the domain boundaries. We must do so since the curl term that moves us from one form to the other vanishes only when integrating over the full domain boundary.

---

<sup>3</sup>It can be useful to here recall the discussion of extensive and intensive fluid properties when studying material tracers in [VOLUME 2](#).

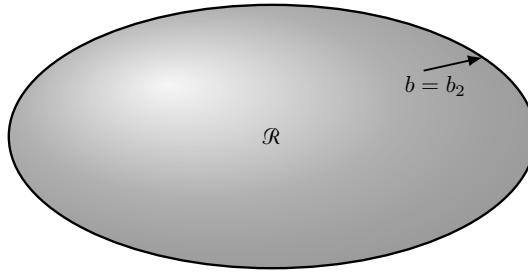


FIGURE 8.2: Integrating PV-substance over regions bounded by constant buoyancy surfaces that do not intersect the ground. Here we show a single buoyancy surface,  $b = b_2$ , bounding the bubble-like fluid region,  $\mathcal{R}$ . Notably, the region inside the bubble generally has a nontrivial buoyancy distribution. The only assumption is that it is wholly contained inside the  $b = b_2$  contour. There is identically zero domain integrated potential vorticity in  $\mathcal{R}$ . Hence, if there is any nontrivial distribution of potential vorticity somewhere in the domain, there must be as much integrated positive values as there are negative.

#### 8.4.2 Region bounded by a single buoyancy surface

Consider a volume of fluid bounded by a single buoyancy surface as shown in the bubble-like region in Figure 8.2. Since the outer boundary of the region is set by a constant  $b$ -surface, we can pull  $b$  outside of the surface integral in equation (8.44c) so that

$$\mathcal{I} = \oint_{\partial\mathcal{R}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} dS = b_2 \oint_{\partial\mathcal{R}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} dS. \quad (8.45)$$

We can now use the divergence theorem to return to the volume integral, only now with  $b$  outside of the integral

$$\mathcal{I} = b_2 \int_{\mathcal{R}} \nabla \cdot \boldsymbol{\omega}_a dV = 0, \quad (8.46)$$

where  $\nabla \cdot \boldsymbol{\omega}_a = 0$  led to the final equality. Equivalently, we can use Stokes' theorem to convert the closed area integral,  $\oint_{\partial\mathcal{R}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} dS$ , to a line integral around the boundary. However, there is no boundary for the closed area since it covers the sphere, thus again showing that  $\mathcal{I} = 0$ .

Yet another way to derive the identity (8.46) is to make use of the alternative expression for  $\mathcal{I}$  given by equation (8.44c)

$$\mathcal{I} = - \oint_{\partial\mathcal{R}} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} dS + 2 \oint_{\partial\mathcal{R}} b \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS. \quad (8.47)$$

Since the domain is bounded by a constant  $b$  surface, the outward normal is parallel to  $\nabla b$  so that the first integral vanishes. Furthermore, since  $b$  is a constant in the second integral we are led to consider

$$\oint_{\partial\mathcal{R}} 2 \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS = \oint_{\partial\mathcal{R}} [\nabla \times (\boldsymbol{\Omega} \times \mathbf{x})] \cdot \hat{\mathbf{n}} dS = \int_{\mathcal{R}} \nabla \cdot [\nabla \times (\boldsymbol{\Omega} \times \mathbf{x})] dV = 0, \quad (8.48)$$

which made use of the divergence theorem and the vanishing divergence of a curl.

The identity (8.46) says that there is zero integrated PV-substance contained within any region bounded by a single buoyancy surface; i.e., a bubble. The result holds whether there are reversible or irreversible processes acting on the buoyancy surface, and it holds if the  $b$ -surface is moving in space. Hence, within the domain there is just as much positive PV-substance as there is negative PV-substance. So if potential vorticity changes locally within the domain, then somewhere else it must experience an oppositely signed change so to leave a zero net

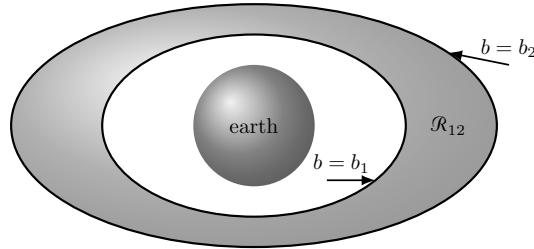


FIGURE 8.3: Integrating PV-substance over regions bounded by constant buoyancy surfaces that do not intersect the ground. Here we show a buoyancy layer or shell,  $\mathcal{R}_{12}$ , bounded by two buoyancy isosurfaces,  $b_1 < b_2$ , surrounding the earth, with neither surface intersecting the ground. There is identically zero domain integrated potential vorticity in  $\mathcal{R}_{12}$ . Hence, if there is any nontrivial distribution of potential vorticity somewhere in the domain, there must be as much integrated positive values as there are negative.

integrated PV-substance. We emphasize that this result holds at each time instance.

#### 8.4.3 Region bounded by two buoyancy surfaces

The identity (8.46) has a corollary, in which we consider a region bounded by two  $b$ -surfaces such as the region  $\mathcal{R}_{12}$  shown in Figure 8.3. The above arguments hold for that region as well, since we can decompose the surface integral into two integrals separately over  $b_1$  and  $b_2$

$$\mathcal{I} = \int_{\mathcal{R}_{12}} \nabla \cdot (\boldsymbol{\omega}_a b) dV = \int_{\mathcal{R}_2} \nabla \cdot (\boldsymbol{\omega}_a b) dV - \int_{\mathcal{R}_1} \nabla \cdot (\boldsymbol{\omega}_a b) dV, \quad (8.49)$$

where the domain  $\mathcal{R}_1$  extends from the ground up to  $b_1$  and  $\mathcal{R}_2$  extends from the ground up to  $b_2$ . Integration over the region below  $b_1$  cancels through the subtraction. Indeed, the region below  $b_1$  could be anything without changing the result. So let that region be filled with fluid throughout (i.e., ignore the earth) to allow us to extend both integrals throughout the spherical region just like in the buoyancy bubble  $\mathcal{R}$  in Figure 8.3. Invoking the buoyancy bubble result we see that both integrals separately vanish. We are thus led to a vanishing integral for the layer

$$\mathcal{I} = \int_{\mathcal{R}_{12}} \nabla \cdot (\boldsymbol{\omega}_a b) dV = 0. \quad (8.50)$$

Again, the key assumption is that no buoyancy surface intersects land, in which case we are able to ignore the presence of land altogether and thus make use of the buoyancy bubble result. The identity (8.50) also follows from the second form of equation (8.44c).

#### 8.4.4 Region bounded by land and a buoyancy surface

We now consider a domain consisting of fluid bounded by a buoyancy surface that intersects (incrops) the ground, such as the region shown in Figure 8.4. This atmospheric example can be turned over to produce an ocean example with buoyancy surfaces outcropping at the ocean surface. Using the vorticity form of the integrated potential vorticity in equation (8.44c) leads to

$$\mathcal{I} = \int_{\mathcal{R}_{AB}} \nabla \cdot (\boldsymbol{\omega}_a b) dV \quad (8.51a)$$

$$= \int_{S_A} \boldsymbol{\omega}_a b \cdot \hat{\mathbf{n}} dS + \int_{S_B} \boldsymbol{\omega}_a b \cdot \hat{\mathbf{n}} dS \quad (8.51b)$$

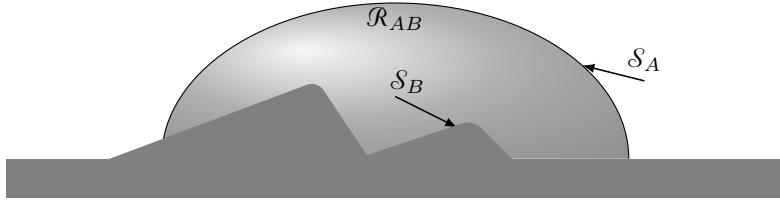


FIGURE 8.4: A fluid region,  $\mathcal{R}_{AB}$ , bounded by two buoyancy surfaces,  $\mathcal{S}_A$  and  $\mathcal{S}_B$ . The upper surface,  $\mathcal{S}_A$ , is defined by a buoyancy isosurface,  $b = b_A$ , with this surface intersecting the ground. The lower surface,  $\mathcal{S}_B$ , is along the ground (which is generally not flat, as shown here) and has a buoyancy that is a function of space and time,  $b = b_B(\mathbf{x}, t)$ .

$$= b_A \int_{\mathcal{S}_A} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_B} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (8.51c)$$

$$= b_A \left[ \int_{\mathcal{S}_A} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_B} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\mathcal{S}_B} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \right] + \int_{\mathcal{S}_B} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (8.51d)$$

$$= b_A \int_{\mathcal{R}_{AB}} \nabla \cdot \boldsymbol{\omega}_a dV + \int_{\mathcal{S}_B} (b - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (8.51e)$$

$$= \int_{\mathcal{S}_B} (b - b_A) (\boldsymbol{\omega} + 2\boldsymbol{\Omega}) \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (8.51f)$$

where we made use of  $\nabla \cdot \boldsymbol{\omega}_a = 0$  to reach the final equality. As both the ground and the ocean surface have buoyancy gradients, they contribute to the PV-substance within the region they bound.

For this domain it can be quite useful to use the second form of the integral in equation (8.44c). For this purpose we write

$$\int_{\mathcal{S}_B} (b - b_A) \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}_B + \mathcal{S}_A} (b - b_A) \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (8.52a)$$

$$= - \int_{\mathcal{S}_B + \mathcal{S}_A} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (8.52b)$$

$$= - \int_{\mathcal{S}_B} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (8.52c)$$

Equation (8.52a) follows from

$$\int_{\mathcal{S}_A} (b - b_A) \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} = 0 \quad \text{since } b = b_A \text{ on } \mathcal{S}_A. \quad (8.53)$$

Equation (8.52b) follows from the divergence theorem

$$\int_{\mathcal{S}_B + \mathcal{S}_A} \nabla \times [(b - b_A) \mathbf{v}] \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{R}_{AB}} \nabla \cdot (\nabla \times [(b - b_A) \mathbf{v}]) dV = 0. \quad (8.54)$$

And equation (8.52c) holds since  $\hat{\mathbf{n}} \times \nabla b = 0$  along  $\mathcal{S}_A$ . The expression (8.52c) is particularly convenient for the case of a **no-slip boundary condition**, whereby the velocity vanishes along  $\mathcal{S}_B$  so that we are left with the rather tidy expression

$$\mathcal{I} = 2 \int_{\mathcal{S}_B} (b - b_A) \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad \text{no-slip condition on } \mathcal{S}_B. \quad (8.55)$$

### 8.4.5 A layer outcropping at the ocean surface

Figure 8.5 depicts a buoyancy layer that outcrops at the ocean surface at both of its ends. Following the derivation in Section 8.4.4 leads to the integrated PV-substance

$$\mathcal{I}_A \equiv \int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_A} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3} (b - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (8.56a)$$

$$\mathcal{I}_B \equiv \int_{\mathcal{S}_2 + \mathcal{S}_B} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}_2} (b - b_B) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (8.56b)$$

with the difference leading to the integrated potential vorticity within the layer  $\mathcal{R}_{AB}$

$$\mathcal{I}_{AB} \equiv \mathcal{I}_A - \mathcal{I}_B = \int_{\mathcal{R}_{AB}} \rho Q dV = \int_{\mathcal{S}_1 + \mathcal{S}_3} (b - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + (b_B - b_A) \int_{\mathcal{S}_2} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (8.57)$$

The relation (8.57) requires information about the absolute vorticity over the region  $\mathcal{S}_2$  that lies outside the outcrop regions. To instead only make use of information over the outcrop areas,  $\mathcal{S}_1$  and  $\mathcal{S}_3$ , we consider buoyancy gradients when considering the contributions from the relative vorticity

$$\begin{aligned} & \int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_A} b \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\mathcal{S}_2 + \mathcal{S}_B} b \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} \\ &= - \int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_A} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_2 + \mathcal{S}_B} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \end{aligned} \quad (8.58)$$

We can drop the integrals along  $\mathcal{S}_A$  and  $\mathcal{S}_B$  since their respective normals are parallel to  $\nabla b$ , in which case

$$\int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_A} b \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\mathcal{S}_2 + \mathcal{S}_B} b \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} = - \int_{\mathcal{S}_1 + \mathcal{S}_3} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (8.59)$$

We are thus led to write the layer integrated PV-substance in the form

$$\mathcal{I}_{AB} = - \int_{\mathcal{S}_1 + \mathcal{S}_3} (\nabla b \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_1 + \mathcal{S}_3} (b - b_A) 2 \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} d\mathcal{S} + (b_B - b_A) \int_{\mathcal{S}_2} 2 \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (8.60)$$

As desired, this alternative formulation only requires information about the flow field and buoyancy field over the outcrop surfaces,  $\mathcal{S}_1$  and  $\mathcal{S}_3$ . For the region between the outcrops, we only need to know its area and outward normal, with  $\hat{\mathbf{n}} \approx \hat{\mathbf{z}}$  an accurate approximation.

### 8.4.6 Further study

Section 4.8 of [Vallis \(2017\)](#) discusses the integrated PV-substance in terms of the vorticity formulation, whereas [Morel et al. \(2019\)](#) introduced the dual perspective based on the buoyancy gradient formulation. [Morel et al. \(2019\)](#) also provide details for the practical diagnosis of potential vorticity in a numerical ocean model or from observational based measurements.

## 8.5 Boundary fluxes of PV-substance

In Section 8.4 we developed kinematic expressions for the PV-substance integrated over a selection of volumes. That discussion illustrated how the volume integrated potential vorticity

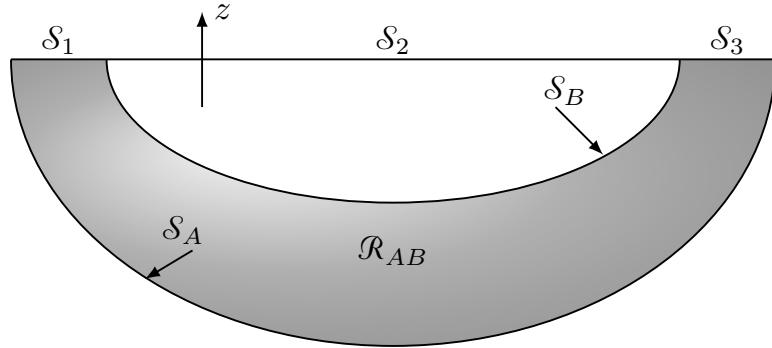


FIGURE 8.5: This figure depicts a buoyancy layer in the ocean that outcrops at both ends of the layer. The boundaries for the layer are given by the following surfaces. Surfaces  $\mathcal{S}_A$  and  $\mathcal{S}_B$  are defined by buoyancy isosurfaces with  $b_A < b_B$ . The sea surface is decomposed into three regions,  $\mathcal{S}_1$ ,  $\mathcal{S}_2$ , and  $\mathcal{S}_3$  according to the outcrop locations of  $\mathcal{S}_A$  and  $\mathcal{S}_B$ .

has contributions only from boundaries; e.g., where an atmospheric region intersects the ground or ocean, and where an oceanic region intersects the ground or the atmosphere. In this section we further our understanding of budgets for PV-substance by examining a buoyancy layer within the ocean that intersects the bottom on one side and the atmosphere on the other (Figure 8.6). We garner further understanding of the physical processes affecting changes to the PV-substance by here unpacking the boundary fluxes.

A buoyancy layer generally moves as it expands and contracts due to both reversible and irreversible processes (waves, currents, mixing). The impermeability theorem means that the total potential vorticity substance for the layer changes only through exchanges at the boundaries, including the bottom (boundary between the solid earth and the fluid) and air-sea boundaries. Removing interior interfaces from the layer PV-substance budget simplifies the budget analysis, as already revealed in Section 8.4. As per the discussion of Section 8.3, the results in this section apply even when there is no materially invariant potential vorticity. All we require is a flux-form budget along with the impermeability theorem, which holds for potential vorticity defined according to an arbitrary smooth scalar field (Section 8.2.2).

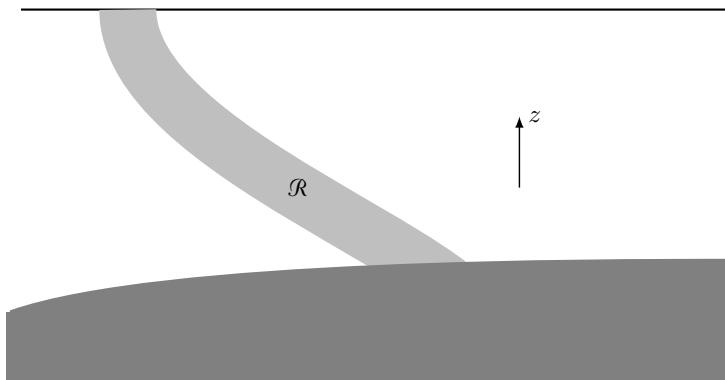


FIGURE 8.6: A buoyancy layer of seawater denoted by  $\mathcal{R}$ , with the layer intersecting bottom topography on one side and the atmosphere on the other. The interior boundaries of the layer are formed by constant buoyancy surfaces.

### 8.5.1 Layer integrated budget

In addition to waves, currents, mixing, and sources affecting the layer interfaces, there is movement of the intersection of the layer with the side boundaries, thus changing the vertical and horizontal extents of these intersections. As a formulation framework, we derive the layer potential vorticity budget making use of the Leibniz-Reynolds transport theorem derived in VOLUME 2. Just as for the layer integrated tracer budget (also examined in VOLUME 2), applying Leibniz-Reynolds to the layer integrated potential vorticity budget renders

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho Q dV \right] = \int_{\mathcal{R}} \partial_t (\rho Q) dV + \oint_{\partial \mathcal{R}} \rho Q \dot{\mathbf{x}} \cdot \hat{\mathbf{n}} dS, \quad (8.61)$$

where  $\mathcal{R}$  is the domain defined by the layer (Figure 8.6),  $\partial \mathcal{R}$  is its boundary, and

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} \quad (8.62)$$

is the velocity for a point on the boundary. Making use of the potential vorticity equation,

$$\partial_t (\rho Q) = -\nabla \cdot \mathbf{J}^Q, \quad (8.63)$$

and the divergence theorem renders

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho Q dV \right] = \oint_{\partial \mathcal{R}} (-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} dS. \quad (8.64)$$

This result holds around the full domain boundary. Now we decompose that boundary into portions defined by layer interfaces and those along the air-sea and land-sea boundaries.

### 8.5.2 Impermeability across interior layer interfaces

Rather than invoking the impermeability theorem derived in Section 8.2, we rederive it within the present context to further our confidence in its use. We thus consider the following for interior layer interfaces, here making use of the Haynes-McIntyre form (8.1) of the potential vorticity flux vector

$$[-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}] \cdot \hat{\mathbf{n}} = \left[ \rho Q (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a - \nabla b \times \mathbf{F} \right] \cdot \hat{\mathbf{n}} \quad (8.65a)$$

$$= \left[ (\boldsymbol{\omega}_a \cdot \nabla b) (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a \right] \cdot \hat{\mathbf{n}} \quad (8.65b)$$

$$= [(\boldsymbol{\omega}_a \cdot \nabla b) (\dot{\mathbf{x}} - \mathbf{v}) + (\partial_t b + \mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a] \cdot \hat{\mathbf{n}} \quad (8.65c)$$

$$= [(\boldsymbol{\omega}_a \cdot \nabla b) \dot{\mathbf{x}} + \boldsymbol{\omega}_a \partial_t b] \cdot \hat{\mathbf{n}}, \quad (8.65d)$$

where

$$\hat{\mathbf{n}} = \nabla b / |\nabla b| \quad (8.66)$$

is the outward unit normal vector pointing to regions of higher buoyancy. Now recall that the velocity of a point fixed on an layer interface has a normal component that satisfies equation (8.10) (here applied to buoyancy surfaces)

$$\dot{\mathbf{x}} \cdot \hat{\mathbf{n}} = -\partial_t b / |\nabla b|. \quad (8.67)$$

This result then leads to the impermeability statement for isopycnal interfaces in the fluid interior

$$(-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = 0. \quad (8.68)$$

We thus conclude that changes to the layer integrated potential vorticity occur only via transfer across the land-sea boundary and the air-sea boundary

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho Q dV \right] = \int_{\text{land-sea}} [-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}] \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\text{air-sea}} [-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (8.69)$$

We now separately consider these two boundaries.

### 8.5.3 Potential vorticity flux at a land-sea boundary

We here evaluate the potential vorticity flux from equation (8.6a) at a land-sea boundary

$$-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}} = \rho Q (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a - \nabla b \times \mathbf{F}. \quad (8.70)$$

At a solid and static boundary, the no-normal flow boundary condition means that  $\hat{\mathbf{n}} \cdot \mathbf{v} = 0$ . Likewise, the velocity of a point along the boundary moves along the tangent to the boundary so that  $\dot{\mathbf{x}} \cdot \hat{\mathbf{n}} = 0$ . Hence, the bottom boundary condition is solely comprised of irreversible processes

$$(-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = (\dot{b} \boldsymbol{\omega}_a - \nabla b \times \mathbf{F}) \cdot \hat{\mathbf{n}}. \quad (8.71)$$

If this boundary flux is positive, then it acts to increase the integrated PV-substance of the region, and conversely if the boundary flux is negative.

In many parts of the ocean bottom, geothermal heating is negligible so that there is no buoyancy input at the bottom, thus leaving just the contribution from friction

$$\text{no geothermal heating} \implies (-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = -(\nabla b \times \mathbf{F}) \cdot \hat{\mathbf{n}} = (\nabla b \times \hat{\mathbf{n}}) \cdot \mathbf{F}. \quad (8.72)$$

Furthermore, in the absence of geothermal heating the buoyancy satisfies a no-flux boundary condition, which can be ensured by having the buoyancy satisfying

$$\text{no geothermal heating} \implies \hat{\mathbf{n}} \cdot \nabla b = 0. \quad (8.73)$$

Buoyancy isolines thus intersect the bottom parallel to the bottom outward normal, as shown in Figure 8.7.<sup>4</sup> Correspondingly,  $(\nabla b \times \hat{\mathbf{n}}) \cdot \mathbf{F}$  projects onto that component of the friction vector pointing parallel to the bottom. Assuming buoyancy increases upward along the sloping bottom, as per a stably stratified fluid, then  $\nabla b \times \hat{\mathbf{n}}$  points counter-clockwise around bowls and clockwise around bumps (see Figure 8.7).

Within the bottom boundary layer, quadratic bottom drag is a common parameterization of the acceleration associated with turbulent frictional processes

$$\mathbf{F} = -C_d |\mathbf{u}| \mathbf{u}, \quad (8.74)$$

where  $C_d$  is a non-dimensional drag coefficient and  $\mathbf{u}$  is the horizontal velocity. In this case the

---

<sup>4</sup>There is ongoing research aimed at determining the thickness of the region over which the boundary condition (8.73) is accurate. The boundary condition presumably holds within a molecular sublayer. But the question is whether larger scale motions near the ocean bottom allow for this condition to hold over a thicker region. See the review chapter by [Polzin and McDougall \(2021\)](#) for discussion.

boundary condition for potential vorticity takes the form

$$(-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = (\nabla b \times \hat{\mathbf{n}}) \cdot \mathbf{F} = -C_d |\mathbf{u}| (\nabla b \times \hat{\mathbf{n}}) \cdot \mathbf{u}. \quad (8.75)$$

We see that the sign of the bottom boundary potential vorticity flux depends on the relative orientation of the bottom flow and the vector  $\nabla b \times \hat{\mathbf{n}}$ . To help understand the sign, consider an abyssal bowl with buoyancy increasing upward along the sloping bottom, in which case  $\nabla b \times \hat{\mathbf{n}}$  points counter-clockwise around the bowl. A bottom boundary flow that is also oriented counter-clockwise carries a positive curvature relative vorticity (Section 3.8). This positive relative vorticity is damped by the bottom friction, which corresponds to the negative potential vorticity source as per equation (8.75). Conversely, a bottom boundary flow that is oriented clockwise around the abyssal bowl carries a negative curvature relative vorticity. This negative relative vorticity is damped by the bottom friction, which corresponds to the positive potential vorticity source as per equation (8.75).

Consider a component to the bottom flow that is parallel to  $\nabla b$ . This flow provides a zero potential vorticity source since  $\nabla b \times \mathbf{F} = 0$  (again, assuming  $\mathbf{F} = -C_d |\mathbf{u}| \mathbf{u}$ ). This result is expected from the discussion in Section 7.5 and Figure 7.4, where we note that friction changes potential vorticity by rotating buoyancy surfaces, with that rotation realized only when friction is not aligned with  $\nabla b$ .

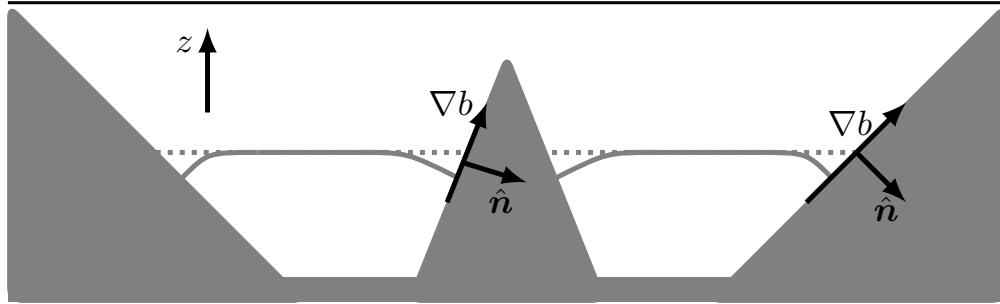


FIGURE 8.7: Depicting a buoyancy isosurface that intersects the bottom. As discussed in our study of tracers in VOLUME 2, in the absence of geothermal heating, a buoyancy isosurface satisfies the no-normal flux bottom boundary condition,  $\hat{\mathbf{n}} \cdot \nabla b = 0$ . This boundary condition requires buoyancy isosurfaces to be orthogonal to the bottom. Assuming buoyancy increases upward along the sloping bottom, then  $\nabla b \times \hat{\mathbf{n}}$  points counterclockwise around bowls and clockwise around bumps (when viewed from above). This structure for the buoyancy surfaces affects how friction impacts on the layer-integrated potential vorticity budget, with details provided in Section 8.5.3.

#### 8.5.4 Potential vorticity flux at the air-sea boundary

For the permeable air-sea boundary, we make use of the kinematic boundary condition derived in VOLUME 2, which leads to

$$\rho \hat{\mathbf{n}} \cdot (\dot{\mathbf{x}} - \mathbf{v}) = \mathcal{Q}_m \quad \text{air-sea boundary}, \quad (8.76)$$

with  $\mathcal{Q}_m$  the mass per time per surface area of matter that crosses the air-sea boundary. We are thus led to the air-sea boundary condition

$$(-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = Q \mathcal{Q}_m + (\dot{b} \boldsymbol{\omega}_a - \nabla b \times \mathbf{F}) \cdot \hat{\mathbf{n}}. \quad (8.77)$$

Besides the irreversible processes, potential vorticity is affected at the air-sea interface by the transfer of matter across the boundary via the term  $Q \mathcal{Q}_m$ . We can think of this term as an advection of potential vorticity across the boundary via the boundary mass transport. More generally, we can think of the full boundary flux (8.77) as acting to stretch/compress the fluid column so to alter vorticity and hence the potential vorticity.

To help interpret the friction term appearing in the flux (8.77), write

$$(b \boldsymbol{\omega}_a - \nabla b \times \mathbf{F}) \cdot \hat{\mathbf{n}} = (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) (\partial_t b + \mathbf{v} \cdot \nabla b) + (\hat{\mathbf{n}} \times \mathbf{F}) \cdot \nabla b \quad (8.78a)$$

$$= (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) [\partial_t b + (\mathbf{v} - \mathbf{v}_E) \cdot \nabla b], \quad (8.78b)$$

where we introduced the *generalized Ekman velocity*

$$\mathbf{v}_E \equiv \frac{\mathbf{F} \times \hat{\mathbf{n}}}{\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}}, \quad (8.79)$$

thus bringing the air-sea boundary potential vorticity flux (8.77) to the form

$$(-\mathbf{J}^Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = Q \mathcal{Q}_m + (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) [\partial_t b + (\mathbf{v} - \mathbf{v}_E) \cdot \nabla b]. \quad (8.80)$$

Note that for the special case of a vertical outward normal,  $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ , and weak relative vorticity,  $f + \zeta \approx f$ , we have

$$\mathbf{v}_E \approx f^{-1} \mathbf{F} \times \hat{\mathbf{z}}, \quad (8.81)$$

which is the Ekman velocity derived in our study of [Ekman mechanics](#) in VOLUME 2. We thus see that the sign of the  $(\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}})$  portion of the potential vorticity flux is determined by whether the surface buoyancy is increasing or decreasing in time following the velocity difference,  $\mathbf{v} - \mathbf{v}_E$ , along the air-sea boundary. We can interpret  $\mathbf{v} - \mathbf{v}_E$  as the inviscid portion of the velocity since it removes that portion arising from friction.<sup>5</sup>

### 8.5.5 Thought experiments

The surface potential vorticity flux (8.77), or its rewritten form in equation (8.80), provide an explicit expression for how surface boundary fluxes affect the potential vorticity budget within a buoyancy layer outcropping at the ocean surface. It contains a wealth of physics that can be explored via thought experiments.

#### Potential vorticity generation in a fluid with zero initial baroclinicity

Consider a fluid region initially with zero baroclinicity and zero flow so that the initial potential vorticity is given by  $f N^2$ , with  $N^2$  the squared buoyancy frequency. The surface potential vorticity flux (8.77) creates potential vorticity via the mass flux term and through heating/cooling. If this term alone affected the potential vorticity, and it did so uniformly in space, then it would alter potential vorticity only via changes in the vertical stratification. More generally, both the mass term and the diabatic term create horizontal buoyancy gradients, which then generate currents and vorticity that generate further contributions to the potential vorticity flux.

<sup>5</sup>The velocity difference,  $\mathbf{v} - \mathbf{v}_E$ , is not equal to the inviscid velocity that would appear in an inviscid (perfect) fluid. Rather,  $\mathbf{v} = \mathbf{v}_E + (\mathbf{v} - \mathbf{v}_E)$  is an interpretational decomposition akin to that used in our study of [Ekman mechanics](#) in VOLUME 2.

### Potential vorticity generation in a fluid that is initially homogeneous

Consider an initially homogenous box of seawater with zero potential vorticity. In this case it is only the buoyancy term,  $\dot{b} f$ , that contributes to initial changes in potential vorticity. Northern hemisphere ( $f > 0$ ) surface cooling ( $\dot{b} < 0$ ) adds negative potential vorticity to the box. Cooling also initiates gravitational instability that mixes the water and in turn spreads the negative potential vorticity boundary source throughout the fluid. Cooling adds structure to the buoyancy field by inflating the formerly zero thickness buoyancy layers, with layer inflation originating from the boundary. Once inflated, the impermeability theorem dictates that the layer integrated PV-substance changes only via boundary interactions, whereas stirring and mixing transport potential vorticity into the fluid interior. Notably, a region with  $f Q < 0$  is locally unstable to symmetric instability, with the generated symmetric instability acting to locally bring the flow towards a state with zero potential vorticity. However, the constraints from impermeability mean that the net PV-substance remains unchanged within a buoyancy layer, even in the presence of mixing.

#### 8.5.6 Is there a preferred form of the PV-substance flux?

Analysis in this section made use of the Haynes-McIntyre form of the PV-substance flux (equation (8.1))

$$\mathbf{J}^Q = \rho \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a + \nabla b \times \mathbf{F}. \quad (8.82)$$

We could have instead chosen to work with the Marshall form (equation (8.32))

$$\mathbf{J}^{marshall} = -[\partial_t \mathbf{v} + \nabla(\mathcal{M} + p/\rho)] \times \nabla b - \boldsymbol{\omega}_a \partial_t b, \quad (8.83)$$

or the modified kinematic form (equation (8.18))

$$\tilde{\mathbf{J}}^{kin} = -\partial_t \mathbf{v}_a \times \nabla b - \boldsymbol{\omega}_a \partial_t b. \quad (8.84)$$

These fluxes differ by a gauge choice and yet they each satisfy the impermeability theorem. Subjective choices determine which one is preferred. Importantly, once chosen, we can use only a single form of the flux throughout the budget analysis in order to remain self-consistent with the form of the total curl that is removed by the divergence operator.

The PV-substance budget, though invariant to the choice of flux, has distinct physical pictures depending on the choice of the flux. As a particularly clear example consider a steady state budget in which the fluxes take the form

$$\mathbf{J}^Q = \rho \mathbf{v} Q - (\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a + \nabla b \times \mathbf{F} \quad (8.85)$$

$$\mathbf{J}^{marshall} = -\nabla(\mathcal{M} + p/\rho) \times \nabla b \quad (8.86)$$

$$\tilde{\mathbf{J}}^{kin} = 0. \quad (8.87)$$

The physical picture for  $\tilde{\mathbf{J}}^{kin}$  is rather trivial, whereby the PV-substance stays constant within buoyancy layers and there are zero PV-substance fluxes across all boundaries of the layer. In contrast, a steady state budget when working with  $\mathbf{J}^Q$  or  $\mathbf{J}^{marshall}$  afford a physical picture of PV-substance entering, leaving, and transported through the buoyancy layers. [Marshall et al. \(2001\)](#) developed a rather elegant analysis framework using  $\mathbf{J}^{marshall}$  for steady budgets, and we explore facets of that approach in Section 7.6 for the special case of a Boussinesq hydrostatic fluid.

Nevertheless, our use of the Haynes-McIntyre PV-substance flux in the present section is motivated by its utility for describing how boundary forcing can change the sign of the potential vorticity. Such forcing exposes the flow to a variety of local instabilities (e.g., symmetric, centrifugal, gravitational). [Thomas et al. \(2008\)](#) offer a pedagogical review for the ocean; [Thomas et al. \(2013\)](#) provides a thorough study of the upper reaches of the Gulf Stream; and [Naveira Garabato et al. \(2019\)](#) provide evidence for such boundary forcing in regions of strong abyssal flows. Each of these studies points to the need to further understand details of the boundary potential vorticity flux and to furthermore ensure it is properly formulated within numerical models (e.g., [Hallberg and Rhines \(1996\)](#)).

### 8.5.7 Further study

The study of boundary potential vorticity fluxes generally requires careful analysis of the multitude of processes active in boundary layers. The interested reader can find ocean examples of these analyses in [Benthuyzen and Thomas \(2012\)](#) and [Wenegrat et al. \(2018\)](#), and the references therein.





## **Part III**

### **Nearly geostrophic flows**

Fluid motion dominated by rotation is characterized by a small Rossby number. To zeroth order in an asymptotic expansion in Rossby number, the flow maintains geostrophic balance , which is a balance between the Coriolis acceleration and pressure gradient acceleration. As seen in our study of geostrophy in VOLUME 2, the geostrophic balance is diagnostic, which means it offers no means to compute time evolution of the motion. To obtain a prognostic equation requires going to next order in Rossby number within the asymptotic expansion. For quasi-geostrophy (QG), the resulting prognostic equation makes use of ageostrophic motions, though only as an intermediate step towards an evolution equation involving just zeroth order geostrophically balanced fields. For planetary geostrophy (PG), the prognostic equation arises from mass (or volume) conservation, whereas fluid motion is a directly determined by evolution of the mass field.

The nuts and bolts of this part of the book involve methods of scale analysis and asymptotic analysis via perturbation series. In Chapter 9, we use these tools to derive equations for both planetary geostrophy and quasi-geostrophy within shallow water flows, and then extend that discussion to continuously stratified flows in Chapters 10 and 11. Planetary geostrophy and quasi-geostrophy are both very useful theoretical models lending insights into ocean and atmosphere fluid mechanics. In particular, planetary geostrophy is often the foundation for theories of large-scale laminar ocean circulation. Quasi-geostrophy serves as a theoretical model for studies of both oceanic and atmospheric flows at or near the deformation radius, particularly when concerned with transient features such as Rossby waves, baroclinic instability, and geostrophic turbulence.

One central property of balanced flow is that knowledge of the relevant balanced version of potential vorticity is sufficient to determine the flow field. A way to appreciate this property is to consider a horizontally non-divergent barotropic flow (Chapter 4), in which the Laplacian of the streamfunction gives the relative vorticity,  $\zeta = \nabla^2\psi$ . Conversely, if we know the vorticity then we can invert the Laplacian (with suitable boundary conditions) to yield the streamfunction and hence the velocity.<sup>6</sup> In our discussion of quasi-geostrophy in Chapter 11, we extend this result to three-dimensional quasi-geostrophic flow, where the prognostic fields are the horizontal components to the geostrophic flow, plus the buoyancy field. Such methods of potential vorticity inversion are routinely used to study atmospheric flows given maps of potential vorticity (see [Hoskins \(1991\)](#) for a review).

#### MATHEMATICS IN THIS PART

The mathematics in this part rely on the Cartesian tensor analysis and vector calculus from VOLUME 1, as well as elements from linear partial differential equations, particularly elliptic equations, also studied in VOLUME 1.

---

<sup>6</sup>We study the Green's function method for elliptic partial differential equations in VOLUME 1.

## Chapter 9

# MODELS OF NEARLY GEOSTROPHIC FLOWS

Planetary rotation, and the corresponding Coriolis acceleration, is a primary feature of geophysical fluid flows. To probe the physics of such flows, it is useful to develop mathematical models where rotation of the reference frame is a prominent facet of the equations of motion. In this chapter we systematically derive such models, thus developing the mechanical equations for [planetary geostrophy](#) (PG) and [quasi-geostrophy](#) (QG), both within the shallow water fluid. We also non-dimensionalize the stratified Boussinesq equations, thus providing the foundations for the equations of stratified planetary geostrophy studied in Chapter 10 and stratified quasi-geostrophy studied in Chapter 11.

### READER'S GUIDE FOR THIS CHAPTER

This chapter makes use of dimensional analysis, scale analysis, and asymptotic methods to derive approximate geostrophic equations, exposing details of the corresponding mathematical derivations. We make use of equations for a single layer of shallow water fluid as derived in Chapter 1, as well as the [Boussinesq ocean](#) equations from VOLUME 2. We follow this work with studies of continuously stratified planetary geostrophy in Chapter 10 and continuously stratified quasi-geostrophy in Chapter 11.

<b>9.1</b>	<b>Loose threads</b>	<b>314</b>
<b>9.2</b>	<b>Scale analysis and the Buckingham-II theorem</b>	<b>314</b>
<b>9.3</b>	<b>Shallow water equations</b>	<b>315</b>
9.3.1	Dimensional scales	315
9.3.2	Physical dimensions	316
9.3.3	Number of non-dimensional parameters	317
9.3.4	Choosing the non-dimensional parameters	317
9.3.5	Assumed values for the non-dimensional parameters	319
9.3.6	Deformation radius and the free surface undulation scale	320
9.3.7	Non-dimensional shallow water equations	320
<b>9.4</b>	<b>Shallow water planetary geostrophy</b>	<b>322</b>
9.4.1	Equations for planetary geostrophy	322
9.4.2	Reduced dynamical degrees of freedom	323
<b>9.5</b>	<b>Shallow water quasi-geostrophy</b>	<b>323</b>
9.5.1	Quasi-geostrophic scaling	323
9.5.2	Outlining the asymptotic method	325
9.5.3	Order $\text{Ro}^0$ asymptotic equations	326
9.5.4	Order $\text{Ro}^1$ asymptotic equations	327
<b>9.6</b>	<b>Elements of shallow water quasi-geostrophy</b>	<b>329</b>
9.6.1	Dimensional potential vorticity and streamfunction	329

9.6.2	Contributions to quasi-geostrophic potential vorticity . . . . .	330
9.6.3	Connecting to background (resting state) potential vorticity . . . . .	331
9.6.4	Connecting to Rossby's shallow water potential vorticity . . . . .	331
9.6.5	Geostrophic flow via potential vorticity inversion . . . . .	332
9.6.6	Evolution of quasi-geostrophic vorticity, velocity, and free surface	333
9.6.7	Unpacking quasi-geostrophic evolution . . . . .	336
9.6.8	Energetics of quasi-geostrophic flows . . . . .	337
9.6.9	Considering topography to be $\mathcal{O}(\text{Ro}^0)$ . . . . .	337
9.6.10	Two layer quasi-geostrophy . . . . .	338
9.6.11	Rigid lid shallow water quasi-geostrophy . . . . .	339
9.6.12	Further study . . . . .	339
<b>9.7</b>	<b>Non-dimensional Boussinesq ocean equations</b> . . . . .	<b>340</b>
9.7.1	Dimensional parameters . . . . .	341
9.7.2	Physical dimensions and non-dimensional parameters . . . . .	341
9.7.3	Choosing the non-dimensional parameters . . . . .	342
9.7.4	Relating the buoyancy scale to the Coriolis acceleration scale . .	343
9.7.5	Richardson number and QG/PG flow regimes . . . . .	343
9.7.6	The Rossby deformation radius . . . . .	344
9.7.7	Assumed values for the non-dimensional parameters . . . . .	345
9.7.8	Non-dimensional Boussinesq equations . . . . .	347
9.7.9	Comments . . . . .	348
<b>9.8</b>	<b>Exercises</b> . . . . .	<b>348</b>

---

## 9.1 Loose threads

- External and internal modes for QG two layers in Section 9.6.10.
- Exercises

## 9.2 Scale analysis and the Buckingham-II theorem

Scale analysis is ubiquitous in physics, with geophysical fluid mechanics no exception. We find it useful to distinguish two classes of scales. An **external scale** is prescribed as part of the physical system under consideration, and it can be considered under “control” of the physicist. Examples include planetary rotation rate, gravitational acceleration, and domain geometry. An **internal scale** emerges from the flow itself, such as the length and time scales of the flow, and as such internal scales are only indirectly controlled by the physicist.

The Buckingham-II theorem provides a useful framework for organizing dimensions and scales. The theorem states that the number of dimensionless parameters characterizing a physical system is a function of the number of dimensional parameters or scales,  $K_{\text{scales}}$  (e.g., scales for the velocity, rotation rate, pressure force, friction force, gravitational acceleration) and the number of physical dimensions,  $K_{\text{dimensions}}$  (e.g., time, length, mass). Precisely, Buckingham-II states that the number of dimensionless parameters is

$$N_{\text{dimensionless}} = K_{\text{scales}} - K_{\text{dimensions}}. \quad (9.1)$$

Different physical systems possessing the same suite of dimensionless parameters are isomorphic. For example, a laboratory study of flow around a cylinder contains two dimensionless parameters: the drag coefficient,  $C_d$ , and the Reynolds number,  $\text{Re}$ . If the problem is scaled up to a building with the same shape, then so long as the values for the dimensionless parameters are the same

(e.g., same drag coefficient and same Reynolds number), one can make use of the laboratory analog for determining suitability of the building architecture. Similar isomorphisms exist between flows in a rotating tank and flows in the ocean and atmosphere.

The Buckingham-II theorem does not provide the form of the dimensionless parameters. Nor does the theorem determine their values. This information comes only after introducing physical prejudices surrounding a regime of chosen interest. Additionally, Buckingham-II does not offer information about how the dimensionless parameters might be related. Instead, any such relations arise from the mechanical and thermodynamical equations describing the system. Consequently, mechanical and thermodynamical equations generally mean that fewer than  $N_{\text{dimensionless}}$  non-dimensional numbers are independent.

We focus in this chapter on the regime of large-scale atmospheric and oceanic flow where the shallow water fluid is close to geostrophic balance. That choice then guides the length and time scales, which in turn determines the size of the dimensionless parameters. In many cases, one is able to identify dimensionless parameters that are large or small in particular regimes, which in turn suggests asymptotic analyses to render equations specific to the regime of interest.

## 9.3 Shallow water equations

A single-layer of inviscid shallow water fluid of thickness  $h$  is governed by the velocity and thickness equations (Chapter 1)

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \mathbf{f} \times \mathbf{u} = -g \nabla \eta \quad (9.2a)$$

$$\partial_t h + \nabla \cdot (h \mathbf{u}) = 0, \quad (9.2b)$$

where  $\mathbf{u}$  is the horizontal velocity that is independent of depth within the layer ( $\partial_z \mathbf{u} = 0$ ). The vertical bounds for the layer are specified by the free surface height,  $z = \eta(x, y, t)$ , and bottom topography,  $z = \eta_b(x, y)$  (see Figure 1.1). They are related according to

$$\eta = \eta_b + h = H + \bar{\eta}_b + \eta' = \bar{h} + \bar{\eta}_b + \eta', \quad (9.3)$$

where  $H = \bar{h}$  is the area average layer thickness,  $\bar{\eta}_b$  is the area average of the bottom topography,  $\eta' = \eta - \bar{\eta}$  is the deviation of the surface height from its area average,  $\bar{\eta} = \bar{\eta}_b + H$ . We also find occasion to write the layer thickness in the form

$$h = \eta - \eta_b = H + (\bar{\eta}_b - \eta_b) + \eta' = \bar{h} - \eta'_b + \eta' = \bar{h} + h', \quad (9.4)$$

where  $\eta'_b = \eta_b - \bar{\eta}_b$  is the deviation of the bottom topography from its area average, and we introduced the thickness deviation

$$h' = h - \bar{h} = \eta' - \eta'_b. \quad (9.5)$$

Dexterity with the geometrical relations (9.3)-(9.5) is assumed in the following.

### 9.3.1 Dimensional scales

We identify the following nine dimensional scales for the shallow water layer.

- LENGTH SCALES

- ★  $H$  = depth scale of the fluid layer. Given that the layer is homogeneous, we take  $H$  equal to the area average layer thickness (see Figure 1.1).
- ★  $L$  = horizontal/lateral length scale of flows under consideration. We assume both horizontal directions to have the same flow length scale. This assumption is not necessarily appropriate on a rotating planet, where flows in the zonal (east-west) direction typically have length scales longer than meridional (north-south) flow scales. Nonetheless, this choice does not preclude the dynamical emergence of anisotropic length scales, and so it serves our uses for the present analysis.
- ★  $R_e$  = prescribed radius of the planet. We include this scale anticipating that for length scales small compared to the earth's radius, the Coriolis parameter may be approximated by a constant ( $f$ -plane) or linear function of latitude ( $\beta$ -plane).
- ★  $\mathcal{H}$  = length scale for deviations of the free surface height,  $\eta$ , relative to its area average,  $\bar{\eta}$ , so that  $\eta' = \eta - \bar{\eta} \sim \mathcal{H}$ .
- ★  $\mathcal{B}$  = prescribed length scale for deviations of the bottom topography,  $\eta_b$ , relative to its area average,  $\bar{\eta}_b$ , so that  $\eta'_b = \eta_b - \bar{\eta}_b \sim \mathcal{B}$ .

- VELOCITY SCALES

- ★  $U$  = velocity scale for horizontal fluid particle motion; i.e., the speed for horizontal flows.
- ★  $c$  = scale for wave speeds. For the shallow water model, the wave speed is given by the shallow water gravity wave (VOLUME 4)

$$c_{\text{grav}} = \sqrt{g H}. \quad (9.6)$$

We introduce the wave speed anticipating the presence of distinct flow regimes depending on whether the fluid particle speed is larger or smaller than the wave speed. The ratio of the fluid particle speed (advection speed) to the wave speed is known as the **Froude number** (Section 9.3.4).

- BODY FORCES: There are two body forces acting on the fluid; one from gravity and one from Coriolis. These two forces are scaled according to the following prescribed parameters:

- ★  $g$  = gravitational acceleration
- ★  $f$  = Coriolis frequency.

If we were interested in other forces, such as electromagnetic or frictional forces, then we add further dimensional parameters corresponding to these forces.

### 9.3.2 Physical dimensions

There are two physical dimensions in the shallow water system: length,  $L$ , and time,  $T$ . Notably, there is no mass in the shallow water system. The reason is that the fluid density is assumed uniform within a shallow water layer so that mass is described by area times thickness

$$M = \int \rho dV [\equiv] L^2 H \rho, \quad (9.7)$$

where  $[\equiv]$  is read “has dimensions”.

### 9.3.3 Number of non-dimensional parameters

The Buckingham-II theorem says there are

$$N_{\text{dimensionless}} = 9 - 2 = 7 \quad (9.8)$$

non-dimensional parameters that characterize the flow within a single shallow water layer. So we expect to find seven non-dimensional parameters as part of the analysis.

How do we know that we properly counted the physical dimensions or the dimensional scales? Fortunately, the process of determining the non-dimensional parameters is largely self-correcting. Namely, if a physical dimension or a physical scale is omitted, then it would appear somewhere in the subsequent analysis. Indeed, such omissions often are only found at the point of a mathematical or physical inconsistency. Hence, the analyst must remain cognizant of the need to self-correct when performing dimensional analysis.

### 9.3.4 Choosing the non-dimensional parameters

There is no unique choice for the non-dimensional parameters. Our choice is guided by experience, interest, and what parameters might be available to experimental control or measurement.

1. VERTICAL TO HORIZONTAL/LATERAL ASPECT RATIO: The ratio of the vertical scale of the layer to the horizontal/lateral scale of the flow<sup>1</sup> defines the aspect ratio

$$\delta_{\text{vertical/horizontal}} = \frac{\text{vertical length scale}}{\text{horizontal length scale}} = \frac{H}{L}. \quad (9.9)$$

2. RATIO OF HORIZONTAL/LATERAL SCALE TO PLANETARY SCALE: The ratio of the horizontal length scale of the flow to the planetary radius is

$$\delta_{\text{horizontal/planet}} = \frac{\text{horizontal length scale}}{\text{planetary length scale}} = \frac{L}{R_e}. \quad (9.10)$$

3. RATIO OF FREE SURFACE UNDULATION TO VERTICAL LENGTH SCALE: The ratio of the free surface undulations that emerge from the flow, to the vertical length scale of the layer is

$$\delta_{\text{free surface/depth}} = \frac{\text{free surface undulation scale}}{\text{vertical length scale}} = \frac{\mathcal{H}}{H}. \quad (9.11)$$

4. RATIO OF BOTTOM TOPOGRAPHY UNDULATION TO VERTICAL LENGTH: The ratio of the prescribed bottom topography undulation scale to the vertical length scale is

$$\delta_{\text{bottom/depth}} = \frac{\text{bottom topography undulation scale}}{\text{vertical length scale}} = \frac{\mathcal{B}}{H}. \quad (9.12)$$

5. FROUDE NUMBER: The Froude number is the ratio of the fluid particle speed emerging from the flow, to the gravity wave speed determined by geophysical parameters. For the shallow water system, this ratio is

$$\text{Fr} = \frac{U}{c} = \frac{U}{\sqrt{g H}}. \quad (9.13)$$

---

<sup>1</sup>The terms “lateral” and “horizontal” are used interchangeably, referring to motion at constant vertical position either on the sphere or on the plane.

6. ROSSBY NUMBER: The Rossby number is the ratio of the fluid particle acceleration scale to the Coriolis acceleration scale, both of which emerge from the flow

$$\text{Ro} = \frac{\text{particle acceleration}}{\text{Coriolis acceleration}}. \quad (9.14)$$

The particle acceleration scale is determined by the local time tendency plus advection

$$\text{particle acceleration} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \sim \frac{U}{T} + \frac{U^2}{L}. \quad (9.15)$$

The local time tendency and advection generally have distinct scales, thus leading to the potential for two Rossby numbers

$$\text{Ro}^{\text{tendency}} = \frac{1}{Tf} \quad \text{and} \quad \text{Ro}^{\text{advection}} = \frac{U}{Lf}. \quad (9.16)$$

In the following we consider these two Rossby numbers to have the same scale, which is realized by advective contributions to material time evolution being comparable to local time changes. We refer to this choice as an **advective time scale**, whereby

$$T = \frac{L}{U} \implies \frac{U^2}{L} = \frac{U}{T}, \quad (9.17)$$

so that there is only one Rossby number

$$\text{Ro} = \frac{1}{fT} = \frac{U}{fL}. \quad (9.18)$$

Another interpretation for the Rossby number is the ratio of the relative vorticity to the planetary vorticity

$$\text{Ro} = \frac{\text{relative vorticity}}{\text{planetary vorticity}}. \quad (9.19)$$

With the relative vorticity scaling as  $U/L$  and the planetary vorticity scaling as  $f$ , we recover the expression (9.18) for the Rossby number.

7. GEOSTROPHIC NUMBER: We define the geostrophic number as the ratio of the Coriolis acceleration scale to the scale of the pressure gradient acceleration

$$\text{Ge} = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}}. \quad (9.20)$$

The Coriolis acceleration scales as

$$\text{Coriolis acceleration} \sim fU, \quad (9.21)$$

whereas the pressure gradient acceleration,  $-g \nabla \eta$ , scales as

$$\text{pressure gradient acceleration} \sim \frac{g\mathcal{H}}{L}, \quad (9.22)$$

so that

$$\text{Ge} = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}} = \frac{fU}{(g/L)\mathcal{H}}. \quad (9.23)$$

### 9.3.5 Assumed values for the non-dimensional parameters

We now enumerate the assumed values for the non-dimensional parameters. These values are based on the preconceived topic of the analysis as determined by a chosen range of external and internal scales. Here, those scales arise from our interest in large scale flows in the atmosphere and ocean.

1. **SMALL VERTICAL TO HORIZONTAL ASPECT RATIO:** The aspect ratio is generally small for large-scale atmospheric and oceanic fluid systems

$$\delta_{\text{vertical/horizontal}} = H/L \ll 1. \quad (9.24)$$

This assumption was made when formulating the shallow water system, which is based on hydrostatic balance (see Section 1.2). We thus retain this assumption as we further scale the shallow water system.

2. **SMALL OR ORDER ONE RATIO OF HORIZONTAL/LATERAL TO PLANETARY SCALES:** The ratio of the lateral length scale to the planet radius is small for quasi-geostrophic systems and order unity for planetary geostrophy

$$\delta_{\text{horizontal/planet}} = L/R_e \ll 1 \quad \text{quasi-geostrophy} \quad (9.25a)$$

$$\delta_{\text{horizontal/planet}} = L/R_e \sim 1 \quad \text{planetary geostrophy}. \quad (9.25b)$$

3. **RATIO OF FREE SURFACE UNDULATION TO VERTICAL LENGTH SCALE:** The ratio  $\mathcal{H}/H$  is implied below by assuming a unit geostrophic number.
4. **RATIO OF BOTTOM TOPOGRAPHY UNDULATION TO VERTICAL LENGTH SCALE:** For quasi-geostrophy, we assume that undulations in the bottom topography are small relative to the vertical length scale, whereas there is no restriction for planetary geostrophy. “Small” in the present context is determined by the Rossby number, in which case

$$\delta_{\text{bottom/depth}} = \mathcal{B}/H = \text{Ro} \quad \text{for quasi-geostrophy}. \quad (9.26)$$

5. **FROUDE NUMBER:** We find that the Froude number is implied by scales assumed for the other non-dimensional numbers.
6. **SMALL ROSSBY NUMBER:** The Rossby number is assumed small

$$\text{Ro} = U/(f L) \ll 1, \quad (9.27)$$

which means that the Coriolis acceleration is a leading order term in the horizontal velocity equation (9.2a).

7. **UNIT GEOSTROPHIC NUMBER:** The geostrophic number is assumed to be order unity

$$\text{Ge} \sim 1. \quad (9.28)$$

This assumption means that the Coriolis acceleration and pressure gradient acceleration scale together

$$f U \sim (g/L) \mathcal{H}. \quad (9.29)$$

Making use of the velocity equation (9.2a), we see that this scaling is consistent only so long as the Rossby number is small,  $\text{Ro} \ll 1$ . Furthermore, this scaling constrains the scale of the free surface undulation,  $\mathcal{H}$ , as we discuss in Section 9.3.6.

### 9.3.6 Deformation radius and the free surface undulation scale

We determine the scale for the free surface height undulation,  $\mathcal{H}$ , by making use of the assumed order unity geostrophic number. For this purpose, start from the geostrophic scaling of Coriolis and pressure gradient accelerations, equation (9.29), to express the free surface undulation scale according to

$$\eta' \sim \mathcal{H} = \frac{f U L}{g} = \text{Ro} \frac{f^2 L^2}{g} = \text{Ro} H \frac{f^2 L^2}{g H} = \text{Ro} H \left[ \frac{L}{L_d} \right]^2. \quad (9.30)$$

In the final equality we introduced the deformation raduis

$$L_d = \sqrt{g H} / f = c_{\text{grav}} / f, \quad (9.31)$$

where  $c_{\text{grav}} = \sqrt{g H}$  is the shallow water gravity wave speed. The deformation radius distinguishes flows where the Coriolis force is important,  $L \geq L_d$ , from those where the Coriolis can be neglected,  $L \ll L_d$ . Since the deformation radius decreases toward the poles, rotational effects are felt by smaller scales in the high latitudes than in the tropics. The shallow water deformation radius is an external scale that we find in Section 9.6.2 acts as a regime boundary between relative vorticity and vortex stretching for the quasi-geostrophic potential vorticity. We again encounter the shallow water deformation radius when discussing shallow water waves in VOLUME 4.

We can use  $L_d$  to rewrite the Froude number as

$$\text{Fr} = U / \sqrt{g H} = U / (f L_d) = \text{Ro} (L / L_d), \quad (9.32)$$

where the second equality wrote the Froude number as the ratio of the advection speed,  $U$ , to the rotation speed,  $f L_d$ . Furthermore, the squared ratio of the deformation radius to the lateral length scale of the flow is termed the Burger number

$$\text{Bu} = (L_d / L)^2. \quad (9.33)$$

Use of the Burger number allows us to write the Froude number in terms of the Rossby number and Burger number

$$\text{Fr} = \text{Ro} / \sqrt{\text{Bu}}. \quad (9.34)$$

Likewise, the free surface height undulation scale can be written

$$\mathcal{H} = H \text{Ro} (L / L_d)^2 = H \text{Ro} / \text{Bu} = H \text{Fr}^2 / \text{Ro}. \quad (9.35)$$

Hence, the ratio of the free surface undulations to the layer thickness (depth) scale is given by

$$\delta_{\text{free surface/depth}} = \mathcal{H} / H = \text{Ro} (L / L_d)^2 = \text{Ro} / \text{Bu} = \text{Fr}^2 / \text{Ro}. \quad (9.36)$$

Again, this scaling is implied by making the dynamical assumption of a unit geostrophic number, which means that the pressure gradient acceleration scales according to the Coriolis acceleration.

### 9.3.7 Non-dimensional shallow water equations

To non-dimensionalize the shallow water equations we introduce non-dimensional variables for time, space, velocity, and Coriolis parameter. Non-dimensional variables are adorned with a

widehat<sup>2</sup>

$$t = T \widehat{t}, \quad (x, y) = L(\widehat{x}, \widehat{y}), \quad \partial_t = \partial_{\widehat{t}}/T, \quad \nabla = \widehat{\nabla}/L, \quad (u, v) = U(\widehat{u}, \widehat{v}), \quad f = f_0 \widehat{f}, \quad (9.37)$$

where  $f_0$  is the Coriolis parameter at the central latitude for the beta-plane ( $\beta$ -plane) approximation introduced in VOLUME 2. We also require non-dimensional variables for the surface and bottom undulations

$$\eta' = \mathcal{H} \widehat{\eta}, \quad \eta'_b = \mathcal{B} \widehat{\eta}_b, \quad h = H + \eta' - \eta'_b = H + \mathcal{H} \widehat{\eta} - \mathcal{B} \widehat{\eta}_b, \quad (9.38)$$

where we used equation (9.4) for the layer thickness. Importantly, we assume that the non-dimensional variables (the widehat variables) are order unity. That assumption is critical for organizing terms in the asymptotic expansion.

### Non-dimensional velocity equation

Introducing the above variables into the shallow water velocity equation (9.2a) renders

$$\frac{U}{T} \frac{\partial \widehat{\mathbf{u}}}{\partial \widehat{t}} + \frac{U^2}{L} (\widehat{\mathbf{u}} \cdot \widehat{\nabla}) \widehat{\mathbf{u}} + f_0 U (\widehat{\mathbf{f}} \times \widehat{\mathbf{u}}) = -\frac{g \mathcal{H}}{L} \widehat{\nabla} \widehat{\eta}. \quad (9.39)$$

As before, we assume the time scale is given by the advection time

$$T = L/U = 1/(\text{Ro } f_0), \quad (9.40)$$

so that dividing by  $f_0 U$  leads to

$$\text{Ro} \left[ \frac{\partial \widehat{\mathbf{u}}}{\partial \widehat{t}} + (\widehat{\mathbf{u}} \cdot \widehat{\nabla}) \widehat{\mathbf{u}} \right] + (\widehat{\mathbf{f}} \times \widehat{\mathbf{u}}) = - \left[ \frac{g H}{f_0 L U} \frac{\text{Ro}}{\text{Bu}} \right] \widehat{\nabla} \widehat{\eta}, \quad (9.41)$$

where we set  $\mathcal{H} = H(\text{Ro/Bu})$  according to equation (9.36). We simplify the factor on the right hand side according to

$$\frac{g H}{f_0 L U} \frac{\text{Ro}}{\text{Bu}} = \frac{g H}{f_0 L U} \frac{U}{f_0 L} \frac{L^2}{L_d^2} = \frac{g H}{f_0 L U} \frac{U}{f_0 L} \frac{L^2 f_0^2}{g H} = 1. \quad (9.42)$$

Hence, the non-dimensional inviscid shallow water velocity equation takes on the rather elegant form

$$\text{Ro} \left[ \frac{\partial \widehat{\mathbf{u}}}{\partial \widehat{t}} + (\widehat{\mathbf{u}} \cdot \widehat{\nabla}) \widehat{\mathbf{u}} \right] + \widehat{\mathbf{f}} \times \widehat{\mathbf{u}} = -\widehat{\nabla} \widehat{\eta}. \quad (9.43)$$

Introducing the non-dimensional material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial \widehat{t}} + \widehat{\mathbf{u}} \cdot \widehat{\nabla} \quad (9.44)$$

brings the velocity equation to

$$\text{Ro} \frac{D \widehat{\mathbf{u}}}{Dt} + \widehat{\mathbf{f}} \times \widehat{\mathbf{u}} = -\widehat{\nabla} \widehat{\eta}. \quad (9.45)$$

<sup>2</sup>The L<sup>A</sup>T<sub>E</sub>X widehat symbol is used for non-dimensional variables, such as the non-dimensional velocity,  $\widehat{\mathbf{u}}$ . The widehat is distinguished from the hat used for unit vectors, such as for the vertical unit vector,  $\widehat{\mathbf{z}}$ . We also use widehats for thickness weighted means in Chapter 14, but that usage is completely distinct from the non-dimensionalization usage in the present chapter.

The velocity equation is consistent with a unit geostrophy number (i.e., Coriolis acceleration balances pressure gradient acceleration) if and only if the Rossby number is small, thus eliminating the material acceleration. We noted this point earlier when studying geostrophic motion in VOLUME 2. Even so, it is reassuring to see it emerge from the process of non-dimensionalization and scaling.

### Non-dimensional thickness equation

The thickness equation (9.2b) can be written

$$\frac{\partial \eta'}{\partial t} + \nabla \cdot [(H + \eta' - \eta'_b) \mathbf{u}] = 0, \quad (9.46)$$

which takes on the non-dimensional form

$$\frac{\mathcal{H}}{T} \frac{\partial \hat{\eta}}{\partial \hat{t}} + \frac{U H}{L} \hat{\nabla} \cdot [(1 + \hat{\eta} \mathcal{H}/H - \hat{\eta}_b \mathcal{B}/H) \hat{\mathbf{u}}] = 0. \quad (9.47)$$

The advective time scaling  $T = L/U$  brings the thickness equation to

$$\frac{\mathcal{H}}{H} \frac{\partial \hat{\eta}}{\partial \hat{t}} + \hat{\nabla} \cdot [(1 + \hat{\eta} \mathcal{H}/H - \hat{\eta}_b \mathcal{B}/H) \hat{\mathbf{u}}] = 0. \quad (9.48)$$

We further specialize this equation in the following, as determined by the assumed horizontal length scale of the flow.

## 9.4 Shallow water planetary geostrophy

We make use of the non-dimensional equations derived in Section 9.3.7 to derive the mechanical equations for planetary geostrophy. We already encountered facets of planetary geostrophy for the continuously stratified fluid in VOLUME 2, and for the shallow water in Sections 5.7, 5.9, and 5.8. We further pursue the continuously stratified theory in Chapter 10, developing a variety of vorticity analyses for use in understanding the large-scale ocean circulation. For these reasons, our presentation here is rather brief.

### 9.4.1 Equations for planetary geostrophy

Planetary geostrophy is realized by dropping the fluid particle acceleration from the momentum equation (9.45), given that it is one order of Rossby number smaller than the Coriolis and pressure gradient accelerations. This assumption means that the velocity equation reduces to the geostrophic balance

$$\hat{\mathbf{f}} \times \hat{\mathbf{u}} = -\hat{\nabla} \hat{\eta}. \quad (9.49)$$

We furthermore assume that the Rossby number and Burger number scale together

$$\text{Ro} \sim \text{Bu} = (L_d/L)^2 \ll 1, \quad (9.50)$$

so that the horizontal length scale for the planetary geostrophic flow is much larger than the deformation radius

$$L \gg L_d. \quad (9.51)$$

This assumption is consistent with dropping material acceleration in the velocity equation. Although the velocity equation is greatly simplified, we make no assumption concerning the thickness equation. Consequently, the free surface and bottom undulations are unconstrained with planetary geostrophic flows, so long as the flow maintains the hydrostatic balance.

In summary, the thickness equation for the planetary geostrophic fluid retains its full unapproximated form, whereas the velocity equation reduces to geostrophy. Reintroducing dimensions leads to the planetary geostrophic equations

$$f \hat{z} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} \quad \text{and} \quad h = \eta - \eta_b. \quad (9.52)$$

Since the Coriolis parameter retains its spatial dependence, and so includes the beta effect, the horizontal velocity field is divergent

$$f \nabla \cdot \mathbf{u} = -\beta (g/f) \partial_x \eta = -\beta v. \quad (9.53)$$

As shown in Exercise 9.1, the shallow water planetary geostrophic equations are equivalent to

$$f \hat{z} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{DQ}{Dt} = 0 \quad \text{with} \quad Q = f/h. \quad (9.54)$$

As seen in Chapter 10, the planetary geostrophic potential vorticity,  $Q = f/h$ , plays a central role in the mechanical interpretation of large scale flows in the ocean.

### 9.4.2 Reduced dynamical degrees of freedom

The layer thickness is the only prognostic field in the shallow water planetary geostrophic equations. From the thickness we diagnose the horizontal pressure gradient and the corresponding horizontal geostrophic velocity. Equivalently, the planetary geostrophic potential vorticity can serve as the single prognostic field from which the horizontal velocity is then diagnosed. This situation contrasts to the full set of shallow water equations, in which there are three prognostic fields: two horizontal velocity components plus the layer thickness.

As seen in our study of shallow water waves in VOLUME 4, the planetary geostrophic equations on a [beta-plane \( \$\beta\$ -plane\)](#) with a flat bottom and zero mean flow, support only a single wave mode: long planetary [Rossby waves](#). This situation contrasts to the shallow water equations, which support horizontal [gravity waves](#) plus the full suite of planetary Rossby waves. Evidently, by reducing the number of prognostic fields, the planetary geostrophic equations reduce the number of dynamical wave modes.

## 9.5 Shallow water quasi-geostrophy

In this section we develop the quasi-geostrophic equations for a single shallow water layer. Doing so requires far more work than for the planetary geostrophic equations. In particular, we use rudimentary asymptotic methods with the Rossby number acting as the relevant small non-dimensional parameter.

### 9.5.1 Quasi-geostrophic scaling

Quasi-geostrophic scaling is based on the following assumptions, with the first and second shared with planetary geostrophy whereas the remaining are distinct.

1. SMALL ROSSBY NUMBER:  $\text{Ro} \ll 1$ , which is fundamental to geostrophic scaling.
2. ADVECTIVE TIME SCALE:  $T \sim L/U$ ; that is, the time scale is determined by horizontal advection of the horizontal flow, which is how we scale time throughout this chapter. Notably, we do not distinguish between the two horizontal directions, so that  $L$  and  $U$  are scales for both the zonal and meridional flow. This assumption will be jettisoned when considering the dynamics of fronts and their stability in VOLUME 4, which then leads to the semi-geostrophic equations.
3. ORDER ROSSBY NUMBER BETA EFFECT:  $|\beta L| \ll |f_0|$ , which means that the Coriolis frequency does not vary much from its central value. To incorporate this assumption into the asymptotics, we expand the non-dimensional Coriolis parameter in terms of the Rossby number<sup>3</sup>

$$\hat{f} = f/f_0 = (1 + \beta y/f_0) \equiv (\hat{f}_0 + \text{Ro} \hat{\beta} \hat{y}). \quad (9.55)$$

Making use of the advective scaling for time as in equation (9.40) renders<sup>4</sup>

$$\hat{\beta} \hat{y} = \beta y/(\text{Ro} f_0) = T \beta y \quad \text{and} \quad \hat{f}_0 = f_0/f_0 = 1. \quad (9.56)$$

The non-dimensional Coriolis parameter (9.55) and the scaling (9.56) are motivated by assuming the horizontal scales of motion are on the same order as the deformation radius, and that the Coriolis parameter does not vary much from its central value. Quasi-geostrophy is thus formulated within the beta plane approximation.

4. BURGER NUMBER ORDER ONE:  $Bu \sim 1$ , which means that the horizontal scales of motion for the quasi-geostrophic flows are on the same order as the deformation radius,  $L \sim L_d$ .
5. ORDER ROSSBY NUMBER FREE SURFACE UNDULATIONS: From equation (9.36), an order unity Burger number means that undulations of the free surface height scale according to the Rossby number:  $\mathcal{H}/H = \text{Ro}$ , so that free surface height undulations are small.
6. ORDER ROSSBY NUMBER BOTTOM TOPOGRAPHY UNDULATIONS: As seen in Section 9.6.4, for the quasi-geostrophic potential vorticity to correspond to the small Rossby number version of the shallow water potential vorticity requires the topography undulations to scale as  $\mathcal{B}/H = \text{Ro}$ . The assumed scaling for the bottom topography undulation pairs with that for the free surface, so that the layer thickness undulations,  $h' = \eta' - \eta'_b$ , also scale as  $\text{Ro}$ . We are thus able to take a sensible Ro expansion of the  $1/h$  factor appearing in the shallow water potential vorticity (Section 9.6.4). It is useful to maintain a direct connection to the shallow water model as doing so helps to ensure that the resulting asymptotic theory is self-consistent. In particular, it ensures that quasi-geostrophic energetics are physically sensible since there is a direct lineage to the shallow water energetics.

The  $\mathcal{B}/H = \text{Ro}$  scaling is consistent with the assumption that the planetary beta effect is small. Together, the two scalings

$$\mathcal{B}/H = \text{Ro} \quad \text{and} \quad \mathcal{H}/H = \text{Ro} \quad (9.57)$$

mean that the effective beta effect (arising from both planetary and topographic variations; see Section 9.6.1 below) are small.

---

<sup>3</sup>One could conceive of another small parameter that scales the beta effect, but the resulting asymptotics would be more difficult to manage given the need to keep track of two small parameters.

<sup>4</sup>Although  $\hat{f}_0 = 1$  in equation (9.56), it is useful to retain this term as a placeholder in the manipulations to follow. In particular, it helps when reintroducing dimensions in Section 9.6.1.

### 9.5.2 Outlining the asymptotic method

To derive equations for the quasi-geostrophic shallow water model, we employ an asymptotic method with the Rossby number as the small parameter. Furthermore, we stop at the first nontrivial order. For this purpose, recall the non-dimensional shallow water equations from Section 9.3.7, and make use of the assumed  $Bu \sim 1$  scaling

$$\text{Ro} \frac{D\hat{\mathbf{u}}}{Dt} + (\hat{\mathbf{f}} \times \hat{\mathbf{u}}) = -\hat{\nabla}\hat{\eta} \quad (9.58a)$$

$$\text{Ro} \left[ \frac{\partial(\hat{\eta} - \hat{\eta}_b)}{\partial t} + \hat{\nabla} \cdot [(\hat{\eta} - \hat{\eta}_b) \hat{\mathbf{u}}] \right] = -\hat{\nabla} \cdot \hat{\mathbf{u}}. \quad (9.58b)$$

We brought the time independent bottom topography,  $\hat{\eta}_b$ , into the time derivative for the thickness equation, as doing so provides some symmetry with  $\hat{\eta}$ .<sup>5</sup>

#### Asymptotic expansion of the prognostic fields

Asymptotic methods are ideally suited for non-dimensional equations since we can unambiguously determine scales via the size of non-dimensional parameters. We here assume the Rossby number to be small, in which case we are led to perform an asymptotic expansion of the prognostic fields in terms of the Rossby number. There are three prognostic fields,  $\hat{u}, \hat{v}, \hat{\eta}$ , and corresponding vertical velocity,  $\hat{w}$ , which we assume can be written as an asymptotic series

$$\hat{u} = \hat{u}_0 + \text{Ro} \hat{u}_1 + \text{Ro}^2 \hat{u}_2 + \dots \quad (9.59a)$$

$$\hat{v} = \hat{v}_0 + \text{Ro} \hat{v}_1 + \text{Ro}^2 \hat{v}_2 + \dots \quad (9.59b)$$

$$\hat{w} = \hat{w}_0 + \text{Ro} \hat{w}_1 + \text{Ro}^2 \hat{w}_2 + \dots \quad (9.59c)$$

$$\hat{\eta} = \hat{\eta}_0 + \text{Ro} \hat{\eta}_1 + \text{Ro}^2 \hat{\eta}_2 + \dots \quad (9.59d)$$

We thus refer to the zeroth, first, second, etc. order of the asymptotic expansion. The three components of the velocity satisfy the non-divergence condition at each order

$$\nabla \cdot \hat{\mathbf{v}}_n = 0 \quad \forall n. \quad (9.60)$$

#### Practical goal

The practical goal of asymptotic analysis is to develop a closed set of prognostic equations for functions appearing in the asymptotic expansions (9.59a)-(9.59d). For our purposes, we are content to stop at the lowest nontrivial order, meaning the point at which there is a prognostic equation that provides a means to move the system forward in time. Motivation for asymptotic analysis is to produce an equation set offering a means to focus analysis on dynamics most active under the regime determined by the chosen non-dimensional parameters. Each higher order in asymptotic expansion generally produces more accurate solutions, and yet requires more complex algebraic manipulations. Hence, pursuit of higher order expansions should be undertaken only after first determining that the lower order equation set remains physically lacking.

---

<sup>5</sup>Also recall our nomenclature, whereby  $\mathbf{u}$  is the horizontal velocity, so that  $-\hat{\nabla} \cdot \hat{\mathbf{u}}$  is the non-dimensionalized horizontal convergence of the horizontal velocity. Also, since  $\eta$  is a function just of the horizontal directions (and time),  $\hat{\nabla}\hat{\eta}$  is the non-dimensionalized horizontal gradient of the free surface.

### Enabling the machinery

At this point we enable the machinery by “turning the crank”. To do so, insert the asymptotic expansions (9.59a)-(9.59d) into the non-dimensional partial differential equations (9.58a) and (9.58b). Since the Rossby number,  $\text{Ro}$ , is arbitrarily small, and all non-dimensional fields are order unity regardless their order, the only means to maintain self-consistency is for terms to balance at equal order in Rossby number. Hence, we do not mix terms from different orders of Rossby number. This point is fundamental to asymptotic methods.

Again, our goal is to establish a set of prognostic equations that allows us to evolve a state that is arbitrarily close to geostrophic balance. We anticipate that at zeroth order, the asymptotic method offers us just the geostrophic balance, which has no prognostic value. Hence, we need to go at least to order  $\text{Ro}^1$ , and hopefully no further since the algebraic tedium increases with order. With that anticipation and hope (and prior knowledge of what is sufficient), we only keep track of terms of order  $\text{Ro}^0$  and  $\text{Ro}^1$ , in which the momentum and continuity equations become

$$\text{Ro} \frac{D_0 \hat{\mathbf{u}}_0}{Dt} + (\hat{f}_0 + \text{Ro} \beta \hat{y}) \hat{\mathbf{z}} \times (\hat{\mathbf{u}}_0 + \text{Ro} \hat{\mathbf{u}}_1) = -\hat{\nabla}(\hat{\eta}_0 + \text{Ro} \hat{\eta}_1) \quad (9.61\text{a})$$

$$\text{Ro} \left[ \frac{\partial \hat{\eta}_0}{\partial \hat{t}} + \hat{\nabla} \cdot [(\hat{\eta}_0 - \hat{\eta}_b) \hat{\mathbf{u}}_0] + \hat{\nabla} \cdot \hat{\mathbf{u}}_1 \right] = -\hat{\nabla} \cdot \hat{\mathbf{u}}_0. \quad (9.61\text{b})$$

Note that the material time derivative in equation (9.61a) makes use of only the zeroth order geostrophic horizontal velocity

$$\frac{D_0}{Dt} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla}. \quad (9.62)$$

#### 9.5.3 Order $\text{Ro}^0$ asymptotic equations

Terms in equations (9.61a) and (9.61b) balancing at order  $\text{Ro}^0$  are given by

$$\hat{f}_0 \times \hat{\mathbf{u}}_0 = -\hat{\nabla} \hat{\eta}_0 \quad (9.63\text{a})$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}}_0 = 0. \quad (9.63\text{b})$$

The zeroth order velocity equation (9.63a) is the  $f$ -plane geostrophic balance. Furthermore, the vertical component to the curl of equation (9.63a) leads to the horizontal non-divergence condition,  $\hat{\nabla} \cdot \hat{\mathbf{u}}_0 = 0$ , which is identical to the zeroth order thickness equation (9.63b). Hence, the zeroth order horizontal velocity is given by  $f$ -plane geostrophy within a single shallow water layer.

#### The geostrophic streamfunction

Given the non-divergence condition (9.63b), the zeroth order velocity field can be written in terms of a **geostrophic streamfunction**

$$\hat{u}_0 = -\partial_y \hat{\psi}_0 \quad \text{and} \quad \hat{v}_0 = \partial_x \hat{\psi}_0 \quad \text{and} \quad \hat{\zeta}_0 = \hat{\nabla}^2 \hat{\psi}_0, \quad (9.64)$$

where the zeroth order streamfunction is the ratio of the zeroth order surface height to zeroth order Coriolis parameter

$$\hat{\psi}_0 = \hat{\eta}_0 / \hat{f}_0, \quad (9.65)$$

and we introduced the non-dimensional sgeostrophic relative vorticity

$$\hat{\zeta}_0 = \partial_{\hat{x}} \hat{v}_0 - \partial_{\hat{y}} \hat{u}_0. \quad (9.66)$$

The corresponding dimensionful quantities are

$$\psi = (U L) \hat{\psi}_0 \quad \text{and} \quad \zeta = (U/L) \hat{\zeta}_0. \quad (9.67)$$

We dropped the asymptotic label on the dimensional geostrophic streamfunction since it is the only streamfunction considered in quasi-geostrophy.<sup>6</sup>

### Horizontally non-divergent flow with an undulating free surface

Recall the discussion of horizontally non-divergent flow in Chapter 4, where the vertical velocity vanishes and the lid pressure renders a surface interface that is rigid and flat. We here also have a horizontally non-divergent flow, and yet it is within a single shallow water layer with a free surface that is not flat and that provides a pressure gradient in geostrophic balance with the Coriolis acceleration. Geostrophy and horizontal non-divergence enforces a free surface (and hence pressure field) that is related to relative vorticity through the Laplacian operator in equation (9.64). Furthermore, in the absence of horizontal convergence, there is no means for the flow to generate a vertical velocity, so that at this order we have

$$\hat{w}_0 = 0 \quad (9.68)$$

for shallow water quasi-geostrophy.

#### 9.5.4 Order $\text{Ro}^1$ asymptotic equations

Time derivatives appear at order  $\text{Ro}^1$ , with the velocity and thickness equations given by

$$\frac{D_0 \hat{\mathbf{u}}_0}{Dt} + \hat{f}_0 \hat{z} \times \hat{\mathbf{u}}_1 + \hat{\beta} \hat{y} \hat{z} \times \hat{\mathbf{u}}_0 = -\hat{\nabla} \hat{\eta}_1 \quad (9.69a)$$

$$\frac{D_0 (\hat{\eta}_0 - \hat{\eta}_b)}{Dt} = -\hat{\nabla} \cdot \hat{\mathbf{u}}_1. \quad (9.69b)$$

These equations appear to be unclosed because the evolution equation for zeroth order (geostrophic) terms is dependent on first order (ageostrophic) terms. However, the ageostrophic terms can be eliminated using two steps. First, we produce the vorticity equation from the momentum equation, which removes the ageostrophic pressure gradient,  $-\hat{\nabla} \hat{\eta}_1$ . Second, combining the vorticity equation and continuity equation eliminates the horizontal convergence of the ageostrophic velocity,  $-\hat{\nabla} \cdot \hat{\mathbf{u}}_1$ . The second step leads to the quasi-geostrophic potential vorticity equation. Although details are specific to the present study of shallow water quasi-geostrophy, similar steps are encountered in other balanced geophysical fluid systems.

---

<sup>6</sup>In some of our discussions, it is useful to introduce “g” or “ag” subscripts to distinguish  $\mathcal{O}(\text{Ro}^0)$  geostrophic terms from  $\mathcal{O}(\text{Ro}^1)$  ageostrophic terms.

### The geostrophic vorticity equation

Taking the curl of the momentum equation (9.69a) eliminates the ageostrophic pressure gradient,  $\widehat{\nabla}\widehat{\eta}_1$ , thus producing the vorticity equation

$$\partial_t \widehat{\zeta}_0 + (\widehat{\mathbf{u}}_0 \cdot \widehat{\nabla}) (\widehat{\zeta}_0 + \widehat{\beta} \widehat{y}) = -\widehat{f}_0 \widehat{\nabla} \cdot \widehat{\mathbf{u}}_1. \quad (9.70)$$

The right hand side term arises from stretching in the presence of planetary rotation, which can be seen by using continuity to express the right hand side as

$$-\widehat{f}_0 \widehat{\nabla} \cdot \widehat{\mathbf{u}}_1 = \widehat{f}_0 \partial_{\widehat{z}} \widehat{w}_1. \quad (9.71)$$

Evidently, to this order in Rossby number, stretching arises just from the planetary vorticity, with stretching due to relative vorticity appearing at a higher order. Since  $\widehat{\beta} \widehat{y}$  is time independent, we can write the vorticity equation (9.70) using the geostrophic material time derivative

$$\frac{D_0 (\widehat{\zeta}_0 + \widehat{\beta} \widehat{y})}{Dt} = -\widehat{f}_0 \widehat{\nabla} \cdot \widehat{\mathbf{u}}_1. \quad (9.72)$$

### The quasi-geostrophic potential vorticity equation

We need one more step to close the system since the evolution of zeroth order vorticity in equations (9.70) and (9.72) depends on vortex stretching induced by convergence of the first order velocity. To eliminate  $\widehat{\nabla} \cdot \widehat{\mathbf{u}}_1$ , we substitute from the thickness equation (9.69b), thus leading to a prognostic equation involving just zeroth order terms

$$\frac{\partial [\widehat{\zeta}_0 + \widehat{\beta} \widehat{y} - \widehat{f}_0 (\widehat{\eta}_0 - \widehat{\eta}_b)]}{\partial \widehat{t}} + \widehat{\mathbf{u}}_0 \cdot \widehat{\nabla} [\widehat{\zeta}_0 + \widehat{\beta} \widehat{y} - \widehat{f}_0 (\widehat{\eta}_0 - \widehat{\eta}_b)] = 0, \quad (9.73)$$

which can be written in the material form

$$\frac{D_0}{Dt} \left[ \widehat{\zeta}_0 + \widehat{\beta} \widehat{y} - \widehat{f}_0 (\widehat{\eta}_0 - \widehat{\eta}_b) \right] = 0. \quad (9.74)$$

Finally, we introduce the geostrophic streamfunction  $\widehat{\psi}_0 = \widehat{\eta}_0 / \widehat{f}_0$  (equation (9.65)) to render

$$\frac{D_0}{Dt} \left[ \widehat{\nabla}^2 \widehat{\psi}_0 + \widehat{\beta} \widehat{y} + \widehat{f}_0 \widehat{\eta}_b - \widehat{f}_0^2 \widehat{\psi}_0 \right] = 0. \quad (9.75)$$

Equation (9.75) is a statement of the material conservation of quasi-geostrophic potential vorticity (in its non-dimensional form), where material evolution is defined by the horizontal geostrophic currents (equation (9.62)). This equation is the culmination of our quest to derive a prognostic equation for the evolution of geostrophic flow. It enables us to evolve the geostrophic velocity and geostrophic free surface by accessing, but not explicitly determining, the leading order ageostrophic motions. Practical use of the quasi-geostrophic method is based on time stepping the quasi-geostrophic potential vorticity, and then inverting the potential vorticity equation to diagnose the streamfunction to then determine the geostrophic velocity and free surface. That inversion requires solving an elliptic partial differential equation with boundary conditions (see VOLUME 1). We present further remarks on inversion in Section 9.6.5.

## 9.6 Elements of shallow water quasi-geostrophy

The asymptotic analysis of Section 9.5 worked with non-dimensional quantities, which are suitable for determining the scales required for organizing terms in the asymptotic expansion. Now that we have worked through the details, we can make use of that effort to further expose the physical content of shallow water quasi-geostrophy theory. For that purpose, it is useful to reintroduce physical dimensions.

### 9.6.1 Dimensional potential vorticity and streamfunction

To introduce physical dimensions, we invert the relations used in Section 9.3.7

$$\hat{t} = t/T \quad (\hat{x}, \hat{y}) = (x, y)/L \quad \partial_{\hat{t}} = T \partial_t \quad \hat{\nabla} = L \nabla \quad (\hat{u}, \hat{v}) = (u, v)/U \quad (9.76a)$$

$$\hat{\eta} = \eta'/\mathcal{H} \quad \hat{\eta}_b = \eta'_b/\mathcal{B} \quad \mathcal{H} = H \text{Ro} \quad \mathcal{B} = H \text{Ro} \quad (9.76b)$$

$$\hat{f}_0 = f_0/f_o \quad \hat{\beta} \hat{y} = \beta y / (\text{Ro } f_o) = (L/U) \beta y \quad \hat{\zeta} = (L/U) \zeta = L^2 \nabla^2 \psi. \quad (9.76c)$$

We often drop asymptotic subscripts on dimensional terms to help reduce clutter, though at times it is useful to introduce a g or ag subscript to distinguish the geostrophic and ageostrophic components (e.g., see Section 9.6.6).

Starting from the non-dimensional quasi-geostrophic potential vorticity (9.74), the reintroduction of dimensions leads to

$$\hat{q} = \hat{\zeta}_0 + \hat{\beta} \hat{y} + \hat{f}_0 \hat{\eta}_b - \hat{f}_0 \hat{\eta}_0 \quad (9.77a)$$

$$= \frac{L}{U} (\zeta + \beta y) + \frac{\eta'_b}{\mathcal{B}} - \frac{\eta'}{\mathcal{H}} \quad (9.77b)$$

$$= \frac{L}{U} (\zeta + \beta y) + \frac{\eta'_b}{H \text{Ro}} - \frac{\eta'}{H \text{Ro}} \quad (9.77c)$$

$$= \frac{L}{U} \left[ \zeta + \beta y - \frac{f_o (\eta' - \eta'_b)}{H} \right] \quad (9.77d)$$

$$= \frac{L}{U} \left[ \zeta + \beta y - \frac{g (\eta' - \eta'_b)}{f_o} \frac{1}{L_d^2} \right] \quad (9.77e)$$

$$= \frac{L}{U} [\zeta + \beta y - L_d^{-2} (\psi - \psi_b)], \quad (9.77f)$$

where  $L_d = \sqrt{g H}/f$  is the shallow water deformation radius (equation (9.31)), and we introduced the geostrophic streamfunction<sup>7</sup>

$$\psi = (g/f_o) \eta'. \quad (9.78)$$

We also wrote the contribution from topography as

$$\psi_b = (g/f_o) \eta'_b, \quad (9.79)$$

which is a static field. We are thus led to the dimensionful quasi-geostrophic potential vorticity for a single shallow water layer

$$q = f_o (1 + \text{Ro } \hat{q}) = (\zeta + f) - L_d^{-2} (\psi - \psi_b) = f + \psi_b L_d^{-2} + (\nabla^2 - L_d^{-2}) \psi. \quad (9.80)$$

<sup>7</sup>The geostrophic streamfunction is arbitrary up to a constant. For example, [Vallis \(2017\)](#) defines the geostrophic streamfunction in his equation (5.69) as  $\psi_{\text{vallis}} = g \eta/f_o = \psi + g H/f_o$ , which differs by the constant  $g H/f_o$ .

We took the liberty of adding the constant,  $f_0$ , to the QG potential vorticity, which does not alter the dynamics but does allow us to introduce the planetary vorticity,  $f = f_0 + \beta y$ . Evidently, the dynamically relevant portions of  $q$  appear at order  $\text{Ro}$ , as expected since the zeroth order theory is  $f$ -plane geostrophy, which provides no prognostic capability.

### 9.6.2 Contributions to quasi-geostrophic potential vorticity

The quasi-geostrophic potential vorticity (9.80) has three main contributions

$$q = f + \zeta - f_0 h'/H, \quad (9.81)$$

where we wrote

$$L_d^{-2} (\psi - \psi_b) = f_0^2 / (g H) (g/f_0) (\eta' - \eta'_b) = f_0 h'/H, \quad (9.82)$$

with  $h'$  the undulations in the layer thickness due to undulations in the free surface and bottom topography. Heuristically, we can connect  $q$  to the shallow water  $Q = (f + \zeta)/h$  by

$$(f + \zeta)/h \approx (f + \zeta) (1 - h'/H) \approx H^{-1} (f + \zeta - f_0 h'/H), \quad (9.83)$$

where we assumed  $\beta y$ ,  $\zeta$ , and  $h'$  are order  $\text{Ro}$  whereas  $f_0$  is order unity. We pursue this expansion more formally in Section 9.6.4.

It is notable that the quasi-geostrophic potential vorticity,  $q$ , is determined by the free surface height, the bottom topography, and the Coriolis parameter. That is, once  $f$ ,  $\eta_b$ , and  $\eta$  are known, then we have  $q$  and thus, through inversion, the geostrophic flow is determined. This is a rather remarkable result that embodies the following variety of physical processes contributing to potential vorticity, and hence to quasi-geostrophic dynamics.

- **PLANETARY VORTICITY:** The contribution  $f = f_0 + \beta y$  arises from planetary vorticity, with the dynamically relevant contribution for quasi-geostrophy arising just from the  $\beta y$  term. The difference,  $q - f$ , is sometimes referred to as the **relative potential vorticity**, in analog to the relative vorticity,  $\zeta = \zeta_a - f$ .
- **GEOSTROPHIC RELATIVE VORTICITY:**  $\zeta = \nabla^2 \psi = (g/f_0) \nabla^2 \eta$  is the relative vorticity of the geostrophic flow. The Laplacian operator emphasizes small spatial scales, so that the relative vorticity is most important at scales at or smaller than the deformation radius.
- **EFFECTIVE BETA:** The contribution from the gradient of planetary vorticity is given by  $\beta y = \mathbf{x} \cdot \nabla f$ . Likewise, the contribution from topography is given by

$$\psi_b L_d^{-2} = f_0 \eta'_b / H = (f_0/H) (\eta_b - \bar{\eta}_b) \approx (f_0/H) \mathbf{x} \cdot \nabla \eta_b, \quad (9.84)$$

where the approximate expression made use of a Taylor series. These two contributions can be combined into an effective beta

$$\beta y + \psi_b L_d^{-2} = \mathbf{x} \cdot \nabla (f + f_0 \eta_b / H) \approx \mathbf{x} \cdot (H - \eta_b) \nabla [f/(H - \eta_b)], \quad (9.85)$$

with the final approximate expression connecting to the effective beta discussed in Section 5.5.1.

- **VERTICAL STRETCHING:** As discussed in Section 1.2, shallow water fluids move as vertical columns that can expand (stretch) and contract (squash). Hence, the term  $(f_0/H) h'$

accounts for the contribution to potential vorticity from column stretching and squashing. It is most important for scales at or larger than those where the relative vorticity is important; i.e., at or larger than the deformation radius.

- FLOW REGIMES RELATIVE TO THE DEFORMATION RADIUS: The term  $(\nabla^2 - L_d^{-2}) \psi$  signals two regimes as determined by the deformation radius. For lateral scales on the order of the deformation radius, both the relative vorticity and vortex stretching make equal contributions to the potential vorticity. For smaller scales, relative vorticity is more important whereas for larger scales vortex stretching dominates.

### 9.6.3 Connecting to background (resting state) potential vorticity

For some purposes, particularly when studying Rossby waves in VOLUME 4, we find it useful to unpack the material time derivative,

$$Dq/Dt = \partial_t q + \hat{z} \cdot (\nabla \psi \times \nabla q), \quad (9.86)$$

to expose the role of the potential vorticity contained in a resting fluid. For this purpose, expand the gradient

$$\nabla q = \nabla(\nabla^2 \psi) - \nabla \psi / L_d^2 + \nabla f + \nabla \psi_b / L_d^2, \quad (9.87)$$

so that

$$\nabla \psi \times \nabla q = \nabla \psi \times \nabla(\nabla^2 \psi) + \nabla \psi \times (\nabla f + \nabla \psi_b / L_d^2). \quad (9.88)$$

The second term on the right hand side can be written in terms of the resting state potential vorticity

$$\nabla f + \nabla \psi_b / L_d^2 = \nabla f + (f_0 / H) \nabla \eta'_b = \nabla f - (f_0 / H) \nabla(H - \eta'_b) \approx H \nabla(f / H_r) = H \nabla Q_r, \quad (9.89)$$

where  $H_r = H - \eta'_b$  is the thickness of the resting fluid, and  $Q_r$  is the potential vorticity of the resting fluid. The approximation in the penultimate step follows from assuming  $\eta'_b / H = \mathcal{O}(\text{Ro})$ , as per the quasi-geostrophic scaling in Section 9.5.1.

Bringing the pieces together leads us to write the material time evolution of quasi-geostrophic potential vorticity as

$$(f_0 / g) Dq/Dt = \partial_t [(L_d^{-2} - \nabla^2) \eta'] - H \hat{z} \cdot (\nabla \eta' \times \nabla Q_r) - (g / f_0) \hat{z} \cdot [\nabla \eta' \times \nabla(\nabla^2 \eta')]. \quad (9.90)$$

The second term on the right hand side arises from geostrophic advection of the potential vorticity in the resting fluid. This linear term is fundamental to the Rossby wave dispersion relation studied in VOLUME 4. The nonlinear term on the right hand side arises from advection of the geostrophic relative vorticity by the geostrophic velocity. This term is generally ignored when studying small amplitude wave fluctuations. Furthermore, it is notable that this term vanishes identically for plane waves, as we show in VOLUME 4.

### 9.6.4 Connecting to Rossby's shallow water potential vorticity

We here determine how potential vorticity for shallow water quasi-geostrophy relates to Rossby's shallow water potential vorticity studied in Chapter 5. For that purpose, recall that the potential vorticity for a single layer of shallow water fluid is (Section 5.3)

$$Q = \frac{f + \zeta}{h}, \quad (9.91)$$

where  $h = H + \eta' - \eta'_b$  (equation (9.4)) is the layer thickness.

To connect to the quasi-geostrophic potential vorticity, we non-dimensionalize the potential vorticity and then perform an asymptotic expansion to order  $\text{Ro}^1$ . For this purpose, use the scaling relations from Section 9.3.7 to write the geostrophic relative vorticity as

$$\zeta = \nabla^2\psi = (g/f_0) \nabla^2\eta' = (g\mathcal{H})/(f_0 L^2) \hat{\nabla}^2\hat{\eta} = (g H \text{Ro})/(f_0 L^2) \hat{\zeta}_0 = f_0 \text{Ro} (L_d/L)^2 \hat{\zeta}_0, \quad (9.92)$$

as well as

$$h = H + H \text{Ro} (\hat{\eta} - \hat{\eta}_b) \quad (9.93a)$$

$$f = f_0 (\hat{f}_0 + \text{Ro} \hat{\beta} \hat{y}). \quad (9.93b)$$

Taking the ratio and expanding to order  $\text{Ro}^1$  leads to

$$\frac{f + \zeta}{h} = \frac{f_0}{H} \left[ \frac{\hat{f}_0 + \text{Ro} (L_d/L)^2 \hat{\zeta}_0 + \text{Ro} \hat{\beta} \hat{y}}{1 + \text{Ro} (\hat{\eta} - \hat{\eta}_b)} \right] \quad (9.94a)$$

$$\approx \frac{f_0}{H} + \frac{f_0 \text{Ro}}{H} \left[ -\hat{f}_0 (\hat{\eta} - \hat{\eta}_b) + (L_d/L)^2 \hat{\zeta}_0 + \hat{\beta} \hat{y} \right] \quad (9.94b)$$

$$= \frac{f_0}{H} + \frac{f_0 \text{Ro}}{H} \left[ (L_d/L)^2 \hat{\zeta}_0 + \hat{\beta} \hat{y} - \hat{f}_0 (\hat{\eta} - \hat{\eta}_b) \right] \quad (9.94c)$$

$$= \frac{f_0}{H} + \frac{f_0 \text{Ro}}{H} \left[ \frac{\zeta}{f_0 \text{Ro}} + \frac{\beta y}{f_0 \text{Ro}} - \frac{\eta'}{\mathcal{H}} + \frac{\eta'_b}{\mathcal{B}} \right] \quad (9.94d)$$

$$= \frac{1}{H} \left[ \zeta + f_0 + \beta y - \frac{f_0 (\eta' - \eta'_b)}{H} \right] \quad (9.94e)$$

$$= \hat{q} U / (H L) \quad (9.94f)$$

$$= q / H. \quad (9.94g)$$

We are thus led to the relation between the shallow water potential vorticity and the quasi-geostrophic shallow water potential vorticity

$$\frac{f + \zeta}{h} = \frac{q}{H} [1 + \mathcal{O}(\text{Ro}^2)]. \quad (9.95)$$

As noted in Section 9.5.1, this connection between the potential vorticities only holds when assuming deviations in both the bottom topography and free surface scale according to the Rossby number,  $\eta' = \mathcal{H} \text{Ro}$  and  $\eta'_b = \mathcal{B} \text{Ro}$ , thus allowing us to combine  $\hat{\eta}$  and  $\hat{\eta}_b$  in the thickness equation (9.93a). That is, we must assume that both the planetary beta and topographic beta are on the order of  $\text{Ro}^1$ .

The identity (9.95) is a consequence of the asymptotic expansion of the velocity and thickness equations. We could alternatively invert our development of quasi-geostrophy by using the quasi-geostrophic potential vorticity as the basis for deriving the governing equations. We pursue that approach to derive the layered quasi-geostrophic equations in Section 9.6.10.

### 9.6.5 Geostrophic flow via potential vorticity inversion

Here are the key pieces to single layer shallow water quasi-geostrophic theory:

$$\mathbf{u} = \hat{\mathbf{z}} \times \nabla\psi \quad \text{and} \quad \zeta = \hat{\mathbf{z}} \cdot \nabla \times \mathbf{u} = \nabla^2\psi \quad \text{and} \quad q = f + \zeta - L_d^{-2} (\psi - \psi_b) \quad (9.96a)$$

$$\psi = (g/f_0) \eta' \quad \text{and} \quad \psi_b = (g/f_0) \eta'_b \quad \text{and} \quad L_d = \sqrt{g H} / f_0. \quad (9.96b)$$

Note that we can choose to use the full free surface height,  $\eta = \eta' + \bar{\eta}$ , and bottom topography,  $\eta_b = \eta'_b - \bar{\eta}_b$ , since doing so only adds a dynamically irrelevant constant to  $\psi$ ,  $\psi_b$ , and  $q$ .

As seen in Section 9.5.4, evolution of the geostrophic flow is determined by material evolution of the quasi-geostrophic potential vorticity following the horizontal geostrophic flow. This evolution equation takes on the dimensional form

$$\frac{Dq}{Dt} = (\partial_t + \mathbf{u} \cdot \nabla) q = \partial_t q + \hat{\mathbf{z}} \cdot (\nabla \psi \times \nabla q) = \partial_t q + J(\psi, q), \quad (9.97)$$

where the final equality introduced the Jacobian operator

$$J(\psi, q) = \hat{\mathbf{z}} \cdot (\nabla \psi \times \nabla q) = \partial_x \psi \partial_y q - \partial_y \psi \partial_x q. \quad (9.98)$$

For a perfect fluid, the quasi-geostrophic potential vorticity is materially constant, whereas more general flows have forcing and dissipation so that

$$\frac{Dq}{Dt} = \text{forcing} - \text{dissipation}. \quad (9.99)$$

As a method to evolve the quasi-geostrophic state forward in time, we first update the potential vorticity to a new time step. Thereafter, we must diagnose the streamfunction,  $\psi = g \eta / f$ , by solving the elliptic boundary value problem for  $\psi$ ,

$$(\nabla^2 - L_d^{-2}) \psi = q - f - \psi_b / L_d^2, \quad (9.100)$$

at which point we have the updated free surface and updated velocity. This same inversion method was outlined for time stepping the two-dimensional non-divergent barotropic model in Section 4.3.6. It is here referred to as **potential vorticity inversion**.

The streamfunction equation (9.100) is a forced **Helmholtz equation**, which is an elliptic partial differential equation.<sup>8</sup> While  $q$  results from taking derivatives of  $\psi$ , the streamfunction results from taking integrals of  $q$ . Consequently, maps of the potential vorticity have more structure at the high wave numbers (small scales) relative to the streamfunction. Writing the potential vorticity in equation (9.100) as

$$(\nabla^2 - L_d^{-2}) (\psi - \psi_b) = q - f - L_d^{-2} \nabla^2 \psi_b \quad (9.101)$$

offers us further interpretations, based on properties of the horizontal Laplacian operator. Namely, regions of relatively low  $\psi - \psi_b$  correspond to regions of relatively high  $q - f - L_d^{-2} \nabla^2 \psi_b$ , and vice versa. For a flat bottom where  $\psi_b$  is a constant, relatively low sea level regions (low pressure regions) correspond to regions of relatively high  $q - f$ . Finally, on the  $f$ -plane with a flat bottom, relatively low pressure regions correspond to relatively high potential vorticity regions.

### 9.6.6 Evolution of quasi-geostrophic vorticity, velocity, and free surface

When making use of the potential vorticity inversion method from Section 9.6.5, we are unconcerned with the velocity equation and the free surface equation. Rather, it is sufficient to invert the potential vorticity equation (9.100) to find the streamfunction, from which both the velocity and the free surface are diagnosed. Even so, there are physical insights resulting from

---

<sup>8</sup>We can mathematically formulate the inversion problem using the Green's function method for elliptic operators studied in VOLUME 1.

analysis of the quasi-geostrophic velocity, vorticity, and free surface. We here discuss these equations and identify a notable ambiguity in their specification.

### Quasi-geostrophic absolute vorticity equation

Reintroducing dimensions to the non-dimensional quasi-geostrophic vorticity equation (9.70) leads to

$$(\partial_t + \mathbf{u}_g \cdot \nabla) (\zeta_g + f) = -f_o \nabla \cdot \mathbf{u}_{ag}, \quad (9.102)$$

which takes on the equivalent form using the geostrophic material time derivative<sup>9</sup>

$$\frac{D_g(\zeta_g + f)}{Dt} = -f_o \nabla \cdot \mathbf{u}_{ag} \quad \text{with} \quad \frac{D_g}{Dt} = \partial_t + \mathbf{u}_g \cdot \nabla. \quad (9.103)$$

In these equations we introduced subscripts to distinguish the geostrophic velocity and corresponding free surface height,

$$f_o \hat{\mathbf{z}} \times \mathbf{u}_g = -g \nabla \eta_g, \quad (9.104)$$

from the ageostrophic velocity,  $\mathbf{u}_{ag}$ , and ageostrophic free surface height,  $\eta_{ag}$ . We have already discussed the vorticity equation (9.103) in its non-dimensional form. There, we emphasized the central role of the horizontal convergence of the ageostrophic flow in providing a source for the geostrophic material time evolution of the geostrophic absolute vorticity.

### Quasi-geostrophic velocity and free surface equations: Version I

The velocity and free surface equations are determined by the order  $Ro^1$  equations derived in Section 9.5.4. These equations take on the following dimensional form

$$\partial_t \mathbf{u}_g + (\mathbf{u}_g \cdot \nabla) \mathbf{u}_g + (f_o + \beta y) \hat{\mathbf{z}} \times \mathbf{u}_g + f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} = -g \nabla (\eta_g + \eta_{ag}) \quad (9.105a)$$

$$\partial_t (\eta_g - \eta_b) + (\mathbf{u}_g \cdot \nabla) (\eta_g - \eta_b) = -H \nabla \cdot \mathbf{u}_{ag}. \quad (9.105b)$$

The ageostrophic velocity generally has a nonzero horizontal convergence. Even so, it is three dimensionally non-divergent since the fluid layer has constant density. Hence, vertical component to the ageostrophic velocity satisfies

$$\mathbf{v}_{ag} = \mathbf{u}_{ag} + \hat{\mathbf{z}} w_{ag} \quad \text{with} \quad \nabla \cdot \mathbf{v}_{ag} = \partial_x u_{ag} + \partial_y v_{ag} + \partial_z w_{ag} = 0. \quad (9.106)$$

Cancelling the geostrophic balance (9.104) from the velocity equation (9.105a), making use of continuity (9.106), and introducing the geostrophic material time derivative (equation (9.103)) leads to

$$\frac{D_g \mathbf{u}_g}{Dt} + (\beta y) \hat{\mathbf{z}} \times \mathbf{u}_g + f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} = -g \nabla \eta_{ag} \quad (9.107a)$$

$$\frac{D_g (\eta_g - \eta_b)}{Dt} = H \partial_z w_{ag}. \quad (9.107b)$$

### Quasi-geostrophic velocity and free surface equations: Version II

As seen in Section 9.5.4, we eliminate the ageostrophic pressure gradient,  $-g \nabla \eta_{ag}$ , by forming the quasi-geostrophic vorticity equation. The resulting source term arises from convergence

<sup>9</sup>It is important to remember our notation for the velocity, in which  $\mathbf{u}_g$  is horizontal so that  $\mathbf{u}_g \cdot \nabla = u_g \partial_x + v_g \partial_y$ . See also equation (9.106).

of the ageostrophic flow,  $-f_o \nabla \cdot \mathbf{u}_{ag}$ , with continuity equating this vorticity source to vortex stretching by the ageostrophic flow in a rotating reference frame,  $-f_o \nabla \cdot \mathbf{u}_{ag} = f_o \partial_z w$  (see equation (9.103)). Evidently, the vorticity equation is unchanged if the ageostrophic free surface is modified by an arbitrary scalar function, since that function is annihilated when taking the curl to form the vorticity equation. Likewise, the ageostrophic velocity is arbitrary up to a horizontally non-divergent velocity, since that extra non-divergent velocity also plays no role in the vorticity equation.

The ambiguity in specifying  $\eta_{ag}$  and  $\mathbf{u}_{ag}$  is constrained, as revealed by taking the divergence of the velocity equation (9.107a) and noting that  $\nabla \cdot \mathbf{u}_g = 0$ , in which case

$$g \nabla^2 \eta_{ag} = 2 J(u_g, v_g) + \beta(y \zeta_g - u_g) + f_o \zeta_{ag}, \quad (9.108)$$

where we used

$$\nabla \cdot (\beta y \hat{\mathbf{z}} \times \mathbf{u}_g) = -\beta y \zeta_g + \beta u_g \quad (9.109a)$$

$$\nabla \cdot (f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag}) = -f_o \zeta_{ag} \quad (9.109b)$$

$$\nabla \cdot \partial_t \mathbf{u}_g = \partial_t (\nabla \cdot \mathbf{u}_g) = 0 \quad (9.109c)$$

$$\nabla \cdot [\partial_t \mathbf{u}_g + (\mathbf{u}_g \cdot \nabla) \mathbf{u}_g] = -2 J(u_g, v_g), \quad (9.109d)$$

with

$$\zeta_{ag} = \partial_x v_{ag} - \partial_y u_{ag} \quad (9.110)$$

the relative vorticity contained in the horizontal ageostrophic flow. The constraint (9.108) means that adding a horizontally non-divergent velocity,  $\tilde{\mathbf{u}}$ , to the ageostrophic velocity,  $\mathbf{u}_{ag}$ , requires a corresponding modification to the free surface via  $\tilde{\eta}$  added to  $\eta_{ag}$ , in which  $\tilde{\eta}$  satisfies the Poisson equation

$$g \nabla^2 \tilde{\eta} = f_o \hat{\mathbf{z}} \cdot (\nabla \times \tilde{\mathbf{u}}). \quad (9.111)$$

Again, these ambiguities arise in quasi-geostrophy since potential vorticity inversion is sufficient to specify evolution of the flow and free surface, with  $\eta_{ag}$  and  $\mathbf{u}_{ag}$  having no affect on the quasi-geostrophic potential vorticity. This freedom arises from a gauge symmetry. Following Section 6.3 of [Holton and Hakim \(2013\)](#), a convenient gauge choice sets  $\eta_{ag} = 0$  so that all ageostrophic effects live within the ageostrophic flow,  $\mathbf{u}_{ag} + \hat{\mathbf{z}} w_{ag}$ . In this case the momentum equation (9.107a) takes on the particularly tidy form

$$\frac{D_g \mathbf{u}_g}{Dt} + (\beta y) \hat{\mathbf{z}} \times \mathbf{u}_g + f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} = 0, \quad (9.112)$$

so that the  $f$ -plane ageostrophic Coriolis acceleration is the only means for the ageostrophic flow to affect the geostrophic flow. The  $\eta_{ag} = 0$  gauge choice amounts to expanding the velocity fields in the asymptotic series (9.59a)–(9.59c), whereas the free surface appears only at zeroth order.<sup>10</sup> The corresponding divergence equation (9.108) reveals that the ageostrophic relative vorticity is diagnosed from the geostrophic flow according to

$$f_o \zeta_{ag} = -2 J(u_g, v_g) - \beta(y \zeta_g - u_g). \quad (9.113)$$

<sup>10</sup>We know of no other gauge choice discussed in the literature.

### 9.6.7 Unpacking quasi-geostrophic evolution

Evolution of the quasi-geostrophic state occurs via the geostrophic material evolution of potential vorticity and then potential vorticity inversion (Section 9.6.5). This perspective is complete and elegant, and it encompasses a wealth of processes. To help expose those processes, and thus to reveal the fine tuned nature of quasi-geostrophic evolution, consider the evolution equation (9.112) for the geostrophic velocity along with equation (9.107b) for the free surface, written in the form

$$\frac{D_g \mathbf{u}_g}{Dt} = -(\beta y) \hat{\mathbf{z}} \times \mathbf{u}_g - f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} \quad (9.114a)$$

$$\frac{D_g \eta_g}{Dt} = \frac{D_g \eta_b}{Dt} + H \partial_z w_{ag}. \quad (9.114b)$$

The quasi-geostrophic flow retains a geostrophically balanced state,  $f_o \hat{\mathbf{z}} \times \mathbf{u}_g = -g \nabla \eta_g$ , at each point in space and for each time instance.<sup>11</sup> Consequently, geostrophic balance is maintained for an observer following a fluid particle moving with the horizontal geostrophic velocity,  $\mathbf{u}_g$ , so that

$$\frac{D_g}{Dt} (f_o \hat{\mathbf{z}} \times \mathbf{u}_g + g \nabla \eta_g) = 0. \quad (9.115)$$

Performing the material time derivatives in equation (9.115), and making use of the evolution equations (9.114a) and (9.114b), lead to the balance

$$\underbrace{f_o \beta y \mathbf{u}_g + g \mathbf{Q}^{(\eta_g)}}_{\text{geostrophic}} + \underbrace{g \nabla (\mathbf{u}_g \cdot \nabla \eta_b) + f_o^2 \mathbf{u}_{ag} + g H \nabla \partial_z w_{ag}}_{\text{ageostrophic}} = 0, \quad (9.116)$$

where we used

$$D_g (\nabla \eta_g) / Dt = \nabla (D_g \eta_g / Dt) + \mathbf{Q}^{(\eta_g)} = \nabla (\mathbf{u}_g \cdot \nabla \eta_b) + H \nabla (\partial_z w_{ag}) + \mathbf{Q}^{(\eta_g)}. \quad (9.117)$$

$\mathbf{Q}^{(\eta_g)}$  is a vector arising from the nonlinear coupling of horizontal shears in the geostrophic flow with gradients in the free surface

$$g \mathbf{Q}^{(\eta_g)} = -g \hat{\mathbf{x}} \partial_x \mathbf{u}_g \cdot \nabla \eta_g - g \hat{\mathbf{y}} \partial_y \mathbf{u}_g \cdot \nabla \eta_g \quad (9.118a)$$

$$= f_o \hat{\mathbf{x}} [\partial_x \mathbf{u}_g \cdot (\hat{\mathbf{z}} \times \mathbf{u}_g)] + f_o \hat{\mathbf{y}} [\partial_y \mathbf{u}_g \cdot (\hat{\mathbf{z}} \times \mathbf{u}_g)] \quad (9.118b)$$

$$= f_o \hat{\mathbf{x}} [\hat{\mathbf{z}} \cdot (\mathbf{u}_g \times \partial_x \mathbf{u}_g)] + f_o \hat{\mathbf{y}} [\hat{\mathbf{z}} \cdot (\mathbf{u}_g \times \partial_y \mathbf{u}_g)], \quad (9.118c)$$

where the second equality used the geostrophic balance. If any process contributes to changes in one of the terms in the balance (9.116), then the other terms compensate to retain the balance and thus to retain geostrophy.

The special case of a flat bottom  $\beta$ -plane is of particular interest, in which case the nonlinear geostrophic term,  $\mathbf{Q}^{(\eta_g)}$ , balances the ageostrophic terms

$$g \mathbf{Q}^{(\eta_g)} = -[f_o^2 \mathbf{u}_{ag} + g H \nabla \partial_z w_{ag}]. \quad (9.119)$$

For example, if  $\mathbf{Q}^{(\eta_g)}$  grows, say from a wave or instability, ageostrophic processes arise to compensate, thus preserving the balance (9.119) and, by extension, preserving geostrophy. From its definition (9.118a), for  $\mathbf{Q}^{(\eta_g)}$  to be nonzero requires the geostrophic flow to not be

<sup>11</sup>Maintaining geostrophic balance at each point in space and time is a constraint that is analogous to the non-divergent flow condition maintained by a Boussinesq ocean. In that case, the pressure field satisfies an elliptic partial differential equation that ensures the flow remains non-divergent.

parallel to its horizontal shear. For example, geostrophic flow along a straight front (say, with  $\mathbf{u}_g = \hat{\mathbf{x}} u_g(y)$  and  $\nabla \eta_g = \hat{\mathbf{y}} \partial_y \eta_g$ ) has  $\mathbf{Q}^{(\eta_g)} = 0$ , whereas  $\mathbf{Q}^{(\eta_g)} \neq 0$  for fronts with curvature. [Hoskins \(1975\)](#) and Section 6.5 of [Holton and Hakim \(2013\)](#) provide a thorough discussion of this term and its role in the dynamics of atmospheric fronts.

### 9.6.8 Energetics of quasi-geostrophic flows

To derive the energetic balances within an unforced quasi-geostrophic shallow water flow, we start by multiplying the potential vorticity equation by the streamfunction

$$\psi \partial_t q + \psi \mathbf{u} \cdot \nabla q = 0. \quad (9.120)$$

The time tendency term can be written

$$\psi \partial_t q = \psi (\partial_t \zeta - L_d^{-2} \partial_t \psi) \quad (9.121a)$$

$$= \psi \nabla \cdot (\nabla \partial_t \psi) - (1/2) L_d^{-2} \partial_t (\psi)^2 \quad (9.121b)$$

$$= \nabla \cdot (\psi \nabla \partial_t \psi) - \nabla \psi \cdot \nabla (\partial_t \psi) - (1/2) L_d^{-2} \partial_t (\psi)^2 \quad (9.121c)$$

$$= \nabla \cdot (\psi \nabla \partial_t \psi) - \partial_t (\mathbf{u} \cdot \mathbf{u} + L_d^{-2} \psi^2)/2, \quad (9.121d)$$

whereas the advection term can be written

$$\psi \mathbf{u} \cdot \nabla q = \psi \nabla \cdot (\mathbf{u} q) = \nabla \cdot (\psi \mathbf{u} q) - \nabla \psi \cdot \mathbf{u} q = \nabla \cdot (\psi \mathbf{u} q), \quad (9.122)$$

where we used

$$\nabla \psi \cdot \mathbf{u} = \nabla \psi \cdot (\hat{\mathbf{z}} \times \nabla \psi) = 0. \quad (9.123)$$

We are thus led to the energy equation for shallow water quasi-geostrophy

$$\partial_t (\mathbf{u} \cdot \mathbf{u} + L_d^{-2} \psi^2)/2 = \nabla \cdot (\psi \nabla \partial_t \psi + \mathbf{u} q \psi). \quad (9.124)$$

The first term on the left hand side is the kinetic energy per mass. The second term is the available potential energy per mass, where the available potential energy is proportional to the free surface height undulation (Section 2.5.6)

$$L_d^{-2} \psi^2 = f_o^2 / (g H) (g^2 / f_o^2) (\eta')^2 = (g/H) (\eta')^2 = c_{\text{grav}}^2 (\eta'/H)^2, \quad (9.125)$$

where the final step introduced the shallow water gravity wave speed  $c_{\text{grav}} = \sqrt{g H}$ .

### 9.6.9 Considering topography to be $\mathcal{O}(\text{Ro}^0)$

In our treatment of quasi-geostrophy, we assumed the topographic undulations to be on the order of  $\text{Ro}$ . Doing so ensured asymptotic consistency by combining  $\eta'$  and  $\eta_b'$  into an order Rossby number fluctuation of the layer thickness. We just encountered this need in Section 9.6.4 when connecting quasi-geostrophic potential vorticity to Ertel potential vorticity (see also Section 9.5.1). What happens if we allow for arbitrarily large topographic undulations? It is not uncommon to examine how an asymptotic theory performs when outside of its formal regime of validity, particularly with the advent of numerical codes to facilitate such studies. In many cases the theories continue to provide provocative, and sometimes physically relevant, information.

To see what happens, consider the quasi-geostrophic PV equation

$$(\partial_t + \mathbf{u} \cdot \nabla) [\zeta + \beta y + L_d^{-2} (\psi_b - \psi)] = 0. \quad (9.126)$$

If  $\psi_b$  order  $\text{Ro}^0$ , whereas the other terms are order  $\text{Ro}$ , then to leading order the potential vorticity is given by the static term,  $L_d^{-2} \psi_b$ , so that material conservation of potential vorticity reduces to

$$\mathbf{u} \cdot \nabla \psi_b = 0 \implies \mathbf{u} \cdot \nabla \eta_b = 0. \quad (9.127)$$

This constraint means that the  $f$ -plane geostrophic flow is constrained to flow along lines of constant topography (isobaths), in which case the geostrophic streamfunction satisfies

$$(f/g) \mathbf{u} \cdot \nabla \eta_b = \hat{\mathbf{z}} \cdot (\nabla \eta \times \nabla \eta_b) \equiv J(\eta, \eta_b) = 0. \quad (9.128)$$

Evidently, the order unity bottom topography undulations provide a constraint on the quasi-geostrophic flow, making the flow align with the bottom topography and in turn aligning surface height undulations with bottom undulations. We uncovered this constraint in our analysis of topographic form stress in Section 5.7.6. We also discussed this flow as a particular realization of the two-dimensional non-divergent barotropic flow in Section 4.2.4.

### 9.6.10 Two layer quasi-geostrophy

In Section 1.4 we developed the equations for an adiabatic stacked shallow water model. We here specialize those equations to a two-layer quasi-geostrophic model, with extensions to multiple layers following straightforwardly. Rather than pursue the formal asymptotic methods used previously, we here make use of our observation in Section 9.6.4 concerning the connection between shallow water and quasi-geostrophic potential vorticities.

To get started, recall from Section 5.3.6 that the shallow water potential vorticity for an arbitrary layer, labelled by the index  $\kappa$ , is given by

$$Q_\kappa = \frac{f + \zeta_\kappa}{h_\kappa}, \quad (9.129)$$

As in Section 9.6.2, we assume  $\beta y$ ,  $\zeta_\kappa$ , and  $h'_\kappa$  scale as  $\text{Ro}$ , in which case we have the quasi-geostrophic potential vorticity for each layer

$$q_\kappa = f + \zeta_\kappa - f_0 h'_\kappa / H, \quad (9.130)$$

where  $\zeta_\kappa$  is here the geostrophic relative vorticity for layer  $\kappa$ .

From our asymptotic analysis earlier in this section, the toughest part of that analysis concerned derivation of the quasi-geostrophic potential vorticity equation. In the present approach, we already have the potential vorticity for each layer via equation (9.130). What we need is the velocity field to advect it. Again, we know what that velocity is: it is the  $f$ -plane geostrophic velocity for each layer. The 2-layer velocity equations are given by equations (1.75a) and (1.75b), with their geostrophic components determined by

$$f_0 \hat{\mathbf{z}} \times \mathbf{u}_1 = -g \nabla (\eta'_b + h'_1 + h'_2) \quad (9.131a)$$

$$f_0 \hat{\mathbf{z}} \times \mathbf{u}_2 = -\nabla \left[ g_{1/2}' (\eta'_b + h'_1 + h'_2) + g_{3/2}' (\eta'_b + h'_2) \right]. \quad (9.131b)$$

In these equations we set the applied atmospheric pressure to a constant, and made use of the

reduced gravities at the layer interfaces are

$$g_{1/2}^r = g(\rho_1 - \rho_{\text{atm}})/\rho_{\text{ref}} \approx g \quad \text{and} \quad g_{3/2}^r = g(\rho_2 - \rho_1)/\rho_{\text{ref}}, \quad (9.132)$$

with the Boussinesq reference density taken as  $\rho_{\text{ref}} = \rho_1$ . Furthermore, we assume  $\rho_{\text{atm}} \ll \rho_{\text{ref}}$  so that the top interface reduced gravity is well approximated by the full gravity. From equations (9.131a) and (9.131b) we can identify the geostrophic streamfunctions

$$\psi_1 = (g/f_o)(\eta'_b + h'_1 + h'_2) \quad (9.133a)$$

$$\psi_2 = (1/f_o) \left[ g_{1/2}^r (\eta'_b + h'_1 + h'_2) + g_{3/2}^r (\eta'_b + h'_2) \right], \quad (9.133b)$$

so that the layer geostrophic velocities are given by

$$\mathbf{u}_1 = \hat{\mathbf{z}} \times \nabla \psi_1 \quad \text{and} \quad \mathbf{u}_2 = \hat{\mathbf{z}} \times \nabla \psi_2. \quad (9.134)$$

We thus proceed with the usual quasi-geostrophic method, whereby evolution is determined by the material time changes of the potential vorticity with advection given by the layer geostrophic flow

$$(\partial_t + \mathbf{u}_k \cdot \nabla) q_k = 0. \quad (9.135)$$

### 9.6.11 Rigid lid shallow water quasi-geostrophy

We studied the horizontally non-divergent barotropic model in Chapter 4, whereby the full velocity field has zero horizontal divergence, thus leading to a rigid surface boundary. Gradients in the lid pressure drive the flow, with the lid pressure required to maintain the non-divergence constraint on the horizontal flow. For a single layer of quasi-geostrophic shallow water fluid, the rigid lid approximation means that the external deformation radius goes to infinity, so that the quasi-geostrophic potential vorticity reduces to the absolute geostrophic vorticity

$$L_d \rightarrow \infty \implies q = \zeta + f. \quad (9.136)$$

That is, the single rigid layer reduces to the non-divergent barotropic model in Chapter 4.

For more than one layer, we follow the rigid lid stacked shallow water model discussed in Section 1.4.5. In this case, the rigid upper surface means that  $\eta'_{1/2} = 0$  so that fluctuations in the upper layer are given just by that of its lower interface

$$h'_1 = \eta'_{1/2} - \eta'_{3/2} = -\eta'_{3/2}. \quad (9.137)$$

As a result, the upper layer potential vorticity is given by

$$q_1 = f + \zeta_1 - f_o h'_1/H = f + \zeta_1 + f_o \eta'_{3/2}/H. \quad (9.138)$$

### 9.6.12 Further study

Section 6.4 of [Holton and Hakim \(2013\)](#) provides an insightful discussion of potential vorticity inversion for example atmospheric flows.

## 9.7 Non-dimensional Boussinesq ocean equations

In this section we non-dimensionalize the continuously stratified Boussinesq ocean equations. As part of this process we identify a variety of non-dimensional numbers that characterize the flow. As for the shallow water system in Section 9.4, continuously stratified planetary geostrophy is rather simple to derive, so that the detailed work in this section is arguably not necessary. However, the details here are essential for systematically deriving the continuously stratified quasi-geostrophic theory pursued in Chapter 11.

Our starting point is the perfect fluid stratified hydrostatic Boussinesq equations from VOLUME 2

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h \varphi \quad (9.139a)$$

$$\frac{\partial \varphi}{\partial z} = b \quad (9.139b)$$

$$\frac{Db}{Dt} = 0 \quad (9.139c)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (9.139d)$$

where  $\mathbf{v} = (\mathbf{u}, w)$  is the three-dimensional velocity written using Cartesian coordinates,  $b = -g(\rho - \rho_0)/\rho$  is the Archimedian buoyancy relative to a constant reference density,  $\rho_0$ , with  $\rho$  the density. We also write  $\varphi = \delta p/\rho$  for the dynamic pressure (dimensions of  $L^2 T^{-2}$ ), and  $\nabla_h = (\partial_x, \partial_y, 0)$  for the horizontal gradient operator. We separate a background vertical buoyancy profile from a space-time fluctuating buoyancy

$$b = \tilde{b}(z) + b'(x, y, z, t), \quad (9.140)$$

and introduce the corresponding background squared buoyancy frequency

$$N^2(z) = \frac{d\tilde{b}(z)}{dz}. \quad (9.141)$$

The background stratification is not determined by the quasi-geostrophic theory. Rather, it is assumed to be a prescribed function.

With the above decomposition, the buoyancy equation (9.139c) takes the form

$$\frac{Db'}{Dt} + w N^2 = 0. \quad (9.142)$$

We also introduce an associated decomposition of the hydrostatic pressure

$$\varphi = \tilde{\varphi}(z) + \varphi'(x, y, z, t) \quad (9.143)$$

where  $\tilde{\varphi}$  is hydrostatically balanced by  $\tilde{b}$

$$\frac{d\tilde{\varphi}}{dz} = \tilde{b}, \quad (9.144)$$

and the fluctuating pressure,  $\varphi'$ , is hydrostatically balanced by  $b'$

$$\frac{\partial \varphi'}{\partial z} = b'. \quad (9.145)$$

### 9.7.1 Dimensional parameters

As for the shallow water discussion in Section 9.3.1, we have the following dimensional parameters for the perfect Boussinesq fluid.

- LENGTH SCALES

- $H$  = length scale of a typical vertical structure in the fluid (e.g., the depth of the ocean pycnocline or height of the atmospheric tropopause). This scale is affected by the prescribed vertical stratification,  $N(z)$ .
- $L$  = horizontal/lateral length scale of the flow (e.g., Gulf Stream rings, atmospheric synoptic weather pattern, ocean gyre).
- $R_e$  = radius of the planet.

- VELOCITY SCALES

- $U$  = horizontal velocity scale for fluid motion.
- $W$  = vertical velocity scale for fluid motion.

- PRESSURE AND BUOYANCY SCALES: Pressure is a contact force, acting on the boundary of an arbitrary fluid region, and buoyancy arises from the gravitational force that acts to raise or lower a fluid element depending on its density relative to the environment. They have scales given by the following.

- $\Phi$  = scale for pressure fluctuations,  $\varphi'$  (dimensions of pressure divided by density = length scale  $\times$  acceleration).
- $B$  = scale of buoyancy fluctuations,  $b'$  (dimensions of acceleration).

- BODY FORCES: There are two body forces acting on the fluid, one from gravity and one from Coriolis.

- $g$  = gravitational acceleration
- $f$  = Coriolis frequency.

Contrary to the shallow water discussion in Section 9.3.1, we do not introduce a wave speed since it does not affect the asymptotics considered here. Also, we do not introduce a scale for the bottom topography undulations, since for planetary geostrophy there is no constraint on topographic undulations. However, for quasi-geostrophy in Chapter 11 we follow the approach used for shallow water in Section 9.3.5 by assuming topographic undulations scale like the Rossby number. Further details of the bottom boundary conditions and their scaling are detailed in Section 11.6.

### 9.7.2 Physical dimensions and non-dimensional parameters

There are two physical dimensions in the Boussinesq system: length,  $L$ , and time,  $T$ . As for the shallow water system, there is no need to consider a mass dimension since mass is determined by the density (buoyancy) and volume. The Buckingham-II theorem then says there are

$$N_{\text{dimensionless}} = 9 - 2 = 7 \quad (9.146)$$

non-dimensional parameters.

### 9.7.3 Choosing the non-dimensional parameters

Following the shallow water discussion in Section 9.3.4, we choose the following non-dimensional parameters.

1. VERTICAL TO HORIZONTAL ASPECT RATIO: The ratio of the vertical length scale to the horizontal length scale of the flow defines the aspect ratio

$$\delta_{\text{vertical/horizontal}} = \frac{\text{vertical length scale}}{\text{horizontal length scale}} = \frac{H}{L}. \quad (9.147)$$

2. RATIO OF HORIZONTAL SCALE TO PLANETARY SCALE: The ratio of the horizontal length scale of the flow to the planetary radius is given by

$$\delta_{\text{horizontal/planet}} = \frac{\text{horizontal length scale}}{\text{planetary length scale}} = \frac{L}{R_e}. \quad (9.148)$$

3. RATIO VERTICAL TO HORIZONTAL VELOCITY SCALES: The ratio of the vertical to horizontal velocity scales is given by

$$\frac{\text{vertical velocity scale}}{\text{horizontal velocity scale}} = \frac{W}{U}. \quad (9.149)$$

4. HYDROSTATIC NUMBER: The hydrostatic number is the ratio of the pressure fluctuation scale to the buoyancy fluctuation scale,

$$\frac{\text{pressure fluctuations}}{\text{buoyancy fluctuations}} = \frac{\Phi/H}{B}. \quad (9.150)$$

5. ROSSBY NUMBER: The Rossby number is the ratio of the fluid particle acceleration scale to the Coriolis acceleration

$$\text{Ro} = \frac{\text{particle acceleration}}{\text{Coriolis acceleration}} = \frac{U}{f L}. \quad (9.151)$$

As in Section 9.3.4, we assume that the time scale is advective

$$T \sim L/U \implies \text{Ro} = U/(f L) = (f T)^{-1}. \quad (9.152)$$

6. GEOSTROPHIC NUMBER: The ratio of the Coriolis acceleration to the pressure gradient acceleration defines the geostrophic number

$$\text{Ge} = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}}. \quad (9.153)$$

The Coriolis acceleration scales as

$$\text{Coriolis acceleration} \sim f U \quad (9.154)$$

whereas the pressure gradient acceleration from the fluctuating pressure,  $\varphi'$ , scales as

$$\text{pressure gradient acceleration} \sim \Phi/L, \quad (9.155)$$

so that

$$\text{Ge} = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}} = \frac{f U}{(\Phi/L)}. \quad (9.156)$$

7. RATIO OF FLUCTUATING STRATIFICATION TO BACKGROUND STRATIFICATION: The ratio of the buoyancy frequency arising from the fluctuating buoyancy,  $B/H$ , to the background squared buoyancy frequency,  $N^2(z)$ , is given by

$$\frac{\text{fluctuating squared buoyancy frequency}}{\text{background squared buoyancy frequency}} = \frac{B/H}{N^2(z)}. \quad (9.157)$$

#### 9.7.4 Relating the buoyancy scale to the Coriolis acceleration scale

The fluctuating buoyancy,  $b'$ , and fluctuating pressure,  $\varphi'$ , have scales related through the hydrostatic balance. Hence, taking a unit hydrostatic number from equation (9.150) renders

$$B = \Phi/H. \quad (9.158)$$

Additionally, assuming geostrophic scaling as per equation (9.156) means that the fluctuating pressure has a scale related to the Coriolis acceleration scale according to

$$\Phi = U f L. \quad (9.159)$$

Evidently, the scale for the fluctuating buoyancy is given by

$$B = f U (L/H). \quad (9.160)$$

We emphasize that the scale for pressure fluctuation,  $\Phi = U f L$ , is distinct from the non-rotating case considered in VOLUME 2, where  $\Phi = U^2$ , which we encountered (for example) when scaling the momentum equations for hydrostatic balance.

#### 9.7.5 Richardson number and QG/PG flow regimes

The Richardson number provides a measure of the stabilizing effects from vertical stratification versus the destabilizing effects from vertical shear

$$\text{Ri} = \frac{\text{vertical stratification}}{\text{vertical shear}}. \quad (9.161)$$

More precisely, the Richardson number is given by the ratio of the squared buoyancy frequency to the squared vertical shear of the horizontal velocity

$$\text{Ri} = \frac{N^2}{|\partial_z \mathbf{u}|^2}. \quad (9.162)$$

In regions where  $\text{Ri} < 1$ , the vertical shear is stronger than the stabilizing effects from vertical stratification. In regions with small Richardson numbers, there is enough kinetic energy in the vertical shear to extract potential energy from the stratification, and this extraction process occurs via a vertical shear instability as studied in VOLUME 4. In contrast, for large-scale highly stratified flow, the Richardson number is quite large, with  $\text{Ri} \sim 100$  common. Large Richardson number flow regimes are where quasi-geostrophy is relevant (Chapter 11).

Given the fundamental role of the Richardson number for stratified fluid motions, it is

useful to introduce it as one of our dimensionless parameters. Namely, we define the vertically dependent Richardson number scale as

$$\text{Ri}(z) = \frac{N^2(z)}{(U/H)^2}, \quad (9.163)$$

where we set the vertical length scale to  $H$ , the horizontal velocity scale to  $U$ , and the squared buoyancy frequency to the background value,  $N^2(z)$ , introduced by equation (9.141). Retaining vertical dependence to the background buoyancy frequency means that the Richardson number scale is also vertically dependent.

The Richardson number scale can be related to the Rossby and Burger numbers through

$$\text{Bu}(z) = \left[ \frac{L_d(z)}{L} \right]^2 = \left[ \frac{N(z) H}{f L} \right]^2 = \frac{U^2 \text{Ri}(z)}{U^2 / (\text{Ro})^2} = (\text{Ro})^2 \text{Ri}(z). \quad (9.164)$$

Evidently, vertical dependence to the prescribed background buoyancy frequency makes the Burger number and Richardson number vertically dependent, as well as the deformation radius.

One further way to write the Burger number is by introducing the angle  $\varphi$  defined by the vertical and horizontal length scales

$$\tan \varphi \equiv H/L \quad (9.165)$$

in which case

$$\text{Bu} = \left[ \frac{L_d}{L} \right]^2 = \left[ \frac{N H}{f L} \right]^2 = (\text{Ro})^2 \text{Ri} = [(N/f) \tan \varphi]^2. \quad (9.166)$$

When  $\tan^2 \varphi$  is set according to the slope of the ocean bottom, then  $[(N/f) \tan \varphi]^2$  is known as the **slope Burger number** ([MacCready and Rhines, 1993](#); [Peterson and Callies, 2022](#)).

The horizontal length scales,  $L$ , for quasi-geostrophic flows are assumed to be on the order of the deformation radius,  $L_d$ , in which case the Burger number is close to unity. The relation (9.164) means that the Richardson number scales as

$$\text{Ri} \sim (\text{Ro})^{-2} \quad \text{quasi-geostrophic flow regime.} \quad (9.167)$$

For atmospheric flows with a Rossby number order 1/10, quasi-geostrophic flow regimes are realized with a Richardson number  $\sim 100$ . For the ocean, the Rossby number can be even smaller, in which case quasi-geostrophic flows are characterized by an even larger Richardson number. For planetary geostrophy, the Burger number is small. Hence, planetary geostrophic flows are characterized by somewhat smaller Richardson numbers than quasi-geostrophic flows.

## 9.7.6 The Rossby deformation radius

The combined effects of buoyancy and rotation yield the richness of continuously stratified planetary geostrophic and quasi-geostrophic motions. Hence, the buoyancy frequency and the Coriolis parameter play central role in characterizing these flow regimes. The ratio of these two frequencies,  $N/f$ , in regions of nontrivial vertical stratification is typically around 100. Hence, rotational inertial oscillations (usually just called **inertial oscillations**) have about 100 times longer period,  $2\pi/f$ , than buoyancy oscillations with period  $2\pi/N$ .

Letting the squared buoyancy frequency,  $N^2$ , refer to a value typical of a particular flow

regime, one can define the Rossby deformation radius

$$L_d = H(N/f). \quad (9.168)$$

As defined, the deformation radius is the vertical length scale multiplied by the ratio of the buoyancy frequency to the Coriolis frequency. The ratio,  $f/N$ , appears frequently in rotating/stratified fluids, and is sometimes called the **Prandtl ratio**

$$\text{Prandtl ratio} = f/N. \quad (9.169)$$

With  $H \approx 1$  km and  $N/f \approx 100$ , the Rossby deformation radius is roughly 100 km. This length scale measures the relative importance of stratification and rotation. Depending on the ratio  $L/L_d$ , we can have large or small stratification fluctuations relative to the background stratification. Furthermore, when studying **baroclinic instability** in VOLUME 4, we find that the deformation radius sets the scale for unstable baroclinic waves leading to baroclinically unstable flow.

For some context, recall the shallow water deformation radius is given by equation (9.31),  $L_d = \sqrt{gH}/f$ , which is the ratio of the gravity wave speed to Coriolis frequency. With  $N = 100$ ,  $f = 10^{-2}$  s<sup>-1</sup> and  $H = 10^3$  m, the shallow water deformation radius is about an order of magnitude larger than the internal deformation radius. This scale difference means that the characteristic length scales, as set by  $L_d$ , are much larger in a single layer of shallow water fluid than in a stratified fluid.

### 9.7.7 Assumed values for the non-dimensional parameters

We now enumerate the assumed values for the non-dimensional parameters, again following the choices made for the shallow water layer in Section 9.3.5. These assumptions are guided by the flow regimes of interest.

1. **SMALL VERTICAL TO HORIZONTAL ASPECT RATIO:** The aspect ratio is generally small for large-scale atmospheric and oceanic flows

$$\delta_{\text{vertical/horizontal}} \ll 1. \quad (9.170)$$

This assumption is part of the **hydrostatic approximation** detailed in VOLUME 2, and as such it is a necessary scaling for any asymptotic theory based on a hydrostatic starting point.

2. **SMALL OR ORDER ONE RATIO OF HORIZONTAL TO PLANETARY SCALES:** The ratio of the horizontal length scale of the flow to the planetary radius is small for quasi-geostrophic systems, whereas the ratio is order unity for planetary geostrophy

$$\delta_{\text{horizontal/planet}} \ll 1 \quad \text{quasi-geostrophy} \quad (9.171a)$$

$$\delta_{\text{horizontal/planet}} \sim 1 \quad \text{planetary geostrophy.} \quad (9.171b)$$

For example, the vertical length scale could be determined by the averaged depth of the pycnocline in the ocean, or the averaged height of the tropopause in the atmosphere. For the horizontal length scale in quasi-geostrophic theory, we might choose the scale of a Gulf Stream ring in the ocean or synoptic weather pattern in the atmosphere. Alternatively, we might choose the lateral scale of an ocean gyre for planetary geostrophy.

3. SMALL RATIO OF VERTICAL TO HORIZONTAL VELOCITY SCALES: The continuity equation implies

$$W/H = U/L, \quad (9.172)$$

so that

$$W = U(H/L). \quad (9.173)$$

As noted above, for a hydrostatic fluid the vertical to horizontal aspect ratio,  $H/L$ , is small, so that the vertical velocity scale is smaller than the horizontal velocity scale. Furthermore, when the fluid is close to geostrophically balanced, the vertical velocity scale is even smaller, by a factor of  $\text{Ro}$ . We see that factor emerge in the following scale analysis.

4. UNIT HYDROSTATIC NUMBER: The hydrostatic balance (9.139b) means that the scales for a buoyancy fluctuation and pressure fluctuation are related by (see equation (9.150))

$$\Phi = H B. \quad (9.174)$$

5. SMALL ROSSBY NUMBER: The Rossby number is assumed small

$$\text{Ro} = U/(f L) = (f T)^{-1} \ll 1, \quad (9.175)$$

where we set the time scale for the motion according to advection,  $T = L/U$ .

6. UNIT GEOSTROPHIC NUMBER: The geostrophic number is assumed to be order unity

$$\text{Ge} \sim 1, \quad (9.176)$$

which means that the Coriolis acceleration and pressure gradient acceleration scale together

$$f U \sim \Phi/L \implies \Phi \sim U f L. \quad (9.177)$$

This scaling is consistent with the momentum equation (9.139a) so long as the Rossby number is small,  $\text{Ro} \ll 1$ .

7. STRATIFICATION FLUCTUATIONS COMPARED TO BACKGROUND STRATIFICATION: Making use of the assumed unit geostrophic number, the ratio of the buoyancy frequency arising from the fluctuating buoyancy to the background buoyancy frequency is given by

$$\frac{B/H}{N^2} = \frac{\Phi}{H^2 N^2} = \frac{f U L}{H^2 N^2} = \frac{U}{f L} \frac{L^2 f^2}{H^2 N^2} = \text{Ro} \frac{L^2}{L_d^2} = \frac{\text{Ro}}{\text{Bu}^2(z)}, \quad (9.178)$$

where we introduced the deformation radius (9.168)  $L_d = H(N/f)$ , which is a function of vertical position through the prescribed background buoyancy frequency,  $N(z)$ . We also introduced the Burger number,  $\text{Bu}(z) = (L_d(z)/L)^2$ , as per equation (9.33). It is important to keep the depth dependence of  $N^2, \text{Bu}(z)$ , and  $L_d(z)$ , when returning to dimensional fields, particularly for the quasi-geostrophic equations derived in Section 11.3.7.

### 9.7.8 Non-dimensional Boussinesq equations

Following the shallow water approach in Section 9.3.7, we introduce non-dimensional variables according to

$$t = T \hat{t} \quad (x, y) = L (\hat{x}, \hat{y}) \quad \partial_t = \partial_{\hat{t}} / T \quad \nabla_h = \hat{\nabla}_h / L \quad \partial_z = \partial_{\hat{z}} / H \quad f = f_0 \hat{f} \quad (9.179a)$$

$$(u, v) = U (\hat{u}, \hat{v}) \quad w = W \hat{w} \quad \varphi' = f_0 U L \hat{\varphi} \quad b' = B \hat{b} = (f_0 U L / H) \hat{b}. \quad (9.179b)$$

For the second equality in the buoyancy scale, we made use of equation (9.160) to connect the buoyancy fluctuation scale to the Coriolis acceleration scale. We also make use of the following relations between scales

$$T = L/U \quad \text{advective scaling for } T \quad (9.180)$$

$$W = U (H/L) \quad \text{continuity scaling for } W \quad (9.181)$$

$$\text{Ro} = U / (f_0 L) = (T f_0)^{-1} \quad \text{advective scaling for } T. \quad (9.182)$$

The first relation assumes the time scale is determined by the advection time,  $T = L/U$ , which then means that the Rossby number is the ratio of the advective frequency,  $1/T$ , to the Coriolis frequency,  $f_0$ . Furthermore, we assume a vertical velocity scale according to the continuity equation,  $W = U (H/L)$ . This continuity scaling for  $W$  is actually an over-estimate, where we find below that  $W$  instead scales like  $W = \text{Ro} U (H/L)$ .

#### Non-dimensional momentum equation

Introducing the dimensionless variables and dimensionful scales into the Boussinesq momentum equation (9.139a) renders

$$\frac{U}{T} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + \frac{U^2}{L} (\hat{\mathbf{u}} \cdot \hat{\nabla}_h) \hat{\mathbf{u}} + \frac{W U}{H} \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} + f_0 U (\hat{\mathbf{f}} \times \hat{\mathbf{u}}) = -f_0 U \hat{\nabla}_h \hat{\varphi}, \quad (9.183)$$

and dividing by  $f_0 U$  leads to

$$\text{Ro} \left[ \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}_h) \hat{\mathbf{u}} + \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} \right] + (\hat{\mathbf{f}} \times \hat{\mathbf{u}}) = -\hat{\nabla}_h \hat{\varphi}. \quad (9.184)$$

The non-dimensional hydrostatic balance is given by

$$\frac{\partial \hat{\varphi}}{\partial \hat{z}} = \hat{b}, \quad (9.185)$$

and the non-dimensional continuity equation is

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0. \quad (9.186)$$

#### Non-dimensional buoyancy equation

The buoyancy equation (9.142) requires a bit more work to non-dimensionalize. The material time derivative takes the form

$$\frac{D b'}{D t} = \frac{B}{T} \frac{D \hat{b}}{D \hat{t}} = \frac{U}{L} \frac{f_0 U L}{H} \frac{D \hat{b}}{D \hat{t}} = \frac{f_0 U^2}{H} \frac{D \hat{b}}{D \hat{t}}, \quad (9.187)$$

where we made use of the advective scaling  $T = L/U$  and continuity scaling  $W = U(H/L)$ . The vertical advection of background stratification is given by

$$N^2 w = N^2 W \hat{w} = N^2 U (H/L) \hat{w} = L_d^2 \frac{U f_o^2}{H L} \hat{w}, \quad (9.188)$$

where we introduced the deformation radius,  $L_d = H(N/f)$ , from equation (9.168). Bringing these two pieces together leads to

$$\text{Ro} \frac{D\hat{b}}{Dt} + \text{Bu} \hat{w} = 0, \quad (9.189)$$

where we introduced the Burger number,  $\text{Bu} = (L_d/L)^2$ .

### 9.7.9 Comments

As stated earlier, the material in this section serves as the starting point for a systematic derivation of the continuously stratified planetary geostrophic equations in Chapter 10, and the continuously stratified quasi-geostrophic equations in Chapter 11. Particularly for the quasi-geostrophic equations, we make use of asymptotic methods as for the shallow water quasi-geostrophic equations in Section 9.5.



## 9.8 Exercises

### EXERCISE 9.1: PV CONSERVATION FOR PLANETARY GEOSTROPHY

Show that the planetary geostrophic equations

$$\mathbf{f} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} \quad \text{with} \quad \eta = \eta_b + h \quad (9.190)$$

are equivalent to

$$\mathbf{f} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{DQ}{Dt} = 0 \quad \text{with} \quad Q = \frac{f}{h}. \quad (9.191)$$

This result shows that the shallow water PG equations may be written as an evolution equation for an approximated version of the shallow water potential vorticity,  $(f + \zeta)/h \approx f/h$ . This limit holds when the Rossby number is small.

### EXERCISE 9.2: CONSTRAINTS ON STEADY STATE PLANETARY GEOSTROPHIC FLOW

Consider a shallow water fluid satisfying the planetary geostrophic equations developed in Section 9.4. Assume the flow is in steady state.

- In what manner does potential vorticity conservation constrain the velocity field?
- Consider an initially zonal geostrophic flow. In what direction (poleward or equatorward) will a fluid parcel deviate when encountering a seamount (i.e., a region of relatively shallow depth)?
- Describe the geostrophic contours (i.e., path of fluid particles following the geostrophic flow) for the case where the ocean sea surface height undulations,  $\eta'$ , are far smaller than undulations in the bottom topography,  $\eta'_b$  (see Figure 1.1 for notation).

- (d) For the special case of an  $f$ -plane, show that the velocity is aligned with isolines of bottom topography.
- (e) For the special case of a flat bottom and latitudinally dependent Coriolis parameter,  $f(y)$ , show that there is no meridional geostrophic velocity. That is, the flow is zonally aligned.





## Chapter 10

# PLANETARY GEOSTROPHIC VORTICITY ANALYSIS

Planetary geostrophy consists of steady and linear frictional geostrophic flow coupled to the non-steady and nonlinear buoyancy evolution. Consequently, the fluid state evolves through the buoyancy equation, whereas the flow is diagnosed from frictional geostrophy and continuity. Furthermore, there is no turbulence in planetary geostrophy since the momentum equation is linear and steady. Evidently, planetary geostrophy is concerned with large-scale flow of a stably stratified laminar fluid whose vorticity is dominated by planetary rotation in the presence of planetary beta and topographic beta.

We already studied various physical properties of planetary geostrophy in VOLUME 2, including the [geostrophic balance](#), vorticity balance via the [Sverdrup balance](#), [thermal wind balance](#), and the [Taylor-Proudman effect](#). We also made use of planetary geostrophy to study western boundary current intensification for a shallow water layer in Section 5.7. We were able to present those studies earlier in the book since, as seen in Section 9.4 for the shallow water, derivation of the planetary geostrophic equations is a very simple task, thus allowing the equations to be plausibly written down without needing any formal asymptotics.

The stratified planetary geostrophic equations form the foundation for theories of the large-scale ocean circulation, with vorticity constraints providing a key reason for the central role of planetary geostrophy. The central role for vorticity motivates a focus in this chapter on the mathematical and physical basis for planetary geostrophic vorticity analysis. It is notable that in a planetary geostrophic flow, we are only concerned with planetary vorticity, and thus ignore relative vorticity. Consequently, we ignore vorticity sources from baroclinicity and tilting. Instead, we focus on how vertical stretching modifies a fluid's planetary vorticity, and how the fluid responds by moving meridionally to adjust its planetary vorticity in response to stretching.

After summarizing the planetary geostrophic equation set, we derive the planetary geostrophic potential vorticity budget and determine how the [impermeability theorem](#) theorem from Section 8.2 appears in planetary geostrophy. We then study a suite of vorticity budgets arrived at through vertical integration of the fluid equations, with particular focus on ocean fluid mechanics. Each of the resulting two-dimensional vorticity budgets offers insights into how large-scale ocean flow is constrained by rotation and the beta effect. In particular, these budgets render insights into how forces and the curl of forces generate vertical flow next to the boundaries as well as meridional flow for the full fluid column.

## CHAPTER GUIDE

We here extend the shallow water discussions from Chapter 9 to develop an understanding of the continuously stratified planetary geostrophic equations. We make use of stratified geophysical fluid dynamics from VOLUME 2, vorticity and the planetary beta effect from Chapter 6, and potential vorticity from Chapter 7. Physical properties of stratified geostrophic mechanics were considered in VOLUME 2, with an understanding of that material assumed in the present chapter.

<b>10.1</b>	<b>Loose threads</b>	<b>352</b>
<b>10.2</b>	<b>Equations for planetary geostrophy</b>	<b>352</b>
10.2.1	Common form of the equations	353
10.2.2	Planetary geostrophic energetics	354
<b>10.3</b>	<b>Planetary geostrophic potential vorticity</b>	<b>355</b>
10.3.1	Derivation	355
10.3.2	Impermeability theorem	356
10.3.3	A kinematic PV flux satisfying impermeability	356
<b>10.4</b>	<b>Depth integrated vorticity budget</b>	<b>357</b>
10.4.1	The $\beta$ -effect, stretching, and meridional transport	357
10.4.2	Bottom kinematics and dynamics	359
10.4.3	Surface kinematics and dynamics	361
10.4.4	Summary of force curls driving depth integrated meridional flow	362
10.4.5	Integral constraints	364
<b>10.5</b>	<b>Sverdrup balance and geostrophic Sverdrup balance</b>	<b>365</b>
<b>10.6</b>	<b>Vorticity of the depth integrated velocity</b>	<b>366</b>
10.6.1	Depth integrated velocity equation	367
10.6.2	Vorticity budget	367
10.6.3	Integral balances for steady flows	368
10.6.4	Comments	369
<b>10.7</b>	<b>Vorticity equation for the depth averaged velocity</b>	<b>369</b>
10.7.1	Relating the depth average velocity to boundary velocities	369
10.7.2	Formulation of the vorticity equation	370
10.7.3	Rigid lid approximation and the role of JEBAR	371
10.7.4	Further study	373

## 10.1 Loose threads

- Exercises

## 10.2 Equations for planetary geostrophy

Just like for the shallow water model in Section 9.4, the planetary geostrophic model for the stratified Boussinesq ocean is a rather simple asymptotic theory. For this case, we assume the horizontal scales are large compared to the deformation radius, so that

$$\text{Ro/Bu} \sim 1 \implies \text{Ro} L^2 \sim L_d^2. \quad (10.1)$$

With this scaling, and with a small Rossby number, the momentum equation (9.184) reduces to geostrophic balance. However, the continuity and buoyancy equations retain their unapproximated Boussinesq form. Hence, in dimensional form, the perfect (adiabatic and inviscid) planetary geostrophic equations for a stratified Boussinesq ocean are

$$\frac{Db'}{Dt} + w N^2 = 0 \quad \text{and} \quad f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h \varphi' \quad \text{and} \quad \frac{\partial \varphi'}{\partial z} = b' \quad \text{and} \quad \nabla \cdot \mathbf{v} = 0. \quad (10.2)$$

### 10.2.1 Common form of the equations

We could just as well write the planetary geostrophic equations (10.2) in terms of the full buoyancy

$$b = \tilde{b}(z) + b', \quad (10.3)$$

and full pressure,

$$p = p_0(z) + \rho \varphi. \quad (10.4)$$

Additionally, it is quite useful to include non-conservative terms such as buoyancy mixing,  $\dot{b}$ , to allow for the study of how stratification evolves, as well as horizontal frictional accelerations and/or boundary accelerations,  $\mathbf{F}$ , to include boundary stress driven circulations through the **Ekman mechanics** studied in VOLUME 2. For these reasons, we take the frictional and diabatic planetary geostrophic equations as the basis for discussions in this chapter

$$f \hat{\mathbf{z}} \times \mathbf{u} = -\rho_0^{-1} \nabla p + \mathbf{F} \quad \text{frictional geostrophy} \quad (10.5a)$$

$$(\partial_t + \mathbf{u} \cdot \nabla_h) b + N^2 w = \dot{b} \quad \text{diabatic buoyancy equation} \quad (10.5b)$$

$$\partial_z p = -\rho g \quad \text{hydrostatic balance} \quad (10.5c)$$

$$\nabla \cdot \mathbf{v} = \nabla_h \cdot \mathbf{u} + \partial_z w = 0 \quad \text{non-divergent flow} \quad (10.5d)$$

$$N^2 = \partial_z b \quad \text{squared buoyancy frequency} \quad (10.5e)$$

$$b = -g(\rho - \rho_0)/\rho_0 \quad \text{Archimedean buoyancy.} \quad (10.5f)$$

Note that the material time derivative in planetary geostrophy makes use of advection by the three velocity components,  $\mathbf{v} = (\mathbf{u}, w)$ , as seen in the buoyancy equation (10.5b), with the horizontal velocity components determined by the frictional geostrophic balance (10.5a). This situation contrasts to quasi-geostrophy, where it is only the horizontal advection by the geostrophic flow that contributes to material time evolution (Section 11.2).

As for the **Boussinesq ocean** equations discussed in VOLUME 2, we sometimes find it convenient to combine the horizontal velocity equation with the hydrostatic balance to write

$$f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla \varphi + b \hat{\mathbf{z}} + \mathbf{F}. \quad (10.6)$$

Furthermore, it is common to assume an equation of state that is independent of pressure, so that material time changes in buoyancy arise only from changes in Conservative Temperature and/or salinity

$$\dot{b} = \frac{\partial b}{\partial S} \dot{S} + \frac{\partial b}{\partial \Theta} \dot{\Theta}. \quad (10.7)$$

The partial derivatives,  $\partial b / \partial \Theta$  and  $\partial b / \partial S$ , are commonly assumed constant in idealized studies.

### 10.2.2 Planetary geostrophic energetics

Since the velocity is diagnostic in planetary geostrophy, it is determined by the buoyancy field. In turn, there is a prognostic equation for potential energy that arises from the buoyancy equation, whereas kinetic energy is diagnostic. The energetics are thus a special case of the Boussinesq energetics studied in VOLUME 2. We here consider just the basics.

#### General considerations

Multiplying the buoyancy equation (10.5b) by  $z$  leads to the potential energy equation

$$\partial_t P + \nabla \cdot (\mathbf{v} P) + w b = -z \dot{b}, \quad (10.8)$$

where we introduced the potential energy per mass relative to the reference density<sup>1</sup>

$$P = -z b = z g (\rho - \rho_0) / \rho_0. \quad (10.9)$$

Equation (10.8) says that the potential energy at a point in the planetary geostrophic fluid is affected by advective transport, buoyancy work, and diabatic processes. As we will see, diabatic processes such as diffusion provide a local source for potential energy, whereas buoyancy work transfers potential energy to kinetic energy.

Projecting the geostrophic/hydrostatic balance equation (10.6) onto the velocity leads to the diagnostic balance between pressure work, buoyancy work, and friction

$$\mathbf{v} \cdot \nabla \varphi - w b = \mathbf{u} \cdot \mathbf{F}. \quad (10.10)$$

We can add this balance to the potential energy equation (10.8) to eliminate the buoyancy work term,  $w b$ , thus yielding

$$\partial_t P + \nabla \cdot (\mathbf{v} P + \mathbf{v} \varphi) = \mathbf{u} \cdot \mathbf{F} - z \dot{b}, \quad (10.11)$$

where we used  $\nabla \cdot \mathbf{v} = 0$  to bring  $\mathbf{v} \varphi$  inside the divergence. We thus see that potential energy at a point in the planetary geostrophic fluid is affected by reversible transport processes from advection and pressure work, along with irreversible processes from friction and material buoyancy changes.

#### Diffusively driven flow

It is revealing to consider the special case of a constant volume domain with static and rigid boundaries and with no boundary fluxes. We also assume that buoyancy is irreversibly modified through diffusion

$$\dot{b} = \nabla \cdot (\kappa \nabla b), \quad (10.12)$$

with  $\kappa > 0$  an isotropic kinematic diffusivity that can be a function of space and time. Integrating the potential energy equation (10.11) over the domain leads to

$$\partial_t \langle P \rangle = \langle \mathbf{u} \cdot \mathbf{F} \rangle + \langle \kappa N^2 \rangle, \quad (10.13)$$

---

<sup>1</sup>We considered this same form for the potential energy in VOLUME 2 as part of the Boussinesq ocean chapter.

where the angle brackets signify volume means. To reach this identity we made use of

$$z \dot{b} = z \nabla \cdot (\kappa \nabla b) = \nabla \cdot (z \kappa \nabla b) - \kappa \partial_z b = \nabla \cdot (z \kappa \nabla b) - \kappa N^2, \quad (10.14)$$

with  $\nabla \cdot (z \kappa \nabla b)$  integrating to zero in the absence of boundary fluxes. The global mean potential energy equation (10.13) indicates that diffusion increases volume mean potential energy for a stably stratified fluid ( $N^2 > 0$ ). In contrast, as shown in our study of energetics in VOLUME 2, friction generally dissipates kinetic energy via  $\langle \mathbf{u} \cdot \mathbf{F} \rangle < 0$ . For planetary geostrophy, equation (10.13) means that friction also dissipates potential energy. We thus find that diffusion is the only source for potential energy, with spatial variations in potential energy leading to motion through the geostrophic balance. In the steady state this diffusively driven flow leads to the global mean balance between diffusion and friction

$$\langle \kappa N^2 \rangle = -\langle \mathbf{u} \cdot \mathbf{F} \rangle \quad \text{steady state.} \quad (10.15)$$

## 10.3 Planetary geostrophic potential vorticity

In Section 7.6 we developed the potential vorticity equation for the hydrostatic Boussinesq ocean in the presence of horizontal friction in the momentum equation and diabatic terms in the buoyancy equation. Here we specialize that result to the case of planetary geostrophic system written in the form of equations (10.5a)-(10.5e).

### 10.3.1 Derivation

Derivation of the potential vorticity equation proceeds much like that for the hydrostatic Boussinesq ocean. The first step requires the planetary geostrophic vorticity equation as determined by taking the curl of the momentum equation (10.5a). As derived in our study of geostrophy in VOLUME 2, the vertical component of this vorticity equation is given by

$$\beta v = f \partial_z w + \hat{\mathbf{z}} \cdot (\nabla_h \times \mathbf{F}) \quad \text{with} \quad \beta = \partial_y f. \quad (10.16)$$

Next, make use of frictional thermal wind balance

$$f \partial_z \mathbf{u} = \hat{\mathbf{z}} \times \nabla_h b - \partial_z (\hat{\mathbf{z}} \times \mathbf{F}) \quad (10.17)$$

as well as the identities

$$N^2 \frac{Df}{Dt} = N^2 \beta v \quad (10.18a)$$

$$f \frac{DN^2}{Dt} = f \frac{\partial \dot{b}}{\partial z} - f \nabla b \cdot \frac{\partial \mathbf{v}}{\partial z} \quad (10.18b)$$

$$f \nabla b \cdot \frac{\partial \mathbf{v}}{\partial z} = f N^2 \frac{\partial w}{\partial z} - \frac{\partial (\hat{\mathbf{z}} \times \mathbf{F})}{\partial z} \cdot \nabla_h b, \quad (10.18c)$$

to render

$$\frac{D(f N^2)}{Dt} = N^2 \beta v + f \frac{\partial \dot{b}}{\partial z} - f \nabla b \cdot \frac{\partial \mathbf{v}}{\partial z} \quad (10.19a)$$

$$= N^2 \left[ f \frac{\partial w}{\partial z} + \hat{\mathbf{z}} \cdot (\nabla_h \times \mathbf{F}) \right] + f \frac{\partial \dot{b}}{\partial z} - f N^2 \frac{\partial w}{\partial z} + \frac{\partial (\hat{\mathbf{z}} \times \mathbf{F})}{\partial z} \cdot \nabla_h b \quad (10.19b)$$

$$= f \frac{\partial \dot{b}}{\partial z} + \nabla b \cdot (\nabla \times \mathbf{F}) \quad (10.19c)$$

$$= \nabla \cdot (f \dot{b} \hat{z} + \mathbf{F} \times \nabla b). \quad (10.19d)$$

We thus identify the planetary geostrophic potential vorticity

$$Q^{\text{pg}} = f N^2, \quad (10.20)$$

which is materially invariant in the absence of diabatic processes and friction

$$\frac{DQ^{\text{pg}}}{Dt} = 0 \quad \text{if } \dot{b} = 0 \text{ and } \mathbf{F} = 0. \quad (10.21)$$

We can write the general budget equation in the form of an Eulerian flux-form expression

$$\partial_t Q^{\text{pg}} + \nabla \cdot \mathbf{J}_{\text{pg}} = 0, \quad (10.22)$$

where the planetary geostrophic potential vorticity flux is given by

$$\mathbf{J}_{\text{pg}} = \mathbf{v} Q^{\text{pg}} - \dot{b} f \hat{z} + \nabla b \times \mathbf{F} + \nabla \times \mathbf{A}. \quad (10.23)$$

The vector,  $\mathbf{A}$ , is an arbitrary gauge function that has no impact on the potential vorticity evolution. Comparing to the hydrostatic Boussinesq ocean expression (7.86), we see that the planetary geostrophic result follows by approximating the absolute vorticity by just the planetary vorticity.

### 10.3.2 Impermeability theorem

Following the discussion in Section 8.2.2, we verify that the potential vorticity flux vector (10.23) satisfies the impermeability theorem for buoyancy isosurfaces. We do so for the particular case of a zero gauge function ( $\mathbf{A} = 0$ ), in which case

$$\mathbf{v}_{\text{pg}} \cdot \nabla b = (\mathbf{J}_{\text{PG}}/Q) \cdot \nabla b = \mathbf{v} \cdot \nabla b - \dot{b} = -\partial_t b, \quad (10.24)$$

so that

$$(\partial_t + \mathbf{v}_{\text{pg}} \cdot \nabla) b = 0. \quad (10.25)$$

Evidently, there is zero flux of PV-substance crossing buoyancy isosurfaces, even in the presence of irreversible processes that allow matter and buoyancy to cross those surfaces. As shown in the next subsection, we identify two more forms of the PV-substance flux vector that also satisfy impermeability, with these alternative forms differing by gauge transformations.

### 10.3.3 A kinematic PV flux satisfying impermeability

Following the discussion of impermeability for the Ertel potential vorticity in Section 8.2.2, we expose a purely kinematic means to derive the impermeability theorem for the planetary geostrophic potential vorticity. This derivation follows by computing the time tendency of the potential vorticity

$$\frac{\partial Q}{\partial t} = \frac{\partial}{\partial t} \nabla \cdot (f b \hat{z}) = \nabla \cdot \left[ f \frac{\partial b}{\partial t} \hat{z} \right] \equiv -\nabla \cdot \tilde{\mathbf{J}}_{\text{pg}}, \quad (10.26)$$

where

$$\tilde{\mathbf{J}}_{\text{pg}} = -f \partial_t b \hat{\mathbf{z}}. \quad (10.27)$$

This form of the PV-substance flux also satisfies impermeability since

$$\tilde{\mathbf{v}}_{\text{pg}} \cdot \nabla b = (\tilde{\mathbf{J}}_{\text{pg}}/Q) \cdot \nabla b = -\partial_t b, \quad (10.28)$$

so that

$$(\partial_t + \tilde{\mathbf{v}}_{\text{pg}} \cdot \nabla) b = 0. \quad (10.29)$$

The PV-substance flux,  $\tilde{\mathbf{J}}_{\text{pg}}$ , vanishes in the steady state, whereas the steady state form of the alternative flux,  $\mathbf{J}_{\text{PG}}$ , is nonzero. Following the discussion in Section 8.3.2, we may choose to introduce a gauge transformation to the kinematic flux,  $\tilde{\mathbf{J}}_{\text{pg}}$ , so that it does not vanish in the steady state. Taking the small Rossby number limit of the flux (8.32) renders

$$\mathbf{J}_Q^{\text{marshall PG}} = -\nabla(g z + \varphi) \times \nabla b - f \partial_t b \hat{\mathbf{z}}. \quad (10.30)$$

This flux differs from  $\tilde{\mathbf{J}}_{\text{pg}}$  by a curl

$$\nabla(g z + \varphi) \times \nabla b = \nabla \times [(g z + \varphi) \nabla b], \quad (10.31)$$

and it also satisfies the impermeability theorem. As discussed in Section 8.5.6, there are a variety of motivations for using one form of the PV-substance flux versus another. Some applications prefer a nonzero steady flux that also does not expose any irreversible processes, with  $\mathbf{J}_Q^{\text{marshall PG}}$  satisfying these desires.

## 10.4 Depth integrated vorticity budget

In a planetary geostrophic flow, vorticity arises just from planetary vorticity since relative vorticity is negligible by comparison. With planetary vorticity a function just of latitude, a budget for the planetary geostrophic vorticity reveals how the curl of forces imparted to the fluid cause meridional motion as the fluid meets the constraints imposed by the vorticity equation. As per our discussion in Section 6.3.6, we refer to a force curl as a “torque” in our study of vorticity sources. However, one must keep in mind that more common usage in physics refers to a torque as affecting changes to angular momentum, with angular momentum generally distinct from vorticity (see Section 3.9 for a discussion of the distinction).

In this section we study the depth integrated vorticity budget for the planetary geostrophic fluid and derive implications for the meridional flow. For this purpose we focus on frictional and boundary accelerations that take the form of a vertical divergence of horizontal turbulent stress vector

$$\mathbf{F} = \partial_z \boldsymbol{\tau}. \quad (10.32)$$

The curl of this stress, as well as pressure forces, provide torques that generate meridional motion as revealed by the planetary geostrophic vorticity budget.

### 10.4.1 The $\beta$ -effect, stretching, and meridional transport

In our study of geostrophy in VOLUME 2, we derived the vorticity equation for planetary geostrophy. We also encountered this equation when deriving the potential vorticity budget in Section 10.3.1. With friction written as a vertical divergence of horizontal turbulent stresses

(equation (10.32)), the vertical component of the planetary geostrophic vorticity equation takes the form

$$\rho_0 \beta v = \partial_z [\rho_0 f w + \hat{z} \cdot (\nabla \times \boldsymbol{\tau})]. \quad (10.33)$$

Vertical integration from the ocean bottom at  $z = \eta_b(x, y)$  to sea surface at  $z = \eta(x, y, t)$  leads to<sup>2</sup>

$$\rho_0 \beta V = \underbrace{\rho_0 f [w(\eta) - w(\eta_b)]}_{\text{column stretching}} + \underbrace{\hat{z} \cdot (\nabla \times \Delta \boldsymbol{\tau})}_{\text{boundary stresses}}, \quad (10.34)$$

where

$$V = \int_{\eta_b}^{\eta} v \, dz \quad (10.35)$$

is the depth-integrated meridional flow, and

$$\Delta \boldsymbol{\tau} = \boldsymbol{\tau}(\eta) - \boldsymbol{\tau}(\eta_b) \quad (10.36)$$

is the difference in boundary stresses applied at the ocean surface and ocean bottom. Note that  $\Delta \boldsymbol{\tau}$  is just a function of horizontal position and time.

For a planetary geostrophic flow, absolute vorticity is approximated by just the planetary vorticity

$$\zeta_a = \zeta + f \approx f. \quad (10.37)$$

As revealed by the vorticity equation (10.34), vorticity sources in a planetary geostrophic fluid lead to meridional motion, with meridional motion the only way a planetary geostrophic fluid can modify its vorticity in response to vorticity sources.

The first term on the right hand side of the vorticity equation (10.34) arises from vertical stretching of the depth integrated column, as measured by differences in the vertical velocity at the ocean surface and bottom. For example, vertical stretching caused by positive surface velocity,  $w(\eta) > 0$ , or a negative bottom velocity,  $w(\eta_b) < 0$ , lead to poleward motion of the fluid column. Conversely, vertical squashing leads to equatorward motion. We emphasize that when studying the motion of the depth integrated flow, we are only concerned with vertical stretching from differences in the boundary vertical velocity rather than the vertical velocity within the fluid interior.

The second term in the vorticity equation (10.34) arises from differences in the vorticity imparted by surface and bottom boundary stresses. Positive vorticity is imparted to the fluid through a positive curl of wind stresses,  $\hat{z} \cdot [\nabla \times \boldsymbol{\tau}(\eta)] > 0$ , or by a negative curl of bottom stresses,  $\hat{z} \cdot [\nabla \times \boldsymbol{\tau}(\eta_b)] < 0$ , with such stress curls leading to poleward motion of the fluid column. The opposite motion occurs from the converse curls.

The  $\beta$ -effect (Section 6.6.2) is a fundamental element of the depth integrated vorticity equation (10.34). Namely, as fluid columns are stretched or squashed, they must move meridionally to maintain vorticity balance for a planetary geostrophic fluid on a rotating spherical earth. The planetary geostrophic vorticity equation restricts attention to vertical stretching through vertical motion (the  $w$  terms) and through the vorticity imparted by the curl of boundary stresses. Notably, the curl of boundary stresses also imparts vertical motion through surface and bottom Ekman layer dynamics (see our study of Ekman mechanics in VOLUME 2). Hence, the right hand side of the vorticity equation (10.34) is fundamentally related to vortex stretching.

---

<sup>2</sup>As a means to unclutter notation, we write  $\nabla$ , rather than  $\nabla_b$ , whenever acting on a field that is just a function of horizontal position, such as  $p_a, p_b, \eta_b$  and  $\Delta \boldsymbol{\tau}$ . We can do so since, for example,  $\nabla p_b = \nabla_b p_b$ , since  $\partial_z p_b = 0$ .

Equation (10.34) is central to mechanical descriptions of large-scale ocean circulation. For many flow regimes, the curl of the surface wind stress dominates, thus allowing us to ignore the vertical velocity terms as well as bottom frictional stresses. Formally, we isolate the wind stress when studying a flat bottom rigid lid model, whereby  $w(\eta) = w(\eta_b) = 0$ . However, there can be nontrivial impacts from bottom pressure torques when flow interacts with sloping topography, with the North Atlantic and Southern Ocean providing important case studies. Other processes can be important in various flow regimes, thus prompting us to derive a full diagnostic framework to identify where these processes are important. To pursue that framework, we make use of the kinematic boundary conditions and the horizontal momentum equation to unpack the vertical velocity terms. Doing so reveals the forces and their curls that drive vertical motion at the boundaries for a planetary geostrophic flow.

#### 10.4.2 Bottom kinematics and dynamics

The bottom no normal flow kinematic boundary condition applied at  $z = \eta_b(x, y)$  is given by

$$w = \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \eta_b(x, y). \quad (10.38)$$

This relation expresses the no-normal flow condition,  $\hat{\mathbf{n}} \cdot \mathbf{v} = 0$ , at the ocean bottom, with

$$\hat{\mathbf{n}} = -\frac{\nabla(z - \eta_b)}{|\nabla(z - \eta_b)|} = -\left[\frac{\hat{\mathbf{z}} - \nabla \eta_b}{\sqrt{1 + \nabla \eta_b \cdot \nabla \eta_b}}\right] \quad (10.39)$$

the outward unit normal to the bottom. The boundary condition constrains the flow so that any horizontal motion next to a sloping bottom that is oriented either up or down the slope must have an associated vertical motion. As we see in this section, such vertical motion next to the bottom boundary arises from force curls acting to stretch or squash a fluid column. In turn, through the vorticity equation (10.34), vertical motion at the bottom leads to meridional motion of the full fluid column.

#### Expressions for bottom vertical velocity

The bottom kinematic boundary condition holds for all dynamical flow regimes. For the particular case of planetary geostrophy, we garner insight into the forces that drive vertical flow near the bottom by making use of the planetary geostrophic momentum equation (10.5a). Evaluating the horizontal components of this equation at the ocean bottom yields<sup>3</sup>

$$\rho_0 f \hat{\mathbf{z}} \times \mathbf{u} = -(\nabla p)_{z=\eta_b} + \mathbf{F}_b, \quad (10.40)$$

where  $(\nabla p)_{z=\eta_b}$  is the horizontal pressure gradient evaluated at the ocean bottom and  $\mathbf{F}_b(x, y, t)$  is the bottom friction. For the special case of a no-slip bottom, all velocity components vanish at  $z = \eta_b$ . In that case, we consider  $\mathbf{u}$  in equation (10.40) to be the horizontal velocity averaged within the bottom boundary layer, and  $w$  the corresponding vertical velocity leaving the boundary layer.

It is convenient to decompose the bottom horizontal velocity into its geostrophic and Ekman<sup>4</sup>

<sup>3</sup>Recall that since  $p_b = p_b(x, y, t)$ , we have  $\nabla p_b = \nabla_b p_b$ . As noted in the footnote on page 358, we drop the  $z$  script to reduce notational clutter.

<sup>4</sup>It can be useful to here review our discussion of [Ekman mechanics](#) from VOLUME 2.

components via

$$\rho_0 f \hat{z} \times \mathbf{u}_g = -(\nabla p)_{z=\eta_b} \quad \text{and} \quad \rho_0 f \hat{z} \times \mathbf{u}_e = \mathbf{F}_b, \quad (10.41)$$

so that

$$\mathbf{u}_g = \frac{\hat{z} \times (\nabla p)_{z=\eta_b}}{\rho_0 f} \quad \text{and} \quad \mathbf{u}_e = -\frac{\hat{z} \times \mathbf{F}_b}{\rho_0 f}. \quad (10.42)$$

For the horizontal horizontal pressure gradient at the ocean bottom, we make use of our study of hydrostatic pressure in VOLUME 2, where we derived

$$(\nabla_h p)_{z=\eta_b} = \nabla_h p_b + g \rho(\eta_b) \nabla_h \eta_b. \quad (10.43)$$

The corresponding bottom vertical velocity components are determined by inserting equations (10.42) and (10.43) into the bottom kinematic boundary condition (10.38)

$$w_g = \frac{\hat{z} \cdot (-\nabla \eta_b \times \nabla p_b)}{\rho_0 f} \quad \text{and} \quad w_e = \frac{\hat{z} \cdot (\nabla \eta_b \times \mathbf{F}_b)}{\rho_0 f}. \quad (10.44)$$

These equations reveal how the curl of pressure forces and boundary frictional forces drive a nonzero vertical motion next to the bottom, while maintaining the bottom kinematic boundary condition (10.38). As seen by these equations, is only the projection of  $\nabla p_b$  and  $\mathbf{F}_b$  onto the isobath direction that contributes to a nonzero vertical velocity. These along-isobath forces are needed to render a horizontal velocity that is itself misaligned with isobaths, thus satisfying the kinematic requirement for vertical motion.<sup>5</sup>

To further understand the bottom pressure term, we write it as

$$w_g = \frac{\hat{z} \cdot (-\nabla \eta_b \times \nabla p_b)}{\rho_0 f} = \frac{\hat{z} \cdot [\nabla \times p_b \nabla \eta_b]}{\rho_0 f}. \quad (10.45)$$

The numerator is the curl of the horizontal projection of the pressure contact force along the bottom,  $p_b \nabla \eta_b$ . This term is the topographic form stress studied in VOLUME 2 for a general fluid and in Section 5.7.6 for the shallow water. We thus conclude that vertical geostrophic motion next to the bottom arises from the curl of the topographic form stress. This is an important result that will appear again within this section as well as in Sections 10.6 and 10.7.

### Comments on the bottom vertical geostrophic velocity

A large part of the bottom pressure gradient driving the horizontal geostrophic flow in equation (10.42) arises from changes in bottom depth. However, that portion of the bottom pressure gradient has no impact on  $w_g$ , since it only drives horizontal flow along isobaths. We see this property by writing

$$p_b = -\rho_0 g \eta_b + p'_b \implies w_g = \frac{\hat{z} \cdot (-\nabla \eta_b \times \nabla p'_b)}{\rho_0 f}. \quad (10.46)$$

When there is misalignment between isolines of bottom pressure and bottom topography, the geostrophic flow in a fluid column crosses isobaths. Correspondingly, with the pressure force misaligned with topographic gradients, the fluid column experiences a twisting action akin to how baroclinicity spins a fluid element if the pressure force does not act through the fluid element's center of mass (see Section 6.4).

---

<sup>5</sup>Note that we derived the expression (10.44) for  $w_g$  in equation (6.196) when studying vorticity mechanics.

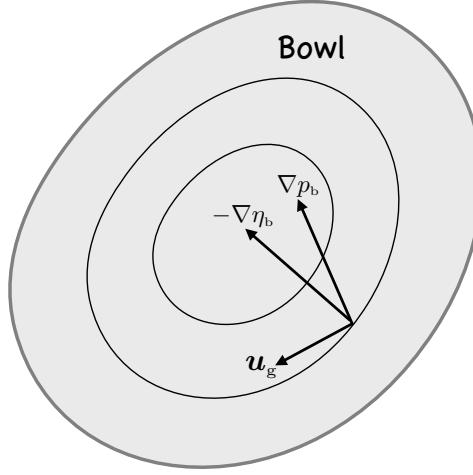


FIGURE 10.1: Depicting how bottom pressure gradients create vertical motion in planetary geostrophic flow next to a sloping bottom according to equations (10.44) and (10.46). Here we show a bowl or depression (local maximum in the depth) with  $-\nabla \eta_b$  increasing inward toward the bowl center. Only those portions of  $\nabla p_b$  and  $\mathbf{F}_b$  that are aligned parallel to the topographic slope contribute to vertical motion. We illustrate here a case where the bottom pressure gradient leads to  $-\nabla \eta_b \cdot \mathbf{u}_g > 0$  so that  $w_g < 0$  in the northern hemisphere and  $w_g > 0$  in the southern hemisphere.

To illustrate the above, consider the topographic bowl in Figure 10.1, with sides steep enough so that the bottom pressure gradient is dominated by the topographic slopes. Along the bottom the pressure increases moving down (increasing depth) towards the bowl center. The corresponding bottom geostrophic flow is anti-cyclonic within the bowl and largely follows isobaths. As already noted, if the geostrophic flow exactly follows isobaths, then there is no corresponding vertical component to the bottom velocity. A vertical velocity arises only in the presence of an anomalous bottom pressure gradient,  $\nabla p'_b$ , that is misaligned with the bottom slope,  $\nabla \eta_b$ . This bottom pressure gradient balances a geostrophic flow that deviates from isobaths thus giving rise to a nonzero  $w_g$ . Similar geometric analysis holds for the bottom friction vector,  $\mathbf{F}_b$ , and how it gives rise to a nonzero vertical Ekman velocity,  $w_e$ .

### What causes misalignment between $p_b$ and $\eta_b$ ?

As we just discussed, misalignment of  $p_b$  and  $\eta_b$  lead to vertical geostrophic motion along the bottom. In Section 10.4.3 we will see a similar relation for vertical geostrophic motion at the ocean surface. But what causes such misalignment? The answer to this question is circular when working within planetary geostrophy since its momentum equation is diagnostic. Even so, we can offer some insight by returning to the depth-integrated vorticity balance (10.34) and rewriting it as an expression for vertical motion

$$\rho_0 f [w(\eta_b) - w(\eta)] = -\rho_0 \beta V + \hat{\mathbf{z}} \cdot (\nabla \times \Delta \boldsymbol{\tau}). \quad (10.47)$$

Hence, vertical motion at the surface and bottom balance meridional motion in the presence of planetary beta, plus the curl of surface and bottom stresses. The absence of planetary beta, and the absence of boundary stress curls, realizes  $w(\eta_b) = w(\eta)$ .

### 10.4.3 Surface kinematics and dynamics

For purposes of large-scale circulation studies using planetary geostrophy, it is generally sufficient to assume a rigid lid upper boundary condition, whereby  $w(\eta) = w(0) = 0$ . Even so, we find it

revealing to present the results for a free surface in which there is the possibility of nonzero surface mass fluxes. This situation is commonly encountered in general circulation models. The surface kinematic boundary condition for a Boussinesq fluid is derived in VOLUME 2, and it is given by

$$w = -Q_m/\rho_0 + (\partial_t + \mathbf{u} \cdot \nabla)\eta \quad \text{at } z = \eta(x, y, t). \quad (10.48)$$

We retain the sea surface time tendency,  $\partial_t\eta$ , even though for transient solutions the time tendency is many orders of magnitude smaller than the typical vertical velocity under the planetary geostrophic regime.<sup>6</sup> Evaluating the horizontal planetary geostrophic momentum equation at the ocean surface renders

$$\rho_0 f \hat{\mathbf{z}} \times \mathbf{u} = -(\nabla_h p)_{z=\eta} + \mathbf{F}_\eta, \quad (10.49)$$

where  $\mathbf{F}_\eta$  is the horizontal friction vector at the surface. In our study of hydrostatic pressure in VOLUME 2, we derived the following expression for the horizontal pressure gradient at the ocean surface

$$(\nabla_h p)_{z=\eta} = \nabla_h p_a + g \rho(\eta) \nabla_h \eta \quad (10.50)$$

is the horizontal pressure gradient at the ocean surface. Like the bottom, we decompose the horizontal velocity into a geostrophic component and an Ekman component via

$$\rho_0 f \hat{\mathbf{z}} \times \mathbf{u}_g = -(\nabla_h p)_{z=\eta} \quad \text{and} \quad \rho_0 f \hat{\mathbf{z}} \times \mathbf{u}_e = \mathbf{F}_\eta, \quad (10.51)$$

so that

$$\mathbf{u}_g = \frac{\hat{\mathbf{z}} \times (\nabla_h p)_{z=\eta}}{\rho_0 f} \quad \text{and} \quad \mathbf{u}_e = -\frac{\hat{\mathbf{z}} \times \mathbf{F}_\eta}{\rho_0 f}. \quad (10.52)$$

The corresponding vertical velocity components are determined by inserting into the surface kinematic boundary condition (10.48)

$$w_{Q\eta} = -\frac{Q_m}{\rho_0} + \frac{\partial \eta}{\partial t} \quad \text{and} \quad w_g = \frac{\hat{\mathbf{z}} \cdot (-\nabla \eta \times \nabla p_a)}{\rho_0 f} \quad \text{and} \quad w_e = \frac{\hat{\mathbf{z}} \cdot (\nabla \eta \times \mathbf{F}_\eta)}{\rho_0 f}, \quad (10.53)$$

where we introduced a vertical velocity,  $w_{Q\eta}$ , associated with the boundary mass flux and transient sea level fluctuations. As for the bottom, the second and third of these equations reveal how the curl of inviscid pressure forces and boundary frictional forces drive a nonzero vertical motion at the ocean surface, all while maintaining the surface kinematic boundary condition (10.48). Furthermore, it is only the projection of  $\nabla p_a$  and  $\mathbf{F}_\eta$  onto the direction parallel to sea surface height contours that contributes to a nonzero vertical velocity. This orientation of the surface forces is needed to render a horizontal velocity that is itself misaligned with sea surface height contours, thus satisfying the kinematics required to render vertical motion.

#### 10.4.4 Summary of force curls driving depth integrated meridional flow

Plugging expressions (10.44), (10.46), and (10.53) into equation (10.34) renders the depth integrated planetary vorticity balance

$$\rho_0 \beta V = f(-Q_m + \rho_0 \partial_t \eta) + \hat{\mathbf{z}} \cdot [\nabla \eta \times (\mathbf{F}_\eta - \nabla p_a) - \nabla \eta_b \times (\mathbf{F}_b - \nabla p'_b) + \nabla \times \Delta \boldsymbol{\tau}]. \quad (10.54)$$

---

<sup>6</sup>See Section 3.3 of [Samelson \(2011\)](#) for more details on this scaling of the planetary geostrophic equations related to this point.

Evidently, the force curls are symmetrically applied at the surface and bottom, which is part of the motivation for exposing the surface terms even though they are generally subdominant. In general, equation (10.54) shows that the depth integrated meridional flow, within the planetary geostrophic regime, is driven by the following processes.

### Surface mass transport plus sea surface fluctuations

The term  $f \rho_o (-Q_m/\rho_o + \partial_t \eta)$  arises from mass transport across the ocean surface plus fluctuations in the sea surface height. For example, as sea surface height increases or as water leaves the ocean surface, they impart a positive surface vertical velocity,  $w(\eta) > 0$ , thus causing column stretching and poleward meridional depth integrated flow. In the steady state, where it is just the mass flux term that contributes, the meridional circulation is known as the Goldsborough-Stommel circulation (see [Huang and Schmitt \(1993\)](#) for a review).

### Curl of turbulent boundary stresses

The term  $\hat{z} \cdot (\nabla \times \Delta \tau)$  arises from the curl of the turbulent wind stress and turbulent bottom stress. The wind stress term is generally larger than the bottom turbulent stress, with many theories for ocean circulation, particularly those with a flat bottom, almost exclusively focused on the role of surface stress in forcing planetary geostrophic vorticity. We comment more on this case in Section 10.5 where we discuss the [Sverdrup balance](#).

### Atmospheric pressure torque

The term

$$\nabla p_a \times \nabla \eta = \nabla \times (p_a \nabla \eta) = -\nabla \times (\eta \nabla p_a) \quad (10.55)$$

arises from differences in lines of constant atmospheric pressure and lines of constant sea surface height. Such misalignments create a torque akin to the baroclinicity detailed in Section 6.4, with these misalignments driving vertical motion and a corresponding depth integrated meridional flow.

### Bottom pressure torque

The term

$$\nabla p_b' \times (-\nabla \eta_b) = \nabla \times (-p_b' \nabla \eta_b) = \nabla \times (\eta_b \nabla p_b') \quad (10.56)$$

arises from differences in lines of constant bottom pressure and lines of constant bottom topography. That is, bottom pressure torque requires a gradient of bottom pressure along isobaths, thus producing a bottom geostrophic flow that deviates from isobaths. As for the atmospheric pressure torques, such misalignments create a torque that drives a depth integrated meridional flow, with this term vanishing when the bottom topography is flat. In many cases with strong flow next to sloping bottoms, this term can contribute more to the vorticity budget than the turbulent bottom stress. Indeed, in some cases it can rival contributions from the surface wind stress. We sketched out such cases for the shallow water when discussing western boundary currents in Section 5.7.6.

### Surface frictional acceleration

The term

$$\hat{z} \cdot (\nabla \eta \times \mathbf{F}_\eta) = \hat{z} \cdot [\nabla \eta \times \partial_z \boldsymbol{\tau}_\eta] \quad (10.57)$$

arises from evaluating the vertical divergence of the frictional stress at the sea surface. A finite volume boundary layer treatment of this term prompts us to integrate the stress divergence over the extent of the surface Ekman layer to render the alternative expression

$$\hat{z} \cdot \int_{-h_{\text{e-surf}}}^{\eta} [\nabla \eta \times \partial_z \boldsymbol{\tau}] dz = \hat{z} \cdot [\nabla \eta \times \boldsymbol{\tau}(\eta)], \quad (10.58)$$

where we assumed  $\boldsymbol{\tau}$  is negligible at the base of the surface Ekman layer,  $z = -h_{\text{e-surf}}(x, y, t)$ . The term  $\hat{z} \cdot [\nabla \eta \times \boldsymbol{\tau}(\eta)]$  creates a torque from that component of the surface turbulent stress that is aligned with isolines of the sea surface height.

### Bottom frictional acceleration

The term

$$\hat{z} \cdot (-\nabla \eta_b \times \mathbf{F}_b) = \hat{z} \cdot \left[ -\nabla \eta_b \times \frac{\partial \boldsymbol{\tau}_b}{\partial z} \right] \quad (10.59)$$

arises from evaluating the vertical divergence of the frictional stress at the ocean bottom. As for the analogous term for the surface, we offer a finite volume Ekman boundary layer treatment to render the alternative expression

$$\hat{z} \cdot \int_{\eta_b}^{h_{\text{e-bot}}} \left[ -\nabla \eta_b \times \frac{\partial \boldsymbol{\tau}}{\partial z} \right] dz = \hat{z} \cdot [-\nabla \eta_b \times \boldsymbol{\tau}(\eta_b)], \quad (10.60)$$

where we assumed  $\boldsymbol{\tau}$  is negligible at the top of the bottom Ekman layer,  $z = -\eta_b + h_{\text{e-bot}}$ . The term  $\hat{z} \cdot [-\nabla \eta_b \times \boldsymbol{\tau}(\eta_b)]$  creates a torque from that component of the bottom turbulent stress that is aligned with isobaths.

### 10.4.5 Integral constraints

The atmospheric and bottom pressure torques appearing in the depth integrated planetary vorticity balance (10.54) satisfy an integral constraint that follows from Stokes' curl theorem. To illustrate this constraint, consider the bottom pressure torque integrated over an arbitrary area along the bottom

$$\int (\nabla \eta_b \times \nabla p'_b) \cdot \hat{z} dA = \int [\nabla \times (\eta_b \nabla p'_b)] \cdot \hat{z} dA \quad (10.61a)$$

$$= \oint \eta_b \nabla p'_b \cdot \hat{t} ds \quad (10.61b)$$

$$= - \oint p'_b \nabla \eta_b \cdot \hat{t} ds. \quad (10.61c)$$

We see that the bottom pressure torque vanishes when integrated around a closed loop that follows either an isobath or a bottom isobar, since the integrand vanishes identically. A similar constraint holds for the atmospheric pressure torque, whereas there is generally no analogous constraint satisfied by the turbulent boundary stresses.

One exception for the turbulent stresses occurs for  $f$ -plane flow ( $\beta = 0$ ) where the depth integrated flow is non-divergent

$$\partial_t \eta - Q_m / \rho_0 = -\nabla \cdot \mathbf{U} = 0, \quad (10.62)$$

and where the interior friction vanishes at the surface and bottom boundaries,  $\mathbf{F}_\eta = \mathbf{F}_b = 0$ .

From equation (10.54), we see that a steady state is realized only if there is a balance between pressure torques and turbulent boundary stresses

$$\hat{\mathbf{z}} \cdot (\nabla \eta \times \nabla p_a - \nabla \eta_b \times \nabla p'_b) = \hat{\mathbf{z}} \cdot (\nabla \times \Delta \boldsymbol{\tau}). \quad (10.63)$$

Consider even further specialization in which the atmospheric pressure torque vanishes, and the bottom turbulent stress is negligible, in which case a steady state balance requires a balance between bottom pressure torques and torques from surface turbulent wind stress

$$-\hat{\mathbf{z}} \cdot (\nabla \eta_b \times \nabla p'_b) = \hat{\mathbf{z}} \cdot [\nabla \times \boldsymbol{\tau}(\eta)]. \quad (10.64)$$

Integration over either an isobath or bottom pressure isobar then requires, for a steady state, the following identity

$$\int [\nabla \times \boldsymbol{\tau}(\eta)] \cdot \hat{\mathbf{z}} dA = \oint \boldsymbol{\tau}(\eta) \cdot \hat{\mathbf{t}} ds = 0 \quad f\text{-plane closed isobath or closed isobar}. \quad (10.65)$$

Deviation from this identity leads to non-steady flow on the  $f$ -plane. In contrast, the  $\beta$ -plane has no such steady state constraint since meridional flow can balance the circulation imparted by turbulent wind stresses.

## 10.5 Sverdrup balance and geostrophic Sverdrup balance

The balance (10.54) exposes the many processes that affect meridional flow in a planetary geostrophic fluid. However, when confronted with minimal information from ocean measurements, we are motivated to examine just the main contributors to this balance.

The **Sverdrup balance** is a simplified form of the balance (10.54), and it was first encountered in our study of geostrophy in VOLUME 2

$$\rho_0 \beta V_{\text{Sverdrup}} = \hat{\mathbf{z}} \cdot [\nabla \times \boldsymbol{\tau}(\eta)] \quad \text{Sverdrup balance}. \quad (10.66)$$

This balance arises from dropping the vertical velocity at both the ocean surface and ocean bottom; ignoring horizontal frictional stresses; and assuming  $\partial_t \eta = -Q_m / \rho_0 = 0$  as per a rigid lid flow in which  $\nabla \cdot \mathbf{U} = 0$ . The Sverdrup balance offers a null hypothesis for the large-scale and low frequency meridional ocean circulation away from sloping sides; i.e., in regions where bottom pressure torques can be ignored.

To derive the Sverdrup balance (10.66), we performed a depth integral of the planetary geostrophic vorticity equation (10.33) from the ocean bottom to the free surface. This integral encompasses both the geostrophic interior and the ageostrophic Ekman flow in the top and bottom Ekman layers. In some treatments we focus exclusively on contributions from the geostrophic interior, in which case the depth integral extends from the top of the bottom Ekman layer,  $z = \eta_{eb}$ , to the bottom of the top Ekman layer,  $z = \eta_{et}$ , thus leading to the depth integrated geostrophic transport

$$V_g \equiv \int_{\eta_{eb}}^{\eta_{et}} v dz. \quad (10.67)$$

Integrating the planetary geostrophic vorticity balance (10.33) over this depth range, and ignoring contributions from friction since we are concerned just with the geostrophic interior, leads to

$$\rho_0 \beta V_g = f [w(\eta_{et}) - w(\eta_{eb})]. \quad (10.68)$$

This equation provides a balance between the depth integrated meridional transport within the geostrophic interior (left hand side), with the vertical vortex stretching within this depth range (right hand side).

We now make use of Ekman mechanics from VOLUME 2 to approximate the vertical velocities in equation (10.68). For this purpose we neglect both the time tendency for the vertical position of the Ekman layer and the slope of the Ekman layer base, in which case the kinematic expression for  $w^{(\dot{\eta}_e)}$  derived in VOLUME 2, and applied at the Ekman base,  $z = \eta_e(x, y, t)$ , is approximated by

$$w^{(\dot{\eta}_e)} = w - (\partial_t z + \mathbf{u} \cdot \nabla z) \approx w \quad \text{at } z = \eta_e(x, y, t), \quad (10.69)$$

where  $z = \eta_e$  is the vertical position of the Ekman layer. We now make use of the Ekman layer mass budget to determine the entrainment velocity,  $w^{(\dot{\eta}_e)}$ . In particular, we derived the following mass budget equation

$$\rho_0 w^{(\dot{\eta}_e)} + Q_m = \rho_0 \partial_t h_e + \nabla_h \cdot \mathbf{M}_{\text{other}} + \hat{\mathbf{z}} \cdot [\nabla_h \times (\boldsymbol{\tau}/f)], \quad (10.70)$$

which we use here for  $w(\eta_{\text{et}})$  to give

$$w(\eta_{\text{et}}) \approx w_{\text{Ekman-top}}^{(\dot{\eta}_e)} = (1/\rho_0) \hat{\mathbf{z}} \cdot [\nabla \times (\boldsymbol{\tau}(\eta)/f)]. \quad (10.71)$$

with a similar treatment for the bottom leading to

$$w(\eta_{\text{eb}}) \approx w_{\text{Ekman-bot}}^{(\dot{\eta}_e)} = (1/\rho_0) f \hat{\mathbf{z}} \cdot [\nabla \times (\boldsymbol{\tau}(\eta_b)/f)]. \quad (10.72)$$

Bringing these results into equation (10.68) leads to

$$\rho_0 \beta V_g = f \hat{\mathbf{z}} \cdot \nabla \times [\boldsymbol{\tau}(\eta)/f - \boldsymbol{\tau}(\eta_b)/f]. \quad (10.73)$$

Since the bottom turbulent stress is generally much smaller than the surface, it is typically ignored, in which case we reach the **geostrophic Sverdrup balance**

$$\rho_0 \beta V_g = f \hat{\mathbf{z}} \cdot \nabla \times [\boldsymbol{\tau}(\eta)/f] \quad \text{geostrophic Sverdrup balance.} \quad (10.74)$$

The geostrophic Sverdrup balance relates the meridional geostrophic transport (left hand side) to the curl of  $\boldsymbol{\tau}/f$  due to upper ocean mechanical stresses from boundary processes (e.g., wind stress and ice-ocean stresses). It differs from the full **Sverdrup balance** in equation (10.66) through presence of the  $f$  outside of the curl and  $1/f$  inside the curl. [Gray and Riser \(2014\)](#) assess the geostrophic Sverdrup balance based on ocean measurements.

## 10.6 Vorticity of the depth integrated velocity

In Section 10.4 we studied the depth integrated vorticity budget for planetary geostrophic flow. We were led to see how boundary torques (i.e., the curl of boundary forces) lead to vertical motion and in turn, through the  $\beta$ -effect, lead to meridional motion of the depth integrated flow. In this section we present another analysis of vorticity in the planetary geostrophic regime, here focusing on vorticity of the depth integrated velocity. Elements of this material were discussed in Section 6.9 without making the planetary geostrophic assumption. By assuming planetary geostrophy we can further constrain the flow by focusing just on vortex stretching.

### 10.6.1 Depth integrated velocity equation

The depth integrated horizontal velocity equation (10.5a) is given by

$$\rho_0 f \hat{z} \times \mathbf{U} = - \int_{\eta_b}^{\eta} \nabla_h p dz + \Delta\tau \quad (10.75)$$

where

$$\mathbf{U} = \int_{\eta_b}^{\eta} \mathbf{u} dz \quad (10.76)$$

is the depth integrated horizontal velocity, and we assumed friction in the form of the vertical divergence of a horizontal turbulent stress as in equation (10.32). For the depth integrated pressure gradient, we follow the decomposition used in our study of form stress in VOLUME 2, in which we write

$$\int_{\eta_b}^{\eta} p dz = \int_{\eta_b}^{\eta} [d(pz) - z dp] = p_a \eta - p_b \eta_b + \mathcal{P}, \quad (10.77)$$

where we used the hydrostatic balance to write  $dp = -g \rho dz$ , which is valid for each fluid column. We also introduced the potential energy per horizontal area of a fluid column

$$\mathcal{P} = \int_{\eta_b}^{\eta} g \rho z dz = (g \rho_0 / 2) (\eta^2 - \eta_b^2) + \int_{\eta_b}^{\eta} g \rho' z dz, \quad (10.78)$$

where

$$\rho' = \rho - \rho_0 \quad (10.79)$$

is the density deviation from the background reference density. These results then lead to the depth integrated horizontal pressure gradient

$$\int_{\eta_b}^{\eta} \nabla_h p dz = \nabla_h \left[ \int_{\eta_b}^{\eta} p dz \right] - p_a \nabla_h \eta + p_b \nabla_h \eta_b \quad (10.80a)$$

$$= \nabla_h [p_a \eta - p_b \eta_b + \mathcal{P}] - p_a \nabla_h \eta + p_b \nabla_h \eta_b \quad (10.80b)$$

$$= \eta \nabla_h p_a - \eta_b \nabla_h p_b + \nabla_h \mathcal{P}, \quad (10.80c)$$

$$= \eta \nabla p_a - \eta_b \nabla p_b + \nabla \mathcal{P}, \quad (10.80d)$$

where the final equality follows since  $p_a$ ,  $\eta_b$ ,  $p_b$ , and  $\mathcal{P}$ , are functions just of horizontal position and time. We are thus led to the depth integrated planetary geostrophic momentum balance

$$\rho_0 f \hat{z} \times \mathbf{U} = -\eta \nabla p_a + \eta_b \nabla p_b - \nabla \mathcal{P} + \Delta\tau. \quad (10.81)$$

The depth integrated balance is here written in terms of gradients in the surface and bottom pressures, the gradient of the potential energy per area, and the difference in turbulent stresses at the top and bottom boundaries,  $\Delta\tau = \tau(\eta) - \tau(\eta_b)$ .

### 10.6.2 Vorticity budget

Taking the curl of the depth integrated balance (10.81) annihilates the potential energy term, thus leaving

$$\rho_0 \beta V = -\rho_0 f \nabla \cdot \mathbf{U} + \hat{z} \cdot \nabla \times [p_a \nabla \eta + \tau(\eta) - p_b \nabla \eta_b - \tau(\eta_b)]. \quad (10.82)$$

From our study of the kinematics for a non-divergent flow in VOLUME 2, we know that the divergence of the depth-integrated flow for a steady Boussinesq ocean is given by

$$\rho_0 \nabla \cdot \mathbf{U} = Q_m, \quad (10.83)$$

so that

$$\rho_0 \beta V = -f Q_m + \hat{\mathbf{z}} \cdot \nabla \times [p_a \nabla \eta + \boldsymbol{\tau}(\eta) - p_b \nabla \eta_b - \boldsymbol{\tau}(\eta_b)]. \quad (10.84)$$

This is the vorticity equation for the depth integrated planetary geostrophic fluid. It is quite similar to the vorticity balance for a shallow water fluid as given by equation (5.106) (which considered zero atmospheric pressure). In the presence of  $\beta$ , meridional mass transport for the fluid column is balanced by surface mass fluxes,  $Q_m \neq 0$ ; the curl of surface form stresses and surface turbulent stresses; and the curl of topographic form stresses and bottom turbulent stresses. This result follows quite naturally when recognizing that the forces acting on a depth integrated fluid column arise from the depth integrated stresses acting on the column sides plus those acting on the top and bottom boundaries. In the absence of interior friction stresses due to horizontal strains (as assumed here), it is only the depth integrated pressure that acts on the column sides, and this term has zero curl. We are thus left just with the curl of the surface and bottom boundary form stresses and turbulent stresses, along with a contribution from boundary mass transport.

The Sverdrup balance (10.66) is a special case of the more complete vorticity budget (10.82). Furthermore, the **topographic topographic Sverdrup balance** results when meridional transport balances the curl of the wind plus bottom pressure form stress

$$\rho_0 \beta V_{\text{topo-Sverdrup}} = \hat{\mathbf{z}} \cdot \nabla \times (\boldsymbol{\tau}(\eta) - p_b \nabla \eta_b) \quad \text{topographic Sverdrup balance.} \quad (10.85)$$

In the presence of flows interacting with topography, where bottom pressure torques are sizable, this balance is generally much more accurate than the Sverdrup balance.

### 10.6.3 Integral balances for steady flows

Following the discussion in Section 6.9.5, we write the vorticity balance (10.82) in the form

$$\rho_0 \nabla \cdot (f \mathbf{U}) = \hat{\mathbf{z}} \cdot \nabla \times [p_a \nabla \eta + \boldsymbol{\tau}(\eta) - p_b \nabla \eta_b - \boldsymbol{\tau}(\eta_b)]. \quad (10.86)$$

Integrating over a horizontal area,  $\mathcal{S}$ , leads to

$$\rho_0 \oint_{\partial\mathcal{S}} f \mathbf{U} \cdot \hat{\mathbf{n}} \, ds = \oint_{\partial\mathcal{S}} [p_a \nabla \eta + \boldsymbol{\tau}(\eta) - p_b \nabla \eta_b - \boldsymbol{\tau}(\eta_b)] \cdot \hat{\mathbf{t}} \, ds. \quad (10.87)$$

To reach this result we used Gauss's divergence theorem for the left hand side and Stokes' curl theorem for the right hand side. The unit vector  $\hat{\mathbf{n}}$  points horizontally outward from the boundary of the area, whereas the unit vector  $\hat{\mathbf{t}}$  is the counter-clockwise oriented tangent to the closed contour around the boundary. For the special case of  $Q_m = 0$  we are afforded a steady state streamfunction for the depth-integrate flow since  $\nabla \cdot \mathbf{U} = 0$ . Choosing the area,  $\mathcal{S}$ , to be bounded by a closed streamline allows us to set  $\mathbf{U} \cdot \hat{\mathbf{n}} = 0$  along that streamline. We thus see that for steady planetary geostrophic flow with  $\nabla \cdot \mathbf{U} = 0$ , any closed streamline of the flow must maintain the following work balance around the streamline

$$\oint_{\partial\mathcal{S}} [p_a \nabla \eta + \boldsymbol{\tau}(\eta) - p_b \nabla \eta_b - \boldsymbol{\tau}(\eta_b)] \cdot \hat{\mathbf{t}} \, ds = 0. \quad (10.88)$$

This equation is a simplified form of equation (6.208) that was formulated for a more general flow. We can rearrange equation (10.88) to display an integrated balance between the work done by boundary pressure form stresses around a closed streamline, and work done by boundary turbulent stresses around the same streamline

$$\oint_{\partial S} (p_a \nabla \eta - p_b \nabla \eta_b) \cdot \hat{\mathbf{t}} \, ds = - \oint_{\partial S} [\boldsymbol{\tau}(\eta) - \boldsymbol{\tau}(\eta_b)] \cdot \hat{\mathbf{t}} \, ds. \quad (10.89)$$

#### 10.6.4 Comments

*Yeager* (2015) connects torques acting on the depth integrated horizontal flow in the North Atlantic to buoyancy forces affecting the Atlantic meridional overturning circulation, thus illustrating how the formulation in this section can be of use for the analysis of an ocean circulation model.

## 10.7 Vorticity equation for the depth averaged velocity

In some numerical models, it is more common to have access to the depth averaged velocity

$$\bar{\mathbf{u}} = \frac{1}{D} \int_{\eta_b}^{\eta} \mathbf{u} \, dz \quad \text{with} \quad D = \eta - \eta_b. \quad (10.90)$$

We thus find it useful to develop a budget for the vorticity of the depth averaged velocity. In this discussion we encounter a distinct means for describing how bottom topography, in the presence of baroclinicity, generates meridional flow.

#### 10.7.1 Relating the depth average velocity to boundary velocities

Before studying the vorticity equation, we here relate the depth averaged velocity,  $\bar{\mathbf{u}}(x, y, t)$ , to the surface velocity,  $\mathbf{u}(x, y, z = \eta, t)$ , and bottom velocity,  $\mathbf{u}(x, y, z = \eta_b, t)$ . This analysis exposes some general features of how the boundary flows are driven away from the depth average.

The starting point is the identity

$$\int_{\eta_b}^{\eta} \mathbf{u} \, dz = (\eta - \eta_b) \mathbf{u}(\eta) - \int_{\eta_b}^{\eta} \frac{\partial \mathbf{u}}{\partial z} (z - \eta_b) \, dz, \quad (10.91)$$

which, along with the analogous identity for the bottom flow, leads to

$$\bar{\mathbf{u}} - \mathbf{u}(\eta) = - \int_{\eta_b}^{\eta} \frac{\partial \mathbf{u}}{\partial z} \left[ \frac{z - \eta_b}{\eta - \eta_b} \right] \, dz \quad \text{and} \quad \bar{\mathbf{u}} - \mathbf{u}(\eta_b) = \int_{\eta_b}^{\eta} \frac{\partial \mathbf{u}}{\partial z} \left[ \frac{\eta - z}{\eta - \eta_b} \right] \, dz. \quad (10.92)$$

Note that  $\mathbf{u}(\eta) - \mathbf{u}(\eta_b) = \int_{\eta_b}^{\eta} (\partial \mathbf{u} / \partial z) \, dz$  serves as a useful check on the manipulations leading to equation (10.92). We see that the difference between the depth averaged flow and the surface flow,  $\bar{\mathbf{u}} - \mathbf{u}(\eta)$ , is determined by the integral of the weighted vertical shear, with the weight linearly decreasing from unity at the surface to zero at the bottom. The minus sign in front of the integral for  $\bar{\mathbf{u}} - \mathbf{u}(\eta)$  follows because if the flow increases in the positive direction from the bottom to the surface, then the depth averaged flow will have a smaller magnitude than the surface flow. The converse weighting holds for computing the difference  $\bar{\mathbf{u}} - \mathbf{u}(\eta_b)$ .

The identities (10.92) hold for arbitrary horizontal velocity fields. Assuming the flow satisfies frictional geostrophy as per equation (10.5a) leads to the frictional thermal wind relation

$$f \mathbf{u} = \rho_o^{-1} \hat{\mathbf{z}} \times \nabla p - \hat{\mathbf{z}} \times \mathbf{F} \implies f \partial_z \mathbf{u} = -(g/\rho_o) \hat{\mathbf{z}} \times \nabla \rho - \hat{\mathbf{z}} \times \partial_z \mathbf{F}, \quad (10.93)$$

so that the velocity differences are given by

$$f [\bar{\mathbf{u}} - \mathbf{u}(\eta)] = \int_{\eta_b}^{\eta} [(g/\rho_o) \hat{\mathbf{z}} \times \nabla \rho - \hat{\mathbf{z}} \times \partial_z \mathbf{F}] \left[ \frac{z - \eta_b}{\eta - \eta_b} \right] dz \quad (10.94a)$$

$$f [\bar{\mathbf{u}} - \mathbf{u}(\eta_b)] = - \int_{\eta_b}^{\eta} [(g/\rho_o) \hat{\mathbf{z}} \times \nabla \rho - \hat{\mathbf{z}} \times \partial_z \mathbf{F}] \left[ \frac{\eta - z}{\eta - \eta_b} \right] dz. \quad (10.94b)$$

Hence, differences between the depth averaged flow and the boundary flows are determined by weighted integrals of the baroclinicity and vertical friction shears.

### 10.7.2 Formulation of the vorticity equation

To develop the vorticity equation, we start by deriving the momentum equation for the depth averaged flow. For that purpose, rearrange the depth integrated momentum budget (10.81) according to

$$\rho_o f \hat{\mathbf{z}} \times \mathbf{U} = -\eta \nabla_h (p_a - p_b) - D \nabla p_b - \nabla \mathcal{P} + \Delta \boldsymbol{\tau}, \quad (10.95)$$

and then divide by the column thickness,  $D = \eta - \eta_b$ , to render

$$\rho_o f \hat{\mathbf{z}} \times \bar{\mathbf{u}} = -(\eta/D) \nabla (p_a - p_b) - \nabla p_b + (1/D) (-\nabla \mathcal{P} + \Delta \boldsymbol{\tau}). \quad (10.96)$$

Taking the curl then leads to

$$\rho_o \nabla \cdot (f \bar{\mathbf{u}}) = -\hat{\mathbf{z}} \cdot \nabla \times [(\eta/D) \nabla (p_a - p_b)] + D^{-2} \hat{\mathbf{z}} \cdot [\nabla \times (D \nabla \mathcal{P})] + \hat{\mathbf{z}} \cdot [\nabla \times (D^{-1} \Delta \boldsymbol{\tau})] \quad (10.97a)$$

$$= \hat{\mathbf{z}} \cdot \nabla \times [(p_a - p_b) \nabla (\eta/D)] - D^{-2} \hat{\mathbf{z}} \cdot [\nabla \times (\mathcal{P} \nabla D)] + \hat{\mathbf{z}} \cdot [\nabla \times (D^{-1} \Delta \boldsymbol{\tau})]. \quad (10.97b)$$

The vorticity budget (10.97) is a bit less tidy than that for the depth integrated budget (10.84). Nonetheless, it offers some novel insights concerning the flow, which can be seen by writing the left hand side in the form

$$\nabla \cdot (f \bar{\mathbf{u}}) = \nabla \cdot [(f/D) \mathbf{U}], \quad (10.98)$$

with  $f/D$  reminiscent of the shallow water potential vorticity for the planetary geostrophic flow (Section 9.4). Motivated by this analog, we write the vorticity equation (10.97) in the form

$$\rho_o \mathbf{U} \cdot \nabla (f/D) = -\rho_o (f/D) \nabla \cdot \mathbf{U} + \hat{\mathbf{z}} \cdot \nabla \times [(p_a - p_b) \nabla (\eta/D)] - D^{-2} \hat{\mathbf{z}} \cdot [\nabla \times (\mathcal{P} \nabla D)] + \hat{\mathbf{z}} \cdot [\nabla \times (D^{-1} \Delta \boldsymbol{\tau})]. \quad (10.99)$$

Contrary to the shallow water case, we here see that even for a perfect planetary geostrophic fluid, the depth-integrated flow does not generally follow contours of constant  $f/D$ . Even so, it is of interest to examine how the processes on the right hand side contribute to flow deviations from  $f/D$  contours. For that purpose we simplify the flow even more by making the rigid lid

approximation.

### 10.7.3 Rigid lid approximation and the role of JEBAR

The rigid lid approximation is commonly made for studies of large-scale circulation. Indeed, it was the basis for many ocean general circulation models following the work of [Bryan \(1969\)](#). A fluid satisfying the rigid lid approximation has a vanishing horizontal divergence for the depth integrated flow

$$\text{rigid lid approximation} \implies \nabla \cdot \mathbf{U} = 0. \quad (10.100)$$

Furthermore, as part of the rigid lid approximation we assume the free surface undulations are much smaller than the resting ocean depth so that

$$|\eta| \ll |\eta_b| \implies \eta/D \approx 0 \quad (10.101a)$$

$$1/(\eta - \eta_b) \approx 1/(-\eta_b) = 1/H. \quad (10.101b)$$

We introduced

$$H = -\eta_b \quad (10.102)$$

to correspond to the literature for rigid lid models and the JEBAR term.<sup>7</sup> Hence, with the rigid lid approximation the vorticity equation (10.99) takes the simplified form

$$\mathbf{U} \cdot \nabla(f/H) = \rho_0^{-1} \hat{z} \cdot [\nabla \times (\mathcal{P} \nabla H^{-1}) + \nabla \times (H^{-1} \Delta \boldsymbol{\tau})]. \quad (10.103)$$

Each term in this equation has dimensions of inverse squared time,  $T^{-2}$ .

#### JEBAR drives deviations from $f/H$ aligned flow

The first term on the right hand side of the vorticity equation (10.103) is referred to as the Joint Effect of Baroclinicity and Relief (JEBAR)

$$\text{JEBAR} \equiv \rho_0^{-1} \hat{z} \cdot [\nabla \times (\mathcal{P} \nabla H^{-1})] = \rho_0^{-1} \hat{z} \cdot [\nabla \mathcal{P} \times \nabla H^{-1}]. \quad (10.104)$$

This name arises since JEBAR is nonzero only in the presence of non-flat topography (“relief”) and (as shown below) when density is not a constant (i.e., for baroclinic flow). In addition to contributions from boundary stresses, equation (10.103) says that misalignment of  $f/H$  contours with the depth-integrated steady rigid lid flow is driven by misalignments of isobaths and isolines of the depth integrated potential energy.

Contributions to JEBAR arise only from the component of the potential energy that deviates from a constant density reference state. To show that property, note that in the rigid lid approximation, the potential energy in a column, as given by equation (10.78), takes the form

$$\mathcal{P} = (1/2) g \rho_0 H^2 + g \int_{-H}^0 \rho' z dz, \quad (10.105)$$

where  $\rho' = \rho - \rho_0$ . With  $\nabla H^2 \times \nabla(1/H) = 0$ , we are left with just the contribution from  $\rho'$

$$\text{JEBAR} = \rho_0^{-1} \hat{z} \cdot \left[ \nabla \left( g \int_{-H}^0 \rho' z dz \right) \times \nabla H^{-1} \right]. \quad (10.106)$$

---

<sup>7</sup>Note that  $H$  is the vertical depth scale, which is a constant, and it is distinct from  $H(x, y) = -\eta_b(x, y)$ , which is a function of horizontal position.

JEBAR vanishes for a homogeneous density field, where  $\rho'$  is a constant, but is nonzero with a nonzero  $\rho'$ . In a Boussinesq ocean, a nonzero  $\rho'$  is associated with baroclinicity (Figure 6.10).

### Relating JEBAR to pressure

JEBAR as given by equation (10.104) has the appearance of the curl of a form stress, and yet it is not. The reason is that  $\mathcal{P}$  is the potential energy of the fluid column rather than bottom pressure. We make this point explicit by recalling the decomposition (10.77), here specialized to the rigid lid in which

$$\mathcal{P} = H(\bar{p} - p_b) \quad \text{with} \quad \bar{p} = H^{-1} \int_{-\infty}^0 p dz = (-\eta_b)^{-1} \int_{\eta_b}^0 p dz. \quad (10.107)$$

Plugging into the vorticity equation (10.103) leads to

$$\rho_o \mathbf{U} \cdot \nabla(f/H) = H^{-1} \hat{\mathbf{z}} \cdot [\nabla H \times \nabla(\bar{p} - p_b)] + \hat{\mathbf{z}} \cdot \nabla \times (H^{-1} \Delta \boldsymbol{\tau}). \quad (10.108)$$

Evidently, JEBAR is nonzero where the depth averaged pressure deviates from the bottom pressure

$$\text{JEBAR} = \rho_o^{-1} \hat{\mathbf{z}} \cdot [\nabla \mathcal{P} \times \nabla H^{-1}] = (\rho_o H)^{-1} \hat{\mathbf{z}} \cdot [\nabla H \times \nabla(\bar{p} - p_b)]. \quad (10.109)$$

We can go one more step in exposing the pressure difference through use of equation (10.92), here applied to hydrostatic pressure, in which case (remember that  $H = -\eta_b$ )

$$\bar{p} - p_b = \int_{\eta_b}^0 \frac{\partial p}{\partial z} \frac{z}{\eta_b} dz = (g/H) \int_{-\infty}^0 z \rho dz = H^{-1} \mathcal{P}, \quad (10.110)$$

so that<sup>8</sup>

$$\nabla_h(\bar{p} - p_b) = -g \rho(z = -H) \nabla_h H + g \int_{-\infty}^0 \nabla_h(\rho z/H) dz. \quad (10.111)$$

The first term arises from slopes in the bottom topography as weighted by the bottom density, whereas the second term arises from the integral of horizontal gradients in the  $z/H$  weighted density.

Alternatively, we can introduce a geostrophic velocity associated with the gradient of the bottom pressure as well as the vertically averaged pressure

$$\rho_o f \hat{\mathbf{z}} \times \mathbf{u}_{gb} = -\nabla p_b \quad \text{and} \quad \rho_o f \hat{\mathbf{z}} \times \bar{\mathbf{u}}_g = -\nabla \bar{p} \implies \nabla(\bar{p} - p_b) = \rho_o f \hat{\mathbf{z}} \times (\mathbf{u}_{gb} - \bar{\mathbf{u}}_g). \quad (10.112)$$

Doing so brings JEBAR from equation (10.109) into the form

$$\text{JEBAR} = (\rho_o H)^{-1} \hat{\mathbf{z}} \cdot [\nabla H \times \nabla(\bar{p} - p_b)] = -f H^{-1} \nabla H \cdot (\bar{\mathbf{u}}_g - \mathbf{u}_{gb}). \quad (10.113)$$

JEBAR thus arises from a nonzero projection onto the bottom slope of the difference between the geostrophic velocity arising from the bottom pressure and the geostrophic velocity arising from the depth averaged pressure.

Equation (10.84) provides the budget for vorticity of the depth integrated flow, in which we find the curl of the topographic form stress leads to vortex stretching. For the vorticity of the depth averaged flow, we instead encounter the JEBAR term in equation (10.108), which is not a pure vortex stretching term. Instead, it accounts for the fact that it is the horizontal

<sup>8</sup>We reintroduced the notation,  $\nabla_h$ , given the presence of  $z$  and  $\rho(z)$  in the integral in equation (10.111).

velocity flowing across isobaths,  $\mathbf{u}(z = \eta_b)$ , rather than depth averaged horizontal velocity,  $\bar{\mathbf{u}}$ , that leads to vortex stretching. Hence, when studying vorticity of the depth averaged velocity,  $\bar{\mathbf{u}}$ , accounting for the role of vortex stretching requires us to include JEBAR.

### Integral balances

Since the depth integrated flow is assumed to be non-divergent in the rigid lid approximation,  $\nabla \cdot \mathbf{U} = 0$ , we know there exists a streamfunction for this flow. Consider a region where there are closed streamlines. Following the steps in Section (10.6.3), we integrate the steady vorticity equation (10.103) around the streamline. Noting that  $\mathbf{U} \cdot \hat{\mathbf{n}} = 0$  along the streamline, where  $\hat{\mathbf{n}}$  is the horizontal unit normal to the streamline, thus renders the steady balance

$$\oint_{\partial S_{\text{streamline}}} (\mathcal{P} \nabla H^{-1} + H^{-1} \Delta \tau) \cdot \hat{\mathbf{t}} \, ds = 0. \quad (10.114)$$

#### 10.7.4 Further study

*Mertz and Wright* (1992) discuss the physics of how JEBAR relates to the curl of the topographic form stress as well as other mathematically equivalent forms. *Cane et al.* (1998) as well as Section 2.5 of *Drijfhout et al.* (2013) discusses how JEBAR can be physically misleading. For this reason, recent studies of vorticity budgets in ocean models generally eschew JEBAR, instead favoring an analysis of vorticity of the depth integrated flow as in Section 10.6 or the depth integrated vorticity from Section 10.4. *Waldman and Giordani* (2023) provide a review of vorticity analysis for the ocean.





# Chapter 11

## FOUNDATIONS OF QUASI-GEOSTROPHY

Quasi-geostrophy (QG) is the canonical example of a balanced model in geophysical fluid mechanics whereby the quasi-geostrophic potential vorticity is the sole prognostic field. All other fields, such as the velocity and buoyancy, are diagnosed from potential vorticity. The process of diagnosing the allied fields requires solving an elliptic boundary value problem to compute the geostrophic streamfunction with potential vorticity acting as the source. This connection between potential vorticity and streamfunction represents an *invertibility* principle, with the mathematical technology required for inversion shared with many other elliptic problems in mathematical physics.<sup>1</sup>

Quasi-geostrophy is an elegant theory of mathematical physics that offers physical and mathematical insights into the workings of geophysical fluid motions where rotation and stratification play leading roles. Our goal in this chapter is to provide a taste of its continuously stratified realization, offering a detailed derivation that builds from Chapters 9 and 10. We also sample some of its physical and mathematical content.

### CHAPTER GUIDE

In this chapter we extend to continuously stratified fluids the shallow water discussion of quasi-geostrophy in Chapter 9. Continuously stratified quasi-geostrophy is not concerned with the processes that affect stratification, but instead with the processes that slightly perturb that stratification. We make use of stratified geophysical fluid dynamics from VOLUME 2, as well as potential vorticity from Chapter 7. This chapter is essential for the study of Rossby waves and baroclinic instability in VOLUME 4.

<b>11.1 Loose threads . . . . .</b>	<b>376</b>
<b>11.2 Non-dimensional equations of quasi-geostrophy . . . . .</b>	<b>376</b>
11.2.1 Zeroth order asymptotic equations . . . . .	377
11.2.2 First order asymptotic equations . . . . .	378
<b>11.3 Dimensionful equations of quasi-geostrophy . . . . .</b>	<b>380</b>
11.3.1 Hydrostatic balance . . . . .	380
11.3.2 Continuity equation . . . . .	380
11.3.3 Geostrophic balance . . . . .	380
11.3.4 Material time derivative . . . . .	381
11.3.5 Buoyancy equation . . . . .	381
11.3.6 Vorticity equation . . . . .	382
11.3.7 Potential vorticity . . . . .	382

<sup>1</sup>In VOLUME 1 we study elliptic partial differential equations and their Green's functions.

---

11.3.8 Velocity equation . . . . .	384
11.3.9 Concerning the ageostrophic state in quasi-geostrophy . . . . .	384
<b>11.4 Constraints on quasi-geostrophic evolution . . . . .</b>	<b>385</b>
11.4.1 A balance between geostrophic and ageostrophic processes . . . . .	385
11.4.2 Vertical motion and the $\omega$ -equation . . . . .	386
11.4.3 Another derivation of the $\omega$ -equation . . . . .	387
<b>11.5 Connecting to Ertel potential vorticity . . . . .</b>	<b>387</b>
11.5.1 Non-dimensionalizing the Ertel potential vorticity . . . . .	388
11.5.2 Material conservation of Ertel PV to order $\text{Ro}^1$ . . . . .	388
<b>11.6 Boundary conditions . . . . .</b>	<b>389</b>
11.6.1 Concerning doubly-periodic QG models . . . . .	389
11.6.2 Buoyancy equation . . . . .	390
11.6.3 Top boundary condition . . . . .	390
11.6.4 Bottom boundary condition . . . . .	391
<b>11.7 Potential vorticity with Dirac delta sheets . . . . .</b>	<b>392</b>
11.7.1 Transition between the interior and the boundaries . . . . .	393
11.7.2 Calculating the stretching term over the extended domain . . . . .	394
11.7.3 The extended potential vorticity . . . . .	394
<b>11.8 Mathematical expressions of the theory . . . . .</b>	<b>395</b>
11.8.1 The Jacobian form of geostropic advection . . . . .	395
11.8.2 The case of constant background buoyancy frequency . . . . .	396
11.8.3 Potential vorticity induction and impermeability . . . . .	396
<b>11.9 Energetics for quasi-geostrophic flow . . . . .</b>	<b>398</b>
11.9.1 Kinetic energy . . . . .	398
11.9.2 Available potential energy . . . . .	399
11.9.3 Exchange of mechanical energy . . . . .	400
11.9.4 Scaling APE and KE . . . . .	400
<b>11.10 Exercises . . . . .</b>	<b>401</b>

---

## 11.1 Loose threads

- Solution needed for Exercise 11.4.
- More exercises

## 11.2 Non-dimensional equations of quasi-geostrophy

In deriving the quasi-geostrophic potential vorticity equation, we proceed much like for the single layer of shallow water fluid in Section 9.5. In particular, quasi-geostrophic scaling from Section 9.5.1 is relevant for both the shallow water and for the continuously stratified fluid. We employ an asymptotic expansion in the Rossby number and stop at the first nontrivial order, which (as for the shallow water) is  $\text{Ro}^1$ . For this purpose, recall the non-dimensional momentum and continuity equations from Section 9.7.8

$$\text{Ro} \left[ \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}_h) \hat{\mathbf{u}} + \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} \right] + \hat{\mathbf{f}} \times \hat{\mathbf{u}} = -\hat{\nabla}_h \hat{\varphi} \quad (11.1a)$$

$$\frac{\partial \hat{\varphi}}{\partial \hat{z}} = \hat{b} \quad (11.1b)$$

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0 \quad (11.1c)$$

$$\frac{\text{Ro}}{\text{Bu}} \frac{D\hat{b}}{Dt} + \hat{w} = 0. \quad (11.1d)$$

We expand all fields in an asymptotic series in Rossby number

$$\hat{u} = \hat{u}_0 + \text{Ro} \hat{u}_1 + \text{Ro}^2 \hat{u}_2 + \dots \quad (11.2a)$$

$$\hat{v} = \hat{v}_0 + \text{Ro} \hat{v}_1 + \text{Ro}^2 \hat{v}_2 + \dots \quad (11.2b)$$

$$\hat{w} = \hat{w}_0 + \text{Ro} \hat{w}_1 + \text{Ro}^2 \hat{w}_2 + \dots \quad (11.2c)$$

$$\hat{b} = \hat{b}_0 + \text{Ro} \hat{b}_1 + \text{Ro}^2 \hat{b}_2 + \dots \quad (11.2d)$$

$$\hat{\varphi} = \hat{\varphi}_0 + \text{Ro} \hat{\varphi}_1 + \text{Ro}^2 \hat{\varphi}_2 + \dots \quad (11.2e)$$

along with the expansion (9.55) for the Coriolis parameter

$$\hat{f} = \hat{f}_0 + \text{Ro} \hat{\beta} \hat{y}, \quad (11.3)$$

and where (equation (9.56))

$$\hat{\beta} \hat{y} = \frac{\beta y}{\text{Ro} f_0} = T \beta y. \quad (11.4)$$

As noted in Section 9.5.2, the velocity field is non-divergent at each order of Rossby number, so that

$$\nabla \cdot \hat{\mathbf{v}}_n = 0 \quad \forall n. \quad (11.5)$$

The Burger number is order unity since the horizontal length scales of the flow considered here are on the order of the deformation radius

$$\text{Bu} \sim 1 \implies L \sim L_d, \quad (11.6)$$

where we introduced the *internal deformation radius* from Section 9.7.6

$$L_d(z) = H [N(z)/f_0]. \quad (11.7)$$

It is important to retain the depth dependence of the Burger number through its dependence on the background stratification  $N^2(z)$

$$\text{Bu}(z) = \left[ \frac{L_d}{L} \right]^2 = N^2(z) \left[ \frac{H}{L f_0} \right]^2, \quad (11.8)$$

which motivates the name *Burger function* for continuously stratified quasi-geostrophy. Importantly, the Burger function does not commute with the vertical derivative operator.

### 11.2.1 Zeroth order asymptotic equations

The zeroth order asymptotic equations are

$$\hat{\mathbf{f}}_0 \times \hat{\mathbf{u}}_0 = -\hat{\nabla}_h \hat{\varphi}_0 \quad (11.9a)$$

$$\partial_{\hat{z}} \hat{\varphi}_0 = \hat{b}_0 \quad (11.9b)$$

$$\hat{\nabla}_h \cdot \hat{\mathbf{u}}_0 + \partial_{\hat{z}} \hat{w}_0 = 0 \quad (11.9c)$$

$$\hat{w}_0 = 0. \quad (11.9d)$$

The first equation represents  $f$ -plane geostrophy, which means that the horizontal velocity has zero divergence

$$\widehat{\nabla}_h \cdot \widehat{\mathbf{u}}_0 = 0. \quad (11.10)$$

Equation (11.9b) means the zeroth order buoyancy determines the zeroth order hydrostatic pressure. Since the horizontal velocity has zero divergence, the continuity equation (11.9c) means that the vertical velocity is depth independent

$$\partial_{\hat{z}} \hat{w}_0 = 0. \quad (11.11)$$

If it vanishes somewhere, such as a solid horizontal bottom boundary, then it vanishes everywhere. This is a manifestation of the [Taylor-Proudman effect](#) studied in VOLUME 2. Indeed, a vanishing  $\hat{w}_0$  is required by the zeroth-order buoyancy equation (11.9d) even if the bottom is not flat. Hence, the non-dimensional velocity has a nonzero contribution only at order  $\text{Ro}^1$

$$\hat{w} = \text{Ro} \hat{w}_1 + \text{Ro}^2 \hat{w}_2 + \dots, \quad (11.12)$$

thus manifesting the vertical stiffening of fluid columns found in rotating fluids. Correspondingly, the dimensionful vertical velocity has the asymptotic expansion

$$w = W \hat{w} = W \text{Ro} (\hat{w}_1 + \text{Ro} \hat{w}_2 + \dots), \quad (11.13)$$

so that to leading to order  $\text{Ro}^1$

$$\hat{w}_1 = \frac{w}{W \text{Ro}}. \quad (11.14)$$

Since the zeroth-order horizontal velocity is non-divergent, we can introduce a geostrophic streamfunction

$$\hat{u}_0 = -\partial_{\hat{y}} \hat{\psi}_0 \quad \text{and} \quad \hat{v}_0 = \partial_{\hat{x}} \hat{\psi}_0 \quad \text{and} \quad \hat{\zeta}_0 = \widehat{\nabla}_h^2 \hat{\psi}_0, \quad (11.15)$$

where the zeroth-order streamfunction is the ratio of the zeroth order pressure to zeroth order Coriolis parameter

$$\hat{\psi}_0 = \hat{\varphi}_0 / \hat{f}_0. \quad (11.16)$$

Note also that the zeroth-order system satisfies the thermal wind balance

$$\hat{\mathbf{f}}_0 \times \partial_{\hat{z}} \hat{\mathbf{u}}_0 = -\widehat{\nabla}_h \hat{b}_0. \quad (11.17)$$

Finally, note that the zeroth order buoyancy is determined by the streamfunction through the hydrostatic balance

$$\hat{b}_0 = \partial_{\hat{z}} \hat{\varphi}_0 = \hat{f}_0 \partial_{\hat{z}} \hat{\psi}_0. \quad (11.18)$$

### 11.2.2 First order asymptotic equations

For a prognostic equation we must consider equations at order  $\text{Ro}^1$

$$\frac{D_0 \hat{\mathbf{u}}_0}{Dt} + \hat{\mathbf{f}}_0 \times \hat{\mathbf{u}}_1 + \hat{\beta} \hat{y} \hat{z} \times \hat{\mathbf{u}}_0 = -\widehat{\nabla}_h \hat{\varphi}_1 \quad (11.19a)$$

$$\partial_{\hat{z}} \hat{\varphi}_1 = \hat{b}_1 \quad (11.19b)$$

$$\widehat{\nabla}_h \cdot \hat{\mathbf{u}}_1 + \partial_{\hat{z}} \hat{w}_1 = 0 \quad (11.19c)$$

$$\frac{1}{\text{Bu}} \frac{\text{D}_0 \hat{b}_0}{\text{Dt}} + \hat{w}_1 = 0. \quad (11.19d)$$

The first order terms are often referred to as the *ageostrophic* components, though note that all contributions higher than zeroth order constitute ageostrophic contributions.

At order  $\text{Ro}^1$ , the material time derivative makes use *only* of the zeroth order horizontal geostrophic velocity

$$\frac{\text{D}_0}{\text{Dt}} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h. \quad (11.20)$$

To close this set of equations, we produce the vorticity equation from the momentum equation, and then combine the vorticity equation and buoyancy equation to produce the quasi-geostrophic potential vorticity equation. In Section 9.5, we performed the same procedure for deriving the shallow water quasi-geostrophic equations.

Taking the curl of the momentum equation (11.19a) eliminates the pressure gradient,  $\hat{\nabla}_h \hat{\varphi}_1$ , thus producing the vorticity equation

$$\partial_t \hat{\zeta}_0 + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \hat{\nabla}_h \cdot \hat{\mathbf{u}}_1. \quad (11.21)$$

We make use of the continuity equation (11.19c) to eliminate the horizontal convergence

$$\partial_t \hat{\zeta}_0 + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = \hat{f}_0 \partial_z \hat{w}_1. \quad (11.22)$$

The right hand side represents the contribution to vorticity evolution from stretching by the ageostrophic vertical velocity acting in a rotating reference frame. We can eliminate the ageostrophic vertical velocity through use of the buoyancy equation (11.19d). When doing so, it is important to keep the depth dependence of the Burger function,  $\text{Bu}(z)$ , according to equation (11.8), with this depth dependence arising from the prescribed background stratification,  $N^2(z)$ . The resulting vorticity equation is

$$\partial_t \hat{\zeta}_0 + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \frac{\partial}{\partial z} \left[ \frac{1}{\text{Bu}} \frac{\text{D}_0 \hat{b}_0}{\text{Dt}} \right]. \quad (11.23)$$

We now use the identity

$$\frac{\partial}{\partial z} \left[ \frac{1}{\text{Bu}} \frac{\text{D}_0 \hat{b}_0}{\text{Dt}} \right] = \frac{\partial}{\partial z} \left[ \frac{1}{\text{Bu}} \left( \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h \right) \hat{b}_0 \right] \quad (11.24a)$$

$$= \frac{\text{D}_0}{\text{Dt}} \left[ \frac{\partial}{\partial z} \left( \frac{\hat{b}_0}{\text{Bu}} \right) \right] + \frac{1}{\text{Bu}} \frac{\partial \hat{\mathbf{u}}_0}{\partial z} \cdot \hat{\nabla}_h \hat{b}_0 \quad (11.24b)$$

$$= \frac{\text{D}_0}{\text{Dt}} \left[ \frac{\partial}{\partial z} \left( \frac{\hat{b}_0}{\text{Bu}} \right) \right], \quad (11.24c)$$

where we set

$$\partial_z \hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h \hat{b}_0 = 0 \quad (11.25)$$

since the zeroth-order velocity maintains thermal wind balance (11.17). Bringing terms together

then leads to the material conservation equation for quasi-geostrophic potential vorticity

$$\frac{D_0}{Dt} \left[ \hat{\zeta}_0 + \hat{\beta} \hat{y} + \hat{f}_0 \frac{\partial}{\partial \hat{z}} \left( \frac{\hat{b}_0}{Bu} \right) \right] = 0. \quad (11.26)$$

## 11.3 Dimensionful equations of quasi-geostrophy

To expose physical elements to the continuously stratified quasi-geostrophic theory, we reintroduce physical dimensions as done for shallow water quasi-geostrophy in Section 9.6. For that purpose, we write

$$\mathbf{u} \equiv \mathbf{u}_g + \mathbf{u}_{ag} = U (\hat{\mathbf{u}}_0 + \text{Ro} \hat{\mathbf{u}}_1) \quad (11.27a)$$

$$w \equiv w_{ag} = \text{Ro} W \hat{w}_1 \quad (11.27b)$$

$$b \equiv b_g + B_{ag} = B (\hat{b}_0 + \text{Ro} \hat{b}_1) \quad (11.27c)$$

$$\varphi \equiv \varphi_g + \varphi_{ag} = f_o U L (\hat{\varphi}_0 + \text{Ro} \hat{\varphi}_1). \quad (11.27d)$$

Reintroducing dimensions is straightforward but a bit tedious.

### 11.3.1 Hydrostatic balance

Hydrostatic balance is maintained for terms at each Rossby number order, so that we have the dimensional equations

$$\partial_z \varphi_g = b_g \quad \text{and} \quad \partial_z \varphi_{ag} = B_{ag}. \quad (11.28)$$

### 11.3.2 Continuity equation

The three dimensional velocity is non-divergent at each order in Rossby number. The zeroth order (geostrophic) flow is horizontally non-divergent so that

$$\nabla \cdot \mathbf{u}_g = 0, \quad (11.29)$$

whereas the next order (ageostrophic) flow satisfies

$$\nabla \cdot \mathbf{v}_{ag} = \nabla \cdot \mathbf{u}_{ag} + \partial_z w_{ag} = 0. \quad (11.30)$$

### 11.3.3 Geostrophic balance

The non-dimensional geostrophic balance for the zeroth order fields

$$\hat{f}_0 \times \hat{\mathbf{u}}_0 = -\hat{\nabla}_h \hat{\varphi}_0 \quad (11.31)$$

takes on the dimensional form

$$\hat{\mathbf{z}} \times \mathbf{u}_g / U = -L \nabla_h \varphi_g / (f_o U L). \quad (11.32)$$

Cancelling factors leads to the expected form of  $f$ -plane geostrophy

$$f_o \hat{\mathbf{z}} \times \mathbf{u}_g = -\nabla_h \varphi_g. \quad (11.33)$$

### 11.3.4 Material time derivative

For the material time derivative operator we write

$$D/Dt = \partial_t + \mathbf{u} \cdot \nabla_h + w \partial_z \quad (11.34a)$$

$$= (1/T) \partial_{\tilde{t}} + (U/L) \hat{\mathbf{u}} \cdot \hat{\nabla}_h + (W/H) \hat{w} \partial_{\tilde{z}} \quad (11.34b)$$

$$= (1/T) (\partial_{\tilde{t}} + \hat{\mathbf{u}} \cdot \hat{\nabla}_h + \hat{w} \partial_{\tilde{z}}) \quad (11.34c)$$

$$= (1/T) (\partial_{\tilde{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h) + (\text{Ro}/T) (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h + \hat{w}_1 \partial_{\tilde{z}}) \quad (11.34d)$$

$$= \partial_t + \mathbf{u}_g \cdot \nabla_h + \mathbf{u}_{ag} \cdot \nabla_h + w_{ag} \partial_z \quad (11.34e)$$

$$\equiv D_g/Dt + \mathbf{u}_{ag} \cdot \nabla_h + w_{ag} \partial_z, \quad (11.34f)$$

where time scales according to advection,  $T = L/U$ , vertical velocity scales according to continuity,  $W = HU/L$ , and we introduced the geostrophic material time derivative operator

$$D_g/Dt \equiv \partial_t + \mathbf{u}_g \cdot \nabla_h. \quad (11.35)$$

### 11.3.5 Buoyancy equation

The buoyancy equation requires a bit more work. For that purpose, split buoyancy into a depth dependent static background and a deviation from the background

$$b = \tilde{b}(z) + b'(x, y, z, t), \quad (11.36)$$

with its vertical derivative

$$\partial_z b = d\tilde{b}/dz + \partial_z b' = N^2 + \partial_z b', \quad (11.37)$$

where  $N^2(z)$  is the squared buoyancy frequency for the prescribed background buoyancy field. In this manner, the adiabatic buoyancy equation is

$$\frac{\partial b'}{\partial t} + \mathbf{u} \cdot \nabla_h b' + w (N^2 + \partial_z b') = 0. \quad (11.38)$$

We non-dimensionalize this equation by making use of the following relations between the scales

$$B = f_0 U L / H \quad W = H (U/L) \quad \text{Ro} = U / (f_0 L) \quad \text{Bu} = (N H)^2 / (f_0 L)^2, \quad (11.39)$$

in which case the buoyancy equation takes the form

$$\partial_t b' + \mathbf{u} \cdot \nabla_h b' + w (N^2 + \partial_z b') = \frac{B}{T} \frac{\partial \hat{b}}{\partial \tilde{t}} + \frac{U B}{L} \hat{\mathbf{u}} \cdot \hat{\nabla}_h \hat{b} + W \hat{w} N^2 + \frac{W B}{H} \hat{w} \partial_{\tilde{z}} \hat{b} = 0. \quad (11.40)$$

We find it useful to divide by  $f_0 N^2$ , so that

$$\frac{1}{f_0 N^2} [\partial_t b' + \mathbf{u} \cdot \nabla_h b' + w (N^2 + \partial_z b')] = \frac{H \text{Ro}^2}{\text{Bu}} [\partial_{\tilde{t}} \hat{b} + \hat{\mathbf{u}} \cdot \hat{\nabla}_h \hat{b} + \hat{w} \partial_{\tilde{z}} \hat{b}] + \text{Ro} \hat{w} = 0. \quad (11.41)$$

The vertical velocity component,  $\hat{w}$ , is nonzero only at order  $\text{Ro}^1$ , so the term  $\text{Ro} \hat{w}$  is order  $\text{Ro}^2$  (recall equation (11.27b)). For the material time derivative term, we drop the vertical velocity term,  $\hat{w} \partial_{\tilde{z}} \hat{b}$ , since it is order  $\text{Ro}^1$  smaller than the other terms in the material time operator. We thus retain only the zeroth order buoyancy contribution,  $\hat{b}_0$ . Reintroducing physical dimensions

then leads to the dimensional form of the quasi-geostrophic buoyancy equation

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) b_g + w_{ag} N^2 = \frac{D_g b_g}{Dt} + w_{ag} N^2 = 0. \quad (11.42)$$

This equation means that the geostrophic transport of the geostrophic buoyancy is affected by a source due to the ageostrophic vertical advection of background buoyancy

$$\frac{D_g b_g}{Dt} = -w_{ag} N^2 \quad \text{with} \quad b_g = f_0 \partial_z \psi. \quad (11.43)$$

### 11.3.6 Vorticity equation

Reintroducing dimensions to the vorticity equation (11.22) yields<sup>2</sup>

$$\partial_t \hat{\zeta}_0 + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) - \hat{f}_0 \partial_{\hat{z}} \hat{w}_1 = T^2 [\partial_t \zeta + \mathbf{u}_g \cdot \nabla_h (\zeta + \beta y)] - \frac{H}{W \text{Ro}} \partial_z w_{ag}. \quad (11.44)$$

The identity  $H/(W \text{Ro}) = f_0 T^2$  leads to the order  $\text{Ro}^1$  vorticity equation

$$\partial_t \zeta_a + J(\psi, \zeta_a) = f_0 \partial_z w_{ag}, \quad (11.45)$$

with the absolute vorticity given by the sum of the geostrophic relative vorticity plus the planetary beta contribution

$$\zeta_a = \zeta_g + \beta y. \quad (11.46)$$

Hence, the absolute geostrophic vorticity is advected by the geostrophic flow, and it has a source (right hand side of equation (11.45)) due to vertical stretching by the ageostrophic flow. Expanding the Jacobian and introducing the geostrophic velocity,  $\mathbf{u}_g = \hat{\mathbf{z}} \times \nabla \psi$ , leads to the geostrophic relative vorticity equation

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) \zeta_g = -\beta v_g + f_0 \partial_z w_{ag}. \quad (11.47)$$

We thus see that the beta effect from the geostrophic flow, plus vertical stretching by the ageostrophic flow, provide local sources for geostrophic relative vorticity.

### 11.3.7 Potential vorticity

From equation (11.26), we identify the non-dimensional quasi-geostrophic potential vorticity

$$\hat{q} = \hat{\zeta}_0 + \hat{\beta} \hat{y} + \hat{f}_0 \frac{\partial (\hat{b}_0 / \text{Bu})}{\partial \hat{z}}. \quad (11.48)$$

Introducing dimensional quantities to the right hand side yields<sup>3</sup>

$$\hat{q} = \frac{L}{U} [\zeta_g + \beta y] + \frac{\partial}{\partial z} \left[ \frac{H b_g}{B \text{Bu}} \right] \quad (11.49a)$$

$$= (1/(f_0 \text{Ro}) (\zeta_g + \beta y) + \frac{H}{B} \frac{\partial}{\partial z} \left[ \frac{b_g}{\text{Bu}} \right]). \quad (11.49b)$$

<sup>2</sup>Recall from equation (11.27b) that to order  $\text{Ro}^1$  we have  $W \hat{w} = \text{Ro} W \hat{w}_1 = w_{ag}$ .

<sup>3</sup>Recall  $\hat{f}_0 = 1$  and  $\hat{\beta} \hat{y} = T \beta y = (L/U) \beta y$ .

The scale for the fluctuating buoyancy is given by equation (9.160),  $B = f_o U L / H$ , and the inverse Burger function is given by equation (11.8),  $Bu^{-1} = [(L f_o) / (H N)]^2$ , so that

$$\hat{q} = (1/(f_o \text{Ro}) (\zeta_g + \beta y) + \frac{H^2}{f_o U L} \frac{L^2 f_o^2}{H^2} \frac{\partial}{\partial z} \left[ \frac{b_g}{N^2} \right]) \quad (11.50a)$$

$$= (1/(f_o \text{Ro}) (\zeta_g + \beta y) + \text{Ro}^{-1} \partial_z (b_g / N^2)). \quad (11.50b)$$

Introducing the geostrophic streamfunction,

$$\mathbf{u}_g = \hat{\mathbf{z}} \times \nabla_h \psi \quad \text{and} \quad \zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}_g) = \nabla_h^2 \psi \quad \text{and} \quad b_g = f_o \partial_z \psi, \quad (11.51)$$

leads to

$$q \equiv f_o \text{Ro} \hat{q} = \beta y + \zeta_g + \frac{\partial}{\partial z} \left[ \frac{f_o^2}{N^2} \frac{\partial \psi}{\partial z} \right]. \quad (11.52)$$

Just as for the shallow water case in Section 9.6.1, the potential vorticity (11.52) scales as  $f_o \text{Ro}$ . The order Ro scaling is expected since it is only at this order that we realize a prognostic set of equations. We sometimes choose to add the constant  $f_o$  to  $q$ , which has no effect on the dynamics but reveals the beta plane planetary vorticity

$$q = \underbrace{f_o + \beta y}_{\text{planetary vorticity}} + \underbrace{\nabla_h^2 \psi}_{\text{relative vorticity}} + \underbrace{\frac{\partial}{\partial z} \left[ \frac{f_o^2}{N^2} \frac{\partial \psi}{\partial z} \right]}_{\text{stretching by } f}. \quad (11.53)$$

Evidently, there are three contributions to the quasi-geostrophic potential vorticity.

- **PLANETARY VORTICITY:** The planetary vorticity contribution,  $f_o + \beta y$ , arises from rotation of the reference frame. As noted above, the  $\beta y$  term is the only dynamically relevant contribution, so that we can equally well drop the  $f_o$  contribution.
- **GEOSTROPHIC RELATIVE VORTICITY:** The vertical component of the geostrophic relative vorticity,  $\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \nabla_h^2 \psi$ , acts to bring out the smaller scale features in the streamfunction.
- **VERTICAL STRETCHING:** The final contribution arises from vertical stretching in the presence of a rotating planet. Equation (11.22) helps to remind us why this term arises from vortex stretching.

Material evolution of quasi-geostrophic potential vorticity follows the horizontal geostrophic flow

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) q = 0. \quad (11.54)$$

Geostrophic material constancy of  $q$  represents a balance, following the geostrophic flow, of time changes for the planetary vorticity, relative vorticity, and vertical stretching. This is a remarkable property of quasi-geostrophic flows. It packs in a wealth of physical processes that act to evolve the geostrophic flow while maintaining an exact geostrophic balance at each point in space and time.

### 11.3.8 Velocity equation

The prognostic velocity equation arises from the first order asymptotic equation (11.19a)

$$\frac{D_0 \hat{\mathbf{u}}_0}{Dt} + \hat{\mathbf{f}}_0 \times \hat{\mathbf{u}}_1 + \hat{\beta} \hat{\mathbf{y}} \hat{\mathbf{z}} \times \hat{\mathbf{u}}_0 = -\hat{\nabla}_h \hat{\varphi}_1. \quad (11.55)$$

Our skills with reintroducing dimensional quantities should be sufficient to write down the dimensional velocity equation by inspection

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) \mathbf{u}_g + \beta y \hat{\mathbf{z}} \times \mathbf{u}_g + f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} = -\nabla_h \varphi_{ag}. \quad (11.56)$$

We can also choose to add the geostrophic balanced flow that holds at order  $\text{Ro}^0$ ,  $f_o \hat{\mathbf{z}} \times \mathbf{u}_g = -\nabla_h \varphi_g$ , so to have the equivalent equation

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) \mathbf{u}_g + \beta y \hat{\mathbf{z}} \times \mathbf{u}_g + f_o \hat{\mathbf{z}} \times (\mathbf{u}_g + \mathbf{u}_{ag}) = -\nabla_h (\varphi_g + \varphi_{ag}). \quad (11.57)$$

Observe that from equation (11.27d) that

$$\varphi_{ag} = f_o U L \text{Ro} \hat{\varphi}_1 = U^2 \hat{\varphi}_1, \quad (11.58)$$

so that the ageostrophic portion of the pressure scales according to pressure in a non-rotating fluid, such as discussed in VOLUME 2 when scaling the hydrostatic balance in a Boussinesq ocean.

### 11.3.9 Concerning the ageostrophic state in quasi-geostrophy

Following the shallow water discussion in Section 9.6.6, we here expose an ambiguity (i.e., *gauge freedom*) associated with the ageostrophic pressure, buoyancy, and velocity within quasi-geostrophy. This gauge freedom arises since the quasi-geostrophic potential vorticity equation remains unchanged upon adding an arbitrary horizontally non-divergent velocity to  $\mathbf{u}_{ag}$ , along with an arbitrary gradient of a scalar to  $\varphi_{ag}$ . The freedom to modify these ageostrophic fields is constrained by taking the divergence of the velocity equation (11.56) to find (compare to equation (9.108) for shallow water quasi-geostrophy)

$$\nabla_h^2 \varphi_{ag} = 2 J(u_g, v_g) + \beta (y \zeta_g - u_g) + f_o \zeta_{ag}, \quad (11.59)$$

where

$$\zeta_{ag} = \partial_x v_{ag} - \partial_y u_{ag} \quad (11.60)$$

is the relative vorticity contained in the horizontal ageostrophic flow. The constraint (11.59) means that adding a horizontally non-divergent velocity,  $\tilde{\mathbf{u}}$ , to the ageostrophic velocity,  $\mathbf{u}_{ag}$ , requires a corresponding modification to the pressure via  $\tilde{\varphi}$  added to  $\varphi_{ag}$ , in which  $\tilde{\varphi}$  satisfies the Poisson equation

$$\nabla_h^2 \tilde{\varphi} = f_o \hat{\mathbf{z}} \cdot (\nabla \times \tilde{\mathbf{u}}). \quad (11.61)$$

Given the gauge freedom, we follow Section 6.3 of [Holton and Hakim \(2013\)](#) by choosing  $\varphi_{ag} = 0$  (and thus  $b_{ag} = 0$ ) so that all ageostrophic effects live within the velocity,  $\mathbf{u}_{ag} + \hat{\mathbf{z}} w_{ag}$ . In this case, the quasi-geostrophic velocity equation (11.56) simplifies to

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) \mathbf{u}_g + \beta y \hat{\mathbf{z}} \times \mathbf{u}_g + f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} = 0. \quad (11.62)$$

Evidently, the Coriolis acceleration,  $f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag}$ , provides the only place that ageostrophic effects couple to the geostrophic velocity equation.

## 11.4 Constraints on quasi-geostrophic evolution

How is it that quasi-geostrophic flow retains a geostrophically balanced state,  $f_o \hat{\mathbf{z}} \times \mathbf{u}_g = -\nabla_h \varphi_g$ , while allowing that state to evolve? The discussion thus far provides an operational means to answer that question via time integration of the quasi-geostrophic potential vorticity equation. Alternatively, let us consider the velocity equation and in so doing to expose a diagnostic balance of geostrophic and ageostrophic flow processes, with this balance ensuring that geostrophy is maintained. We derive this balance closely following that considered for shallow water quasi-geostrophy in Section 9.6.7.<sup>4</sup>

### 11.4.1 A balance between geostrophic and ageostrophic processes

Geostrophic balance is maintained for an observer following a fluid particle moving with the horizontal geostrophic velocity, so that

$$\frac{D_g}{Dt} (f_o \hat{\mathbf{z}} \times \mathbf{u}_g + \nabla_h \varphi_g) = 0. \quad (11.63)$$

Performing the material time derivative, and making use of equation (11.62) for the geostrophic velocity and equation (11.43) for buoyancy, here written as

$$\frac{D_g \mathbf{u}_g}{Dt} = -(\beta y) \hat{\mathbf{z}} \times \mathbf{u}_g - f_o \hat{\mathbf{z}} \times \mathbf{u}_{ag} \quad (11.64a)$$

$$\frac{D_g b_g}{Dt} = -w_{ag} N^2, \quad (11.64b)$$

leads to the balance

$$f_o \beta y \mathbf{u}_g + f_o^2 \mathbf{u}_{ag} + \frac{D_g (\nabla_h \varphi_g)}{Dt} = 0. \quad (11.65)$$

We can derive a diagnostic balance (i.e., an equation without a time derivative) by taking the vertical derivative of equation (11.65) and using the hydrostatic balance and the buoyancy equation. For that purpose we need

$$\partial_z [D_g (\nabla_h \varphi_g) / Dt] = D_g (\nabla_h b_g) / Dt + (\partial_z \mathbf{u}_g \cdot \nabla_h) \nabla_h \varphi_g \quad (11.66a)$$

$$= \nabla_h (D_g b_g / Dt) + \mathbf{Q} + (\partial_z \mathbf{u}_g \cdot \nabla_h) \nabla_h \varphi_g \quad (11.66b)$$

$$= -\nabla_h (N^2 w_{ag}) + \mathbf{Q} + (\partial_z \mathbf{u}_g \cdot \nabla_h) \nabla_h \varphi_g, \quad (11.66c)$$

where we introduced the vector arising from the coupling of horizontal shears in the geostrophic flow with horizontal gradients in the geostrophic buoyancy

$$\mathbf{Q} = -\hat{\mathbf{x}} \partial_x \mathbf{u}_g \cdot \nabla_h b_g - \hat{\mathbf{y}} \partial_y \mathbf{u}_g \cdot \nabla_h b_g. \quad (11.67)$$

---

<sup>4</sup>Enforcing geostrophy even while the flow evolves is analogous to constraints arising from non-divergent flow condition,  $\nabla \cdot \mathbf{v} = 0$ , in an evolving flow. As seen in our study of the Boussinesq ocean in VOLUME 2, pressure in a Boussinesq ocean enforces flow non-divergence at each point in space and time.

We can further simplify by use of thermal wind to write

$$\mathbf{Q} = -\hat{x} \partial_x \mathbf{u}_g \cdot (f_o \hat{z} \times \partial_z \mathbf{u}_g) - \hat{y} \partial_y \mathbf{u}_g \cdot (f_o \hat{z} \times \partial_z \mathbf{u}_g) \quad (11.68a)$$

$$= f_o \hat{x} (-\partial_x u \partial_z v + \partial_x v \partial_z u) + f_o \hat{x} (-\partial_y u \partial_z v + \partial_y v \partial_z u), \quad (11.68b)$$

and geostrophy to write

$$(\partial_z \mathbf{u}_g \cdot \nabla_h) \nabla_h \varphi_g = f_o (\partial_z \mathbf{u}_g \cdot \nabla_h) (-v_g \hat{x} + u_g \hat{y}) \quad (11.69a)$$

$$= f_o \hat{x} (-\partial_x u \partial_z v + \partial_x v \partial_z u) + f_o \hat{x} (-\partial_y u \partial_z v + \partial_y v \partial_z u) \quad (11.69b)$$

$$= \mathbf{Q}, \quad (11.69c)$$

so that

$$\partial_z [D_g(\nabla_h \varphi)/Dt] = 2 \mathbf{Q}. \quad (11.70)$$

Bringing terms together leads to the diagnostic balance

$$\underbrace{f_o \beta y \partial_z \mathbf{u}_g}_{\text{geostrophic}} + \underbrace{2 \mathbf{Q} + f_o^2 \partial_z \mathbf{u}_{ag} - N^2 \nabla_h w_{ag}}_{\text{ageostrophic}} = 0, \quad (11.71)$$

where we noted that the buoyancy frequency is a function only of the vertical. Equation (11.71) summarizes a wealth of geostrophic and ageostrophic processes that, when taken together, maintain geostrophy and thermal wind for evolving quasi-geostrophic flows. Note that when considered in isolation, each process acts to move the flow away from geostrophic balance (e.g., see [Hoskins \(1975\)](#) and Section 6.5 of [Holton and Hakim \(2013\)](#), who emphasize the nonlinear geostrophic term,  $\mathbf{Q}$ ). It is only when all terms are considered together that they render an evolving flow respecting geostrophy.

### 11.4.2 Vertical motion and the $\omega$ -equation

As we discovered in this chapter, for quasi-geostrophy the vertical component to the velocity is non-zero only at first order in Rossby number, hence it is part of the ageostrophic flow. In contrast, the zeroth order flow is horizontally non-divergent and geostrophically balanced. To evolve the horizontal geostrophic flow it is not necessary to explicitly compute the ageostrophic vertical velocity. However, there are cases where vertical ageostrophic velocity is of interest. Since the vertical motion is relatively small, it is important to formulate the calculation of this motion in a manner that avoids computing small differences between large numbers. Namely, even if we knew the horizontal ageostrophic flow, vertically integrating the continuity equation,  $\partial_z w_{ag} = -(\partial_x u_{ag} + \partial_y v_{ag})$ , is prone to errors since the horizontal convergence is typically the small difference of relatively large numbers. We here derive a more suitable means to diagnose  $w_{ag}$ , with the diagnostic equation known as the  $\omega$ -equation. The name for this equation originates from the atmospheric community where  $\omega$  is the common symbol for mass transport across pressure surfaces (using pressure vertical coordinates rather than geopotential). Here, we make use of the Boussinesq ocean equations so that the vertical velocity component is across geopotential surfaces.

It turns out that we have done most of the work needed to derive an equation for  $w_{ag}$  as part of our derivation of the balance (11.71). Namely, by simply taking horizontal the divergence of (11.71) and noting that  $\nabla \cdot \mathbf{u}_g = 0$ , we have

$$(f_o^2 \partial_{zz} + N^2 \nabla_h^2) w_{ag} = (\nabla_h + \hat{z} \partial_z) \cdot (2 \mathbf{Q} + \hat{z} f_o \beta v_g), \quad (11.72)$$

which corresponds to the second form of the  $\omega$ -equation first derived by [Hoskins et al. \(1978\)](#). We identify the following convenient features of this  $\omega$ -equation. First, the right hand side is determined solely from knowledge of the geostrophic flow, which contrasts to the approach from continuity. Second, the linear operator acting on  $w_{\text{ag}}$  is elliptic, so that  $w_{\text{ag}}$  is a smoothed version of the geostrophic forcing on the right hand side of equation (11.72). Third, maps of the vector field,  $\mathbf{Q} + \hat{\mathbf{z}} f_0 \beta v_g$ , provide insights into regions where  $w_{\text{ag}}$  is prone to have large values. For example, where the divergence on the right hand side is positive, the vertical velocity is negative.<sup>5</sup>

### 11.4.3 Another derivation of the $\omega$ -equation

One commonly finds an alternative derivation in the literature. Here, we write the dimensional buoyancy equation (11.42) and vorticity equation (11.45) in the form

$$f_0 \partial_{tz} \psi + J(\psi, b_g) + N^2 w_{\text{ag}} = 0 \quad (11.73a)$$

$$\partial_t (\nabla_h^2 \psi) + J(\psi, \nabla_h^2 \psi + \beta y) - f_0 \partial_z w_{\text{ag}} = 0. \quad (11.73b)$$

Taking the horizontal Laplacian of equation (11.73a) and subtracting it from  $f_0$  times the vertical derivative of equation (11.73b) allows us to cancel the time derivative and thus to render the Boussinesq form of the quasi-geostrophic  $\omega$ -equation

$$(N^2 \nabla_h^2 + f_0^2 \partial_{zz}) w_{\text{ag}} = f_0 \partial_z [J(\psi, \zeta_g + \beta y)] - \nabla_h^2 J(\psi, b_g). \quad (11.74)$$

The operator on the left hand side is a generalized Laplacian and all terms on the right hand side are known from the geostrophic streamfunction, including the vorticity,  $\zeta_g = \nabla_h^2 \psi$ , and buoyancy,  $b_g = f_0 \partial_z \psi$ . This equation is thus in the form of a generalized Poisson equation, whose solution renders an expression for the vertical velocity valid to order  $\text{Ro}^1$ . Note that the right hand side can be written as the convergence of a flux

$$f_0 \partial_z [J(\psi, \zeta_g + \beta y)] - \nabla_h^2 J(\psi, b_g) = -\nabla_h \cdot [\nabla_h J(\psi, b_g)] + f_0 \partial_z [J(\psi, \zeta_g + \beta y)] \quad (11.75a)$$

$$= -\nabla \cdot \mathbf{G}, \quad (11.75b)$$

where the geostrophic forcing vector is given by

$$\mathbf{G} = \nabla_h J(\psi, b_g) - f_0 \partial_z [J(\psi, \zeta_g + \beta y)] \hat{\mathbf{z}}. \quad (11.76)$$

It is a useful exercise to show that equations (11.74) and (11.72) are identical.

## 11.5 Connecting to Ertel potential vorticity

Following our treatment for the shallow water system in Section 9.6.4, we here determine how quasi-geostrophic potential vorticity relates to the Ertel potential vorticity from Chapter 7. For this purpose, consider the continuously stratified hydrostatic Boussinesq fluid and make use of the Ertel potential vorticity derived in Exercise 7.1

$$Q = (\boldsymbol{\omega} + \hat{\mathbf{z}} f) \cdot \nabla b = \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + (\zeta + f) \frac{\partial b}{\partial z}. \quad (11.77)$$

---

<sup>5</sup>Elliptic operators generally swap signs, which can be readily seen upon assuming a single Fourier mode solution to  $w_{\text{ag}}$ .

In a perfect fluid we have the material conservation

$$(\partial_t + \mathbf{u} \cdot \nabla_h + w \partial_z) Q = 0. \quad (11.78)$$

Our strategy in this section is to non-dimensionalize both  $Q$  and the material time operator, and then to organize terms in equation (11.78) according to the Rossby number. We then show that material conservation of Ertel potential vorticity, when expanded asymptotically to order  $\text{Ro}^1$ , leads to the geostrophic material conservation of quasi-geostrophic potential vorticity. The continuous stratification makes the derivation more involved than for the shallow water potential vorticity in Section 9.6.4, prompting us to expose the details.

### 11.5.1 Non-dimensionalizing the Ertel potential vorticity

As above for the buoyancy, we are led to write the Ertel potential vorticity in the form

$$\frac{Q - f_o N^2}{f_o N^2} = \frac{1}{f_o N^2} \left[ \frac{\partial u}{\partial z} \frac{\partial b'}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b'}{\partial x} \right] + \frac{1}{N^2} \frac{\partial b'}{\partial z} + \frac{\beta y + \zeta}{f_o} \left[ 1 + \frac{1}{N^2} \frac{\partial b'}{\partial z} \right]. \quad (11.79)$$

with non-dimensionalization leading to

$$\frac{1}{f_o N^2} \left[ \frac{\partial u}{\partial z} \frac{\partial b'}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b'}{\partial x} \right] = \frac{B U}{f_o N^2 H L} \left[ \frac{\partial \hat{u}}{\partial \hat{z}} \frac{\partial \hat{b}}{\partial \hat{y}} - \frac{\partial \hat{v}}{\partial \hat{z}} \frac{\partial \hat{b}}{\partial \hat{x}} \right] = \frac{\text{Ro}^2}{\text{Bu}} \left[ \frac{\partial \hat{u}}{\partial \hat{z}} \frac{\partial \hat{b}}{\partial \hat{y}} - \frac{\partial \hat{v}}{\partial \hat{z}} \frac{\partial \hat{b}}{\partial \hat{x}} \right] \quad (11.80a)$$

$$\frac{1}{N^2} \frac{\partial b'}{\partial z} = \frac{B}{H N^2} \frac{\partial \hat{b}}{\partial \hat{z}} = \frac{\text{Ro}}{\text{Bu}} \frac{\partial \hat{b}}{\partial \hat{z}} \quad (11.80b)$$

$$\frac{\beta y + \zeta}{f_o} = \text{Ro} (\hat{\beta} \hat{y} + \hat{\zeta}). \quad (11.80c)$$

The order  $\text{Ro}^2$  terms appearing in equation (11.80a) are dropped since they do not contribute to the quasi-geostrophic potential vorticity, which involve terms only up to order  $\text{Ro}^1$ . For the order  $\text{Ro}^1$  term, we only retain the zeroth order buoyancy,  $\hat{b}_0 = b_g/B$ , and likewise we just retain the zeroth order relative vorticity,  $\hat{\zeta}_0 = (L/U) \zeta_g$ . Hence, the Ertel potential vorticity is given by

$$Q = N^2 (f_o + q_*) + \mathcal{O}(\text{Ro}^2) \quad (11.81)$$

where  $q_*$  is the order  $\text{Ro}^1$  term

$$q_* = \text{Ro} f_o \left[ \frac{1}{\text{Bu}} \frac{\partial \hat{b}_0}{\partial \hat{z}} + \hat{\beta} \hat{y} + \hat{\zeta}_0 \right] = \frac{f_o}{N^2} \frac{\partial b_g}{\partial z} + \beta y + \zeta_g. \quad (11.82)$$

### 11.5.2 Material conservation of Ertel PV to order $\text{Ro}^1$

The material conservation of Ertel PV now takes the form

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h + w_{ag} \partial_z) (f_o N^2 + q_* N^2) = f_o w_{ag} \partial_z N^2 + N^2 D_g q_*/Dt = 0. \quad (11.83a)$$

We dropped the  $w_{ag} \partial_z$  contribution to the advection of  $q_*$  since ageostrophic vertical advection of  $q_*$  is an order  $\text{Ro}^2$  term. To eliminate the vertical ageostrophic velocity we make use of the buoyancy equation (11.42) so that

$$\frac{D_g q_*}{Dt} + \frac{w_{ag}}{N^2} \frac{\partial N^2}{\partial z} = \frac{D_g q_*}{Dt} - \frac{f_o}{N^4} \frac{D_g b_g}{Dt} \frac{\partial N^2}{\partial z} = 0. \quad (11.84)$$

Writing

$$\frac{\partial}{\partial z} \left[ \frac{1}{N^2} \right] = -\frac{1}{N^4} \frac{\partial N^2}{\partial z} \quad (11.85)$$

leads to

$$\frac{D_g q_*}{Dt} + f_o \left[ \frac{\partial N^{-2}}{\partial z} \right] \frac{D_g b_g}{Dt} = 0. \quad (11.86)$$

Since the geostrophic material time derivative operator only involves horizontal advection, we can merge these two terms to render

$$\frac{D_g q_*}{Dt} + f_o \left[ \frac{\partial N^{-2}}{\partial z} \right] \frac{D_g b_g}{Dt} = \frac{D_g}{Dt} \left[ q_* + f_o b_g \left( \frac{\partial N^{-2}}{\partial z} \right) \right] \quad (11.87a)$$

$$= \frac{D_g}{Dt} \left[ \beta y + \zeta + \frac{f_o}{N^2} \left( \frac{\partial b_g}{\partial z} \right) + f_o b_g \left( \frac{\partial N^{-2}}{\partial z} \right) \right] \quad (11.87b)$$

$$= \frac{D_g}{Dt} \left[ \beta y + \zeta + f_o \frac{\partial}{\partial z} \left( \frac{b_g}{N^2} \right) \right] \quad (11.87c)$$

$$= \frac{D_g q}{Dt} \quad (11.87d)$$

$$= 0. \quad (11.87e)$$

In the penultimate step we introduced the quasi-geostrophic potential vorticity given by equation (11.53)

$$q = \beta y + \zeta + f_o \frac{\partial}{\partial z} \left[ \frac{b_g}{N^2} \right] = q_* + f_o b_g \frac{\partial N^{-2}}{\partial z}. \quad (11.88)$$

We have thus established how material conservation of Ertel potential vorticity, when expanded asymptotically to order  $\text{Ro}^1$ , leads to the geostrophic material conservation of quasi-geostrophic potential vorticity.

## 11.6 Boundary conditions

We need boundary conditions on the geostrophic streamfunction,  $\psi$ , to invert the elliptic quasi-geostrophic PV equation (11.53), with boundary conditions the focus of this section.

### 11.6.1 Concerning doubly-periodic QG models

For lateral boundaries, we may choose to set the normal component of the flow to zero for the inviscid case, in which  $\psi$  is a constant along material boundaries<sup>6</sup>. More commonly (at least for idealized studies), we remove the lateral boundaries altogether by assuming a doubly periodic domain. The value of the Coriolis parameter does not appear in the potential vorticity equation (11.54). Rather, it is only  $\partial_y f = \beta$  that appears. So even though the Coriolis parameter,  $f = f_o + \beta y$ , is not meridionally periodic, we can still consider a horizontally doubly-periodic domain in which there is periodicity in flow properties in both the zonal and meridional directions.

Furthermore, we can study quasi-geostrophic flows in doubly periodic domains in the presence of certain idealized background flows. Namely, for a prescribed depth-independent thermal wind flow, the corresponding buoyancy that supports this flow plays no role in the

---

<sup>6</sup>We derived this result in VOLUME 2 when studying the kinematics of a non-divergent flow.

quasi-geostrophic potential vorticity. For example, if the buoyancy has the form

$$b = \underbrace{M^2 y + N^2 z}_{\text{prescribed background}} + b'(x, y, z, t), \quad (11.89)$$

with  $M$  and  $N$  constant frequencies, then the corresponding quasi-geostrophic potential vorticity is independent of  $M$ . In this manner, the background buoyancy, though not meridionally periodic, prescribes a background thermal wind shear that is constant and so is trivially periodic.

### 11.6.2 Buoyancy equation

To develop the vertical boundary conditions on the streamfunction, consider the quasi-geostrophic buoyancy equation (11.42)

$$D_g b_g / Dt + w_{ag} N^2 = (\partial_t + \mathbf{u}_g \cdot \nabla_h) b_g + w_{ag} N^2 = 0, \quad (11.90)$$

where the buoyancy is related to the geostrophic streamfunction via

$$b_g = f_0 \partial_z \psi, \quad (11.91)$$

which leads to the equivalent form of the buoyancy equation

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) (f_0 \partial_z \psi) + w_{ag} N^2 = 0. \quad (11.92)$$

We now consider how the buoyancy equation appears at the top and bottom boundaries.

### 11.6.3 Top boundary condition

Assuming the top boundary is a material surface, we are led to the kinematic boundary condition

$$w = (\partial_t + \mathbf{u}_g \cdot \nabla_h) \eta. \quad (11.93)$$

How does this vertical velocity compare to that associated with motion in the fluid interior? To answer that question we assume an advective scaling for time derivatives so that the vertical velocity at the free surface scales as

$$w_{\text{surf}} \sim U \eta / L \sim U p / (\rho_0 g L) \sim U^2 f_0 / g, \quad (11.94)$$

where  $L$  is the horizontal length scale, and we scaled the surface pressure gradient according to the free surface gradient. A point in the fluid interior has a vertical velocity that scales according to the buoyancy equation (11.90),

$$w_{\text{interior}} \sim U b / (N^2 L) \sim f_0 U^2 / (H N^2), \quad (11.95)$$

where  $H$  is the vertical scale over which the thermal wind shear is sizable, and which scales the buoyancy frequency according to

$$N^2 \sim (g / \rho_0) \Delta \rho / H, \quad (11.96)$$

where  $\Delta\rho \ll \rho_0$  is the scale for the density difference setting the size of the buoyancy frequency. The ratio of the two vertical velocities is thus given by

$$w_{\text{surf}}/w_{\text{interior}} = H N^2/g = \Delta\rho/\rho_0 \ll 1. \quad (11.97)$$

Evidently, we can set the surface boundary condition to that of a rigid lid, in which  $w_{\text{ag}} = 0$ . In this case, the top boundary condition reduces to geostrophic advection of boundary buoyancy

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) (f_0 \partial_z \psi) = (\partial_t + \mathbf{u}_g \cdot \nabla_h) b_g = 0 \quad \text{at top boundary where } w_{\text{ag}} = 0. \quad (11.98)$$

That is, the geostrophic buoyancy at the surface is a material constant when advected by the surface geostrophic flow.

#### 11.6.4 Bottom boundary condition

With a nonzero slope in the bottom topography,  $\nabla \eta_b \neq 0$ , the bottom no normal flow kinematic boundary condition says that velocity is constrained so that

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \implies w = \mathbf{u} \cdot \nabla \eta_b, \quad (11.99)$$

where  $\hat{\mathbf{n}} = -\nabla(z - \eta_b)/|\nabla(z - \eta_b)|$  is the boundary's outward normal. Expanding this kinematic boundary condition leads to

$$w_{\text{ag}} = (\mathbf{u}_g + \mathbf{u}_{\text{ag}}) \cdot \nabla \eta_b, \quad (11.100)$$

which means there is vertical ageostrophic motion at the bottom so long as the horizontal motion is not aligned with isobaths. To be clear on the implications of this boundary condition, it is useful to examine the asymptotics by non-dimensionalizing<sup>7</sup>

$$\mathbf{u}_g = U \hat{\mathbf{u}}_0 \quad (11.101a)$$

$$w_{\text{ag}} = W \text{Ro} \hat{w}_1 \quad (11.101b)$$

$$\mathbf{u}_{\text{ag}} = U \text{Ro} \hat{\mathbf{u}}_1 \quad (11.101c)$$

$$\nabla_h \eta_b = (\mathcal{B}/L) \hat{\nabla}_h \hat{\eta}_b = (H \text{Ro}/L) \hat{\nabla}_h \hat{\eta}_b, \quad (11.101d)$$

which brings the kinematic boundary condition (11.100) to

$$f_0 H \text{Ro}^2 (\hat{\mathbf{u}}_0 + \text{Ro} \hat{\mathbf{u}}_1) \cdot \hat{\nabla}_h \hat{\eta}_b = W \text{Ro} \hat{w}_1 \implies (\hat{\mathbf{u}}_0 + \text{Ro} \hat{\mathbf{u}}_1) \cdot \hat{\nabla}_h \hat{\eta}_b = \hat{w}_1. \quad (11.102)$$

Asymptotic consistency implies that

$$\hat{\mathbf{u}}_0 \cdot \hat{\nabla}_h \hat{\eta}_b = \hat{w}_1 \implies \mathbf{u}_g \cdot \nabla \eta_b = w_{\text{ag}}. \quad (11.103)$$

Hence, for quasi-geostrophic flow, any projection of the horizontal geostrophic velocity in a direction not aligned with isobaths leads to an ageostrophic vertical velocity at the bottom.

Use of the bottom kinematic boundary condition (11.103) in the buoyancy equation (11.92) leads to the bottom boundary evolution of buoyancy

$$\partial_t (f_0 \partial_z \psi) + \mathbf{u}_g \cdot \nabla_h [f_0 \partial_z \psi + N^2 \eta_b] = 0 \quad \text{at } z = \bar{\eta}_b, \quad (11.104)$$

<sup>7</sup>The bottom topography slope is non-dimensionalized according to the shallow water discussion in Section 9.3.4.

which can be written in terms of the geostrophic buoyancy

$$\partial_t b_g + \mathbf{u}_g \cdot \nabla_h (b_g + N^2 \eta_b) = 0 \quad \text{at } z = \bar{\eta}_b, \quad (11.105)$$

The boundary condition is evaluated at the horizontally averaged bottom position,  $z = \bar{\eta}_b$ , since the more precise boundary location,  $z = \eta_b(x, y)$ , is higher order in Rossby number. Since  $N^2 \eta_b$  is time independent, the boundary condition (11.105) can be written as a material conservation statement for buoyancy at the bottom

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h)(b_g + N^2 \eta_b) = \frac{D_g(b_g + N^2 \eta_b)}{Dt} = 0 \quad \text{at } z = \bar{\eta}_b. \quad (11.106)$$

When  $\eta_b$  is a constant, and since  $N^2$  is just a function of  $z$ , then the bottom boundary condition becomes a statement about the material conservation of buoyancy along the bottom boundary

$$\frac{D_g b_g}{Dt} = 0 \quad \text{at } z = \bar{\eta}_b = 0, \quad (11.107)$$

which is the bottom analog of the top boundary condition (11.98) that holds when making the rigid lid approximation. The more general case of a nonzero bottom slope renders a source for bottom buoyancy

$$\frac{D_g b_g}{Dt} = -N^2 \mathbf{u}_g \cdot \nabla_h \eta_b = -N^2 w_{ag} \quad \text{at } z = \bar{\eta}_b. \quad (11.108)$$

In this manner we see that a material evolution of bottom buoyancy is associated with vertical ageostrophic flow at the bottom, and correspondingly to horizontal geostrophic flow that deviates from the contours of constant topography.

Finally, since  $N^2$  is a function only of  $z$ , this boundary condition can be written in terms of the stretching term appearing in the potential vorticity

$$\frac{D_g}{Dt} \left[ \frac{f_o b_g}{N^2} \right] = -f_o \mathbf{u}_g \cdot \nabla_h \eta_b \quad \text{at } z = \bar{\eta}_b. \quad (11.109)$$

Evidently, material evolution of vertical stretching is coupled to bottom geostrophic flow that crosses lines of constant topography.

## 11.7 Potential vorticity with Dirac delta sheets

Based on the boundary conditions derived in Section 11.6, we know that quasi-geostrophic theory is comprised of the material conservation of potential vorticity in the interior, along with material evolution of buoyancy at the boundaries, with the material evolution determined by the horizontal geostrophic velocity

$$\frac{D_g q}{Dt} = 0 \quad \text{for } \bar{\eta}_b < z < 0 \quad (11.110a)$$

$$\frac{D_g b_g}{Dt} = 0 \quad \text{for } z = 0 \quad (11.110b)$$

$$\frac{D_g b_g}{Dt} = -N^2 \mathbf{u}_g \cdot \nabla_h \eta_b \quad \text{for } z = \bar{\eta}_b, \quad (11.110c)$$

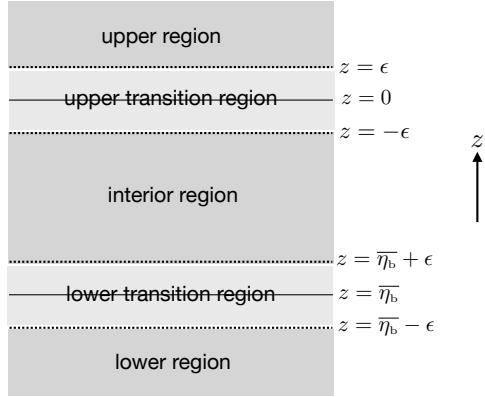


FIGURE 11.1: Geometry for the transition layers associated with the Dirac delta sheets at the upper and lower boundaries of the domain.

where the geostrophic velocity, buoyancy, material time derivative, and quasi-geostrophic potential vorticity are written in terms of the geostrophic streamfunction

$$\mathbf{u}_g = \hat{\mathbf{z}} \times \nabla \psi \quad (11.111a)$$

$$b_g = f_o \partial_z \psi \quad (11.111b)$$

$$D_g/Dt = \partial_t + \mathbf{u}_g \cdot \nabla \quad (11.111c)$$

$$q = f_o + \beta y + \nabla_h^2 \psi + \frac{\partial}{\partial z} \left[ \frac{f_o^2}{N^2} \frac{\partial \psi}{\partial z} \right] = f_o + \beta y + \nabla_h^2 \psi + \frac{\partial}{\partial z} \left[ \frac{f_o b_g}{N^2} \right]. \quad (11.111d)$$

Studying the dynamics of boundary buoyancy constitutes the focus of *surface quasi-geostrophy*, as exemplified by [Held et al. \(1995\)](#) and whose mathematical foundations were examined by [Yassin and Griffies \(2022\)](#). In this section we focus on how to understand the transition from the interior to the boundaries. We follow the method of [Bretherton \(1966\)](#), where he noted that it can be conceptually convenient to bring the boundary conditions into the potential vorticity, much like when studying Green's functions for elliptic operators in VOLUME 1. In this manner, the boundary and interior terms can be placed on equal footing and thus their contribution to the potential vorticity directly compared.<sup>8</sup>

### 11.7.1 Transition between the interior and the boundaries

The transition from the interior to the boundary occurs over an infinitesimal region surrounding each boundary. There is a corresponding jump in the buoyancy as it moves from the domain interior to the boundaries. To help understand the nature of the jump it is useful to expand the infinitesimal thickness into a tiny but finite layer whose thickness,  $\Delta z = \epsilon$ , is later taken to vanish. For notational brevity introduce

$$S = b_g/N^2, \quad (11.112)$$

and, with reference to the geometry in Figure 11.1, the full domain extent of this function is written

$$S^{\text{ext}}(z) = \underbrace{S(\bar{\eta}_b - \epsilon) \mathcal{H}(\bar{\eta}_b - \epsilon - z)}_{\text{lower region}} + \underbrace{S(z) [\mathcal{H}(z - \bar{\eta}_b - \epsilon) - \mathcal{H}(z + \epsilon)]}_{\text{interior region}} + \underbrace{S(\epsilon) \mathcal{H}(z - \epsilon)}_{\text{upper region}}, \quad (11.113)$$

<sup>8</sup>See also Section 5.4.3 of [Vallis \(2017\)](#) for general comments on this approach.

where  $\mathcal{H}$  is the **Heaviside step function** studied in VOLUME 1. Values within the outer edge of the transition regions (i.e.,  $S(\epsilon)$  and  $S(\bar{\eta}_b - \epsilon)$ ) are constructs whose values must be set consistent with the boundary conditions (11.110b) and (11.110c). We furthermore choose these values to be static, since we have no way to determine their evolution since they are outside the domain. The upper boundary satisfies the homogeneous boundary condition

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) b_g = 0, \quad (11.114)$$

which allows us to set

$$b_g(\epsilon) = 0. \quad (11.115)$$

Likewise, the lower boundary condition

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) b_g = -N^2 \mathbf{u}_g \cdot \nabla_h \eta_b, \quad (11.116)$$

is consistent with the choice

$$S(\bar{\eta}_b - \epsilon) = -\eta_b. \quad (11.117)$$

### 11.7.2 Calculating the stretching term over the extended domain

In VOLUME 1 we showed that the derivative of a Heaviside step function yields a **Dirac delta**, so that

$$\begin{aligned} \partial_z S^{\text{ext}}(z) &= \partial_z S(z) [\mathcal{H}(z - \bar{\eta}_b - \epsilon) - \mathcal{H}(z + \epsilon)] + S(z) \delta(z - \bar{\eta}_b - \epsilon) \\ &\quad - S(\bar{\eta}_b - \epsilon) \delta(\bar{\eta}_b - \epsilon - z) + S(\epsilon) \delta(z - \epsilon) - S(z) \delta(z + \epsilon), \end{aligned} \quad (11.118)$$

whose  $\epsilon \rightarrow 0$  limit is

$$\partial_z S^{\text{ext}}(z) = \partial_z S(z) [\mathcal{H}(z - \bar{\eta}_b) - \mathcal{H}(z)] + [S(z) - S(\bar{\eta}_b - \epsilon)] \delta(z - \bar{\eta}_b) + [S(\epsilon) - S(z)] \delta(z). \quad (11.119)$$

Reintroducing the definition of  $S$  as per equation (11.112) yields the stretching term contribution to the potential vorticity, now defined over the full vertical extent of the domain (i.e., the domain interior and its boundaries)

$$\begin{aligned} \partial_z (f_o b_g / N^2)^{\text{ext}} &= \partial_z (f_o b_g / N^2) [\mathcal{H}(z - \bar{\eta}_b) - \mathcal{H}(z)] \\ &\quad + \delta(z - \bar{\eta}_b) \left[ \frac{f_o b_g}{N^2} \right]_{z=\bar{\eta}_b^-}^{z=\bar{\eta}_b^+} + \delta(z) \left[ \frac{f_o b_g}{N^2} \right]_{z=0^-}^{z=0^+}, \end{aligned} \quad (11.120)$$

where the square bracket terms measure the jump in the stretching term taking place across the respective boundaries. The Heaviside term is nonzero only within the interior of the domain, whereas the two Dirac delta terms fire at their respective boundaries.

### 11.7.3 The extended potential vorticity

Making use of the extended stretching term (11.120) for the potential vorticity allows us to collapse the three equations (11.110a)-(11.110c) into the single equation valid over the domain interior as well as the domain boundaries

$$\frac{D_g q^{\text{ext}}}{Dt} = 0 \quad \text{for } \bar{\eta}_b \leq z \leq 0. \quad (11.121)$$

Bringing the above results together leads to the extended potential vorticity<sup>9</sup>

$$q^{\text{ext}} = \underbrace{f_o + \beta y + \nabla_h^2 \psi + \frac{\partial}{\partial z} \left[ \frac{f_o b_g}{N^2} \right]}_{\text{interior}} + \underbrace{\delta(z) \left[ \frac{f_o b_g}{N^2} \right]_{z=0^-}^{z=0^+} + \delta(z - \bar{\eta}_b) \left[ \frac{f_o b_g}{N^2} \right]_{z=\bar{\eta}_b^-}^{z=\bar{\eta}_b^+}}_{\text{boundary contributions}}. \quad (11.122)$$

Note that the Dirac deltas each have dimensions of inverse length<sup>10</sup>, thus making this equation dimensionally consistent. We can make use of the transition region values (11.115) and (11.117) to evaluate the jump conditions

$$\left[ \frac{f_o b_g}{N^2} \right]_{z=0^-}^{z=0^+} = \left[ \frac{f_o b_g}{N^2} \right]_{z=0} \quad (11.123a)$$

$$\left[ \frac{f_o b_g}{N^2} \right]_{z=\bar{\eta}_b^-}^{z=\bar{\eta}_b^+} = \left[ \frac{f_o b_g}{N^2} \right]_{z=\bar{\eta}_b} + f_o \eta_b, \quad (11.123b)$$

in which case the extended potential vorticity is

$$q^{\text{ext}} = \underbrace{f_o + \beta y + \nabla_h^2 \psi + \frac{\partial}{\partial z} \left[ \frac{f_o b_g}{N^2} \right]}_{\text{interior}} + \underbrace{\delta(z) (f_o b_g / N^2) + f_o \delta(z - \bar{\eta}_b) (b_g / N^2 + \eta_b)}_{\text{boundary contributions}}. \quad (11.124)$$

For some studies it is useful to ignore the buoyancy contribution at the two boundaries by setting them to zero, in which case the potential vorticity is

$$q^{\text{ext}} = \underbrace{f_o + \beta y + \nabla_h^2 \psi + \frac{\partial}{\partial z} \left[ \frac{f_o b_g}{N^2} \right]}_{\text{interior}} + \underbrace{f_o \eta_b \delta(z - \bar{\eta}_b)}_{\text{bottom topog}}. \quad (11.125)$$

Dropping the boundary buoyancy contributions allows one to focus on contributions from the bottom topography and interior processes, in isolation from the boundary buoyancy.

## 11.8 Mathematical expressions of the theory

In this section we sample various mathematical expressions of quasi-geostrophic theory.

### 11.8.1 The Jacobian form of geostropic advection

The geostrophic velocity, as a horizontally non-divergent field, can be written in terms of the quasi-geostrophic streamfunction

$$\mathbf{u}_g = \hat{\mathbf{z}} \times \nabla \psi. \quad (11.126)$$

We can thus write the following equivalent forms for the material time derivative of quasi-geostrophic PV

$$\frac{Dq}{Dt} = \partial_t q + \mathbf{u}_g \cdot \nabla_h q \quad (11.127a)$$

$$= \partial_t q + (\hat{\mathbf{z}} \times \nabla \psi) \cdot \nabla q \quad (11.127b)$$

<sup>9</sup>We drop the Heaviside terms in equation (11.122) for brevity.

<sup>10</sup>In our study of the Dirac delta in VOLUME 1, we noted that its physical dimensions are the inverse of the dimensions of its argument.

$$= \partial_t q + (\nabla \psi \times \nabla q) \cdot \hat{z} \quad (11.127c)$$

$$= \partial_t q + J(\psi, q). \quad (11.127d)$$

The final equality introduced the Jacobian operator

$$J(\psi, q) = (\nabla \psi \times \nabla q) \cdot \hat{z}, \quad (11.128)$$

which is a notation commonly used in the geophysical fluids literature.<sup>11</sup>

For a perfect fluid, in which  $Dq/Dt = 0$ , a steady state (zero Eulerian time derivative) is realized when

$$\mathbf{u}_g \cdot \nabla_h q = (\nabla \psi \times \nabla q) \cdot \hat{z} = J(\psi, q) = 0, \quad (11.129)$$

which means that the steady velocity is parallel to surfaces of constant  $q$ . We are ensured that these equalities hold if the streamfunction is a function only of the potential vorticity

$$\psi = \psi(q) \implies (\nabla \psi \times \nabla q) \cdot \hat{z} = J(\psi, q) = 0. \quad (11.130)$$

As the steady state is of physical interest, this functional relation between streamfunction and potential vorticity commonly arises in applications.

### 11.8.2 The case of constant background buoyancy frequency

Consider the quasi-geostrophic potential vorticity for the special case of a constant background buoyancy frequency,  $N^2 = \text{constant}$ , in which the relative potential vorticity (11.52) takes on the form

$$q - \beta y = \nabla_h^2 \psi + \frac{\partial}{\partial z} \left[ \frac{f_o^2}{N^2} \frac{\partial \psi}{\partial z} \right] \quad (11.131a)$$

$$= \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{f_o^2}{N^2} \frac{\partial^2 \psi}{\partial z^2} \quad (11.131b)$$

$$= [\partial_{xx} + \partial_{yy} + \partial_{zz}] \psi. \quad (11.131c)$$

For the final equality we introduced the vertical coordinate

$$z = (f/N) \tilde{z} \iff \tilde{z} = (N/f) z, \quad (11.132)$$

where  $f/N$  is the Prandtl ratio introduced by equation 9.169. Since  $|N/f| \gg 1$  the stably stratified flows considered in QG,  $\tilde{z}$  is a *stretched* vertical coordinate so that the Laplacian operator acting on  $\psi$  is anisotropic. The linear operator acting on  $\psi$  remains elliptic even in the more general case of a depth dependent stratification, thus warranting the use of elliptical solvers when performing the inversion numerically.

### 11.8.3 Potential vorticity induction and impermeability

Outside of boundaries, the potential vorticity equation is the sole prognostic equation required to evolve the quasi-geostrophic flow. Consequently, one often ignores the quasi-geostrophic velocity and buoyancy equations. Even so, we found it useful to consider the buoyancy equation in Section 11.3.5 as part of connecting quasi-geostrophic potential vorticity to Ertel potential

<sup>11</sup>Recall that we also encountered the Jacobian form for horizontally non-divergent two-dimensional advection in Section 4.3.4, as part of our study of the non-divergent barotropic flows.

vorticity. Likewise, there are occasions when it is useful to examine the velocity equation, with a similar discussion provided in Section 9.6.6. In this section we directly connect the velocity and buoyancy equations and then reveal their connection to the potential vorticity equation.

### Combining the velocity and buoyancy equations

Consider the quasi-geostrophic velocity and buoyancy equations derived in Section 11.3

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) \mathbf{u}_g + \hat{\mathbf{z}} \times (\beta y \mathbf{u}_g + f_o \mathbf{u}_{ag}) = -\nabla_h \varphi_{ag} + \mathbf{F} \quad (11.133a)$$

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) b_g + w_{ag} N^2 = \dot{b} \quad (11.133b)$$

$$f_o \hat{\mathbf{z}} \times \mathbf{u}_g + \nabla_h \varphi_g = 0 \quad (11.133c)$$

$$\nabla \cdot \mathbf{u}_g = \partial_x u_g + \partial_y v_g = 0 \quad (11.133d)$$

$$\nabla \cdot \mathbf{v}_{ag} = \nabla_h \cdot \mathbf{u}_{ag} + \partial_z w_{ag} = 0, \quad (11.133e)$$

where we added a non-conservative force per mass,  $\mathbf{F}$  (e.g., friction, wind stress), and irreversible buoyancy source,  $\dot{b}$  (e.g., diffusion, boundary fluxes). Taking  $-\hat{\mathbf{z}} \times$  on the velocity equation<sup>12</sup> and multiplying the buoyancy equation by  $f_o/N^2$  leads to

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h)(-\hat{\mathbf{z}} \times \mathbf{u}_g) + \beta y \mathbf{u}_g + f_o \mathbf{u}_{ag} = \hat{\mathbf{z}} \times \nabla_h \varphi_{ag} - \hat{\mathbf{z}} \times \mathbf{F} \quad (11.134a)$$

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h)(f_o b_g / N^2) + f_o w_{ag} = f_o \dot{b} / N^2. \quad (11.134b)$$

Introduce the following vector fields

$$\mathbf{D} \equiv -\hat{\mathbf{z}} \times \mathbf{u}_g + (f_o / N^2) b_g \hat{\mathbf{z}} = \nabla_h \psi + (f_o / N)^2 \partial_z \psi \hat{\mathbf{z}} \quad (11.135a)$$

$$\mathbf{R} \equiv -\hat{\mathbf{z}} \times \mathbf{F} + (f_o / N^2) \dot{b} \hat{\mathbf{z}}, \quad (11.135b)$$

with  $\mathbf{D}$  built from both the velocity and buoyancy fields and  $\mathbf{R}$  built from the corresponding non-conservative tendencies. These vectors allow us to combine the velocity and buoyancy equations into a single vector equation

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) \mathbf{D} + \beta y \mathbf{u}_g + f_o \mathbf{v}_{ag} = \hat{\mathbf{z}} \times \nabla_h \varphi_{ag} + \mathbf{R}. \quad (11.136)$$

### Potential vorticity induction

The divergence of  $\mathbf{D}$  yields the relative quasi-geostrophic potential vorticity

$$\nabla \cdot \mathbf{D} = q - (f_o + \beta y), \quad (11.137)$$

where we made use of the expression (11.53) for the potential vorticity. In analogy to Gauss's law of electromagnetism, we refer to  $\mathbf{D}$  as the quasi-geostrophic potential vorticity *induction vector*.<sup>13</sup> Additionally, the potential vorticity equation can be written (see Exercise 11.3)

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) q = \nabla \cdot \mathbf{R}, \quad (11.138)$$

<sup>12</sup>The operation  $-\hat{\mathbf{z}} \times$  acts to rotate by  $\pi/2$  in the clockwise direction.

<sup>13</sup>This connection between potential vorticity dynamics and electromagnetism was pointed out by [Schneider et al. \(2003\)](#) and further examined by [Maddison and Marshall \(2013\)](#).

so that the quasi-geostrophic potential vorticity flux vector

$$\mathbf{J}^q = \mathbf{u}_g q - \mathbf{R} \quad (11.139)$$

allows us to write the potential vorticity equation in the Eulerian flux-form

$$\partial_t q = -\nabla \cdot \mathbf{J}^q. \quad (11.140)$$

### Kinematic PV flux and impermeability

Taking the Eulerian time derivative of the potential vorticity induction equation (11.137) renders

$$\partial_t q = \partial_t (\nabla \cdot \mathbf{D}) = \nabla \cdot (\partial_t \mathbf{D}) \equiv -\nabla \cdot \mathbf{J}^{q\text{kin}}, \quad (11.141)$$

where we introduced the kinematic form of the potential vorticity flux

$$\mathbf{J}^{q\text{kin}} \equiv -\partial_t \mathbf{D} = \mathbf{J}^q + \nabla \times \mathbf{A}, \quad (11.142)$$

with  $\mathbf{A}$  a gauge function. This equation is analogous to the kinematic Ertel PV flux discussed in Section 8.2.2. We determine the explicit expression for the gauge function in Exercise 11.4.

## 11.9 Energetics for quasi-geostrophic flow

Consider a quasi-geostrophic fluid configured with flat upper (rigid lid) and lower boundaries, and assume the lateral boundaries are periodic or material solid boundaries. These restrictive assumptions allow us to more readily study energetics within the fluid, sans the impacts from nontrivial boundary effects. To study energetics we make use of the quasi-geostrophic buoyancy equation (11.42) and relative vorticity equation (11.47)

$$(\partial_t + \mathbf{u} \cdot \nabla_h) b = -w N^2 \quad (11.143a)$$

$$(\partial_t + \mathbf{u} \cdot \nabla_h) \zeta = -\beta v + f_o \partial_z w, \quad (11.143b)$$

where all labels are dropped from the variables to reduce clutter, and where

$$\frac{D}{Dt} = \partial_t + \mathbf{u} \cdot \nabla_h \quad b = f_o \partial_z \psi \quad u = -\partial_y \psi \quad v = \partial_x \psi \quad \zeta = \nabla_h^2 \psi. \quad (11.144)$$

### 11.9.1 Kinetic energy

The kinetic energy per mass for the total fluid domain is given by the integral

$$\mathcal{K} = \frac{1}{2} \int \mathbf{u} \cdot \mathbf{u} dV = \frac{1}{2} \int \nabla_h \psi \cdot \nabla_h \psi dV, \quad (11.145)$$

and its time derivative is

$$\frac{d\mathcal{K}}{dt} = \int \nabla_h \psi \cdot \nabla_h (\partial_t \psi) dV. \quad (11.146)$$

For the kinetic energy time derivative we noted that the fluid domain has a constant volume to allow the time derivative to move inside the integral without introducing boundary terms.

Manipulation renders

$$\frac{d\mathcal{K}}{dt} = \int \nabla_h \psi \cdot \nabla_h (\partial_t \psi) dV \quad (11.147a)$$

$$= \int [\nabla_h \cdot [\psi \nabla_h (\partial_t \psi)] - \psi \partial_t (\nabla_h^2 \psi)] dV \quad (11.147b)$$

$$= - \int \psi \partial_t \zeta dV, \quad (11.147c)$$

where we dropped the lateral boundary term and introduced relative vorticity. Use of the vorticity equation (11.143b) yields

$$\frac{d\mathcal{K}}{dt} = - \int \psi \partial_t \zeta dV = \int \psi [\mathbf{u} \cdot \nabla_h \zeta + \beta v - f_o \partial_z w] dV. \quad (11.148)$$

The first and second terms vanish since

$$\int \psi (\mathbf{u} \cdot \nabla_h \zeta + \beta v) dV = \int \psi \mathbf{u} \cdot \nabla_h (\zeta + f) dV \quad (11.149a)$$

$$= \int \psi \nabla_h \cdot (\mathbf{u} \zeta_a) dV \quad (11.149b)$$

$$= \int [\nabla_h \cdot (\psi \mathbf{u} \zeta_a) - \nabla_h \psi \cdot \mathbf{u} \zeta_a] dV = 0. \quad (11.149c)$$

The final equality holds since the boundary term vanishes, and  $\mathbf{u} \cdot \nabla_h \psi = 0$  since  $\psi$  is the streamfunction for the horizontal geostrophic flow. We are thus left with the expression for the kinetic energy evolution

$$\frac{d\mathcal{K}}{dt} = -f_o \int \psi \partial_z w dV. \quad (11.150)$$

Since the top and bottom are assumed flat and rigid, the vertical velocity vanishes on these boundaries so that

$$\frac{d\mathcal{K}}{dt} = -f_o \int \psi \partial_z w dV = -f_o \int [\partial_z (w \psi) - w \partial_z \psi] dV = f_o \int w \partial_z \psi dV. \quad (11.151)$$

Introducing the quasi-geostrophic buoyancy through  $b = f_o \partial_z \psi$  leads to

$$\frac{d\mathcal{K}}{dt} = \int w b dV. \quad (11.152)$$

Kinetic energy thus increases when vertical motion is positively correlated with buoyancy. For example, upward motion ( $w > 0$ ) of a positive buoyancy anomaly (relatively light water has  $b > 0$ ) increases kinetic energy, as does downward motion of a negative buoyancy anomaly. This behavior is also reflected in the energetics of the unapproximated fluid flow, as studied in VOLUME 2.

## 11.9.2 Available potential energy

Available potential energy was introduced in VOLUME 2 in the context of the Boussinesq ocean, where we derived the following approximate form

$$\mathcal{A}_{\text{bouss}} \approx \int \frac{1}{2} (b/N)^2 dV = \int \frac{1}{2} [(f_o/N) \partial_z \psi]^2 dV, \quad (11.153)$$

where we set  $b = f_o \partial\psi/\partial z$  for the second equality. Taking a time derivative leads to

$$\frac{d\mathcal{A}}{dt} = \int (f_o/N)^2 \partial_z \psi \partial_{tz} \psi dV = \int (f_o/N)^2 \partial_z \psi [-w N^2 - \nabla_h \cdot (\mathbf{u} b)] dV, \quad (11.154)$$

where we used the buoyancy equation (11.143a) for the second equality. The second term vanishes since

$$\int (f_o/N)^2 \partial_z \psi [\nabla_h \cdot (\mathbf{u} b)] dV = \int [(f_o/N)^2 \partial_z \psi] \mathbf{u} \cdot \nabla_h (\partial_z \psi) dV \quad (11.155a)$$

$$= \frac{1}{2} \int \nabla_h \cdot [\mathbf{u} ((f_o/N) \partial_z \psi)^2] dV \quad (11.155b)$$

$$= 0. \quad (11.155c)$$

Consequently, the quasi-geostrophic APE has a time derivative given by

$$\frac{d\mathcal{A}}{dt} = - \int w f_o \partial_z \psi dV = - \int w b dV, \quad (11.156)$$

so that the APE evolves oppositely to the kinetic energy.

### 11.9.3 Exchange of mechanical energy

We refer to the term

$$\text{buoyancy work} = \int w f_o \partial_z \psi dV = \int w b dV, \quad (11.157)$$

as the buoyancy work conversion term. It has the same form as that encountered for the conversion between potential energy and kinetic energy in the unapproximated fluid equations studied in VOLUME 2.

The evolution of kinetic energy involves the relative vorticity equation, whereas evolution of the APE involves the buoyancy equation. However, their sum remains constant in time. The reason for the exact exchange of energy is that, when the kinetic energy increases through buoyancy work, the available potential energy decreases so that

$$\frac{d(\mathcal{K} + \mathcal{A})}{dt} = 0. \quad (11.158)$$

This is a relatively simple mechanical energy budget equation reminiscent of a classical point particle studied in VOLUME 1. In particular, note the absence of a pressure work term that appears in the mechanical energy budget for other fluids, such as the Euler equation and the Boussinesq ocean (both studied in VOLUME 2). We anticipate the absence of pressure work since knowledge of potential vorticity is sufficient to know all quasi-geostrophic dynamical fields, and yet pressure plays no explicit role in potential vorticity evolution.

### 11.9.4 Scaling APE and KE

The scale for kinetic energy in a quasi-geostrophic flow is given by

$$\mathcal{K} = \int \frac{1}{2} (\nabla_h \psi \cdot \nabla_h \psi) dV \sim L^{-2} \Psi^2 V, \quad (11.159)$$

and the scale for the available potential energy is

$$\mathcal{A} = \frac{1}{2} \int \left[ \frac{f_o}{N} \frac{\partial \psi}{\partial z} \right]^2 dV \sim H^{-2} (f_o/N)^2 \Psi^2 V = L_d^{-2} \Psi^2 V, \quad (11.160)$$

where we wrote  $\Psi$  for the streamfunction scale,  $V$  for the domain volume, and  $L_d = H(N/f_o)$  is the deformation radius (see equation (11.7)). Taking the ratio yields

$$\frac{\mathcal{K}}{\mathcal{A}} \sim \left[ \frac{L_d}{L} \right]^2 = \left[ \frac{H}{L} \right]^2 \left[ \frac{N}{f_o} \right]^2 = Bu. \quad (11.161)$$

Hence, the **Burger number** is the ratio of the scale for quasi-geostrophic kinetic energy to the scale for quasi-geostrophic available potential energy. A large Burger number means that the horizontal scales of the flow are smaller than the deformation radius, in which case the quasi-geostrophic dynamics is dominated by its kinetic energy. In contrast, for scales larger than the deformation radius (not much larger, as then the flow would not satisfy quasi-geostrophic scaling), the Burger number is less than unity, in which case the quasi-geostrophic dynamics is dominated by available potential energy.



## 11.10 Exercises

### EXERCISE 11.1: A VARIETY OF POTENTIAL VORTICITIES

Give the mathematical expressions for dimensionful potential vorticity in the following fluid models. Define all terms in the respective expressions. Give the physical dimensions and/or SI units for the potential vorticity. Hint: the answers can be found somewhere in this book.

- (a) Ertel potential vorticity for compressible fluid in a rotating reference frame using potential temperature as the tracer.
- (b) Ertel potential vorticity for a Boussinesq ocean in a rotating reference frame using Archimedean buoyancy as the tracer.
- (c) Single shallow water layer on a beta plane
- (d) Continuously stratified planetary geostrophy
- (e) Continuously stratified quasi-geostrophy on a beta plane

### EXERCISE 11.2: QUASI-GEOSTROPHIC PV EVOLUTION WITH VERTICAL FRICTION

The first part of this exercise involves elements of the asymptotic method used for deriving the quasi-geostrophic equations, only now with the advent of a non-zero friction. Use is made to incorporate the non-dimensionalization detailed in our study of **Ekman mechanics** in VOLUME 2, which provides a detailed discussion of the Ekman number and Ekman layers. The second part of this exercise makes use of the **thermal wind balance** to connect horizontal buoyancy transfer to the vertical viscous transfer of horizontal momentum.

- (a) Derive the material evolution equation for quasi-geostrophic potential vorticity in a continuously stratified Boussinesq fluid in the presence of friction,  $\mathbf{F}$ . Assume the Ekman number is on the order of the Rossby number, so that the zeroth order asymptotic solution satisfies the usual inviscid  $f$ -plane geostrophic balance. Friction only appears in the first order equations.

(b) Assume friction arises just from vertical shears in the horizontal velocity, so that

$$\mathbf{F} = \frac{\partial}{\partial z} \left[ \nu \frac{\partial \mathbf{u}}{\partial z} \right], \quad (11.162)$$

where  $\nu = \nu(z)$  is a vertical eddy viscosity that is a function of depth (dimensions of squared length per time). Also assume an approximate form of quasi-geostrophic potential vorticity in which we drop relative vorticity (i.e., quasi-geostrophic potential vorticity is dominated by planetary vorticity and stretching). Determine the form for the vertical eddy viscosity so that the approximate form of quasi-geostrophic PV is laterally diffused via

$$\frac{Dq^{\text{approx}}}{Dt} = A \nabla_h^2 q^{\text{approx}}, \quad (11.163)$$

where  $A$  is a constant eddy diffusivity for the potential vorticity.

Hint: to leading order, the friction operator is a function just of the geostrophic velocity.

**EXERCISE 11.3: QGPV FLUX-FORM EQUATION WITH NON-CONSERVATIVE PROCESSES**

Derive the quasi-geostrophic potential vorticity equation (11.138). Show all the relevant steps.

Hint: the key step requires showing that

$$\nabla \cdot [(\mathbf{u}_g \cdot) \mathbf{D}] = (\mathbf{u}_g \cdot) \nabla \cdot \mathbf{D}. \quad (11.164)$$

To do so, it is useful to express  $\mathbf{u}_g$  and  $b_g$  in terms of the geostrophic streamfunction.

**EXERCISE 11.4: QGPV GAUGE FUNCTION**

Derive the gauge function,  $\mathbf{A}$ , that connects the two forms of the quasi-geostrophic flux vector as per equation (11.142).



## **Part IV**

### **Generalized vertical coordinates**

We have made use of two kinematic perspectives in this book: the Eulerian perspective as revealed by observers in a spatially fixed reference frame, and the Lagrangian perspective revealed by observers fixed on moving fluid particles. For much of fluid mechanics, these are the only two reference frames considered. In this part of the book we introduce a third that is motivated by the physics of rotating and stratified fluids. This GVC (GVC) approach is neither Eulerian nor Lagrangian, though it builds from both and thus shares much with each. Even so, it requires some novel mathematical and physical notions, thus warranting its own special treatment in this part of the book.

The flow of geophysical fluids is affected by the Coriolis acceleration arising from planetary rotation, the gravitational stratification arising in a fluid with non-homogeneous density, and interactions with the solid-earth boundaries that introduces form stresses and boundary layer turbulent processes. In describing geophysical fluid flows, we can be motivated to choose coordinates designed to simplify theoretical analyses and/or enhance the fidelity of numerical simulations. A GVC offers a mathematical framework with these motivations in mind. Specifically, a GVC is suited to describing fluid mechanics according to monotonically stacked coordinate surfaces defined by the generalized vertical coordinate,  $\sigma = \sigma(x, y, z, t)$ . The introduction of generalized vertical coordinates originates from [Starr \(1945\)](#) and [Kasahara \(1974\)](#) for atmospheric modeling and [Bleck \(1978\)](#) for ocean modeling. There is a growing use of GVC-based numerical ocean (e.g., [Griffies et al. \(2020\)](#)) and atmospheric models. This usage prompts the need to master their use in theoretical studies as well as the formulation of numerical models, thus motivating this part of the book.

There are two key mathematical features of generalized vertical coordinates that require extra mathematical precision beyond the geopotential vertical coordinate. First, surfaces of constant generalized vertical coordinate are space and time dependent, so that observers on a fixed  $\sigma$ -surface are non-inertial, and the lateral directions on these surfaces are not strictly horizontal. Second, generalized vertical coordinates are non-orthogonal, with one of its basis vectors aligned with the gravity field whereas two basis vectors are aligned along the undulating  $\sigma$ -surface (we illustrate this property in Figure 12.3). The time dependent and non-orthogonal properties of generalized vertical coordinates contrast to the static and locally orthogonal coordinates, such as Cartesian, cylindrical-polar, or spherical). To handle such generality, we find it useful to employ the general tensor machinery from VOLUME 1.

By aligning one of its basis vectors with the gravitational acceleration (as well as the local vertical component of planetary rotation), generalized vertical coordinates allow for the mathematical treatment of gravity and planetary rotation in a manner directly analogous to geopotential coordinates. That is, gravity defines the vertical direction for all GVCs, in the same way it does for geopotential coordinates. Correspondingly, GVCs also use the same contravariant components for the horizontal velocity,  $\mathbf{u} = \hat{\mathbf{x}} u + \hat{\mathbf{y}} v$ , as used with geopotential coordinates. This orientation is a great advantage of generalized vertical coordinates. However, one might question whether this benefit is worth the cost of having to learn how to handle non-orthogonal coordinates. For large-scale geophysical fluid mechanics, we contend that the benefit greatly outweighs the cost. However, some who study the physics of ocean boundary layers next to a constant sloping bottom have chosen to use orthogonal slope-normal coordinates ([Phillips \(1970\)](#), [Wunsch \(1970\)](#), [Garrett et al. \(1993\)](#), [Callies \(2018\)](#), and [Holmes et al. \(2019a\)](#), others). In Exercise 12.7 we derive the planetary geostrophic equations in the rotated orthogonal coordinate, thus illustrating how the gravitational acceleration and locally vertical planetary rotation vector are projected onto more than a single basis vector.<sup>14</sup> Although the use of

---

<sup>14</sup>There is another case where locally orthogonal coordinate are useful. Namely, for the study of anisotropic

slope-normal coordinates is widespread for idealized theoretical studies of boundary layer flows, [Peterson and Callies \(2022\)](#) illustrate the utility of a particular terrain following generalized vertical coordinate as a contrast to the slope-normal coordinate.

Taking the perspective that generalized vertical coordinates are valuable for geophysical fluid mechanics, our goal in this part of the book is to develop their mathematics and to use these coordinates in formulating the kinematics and dynamics of stratified geophysical fluid flows. The following provides a summary of the chapters where this work is presented.

- **MATHEMATICAL FOUNDATIONS:** In Chapter 12 we establish the mathematics of generalized vertical coordinates, with particular attention given to identifying aspects arising from their time dependence and non-orthogonality. Notably, the non-orthogonality makes it crucial to distinguish covariant from contravariant representations of vector fields.
- **GEOPHYSICAL FLUID MECHANICS:** We then focus in Chapter 13 on developing the kinematics of generalized vertical coordinates. Dia-surface transport plays a central role in the kinematics, thus warranting a deep dive into its many facets. We then study the suite of geophysical fluid dynamical equations using generalized vertical coordinates.
- **CONTINUOUS AND LAYERED ISOPYCNAL EQUATIONS:** In Chapter 14 we focus on the hydrostatic ocean Boussinesq equations as formulated using buoyancy as the vertical coordinate. This work provides a mathematical and physical basis for *isopycnal* models of the ocean and isentropic models of the atmosphere. We then specialize the continuous equations to their vertically discrete (layered) form and explore [thickness weighted averaging \(TWA\)](#) within the perfect fluid stacked shallow water model.

#### MATHEMATICS IN THIS PART

The mathematics in this chapter rely extensively on the tensor analysis from VOLUME 1, including Cartesian tensors and general tensors.

---

tracer diffusion within the ocean interior, in which diffusive fluxes from mesoscale turbulence is suitably oriented according to neutral directions. A full study of neutral diffusion provided in Section 18.4.



# Chapter 12

## FOUNDATIONS IN TENSOR CALCULUS

In this chapter we present the tensor calculus foundations for describing fluid flow using a GVC (GVC). Figure 12.1 offers a schematic of how these coordinates monotonically partition the vertical direction. Such coordinates are of particular use for stratified fluid mechanics, where it is often convenient to use a vertical coordinate that is distinct from, but uniquely and invertibly related to, the geopotential vertical coordinate,  $z$ . To allow us to focus on the vertical coordinate, we here make use of Cartesian coordinates for the lateral/horizontal directions.<sup>1</sup>

### CHAPTER GUIDE

We make use of the general tensor analysis detailed in VOLUME 1, with the maths here forming the foundations for subsequent chapters on generalized vertical coordinates.

<b>12.1 Loose threads</b>	408
<b>12.2 Cartesian and generalized vertical coordinates</b>	409
12.2.1 Notation conventions	409
12.2.2 Notation for the vertical position of a GVC surface	410
<b>12.3 Example generalized vertical coordinates</b>	410
12.3.1 Pressure coordinates	410
12.3.2 Terrain following coordinates	411
12.3.3 Rescaled geopotential and pressure coordinates	411
12.3.4 Bottom slope oriented coordinates	412
12.3.5 Isopycnal or buoyancy coordinates	412
<b>12.4 Spatial basis vectors</b>	412
12.4.1 An important detail regarding the transformation matrix	413
12.4.2 Expressions for the basis vectors	413
<b>12.5 Basis one-forms</b>	414
12.5.1 More on the inverse transformation matrix	414
12.5.2 Basis one-forms	415
12.5.3 Verifying the bi-orthogonality relation	415
<b>12.6 Triple product identities</b>	415
<b>12.7 Position of a point in space</b>	417
<b>12.8 Transforming components of a vector</b>	418
<b>12.9 Representing the velocity vector</b>	419
12.9.1 Contravariant components	419
12.9.2 Covariant components	419

<sup>1</sup>A coordinate transformation to general orthogonal curvilinear coordinates, such as spherical coordinates, is straightforward relative to the niceties arising from the non-orthogonal nature of generalized vertical coordinates.

---

12.9.3	Introducing the material time derivative . . . . .	420
12.9.4	Equivalence to the Cartesian velocity representation . . . . .	420
<b>12.10</b>	<b>Metric tensor using generalized vertical coordinates . . . . .</b>	<b>420</b>
12.10.1	Representation of the metric tensor and its inverse . . . . .	421
12.10.2	Jacobian of transformation . . . . .	421
12.10.3	Covariant and contravariant representations . . . . .	422
12.10.4	Metric tensor for points living on a constant $\sigma$ -surface . . . . .	422
<b>12.11</b>	<b>Volume element and the Levi-Civita tensor . . . . .</b>	<b>423</b>
<b>12.12</b>	<b>Partial derivative operators . . . . .</b>	<b>423</b>
12.12.1	Analytical derivation . . . . .	423
12.12.2	Geometrical derivation . . . . .	424
12.12.3	The gradient as a tensor operator . . . . .	425
<b>12.13</b>	<b>Material time derivative . . . . .</b>	<b>425</b>
<b>12.14</b>	<b>Divergence of a vector and the divergence theorem . . . . .</b>	<b>426</b>
<b>12.15</b>	<b>The diffusion operator . . . . .</b>	<b>426</b>
12.15.1	Continuous expression . . . . .	426
12.15.2	Layer thickness weighted diffusion operator . . . . .	427
<b>12.16</b>	<b>Vorticity . . . . .</b>	<b>428</b>
12.16.1	Contravariant components . . . . .	428
12.16.2	Transforming from Cartesian coordinates . . . . .	428
<b>12.17</b>	<b>Circulation . . . . .</b>	<b>429</b>
<b>12.18</b>	<b>Exercises . . . . .</b>	<b>430</b>

---

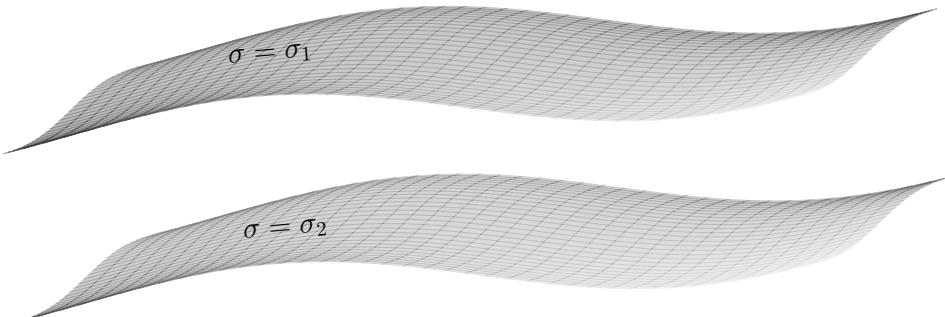


FIGURE 12.1: Schematic to illustrate the geometry of two surfaces of constant GVC with values,  $\sigma(x, y, z, t) = \sigma_1$  and  $\sigma(x, y, z, t) = \sigma_2$ . The surfaces undulate in space and time and are assumed to layer the fluid in a manner that the surface normal,  $\hat{\mathbf{n}} = \nabla\sigma/|\nabla\sigma|$ , always has a non-zero projection onto the vertical:  $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}} = \partial_z\sigma/|\nabla\sigma| \neq 0$ . That is, the surfaces never become vertical nor do they overturn. It also means that there is a 1-to-1 invertible relation between  $\sigma$  and geopotential, so that there is a monotonic relation in which specifying  $(x, y, \sigma(x, y, z, t))$  is sufficient to yield a unique  $z$ .

## 12.1 Loose threads

- Compute geodesics for motion restricted to an isopycnal, using the metric from Section 12.10.4. Will need to compute the Christoffel symbols; could be a bunch of work but one question is whether there is a relation to neutral directions that appears.
- Consider a homework that works through the  $s = z f_0/f$  generalized vertical coordinate. What does  $w^*$  look like? This inspired by pycnocnestic vertical coordinate in [Fraser et al. \(2026\)](#).

## 12.2 Cartesian and generalized vertical coordinates

In this section we establish some of the basic notational conventions for Cartesian coordinates and generalized vertical coordinates.

### 12.2.1 Notation conventions

We make use of the symbol,  $\sigma$ , for a generalized vertical coordinate (**GVC**) as in Figure 12.1. The coordinate triad,  $(x, y, \sigma)$ , is *not* orthogonal, with lack of orthogonality the central property of **GVCs** that influences nearly all aspects of their calculus. To help develop the mathematics for transforming between Cartesian coordinates and **GVCs**, it is important to distinguish the two coordinate systems, and we make use of a few options that are motivated by the usage. Namely, we write the time coordinate and spatial Cartesian coordinates according to

$$\xi^\alpha = (\xi^0, \xi^a) = (\xi^0, \xi^1, \xi^2, \xi^3) = (t, x, y, z) \text{ with } \alpha = 0, 1, 2, 3, \text{ and } a = 1, 2, 3. \quad (12.1)$$

The Latin coordinate label,  $a$ , runs over the spatial coordinates 1, 2, 3, whereas the Greek label,  $\alpha$ , also includes the time coordinate that has  $\alpha = 0$ . The corresponding generalized vertical coordinates are denoted with an overbar

$$\bar{\xi}^\alpha = (\bar{\xi}^0, \bar{\xi}^1, \bar{\xi}^2, \bar{\xi}^3) = (\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (12.2)$$

The 1-to-1 coordinate transformation between Cartesian and **GVC** coordinates is written

$$\bar{\xi}^0 = \xi^0 \iff \bar{t} = t \quad (12.3a)$$

$$\bar{\xi}^1 = \xi^1 \iff \bar{x} = x \quad (12.3b)$$

$$\bar{\xi}^2 = \xi^2 \iff \bar{y} = y \quad (12.3c)$$

$$\bar{\xi}^3 = \sigma(x, y, z, t), \quad (12.3d)$$

with the final relation expressing the generalized vertical coordinate as a function of space and time. We ordered the coordinates appearing in  $\sigma(x, y, z, t)$  with time in the last position, which is the conventional ordering in this book for functions of space and time. We continue this ordering even though the zeroth coordinate is time.

The coordinate transformation between Cartesian space-time and **GVC** space-time is invertible so that we can write

$$\xi^0 = \bar{\xi}^0 \quad (12.4a)$$

$$\xi^1 = \bar{\xi}^1 \quad (12.4b)$$

$$\xi^2 = \bar{\xi}^2 \quad (12.4c)$$

$$\xi^3 = \bar{\xi}^3(\bar{x}, \bar{y}, \sigma, \bar{t}) = \xi^3(x, y, \sigma, t). \quad (12.4d)$$

The relation

$$\xi^3 = \bar{\xi}^3(\bar{\xi}^\alpha) = \xi^3(x, y, \sigma, t) \quad (12.5)$$

provides the vertical position of a given  $\sigma$  surface.

### 12.2.2 Notation for the vertical position of a GVC surface

Since  $\xi^3 = z$ , one commonly writes

$$z = z(x, y, \sigma, t). \quad (12.6)$$

However, this expression is prone to confusion since the meaning of  $z$  is overloaded.<sup>2</sup> Namely, one meaning ascribes to  $z$  a particular value of the vertical position; say  $z = -100$  m. The other meaning, as on the right hand side of equation (12.6), is for  $z$  as the geopotential position of a particular  $\sigma$  surface, with this position determined by setting the horizontal position,  $(x, y)$ , the generalized vertical coordinate,  $\sigma$ , and the time. It takes practice to routinely distinguish when  $z$  refers to a particular vertical position or refers to a coordinate function. The distinction is important for both fundamentals and practices.

Those who routinely work with generalized vertical coordinates typically have no problem with the overloaded meaning for  $z$ . Indeed, after reading this chapter we should be able to hold the two meanings in our mind without confusion. Even so, to help build the necessary brain muscle, we commonly write the vertical position of a generalized vertical coordinate as

$$z = \eta(x, y, \sigma, t). \quad (12.7)$$

The symbol,  $\eta$ , is used throughout this book to represent the vertical position of a surface, such as the ocean free surface ( $z = \eta(x, y, t)$ ), solid-earth topography ( $z = \eta_b(x, y)$ ), or, as used here, for the vertical position of a specific generalized vertical coordinate surface labeled by  $\sigma$ . Hence, for example, the vertical position of a pressure surface of chosen value  $p$  is given by the functional relation

$$\xi^3 = z = \eta(\bar{x}, \bar{y}, p, \bar{t}) = \eta(x, y, p, t). \quad (12.8)$$

We make use of the  $\eta$  nomenclature where it seems useful, but gradually sprinkle in more use of the  $z(x, y, \sigma, t)$  notation since it is natural for many of the formulations.

## 12.3 Example generalized vertical coordinates

Before further diving into the maths, we here offer some examples of generalized vertical coordinates commonly used to study geophysical fluid flows. Figure 12.2 displays three of the coordinates described below.

### 12.3.1 Pressure coordinates

For hydrostatic compressible fluids, such as the large-scale atmosphere, pressure is a convenient choice since it absorbs the appearance of density in many formula, such as mass continuity as discussed in Section 13.10.2. Hence, a natural expression of the compressible hydrostatic equations of motion makes use of pressure rather than geopotential for the vertical coordinate, in which case we set  $\sigma = p(x, y, z, t)$ .

---

<sup>2</sup>We use the term *overloaded* as in computer programming, where one often finds an operator given more than one usage depending on the inputs.

### 12.3.2 Terrain following coordinates

The terrain following coordinate used in Boussinesq ocean studies is given by

$$\sigma = \frac{z - \eta_s}{-\eta_b + \eta_s} \quad \text{terrain following Boussinesq coordinate,} \quad (12.9)$$

where  $z = \eta_s(x, y, t)$  is the vertical position of the ocean free surface. We illustrate these coordinates in Figure 12.2. The terrain following coordinate is non-dimensional and extends from  $\sigma = 0$  at the ocean surface ( $z = \eta_s(x, y, t)$ ) to  $\sigma = -1$  at the ocean bottom ( $z = \eta_b(x, y)$ ). For rigid lid ocean models with  $\eta_s = 0$ , the terrain following coordinate becomes time independent

$$\sigma = -\frac{z}{\eta_b} \quad \text{terrain following rigid lid ocean.} \quad (12.10)$$

Finally, for a non-Boussinesq fluid it is more convenient to use pressure to define the terrain following coordinate so that

$$\sigma = \frac{p - p_a}{p_b - p_a} \quad \text{terrain following pressure coordinate.} \quad (12.11)$$

In this equation,  $p_a$  is the pressure applied at the ocean surface and  $p_b$  is the pressure at the ocean bottom. For use in studying the atmosphere we might set  $p_a$  as the top of the atmosphere pressure, which is typically assumed to be zero as in [Phillips \(1957\)](#).

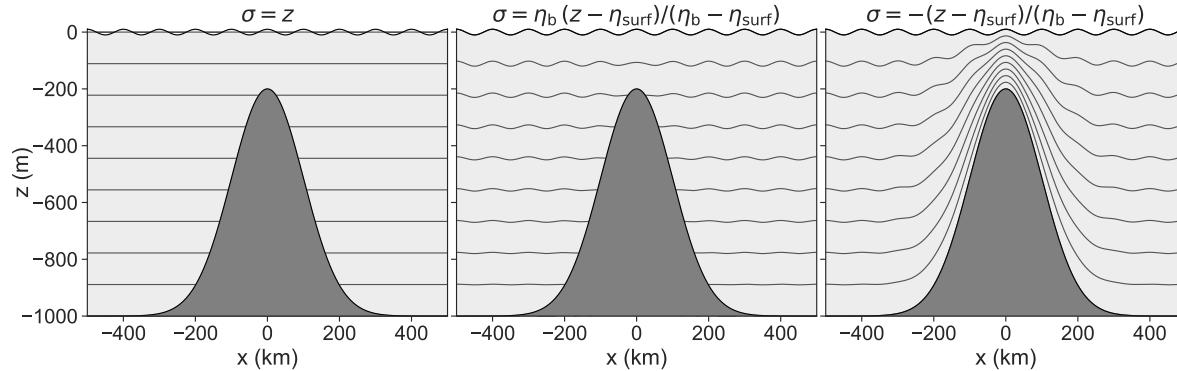


FIGURE 12.2: Three examples of generalized vertical coordinates in the presence of a Gaussian hill and a free surface, with a highly exaggerated free surface amplitude of 20 m allowing for its visualization (2 m is a more physically realistic amplitude for large-scale ocean flows). Left panel: geopotential coordinates, which are horizontal throughout the domain. Middle panel:  $\sigma = z^* = \eta_b(z - \eta_s)/(\eta_b - \eta_s)$ , which ranges from  $\eta_b \leq z^* \leq 0$  and is quasi-horizontal and yet incorporates motion of the free surface. Right panel: terrain following coordinate,  $\sigma = -(z - \eta_s)/(\eta_b - \eta_s)$ , which ranges from  $-1 \leq \sigma \leq 0$  and incorporates undulations of the bottom topography and free surface. Note that near the ocean surface there can be regions where the free surface undulations cause water to fall outside of the domain of the geopotential vertical coordinate domain, with such situations problematic for a numerical model.

### 12.3.3 Rescaled geopotential and pressure coordinates

Numerical models making use of terrain-following coordinates from Section 12.3.2 can suffer from nontrivial errors in the horizontal pressure gradient calculation, which we discuss in Section 13.14. To partially remedy this problem, it is common to follow [Adcroft and Campin \(2004\)](#) (see [Black \(1994\)](#) for the atmosphere analog) by working with a nearly horizontal coordinate

that absorbs the free surface motion

$$\sigma = z^* = \frac{\eta_b(z - \eta_s)}{\eta_b - \eta_s} \quad \text{rescaled geopotential ocean coordinate,} \quad (12.12)$$

where  $z^*$  is the symbol used in the literature. We illustrate these coordinates in Figure 12.2. Like the terrain following coordinate (12.9),  $z^*$  has a time-independent range, only here it is given by the dimensional range

$$\eta_b \leq z \leq \eta_s \implies \eta_b \leq z^* \leq 0. \quad (12.13)$$

The corresponding coordinate for a non-Boussinesq hydrostatic fluid is

$$\sigma = p^* = \frac{p_{bo}(p - p_a)}{p_b - p_a} \quad \text{rescaled pressure coordinate,} \quad (12.14)$$

where  $p_{bo}$  is the bottom pressure for a resting fluid.

### 12.3.4 Bottom slope oriented coordinates

For their studies of ocean mixing along a constant sloped bottom, [Peterson and Callies \(2022\)](#) make use of an alternative to the traditional terrain following coordinates from Section 12.3.2, here defining a bottom slope oriented coordinate (recall  $\eta_b = \eta_b(x, y)$ )

$$\sigma = z - \mathbf{x} \cdot \nabla \eta_b = z - x \partial \eta_b / \partial x - y \partial \eta_b / \partial y, \quad (12.15)$$

with  $\nabla \eta_b$  the slope of the bottom topography.

### 12.3.5 Isopycnal or buoyancy coordinates

Buoyancy surfaces are material when there is no mixing. Hence, for the study of perfect fluid mechanics it is quite convenient to use the Archimedian buoyancy,  $b$ , as the vertical coordinate,  $\sigma = b(x, y, z, t)$ . Equivalently, one may choose the potential density as the vertical coordinate. We have much more to say about isopycnal vertical coordinates in Chapter 14 when developing the equations for isopycnal ocean models.

## 12.4 Spatial basis vectors

Write the three Cartesian basis vectors,  $e_a$ , as

$$(e_1, e_2, e_3) = (\hat{x}, \hat{y}, \hat{z}). \quad (12.16)$$

We now consider the transformation of these basis vectors into their corresponding GVC (GVC) representation. This transformation is given by

$$e_{\bar{a}} = \Lambda^a_{\bar{a}} e_a, \quad (12.17)$$

where the transformation matrix is<sup>3</sup>

$$\Lambda^a_{\bar{a}} = \begin{bmatrix} \partial x / \partial \bar{x} & \partial x / \partial \bar{y} & \partial x / \partial \sigma \\ \partial y / \partial \bar{x} & \partial y / \partial \bar{y} & \partial y / \partial \sigma \\ \partial z / \partial \bar{x} & \partial z / \partial \bar{y} & \partial z / \partial \sigma \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial \eta / \partial \bar{x} & \partial \eta / \partial \bar{y} & \partial \eta / \partial \sigma \end{bmatrix}, \quad (12.18)$$

where the second equality wrote  $z = \eta(x, y, \sigma, t)$  for the vertical position of a  $\sigma$  surface (see Section 12.2). The diagonal unit values for the space-space components arise since a horizontal position is the same for Cartesian coordinates and generalized vertical coordinates, and the horizontal directions are orthogonal. Likewise, the time coordinate does not change when changing  $\bar{x}$ ,  $\bar{y}$ , or  $\sigma$ . Additionally,  $\partial x / \partial \sigma = \partial y / \partial \sigma = 0$  since the horizontal position remains unchanged when moving across a GVC surface. In contrast, non-zero values for  $\partial \eta / \partial \bar{x}$  and  $\partial \eta / \partial \bar{y}$  arise since we generally change vertical position when moving horizontally along a sloped  $\sigma$  surface. Finally, the element  $\partial \eta / \partial \sigma$  is nonzero due to vertical coordinate stratification of the fluid when represented using generalized vertical coordinates.

#### 12.4.1 An important detail regarding the transformation matrix

To further detail how to produce elements of the transformation matrix (12.18), it is crucial to ensure that the proper variables are held fixed when performing the partial derivatives. For example, consider the top row

$$\Lambda^1_{\bar{a}} = [ [\partial x / \partial \bar{x}]_{\bar{y}, \sigma} \quad [\partial x / \partial \bar{y}]_{\bar{x}, \sigma} \quad [\partial x / \partial \sigma]_{\bar{x}, \bar{y}} ] \quad (12.19)$$

Since  $x = \bar{x}$ , all elements vanish except for the first. Namely,  $[\partial x / \partial \bar{y}]_{\bar{x}, \sigma} = 0$  since  $x$  cannot change when  $\bar{x}$  is fixed. The same idea leads to the results for  $y$  derivatives.

#### 12.4.2 Expressions for the basis vectors

Use of the transformation matrix (12.18) renders the spatial components of the GVC basis vectors

$$\mathbf{e}_{\bar{1}} = \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial \eta / \partial \bar{x}) \quad (12.20a)$$

$$\mathbf{e}_{\bar{2}} = \hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial \eta / \partial \bar{y}) \quad (12.20b)$$

$$\mathbf{e}_{\bar{3}} = \hat{\mathbf{z}} (\partial \eta / \partial \sigma). \quad (12.20c)$$

The basis vectors  $\mathbf{e}_{\bar{1}}$  and  $\mathbf{e}_{\bar{2}}$  have a vertical component due to sloping GVC surfaces. These basis vectors lie within the tangent plane of the GVC surface. The basis vector  $\mathbf{e}_{\bar{3}}$  is purely vertical and has a non-unit magnitude due to the inverse vertical stratification,

$$\partial \eta / \partial \sigma = (\partial \sigma / \partial \eta)^{-1}. \quad (12.21)$$

The left panel of Figure 12.3 illustrates the basis vectors.

As an example, consider the rigid lid terrain following vertical coordinate (12.10), where  $\sigma = -z/\eta_b$ . In this case, the vertical position of a generalized vertical surface is given by

---

<sup>3</sup>Equation (12.18) makes use of a common overloaded meaning for  $\Lambda^a_{\bar{a}}$ . First, it can mean the element in the  $a$ -row and  $\bar{a}$ -column of the transformation matrix,  $\Lambda$ . Second, it can be the matrix itself, with the exposed indices useful to reveal how the matrix is used within tensorial manipulations. It is this second meaning that is used for equation (12.18).

$\eta = -\sigma \eta_b$  so that the basis vectors are

$$\mathbf{e}_{\bar{1}} = \hat{\mathbf{x}} - \hat{\mathbf{z}} \sigma (\partial \eta_b / \partial \bar{x}) \quad \text{and} \quad \mathbf{e}_{\bar{2}} = \hat{\mathbf{y}} - \hat{\mathbf{z}} \sigma (\partial \eta_b / \partial \bar{y}) \quad \text{and} \quad \mathbf{e}_{\bar{3}} = -\hat{\mathbf{z}} \eta_b. \quad (12.22)$$

Notice how  $\mathbf{e}_{\bar{3}}$  has physical dimensions of length, whereas  $\mathbf{e}_{\bar{1}}$  and  $\mathbf{e}_{\bar{2}}$  are non-dimensional. For the bottom slope oriented coordinate (12.15), with  $\sigma = z - \mathbf{x} \cdot \nabla \eta_b$ , we have  $\eta = \sigma + \mathbf{x} \cdot \nabla \eta_b$  and the corresponding basis vectors

$$\mathbf{e}_{\bar{1}} = \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial \eta_b / \partial \bar{x}) \quad \text{and} \quad \mathbf{e}_{\bar{2}} = \hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial \eta_b / \partial \bar{y}) \quad \text{and} \quad \mathbf{e}_{\bar{3}} = \hat{\mathbf{z}}. \quad (12.23)$$

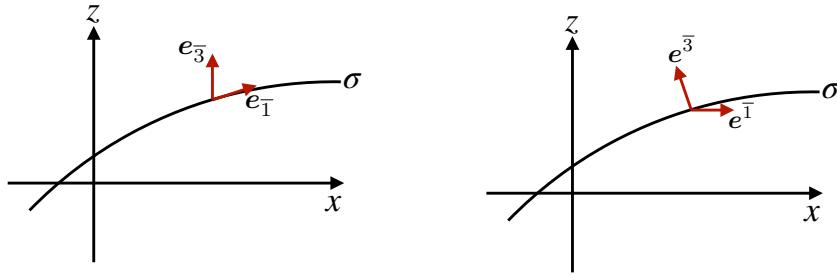


FIGURE 12.3: Illustrating the basis vectors (left panel) and basis one-forms (right panel) for generalized vertical coordinates. The  $\mathbf{e}_{\bar{3}}$  basis vector is vertical whereas  $\mathbf{e}_{\bar{1}}$  and  $\mathbf{e}_{\bar{2}}$  lie within the tangent plane to the  $\sigma$  surface. As a complement, the basis one-form,  $\mathbf{e}^{\bar{3}}$ , is normal to the  $\sigma$  surface whereas the basis one-forms,  $\mathbf{e}^{\bar{1}}$  and  $\mathbf{e}^{\bar{2}}$ , are horizontal.

## 12.5 Basis one-forms

In addition to basis vectors, we make use of *one-forms*, which are dual to vectors. The basis one-forms are obtained by transforming from Cartesian into GVCs through use of the inverse transformation matrix

$$\mathbf{e}^{\bar{a}} = \Lambda^{\bar{a}}_a \mathbf{e}^a, \quad (12.24)$$

where the inverse transformation matrix takes the form

$$\Lambda^{\bar{a}}_a = \begin{bmatrix} \partial \bar{x} / \partial x & \partial \bar{x} / \partial y & \partial \bar{x} / \partial z \\ \partial \bar{y} / \partial x & \partial \bar{y} / \partial y & \partial \bar{y} / \partial z \\ \partial \bar{\sigma} / \partial x & \partial \bar{\sigma} / \partial y & \partial \bar{\sigma} / \partial z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial \sigma / \partial x & \partial \sigma / \partial y & \partial \sigma / \partial z \end{bmatrix}. \quad (12.25)$$

As for the transformation matrix (12.18), the unit diagonal values arise since a horizontal position in Cartesian and GVCs is the same and the horizontal directions are orthogonal. Likewise,  $\partial \bar{x} / \partial z = \partial \bar{y} / \partial z = 0$  since the horizontal position on a GVC surface remains unchanged when moving across a depth surface. The nonzero values for  $\partial \sigma / \partial x$ ,  $\partial \sigma / \partial y$ , and  $\partial \sigma / \partial z$ , arise in the presence of horizontal and vertical stratification of the generalized vertical coordinate.

### 12.5.1 More on the inverse transformation matrix

When computing elements of the inverse transformation matrix (12.25), it is crucial to ensure that the proper variables are held fixed. For example, consider the top row where we compute

$$\Lambda^{\bar{1}}_a = [ [\partial \bar{x} / \partial x]_{y,z} \quad [\partial \bar{x} / \partial y]_{x,z} \quad [\partial \bar{x} / \partial z]_{x,y} ]. \quad (12.26)$$

Just as for the transformation matrix (12.19), since  $x = \bar{x}$ , all but the first element vanish in equation (12.26). Namely,  $[\partial\bar{x}/\partial y]_{x,z} = 0$  since the  $\bar{x}$  cannot change when  $x$  is fixed. The same idea holds for the  $\bar{y}$  row.

### 12.5.2 Basis one-forms

Use of the inverse transformation matrix (12.25) renders the spatial components of the GVC basis one-forms

$$\mathbf{e}^{\bar{1}} = \hat{\mathbf{x}} \quad (12.27a)$$

$$\mathbf{e}^{\bar{2}} = \hat{\mathbf{y}} \quad (12.27b)$$

$$\mathbf{e}^{\bar{3}} = \mathbf{e}^a \partial_a \sigma = \hat{\mathbf{x}} (\partial\sigma/\partial x) + \hat{\mathbf{y}} (\partial\sigma/\partial y) + \hat{\mathbf{z}} (\partial\sigma/\partial z) = \nabla\sigma. \quad (12.27c)$$

The left panel of Figure 12.3 illustrates the basis one-forms.

As an example, consider again the rigid lid terrain following coordinate (12.10),  $\sigma = -z/\eta_b$ , in which case

$$\mathbf{e}^{\bar{3}} = \nabla\sigma = -(1/\eta_b) [\hat{\mathbf{z}} - (z/\eta_b) \nabla\eta_b]. \quad (12.28)$$

Similarly, the bottom slope oriented coordinate (12.15), with  $\sigma = z - \mathbf{x} \cdot \nabla\eta_b$ , has

$$\mathbf{e}^{\bar{3}} = \hat{\mathbf{z}} - \nabla\eta_b - \hat{\mathbf{x}} (\mathbf{x} \cdot \partial_x \nabla\eta_b) - \hat{\mathbf{y}} (\mathbf{x} \cdot \partial_y \nabla\eta_b) = \hat{\mathbf{z}} - \nabla\eta_b - (\mathbf{x} \cdot \nabla) \nabla\eta_b. \quad (12.29)$$

In the case where the bottom slope is constant in both directions then this result simplifies to

$$\mathbf{e}^{\bar{3}} = \hat{\mathbf{z}} - \nabla\eta_b. \quad (12.30)$$

### 12.5.3 Verifying the bi-orthogonality relation

The basis one-forms satisfy the bi-orthogonality relation with the basis vectors

$$\mathbf{e}^{\bar{a}} \cdot \mathbf{e}_{\bar{b}} = \delta^{\bar{a}}_{\bar{b}}. \quad (12.31)$$

This identity is trivial to verify for all  $\bar{a} = 1, 2, 3$ .

## 12.6 Triple product identities

We find various occasions to make use of a suite of triple product identities that hold for generalized vertical coordinates. For this purpose we write  $\sigma$  as a composite function

$$\sigma = \sigma(x, y, z, t) = \sigma[x, y, z(\bar{t}, \bar{x}, \bar{y}, \sigma), t], \quad (12.32)$$

with  $\eta(\bar{x}, \bar{y}, \sigma, \bar{t})$  written as  $z(\bar{x}, \bar{y}, \sigma, \bar{t})$  as it here eases the manipulations. Use of the chain rule leads to the space-time differential increment

$$d\sigma = dt \left[ \frac{\partial\sigma}{\partial t} \right]_{x,y,z} + dx \left[ \frac{\partial\sigma}{\partial x} \right]_{t,y,z} + dy \left[ \frac{\partial\sigma}{\partial y} \right]_{t,x,z} + dz \left[ \frac{\partial\sigma}{\partial z} \right]_{t,x,y}. \quad (12.33)$$

Likewise, writing  $z = z[\bar{t}, \bar{x}, \bar{y}, \sigma]$  leads to the space-time differential increment  $dz$

$$dz = d\bar{t} \left[ \frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} + d\bar{x} \left[ \frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} + d\bar{y} \left[ \frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} + d\sigma \left[ \frac{\partial z}{\partial \sigma} \right]_{\bar{t}, \bar{x}, \bar{y}}. \quad (12.34)$$

We note the identities

$$\left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \sigma} \right]_{\bar{t}, \bar{x}, \bar{y}} = 1 \quad d\bar{t} = dt \quad d\bar{x} = dx \quad d\bar{y} = dy, \quad (12.35)$$

which follow since  $t = \bar{t}$ ,  $x = \bar{x}$ , and  $y = \bar{y}$ . Substituting equation (12.34) into equation (12.33) and making use of the identities (12.35) yields

$$0 = dt \left[ \left[ \frac{\partial \sigma}{\partial t} \right]_{x, y, z} + \left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} \right] \\ + dx \left[ \left[ \frac{\partial \sigma}{\partial x} \right]_{t, y, z} + \left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} \right] + dy \left[ \left[ \frac{\partial \sigma}{\partial y} \right]_{t, x, z} + \left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} \right]. \quad (12.36)$$

For this equation to hold with general increments  $dt$ ,  $dx$ , and  $dy$  requires that each bracketed term vanish, which in turn leads to the following set of triple product identities

$$\left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} = - \left[ \frac{\partial \sigma}{\partial t} \right]_{x, y, z} \quad (12.37a)$$

$$\left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} = - \left[ \frac{\partial \sigma}{\partial x} \right]_{t, y, z} \quad (12.37b)$$

$$\left[ \frac{\partial \sigma}{\partial z} \right]_{t, x, y} \left[ \frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} = - \left[ \frac{\partial \sigma}{\partial y} \right]_{t, x, z}. \quad (12.37c)$$

If the vertical stratification,  $\partial \sigma / \partial z$ , is non-zero, the triple product identities are equivalent to

$$\left[ \frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} = - \frac{[\partial \sigma / \partial t]_{x, y, z}}{[\partial \sigma / \partial z]_{t, x, y}} = - \left[ \frac{\partial \sigma}{\partial t} \right]_{x, y, z} \left[ \frac{\partial z}{\partial \sigma} \right]_{t, x, y} \quad (12.38a)$$

$$\left[ \frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} = - \frac{[\partial \sigma / \partial x]_{t, y, z}}{[\partial \sigma / \partial z]_{t, x, y}} = - \left[ \frac{\partial \sigma}{\partial x} \right]_{t, y, z} \left[ \frac{\partial z}{\partial \sigma} \right]_{t, x, y} \quad (12.38b)$$

$$\left[ \frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} = - \frac{[\partial \sigma / \partial y]_{t, x, z}}{[\partial \sigma / \partial z]_{t, x, y}} = - \left[ \frac{\partial \sigma}{\partial y} \right]_{t, x, z} \left[ \frac{\partial z}{\partial \sigma} \right]_{t, x, y}. \quad (12.38c)$$

Since  $t = \bar{t}$ ,  $x = \bar{x}$ , and  $y = \bar{y}$  we can write these identities in the more succinct form

$$\left[ \frac{\partial z}{\partial \bar{t}} \right]_\sigma = \left[ \frac{\partial \eta}{\partial \bar{t}} \right]_\sigma = - \frac{[\partial \sigma / \partial t]_z}{[\partial \sigma / \partial z]} \quad (12.39a)$$

$$\left[ \frac{\partial z}{\partial \bar{x}} \right]_\sigma = \left[ \frac{\partial \eta}{\partial \bar{x}} \right]_\sigma = - \frac{[\partial \sigma / \partial x]_z}{[\partial \sigma / \partial z]} \quad (12.39b)$$

$$\left[ \frac{\partial z}{\partial \bar{y}} \right]_\sigma = \left[ \frac{\partial \eta}{\partial \bar{y}} \right]_\sigma = - \frac{[\partial \sigma / \partial y]_z}{[\partial \sigma / \partial z]}, \quad (12.39c)$$

where we reintroduced the notation  $\eta(x, y, \sigma, t) = z(x, y, \sigma, t)$ . These identities are quite useful for manipulating equations involving GVCs. In particular, equations (12.39b) and (12.39c)

provide alternate expressions for the slope of  $\sigma$  isosurfaces relative to the horizontal plane (see Section 12.12).

## 12.7 Position of a point in space

We are familiar with locating a point in space using Cartesian coordinates. We can also specify the position using generalized vertical coordinates by making use of the basis vectors (12.20a)-(12.20c)

$$\mathcal{P} = \xi^{\bar{a}} e_{\bar{a}} \quad (12.40a)$$

$$= \bar{x} [\hat{x} + (\partial\eta/\partial\bar{x}) \hat{z}] + \bar{y} [\hat{y} + (\partial\eta/\partial\bar{y}) \hat{z}] + \sigma (\partial\eta/\partial\sigma) \hat{z} \quad (12.40b)$$

$$= \hat{x} \bar{x} + \hat{y} \bar{y} + \hat{z} [\bar{x} (\partial\eta/\partial\bar{x}) + \bar{y} (\partial\eta/\partial\bar{y}) + \sigma (\partial\eta/\partial\sigma)] \quad (12.40c)$$

$$= \hat{x} \bar{x} + \hat{y} \bar{y} + \hat{z} \xi^{\bar{a}} \partial_{\bar{a}}\eta. \quad (12.40d)$$

We identify the following properties as a means to help understand these expressions, with Figure 12.4 offering a schematic.

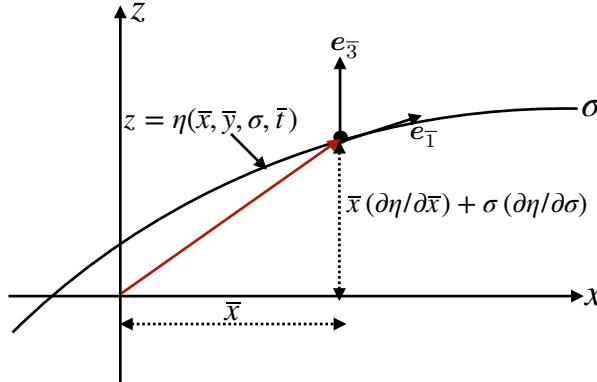


FIGURE 12.4: The position of a point in space as represented using a generalized vertical coordinate following equation (12.41a). For this example,  $\bar{y} = 0$  so that the horizontal position is determined by the coordinate  $\bar{x} = x$ , whereas the vertical position is determined by  $\bar{x}(\partial z/\partial\bar{x}) + \sigma(\partial z/\partial\sigma) = \bar{x}(\partial\eta/\partial\bar{x}) + \sigma(\partial\eta/\partial\sigma)$ .

- The expression (12.40b) has horizontal positions,  $\bar{x}$  and  $\bar{y}$ , multiplying the basis vectors,  $e_{\bar{x}}$  and  $e_{\bar{y}}$ , with these vectors tangent to a surface of constant generalized vertical coordinate ( $\sigma$ -surface) as in Figure 12.4. Likewise, the third term,  $\sigma(\partial\eta/\partial\sigma) \hat{z}$ , positions the point vertically according to the value of  $\sigma$ , along with its inverse stratification,  $\partial\eta/\partial\sigma = \partial z/\partial\sigma$ .
- Consider the case of  $y = \bar{y} = 0$  so that the position is given by

$$\mathcal{P} = \bar{x} \hat{x} + \hat{z} [\bar{x} (\partial\eta/\partial\bar{x}) + \sigma (\partial\eta/\partial\sigma)] \quad (12.41a)$$

$$= \bar{x} \hat{x} + \hat{z} (\partial\eta/\partial\sigma) [\bar{x} (\partial\sigma/\partial z)_x (\partial\eta/\partial\bar{x})_\sigma + \sigma] \quad (12.41b)$$

$$= \bar{x} \hat{x} + \hat{z} (\partial\eta/\partial\sigma) [-\bar{x} (\partial\sigma/\partial x)_z + \sigma], \quad (12.41c)$$

where we used the triple product identity (12.39b) for the final equality. Consequently, a horizontal position vector is realized using generalized vertical coordinates with  $\sigma = \bar{x}(\partial\sigma/\partial x)$ . That is, a horizontal position vector crosses surfaces of constant generalized vertical coordinate when the  $\sigma$ -surface has a nonzero horizontal slope.

- The projection of the position vector onto the basis one-forms leads to

$$\mathcal{P} \cdot e^{\bar{b}} = \xi^{\bar{a}} e_{\bar{a}} \cdot e^{\bar{b}} = \xi^{\bar{b}}. \quad (12.42)$$

This result follows from the bi-orthogonality relation (12.31). So the projection of the position vector onto a basis one-form picks out the corresponding coordinate value.

- Equation (12.7) provides the spatial dependence for the vertical position of the surface of constant generalized vertical coordinate

$$z = z(\xi^{\bar{a}}) = \eta(\xi^{\bar{a}}). \quad (12.43)$$

At any particular time instance we can perform a Taylor series about a reference geopotential  $z_0 = \eta_0$ , so that

$$\eta(\xi^{\bar{a}}) \approx \eta_0 + \xi^{\bar{a}} \partial_{\bar{a}} \eta. \quad (12.44)$$

We can thus write the position (12.40d) in the form

$$\mathcal{P} = \hat{x} \bar{x} + \hat{y} \bar{y} + \hat{z} [\eta - \eta_0]. \quad (12.45)$$

Taking the reference geopotential as  $\eta_0 = 0$  recovers the Cartesian expression. Since the position vector is a geometric object, it is reassuring that the generalized vertical coordinate representation is the same as the Cartesian representation; it is merely a reorganization of the basis vectors and corresponding coordinates.

## 12.8 Transforming components of a vector

Consider a vector field,  $\mathbf{F}$ , with Cartesian representation

$$\mathbf{F} = F^a e_a = F^x \hat{x} + F^y \hat{y} + F^z \hat{z}. \quad (12.46)$$

The corresponding generalized vertical coordinate components are related through the transformation matrix

$$F^{\bar{a}} = \Lambda^{\bar{a}}{}_a F^a. \quad (12.47)$$

Making use of the transformation matrix (12.25) yields the relations between GVC components and Cartesian components

$$F^{\bar{1}} = F^1 \quad \text{and} \quad F^{\bar{2}} = F^2 \quad \text{and} \quad F^{\bar{3}} = \nabla \sigma \cdot \mathbf{F}, \quad (12.48)$$

where we wrote

$$\nabla \sigma \cdot \mathbf{F} = (\partial \sigma / \partial x) F^1 + (\partial \sigma / \partial y) F^2 + (\partial \sigma / \partial z) F^3. \quad (12.49)$$

The vector field thus can be represented in generalized vertical coordinates as

$$\mathbf{F} = F^{\bar{a}} e_{\bar{a}} = F^1 e_{\bar{1}} + F^2 e_{\bar{2}} + (\nabla \sigma \cdot \mathbf{F}) e_{\bar{3}}. \quad (12.50)$$

Similarly, the covariant components transform as  $F_{\bar{a}} = \Lambda^a{}_{\bar{a}} F_a$ , where use of the inverse transformation matrix (12.25) renders

$$F_{\bar{1}} = F_1 + (\partial z / \partial \bar{x}) F_3 = F_1 + (\partial \eta / \partial \bar{x}) F_3 \quad (12.51a)$$

$$F_{\bar{2}} = F_2 + (\partial z / \partial \bar{y}) F_3 = F_2 + (\partial \eta / \partial \bar{y}) F_3 \quad (12.51b)$$

$$F_{\bar{3}} = (\partial z / \partial \sigma) F_3 = (\partial \eta / \partial \sigma) F_3, \quad (12.51c)$$

and the expression for the vector field

$$\mathbf{F} = F_{\bar{a}} \mathbf{e}^{\bar{a}} = [F_1 + (\partial \eta / \partial \bar{x}) F_3] \mathbf{e}^{\bar{1}} + [F_2 + (\partial \eta / \partial \bar{y}) F_3] \mathbf{e}^{\bar{2}} + (\partial \eta / \partial \sigma) F_3 \mathbf{e}^{\bar{3}}. \quad (12.52)$$

Recall also that for Cartesian coordinates the contravariant and covariant components to a vector are identical:  $F^a = F_a$ .

## 12.9 Representing the velocity vector

As an example of the results from Section 12.8, we here represent the velocity vector,  $\mathbf{v}$ , considering both covariant and contravariant representations. As for the position of a point in space as detailed in Section 12.7, we are assured that both the Cartesian and GVC representations lead to the same velocity vector since the velocity is an objective geometric object. In Section 12.9.4 we verify that the transformation formalism indeed respects this equivalence, with the GVC representation equivalent to the Cartesian representation

$$\mathbf{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}}. \quad (12.53)$$

### 12.9.1 Contravariant components

Following Section 12.8, we have the contravariant velocity components using generalized vertical coordinates

$$v^{\bar{1}} = u \quad \text{and} \quad v^{\bar{2}} = v \quad \text{and} \quad v^{\bar{3}} = v^a \partial_a \sigma = \mathbf{v} \cdot \nabla \sigma. \quad (12.54)$$

Use of the basis vectors (12.20a)-(12.20c) then leads to

$$\mathbf{v} = v^{\bar{a}} \mathbf{e}_{\bar{a}} \quad (12.55a)$$

$$= u \mathbf{e}_{\bar{x}} + v \mathbf{e}_{\bar{y}} + (v^a \partial_a \sigma) \mathbf{e}_\sigma \quad (12.55b)$$

$$= u [\hat{\mathbf{x}} + (\partial \eta / \partial \bar{x}) \hat{\mathbf{z}}] + v [\hat{\mathbf{y}} + (\partial \eta / \partial \bar{y}) \hat{\mathbf{z}}] + (\mathbf{v} \cdot \nabla \sigma) (\partial \eta / \partial \sigma) \hat{\mathbf{z}}. \quad (12.55c)$$

We emphasize that the contravariant representation of the horizontal velocity is the same for both geopotential vertical coordinates and generalized vertical coordinates. This fundamental property is often missed in the literature.

### 12.9.2 Covariant components

The covariant velocity components are given by

$$v_{\bar{1}} = u + (\partial \eta / \partial \bar{x}) w \quad \text{and} \quad v_{\bar{2}} = v + (\partial \eta / \partial \bar{y}) w \quad \text{and} \quad v_{\bar{3}} = (\partial \eta / \partial \sigma) w. \quad (12.56)$$

The one-form basis (12.27a)-(12.27c) thus leads to the velocity vector

$$\mathbf{v} = v_{\bar{a}} \mathbf{e}^{\bar{a}} = [u + (\partial \eta / \partial \bar{x}) w] \hat{\mathbf{x}} + [v + (\partial \eta / \partial \bar{y}) w] \hat{\mathbf{y}} + w (\partial \eta / \partial \sigma) \nabla \sigma. \quad (12.57)$$

### 12.9.3 Introducing the material time derivative

The material time derivagitive of the generalized vertical coordinate is

$$\frac{D\sigma}{Dt} = \frac{\partial\sigma}{\partial t} + \mathbf{v} \cdot \nabla\sigma = \dot{\sigma}, \quad (12.58)$$

with  $\dot{\sigma}$  symbolizing any process contributing to motion across  $\sigma$  isosurfaces (as fully explained in Section 13.4). Using the expression (12.58) in the velocity vector expression (12.55c) leads to

$$\mathbf{v} = u [\hat{\mathbf{x}} + (\partial\eta/\partial\bar{x}) \hat{\mathbf{z}}] + v [\hat{\mathbf{y}} + (\partial\eta/\partial\bar{y}) \hat{\mathbf{z}}] + (\mathbf{v} \cdot \nabla\sigma) (\partial\eta/\partial\sigma) \hat{\mathbf{z}} \quad (12.59a)$$

$$= u [\hat{\mathbf{x}} + (\partial\eta/\partial\bar{x}) \hat{\mathbf{z}}] + v [\hat{\mathbf{y}} + (\partial\eta/\partial\bar{y}) \hat{\mathbf{z}}] + (\dot{\sigma} - \partial\sigma/\partial t) (\partial\eta/\partial\sigma) \hat{\mathbf{z}} \quad (12.59b)$$

$$= u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + [\partial\eta/\partial\bar{t} + \mathbf{u} \cdot \nabla_{\text{h}} z + (\partial\eta/\partial\sigma)\dot{\sigma}] \hat{\mathbf{z}}, \quad (12.59c)$$

where the final equality made use of the triple product (12.38a):  $(\partial\sigma/\partial t)(\partial\eta/\partial\sigma) = -\partial\eta/\partial\bar{t}$ . In the steady state and in the absence of material changes to  $\sigma$ , the three dimensional flow lies within a surface of constant  $\sigma$ , whereby  $\mathbf{v} \cdot \nabla\sigma = 0$  and

$$\mathbf{v} = u [\hat{\mathbf{x}} + (\partial\eta/\partial\bar{x}) \hat{\mathbf{z}}] + v [\hat{\mathbf{y}} + (\partial\eta/\partial\bar{y}) \hat{\mathbf{z}}] \quad \text{if } \partial_t\sigma = 0 \text{ and } \dot{\sigma} = 0. \quad (12.60)$$

However, in general there are transient fluctuations and material changes so that  $\mathbf{v} \cdot \nabla\sigma \neq 0$ .

### 12.9.4 Equivalence to the Cartesian velocity representation

Use of the triple product identities (12.39b)-(12.39c) allows us to manipulate both expressions (12.55c) and (12.57) to recover the Cartesian expression

$$\mathbf{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}}. \quad (12.61)$$

Another way to see this identity is to note that in equation (12.59c), the vertical component is an expression for the material time derivative of the vertical position

$$w = \frac{Dz}{Dt} = \frac{\partial\eta}{\partial\bar{t}} + \mathbf{u} \cdot \nabla_{\text{h}}\eta + \frac{\partial\eta}{\partial\sigma}\dot{\sigma}. \quad (12.62)$$

We derive this identity in Section 13.5 where we discuss further kinematic results using generalized vertical coordinates.

## 12.10 Metric tensor using generalized vertical coordinates

The metric tensor provides the means to measure the distance between two points in space. One can represent the metric using any variety of coordinates, and here we examine its representation using generalized vertical coordinates. We are interested in two metric tensors, one for points living in three dimensional Euclidean space and the second for points restricted to a particular surface of constant generalized vertical coordinate.

### 12.10.1 Representation of the metric tensor and its inverse

For three dimensional Euclidean space, the generalized vertical coordinate representation of the metric tensor is given by

$$g_{\bar{a}\bar{b}} = \mathbf{e}_{\bar{a}} \cdot \mathbf{e}_{\bar{b}} = \begin{bmatrix} 1 + (\partial z / \partial \bar{x})^2 & (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) & (\partial z / \partial \bar{x})(\partial z / \partial \sigma) \\ (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) & 1 + (\partial z / \partial \bar{y})^2 & (\partial z / \partial \bar{y})(\partial z / \partial \sigma) \\ (\partial z / \partial \bar{x})(\partial z / \partial \sigma) & (\partial z / \partial \bar{y})(\partial z / \partial \sigma) & (\partial z / \partial \sigma)^2 \end{bmatrix}, \quad (12.63)$$

with the triple product identities (12.39b) and (12.39c) bringing the metric tensor into the form

$$g_{\bar{a}\bar{b}} = \begin{bmatrix} 1 + [(\partial \sigma / \partial x)(\partial z / \partial \sigma)]^2 & (\partial \sigma / \partial x)(\partial \sigma / \partial y)(\partial z / \partial \sigma)^2 & -(\partial \sigma / \partial x)(\partial z / \partial \sigma)^2 \\ (\partial \sigma / \partial x)(\partial \sigma / \partial y)(\partial z / \partial \sigma)^2 & 1 + [(\partial \sigma / \partial y)(\partial z / \partial \sigma)]^2 & -(\partial \sigma / \partial y)(\partial z / \partial \sigma)^2 \\ -(\partial \sigma / \partial x)(\partial z / \partial \sigma)^2 & -(\partial \sigma / \partial y)(\partial z / \partial \sigma)^2 & (\partial z / \partial \sigma)^2 \end{bmatrix}. \quad (12.64)$$

The representation of the inverse metric tensor is given by the somewhat simpler expression

$$g^{\bar{a}\bar{b}} = \mathbf{e}^{\bar{a}} \cdot \mathbf{e}^{\bar{b}} = \begin{bmatrix} 1 & 0 & \partial \sigma / \partial x \\ 0 & 1 & \partial \sigma / \partial y \\ \partial \sigma / \partial x & \partial \sigma / \partial y & |\nabla \sigma|^2 \end{bmatrix}, \quad (12.65)$$

where we used the basis one-forms given by equations (12.27a)-(12.27c). Proof that

$$g^{\bar{a}\bar{b}} g_{\bar{b}\bar{c}} = \delta^{\bar{a}}_{\bar{c}} \quad (12.66)$$

requires use of the triple product identities (12.39b) and (12.39c). Note that an additional means to derive the metric tensor (12.63) is given by writing the squared line element as<sup>4</sup>

$$ds^2 = dx^2 + dy^2 + dz^2 \quad (12.67a)$$

$$= dx^2 + dy^2 + [(\partial z / \partial \bar{x}) dx + (\partial z / \partial \bar{y}) dy + (\partial z / \partial \sigma) d\sigma]^2, \quad (12.67b)$$

from which the metric tensor (12.63) is revealed upon expanding the quadratic term and then rearranging.

### 12.10.2 Jacobian of transformation

The determinant of the generalized vertical coordinate representation of the three-dimensional metric tensor (12.63) is

$$\det(g_{\bar{a}\bar{b}}) = (\partial z / \partial \sigma)^2 = (\partial \eta / \partial \sigma)^2, \quad (12.68)$$

so that the Jacobian of transformation is the *specific thickness*

$$\frac{\partial(x, y, z)}{\partial(\bar{x}, \bar{y}, \sigma)} = \frac{\partial z}{\partial \sigma} = \frac{\partial \eta}{\partial \sigma}. \quad (12.69)$$

The coordinate transformation from Cartesian coordinates to generalized vertical coordinates is invertible only so long as the Jacobian remains nonzero and single-signed, meaning the fluid retains a monotonic vertical stratification of the  $\sigma$ -surfaces. The invertible relation between  $z$  and  $\sigma$  means that each point in the vertical can be uniquely specified by either of the two

---

<sup>4</sup>The traditional notation in physics writes the squared line element as  $ds^2 = (ds)^2$ . Likewise,  $dx^2 = (dx)^2$ , etc.

vertical coordinates. For example, using pressure as the generalized vertical coordinate in a hydrostatic fluid yields the Jacobian

$$\frac{\partial z}{\partial \sigma} = \frac{\partial z}{\partial p} = -\frac{1}{\rho g}, \quad (12.70)$$

which is indeed single-signed since the gravitational acceleration,  $g > 0$ , and the mass density is also positive,  $\rho > 0$ .

### 12.10.3 Covariant and contravariant representations

In addition to providing the means to measure distances between two points in space, the metric tensor allows us to convert between the covariant and contravariant representations of a tensor; i.e., to raise and lower tensor indices. For a vector we have the identity

$$F_{\bar{a}} = g_{\bar{a}\bar{b}} F^{\bar{b}}. \quad (12.71)$$

Through the metric tensor, we are afforded the ability to work with either the covariant or contravariant representation of a tensor as determined by convenience.

### 12.10.4 Metric tensor for points living on a constant $\sigma$ -surface

In some cases we are interested in describing motion restricted to a constant  $\sigma$ -surface, with points on this surface determined by

$$\sigma(x, y, z) = \sigma_0, \quad (12.72)$$

for some constant,  $\sigma_0$ . Taking the differential of this identity leads to

$$d\sigma = 0 = (\partial_a \sigma) dx^a. \quad (12.73)$$

Introducing the slope of the  $\sigma$ -surface relative to the horizontal plane (recall that  $\partial_z \sigma \neq 0$ )

$$S = \nabla_\sigma z = -\nabla_h \sigma / \partial_z \sigma, \quad (12.74)$$

allows us to eliminate the vertical increment via

$$dz = S_x dx + S_y dy. \quad (12.75)$$

This identity provides the vertical distance between two infinitesimally close points that live on a constant  $\sigma$ -surface.

To find the metric tensor we write the distance in three dimensional space between two infinitesimally close points that live a constant  $\sigma$ -surface

$$dx \cdot dx = dx^a \delta_{ab} dx^b \quad (12.76a)$$

$$= (dx)^2 + (dy)^2 + (S_x dx + S_y dy)^2 \quad (12.76b)$$

$$= (dx \quad dy) \begin{pmatrix} 1 + S_x^2 & S_x S_y \\ S_x S_y & 1 + S_y^2 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}, \quad (12.76c)$$

which then leads to the metric tensor

$$\tilde{g}_{ab} = \begin{pmatrix} 1 + S_x^2 & S_x S_y \\ S_x S_y & 1 + S_y^2 \end{pmatrix}, \quad (12.77)$$

where the tensor labels extend over  $a, b = 1, 2$ . It is straightforward to show that the inverse metric is given by

$$\tilde{g}^{ab} = \frac{1}{1 + S_x^2 + S_y^2} \begin{pmatrix} 1 + S_y^2 & -S_x S_y \\ -S_x S_y & 1 + S_x^2 \end{pmatrix}. \quad (12.78)$$

## 12.11 Volume element and the Levi-Civita tensor

As noted in Section 12.10.2, when represented using generalized vertical coordinates, the square root of the metric tensor determinant (12.63) provides the Jacobian of transformation

$$\sqrt{\det(g_{\bar{a}\bar{b}})} = \partial z / \partial \sigma = \partial \eta / \partial \sigma. \quad (12.79)$$

In turn, the volume element is

$$dV = dx dy dz = (\partial z / \partial \sigma) dx dy d\sigma. \quad (12.80)$$

The covariant Levi-Civita tensor has the representation using generalized vertical coordinates

$$\varepsilon_{\bar{a}\bar{b}\bar{c}} = (\partial z / \partial \sigma) \epsilon_{\bar{a}\bar{b}\bar{c}} \quad \varepsilon^{\bar{a}\bar{b}\bar{c}} = (\partial z / \partial \sigma)^{-1} \epsilon^{\bar{a}\bar{b}\bar{c}}, \quad (12.81)$$

where  $\epsilon$  is the permutation symbol with its components independent of coordinate representation.

## 12.12 Partial derivative operators

We here consider the partial derivative operators and their transformation between coordinate systems. These identities are used throughout GVC calculus. Given the importance of these expressions, we offer two derivations, with the geometric derivation in Section 12.12.2 requiring minimal use of tensor formalism.

### 12.12.1 Analytical derivation

The partial derivative operators in generalized vertical coordinates is determined via  $\partial_{\bar{a}} = \Lambda^a_{\bar{a}} \partial_a$ . Including also the time derivative leads to the relations

$$\partial_{\bar{t}} = \partial_t + (\partial z / \partial \bar{t}) \partial_z = \partial_t + (\partial \eta / \partial \bar{t}) \partial_z \quad (12.82a)$$

$$\partial_{\bar{x}} = \partial_x + (\partial z / \partial \bar{x}) \partial_z = \partial_x + (\partial \eta / \partial \bar{x}) \partial_z \quad (12.82b)$$

$$\partial_{\bar{y}} = \partial_y + (\partial z / \partial \bar{y}) \partial_z = \partial_y + (\partial \eta / \partial \bar{y}) \partial_z \quad (12.82c)$$

$$\partial_{\sigma} = (\partial z / \partial \sigma) \partial_z = (\partial \eta / \partial \sigma) \partial_z. \quad (12.82d)$$

We can make use of the triple product identities (12.39b) and (12.39c) to express the slope of a constant GVC surface in the equivalent manners

$$S = \nabla_{\bar{\sigma}} \eta = \nabla_{\bar{\sigma}} z = -(\partial \sigma / \partial z)^{-1} \nabla_z \sigma \quad (12.83)$$

where we introduced the standard shorthand notation

$$\nabla_{\sigma} = \hat{x} \partial/\partial \bar{x} + \hat{y} \partial/\partial \bar{y} \quad \text{and} \quad \nabla_h = \hat{x} \partial/\partial x + \hat{y} \partial/\partial y. \quad (12.84)$$

It is common to transform between the horizontal gradient operators, in which case we write

$$\nabla_{\sigma} = \nabla_h + (\nabla_{\sigma} z) \partial_z \equiv \nabla_h + \mathbf{S} \partial_z. \quad (12.85)$$

We emphasize that  $\nabla_{\sigma}$  is merely a shorthand for the two partial derivative operators and that it only has components in the horizontal directions. Furthermore, the  $\sigma$  subscript is not a tensor index.

### 12.12.2 Geometrical derivation

We now complement the previous analytical derivation of the lateral derivative operator by offering a geometric derivation. As we see, this operator is computed by taking the difference of a function along surfaces of constant generalized vertical coordinate, but with the lateral distance computed in the horizontal direction as shown in Figure 12.5. This particular feature of the horizontal derivative operator is a key aspect of the non-orthogonality property of generalized vertical coordinates.

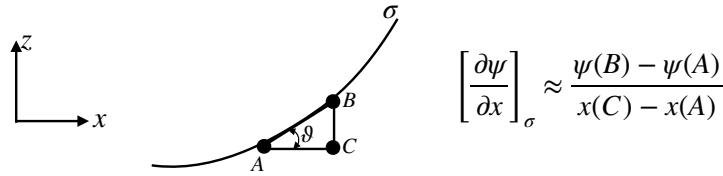


FIGURE 12.5: A surface of constant generalized vertical coordinate,  $\sigma$ , along with a local tangent plane with a slope  $\tan \vartheta$  with respect to the horizontal plane. This figure illustrates the identities (12.87a)-(12.87d), with these identities relating a lateral derivative taken along the GVC surface to horizontal and vertical derivatives taken along orthogonal Cartesian axes.

Consider the geometry shown in Figure 12.5, which shows a generalized vertical coordinate surface (constant  $\sigma$  surface) along with a sample tangent plane with a slope

$$S^x = \frac{\text{rise}}{\text{run}} = \tan \vartheta = \frac{z(B) - z(C)}{x(C) - x(A)} \approx \left[ \frac{\partial z}{\partial x} \right]_{\sigma} = - \frac{(\partial \sigma / \partial x)_z}{(\partial \sigma / \partial z)} \quad (12.86)$$

relative to the horizontal. We readily verify the following identities based on finite difference operations for an arbitrary function

$$\left[ \frac{\partial \psi}{\partial x} \right]_{\sigma} \approx \frac{\psi(B) - \psi(A)}{x(C) - x(A)} \quad (12.87a)$$

$$= \frac{\psi(C) - \psi(A)}{x(C) - x(A)} + \frac{\psi(B) - \psi(C)}{x(C) - x(A)} \quad (12.87b)$$

$$= \frac{\psi(C) - \psi(A)}{x(C) - x(A)} + \left[ \frac{z(B) - z(C)}{x(C) - x(A)} \right] \frac{\psi(B) - \psi(C)}{z(B) - z(C)} \quad (12.87c)$$

$$= \left[ \frac{\partial \psi}{\partial x} \right]_z + S^x \left[ \frac{\partial \psi}{\partial z} \right]_x. \quad (12.87d)$$

Taking the continuum limit then leads to the relations between horizontal derivatives computed

on constant  $\sigma$  surfaces to those computed on constant  $z$  surfaces

$$\left[ \frac{\partial}{\partial x} \right]_{\sigma} = \left[ \frac{\partial}{\partial x} \right]_z + \left[ \frac{\partial z}{\partial x} \right]_{\sigma} \frac{\partial}{\partial z} = \left[ \frac{\partial}{\partial x} \right]_z + \left[ \frac{\partial \eta}{\partial x} \right]_{\sigma} \frac{\partial}{\partial z} \quad (12.88a)$$

$$\left[ \frac{\partial}{\partial y} \right]_{\sigma} = \left[ \frac{\partial}{\partial y} \right]_z + \left[ \frac{\partial z}{\partial y} \right]_{\sigma} \frac{\partial}{\partial z} = \left[ \frac{\partial}{\partial y} \right]_z + \left[ \frac{\partial \eta}{\partial y} \right]_{\sigma} \frac{\partial}{\partial z}, \quad (12.88b)$$

which can be written in the shorthand vector notation

$$\nabla_{\sigma} = \hat{x} \left[ \frac{\partial}{\partial x} \right]_{\sigma} + \hat{y} \left[ \frac{\partial}{\partial y} \right]_{\sigma} = \nabla_z + (\nabla_{\sigma} z) \partial_z = \nabla_z + (\nabla_{\sigma} \eta) \partial_z. \quad (12.89)$$

### 12.12.3 The gradient as a tensor operator

The gradient is given by the equivalent expressions

$$\nabla = e^a \partial_a = e^{\bar{a}} \partial_{\bar{a}}. \quad (12.90)$$

The gradient has the following Cartesian coordinate expression

$$\nabla = \hat{x} \partial_x + \hat{y} \partial_y + \hat{z} \partial_z, \quad (12.91)$$

and the equivalent generalized vertical coordinate expression

$$\nabla = \hat{x} \partial_{\bar{x}} + \hat{y} \partial_{\bar{y}} + (\nabla \sigma) \partial_{\sigma}. \quad (12.92)$$

In Exercise 12.3 we verify the equality between equations (12.91) and (12.92) by making use of equations (12.82b)-(12.82d) for the partial derivatives and equations (12.27a)-(12.27c) for the one-form basis.

## 12.13 Material time derivative

Making use of the relations for the partial derivative operators in Section 12.12 allows us to write the material time derivative in the following equivalent forms

$$\frac{D}{Dt} = \left[ \frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (12.93a)$$

$$= \left[ \frac{\partial}{\partial t} \right]_{\sigma} - (\partial \eta / \partial \bar{t}) \partial_z + \mathbf{u} \cdot [\nabla_{\sigma} - (\nabla_{\sigma} \eta) \partial_z] + w \partial / \partial z \quad (12.93b)$$

$$= \left[ \frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + [w - \mathbf{u} \cdot \nabla_{\sigma} \eta - \partial \eta / \partial \bar{t}] (\partial \sigma / \partial z) \partial / \partial \sigma \quad (12.93c)$$

$$= \left[ \frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + \frac{D\sigma}{Dt} \frac{\partial}{\partial \sigma} \quad (12.93d)$$

$$= \left[ \frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} \frac{\partial}{\partial z}. \quad (12.93e)$$

The equality (12.93d) made use of the identity (12.62), which is itself derived in Section 13.5 where we discuss further kinematic results using generalized vertical coordinates. Besides differences in the spatial operators, it is important to note that the time derivative operators are computed on constant geopotential surfaces and constant  $\sigma$ -surfaces, respectively. However,

the horizontal velocity component is the *same* for both forms of the material time derivative

$$(u, v) = \frac{D(x, y)}{Dt}. \quad (12.94)$$

## 12.14 Divergence of a vector and the divergence theorem

Making use of the general expression from VOLUME 1 for the covariant divergence of a vector renders the generalized vertical coordinate expression

$$\nabla_{\bar{a}} F^{\bar{a}} = [\det(g_{\bar{a}\bar{b}})]^{-1/2} \partial_{\bar{a}} \left[ [\det(g_{\bar{a}\bar{b}})]^{1/2} F^{\bar{a}} \right] = (\partial z / \partial \sigma)^{-1} \partial_{\bar{a}} [(\partial z / \partial \sigma) F^{\bar{a}}]. \quad (12.95)$$

Recall that the GVC vector components,  $F^{\bar{a}}$ , are related to the Cartesian components in equation (12.48), and the GVC components of the partial derivative operator,  $\partial_{\bar{a}}$ , are related to the Cartesian operator in equation (12.84).

When making use of the divergence theorem, we require the product of the volume element and the covariant divergence. For generalized vertical coordinates we make use of the volume element (12.80) so that

$$(\nabla_{\bar{a}} F^{\bar{a}}) dV = \partial_{\bar{a}} [(\partial z / \partial \sigma) F^{\bar{a}}] d\bar{x} d\bar{y} d\sigma, \quad (12.96)$$

which reduces to a boundary integral when integrating over a volume.

## 12.15 The diffusion operator

As an explicit example of the covariant divergence operator (12.95), we here consider the diffusion operator discussed in Chapter 16. The derivation here recovers much of what we just discussed in Section 12.14, yet we make use of a bit less tensor formalism though at the cost of more algebra.

### 12.15.1 Continuous expression

The diffusion operator is the convergence of the diffusive flux

$$\mathcal{R} = -\nabla \cdot \mathbf{J}, \quad (12.97)$$

where  $\mathbf{J}$  is the tracer flux vector. We here convert the pieces of this operator from Cartesian coordinates into generalized vertical coordinates, making use of the transformation of partial derivative operators given in Section 12.12. Also, we make use of the shorthand  $z(x, y, \sigma, t)$  rather than  $\eta(x, y, \sigma, t)$

$$-\mathcal{R} = \nabla \cdot \mathbf{J} \quad (12.98a)$$

$$= \nabla_h \cdot \mathbf{J}^h + \partial_z J^z \quad (12.98b)$$

$$= (\nabla_{\bar{\sigma}} - \nabla_{\bar{\sigma}} z \partial_z) \cdot \mathbf{J}^h + (\sigma_z) \partial_{\sigma} J^z \quad (12.98c)$$

$$= \sigma_z [z_{\sigma} \nabla_{\bar{\sigma}} \cdot \mathbf{J}^h + (\hat{z} \partial_{\sigma} - \nabla_{\bar{\sigma}} z \partial_{\sigma}) \cdot \mathbf{J}] \quad (12.98d)$$

$$= \sigma_z [\nabla_{\bar{\sigma}} \cdot (z_{\sigma} \mathbf{J}^h) - \mathbf{J}^h \cdot \nabla_{\bar{\sigma}} (z_{\sigma}) + \partial_{\sigma} J^z - \partial_{\sigma} (\nabla_{\bar{\sigma}} z \cdot \mathbf{J}) + \mathbf{J} \cdot \partial_{\sigma} (\nabla_{\bar{\sigma}} z)] \quad (12.98e)$$

$$= \sigma_z [\nabla_{\bar{\sigma}} \cdot (z_{\sigma} \mathbf{J}^h) + \partial_{\sigma} J^z - \partial_{\sigma} (\nabla_{\bar{\sigma}} z \cdot \mathbf{J}^h)] \quad (12.98f)$$

$$= \sigma_z (\nabla_{\bar{\sigma}} \cdot (\partial_{\sigma} z \mathbf{J}^h) + \partial_{\sigma} [(\hat{z} - \nabla_{\bar{\sigma}} z) \cdot \mathbf{J}]) \quad (12.98g)$$

$$= \sigma_z [\nabla_{\mathbf{h}} \cdot (z_\sigma \mathbf{J}^h) + \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J})], \quad (12.98h)$$

where we used

$$z_\sigma \nabla \sigma = \hat{\mathbf{z}} - \nabla_{\mathbf{h}} z \quad (12.99)$$

to reach the final equality, and made use of the shorthand

$$z_\sigma = \partial z / \partial \sigma \quad \text{and} \quad \sigma_z = \partial \sigma / \partial z = (z_\sigma)^{-1}. \quad (12.100)$$

The coordinate transformations in Section 12.8 for vector components reveal that the expression (12.98h) is identical to equation (12.95) derived using formal tensor methods. Likewise, multiplying by the volume element

$$dV = dx dy dz = dx dy z_\sigma d\sigma, \quad (12.101)$$

leads to

$$-\mathcal{R} dV = [\nabla_{\mathbf{h}} \cdot (z_\sigma \mathbf{J}^h) + \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J})] dx dy d\sigma, \quad (12.102)$$

which is identical to the expression (12.96).

## 12.15.2 Layer thickness weighted diffusion operator

Consider a prescribed increment,  $\delta\sigma$ , separating two  $\sigma$  isosurfaces. This increment commutes with the horizontal operator  $\nabla_{\mathbf{h}}$ , acting within the layer. We can thus formally consider the following layer-integrated or thickness weighted form of the diffusion operator

$$-\mathcal{R} \delta V = [\nabla_{\mathbf{h}} \cdot (\delta\sigma z_\sigma \mathbf{J}^h) + \delta\sigma \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J})] \delta x \delta y \quad (12.103a)$$

$$= \frac{1}{\delta z} [\nabla_{\mathbf{h}} \cdot (\delta\sigma z_\sigma \mathbf{J}^h) + \delta\sigma \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J})] \delta x \delta y \delta z \quad (12.103b)$$

$$= \frac{1}{h} [\nabla_{\mathbf{h}} \cdot (h \mathbf{J}^h) + \Delta_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J})] \delta x \delta y h, \quad (12.103c)$$

where we introduced the infinitesimal layer thickness

$$h = \delta z = z_\sigma \delta\sigma \quad (12.104)$$

and the non-dimensional differential operator

$$\Delta_\sigma \equiv \delta\sigma \frac{\partial}{\partial \sigma}. \quad (12.105)$$

Cancelling the volume element on both sides leads to the diffusion operator

$$\mathcal{R} = -\frac{1}{h} [\nabla_{\mathbf{h}} \cdot (h \mathbf{J}^h) + \Delta_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J})]. \quad (12.106)$$

This form is commonly found in the numerical modeling literature when considering generalized vertical coordinate models.

We make the following comments concerning the diffusion operator in equation (12.106).

- Our introduction of the layer thickness  $h = z_\sigma \delta\sigma$  is treated a bit more formally in Sections 13.10 and 13.11 by considering a vertical integral over a coordinate layer. Even so, the resulting diffusion operator is the same as that derived here.

- The thickness weighted flux,  $h \mathbf{J}^h$ , is oriented within the horizontal plane. However, its contribution to the diffusion operator is computed by taking its convergence using the operator  $\nabla_{\sigma}$  rather than the horizontal operator  $\nabla_h$ . This distinction is fundamental to how operators, such as advection and diffusion, appear using generalized vertical coordinates.
- The flux  $z_\sigma \nabla \sigma \cdot \mathbf{J}$  is commonly referred to as the dia-surface subgrid scale flux.
- For the special case of a diffusive flux with zero component parallel to  $\nabla \sigma$ , the diffusion operator reduces to

$$\mathcal{R} = -\frac{1}{h} [\nabla_{\sigma} \cdot (h \mathbf{J}^h)] \quad \text{if } \nabla \sigma \cdot \mathbf{J} = 0. \quad (12.107)$$

The neutral diffusion operator of Section 18.4.4 is an example of such an operator, with  $\sigma$  in that case given by the locally referenced potential density.

## 12.16 Vorticity

Vorticity is the curl of the velocity

$$\omega = \text{curl}(\mathbf{v}) = \mathbf{e}_a \varepsilon^{abc} \partial_b v_c = \mathbf{e}_{\bar{a}} \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}}, \quad (12.108)$$

where we made use of the expression for the curl from VOLUME 1. In this section we unpack this expression to write the vorticity using generalized vertical coordinates.

### 12.16.1 Contravariant components

The contravariant components of the vorticity are given by via

$$\omega^{\bar{a}} = \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}} = (\partial z / \partial \sigma)^{-1} \epsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}}, \quad (12.109)$$

where we made use of equation (12.81) to introduce the permutation symbol. Expanding the components leads to

$$\omega^{\bar{1}} = (\partial \sigma / \partial z) (\partial_{\bar{2}} v_{\bar{3}} - \partial_{\bar{3}} v_{\bar{2}}) \quad (12.110a)$$

$$\omega^{\bar{2}} = (\partial \sigma / \partial z) (\partial_{\bar{3}} v_{\bar{1}} - \partial_{\bar{1}} v_{\bar{3}}) \quad (12.110b)$$

$$\omega^{\bar{3}} = \omega^\sigma = (\partial \sigma / \partial z) (\partial_{\bar{1}} v_{\bar{2}} - \partial_{\bar{2}} v_{\bar{1}}), \quad (12.110c)$$

so that the curl is given by

$$\omega = \mathbf{e}_{\bar{a}} \omega^{\bar{a}} = (\partial \sigma / \partial z) [\mathbf{e}_{\bar{1}} (\partial_{\bar{2}} v_{\bar{3}} - \partial_{\bar{3}} v_{\bar{2}}) + \mathbf{e}_{\bar{2}} (\partial_{\bar{3}} v_{\bar{1}} - \partial_{\bar{1}} v_{\bar{3}}) + \mathbf{e}_{\bar{3}} (\partial_{\bar{1}} v_{\bar{2}} - \partial_{\bar{2}} v_{\bar{1}})], \quad (12.111)$$

where equations (12.20a)-(12.20c) provide expressions for the basis vectors written using generalized vertical coordinates.

### 12.16.2 Transforming from Cartesian coordinates

Equation (12.111) is written solely with the generalized vertical coordinates. An alternative approach connects performs a transformation from the Cartesian vorticity components using the transformation matrix

$$\omega^{\bar{a}} = \Lambda^{\bar{a}}_a \omega^a, \quad (12.112)$$

where  $\omega^a$  are the Cartesian components

$$\omega = \hat{x} \left[ \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right] + \hat{y} \left[ \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right] + \hat{z} \left[ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right]. \quad (12.113)$$

Making use of the transformation matrix  $\Lambda^{\bar{a}}_a$  from equation (12.25) yields (just as in Section 12.8)

$$\omega^{\bar{x}} = \omega^x = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \quad \text{and} \quad \omega^{\bar{y}} = \omega^y = \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \quad \text{and} \quad \omega^\sigma = \boldsymbol{\omega} \cdot \nabla \sigma. \quad (12.114)$$

Note that for isopycnal coordinates in a Boussinesq fluid,  $\omega^\sigma$  equals to the potential vorticity when the vorticity is the absolute vorticity (Section 14.3). That is, the potential vorticity is the isopycnal component of the absolute vorticity.

## 12.17 Circulation

The velocity **circulation** is given by the closed oriented path integral of the velocity projected into the direction of the path

$$\mathcal{C} \equiv \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{x} \quad (12.115)$$

where  $d\mathbf{x}$  is the vector line element along the path and  $\partial\mathcal{S}$  is the closed path defining the boundary to a two-dimensional surface,  $\mathcal{S}$ . **Stokes' theorem** leads to the identity

$$\mathcal{C} = \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{x} = \int_{\mathcal{S}} (\nabla \times \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (12.116)$$

where  $\hat{\mathbf{n}}$  is the outward normal vector orienting the area element  $d\mathcal{S}$  according to the right-hand rule applied to the bounding circuit. These results manifest Galilean space-time **covariance**, so that they hold for an arbitrary coordinate representation.

As a particular case, consider the circulation around a closed path on a constant  $\sigma$  surface, in which

$$\hat{\mathbf{n}} = \frac{\nabla \sigma}{|\nabla \sigma|} \quad (12.117)$$

is the outward normal and

$$\boldsymbol{\omega} \cdot \hat{\mathbf{n}} = \frac{\omega^\sigma}{|\nabla \sigma|} \quad (12.118)$$

where  $\omega^\sigma = \boldsymbol{\omega} \cdot \nabla \sigma$  (equation (12.114)). So long as the vertical stratification remains non-zero ( $\partial \sigma / \partial z \neq 0$ ) we can write the area factor in the form

$$\frac{d\mathcal{S}}{|\nabla \sigma|} = \frac{d\mathcal{S}}{\sqrt{(\partial \sigma / \partial x)^2 + (\partial \sigma / \partial y)^2 + (\partial \sigma / \partial z)^2}} \quad (12.119a)$$

$$= \frac{d\mathcal{S}}{|\partial \sigma / \partial z| \sqrt{[(\partial \sigma / \partial x)/(\partial \sigma / \partial z)]^2 + [(\partial \sigma / \partial y)/(\partial \sigma / \partial z)]^2 + 1}} \quad (12.119b)$$

$$= \frac{d\mathcal{S}}{|\partial \sigma / \partial z| \sqrt{1 + \tan^2 \vartheta}} \quad (12.119c)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| |\cos \vartheta| d\mathcal{S} \quad (12.119d)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| dA. \quad (12.119e)$$

The equality (12.119c) introduces the angle,  $\vartheta$ , between the boundary surface and the horizontal plane as in Figure 12.5. The squared slope of this surface given by

$$\tan^2 \vartheta = \frac{\nabla_h \sigma \cdot \nabla_h \sigma}{(\partial \sigma / \partial z)^2} = \nabla_{\text{fr}} z \cdot \nabla_{\text{fr}} z. \quad (12.120)$$

The equality (12.119d) made use of a trigonometric identity, and the equality (12.119e) introduced the horizontal projection of the area,

$$dA = |\cos \vartheta| dS. \quad (12.121)$$

Bringing these results together leads to the expression for circulation around a closed loop on a constant  $\sigma$  surface

$$\mathcal{C}_{\sigma-\text{surface}} = \int_S (\omega \cdot \nabla \sigma) |\partial z / \partial \sigma| dA. \quad (12.122)$$



## 12.18 Exercises

### EXERCISE 12.1: TRIPLE PRODUCT AND TRANSFORMATION MATRIX

Use the triple product identities (12.39b)-(12.39c) to verify that equation (12.71) agrees with the transformation matrix approach detailed in Section 12.8. It is sufficient to show the agreement with just  $F_{\bar{1}}$ .

### EXERCISE 12.2: VECTOR CROSS PRODUCT OF BASIS VECTORS

Verify the following relation (VOLUME 1) for the cross product of two basis vectors using generalized vertical coordinates

$$\mathbf{e}_{\bar{a}} \times \mathbf{e}_{\bar{b}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \mathbf{e}^{\bar{c}} \implies \mathbf{e}_{\bar{a}} \times \mathbf{e}_{\bar{b}} = (\partial z / \partial \sigma) \epsilon_{\bar{a}\bar{b}\bar{c}} \mathbf{e}^{\bar{c}}. \quad (12.123)$$

### EXERCISE 12.3: CHECKING THE REPRESENTATIONS OF THE GRADIENT OPERATOR

Confirm that the two Cartesian expression (12.91) for the gradient operator is indeed identical to the generalized vertical coordinate representation (12.92). Hint: make use of equations (12.82b)-(12.82d) for the partial derivatives and equations (12.27a)-(12.27c) for the one-form basis. Hint: also make use of the triple product identities from Section 12.10.3.s

### EXERCISE 12.4: CHECKING THE GVC REPRESENTATION OF THE METRIC TENSOR

We here verify the identity (12.66) through use of the generalized vertical coordinate representation of the metric tensor and its inverse as given in Section 12.10.1.

- (a) Show that  $g^{\bar{1}\bar{b}} g_{\bar{b}\bar{1}} = 1$ .
- (b) Show that  $g^{\bar{1}\bar{b}} g_{\bar{b}\bar{2}} = 0$ .

### EXERCISE 12.5: TRANSFORMATION OF THE METRIC TENSOR

We computed the inverse metric tensor in equation (12.65) through the identity

$$g^{\bar{a}\bar{b}} = \mathbf{e}^{\bar{a}} \cdot \mathbf{e}^{\bar{b}}, \quad (12.124)$$

and used the basis one-forms for generalized vertical coordinates given by equations (12.27a)-(12.27c). Here we make use of the transformation matrix and verify the identity

$$g^{\bar{a}\bar{b}} = (\Lambda^{\bar{a}}_a \mathbf{e}^a) \cdot (\Lambda^{\bar{b}}_b \mathbf{e}^b), \quad (12.125)$$

where  $\Lambda^{\bar{a}}_a$  is the inverse transformation matrix (12.25) and  $\mathbf{e}^b$  are the Cartesian basis one-forms.

- (a) Show that  $g^{\bar{1}\bar{1}} = 1$ .
- (b) Show that  $g^{\bar{1}\bar{3}} = \partial\sigma/\partial x$ .

#### EXERCISE 12.6: METRIC TENSOR FOR RIGID LID TERRAIN FOLLOWING COORDINATES

Consider the specific case of rigid lid terrain following coordinates from Section 12.3.2, for which the generalized vertical coordinate is  $\sigma = -z/\eta_b$ . Hint: see Appendix B from [Callies and Ferrari \(2018\)](#).

- (a) Write the metric tensor (12.64).
- (b) Write the inverse metric tensor (12.65).

#### EXERCISE 12.7: RIGIDLY ROTATED EQUATIONS FOR PLANETARY GEOSTROPHY

On page 404 we identified a key reason to favor generalized vertical coordinates, which are non-orthogonal, rather than rigidly rotated orthogonal coordinates. Namely, the problem with rotated orthogonal coordinates is that the hydrostatic pressure gradient appears in more than a single component equation of motion. Here we illustrate the mathematics of this rotation by considering the unforced perfect fluid planetary geostrophic equations from VOLUME 2. In fact, our only concern is with the velocity equation, here written as

$$\rho_0 f (\hat{\mathbf{z}} \times \mathbf{u}) = -\nabla p' - \rho' g \hat{\mathbf{z}}, \quad (12.126)$$

where

$$p' = p - p_0 \quad \text{and} \quad \rho' = \rho - \rho_0 \quad \text{with} \quad dp_0/dz = -\rho_0 g \quad \text{and} \quad \partial p'/\partial z = -\rho' g. \quad (12.127)$$

Our goal in this exercise is to examine the planetary geostrophic momentum equation after performing a rigid coordinate rotation as in Figure 12.6. That is, we perform a rotation of the Cartesian coordinate system in a counter-clockwise direction through an angle,  $\varphi$ , relative to the  $\hat{\mathbf{x}}$  axis.

A version of the rotated equations (with friction included) are sometimes used to study bottom boundary layer physics next to a sloped bottom, such as [Phillips \(1970\)](#), [Garrett et al. \(1993\)](#), [Callies \(2018\)](#), [Holmes et al. \(2019a\)](#), and [Wenegrat and Thomas \(2020\)](#). Notably, these studies generally drop certain terms in the rotated equations based on specifics of the dynamical regime under consideration. For this exercise we keep all terms arising from the rotation since the goal here is mathematical. That is, we wish to see the full rotated equation set so that it can be examined prior to making any dynamical assumptions that target a specific application.

- (a) Write the basis vectors,  $\mathbf{e}_{\bar{a}}$ , for the rotated Cartesian coordinate system in terms of the unrotated basis vectors,  $\mathbf{e}_a = (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ .
- (b) Write the transformation matrix,  $\Lambda^a_{\bar{a}}$ , and inverse transformation matrix,  $\Lambda^{\bar{a}}_a$ , that give rise to this rotation.
- (c) Write the representation of the velocity vector using the rotated basis, and express this representation in terms of the unrotated velocity components,  $v^a = (u, v, w)$ . That is, compute  $v^{\bar{a}} = \Lambda^{\bar{a}}_a v^a$ .
- (d) Write the unrotated basis vectors,  $\mathbf{e}_a$ , in terms of the rotated basis vectors,  $\mathbf{e}_{\bar{a}}$ , according

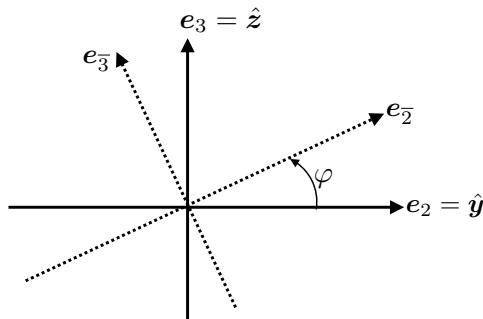


FIGURE 12.6: A schematic for Exercise 12.7 to illustrate the rigid rotation of the Cartesian coordinate system in an counter-clockwise sense around the  $\hat{x}$  direction (which points out of the page) through an angle  $\varphi$ . The unrotated basis vectors are written  $e_a = (\hat{x}, \hat{y}, \hat{z})$ , and the rotated basis vectors are  $e_{\bar{a}}$ .

to the transformation  $e_a = \Lambda^{\bar{a}}_a e_{\bar{a}}$ .

- (e) Write the representation of the velocity vector using the unrotated basis, and express this representation in terms of the rotated velocity components,  $v^{\bar{a}}$ . That is, compute  $v^a = \Lambda^a_{\bar{a}} v^{\bar{a}}$ .
- (f) Write  $\hat{z} \times \mathbf{u}$  using the rotated coordinates.
- (g) Write  $\nabla p'$  using the rotated coordinates.
- (h) Write the equation of motion (12.126) in the rotated coordinate system.
- (i) Comment on the potential problems with this approach, as relates to the form of the hydrostatic balance.



# Chapter 13

## MECHANICAL EQUATIONS

In this chapter we develop the kinematical and dynamical equations for a moving geophysical fluid as formulated using generalized vertical coordinates. The key part of the kinematic formulation concerns the *dia-surface*, which measures the motion of fluid through a surface. The discussion here unifies ideas developed for [kinematic boundary conditions](#) with transport across an arbitrary surface in the fluid interior, and we do so within the context of a [GVC \(GVC\)](#), whose mathematics was studied in Chapter 12. In particular, we make use of the dia-surface transport formulation to express the material time derivative operator using generalized vertical coordinates. This form for the material time derivative allows us to decompose the vertical velocity into motion relative to a moving [GVC](#) surface. In turn, we are afforded a means to reinterpret the velocity vector and corresponding particle trajectories.

Once we have explored the kinematic description of motion using generalized vertical coordinates, we then work through derivations of the [GVC](#) version of mass continuity as well as the tracer equation, and derive their layer integrated versions appropriate for vertically discrete layered models. Thereafter, we derive the dynamical equations for momentum, vorticity, and potential vorticity using generalized vertical coordinates. The suite of mechanical equations provide the foundations for many numerical models of the atmosphere and ocean. Besides being essential for developing methods for numerical simulations, understanding the physical and mathematical basis of these equations supports the analysis of simulations.

### CHAPTER GUIDE

We introduced mathematical properties of a [GVC \(GVC\)](#) in Chapter 12, including the calculus using these time dependent non-orthogonal coordinates. Following the treatment in Chapter 12, we here use the symbol  $\sigma$  to denote a generalized vertical coordinate, where  $\sigma$  has functional dependence  $\sigma(x, y, z, t)$ .

<b>13.1</b>	<b>Loose threads</b>	<b>434</b>
<b>13.2</b>	<b>Example generalized vertical coordinate surfaces</b>	<b>435</b>
13.2.1	Ocean free surface	435
13.2.2	Solid earth boundary	435
13.2.3	Ocean mixed layer base	436
13.2.4	Interior generalized vertical coordinate surfaces	436
<b>13.3</b>	<b>Specific thickness</b>	<b>436</b>
<b>13.4</b>	<b>The dia-surface transport</b>	<b>438</b>
13.4.1	Flow normal to the GVC surface	438
13.4.2	Accounting for movement of the surface	438

13.4.3	We only care about divergent surface motion . . . . .	439
13.4.4	Dia-surface transport in terms of $D\sigma/Dt$ . . . . .	440
13.4.5	Defining the dia-surface transport . . . . .	440
13.4.6	Expressions for the dia-surface velocity component . . . . .	441
13.4.7	Relating $w$ and $w^{(\delta)}$ . . . . .	442
13.4.8	Dia-surface velocity component for water mass analysis . . . . .	442
13.4.9	Area integrated dia-surface transport for non-divergent flows . . . . .	442
13.5	Material time derivative . . . . .	444
13.6	Vertical velocity and dia-surface velocity . . . . .	445
13.6.1	Decomposing the vertical velocity . . . . .	445
13.6.2	Another form of the vertical velocity decomposition . . . . .	445
13.7	The velocity vector and fluid particle trajectories . . . . .	447
13.7.1	Coordinate representations of the velocity vector . . . . .	447
13.7.2	Fluid particle trajectories . . . . .	448
13.8	Subduction across the mixed layer base . . . . .	449
13.9	Mass continuity . . . . .	450
13.9.1	Cartesian coordinates . . . . .	450
13.9.2	Generalized vertical coordinates . . . . .	451
13.10	Layer integrated mass continuity . . . . .	452
13.10.1	Non-Boussinesq (compressible) fluids . . . . .	452
13.10.2	Mass continuity using pressure coordinates . . . . .	454
13.10.3	Boussinesq ocean (non-divergent flow) . . . . .	455
13.10.4	Rescaled geopotential coordinates . . . . .	455
13.11	Layer integrated tracer equation . . . . .	456
13.12	OVERTURNING CIRCULATION IN THE MERIDIONAL- $\sigma$ PLANE . . . . .	457
13.12.1	OVERTURNING STREAMFUNCTION . . . . .	457
13.12.2	PROVING THAT $\Psi(y, \sigma, t)$ IS A STREAMFUNCTION . . . . .	458
13.13	EQUATIONS OF MOTION . . . . .	459
13.13.1	MASS AND TRACER EQUATIONS . . . . .	460
13.13.2	MOMENTUM EQUATION . . . . .	460
13.13.3	FLUX-FORM HORIZONTAL MOMENTUM EQUATION . . . . .	461
13.13.4	VECTOR-INVARIANT HORIZONTAL MOMENTUM EQUATION . . . . .	461
13.13.5	HYDROSTATIC FLOW WITH CONSTANT GRAVITATIONAL ACCELERATION . . . . .	462
13.14	CONCERNING THE PRESSURE FORCE . . . . .	462
13.14.1	COMPUTING THE HORIZONTAL PRESSURE GRADIENT . . . . .	463
13.14.2	INTEGRATED PRESSURE CONTACT STRESS ON THE CELL FACES . . . . .	464
13.14.3	NET VERTICAL PRESSURE FORCE . . . . .	465
13.14.4	NET HORIZONTAL PRESSURE FORCE . . . . .	465
13.14.5	COMMENTS . . . . .	466
13.15	HYDROSTATIC VORTICITY AND POTENTIAL VORTICITY . . . . .	466
13.15.1	BASIC MANIPULATIONS . . . . .	466
13.15.2	VORTICITY AND POTENTIAL VORTICITY EQUATION . . . . .	467
13.15.3	POTENTIAL VORTICITY FOR THE BOUSSINESQ OCEAN . . . . .	468
13.15.4	COMMENTS . . . . .	469
13.16	EXERCISES . . . . .	469

---

## 13.1 Loose threads

- Compare to the slope orthogonal coordinates of [Holmes et al. \(2019a\)](#).

## 13.2 Example generalized vertical coordinate surfaces

In Section 12.3 we provided some specific examples of a generalized vertical coordinates used to partition the ocean domain. Here we offer further examples as illustrated in Figure 13.1, focusing on aspects of specific coordinate surfaces that prove of use for the kinematics of this chapter.

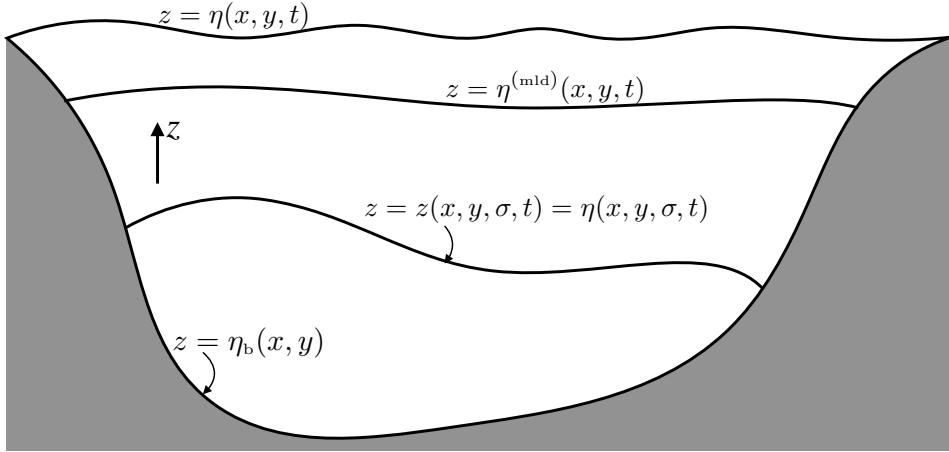


FIGURE 13.1: Example surfaces of a constant generalized vertical coordinate,  $\sigma(x, y, z, t)$ , in an ocean basin. Each of these surfaces is assumed to be smooth and to not overturn. The ocean free surface can be represented mathematically by  $\sigma(x, y, z, t) = z - \eta(x, y, t) = 0$ ; the ocean mixed layer base can be written  $\sigma(x, y, z, t) = z - \eta^{(mld)}(x, y, t) = 0$ ; and the solid earth bottom  $\sigma(x, y, z) = z - \eta_b(x, y) = 0$ . Likewise, the vertical position of an interior generalized vertical coordinate surface can be written  $z - \eta(x, y, \sigma, t) = \text{constant}$ , where  $\eta(x, y, \sigma, t)$  is a function of horizontal position and time for the surface defined by a particular value of the generalized vertical coordinate.

### 13.2.1 Ocean free surface

We considered the **kinematic boundary condition** at the ocean free surface in VOLUME 2, across which water and tracer can cross via precipitation, evaporation, river runoff, and sea ice melt. Momentum exchange across the boundary arises from stresses between the ocean and atmosphere or ice. The ocean free surface can be represented mathematically by the identity

$$\sigma(x, y, z, t) = z - \eta(x, y, t) = 0 \quad \text{ocean free surface.} \quad (13.1)$$

This identity holds so long as we assume the surface height,  $\eta$ , is a smooth and well defined surface that contains no overturns at the scales of interest. This assumption requires the filtering of breaking surface waves from the mathematical description.

### 13.2.2 Solid earth boundary

We may mathematically describe the solid Earth lower boundary by using the time independent generalized vertical coordinate

$$\sigma(x, y, z) = z - \eta_b(x, y) = 0 \quad \text{ocean bottom.} \quad (13.2)$$

We typically assume that there is no fluid mass transport through the solid Earth. However, in the case of geothermal heating, we may consider an exchange of heat between the ocean and

the solid Earth. Momentum is exchanged between the solid Earth and ocean fluid through the action of stresses.

### 13.2.3 Ocean mixed layer base

Let

$$\sigma = z - \eta^{\text{mld}}(x, y, t) = 0 \quad (13.3)$$

represent the vertical position of the ocean mixed layer base, with the corresponding normal vector

$$\hat{\mathbf{n}}^{(\text{mld})} = \frac{\nabla(z - \eta^{\text{mld}})}{|\nabla(z - \eta^{\text{mld}})|} = \frac{\nabla(\hat{z} - \nabla\eta^{\text{mld}})}{\sqrt{1 + |\nabla\eta^{\text{mld}}|^2}}. \quad (13.4)$$

This example is relevant for the study of ocean ventilation, whereby we are interested in measuring the transport of fluid that enters the ocean interior across the mixed layer base (see Section 13.8).

### 13.2.4 Interior generalized vertical coordinate surfaces

Within the fluid interior, transport across surfaces of constant generalized vertical coordinate,  $\sigma = \sigma(x, y, z, t)$ , constitutes the dia-surface transport affecting budgets of mass, tracer, and momentum within layers bounded by two generalized vertical coordinate surfaces. A canonical example is provided by isopycnal layers formed by surfaces of constant potential density (equivalently, constant buoyancy surfaces) as used in isopycnal ocean models as well as theoretical descriptions of perfect fluid dynamics. The vertical position of this surface is written in one of two equivalent manners

$$z = z(x, y, \sigma, t) = \eta(x, y, \sigma, t). \quad (13.5)$$

The first expression exposes the functional dependence of the vertical position of the  $\sigma$  surface at a horizontal position and time. In Section 12.2 we discussed the potential for confusion between writing  $z$  as a particular vertical position versus a function, thus motivating  $z = \eta(x, y, \sigma, t)$ . However, by now we should have sufficient experience with generalized vertical coordinates so that we can well distinguish when  $z$  refers to a particular vertical position versus  $z(x, y, \sigma, t)$  as a coordinate function. For this reason we only infrequently use the nomenclature  $z = \eta(x, y, \sigma, t)$  in this chapter.

## 13.3 Specific thickness

As mentioned in Section 12.10.2, a surface of constant generalized vertical coordinate can be successfully used to partition the vertical so long as the transformation between the generalized vertical coordinate and the geopotential coordinate is invertible. The Jacobian of transformation is given by

$$\frac{\partial z}{\partial \sigma} = z_{,\sigma}, \quad (13.6)$$

which must be single signed for suitable generalized vertical coordinates. This constraint means that we do not allow the surfaces to overturn, which is the same assumption made about the ocean surface,  $z = \eta(x, y, t)$ , and solid earth bottom,  $z = \eta_b(x, y)$ . This restriction places a limitation on the ability of certain GVC models (e.g., isopycnal models) to describe non-hydrostatic processes, such as the overturning common in Kelvin-Helmholtz billows and

gravitational convection. Note that for both the solid earth bottom and ocean free surface

$$\frac{\partial z}{\partial \sigma} = 1, \quad (13.7)$$

with this relation also holding for the geopotential coordinate,  $\sigma = z$ .

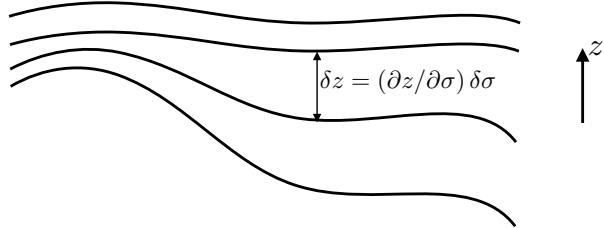


FIGURE 13.2: Illustrating the thickness between surfaces of constant generalized vertical coordinates,  $\delta z = (\partial z / \partial \sigma) \delta \sigma$ . In regions with larger magnitude for the specific thickness,  $\partial z / \partial \sigma$ , or equivalently smaller vertical stratification of the  $\sigma$  surfaces,  $\partial \sigma / \partial z$ , the layers are thicker so that isolines of constant  $\sigma$  are further apart. The converse holds where  $\partial z / \partial \sigma$  is small (equivalently  $\partial \sigma / \partial z$  is large).

We refer to the Jacobian of transform as the *specific thickness* and sometimes find it useful to write it in one of the following manners

$$h = z_{,\sigma} = \frac{\partial z}{\partial \sigma}. \quad (13.8)$$

This name is motivated by noting that the vertical thickness of an infinitesimal layer of coordinate thickness  $\delta \sigma$  is given by

$$\delta z = \frac{\partial z}{\partial \sigma} \delta \sigma = h \delta \sigma, \quad (13.9)$$

with Figure 13.2 providing an example with finitely thick layers. For example, if  $\sigma = b(x, y, z, t)$  (buoyancy or potential density as in isopycnal models), then the thickness of a buoyancy layer is given by

$$\delta z = \frac{\partial z}{\partial b} \delta b = N^{-2} \delta b, \quad (13.10)$$

with

$$N^2 = \frac{\partial b}{\partial z} \quad (13.11)$$

the squared **buoyancy frequency** in a **Boussinesq ocean**. For a hydrostatic flow using pressure as the vertical coordinate, the thickness of a pressure layer is

$$\delta z = \frac{\partial z}{\partial p} \delta p = -\frac{1}{\rho g} \delta p \quad (13.12)$$

where we used the hydrostatic relation

$$\frac{\partial p}{\partial z} = -\rho g \quad (13.13)$$

with  $g$  the constant acceleration due to effective gravity. Note that we assume the layer thickness is positive,  $\delta z > 0$ . For this purpose, with hydrostatic pressure we might choose to consider negative pressure increments,  $\delta p < 0$ , as this corresponds to vertically upward movement in a

fluid column.

## 13.4 The dia-surface transport

In this section we develop the concept of dia-surface transport and derive its expression in terms of the material time derivative of the generalized vertical coordinate.

### 13.4.1 Flow normal to the GVC surface

At an arbitrary point on a surface of constant generalized vertical coordinate (see Figure 13.3), the rate at which fluid moves in the direction normal to the surface is given by

$$\text{RATE OF FLUID FLOW IN DIRECTION } \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}}, \quad (13.14)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla\sigma}{|\nabla\sigma|}, \quad (13.15)$$

is the surface unit normal pointing in a direction of increasing  $\sigma$ . Two examples are useful to ground this expression in common experience. For the ocean free surface,  $\sigma = z - \eta(x, y, t) = 0$ , the unit normal takes the form

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla\eta}{\sqrt{1 + |\nabla\eta|^2}}, \quad (13.16)$$

whereas at the solid Earth bottom,  $\sigma = z - \eta_b(x, y) = 0$ ,

$$\hat{\mathbf{n}} = -\frac{\nabla(z - \eta_b)}{|\nabla(z - \eta_b)|} = \frac{-\hat{\mathbf{z}} + \nabla\eta_b}{\sqrt{1 + |\nabla\eta_b|^2}}. \quad (13.17)$$

Use of the material time derivative

$$\frac{D\sigma}{Dt} = \frac{\partial\sigma}{\partial t} + \mathbf{v} \cdot \nabla\sigma \quad (13.18)$$

in equation (13.14) leads to the identity

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \frac{1}{|\nabla\sigma|} \left[ \frac{D\sigma}{Dt} - \frac{\partial\sigma}{\partial t} \right]. \quad (13.19)$$

Hence, the component to the velocity of a fluid particle that is normal to a generalized vertical coordinate surface is proportional to the difference between the material time derivative of the surface coordinate and its partial time derivative.

### 13.4.2 Accounting for movement of the surface

A generalized vertical coordinate surface is generally moving. So to diagnose the net transport of fluid penetrating the surface requires us to subtract the velocity of the surface,  $\mathbf{v}^{(\sigma)}$ , from the velocity of a fluid particle. We are thus led to

$$\text{RATE THAT FLUID CROSSES A MOVING GVC SURFACE} = \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}). \quad (13.20)$$

We next develop a kinematic property of the surface velocity, or more precisely the normal

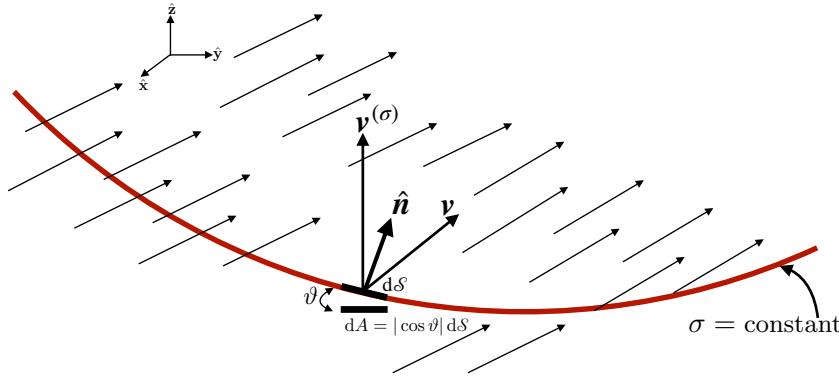


FIGURE 13.3: A surface of constant generalized vertical coordinate,  $\sigma = \text{constant}$ , within a fluid. The normal direction,  $\hat{\mathbf{n}} = \nabla\sigma/|\nabla\sigma|$ , points in the direction of increasing  $\sigma$ . We show an example velocity vector for a fluid particle,  $\mathbf{v}$ , at a point on the surface as well as the velocity,  $\mathbf{v}^{(\sigma)}$ , of a point that lives on the surface. Note that kinematics is only concerned with the normal component to the surface velocity,  $\mathbf{v}^{(\sigma)} \cdot \hat{\mathbf{n}}$ , as per equation (13.25). We require dynamical information to obtain information about the tangential component of  $\mathbf{v}^{(\sigma)}$ , but such information is not required for this chapter. Following equation (13.30), the horizontal projection of the surface area element is given by  $dA = |\cos \vartheta| dS$ , where  $\vartheta$  is the angle between the surface and the horizontal and  $dA = dx dy$ .

component to that velocity. For that purpose, consider an infinitesimal increment in both space and time under which  $\sigma$  undergoes an infinitesimal change

$$\delta\sigma = \delta\mathbf{x} \cdot \nabla\sigma + \delta t \partial_t\sigma. \quad (13.21)$$

Now restrict attention to a point fixed on a constant  $\sigma$  surface, in which

$$\delta\sigma = \delta\mathbf{x}^{(\sigma)} \cdot \nabla\sigma + \delta t \partial_t\sigma = 0, \quad (13.22)$$

where  $\delta\mathbf{x}^{(\sigma)}$  is a differential increment following the moving surface. We define the velocity of that point as

$$\mathbf{v}^{(\sigma)} = \frac{\delta\mathbf{x}^{(\sigma)}}{\delta t}, \quad (13.23)$$

in which case equation (13.22) implies that at each point within the fluid,

$$\frac{\partial\sigma}{\partial t} + \mathbf{v}^{(\sigma)} \cdot \nabla\sigma = 0. \quad (13.24)$$

We can likewise write this equation as one for the normal component of the surface velocity

$$\mathbf{v}^{(\sigma)} \cdot \hat{\mathbf{n}} = -\frac{1}{|\nabla\sigma|} \frac{\partial\sigma}{\partial t}. \quad (13.25)$$

Evidently, the normal component to the surface velocity vanishes when the surface is locally static.

### 13.4.3 We only care about divergent surface motion

For the kinematics of fluid motion relative to a surface of constant generalized vertical coordinates, we are only concerned with the normal component to the surface velocity,  $\mathbf{v}^{(\sigma)} \cdot \hat{\mathbf{n}}$ . That is, we are only concerned with divergent motion of the surface, defined as motion parallel

to the surface normal direction,  $\hat{\mathbf{n}}$ . We have no concern for rotational or tangential motion, which is motion perpendicular to  $\hat{\mathbf{n}}$ . Even so, some authors, by fiat, choose to set to zero the tangential component of the surface motion. In fact, specification of the tangential surface velocity component is generally not available without extra information about the surface motion, nor is its specification necessary for developing kinematic properties of fluid motion relative to arbitrary generalized vertical coordinate surfaces. Hence, we make no assumption about tangential motion of the surface.

#### 13.4.4 Dia-surface transport in terms of $D\sigma/Dt$

Using expression (13.25) in equation (13.20) leads to the net flux of fluid crossing the surface of constant generalized vertical coordinate

$$\hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = \frac{1}{|\nabla\sigma|} \frac{D\sigma}{Dt} \iff \nabla\sigma \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = \frac{D\sigma}{Dt}. \quad (13.26)$$

Hence, if the material time derivative of a generalized vertical coordinate vanishes, then there is zero fluid crossing the surface with the *barycentric velocity*. That is, this result holds for motion of the fluid as defined by the barycentric velocity,  $\mathbf{v}$  (VOLUME 2). Hence, for multi-component fluids,  $D\sigma/Dt = 0$  can still, in principle, be associated with tracer exchange or heat exchange across the surface via diffusion, so long as the net matter crossing the surface is zero.

#### 13.4.5 Defining the dia-surface transport

The area normalizing the volume flux in equation (13.26) is the area  $dS$  of an infinitesimal patch on the surface of constant generalized vertical coordinate with outward unit normal  $\hat{\mathbf{n}}$ . We now follow the trigonometry discussed in Section 12.17 to introduce the horizontal projection of this area,  $dA$ , which is more convenient to work with for many purposes. So long as the vertical stratification remains non-zero ( $\partial\sigma/\partial z \neq 0$ ) we can write the area factor in the form

$$\frac{dS}{|\nabla\sigma|} = \frac{dS}{\sqrt{(\partial\sigma/\partial x)^2 + (\partial\sigma/\partial y)^2 + (\partial\sigma/\partial z)^2}} \quad (13.27a)$$

$$= \frac{dS}{|\partial\sigma/\partial z| \sqrt{[(\partial\sigma/\partial x)/(\partial\sigma/\partial z)]^2 + [(\partial\sigma/\partial y)/(\partial\sigma/\partial z)]^2 + 1}} \quad (13.27b)$$

$$= \frac{dS}{|\partial\sigma/\partial z| \sqrt{1 + \tan^2 \vartheta}} \quad (13.27c)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| |\cos \vartheta| dS \quad (13.27d)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| dA. \quad (13.27e)$$

The equality (13.27c) introduced the angle,  $\vartheta$ , between the boundary surface and the horizontal plane (Figure 13.3). The squared slope of this surface given by (see Section 12.12)

$$\tan^2 \vartheta = \frac{\nabla_h \sigma \cdot \nabla_h \sigma}{(\partial\sigma/\partial z)^2} = \nabla_{bv} z \cdot \nabla_{bv} z. \quad (13.28)$$

The equality (13.27d) made use of a trigonometric identity so that

$$|\cos \vartheta|^{-1} = |z_{,\sigma} \nabla \sigma|. \quad (13.29)$$

Furthermore, the equality (13.27e) introduced the horizontal projection of the area,

$$dA = |\cos \vartheta| d\mathcal{S}. \quad (13.30)$$

We now introduce the *dia-surface velocity component* for the GVC coordinate

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = z_{,\sigma} \dot{\sigma}, \quad (13.31)$$

which measures the volume of fluid passing through the surface, per unit horizontal area, per unit time

$$w^{(\dot{\sigma})} \equiv \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \frac{d\mathcal{S}}{dA} \quad (13.32)$$

$$= \frac{(\text{VOLUME/TIME}) \text{ FLUID THROUGH SURFACE}}{\text{HORIZONTAL AREA OF SURFACE}}, \quad (13.33)$$

so that

$$w^{(\dot{\sigma})} dA \equiv \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) d\mathcal{S}. \quad (13.34)$$

We refer to  $w^{(\dot{\sigma})}$  as the dia-surface velocity component since it measures flow rate of fluid through the surface. We can think of  $w^{(\dot{\sigma})}$  as the “vertical” velocity which, when multiplied by the horizontal area element, measures the transport of fluid that crosses the surface in the normal direction.

#### 13.4.6 Expressions for the dia-surface velocity component

Making use of various identities derived above, as well as the transformation of partial derivative operators in Section 12.12, allows us to write the dia-surface velocity component in the following equivalent forms

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} \quad (13.35a)$$

$$= z_{,\sigma} |\nabla \sigma| \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \quad (13.35b)$$

$$= z_{,\sigma} \nabla \sigma \cdot \mathbf{v} - z_{,\sigma} |\nabla \sigma| \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \quad (13.35c)$$

$$= (\hat{\mathbf{z}} - \nabla_{\mathbf{h}} z) \cdot \mathbf{v} + z_{,\sigma} \partial_t \sigma \quad (13.35d)$$

$$= (\hat{\mathbf{z}} - \nabla_{\mathbf{h}} z) \cdot \mathbf{v} - \frac{\partial z}{\partial t} \quad (13.35e)$$

$$= w - (\partial_t + \mathbf{u} \cdot \nabla_{\mathbf{h}}) z, \quad (13.35f)$$

where  $\partial z / \partial t = (\partial z / \partial t)_{\sigma}$  is the time derivative for the depth of the  $\sigma$  surface. We also made use of the identity (see equations (12.39b) and (12.39c))

$$\nabla_{\mathbf{h}} z = -z_{,\sigma} \nabla_{\mathbf{h}} \sigma = -\nabla_{\mathbf{h}} \sigma / \partial_z \sigma \quad (13.36)$$

to express the slope of the  $\sigma$  surface as projected onto the horizontal direction plane, as well as the triple product identity (12.39a) for the time derivative

$$\left[ \frac{\partial z}{\partial t} \right]_{\sigma} = - \frac{[\partial \sigma / \partial t]_z}{[\partial \sigma / \partial z]}. \quad (13.37)$$

### 13.4.7 Relating $w$ and $w^{(\dot{\sigma})}$

Equation (13.35f) directly relates the dia-surface velocity component to the vertical component of the fluid particle velocity

$$w = \frac{Dz}{Dt} \longleftrightarrow w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = w - (\partial_t + \mathbf{u} \cdot \nabla_{\sigma})z. \quad (13.38)$$

When the  $\sigma$ -surface is static, so that it occupies a constant vertical position,  $\partial z / \partial t = 0$ , then the dia-surface velocity component reduces to

$$w^{(\dot{\sigma})} = w - \mathbf{u} \cdot \nabla_{\sigma} z \quad \text{static } \sigma\text{-surface}, \quad (13.39)$$

whereas if the  $\sigma$ -surface is flat, then the dia-surface velocity component measures the flux of fluid moving vertically relative to the motion of the generalized vertical coordinate surface. Finally, if the surface is flat and static, the dia-surface velocity component becomes the vertical velocity component

$$w^{(\dot{\sigma})} = w = \frac{Dz}{Dt} \quad \sigma\text{-surface static and flat}, \quad (13.40)$$

which is the case for the geopotential vertical coordinate.

### 13.4.8 Dia-surface velocity component for water mass analysis

In some literature presentations, the dia-surface velocity component is taken to be

$$w^{\text{dia}} = \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = \frac{1}{|\nabla \sigma|} \frac{D\sigma}{Dt}. \quad (13.41)$$

As seen in Chapter 20, the reason to prefer equation (13.41) for watermass transformation analysis is that avoids the need to assume vertically stable stratification for surfaces of constant  $\sigma$ . Dropping that assumption allows us to consider water mass transformation between arbitrarily oriented elements of seawater, even those that are gravitationally unstable, with  $w^{\text{dia}}$  the appropriate object to measure flow across such arbitrarily oriented surfaces.

### 13.4.9 Area integrated dia-surface transport for non-divergent flows

We close this section by further emphasizing the distinction, in time dependent flows, between dia-surface transport and flow normal to a surface. For this purpose consider a non-divergent flow whereby  $\nabla \cdot \mathbf{v} = 0$ . Non-divergence means that for any closed surface within the fluid interior, the following identity holds via the divergence theorem

$$0 = \int_{\mathcal{R}} \nabla \cdot \mathbf{v} dV = \oint_{\partial \mathcal{R}} \hat{\mathbf{n}} \cdot \mathbf{v} dS. \quad (13.42)$$

Notably, only in the case of a static surface can we conclude that there is no net flow across the surface. For surfaces that move, there is generally a nonzero net dia-surface transport.

As a specific example, consider a fluid region such as shown in Figure 13.4, which is bounded below by the solid-earth bottom and above by a constant  $\sigma$ -surface. Since the solid-earth bottom is static and there is no-normal flow through the bottom, the identity (13.42) means

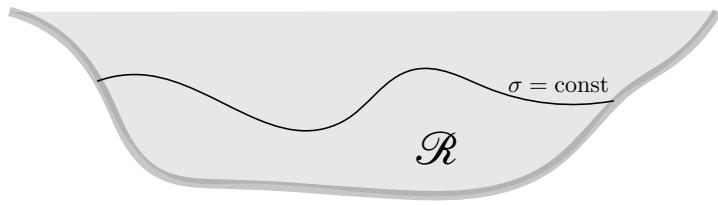


FIGURE 13.4: A constant generalized vertical coordinates surface,  $\sigma = \text{constant}$ , within a fluid region that intersects the bottom. The region,  $\mathcal{R}$ , is bounded above by the  $\sigma$ -surface and below by the solid-earth. Along the constant  $\sigma$ -surface a non-divergent flow satisfies  $\int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot \mathbf{v} dS = 0$ .

that the area integrated flow normal to the GVC vanishes

$$\int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot \mathbf{v} dS = 0. \quad (13.43)$$

But what does this identity imply about the area integrated dia-surface velocity? For the case of a geopotential vertical coordinate,  $\sigma = z$ , it means that the area integrated vertical velocity vanishes across any geopotential surface below the ocean free surface,  $\int_{z=\text{const}} w dA = 0$  (see VOLUME 2). What about other generalized vertical coordinates?

To address this question consider the general result

$$\int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) dS = \int_{\sigma=\text{const}} w^{\text{dia}} dS = \int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA, \quad (13.44)$$

where again  $dA = dx dy$ . Now make use of the property (13.43) for non-divergent flows as well as the identity (13.25) to render

$$\int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA = - \int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot \mathbf{v}^{(\sigma)} dS \quad (13.45a)$$

$$= \int_{\sigma=\text{const}} \frac{\partial_t \sigma}{|\nabla \sigma|} dS \quad (13.45b)$$

$$= \int_{\sigma=\text{const}} \frac{\partial \sigma}{\partial t} \left| \frac{\partial z}{\partial \sigma} \right| dA \quad (13.45c)$$

$$= - \int_{\sigma=\text{const}} \left[ \frac{\partial z}{\partial t} \right]_\sigma dA. \quad (13.45d)$$

The final equality holds if  $\partial z / \partial \sigma > 0$ , whereas we swap signs when the vertical stratification is  $\partial z / \partial \sigma < 0$ . We can go one further step by noting that the time derivative is computed with  $\sigma$  constant, as is the horizontal area integral. Hence, we can pull the time derivative outside the integral to render

$$\int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA = - \left[ \frac{\partial}{\partial t} \right]_\sigma \int_{\sigma=\text{const}} z dA. \quad (13.46)$$

Evidently, for a non-divergent flow the integrated dia-surface transport across a  $\sigma$ -surface equals to minus the time tendency for the area integrated vertical position of that surface. Hence, there is an area integrated dia-surface transport across the  $\sigma$ -surface so long as there is a volume change for the region beneath the surface.

For the case of an isopycnal surface in a perfect fluid, there is no change in the volume beneath any interior isopycnal surface since no flow crosses the isopycnal, in which case we recover the expected result  $\int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA = 0$ . However, this result does not hold for other

coordinates, such as the rescaled vertical coordinate discussed in Sections 12.3.3 and 13.10.4, in which case

$$z^* = H \frac{z - \eta}{H + \eta} \quad (13.47a)$$

$$\frac{\partial z}{\partial z^*} = 1 + H/\eta > 0 \quad (13.47b)$$

$$\left[ \frac{\partial z}{\partial t} \right]_{z^*} = \frac{\partial \eta}{\partial t} (1 + z^*/H), \quad (13.47c)$$

so that

$$\int_{z^*=\text{const}} w^{(z^*)} dA = \int_{z^*=\text{const}} (\partial \eta / \partial t) (1 + z^*/H) dA, \quad (13.48)$$

which is generally nonzero. For example, consider a flat bottom so that

$$\int_{z^*=\text{const}} w^{(z^*)} dA = (1 + z^*/H) \int_{z^*=\text{const}} (\partial \eta / \partial t) dA = (1 + z^*/H) \int_{z^*=\text{const}} (Q_m / \rho_o) dA, \quad (13.49)$$

where  $Q_m$  is the surface mass flux and we made use of the free surface equation

$$\partial_t \eta = Q_m / \rho_o - \nabla \cdot \mathbf{U}, \quad (13.50)$$

holding for a non-divergent flow (VOLUME 2). In this case the area integrated dia-surface transport across a  $z^*$  surface is proportional to the area integrated surface mass flux.

## 13.5 Material time derivative

The expression (13.31) for  $w^{(\sigma)}$  brings the material time derivative operator into the following equivalent forms

$$\frac{D}{Dt} = \left[ \frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_h + w \frac{\partial}{\partial z} \quad (13.51a)$$

$$= \left[ \frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma + \frac{D\sigma}{Dt} \frac{\partial}{\partial \sigma} \quad (13.51b)$$

$$= \left[ \frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma + w^{(\sigma)} \frac{\partial}{\partial z}. \quad (13.51c)$$

Note that the chain-rule means that

$$\frac{\partial}{\partial \sigma} = \frac{\partial z}{\partial \sigma} \frac{\partial}{\partial z}, \quad (13.52)$$

thus providing a relationship between the two vertical coordinate partial derivatives. Furthermore, recall that subscripts in the above derivative operators denote variables held fixed when taking the partial derivatives.

We highlight the special case of no fluid particles crossing the generalized coordinate surface. This situation occurs for a perfect fluid with  $\sigma$  equal to the buoyancy or isopycnal coordinate. For perfect fluid flow, the material time derivative in equation (13.51c) only has a horizontal two-dimensional advective component,  $\mathbf{u} \cdot \nabla_\sigma$ . This result does *not* mean that fluid particle motion in a perfect fluid is strictly horizontal. Indeed, it generally is not, as the form given by equation (13.51a) makes clear. Rather, it means that the advective transport of fluid properties

occurs along surfaces of constant buoyancy, and such transport is measured by the convergence of horizontal advective fluxes as measured along these constant buoyancy surfaces.

## 13.6 Vertical velocity and dia-surface velocity

Making use of the material time derivative operator in the form of equation (13.51c) affords us an opportunity to emphasize both the differences and similarities between the vertical velocity component and the dia-surface velocity component. Namely, the vertical velocity component takes on the equivalent forms

$$w = \frac{Dz}{Dt} = \left[ \frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \left[ -\frac{\partial \sigma}{\partial t} - \mathbf{u} \cdot \nabla_{\sigma} \sigma + \frac{D\sigma}{Dt} \right], \quad (13.53)$$

and the corresponding expressions for the dia-surface velocity component are given by

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = \frac{\partial z}{\partial \sigma} \left[ \frac{\partial \sigma}{\partial t} + \mathbf{u} \cdot \nabla_{\sigma} \sigma + w \frac{\partial \sigma}{\partial z} \right] = - \left[ \frac{\partial z}{\partial t} \right]_{\sigma} - \mathbf{u} \cdot \nabla_{\sigma} z + w. \quad (13.54)$$

Whereas the vertical velocity component,  $w$ , measures the transport crossing  $z$  surfaces, which are static and horizontal, the dia-surface velocity component,  $w^{(\dot{\sigma})}$ , measures the transport crossing  $\sigma$  surfaces, which are generally moving and sloped. It is notable that the area normalization used in equation (13.33) for the dia-surface velocity component means that it appears only in the expression for the vertical velocity (i.e.,  $w^{(\dot{\sigma})}$  is not in the expressions for the horizontal velocity components,  $u$  or  $v$ ). However, as we see in the following, the appearance of  $w^{(\dot{\sigma})}$  in the  $w$  equation does not necessarily mean that  $w^{(\dot{\sigma})}$  corresponds to vertical particle motion. Instead, a nonzero  $w^{(\dot{\sigma})}$  can arise from vertical motion of the  $\sigma$ -surface even while a fluid particle remains fixed in space.

### 13.6.1 Decomposing the vertical velocity

The expression

$$w = \left[ \frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} \quad (13.55)$$

decomposes the vertical velocity of a fluid particle into (i) time changes to the vertical position of the  $\sigma$ -surface at a particular horizontal point, (ii) lateral particle motion projected onto a sloped  $\sigma$ -surface, (iii) motion that crosses a  $\sigma$ -surface. Importantly, the three terms are coupled. For example, consider the case of  $\sigma$  defined by isopycnals, in which case irreversible mixing ( $w^{(\dot{\sigma})} \neq 0$ ) changes the configuration of  $\sigma$  surfaces by changing both their height,  $(\partial z / \partial t)_{\sigma}$ , and their slope,  $\nabla_{\sigma} z$ .

### 13.6.2 Another form of the vertical velocity decomposition

Consider the velocity for a point on the surface,  $\mathbf{v}^{(\sigma)}$ , which satisfies (Section 13.4.2)

$$\frac{\partial \sigma}{\partial t} + \mathbf{v}^{(\sigma)} \cdot \nabla \sigma = 0. \quad (13.56)$$

Making use of the triple product identities from Section 12.6

$$\frac{\partial z}{\partial \sigma} \nabla \sigma = -\nabla_{\mathbf{v}} z + \hat{\mathbf{z}} \quad \text{and} \quad \frac{\partial z}{\partial \sigma} \left[ \frac{\partial \sigma}{\partial t} \right]_z = - \left[ \frac{\partial z}{\partial t} \right]_\sigma \quad (13.57)$$

brings equation (13.56) into the form

$$\left[ \frac{\partial z}{\partial t} \right]_\sigma = (\hat{\mathbf{z}} - \nabla_{\mathbf{v}} z) \cdot \mathbf{v}^{(\sigma)} \implies \hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} = \left[ \frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u}^{(\sigma)} \cdot \nabla_{\mathbf{v}} z, \quad (13.58)$$

where  $\mathbf{u}^{(\sigma)}$  is the horizontal component to the surface velocity,  $\mathbf{v}^{(\sigma)}$ . This equation shows that the vertical component to the  $\sigma$ -surface velocity is given by the sum of the changes to the vertical position of the surface, plus the projection of the horizontal motion of the surface onto the slope of the surface. Additionally, even if the  $\sigma$ -surface has no component of velocity in the vertical, the vertical position of the  $\sigma$ -surface, as measured at a horizontal point, generally changes if the surface is sloped and moves horizontally pass that point

$$\left[ \frac{\partial z}{\partial t} \right]_\sigma = -\mathbf{u}^{(\sigma)} \cdot \nabla_{\mathbf{v}} z \quad \text{if } \hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} = 0. \quad (13.59)$$

Returning to the general result (13.58) allows us to write

$$\left[ \frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_{\mathbf{v}} z = \hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\mathbf{v}} z. \quad (13.60)$$

Furthermore, return to the fundamental definition of the dia-surface velocity component detailed in Section 13.4, in which we showed that

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = \frac{\partial z}{\partial \sigma} \nabla \sigma \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = (-\nabla_{\mathbf{v}} z + \hat{\mathbf{z}}) \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}). \quad (13.61)$$

This expression, along with equation (13.60), leads to the rather elaborate decomposition of the vertical velocity component according to motion of a generalized vertical coordinate surface

$$w = \underbrace{\left[ \hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\mathbf{v}} z \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\mathbf{v}}) z} + \underbrace{\left[ \hat{\mathbf{z}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) - (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\mathbf{v}} z \right]}_{w^{(\dot{\sigma})}}. \quad (13.62)$$

Terms in the first bracket compute vertical particle motion relative to the  $\sigma$ -surface. The dia-surface contribution from the second bracket removes the contribution from  $\sigma$ -surface motion to leave just the vertical motion of the particle. All terms on the right hand side cancel, except for  $\hat{\mathbf{z}} \cdot \mathbf{v} = w$ , thus trivially revealing  $w = w$ . The decomposition of  $w$  is rather pedantic when viewed in the unpacked form of equation (13.62). Even so, let us consider some special cases to offer further interpretation.

- NO HORIZONTAL CONTRIBUTION: Consider the case where the horizontal velocity of a fluid particle matches that of the  $\sigma$ -surface:  $\mathbf{u} = \mathbf{u}^{(\sigma)}$ . Alternatively, consider the case with flat  $\sigma$ -surfaces so that  $\nabla_{\mathbf{v}} z = 0$ . In either case the vertical velocity is given by

$$w = \underbrace{\left[ \hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\mathbf{v}}) z} + \underbrace{\left[ \hat{\mathbf{z}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \right]}_{w^{(\dot{\sigma})}}. \quad (13.63)$$

The first contribution is from vertical motion of the  $\sigma$ -surface. The second contribution adjusts for the vertical motion of the particle relative to the  $\sigma$ -surface, leaving behind just the vertical motion of the particle. This rather trivial case exemplifies the contributions from the two pieces of the vertical velocity.

- ZERO VERTICAL PARTICLE MOTION: Consider the case where  $w = 0$  so that

$$w = 0 \quad (13.64a)$$

$$= \left[ \frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\mathbf{h}\sigma} z + w^{(\dot{\sigma})} \quad (13.64b)$$

$$= \underbrace{\left[ \hat{z} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\mathbf{h}\sigma} z \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\mathbf{h}\sigma}) z} + \underbrace{\left[ -\hat{z} \cdot \mathbf{v}^{(\sigma)} - (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\mathbf{h}\sigma} z \right]}_{w^{(\dot{\sigma})}}. \quad (13.64c)$$

The final expression is trivial since each term in one bracket identically cancels terms in the other bracket. The penultimate expression reveals the balance between dia-surface transport and motion relative to the  $\sigma$  surface

$$-w^{(\dot{\sigma})} = \left[ \frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\mathbf{h}\sigma} z \quad \text{if } w = 0. \quad (13.65)$$

A particularly simple realization of this balance holds for  $\sigma$  given by isopycnals and where the isosurfaces are horizontal. In the presence of uniform mixing, the flat isopycnals stay flat and there is correspondingly no vertical motion of fluid particles even as the vertical stratification is modified. In contrast, the vertical position of an isopycnal surface changes according to the dia-surface velocity component  $(\partial z / \partial t)_{\sigma} = -w^{(\dot{\sigma})} \neq 0$ . This case illustrates that  $w^{(\dot{\sigma})} \neq 0$  can still occur even when there is zero fluid particle motion since  $w^{(\dot{\sigma})} \neq 0$  can arise from motion of a  $\sigma$ -surface alone.

## 13.7 The velocity vector and fluid particle trajectories

In this section we explore aspects of the velocity vector written using both Cartesian coordinates and generalized vertical coordinates. We complement that analytical discussion with a geometrical interpretation of fluid particle motion.

### 13.7.1 Coordinate representations of the velocity vector

Recall from Section 13.6 the vertical velocity component can be written in two forms given by equation (13.53). We focus on the expression

$$w = \left[ \frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\mathbf{h}\sigma} z + w^{(\dot{\sigma})}, \quad (13.66)$$

so that the full velocity vector can be written in the following equivalent manners

$$\mathbf{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}} \quad (13.67a)$$

$$= u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + \left[ (\partial z / \partial t)_{\sigma} + \mathbf{u} \cdot \nabla_{\mathbf{h}\sigma} z + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}} \quad (13.67b)$$

$$= u [\hat{\mathbf{x}} + \hat{z} (\partial z / \partial x)_{\sigma}] + v [\hat{\mathbf{y}} + \hat{z} (\partial z / \partial y)_{\sigma}] + \left[ (\partial z / \partial t)_{\sigma} + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}}. \quad (13.67c)$$

$$= u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_\sigma] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_\sigma] + [(w - \nabla_{\mathbf{v}} z) \cdot \mathbf{v}] \hat{\mathbf{z}} \quad (13.67d)$$

$$= u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_\sigma] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_\sigma] + (\partial z / \partial \sigma) (\nabla \sigma \cdot \mathbf{v}) \hat{\mathbf{z}}. \quad (13.67e)$$

From Section 12.9, we can connect equation (13.67e) to the contravariant representation of the velocity vector using generalized vertical coordinates

$$\mathbf{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}} = v^{\bar{1}} \mathbf{e}_{\bar{1}} + v^{\bar{2}} \mathbf{e}_{\bar{2}} + v^{\bar{3}} \mathbf{e}_{\bar{3}}, \quad (13.68)$$

where we introduced the generalized vertical coordinate basis vectors,  $\mathbf{e}_{\bar{\alpha}}$ , given by equations (12.20a)-(12.20c)

$$\mathbf{e}_{\bar{1}} = \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_\sigma \quad \text{and} \quad \mathbf{e}_{\bar{2}} = \hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_\sigma \quad \text{and} \quad \mathbf{e}_{\bar{3}} = \hat{\mathbf{z}} (\partial z / \partial \sigma), \quad (13.69)$$

and the contravariant components to velocity vector (equation (12.54))

$$v^{\bar{1}} = u \quad \text{and} \quad v^{\bar{2}} = v \quad \text{and} \quad v^{\bar{3}} = \nabla \sigma \cdot \mathbf{v} \quad (13.70)$$

### 13.7.2 Fluid particle trajectories

To help further understand the velocity as represented using generalized vertical coordinates, consider the following three cases as illustrated in Figure 13.5.

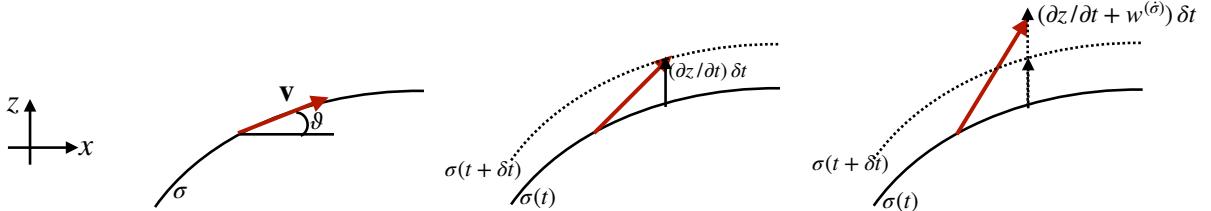


FIGURE 13.5: This schematic shows the various contributions to the fluid particle velocity (red vector) when written relative to motion of a particular generalized vertical coordinate surface. The fluid particle sits at the tail of the velocity vector at time  $t$  and at the head at time  $t + \delta t$ . The left panel is for the case of a static and material  $\sigma$ -surface so that the particle remains on the  $\sigma$ -surface and has a velocity vector given by equation (13.72). The slope of the  $\sigma$ -surface in the  $\hat{\mathbf{x}}$ -direction is given by  $\tan \vartheta = (\partial z / \partial x)_\sigma$ . The middle panel is for a non-steady material  $\sigma$ -surface whereby the velocity of a particle takes on the form (13.73), with the particle remaining on the moving  $\sigma$ -surface. The right panel shows the case of a non-steady and non-material  $\sigma$ -surface with velocity (13.74). In this final case the particle position departs from the original  $\sigma$ -surface due to the nonzero dia-surface velocity component,  $w^{(\dot{\sigma})} \neq 0$ . It is not known *a priori* whether this departure is due to particle motion, motion of the surface, or both. Notably, the horizontal position of the particle remains identical for each of the three cases. It is only the vertical position that is modified according to the slope of the  $\sigma$ -surface (left panel), motion of the  $\sigma$ -surface (middle panel), and motion crossing the  $\sigma$ -surface (right panel).

- STEADY AND MATERIAL  $\sigma$ -SURFACE: The velocity vector is aligned with the instantaneous  $\sigma$ -surface ( $\mathbf{v} \cdot \nabla \sigma = 0$ ) when the  $\sigma$ -surface is steady ( $\partial \sigma / \partial t = 0$ ) and material ( $D\sigma / Dt = 0$ ). Hence, we can diagnose the vertical velocity component in terms of the horizontal via

$$w \partial \sigma / \partial z = -\mathbf{u} \cdot \nabla_h \sigma \implies w = \mathbf{u} \cdot \nabla_{\mathbf{v}} z, \quad (13.71)$$

where we used the triple product identities (12.39b) and (12.39c) for the final equality. The velocity vector thus takes on the form

$$\partial \sigma / \partial t = D\sigma / Dt = 0 \implies \mathbf{v} = u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_\sigma] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_\sigma]. \quad (13.72)$$

In this case, the fluid particle motion remains on the static  $\sigma$ -surface, so that the fluid velocity vector is determined by the horizontal velocity plus the slope of the  $\sigma$  surface.

- NON-STEADY AND MATERIAL  $\sigma$ -SURFACE: Next consider material  $\sigma$  surfaces ( $D\sigma/Dt = 0$ ) that move ( $\partial_t \sigma \neq 0$ ), in which case the velocity vector is

$$D\sigma/Dt = 0 \implies \mathbf{v} = u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z/\partial x)_\sigma] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z/\partial y)_\sigma] + (\partial z/\partial t)_\sigma \hat{\mathbf{z}}. \quad (13.73)$$

To remain on the moving surface, the fluid particle must move vertically by the extra amount  $(\partial z/\partial t)_\sigma \delta t \hat{\mathbf{z}}$  relative to the case of a static  $\sigma$ -surface.

- NON-STEADY AND NON-MATERIAL  $\sigma$ -SURFACE: The general case with a non-material and non-steady  $\sigma$  also requires the dia-surface velocity component,  $w^{(\dot{\sigma})}$ , which is diagnosed based on the material time derivative of  $\sigma$  and the inverse stratification,  $w^{(\dot{\sigma})} = (\partial z/\partial \sigma) D\sigma/Dt$ :

$$\mathbf{v} = u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z/\partial x)_\sigma] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z/\partial y)_\sigma] + \left[ (\partial z/\partial t)_\sigma + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}}. \quad (13.74)$$

The contribution from  $w^{(\dot{\sigma})}$  measures the vertical motion of the particle relative to the moving  $\sigma$ -surface. Hence, the sum,  $(\partial z/\partial t)_\sigma + w^{(\dot{\sigma})}$ , measures the vertical motion of the particle relative to a fixed origin. As emphasized in Section 13.6, a non-zero  $w^{(\dot{\sigma})}$  arises from motion of the fluid particle relative to the  $\sigma$ -surface, and this relative motion does not necessarily mean that the particle moves; e.g., recall the example discussed in Section 13.6.2 with a static particle and moving  $\sigma$ -surface.

## 13.8 Subduction across the mixed layer base

As a brief example of the formalism, consider the generalized vertical coordinate defined according to the mixed layer base as in equation (13.3). The dia-surface mass transport across this surface leads us to define the subduction

$$-\mathcal{S}^{(\text{subduct})} \equiv \rho dA \left[ \frac{D(z - \eta^{\text{mld}})}{Dt} \right] \quad \text{at } z = \eta^{\text{mld}}(x, y, t), \quad (13.75)$$

where the mass transport,  $\mathcal{S}^{(\text{subduct})}$  (dimensions of mass per time), is positive for fluid moving downward beneath the mixed layer base into the pycnocline (subduction) and negative for water moving into the mixed layer (obduction). The area element,  $dA$ , is the horizontal projection of the area on the mixed layer base. Expanding the material time derivative leads to

$$\left[ \frac{\mathcal{S}^{(\text{subduct})}}{\rho dA} \right] = -w + [\partial_t + \mathbf{u} \cdot \nabla] \eta^{\text{mld}} \quad \text{at } z = \eta^{\text{mld}}(x, y, t), \quad (13.76)$$

where again we define

$$\mathcal{S}^{(\text{subduct})} > 0 \quad \text{subduction (mixed layer loses mass)} \quad (13.77)$$

$$\mathcal{S}^{(\text{subduct})} < 0 \quad \text{obduction (mixed layer gains mass).} \quad (13.78)$$

We illustrate the kinematics of subduction in Figure 13.6. For the particular case of positive subduction (water moves from mixed layer to the interior), equation (13.76) reveals the following

three kinematic means for that to happen.

$$w < 0 \quad \text{vertically downward flow through mixed layer base} \quad (13.79a)$$

$$\partial_t \eta^{\text{mld}} > 0 \quad \text{mixed layer base shoals, thus decreasing MLD thickness} \quad (13.79b)$$

$$\mathbf{u} \cdot \nabla \eta^{\text{mld}} > 0 \quad \text{horizontal flow crossing sloped MLD boundary.} \quad (13.79c)$$

These pieces to the subduction provide kinematic answers to how water crossed the base of the mixed layer. We need dynamical information to determine why it crossed.

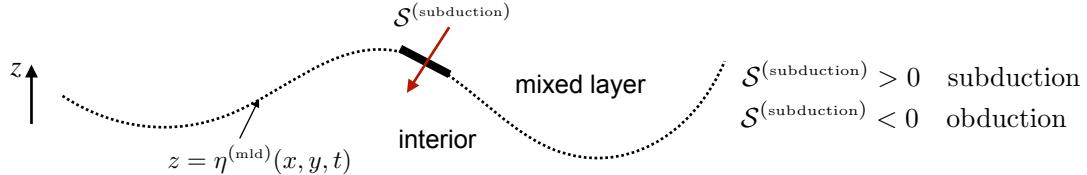


FIGURE 13.6: Illustrating the subduction as defined by equation (13.76), which measures the mass transport across the base of the ocean mixed layer.  $S^{(\text{subduct})} > 0$  for water that enters the ocean interior, in which we say that water subducts from the mixed layer to the ocean interior, thus depleting the seawater mass contained in the mixed layer. Conversely,  $S^{(\text{subduct})} < 0$  when water enters the mixed layer from below, in which we say that water is obducted from the interior into the mixed layer, thus adding seawater mass to the mixed layer. From equations (13.79a)-(13.79c) we see three kinematic means for water to cross  $z = \eta^{\text{mld}}$ : vertical flow, horizontal flow across a sloping mixed layer base, and vertical expansion or contraction of the mixed layer thickness.

## 13.9 Mass continuity

We here derive the Eulerian expression for mass continuity using generalized vertical coordinates. We then specialize to non-divergent flows, in which mass conservation is converted to volume conservation. To start, recall that mass conservation for a fluid element states that To develop the Eulerian expressions we first consider the case of Cartesian coordinates.

$$\rho \delta V = \rho \delta x \delta y \delta z = \rho \delta x \delta y z_{,\sigma} \delta \sigma \quad (13.80)$$

is constant following a fluid element.

### 13.9.1 Cartesian coordinates

Consider the expression

$$\frac{1}{\rho \delta V} \frac{D(\rho \delta V)}{Dt} = \frac{1}{\rho} \frac{D\rho}{Dt} + \frac{1}{\delta V} \frac{D(\delta V)}{Dt}. \quad (13.81)$$

Now make use of Cartesian coordinates to write the volume

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \frac{1}{\delta x \delta y \delta z} \frac{D(\delta x \delta y \delta z)}{Dt} \quad (13.82a)$$

$$= \frac{1}{\delta x} \frac{D(\delta x)}{Dt} + \frac{1}{\delta y} \frac{D(\delta y)}{Dt} + \frac{1}{\delta z} \frac{D(\delta z)}{Dt} \quad (13.82b)$$

$$= \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} + \frac{\delta w}{\delta z} \quad (13.82c)$$

$$= \nabla \cdot \mathbf{v}. \quad (13.82d)$$

Setting  $D(\rho \delta V)/Dt = 0$  leads to the expression from VOLUME 2 for the fluid mass continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \quad (13.83)$$

### 13.9.2 Generalized vertical coordinates

We follow the above procedure but now use generalized vertical coordinates so that

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \frac{1}{\delta x \delta y z_{,\sigma} \delta \sigma} \frac{D(\delta x \delta y z_{,\sigma} \delta \sigma)}{Dt} \quad (13.84a)$$

$$= \frac{1}{\delta x} \frac{D(\delta x)}{Dt} + \frac{1}{\delta y} \frac{D(\delta y)}{Dt} + \frac{1}{z_{,\sigma}} \frac{D(z_{,\sigma})}{Dt} + \frac{1}{\delta \sigma} \frac{D(\delta \sigma)}{Dt} \quad (13.84b)$$

$$= \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} + \frac{1}{z_{,\sigma}} \frac{D(z_{,\sigma})}{Dt} + \frac{\delta(\dot{\sigma})}{\delta \sigma} \quad (13.84c)$$

$$= \nabla_{\mathbf{v}} \cdot \mathbf{u} + \frac{1}{z_{,\sigma}} \frac{D(z_{,\sigma})}{Dt} + \frac{\partial \dot{\sigma}}{\partial \sigma}, \quad (13.84d)$$

where we introduced the shorthand

$$\dot{\sigma} = D\sigma/Dt. \quad (13.85)$$

We also set

$$\frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} = \nabla_{\mathbf{v}} \cdot \mathbf{u} \quad (13.86)$$

since we are working with generalized vertical coordinates so that we consider infinitesimal displacements occurring on constant  $\sigma$  surfaces. We are thus led to

$$\frac{1}{\rho \delta V} \frac{D(\rho \delta V)}{Dt} = \nabla_{\mathbf{v}} \cdot \mathbf{u} + \frac{1}{z_{,\sigma}} \frac{Dz_{,\sigma}}{Dt} + \frac{\partial \dot{\sigma}}{\partial \sigma} + \frac{1}{\rho} \frac{D\rho}{Dt} = 0. \quad (13.87)$$

Now use the material time derivative in the form (13.51b) to derive the flux-form expression of mass conservation

$$\frac{\partial(\rho z_{,\sigma})}{\partial t} + \nabla_{\mathbf{v}} \cdot (\rho z_{,\sigma} \mathbf{u}) + \frac{\partial(\rho z_{,\sigma} \dot{\sigma})}{\partial \sigma} = 0, \quad (13.88)$$

where the time derivative is computed holding  $\sigma$  fixed. We can furthermore introduce the dia-surface velocity component

$$w^{(\dot{\sigma})} = z_{,\sigma} \dot{\sigma} \quad (13.89)$$

so that mass continuity takes the form

$$\frac{\partial(\rho z_{,\sigma})}{\partial t} + \nabla_{\mathbf{v}} \cdot (\rho z_{,\sigma} \mathbf{u}) + \frac{\partial(\rho w^{(\dot{\sigma})})}{\partial \sigma} = 0. \quad (13.90)$$

Alternatively, we can reintroduce the material time derivative operator to write the mass continuity equation (13.88) in the form

$$\frac{1}{\rho z_{,\sigma}} \frac{D(\rho z_{,\sigma})}{Dt} = -(\nabla_{\mathbf{v}} \cdot \mathbf{u} + \partial \dot{\sigma} / \partial \sigma), \quad (13.91)$$

where we used equation (13.51b) to write

$$\frac{D}{Dt} = \left[ \frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_{\! \sigma} + \dot{\sigma} \frac{\partial}{\partial \sigma}. \quad (13.92)$$

## 13.10 Layer integrated mass continuity

The formulation thus far has been continuous, with the only assumption made that the specific thickness,  $h = \partial z / \partial \sigma$ , is single signed. We here consider a discrete increment in the generalized vertical coordinate,

$$\sigma - \delta\sigma/2 \leq \sigma' \leq \sigma + \delta\sigma/2, \quad (13.93)$$

and formulate the mass budget over this layer whose thickness is given by

$$h \equiv \int_{z(\sigma-\delta\sigma/2)}^{z(\sigma+\delta\sigma/2)} dz = \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \frac{\partial z}{\partial \sigma} d\sigma, \quad (13.94)$$

and whose mass per horizontal area is

$$\delta m = \int_{z(\sigma-\delta\sigma/2)}^{z(\sigma+\delta\sigma/2)} \rho dz = \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho z_{,\sigma} d\sigma = \bar{\rho} h, \quad (13.95)$$

where  $\bar{\rho}$  is the layer averaged density. Note that for a Boussinesq ocean the mass per area equals to the layer thickness times the reference density,  $\rho_0$ ,

$$\delta m = \rho_0 h \quad \text{Boussinesq ocean.} \quad (13.96)$$

As defined by equation (13.94) and illustrated in Figure 13.2, the thickness of a layer is relatively large in regions where  $\partial\sigma/\partial z$  is small; i.e., in regions where  $\sigma$  is weakly stratified in the vertical. Conversely, the layer thickness is relatively small where the vertical stratification is large. Furthermore, if the specific thickness is negative, then the layer thickness remains positive by choosing  $\delta\sigma < 0$ . For example, in a stably stratified fluid with  $\sigma$  given by potential density,  $\partial\sigma/\partial z = -(g/\rho_0) N^2 < 0$  so that we take  $\delta\sigma < 0$  to move vertically upward in the water column to regions of lower potential density. The same situation holds when  $\sigma$  is the hydrostatic pressure in which  $\partial p/\partial z = -\rho g$  (Section 13.10.2).

The formulation in this section, and its companion for tracers in Section 13.11, holds across all generalized vertical coordinates, even incorporating the trivial case of geopotential coordinates ( $\sigma = z$ ) whereby the specific thickness is unity. Application of the resulting layer integrated kinematics include the development of discrete equations for numerical layered models (see [Griffies et al. \(2020\)](#) for a review), as well as [shallow water models](#).

### 13.10.1 Non-Boussinesq (compressible) fluids

Performing a layer integral of the specific thickness equation (13.90) renders

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \left[ \frac{\partial(\rho z_{,\sigma})}{\partial t} + \nabla_{\! \sigma} \cdot (\rho z_{,\sigma} \mathbf{u}) + \frac{\partial(\rho w^{(\dot{\sigma})})}{\partial \sigma} \right] d\sigma = 0. \quad (13.97)$$

The dia-surface term integrates to a finite difference across the layer

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \left[ \frac{\partial(\rho z_{,\sigma})}{\partial t} + \nabla_{\mathbf{v}} \cdot (\rho z_{,\sigma} \mathbf{u}) \right] = -\Delta_{\sigma}(\rho w^{(\dot{\sigma})}), \quad (13.98)$$

where we introduced the dimensionless finite difference operator for properties defined at the layer interface

$$\Delta_{\sigma}(A) = A(\sigma + \delta\sigma/2) - A(\sigma - \delta\sigma/2). \quad (13.99)$$

The time derivative and horizontal space derivative commute with the layer integral, since the limits are specified fixed values for the layer increment,  $\delta\sigma$ , and the derivatives are computed with  $\sigma$  fixed. Hence, layer mass continuity takes the form

$$\left[ \frac{\partial}{\partial t} \right]_{\sigma} \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho z_{,\sigma} d\sigma + \nabla_{\mathbf{v}} \cdot \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho \mathbf{u} z_{,\sigma} d\sigma = -\Delta_{\sigma}(\rho w^{(\dot{\sigma})}). \quad (13.100)$$

The first term involves the layer averaged density times the layer thickness as per equation (13.95). The second term involves the layer averaged density-weighted velocity, which is the layer averaged horizontal mass flux

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho \mathbf{u} z_{,\sigma} d\sigma = h \bar{\rho} \bar{\mathbf{u}}. \quad (13.101)$$

We are thus led to the layer integrated continuity equation

$$\left[ \frac{\partial(h \bar{\rho})}{\partial t} \right]_{\sigma} + \nabla_{\mathbf{v}} \cdot (h \bar{\rho} \bar{\mathbf{u}}) + \Delta_{\sigma}(\rho w^{(\dot{\sigma})}) = 0. \quad (13.102)$$

When evolving the fields in a discrete numerical model, we have direct access to information only about layer averaged fields. So how do we estimate the depth average of the horizontal advective flux,  $\bar{\rho} \bar{\mathbf{u}}$ , appearing in equation (13.102)? One method interprets all fields as their layer averaged values so that  $\bar{\rho} \bar{\mathbf{u}} = \bar{\rho} \bar{\mathbf{u}}$ , thus considering uncomputed sub-layer correlations  $\bar{\rho}' \bar{\mathbf{u}'}$  as part of the truncation error. Alternately, we note that non-Boussinesq (compressible) hydrostatic flows can be described by a pressure-based vertical coordinate in which case the layer mass per horizontal area is proportional to a prescribed increment in pressure

$$\delta m = \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho z_{,\sigma} d\sigma = \bar{\rho} h = -g^{-1} \delta p. \quad (13.103)$$

Correspondingly, the layer integrated horizontal mass flux equals to the mass increment times the pressure-layer averaged velocity

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho \mathbf{u} z_{,\sigma} d\sigma = -g^{-1} \int_{p-\delta p/2}^{p+\delta p/2} \mathbf{u} dp = -g^{-1} \bar{\mathbf{u}} \delta p = h \bar{\rho} \bar{\mathbf{u}}. \quad (13.104)$$

With either of the above two methods, we are led to the same layer integrated continuity equation, which we write in the generic form that drops overbars

$$\left[ \frac{\partial(h \rho)}{\partial t} \right]_{\sigma} + \nabla_{\mathbf{v}} \cdot (h \rho \mathbf{u}) + \Delta_{\sigma}(\rho w^{(\dot{\sigma})}) = 0. \quad (13.105)$$

We illustrate contributions to this layer mass budget in Figure 13.7.

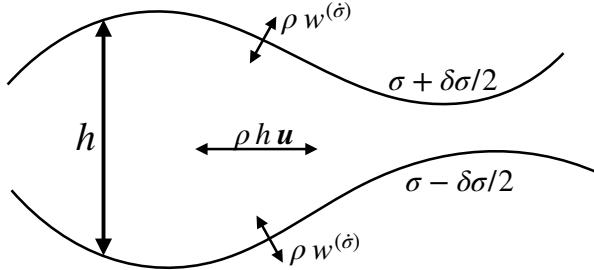


FIGURE 13.7: Illustrating the terms contributing to changes in layer mass according to the layer integrated continuity equation (13.105). The discrete layer is shown here with bounding interfaces at  $\sigma - \delta\sigma/2$  and  $\sigma + \delta\sigma/2$ . Within a layer there is a horizontal redistribution due to horizontal advective transport. Additionally, matter can cross the layer due to dia-surface transport via  $w^{(\delta)}$ .

### 13.10.2 Mass continuity using pressure coordinates

Let us here consider in some detail the special case of pressure coordinates in a hydrostatic flow, and thus derive mass continuity using these coordinates.

#### Method I

The thickness of a hydrostatic pressure layer (equation (13.94)) takes on the following form

$$h = \int_{p-\delta p/2}^{p+\delta p/2} \frac{\partial z}{\partial p} dp = - \int_{p-\delta p/2}^{p+\delta p/2} \frac{dp}{\rho g}, \quad (13.106)$$

so that its mass per unit area is

$$\int_{p-\delta p/2}^{p+\delta p/2} \rho \frac{\partial z}{\partial p} dp = -\delta p/g. \quad (13.107)$$

The mass continuity equation (13.105) thus becomes

$$\frac{\partial(\delta p)}{\partial t} + \nabla_{hp} \cdot (\mathbf{u} \delta p) + \Delta_p (\dot{p}) = 0. \quad (13.108)$$

The partial time derivative vanishes since it is computed by holding pressure fixed so that the pressure increment has a zero time tendency

$$\left[ \frac{\partial(\delta p)}{\partial t} \right]_p = 0. \quad (13.109)$$

Likewise,  $\nabla_{hp}(\delta p) = 0$ . Thus, we can divide by  $\delta p$  to render the continuity equation

$$\nabla_{hp} \cdot \mathbf{u} + \frac{\partial \dot{p}}{\partial p} = 0 \quad \text{compressible hydrostatic.} \quad (13.110)$$

This equation is isomorphic to the continuity equation for non-divergent flows written using geopotential coordinates

$$\nabla_h \cdot \mathbf{u} + \frac{\partial w}{\partial z} = \nabla_h \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad \text{non-divergent flow,} \quad (13.111)$$

where  $w = z$  is the vertical component to the velocity vector. For both pressure coordinates, describing non-Boussinesq fluids, and depth coordinates, describing Boussinesq fluids, the continuity equation is a diagnostic relation (i.e., no time derivatives) rather than prognostic (i.e., containing time derivatives).

## Method II

For the second method we make use of the approach detailed in Section 13.9.2, which starts from

$$\frac{D(\rho \delta V)}{Dt} = 0. \quad (13.112)$$

In pressure coordinates the volume of the fluid element takes the form

$$\delta V = \delta x \delta y \delta z = \delta x \delta y \left[ \frac{\partial z}{\partial p} \right] \delta p = -(\rho g)^{-1} \delta x \delta y \delta p. \quad (13.113)$$

Consequently,

$$0 = \frac{D(\rho \delta V)}{Dt} = g^{-1} \left( \frac{D(\delta x \delta y \delta p)}{Dt} \right), \quad (13.114)$$

so that

$$0 = \frac{1}{\delta x \delta y \delta p} \left( \frac{D(\delta x \delta y \delta p)}{Dt} \right) = \nabla_{hp} \cdot \mathbf{u} + \frac{\partial p}{\partial p}. \quad (13.115)$$

The second step made use of the isomorphism between this result and that for equation (13.111) that holds for a geopotential vertical coordinate.

### 13.10.3 Boussinesq ocean (non-divergent flow)

Specializing to a non-divergent flow of the Boussinesq ocean, where fluid elements conserve their volume, yields the layer thickness equation

$$\frac{\partial h}{\partial t} + \nabla_h \cdot (h \mathbf{u}) + \Delta_\sigma w^{(\dot{\sigma})} = 0. \quad (13.116)$$

Further specializing to the case of zero dia-surface transport leads to

$$\frac{\partial h}{\partial t} + \nabla_h \cdot (h \mathbf{u}) = 0 \quad \text{no dia-surface transport.} \quad (13.117)$$

This case is commonly studied for stratified perfect fluids flows using isopycnal coordinates, in which isopycnal surfaces are material (Section 14.2).

### 13.10.4 Rescaled geopotential coordinates

In Section 12.3.3 we introduced the rescaled geopotential coordinate

$$z^* = \frac{H(z - \eta)}{H + \eta} = \frac{\eta_b(z - \eta)}{\eta_b - \eta} \quad \text{and} \quad \eta_b(x, y) \leq z^* \leq 0, \quad (13.118)$$

which is commonly used in Boussinesq ocean models, where  $z = \eta(x, y, t)$  is the ocean free surface and  $z = \eta_b(x, y) = -H(x, y)$  is the ocean bottom. The thickness of a coordinate layer is given by

$$h = dz = \frac{\partial z}{\partial z^*} dz^* = (1 + \eta/H) dz^* = (1 - \eta/\eta_b) dz^*. \quad (13.119)$$

The depth integrated column thickness and depth integrated coordinate thickness are given by

$$\int_{\eta_b}^{\eta} dz = \eta - \eta_b = \eta + H \quad \text{and} \quad \int_{z^*(\eta_b)}^{z^*(\eta)} dz^* = -\eta_b = H. \quad (13.120)$$

Correspondingly, the depth integrated thickness equation is given by the depth integrated volume budget

$$\frac{\partial \eta}{\partial t} + \nabla \cdot \mathbf{U} + [w_{z^*=0}^{(\dot{\sigma})} - w_{z^*=\eta_b}^{(\dot{\sigma})}] = 0. \quad (13.121)$$

We typically assume no fluid flow through the solid-earth bottom, so that  $w_{z^*=\eta_b}^{(\dot{\sigma})} = 0$ , whereas

$$-\rho_o w_{z^*=0}^{(\dot{\sigma})} = Q_m \quad (13.122)$$

is the mass flux crossing the ocean free surface as studied in VOLUME 2.

## 13.11 Layer integrated tracer equation

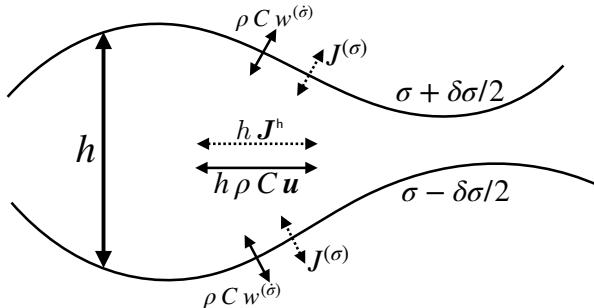


FIGURE 13.8: Illustrating the terms contributing to changes in layer tracer content according to the layer integrated tracer equation (13.127). The layer is shown here with bounding interfaces at  $\sigma - \delta\sigma/2$  and  $\sigma + \delta\sigma/2$ . Within a layer there is a redistribution of tracer due to horizontal advective and subgrid scale tracer fluxes. Additionally, matter can cross the layer due to dia-surface transport via  $\rho C w^{(\dot{\sigma})}$  and subgrid tracer transport  $J^{(\sigma)}$ .

The tracer equation is given by

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}, \quad (13.123)$$

where  $\mathbf{J}$  is a subgrid scale flux. Now introduce the material time derivative operator in the form (13.51b) to have

$$\rho \left[ \frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla_v C + \dot{\sigma} \partial_\sigma C \right] = -\nabla \cdot \mathbf{J}, \quad (13.124)$$

Multiplying by the specific thickness and making use of the mass conservation equation (13.90)

renders the flux-form conservation equation

$$\frac{\partial(z_{,\sigma} \rho C)}{\partial t} + \nabla_{\mathbf{v}} \cdot (z_{,\sigma} \rho C \mathbf{u}) + \frac{\partial(\rho C w^{(\dot{\sigma})})}{\partial \sigma} = - \left[ \nabla_{\mathbf{v}} \cdot (z_{,\sigma} \mathbf{J}^h) + \frac{\partial(z_{,\sigma} \nabla \sigma \cdot \mathbf{J})}{\partial \sigma} \right], \quad (13.125)$$

where we made use of expression (12.106) for the subgrid scale operator. Now perform a layer integral as detailed in Section 13.10 and use the layer mass continuity equation (13.105) to yield the layer integrated tracer equation

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_{\mathbf{v}} \cdot (h \rho C \mathbf{u}) + \Delta_{\sigma}(\rho C w^{(\dot{\sigma})}) = - [\nabla_{\mathbf{v}} \cdot (h \mathbf{J}^h) + \Delta_{\sigma}(z_{,\sigma} \nabla \sigma \cdot \mathbf{J})]. \quad (13.126)$$

Alternatively, we can bring all terms to the left hand side to yield

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_{\mathbf{v}} \cdot (h \rho C \mathbf{u} + h \mathbf{J}^h) + \Delta_{\sigma}(\rho C w^{(\dot{\sigma})} + J^{(\sigma)}) = 0 \quad (13.127)$$

where we wrote

$$J^{(\sigma)} = z_{,\sigma} \nabla \sigma \cdot \mathbf{J}. \quad (13.128)$$

We illustrate contributions to the layer tracer budget (13.127) in Figure 13.8. Note that we interpret these layer integrated fields and fluxes as per the discussion in Section 13.10.1.

## 13.12 Overturning circulation in the meridional- $\sigma$ plane

In VOLUME 2 we studied the meridional-depth streamfunction. Here we introduce a streamfunction defined according to generalized vertical coordinate,  $\sigma(x, y, z, t)$ . This generalization is useful for studying the zonally integrated circulation partitioned according to  $\sigma$  surfaces rather than  $z$  surfaces, in particular when  $\sigma$  is potential density or specific entropy. We make use of Figure 13.9 in the following. We can make use of the Boussinesq ocean, where the flow field is non-divergent. Alternatively, we make use of a non-Boussinesq fluid when the flow is steady. As shown in this section, we are afforded a meridional- $\sigma$  streamfunction only when  $\sigma$  is time independent.

### 13.12.1 Overturning streamfunction

Start from our discussion in VOLUME 2 for the meridional-depth streamfunction, here generalized to

$$\Psi(y, \sigma, t) = - \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} \left[ \int_{\eta_b(x', y)}^{z=\eta(x', y, \sigma, t)} v(x', y, z', t) dz' \right] dx'. \quad (13.129)$$

As defined,  $\Psi(y, \sigma, t)$  makes use of  $z = \eta(x', y, \sigma, t)$  for the upper bound on the vertical integral, where  $\eta(x', y, \sigma, t)$  is the vertical position of a generalized vertical coordinate surface with value  $\sigma$ , such as depicted in Figure 13.9. In this manner,  $\Psi(y, \sigma, t)$ , is a function of latitude,  $\sigma$ , and time. Our job in the next subsection is to prove that  $\Psi(y, \sigma, t)$  indeed serves as a streamfunction for the zonally integrated flow, with the zonal integral along constant  $\sigma$  surfaces rather than constant  $z$  surfaces.

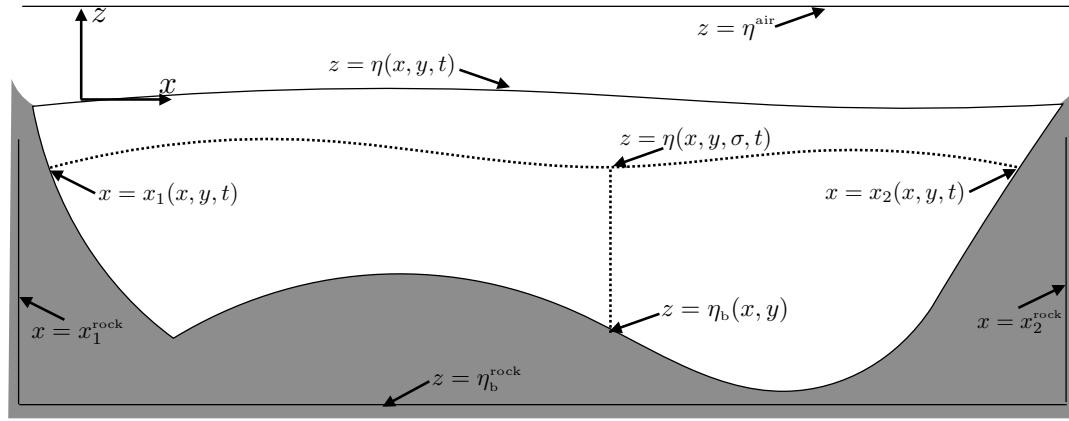


FIGURE 13.9: Geometry for computing the meridional streamfunction, here generalizing the meridional-depth streamfunction geometry from VOLUME 2 to meridional- $\sigma$ . The zonal boundaries are written  $x = x_1(y, z, t)$  and  $x = x_2(y, z, t)$ , which are generally functions of latitude and vertical position as well as time. The bottom is written as  $z = \eta_b(x, y)$  and the vertical position of an arbitrary surface is written  $z = \eta(x, y, \sigma, t)$ , where we assume this surface is monotonic in the vertical as per the usual assumption for generalized vertical coordinates. Note that for shorthand, we often find it convenient to write  $z = \eta_\sigma(x, y, t)$ , particularly when suppressing the dependencies on horizontal position and time. We also display the constant zonal positions,  $x_{1,2}^{\text{rock}}$ , which are fully within the rock, as well as the bottom position,  $\eta_b^{\text{rock}}$ , which is also within the rock.

### 13.12.2 Proving that $\Psi(y, \sigma, t)$ is a streamfunction

To prove that  $\Psi(y, \sigma, t)$  is indeed a streamfunction, we proceed much like in VOLUME 2 for the meridional-depth streamfunction,  $\Psi(y, z, t)$ , with the key new piece in the derivation concerning the space-time dependence of the  $z = \eta(x, y, \sigma, t)$  surface. The vertical derivative of the streamfunction is given by

$$\frac{\partial \Psi}{\partial z} = -\frac{\partial}{\partial z} \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} \left[ \int_{\eta_b(x', y)}^{z=\eta(x', y, \sigma, t)} v(x', y, z', t) dz' \right] dx' \quad (13.130a)$$

$$= - \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} \left[ \frac{\partial}{\partial z} \int_{\eta_b(x', y)}^{z=\eta(x', y, \sigma, t)} v(x', y, z', t) dz' \right] dx' \quad (13.130b)$$

$$= - \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} v(x', y, z = \eta(x', y, \sigma, t)) dx' \quad (13.130c)$$

$$= -V(y, \sigma, t). \quad (13.130d)$$

In these steps we used Leibniz's rule and noted that only the upper integration limit is a function of  $z$ . Furthermore, the upper limit on the vertical integral is evaluated at the vertical position of the  $\sigma$  surface. Hence, the zonal integral is defined while keeping the vertical position on the  $\sigma$  surface rather than on a constant geopotential surface.

For the meridional derivative we have

$$\frac{\partial \Psi}{\partial y} = - \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} \left[ \frac{\partial}{\partial y} \int_{\eta_b(x', y)}^{z=\eta(x', y, \sigma, t)} v(x', y, z', t) dz' \right] dx'. \quad (13.131)$$

Focusing on the vertical integral yields (dropping various coordinate dependencies when not

essential)

$$\frac{\partial}{\partial y} \int_{\eta_b(x',y)}^{z=\eta(x',y,\sigma)} v(x',y,z') dz' = v(\eta_\sigma) \partial_y \eta_\sigma - v(\eta_b) \partial_y \eta_b + \int_{\eta_b(x',y)}^{z=\eta(x',y,\sigma)} \partial_y v(x',y,z') dz'. \quad (13.132)$$

Focus again on the vertical integral and make use of the non-divergence condition to yield

$$\int_{\eta_b(x',y)}^{z=\eta(x',y,\sigma)} \partial_y v(x',y,z') dz' = - \int_{\eta_b(x',y)}^{z=\eta(x',y,\sigma)} [\partial_{x'} u(x',y,z') + \partial_{z'} w(x',y,z')] dz'. \quad (13.133)$$

Leibniz's rule on the  $\partial_{x'} u(x',y,z')$  term then brings us to

$$\begin{aligned} \frac{\partial}{\partial y} \int_{\eta_b(x',y)}^{z=\eta(x',y,\sigma)} v(x',y,z') dz' &= -[w(\eta_\sigma) - \mathbf{u}(\eta_\sigma) \cdot \nabla \eta_\sigma] + [w(\eta_b) - \mathbf{u}(\eta_b) \cdot \nabla \eta_b] \\ &\quad - \frac{\partial}{\partial x} \int_{\eta_b(x',y)}^{z=\eta(x',y,\sigma)} u(x',y,z') dz'. \end{aligned} \quad (13.134)$$

Now  $w(\eta_b) - \mathbf{u}(\eta_b) \cdot \nabla \eta_b = 0$  from the bottom no-flow **kinematic boundary condition**. Furthermore, from equation (13.35f) for the dia-surface velocity, we have

$$w(\eta_\sigma) - \mathbf{u}(\eta_\sigma) \cdot \nabla \eta_\sigma = w^{(\dot{\sigma})} + \partial_t \eta_\sigma. \quad (13.135)$$

Bringing these results together then renders

$$\frac{\partial \Psi}{\partial y} = \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} [w^{(\dot{\sigma})}(x',y,z=\eta_\sigma) + \partial_t \eta_\sigma] dx', \quad (13.136)$$

where we set

$$u(x_1^{\text{rock}}) = u(x_2^{\text{rock}}) = 0. \quad (13.137)$$

We conclude that  $\Psi(y, \sigma, t)$  is a streamfunction for the special case where  $\sigma$  is time independent, in which case

$$\frac{\partial \Psi}{\partial y} = \int_{x_1^{\text{rock}}}^{x_2^{\text{rock}}} w^{(\dot{\sigma})}(x',y,z=\eta_\sigma) dx' = W^{(\dot{\sigma})}(y, \sigma, t). \quad (13.138)$$

We can understand the need for time independence since that ensures that the flow underneath the  $\sigma$  surface is non-divergent, much like the case for a shallow water model in steady state. In the literature, one can find  $\Psi(y, \sigma, t)$  referred to as a streamfunction even when the flow has time dependence, in which case extra caution is needed if inferring the associated flow patterns.

## 13.13 Equations of motion

We here derive the equations of motion based on generalized vertical coordinates. The scalar equations were already discussed in Sections 13.9, 13.10, and 13.11, so we only summarize these equations here.

### 13.13.1 Mass and tracer equations

The mass and tracer equations were derived in Sections 13.9, 13.10, and 13.11, with their continuous vertical coordinate formulation given by

$$\frac{\partial(\rho h)}{\partial t} + \nabla_{\sigma} \cdot (\rho h \mathbf{u}) + \partial_{\sigma}(\rho h \dot{\sigma}) = 0 \quad (13.139a)$$

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_{\sigma} \cdot (h \rho C \mathbf{u} + h \mathbf{J}^h) + \partial_{\sigma}(h \dot{\sigma} C + h \nabla \sigma \cdot \mathbf{J}) = 0. \quad (13.139b)$$

Compatibility is maintained between the mass continuity equation (13.139a) and the tracer equation (13.139b) so long as the tracer equation reduces to the mass equation upon setting the tracer concentration to a spatial constant. Hence, for compatibility we must have the subgrid fluxes,  $\mathbf{J}$ , vanish when the tracer is a spatial constant. For example, diffusive fluxes, which are proportional to the tracer gradient, respect this constraint. These properties originate from our discussion of mass budgets and the **barycentric velocity** in VOLUME 2.

### 13.13.2 Momentum equation

From VOLUME 2, the horizontal and vertical components to the momentum equation are

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \times \mathbf{u} = -\rho \nabla_h \Phi - \nabla_h p + \rho \mathbf{F}^h \quad (13.140a)$$

$$\rho \frac{Dw}{Dt} = -\rho \frac{\partial \Phi}{\partial z} - \frac{\partial p}{\partial z} + \rho F^z. \quad (13.140b)$$

The simple form of the geopotential sets  $\Phi = g z$ , so that the horizontal gradient of the geopotential vanishes

$$\Phi = g z \implies \nabla_h \Phi = 0. \quad (13.141)$$

However, this gradient is nonzero in the presence of astronomical tide.

#### Horizontal momentum equation

We transform the horizontal derivatives from geopotential coordinates to generalized vertical coordinates according to (see equation (12.85))

$$\nabla_h = \nabla_{\sigma} - (\nabla_{\sigma} z) \partial_z, \quad (13.142)$$

thus leading to the horizontal momentum equation

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \times \mathbf{u} = -\rho [\nabla_{\sigma} - (\nabla_{\sigma} z) \partial_z] \Phi - [\nabla_{\sigma} - (\nabla_{\sigma} z) \partial_z] p + \rho \mathbf{F}^h. \quad (13.143)$$

In Section 13.13.5 we present some special cases for this equation that simplify the pressure and geopotential terms.

#### Vertical momentum equation

The vertical momentum equation is transformed into

$$\rho \frac{Dw}{Dt} = -\frac{\partial \sigma}{\partial z} \left[ \rho \frac{\partial \Phi}{\partial \sigma} + \frac{\partial p}{\partial \sigma} \right] + \rho F^z, \quad (13.144)$$

with the hydrostatic form given by

$$\frac{\partial p}{\partial \sigma} = -\rho \frac{\partial \Phi}{\partial \sigma}. \quad (13.145)$$

### 13.13.3 Flux-form horizontal momentum equation

Using Cartesian horizontal coordinates and generalized vertical coordinates, the horizontal momentum equation includes a contribution from the acceleration that has a form similar to that for a tracer (Section 13.11)

$$h \rho \frac{Du}{Dt} = \left[ \frac{\partial(h \rho u)}{\partial t} \right]_{\sigma} + \nabla_{\mathbf{h}\sigma} \cdot (h \rho u \mathbf{u}) + \partial_{\sigma}(h \rho u \dot{\sigma}) \quad (13.146a)$$

$$h \rho \frac{Dv}{Dt} = \left[ \frac{\partial(h \rho v)}{\partial t} \right]_{\sigma} + \nabla_{\mathbf{h}\sigma} \cdot (h \rho v \mathbf{u}) + \partial_{\sigma}(h \rho v \dot{\sigma}). \quad (13.146b)$$

We provide a  $\sigma$  subscript on the time derivative operator to signal that this derivative is taken with  $\sigma$  held fixed. With spherical coordinates there are additional terms appearing on the right hand side (VOLUME 2). In particular, there is a metric term that contains the vertical velocity component,  $w = Dz/Dt$ . The appearance of  $w$  is awkward since the vertical velocity is not naturally computed using generalized vertical coordinates. This limitation is overcome through use of the vector-invariant velocity equation derived in Section 13.13.4.

### 13.13.4 Vector-invariant horizontal momentum equation

The vector invariant form of the velocity equation eliminates the metric terms that appear in the non-Cartesian flux-form equations. The vector-invariant form is also suited for deriving the vorticity equation considered in Section 13.15. Here, we start with the material time derivative in the form (13.51c) appropriate for generalized vertical coordinates, in which case the horizontal acceleration is given by

$$\frac{D\mathbf{u}}{Dt} = \left[ \frac{\partial \mathbf{u}}{\partial t} \right]_{\sigma} + (\mathbf{u} \cdot \nabla_{\mathbf{h}\sigma}) \mathbf{u} + (\dot{\sigma} \partial_{\sigma}) \mathbf{u}. \quad (13.147)$$

Now make use of the vector identity (VOLUME 1)

$$(\mathbf{u} \cdot \nabla_{\mathbf{h}\sigma}) \mathbf{u} = \nabla_{\mathbf{h}\sigma} K + (\nabla_{\mathbf{h}\sigma} \times \mathbf{u}) \times \mathbf{u}, \quad (13.148)$$

where

$$K = \mathbf{u} \cdot \mathbf{u}/2 \quad (13.149)$$

is the kinetic energy per mass of the horizontal flow. Introducing the generalized vertical coordinate version of the relative vorticity

$$\tilde{\zeta} \equiv \hat{\mathbf{z}} \cdot (\nabla_{\mathbf{h}\sigma} \times \mathbf{u}) = \left[ \frac{\partial v}{\partial x} \right]_{\sigma} - \left[ \frac{\partial u}{\partial y} \right]_{\sigma} \quad (13.150)$$

renders

$$\frac{D\mathbf{u}}{Dt} = \left[ \frac{\partial \mathbf{u}}{\partial t} \right]_{\sigma} + \nabla_{\mathbf{h}\sigma} K + \tilde{\zeta} \hat{\mathbf{z}} \times \mathbf{u} + \dot{\sigma} \partial_{\sigma} \mathbf{u}, \quad (13.151)$$

so that the horizontal momentum equation takes the vector-invariant form

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + \dot{\sigma} \frac{\partial \mathbf{u}}{\partial \sigma} + (2\boldsymbol{\Omega} + \hat{\mathbf{z}} \tilde{\zeta}) \times \mathbf{u} = -\nabla_{\!h} K - \nabla_{\!h} \Phi - (1/\rho) \nabla_{\!h} p + \mathbf{F}^h, \quad (13.152)$$

where again  $\nabla_{\!h} = \nabla_{\!h} - (\nabla_{\!h} z) \partial_z$  as per equation (13.142). This equation is form-invariant regardless the horizontal coordinates, thus motivating the name *vector-invariant*.<sup>1</sup>

### 13.13.5 Hydrostatic flow with constant gravitational acceleration

There are many special cases that simplify various terms in the momentum equation. For example, when considering a geopotential in the simple form,  $\Phi = g z$  with  $g$  assumed to be a constant effective gravitational acceleration, then the horizontal momentum equation (13.143) becomes

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \times \mathbf{u} = -[\nabla_{\!h} - (\nabla_{\!h} z) \partial_z] p + \rho \mathbf{F}^h. \quad (13.153)$$

Furthermore, assuming the flow maintains a hydrostatic balance (and corresponding simplification of the Coriolis acceleration as per the [Traditional Approximation](#)) allows us to write  $\partial p / \partial z = -g \rho$  so that

$$\rho \frac{D\mathbf{u}}{Dt} + \rho f \hat{\mathbf{z}} \times \mathbf{u} = -[\nabla_{\!h} p + \rho \nabla_{\!h} \Phi] + \rho \mathbf{F}^h, \quad (13.154)$$

which also takes on the vector-invariant form

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + \dot{\sigma} \frac{\partial \mathbf{u}}{\partial \sigma} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_{\!h} (K + \Phi) - (1/\rho) \nabla_{\!h} p + \mathbf{F}^h. \quad (13.155)$$

This form is commonly used for hydrostatic models of the ocean and atmosphere, such as discussed in [Griffies et al. \(2020\)](#).

## 13.14 Concerning the pressure force

As studied in VOLUME 2, the pressure force acting on a fluid region is given by the integral

$$\mathbf{F}^{\text{press}} = - \oint_{\partial\mathcal{R}} p \hat{\mathbf{n}} d\mathcal{S} = - \int_{\mathcal{R}} \nabla p dV, \quad (13.156)$$

where the second equality follows from the [divergence theorem](#) applied to a scalar field. We refer to the right-most expression as the pressure gradient body force, and this expression is the basis for the discussion in Sections 13.13.5 and 13.14.1. In this formulation, the pressure force at a point is oriented down the pressure gradient, so that the net pressure force acting on a region is the volume integral of the pressure gradient.

The middle expression in equation (13.156) formulates the pressure force acting on a region as the area integrated pressure contact stress acting on the region boundaries, with the orientation of the force determined by the inward normal at each point on the boundary. Much of this section is concerned with the contact stress expression as the basis for computing the pressure force acting on a finite region as shown in Figure 13.11. The contact stress perspective was taken by [Lin \(1997\)](#) and [Adcroft et al. \(2008\)](#) in their finite volume approach to computing the pressure force acting on a finite numerical model grid cell.

---

<sup>1</sup>See Section 4.4.4 of [Griffies \(2004\)](#) for a detailed derivation using arbitrary horizontal coordinates.

### 13.14.1 Computing the horizontal pressure gradient

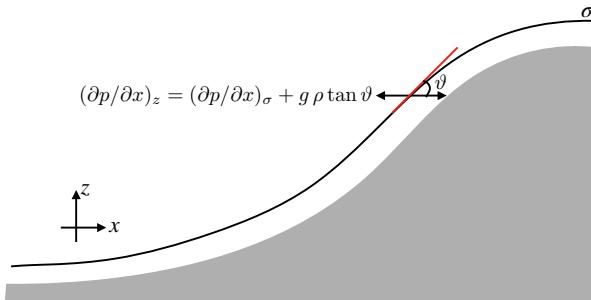


FIGURE 13.10: Illustrating how the horizontal pressure gradient is decomposed into two terms, one aligned with the surface of constant  $\sigma$ , and another associated with the slope of the  $\sigma$ -surface relative to the horizontal,  $\tan \vartheta = (\partial z / \partial x)_\sigma$ . We here consider the decomposition using terrain following vertical coordinates, where the vertical coordinate is aligned according to the solid-earth bottom (shaded region). Specifically, for terrain following Boussinesq ocean models we set  $\sigma = (z - \eta)/(-\eta_b + \eta)$ , where  $z = \eta(x, y, t)$  is the ocean free surface and  $z = \eta_b(x, y)$  is the ocean bottom topography. Terrain-following atmospheric models have a similar definition, often using pressure rather than geopotential so that  $\sigma = (p - p_a)/(p_b - p_a)$ , where  $p$  is the pressure,  $p_a = p_a(x, y, t)$  is the pressure applied at the top of the atmosphere (typically assumed to be zero), and  $p_b = p_b(x, y, t)$  is the pressure at the bottom of the atmosphere.

The horizontal pressure gradient is aligned perpendicular to the local gravitational direction, and it is generally among the dominant horizontal forces acting on a fluid element. Hence, its accurate representation in numerical models is crucial for the physical integrity of a simulation. Unfortunately, decomposition of the horizontal pressure gradient into two terms according to the transformation (13.142) can lead to numerical difficulties. For example, with a simple geopotential and a hydrostatic flow, equation (13.154) shows that the horizontal pressure gradient takes the form

$$\nabla_h p = \nabla_{\sigma} p + \rho \nabla_{\sigma} \Phi = \nabla_{\sigma} p + g \rho \nabla_{\sigma} z, \quad (13.157)$$

with this decomposition illustrated in Figure 13.10 for the case of terrain following vertical coordinates. Numerical difficulties occur when the two terms on the right hand side have comparable magnitude but distinct signs. We are thus confronted with computing the small difference between two large numbers, and that situation generally exposes a numerical simulation to nontrivial truncation errors. Unfortunately, these errors can corrupt the integrity of the computed pressure forces and in turn contribute to spurious flow. An overview of this issue for ocean models is given by [Haney \(1991\)](#), [Mellor et al. \(1998\)](#), [Griffies et al. \(2000\)](#), with advances offered by [Lin \(1997\)](#), [Shchepetkin and McWilliams \(2002\)](#), and [Adcroft et al. \(2008\)](#).

In the remainder of this section, we outline a finite volume method for computing the pressure force as proposed by [Lin \(1997\)](#) for atmosphere models and [Adcroft et al. \(2008\)](#) for ocean models. This approach starts from the middle expression in equation (13.156) for the pressure force; i.e., it formulates the pressure force as the area integral of the pressure contact force rather than the volume integral of the pressure gradient force. Throughout this discussion we consider the proto-typical model grid cell as depicted in Figure 13.11, and we make use of the discussion of pressure [form stress](#) in VOLUME 2.

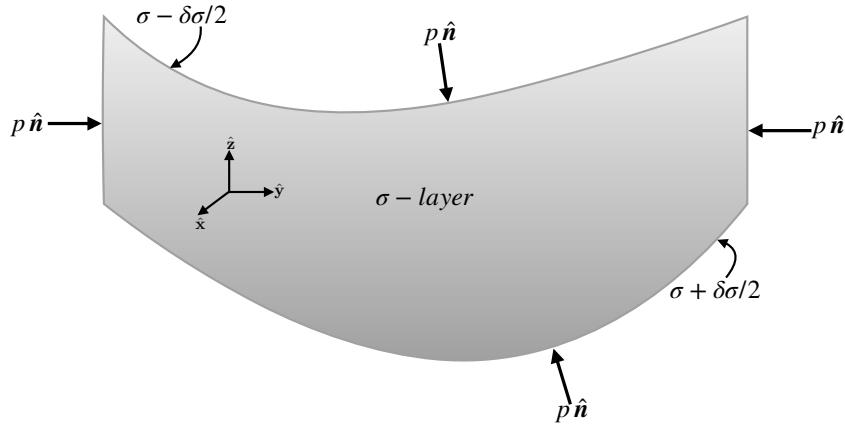


FIGURE 13.11: Schematic of pressure forces acting on the boundaries of a finite fluid region such as a discrete model grid cell. In generalized vertical coordinate models, the side faces are vertical, so that pressure acts only in the horizontal directions. The top and bottom faces are defined by surfaces of constant generalized vertical coordinates,  $\sigma(x, y, z, t) = \text{constant}$ , and we assume that these surfaces have an outward normal that has a nonzero projection into the vertical. As a result, we can write the depth of a point on the top and bottom surfaces as  $z = \eta(x, y, \sigma, t)$ . Because of the slope of the top and bottom surfaces, the surface normal vector has both a horizontal and vertical component. The net pressure force vector acting on the grid cell is given by the area integral of the pressure contact stress that acts on the grid cell boundary.

### 13.14.2 Integrated pressure contact stress on the cell faces

The inward normal on the grid cell vertical side boundaries points in the horizontal direction. For example, on the left side of Figure 13.11 the pressure force acts in the positive  $\hat{y}$  direction

$$\mathbf{F}_{\text{left}}^{\text{press}} = \hat{y} \int_{\text{left}} p \, dx \, dz \quad (13.158)$$

whereas pressure force on the right wall acts in the opposite direction

$$\mathbf{F}_{\text{right}}^{\text{press}} = -\hat{y} \int_{\text{right}} p \, dx \, dz. \quad (13.159)$$

Similar expressions appear for the front and back vertical boundaries acting in the  $\pm \hat{x}$  directions.

Since the top and bottom boundaries of the grid cell are sloped, there is a pressure force acting on this surface directed in the horizontal and vertical directions. To unpack this force, write the vertical position of a point on the top interface as  $z = \eta(x, y, t)$  so that the outward normal is given by

$$\hat{n} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{z} - \nabla\eta}{\sqrt{1 + |\nabla\eta|^2}}. \quad (13.160)$$

Following our discussion of dia-surface transport in Section 13.4.5, we know that the product of the normal direction and the area element can be written

$$\hat{n} \, dS = (\hat{z} - \nabla\eta) \, dA, \quad (13.161)$$

where  $dA = dx \, dy$  is the horizontal projection of the area element (see Figure 13.3). Hence, the net pressure force acting on the top face is given by

$$\mathbf{F}_{\text{top}}^{\text{press}} = \int_{\text{top}} p (-\hat{z} + \nabla_h z) \, dx \, dy, \quad (13.162)$$

where we set  $z = \eta$  in the second and third terms and recall that the horizontal slope vector for the top surface is given by

$$\nabla_{\text{h}} z = \hat{\mathbf{x}} (\partial z / \partial x)_{\sigma} + \hat{\mathbf{y}} (\partial z / \partial y)_{\sigma}, \quad (13.163)$$

where the  $\sigma$  subscript emphasizes that the horizontal derivative is taken with  $\sigma$  held constant. Notice that the pressure acts in the positive horizontal direction if the top surface slopes upward (surface shoaling) when directed in either of the two horizontal directions. Pressure acting on the bottom face has the same appearance yet with opposite signs

$$\mathbf{F}_{\text{bott}}^{\text{press}} = \int_{\text{bott}} p (\hat{\mathbf{z}} - \nabla_{\text{h}} z) dx dy. \quad (13.164)$$

The pressure acts in the positive horizontal direction if the bottom surface slopes downward (surface deepens) when directed in either of the two horizontal directions. As discussed in VOLUME 2, the horizontal pressure acting on a sloped surface is known as **form stress**. Here the sloped surface is defined by a constant generalized vertical coordinate.

### 13.14.3 Net vertical pressure force

Bringing the pieces together leads to the net vertical pressure force acting on the grid cell

$$\mathbf{F}_{\text{vertical}}^{\text{press}} = -\hat{\mathbf{z}} \left[ \int_{\text{top}} p dx dy - \int_{\text{bott}} p dx dy \right]. \quad (13.165)$$

If the flow maintains a hydrostatic balance, then this vertical force is given by the weight of fluid within the cell

$$\mathbf{F}_{\text{vertical}}^{\text{press}} = \hat{\mathbf{z}} Mg, \quad (13.166)$$

where  $M$  is the mass of fluid in the grid cell. The net vertical hydrostatic pressure force acts vertically upward since hydrostatic pressure at the cell bottom is greater than at the cell top. This upward hydrostatic pressure force acting on the cell balances the downward weight of fluid in the cell.

### 13.14.4 Net horizontal pressure force

The net meridional pressure force is given by the forces acting on the sides as well as those acting on the sloped top and bottom boundaries. Taking the specific case of the  $\hat{\mathbf{y}}$  pressure force yields

$$\begin{aligned} \hat{\mathbf{y}} \cdot \mathbf{F}^{\text{press}} &= \left[ \int_{\text{left}} p dx dz - \int_{\text{right}} p dx dz \right] \\ &\quad + \left[ \int_{\text{top}} p (\partial z / \partial y)_{\sigma} dx dy - \int_{\text{bott}} p (\partial z / \partial y)_{\sigma} dx dy \right]. \end{aligned} \quad (13.167)$$

We can write this expression in a more compact form by orienting the integration in a counter-clockwise manner around the cell boundaries, and making use of the identity  $(\partial z / \partial y)_{\sigma} dy = dz$  on the top and bottom faces, so that

$$\hat{\mathbf{y}} \cdot \mathbf{F}^{\text{press}} = - \oint p dx dz. \quad (13.168)$$

For some purposes it is more convenient to work with the geopotential,  $\Phi = g z$ , than the pressure. In this case we can write the pressure force as

$$\hat{\mathbf{y}} \cdot \mathbf{F}^{\text{press}} = - \oint p \, dx \, dz = - \oint dx [d(pz) - z \, dp] = g^{-1} \oint \Phi \, dx \, dp, \quad (13.169)$$

where  $\oint dx d(pz) = 0$ . This form is useful with non-Boussinesq (compressible) flows, in which pressure is a natural vertical coordinate (e.g., see the caption to Figure 13.10). Notably, equation (13.169) holds whether the flow is hydrostatic or non-hydrostatic.

### 13.14.5 Comments

A numerical realization of the integrated contact pressure stress requires a representation of the pressure field along the boundaries of the grid cell. A variety of methods are available with differing accuracies. [Adcroft et al. \(2008\)](#) are notable in proposing an analytic form that allows for an exact integration along the cell faces in special cases, and a highly accurate numerical integration in other cases. In general, this method for computing pressure forces is highly suited to generalized vertical coordinate grid cells, which was the motivation offered by [Lin \(1997\)](#) in the context of terrain following atmospheric models.

## 13.15 Hydrostatic vorticity and potential vorticity

Generalized vertical coordinates are most commonly used to study flow that maintains the hydrostatic balance. We are thus motivated to develop the evolution equation for the vertical component of vorticity,  $\tilde{\zeta}$ , in a hydrostatic flow as written using generalized vertical coordinates. By extension, we derive the budget equation for the corresponding potential vorticity.

### 13.15.1 Basic manipulations

Recall the vector-invariant velocity equation given by equation (13.155)

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + \dot{\sigma} \frac{\partial \mathbf{u}}{\partial \sigma} + \tilde{\zeta}_a \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_{\text{h}}(K + \Phi) - (1/\rho) \nabla_{\text{h}} p + \mathbf{F}^h, \quad (13.170)$$

where  $\tilde{\zeta}_a = \tilde{\zeta} + f$  is the absolute vorticity. Taking the curl of this equation and projecting onto the vertical direction leads to the vorticity equation

$$\frac{D\tilde{\zeta}_a}{Dt} = -\tilde{\zeta}_a \nabla_{\text{h}} \cdot \mathbf{u} + \frac{\hat{\mathbf{z}} \cdot (\nabla_{\text{h}} \rho \times \nabla_{\text{h}} p)}{\rho^2} + \hat{\mathbf{z}} \cdot \left[ \frac{\partial \mathbf{u}}{\partial \sigma} \times \nabla_{\text{h}} \dot{\sigma} + \nabla_{\text{h}} \times \mathbf{F}^h \right] \quad (13.171)$$

where we noted that the planetary vorticity,  $f$ , is independent of time and vertical position.

### Making use of the mass conservation equation

Mass conservation in the form of equation (13.91)

$$\frac{1}{\rho h} \frac{D(\rho h)}{Dt} = -(\nabla_{\text{h}} \cdot \mathbf{u} + \partial \dot{\sigma} / \partial \sigma), \quad (13.172)$$

renders

$$\rho h \frac{D}{Dt} \left[ \frac{\tilde{\zeta}_a}{\rho h} \right] = \frac{\hat{z} \cdot (\nabla_{\text{h}} \rho \times \nabla_{\text{h}} p)}{\rho^2} + \tilde{\zeta}_a \frac{\partial \dot{\sigma}}{\partial \sigma} + \hat{z} \cdot \left[ \frac{\partial \mathbf{u}}{\partial \sigma} \times \nabla_{\text{h}} \dot{\sigma} + \nabla_{\text{h}} \times \mathbf{F}^h \right]. \quad (13.173)$$

### Massaging the $\dot{\sigma}$ terms

The terms containing  $\dot{\sigma}$  can be written in the form

$$\tilde{\zeta}_a \partial_\sigma \dot{\sigma} + \hat{z} \cdot (\partial_\sigma \mathbf{u} \times \nabla_{\text{h}} \dot{\sigma}) = \tilde{\zeta}_a \partial_\sigma \dot{\sigma} + \hat{z} \cdot [-\nabla_{\text{h}} \times (\dot{\sigma} \partial_\sigma \mathbf{u}) + \dot{\sigma} \nabla_{\text{h}} \times \partial_\sigma \mathbf{u}] \quad (13.174a)$$

$$= \tilde{\zeta}_a \partial_\sigma \dot{\sigma} + \dot{\sigma} \partial_\sigma \tilde{\zeta}_a - \hat{z} \cdot [\nabla_{\text{h}} \times (\dot{\sigma} \partial_\sigma \mathbf{u})] \quad (13.174b)$$

$$= \partial_\sigma (\dot{\sigma} \tilde{\zeta}_a) - \hat{z} \cdot [\nabla_{\text{h}} \times (\dot{\sigma} \partial_\sigma \mathbf{u})]. \quad (13.174c)$$

### 13.15.2 Vorticity and potential vorticity equation

The above results allow us to write equation (13.173) in the form

$$\rho h \frac{DQ}{Dt} = \frac{\hat{z} \cdot (\nabla_{\text{h}} \rho \times \nabla_{\text{h}} p)}{\rho^2} + \partial_\sigma (\dot{\sigma} \tilde{\zeta}_a) + \nabla_{\text{h}} \cdot [\hat{z} \times \dot{\sigma} \partial_\sigma \mathbf{u} - \hat{z} \times \mathbf{F}^h], \quad (13.175)$$

where we introduced the potential vorticity defined according to the generalized vertical coordinates

$$Q = \frac{\tilde{\zeta}_a}{\rho h}. \quad (13.176)$$

The potential vorticity equation (13.175) has a generally nonzero baroclinicity

$$\frac{\hat{z} \cdot (\nabla_{\text{h}} \rho \times \nabla_{\text{h}} p)}{\rho^2}, \quad (13.177)$$

so that the potential vorticity (13.176) is generally not materially invariant even if  $\dot{\sigma} = 0$  and  $\mathbf{F}^h = 0$ . Finally, note that it is sometimes convenient to make use of the potential vorticity (13.176) in the horizontal velocity equation (13.170) so that

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + \dot{\sigma} \frac{\partial \mathbf{u}}{\partial \sigma} + (h \rho Q) \hat{z} \times \mathbf{u} = -\nabla_{\text{h}} (K + \Phi) - (1/\rho) \nabla_{\text{h}} p + \mathbf{F}^h. \quad (13.178)$$

### Pressure coordinates

The baroclinicity (13.177) vanishes when choosing  $\sigma = p$ , as noted when studying baroclinicity in Chapter 6. However, pressure is not a useful scalar for defining potential vorticity since  $\dot{\sigma} = \dot{p}$  does not generally vanish for a perfect fluid. Namely, a nonzero  $\dot{p}$  merely signals vertical motion, so that  $\dot{p} \neq 0$  for both real and perfect fluids. Hence, even though the baroclinicity vanishes by choosing  $\sigma = p$ , the terms with  $\dot{\sigma}$  do not.

### Flux-form potential vorticity budget

Just like we did in Section 13.13.3 for the velocity equation, we can make use of the thickness equation (13.139a) to bring the material time derivative in equation (13.175) into the form

$$\rho h \frac{DQ}{Dt} = \rho h (\partial_t + \mathbf{u} \cdot \nabla_{\text{h}} + \dot{\sigma} \partial_\sigma) Q + Q [\partial_t (\rho h) + \nabla_{\text{h}} \cdot (\rho h \mathbf{u}) + \partial_\sigma (\rho h \dot{\sigma})] \quad (13.179a)$$

$$= \partial_t(\rho h Q) + \nabla_{\mathbf{v}} \cdot (\rho h \mathbf{u} Q) + \partial_\sigma(\rho h \dot{\sigma} Q). \quad (13.179b)$$

Since  $\tilde{\zeta}_a = h \rho Q$ , we see that the term

$$\partial_\sigma(\rho h \dot{\sigma} Q) = \partial_\sigma(\dot{\sigma} \tilde{\zeta}_a), \quad (13.180)$$

also appears on the right hand side of equation (13.175). Hence, it cancels from the flux form potential vorticity equation

$$\left[ \frac{\partial(\rho h Q)}{\partial t} \right]_\sigma = \frac{\hat{\mathbf{z}} \cdot (\nabla_{\mathbf{v}} \rho \times \nabla_{\mathbf{v}} p)}{\rho^2} - \nabla_{\mathbf{v}} \cdot [\rho h \mathbf{u} Q + \hat{\mathbf{z}} \times \dot{\sigma} \partial_\sigma \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}^h], \quad (13.181)$$

which is equivalent to the absolute vorticity equation

$$\left[ \frac{\partial \tilde{\zeta}_a}{\partial t} \right]_\sigma = \frac{\hat{\mathbf{z}} \cdot (\nabla_{\mathbf{v}} \rho \times \nabla_{\mathbf{v}} p)}{\rho^2} - \nabla_{\mathbf{v}} \cdot [\mathbf{u} \tilde{\zeta}_a + \hat{\mathbf{z}} \times \dot{\sigma} \partial_\sigma \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}^h]. \quad (13.182)$$

As a check, note that setting  $\sigma = z$  so that  $h = 1$  reduces the vorticity equation (13.182) to the vertical component of the vorticity equation

$$\rho \frac{D(\omega_a/\rho)}{Dt} = (\omega_a \cdot \nabla) \mathbf{v} + \mathbf{B} + \nabla \times \mathbf{F}. \quad (13.183)$$

### 13.15.3 Potential vorticity for the Boussinesq ocean

From Chapter 6, the [Boussinesq ocean](#) vorticity budget has a vertical component to the absolute vorticity that is unaffected by baroclinicity. This property holds in the present context, as seen by returning to the vector-invariant velocity equation (13.170) and setting the factor  $1/\rho$  multiplying the pressure gradient to  $1/\rho_0$  as part of the Boussinesq ocean

$$(1/\rho) \nabla_{\mathbf{v}} p \longrightarrow (1/\rho_0) \nabla_{\mathbf{v}} p, \quad (13.184)$$

in which  $\rho_0$  is a constant. In this case the  $\nabla_{\mathbf{v}} \times$  operation annihilates pressure and we are left with no vertical component to the baroclinicity. We are thus led to define the Boussinesq potential vorticity

$$Q = \frac{\tilde{\zeta}_a}{h} \quad (13.185)$$

which satisfies the material and flux-form evolution equations

$$h \frac{DQ}{Dt} = \partial_\sigma(\dot{\sigma} h Q) + \nabla_{\mathbf{v}} \cdot [\hat{\mathbf{z}} \times \dot{\sigma} \partial_\sigma \mathbf{u} - \hat{\mathbf{z}} \times \mathbf{F}^h] \quad (13.186)$$

$$\left[ \frac{\partial(h Q)}{\partial t} \right]_\sigma = -\nabla_{\mathbf{v}} \cdot [h \mathbf{u} Q + \hat{\mathbf{z}} \times \dot{\sigma} \partial_\sigma \mathbf{u} + \hat{\mathbf{z}} \times \mathbf{F}^h]. \quad (13.187)$$

We again emphasize that  $\dot{\sigma}$  is generally non-zero, even for a perfect fluid, so that potential vorticity as defined via  $\sigma$  is not generally a material constant for a perfect fluid. It is only when  $\dot{\sigma} = 0$  for a perfect fluid (e.g.,  $\sigma$  is buoyancy or specific entropy) that we recover the desirable perfect fluid properties of potential vorticity. We develop the theory for this case in Section 14.3.

### 13.15.4 Comments

We emphasize that the potential vorticity as given by either equations (13.176) or equation (13.185) are generally not materially invariant since they have a non-zero baroclinicity that remains even for a perfect fluid. The key point is that the vertical coordinate must itself be materially invariant, as per the case of a buoyancy coordinate in a perfect fluid that renders the Ertel potential vorticity. This point is easily lost in the details of generalized vertical coordinate numerical models (e.g., [Griffies et al. \(2020\)](#)). So although the numerical model might have a variable referred to as “potential vorticity”, it refers to the Ertel potential vorticity only when the layer specific thickness measures the buoyancy stratification. The special case of  $\sigma = z$  offers a clear case in point, in which  $h = 1$  so that  $Q$  reduces to the absolute vorticity.



## 13.16 Exercises

### EXERCISE 13.1: PRACTICE WITH BOUSSINESQ TERRAIN FOLLOWING COORDINATES

In this exercise we derive some equations using the terrain following coordinate from Section 12.3.2

$$\sigma = \frac{z - \eta_s}{-\eta_b + \eta_s}, \quad (13.188)$$

with this coordinate suited to a Boussinesq ocean.

- (a) Write the expression for the vertical grid cell increment,  $dz$ , in terms of  $d\sigma$ .
- (b) Write the time tendency equation for  $dz$ .
- (c) Compute the vertical integral,  $\int_{\sigma(\eta_b)}^{\sigma(\eta_s)} d\sigma$ .

### EXERCISE 13.2: PRACTICE WITH BOUSSINESQ BOTTOM SLOPE COORDINATES

In this exercise we repeat Exercise 13.1 only here with the bottom slope coordinate from Section 12.3.4

$$\sigma = z - \mathbf{x} \cdot \nabla \eta_b. \quad (13.189)$$

These coordinates were used by [Peterson and Callies \(2022\)](#) in their study of turbulence along a sloping bottom.

- (a) Write the expression for the vertical grid cell increment,  $dz$ , in terms of  $d\sigma$ .
- (b) Write the time tendency equation for  $dz$ .
- (c) Compute the vertical integral,  $\int_{\sigma(\eta_b)}^{\sigma(\eta_s)} d\sigma$ .

### EXERCISE 13.3: PRACTICE WITH NON-BOUSSINESQ TERRAIN FOLLOWING COORDINATES

In this exercise we repeat Exercise 13.1, here using the pressure-based terrain following vertical coordinate,

$$\sigma = \frac{p - p_a}{p_b - p_a}, \quad (13.190)$$

for a fluid in approximate hydrostatic balance so that

$$\partial_z p = -\rho g. \quad (13.191)$$

In equation (13.190) we introduced the bottom pressure,  $p_b(x, y, t)$ , and the applied surface pressure,  $p_a(x, y, t)$ . For an atmosphere we might set  $p_a = 0$  for the top of the atmosphere pressure.

- (a) Write the expression for the density-weighted vertical grid cell increment,  $\rho dz$  (also the mass per horizontal area), in terms of  $d\sigma$ .
- (b) Write the time tendency equation for  $\rho dz$ .
- (c) Compute the vertical integral,  $\int_{\sigma(\eta_b)}^{\sigma(\eta_s)} d\sigma$ .
- (d) Compute the vertical integral,  $\int_{\eta_b}^{\eta_s} \rho dz$ , thus determining the total mass per unit horizontal area for the fluid column.

## EXERCISE 13.4: PRACTICE WITH NON-BOUSSINESQ RESCALED PRESSURE COORDINATES

In this exercise we repeat Exercise 13.3, here using the rescaled pressure vertical coordinate from Section 12.3.3,

$$\sigma = \frac{p_{bo}(p - p_a)}{p_b - p_a} \quad (13.192)$$

for a fluid in approximate hydrostatic balance so that

$$\partial_z p = -\rho g. \quad (13.193)$$

In equation (13.192), we introduced the bottom pressure,  $p_b(x, y, t)$ , the applied surface pressure,  $p_a(x, y, t)$ , and the bottom pressure for a resting fluid,  $p_{bo}$ .

- (a) Write the expression for the density-weighted vertical grid cell increment,  $\rho dz$  (also the mass per horizontal area), in terms of  $d\sigma$ .
- (b) Write the time tendency equation for  $\rho dz$ .
- (c) Compute the vertical integral,  $\int_{\sigma(\eta_b)}^{\sigma(\eta_s)} d\sigma$ .
- (d) Compute the vertical integral,  $\int_{\eta_b}^{\eta_s} \rho dz$ , thus determining the total mass per unit horizontal area for the fluid column.

## EXERCISE 13.5: CHECKING THE VORTICITY EQUATION

Verify that for  $2\Omega = f\hat{z}$  the choice  $\sigma = z$  reduces the vorticity equation (13.182) to the vertical component of the vorticity equation (13.183).



## Chapter 14

# CONTINUOUS AND LAYERED ISOPYCNAL MODELS

Away from turbulent boundary layers, flows in stratified fluids are oriented according to buoyancy surfaces. For this reason, buoyancy in the ocean or specific entropy in the atmosphere play a key role in theoretical and numerical models of ocean and atmosphere circulation. In this chapter we study the hydrostatic Boussinesq ocean equations using buoyancy as the vertical coordinate, deriving their momentum equations as well as their vorticity and potential vorticity equations. The resulting primitive equations form the basis for *isopycnal* models of the ocean as well as isentropic models of the atmosphere.

We pay particular attention to the needs of vertically integrating the equations over discrete layers, as required to develop numerical layered isopycnal models. Since we are considering a perfect fluid, the vertically discrete layered isopycnal equations are identical to the stacked shallow water equations. We further an understanding of the stacked shallow water model by deriving, in some detail, the [thickness weighted averaging \(TWA\)](#) shallow water equations. The [TWA](#) equations provide a versatile mathematical and numerical framework for studying the physics of eddy and mean flow interactions.

### READER'S GUIDE FOR THIS CHAPTER

Throughout this chapter we expose details for the practitioner interested in the mathematical physics of momentum, vorticity and potential vorticity as realized using buoyancy as a vertical coordinate. We assume a linear equation of state, thus allowing for buoyancy to provide a full expression of thermodynamics. The derivations require an understanding of the generalized vertical coordinate mathematics in Chapter 12 and geophysical fluid mechanics in Chapter 13. It is notable that the contravariant form of the horizontal velocity,  $\mathbf{u}$ , appears in both geopotential vertical coordinate and generalized vertical coordinate formulations, thus motivating its use throughout this chapter. In the second half of this chapter, we assume a working knowledge of the [shallow water model](#) from Chapters 1 and 2.

14.1	Loose threads . . . . .	472
14.2	Layered isopycnal primitive equations . . . . .	473
14.2.1	Montgomery potential and the pressure force . . . . .	473
14.2.2	Material time derivative . . . . .	475
14.2.3	Layer thickness and specific thickness . . . . .	475
14.2.4	Summary of the isopycnal ocean equations . . . . .	477
14.2.5	Thickness weighted velocity equation . . . . .	477
14.2.6	Vector-invariant horizontal velocity equation . . . . .	477
14.2.7	Connection to the stacked shallow water equations . . . . .	478
14.2.8	Diapycnal transfer . . . . .	478

14.2.9	Momentum transfer . . . . .	479
14.2.10	Vanished isopycnal layers . . . . .	479
<b>14.3</b>	<b>Potential vorticity using isopycnal coordinates . . . . .</b>	<b>479</b>
14.3.1	The vorticity equation . . . . .	480
14.3.2	Derivation of the potential vorticity equation . . . . .	480
14.3.3	Potential vorticity with irreversible processes . . . . .	481
14.3.4	Comments . . . . .	483
<b>14.4</b>	<b>Thickness weighted shallow water . . . . .</b>	<b>483</b>
14.4.1	Stacked shallow water versus continuous isopycnal . . . . .	483
14.4.2	Concerning the gradient operator . . . . .	484
14.4.3	The unaveraged thickness weighted equations . . . . .	484
<b>14.5</b>	<b>Thickness transport by the bolus velocity . . . . .</b>	<b>485</b>
14.5.1	Rectified effects . . . . .	485
14.5.2	An undulating shallow water fluid layer . . . . .	485
14.5.3	Stokes drift . . . . .	486
14.5.4	Linearized thickness perturbations . . . . .	487
14.5.5	Correlation between thickness and velocity . . . . .	488
14.5.6	Is the bolus velocity needed for the mean-field equations? . . . . .	488
<b>14.6</b>	<b>A summary of averaging operators . . . . .</b>	<b>488</b>
14.6.1	Reynolds average . . . . .	489
14.6.2	Ensemble average . . . . .	489
14.6.3	The algebra of thickness weighted averages . . . . .	490
14.6.4	Comments . . . . .	491
<b>14.7</b>	<b>TWA equations for thickness and tracer . . . . .</b>	<b>491</b>
14.7.1	TWA thickness equation . . . . .	491
14.7.2	Tracer equation . . . . .	492
14.7.3	Defining $w^\sharp$ via continuity with $\hat{u}$ . . . . .	492
<b>14.8</b>	<b>Horizontal momentum equation . . . . .</b>	<b>493</b>
14.8.1	Kinetic stress and Reynolds stress . . . . .	493
14.8.2	Thickness and pressure gradient correlation . . . . .	494
14.8.3	Unpacking the thickness and pressure gradient correlation . . . . .	494
14.8.4	Zonal and meridional Eliassen-Palm fluxes: version I . . . . .	498
14.8.5	Zonal and meridional Eliassen-Palm fluxes: version II . . . . .	499
14.8.6	Interfacial stresses from geostrophic eddies . . . . .	500
14.8.7	Comments . . . . .	501
<b>14.9</b>	<b>Vorticity and potential vorticity . . . . .</b>	<b>501</b>
14.9.1	Derivation . . . . .	501
14.9.2	Concerning the mean field potential vorticity . . . . .	502
14.9.3	Comments . . . . .	502
<b>14.10</b>	<b>Vorticity fluxes for non-divergent barotropic flow . . . . .</b>	<b>503</b>
<b>14.11</b>	<b>Exercises . . . . .</b>	<b>504</b>

---

## 14.1 Loose threads

- Formulate the TWA energy equations as in [Loose et al. \(2022\)](#) or [Loose et al. \(2023\)](#).
- Formulate the thickness weighted tracer variance equations as in Appendix A of [Pudig et al. \(2026\)](#).

## 14.2 Layered isopycnal primitive equations

Rather than specializing the generalized vertical coordinate equations provided in Section 13.13, we find it pedagogical to start from the equations written using the geopotential vertical coordinate (VOLUME 2)

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h \varphi + \mathbf{F}^h \quad \text{horizontal momentum} \quad (14.1a)$$

$$\frac{\partial \varphi}{\partial z} = b \quad \text{hydrostatic} \quad (14.1b)$$

$$\nabla_h \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad \text{continuity} \quad (14.1c)$$

$$\frac{Db}{Dt} = \dot{b} \quad \text{thermodynamics} \quad (14.1d)$$

$$\frac{DC}{Dt} = \dot{C}. \quad \text{tracers,} \quad (14.1e)$$

In these equations,  $\mathbf{v} = (\mathbf{u}, w)$  is the velocity field with  $\mathbf{u}$  its horizontal component,  $\varphi$  is the dynamic pressure,  $b$  is the Archimedean buoyancy,  $C$  is an arbitrary tracer concentration,  $\dot{b}$  and  $\dot{C}$  arise from processes leading to material time changes, and  $\mathbf{F}^h$  is an acceleration arising from friction and/or boundary stresses. A discrete realization of the isopycnal layer-integrated form of these equations is depicted in Figure 14.1, with the remainder of this section detailing the continuum formulation using isopycnal generalized vertical coordinates.

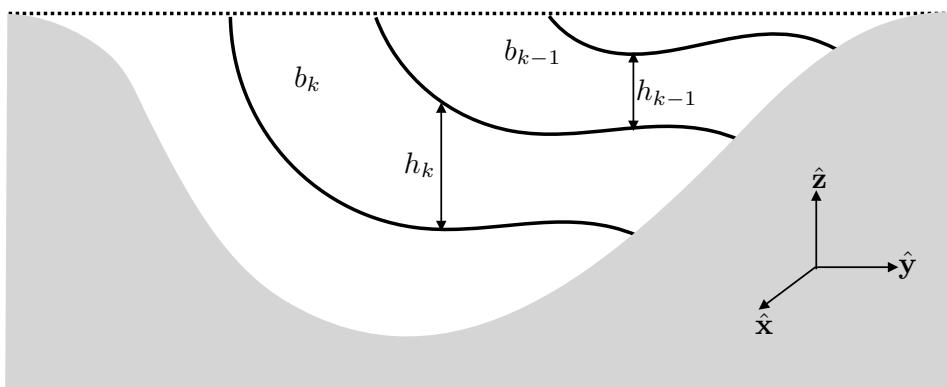


FIGURE 14.1: Schematic of an isopycnal model, formulated as stacked shallow water layers that generally allow for the transfer of matter and energy across the layer interfaces as well as across the ocean surface and ocean bottom. Discrete layer thicknesses are denoted  $h_k$  with corresponding layer buoyancy,  $b_k$ . The dark gray region is land.

### 14.2.1 Montgomery potential and the pressure force

We here consider the horizontal pressure force appearing in isopycnal models, in which we uncover the importance of the Montgomery potential.

### Horizontal pressure gradient force

Throughout this chapter we make use of the horizontal derivatives on constant buoyancy surfaces (derived in Section 12.12), written here in the form

$$\nabla_{hb} = \hat{x} \left[ \frac{\partial}{\partial x} \right]_b + \hat{y} \left[ \frac{\partial}{\partial y} \right]_b. \quad (14.2)$$

Following the discussion in Section 13.13.2, the horizontal pressure gradient transforms as

$$\nabla_h \varphi = \nabla_{hb} \varphi - \frac{\partial \varphi}{\partial z} \nabla_{hb} z = \nabla_{hb} \varphi - b \nabla_{hb} z = \nabla_{hb} (\varphi - bz) = \nabla_{hb} M, \quad (14.3)$$

where

$$M = \varphi - bz \quad (14.4)$$

defines the [Montgomery potential](#). Furthermore, as shown in a few paragraphs below, the Montgomery potential satisfies the buoyancy coordinate form of the hydrostatic balance. Evidently, the Montgomery potential plays a role for isopycnal coordinates that is directly analogous to pressure in geopotential coordinates. Correspondingly, the Montgomery potential is the geostrophic streamfunction in buoyancy coordinates (see Section 14.2.4).

### Concerning the horizontal pressure gradient force for numerical models

In deriving equation (14.3), a critical step concerns our ability to bring buoyancy inside the  $\nabla_{hb}$  gradient operator, which follows from  $\nabla_{hb} b = 0$ . As a result, we can combine two terms into the single Montgomery potential. We contrast this key step with that found in the more general case considered in Section 13.14 and highlighted in Figure 13.10. For the general case, we cannot always combine the two terms, thus leading to difficulties with accurate numerical realizations of the horizontal pressure gradient. Evidently, the ability to make use of the Montgomery potential offers a distinct advantage for isopycnal coordinates over alternative approaches to studying stratified perfect fluid flows.

The property  $\nabla_{hb} b = 0$  is available only under certain cases that utilize an idealized equation of state. In more realistic cases, the buoyancy determining the hydrostatic pressure (i.e., the *mass buoyancy*) is defined locally with the *in situ* density, whereas the generalized vertical coordinate must be defined globally. As a result, use of a realistic equation of state leads to two terms contributing to the pressure gradient in a manner similar to terrain-following models (Figure 13.10). [Sun et al. \(1999\)](#), [Hallberg \(2005\)](#), and [Adcroft et al. \(2008\)](#) discuss this issue in the context of numerical isopycnal ocean modeling. For the present chapter we ignore this detail and so assume a simplified equation of state so that  $\nabla_{hb} b = 0$ .

### Hydrostatic balance

Supporting our use of the Montgomery potential as a pressure field, the hydrostatic balance takes the form

$$\frac{\partial M}{\partial b} = \frac{\partial \varphi}{\partial b} - b \frac{\partial z}{\partial b} - z = \frac{\partial \varphi}{\partial z} \frac{\partial z}{\partial b} - b \frac{\partial z}{\partial b} - z = -z, \quad (14.5)$$

where we made use of the hydrostatic balance  $\partial \varphi / \partial z = b$  (equation (14.1b)). This result further supports considering the Montgomery potential as the buoyancy coordinate version of pressure.

### 14.2.2 Material time derivative

As seen in Section 13.5, there are two equivalent forms for the material time derivative

$$\frac{D}{Dt} = \left[ \frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_h + w \frac{\partial}{\partial z} \quad \text{geopotential form} \quad (14.6a)$$

$$= \left[ \frac{\partial}{\partial t} \right]_b + \mathbf{u} \cdot \nabla_{hb} + w^{(b)} \frac{\partial}{\partial z} \quad \text{isopycnal form,} \quad (14.6b)$$

where

$$w^{(b)} = \frac{\partial z}{\partial b} \frac{Db}{Dt} \quad (14.7)$$

is the diapycnal velocity component that measures the rate that fluid crosses buoyancy surfaces (Section 13.4). Besides differences in the spatial operators, it is important to note that the time derivative operators in equations (14.6a) and (14.6b) are computed on constant geopotential and constant buoyancy surfaces, respectively. However, the horizontal velocity component is the *same* for both forms of the material time derivative

$$\mathbf{u} = \hat{\mathbf{x}} u + \hat{\mathbf{y}} v = (D/Dt)(\hat{\mathbf{x}} x + \hat{\mathbf{y}} y). \quad (14.8)$$

### 14.2.3 Layer thickness and specific thickness

The continuity equation,  $\nabla_h \cdot \mathbf{u} + \partial_z w = 0$ , is an expression of volume conservation. We already derived the generalized vertical coordinate version of this equation in Section 13.10.3, and thus quote the isopycnal layer thickness result here

$$\left[ \frac{\partial h}{\partial t} \right]_b + \nabla_{hb} \cdot (h \mathbf{u}) + \Delta_b w^{(b)} = 0. \quad (14.9)$$

The field,  $h$ , measures the isopycnal layer thickness (with dimensions of length) and is given by the vertical integral over a layer

$$h = \int_{z(b-\delta b/2)}^{z(b+\delta b/2)} dz = \int_{b-\delta b/2}^{b+\delta b/2} \frac{\partial z}{\partial b} db = \int_{b-\delta b/2}^{b+\delta b/2} h db = \int_{b-\delta b/2}^{b+\delta b/2} N^{-2} db = \bar{h} db. \quad (14.10)$$

The specific thickness,  $h$ , is the thickness per buoyancy, and it equals to the inverse squared buoyancy frequency

$$h = \frac{\partial z}{\partial b} = N^{-2}, \quad (14.11)$$

with its layer averaged value

$$\bar{h} = h/\delta b. \quad (14.12)$$

Furthermore, the dia-surface transport velocity is given by

$$w^{(b)} = h \dot{b}. \quad (14.13)$$

Its difference across layer interfaces,

$$\Delta_b w^{(b)} = \int_{b-\delta b/2}^{b+\delta b/2} \frac{\partial w^{(b)}}{\partial b} db = w^{(b)}(b + \delta b/2) - w^{(b)}(b - \delta b/2) \quad (14.14)$$

measures the amount of fluid that diverges from the layer through cross-layer transport.

In the limit that  $\delta b \rightarrow 0$ , we find that the non-dimensional vertical difference operator can be written in one of the following equivalent manners

$$\lim_{\delta b \rightarrow 0} \Delta_b = \delta b \frac{\partial}{\partial b} = \delta b \frac{\partial z}{\partial b} \frac{\partial}{\partial z} = \delta z \frac{\partial}{\partial z} = h \frac{\partial}{\partial z}. \quad (14.15)$$

The relations are useful in moving between discrete and continuous formulations of the isopycnal equations.

### Specific thickness equation

Inserting  $h = \bar{h} \delta b$  into the thickness equation (14.9) leads to

$$\left[ \frac{\partial \bar{h}}{\partial t} \right]_b + \nabla_{hb} \cdot (\bar{h} \mathbf{u}) + \partial_b w^{(b)} = 0, \quad (14.16)$$

where we pulled the buoyancy increment,  $\delta b$ , outside of the time and horizontal derivative operators since  $\delta b$  is a fixed number for a chosen layer. We also used the identity (14.15) relating the difference operator to a differential operator

$$\delta b \partial_b = \Delta_b. \quad (14.17)$$

For a vertically continuous treatment, equation (14.16) can be written with  $h$  rather than the discrete layer averaged value

$$\left[ \frac{\partial h}{\partial t} \right]_b + \nabla_{hb} \cdot (h \mathbf{u}) + \partial_b w^{(b)} = 0. \quad (14.18)$$

It is generally more convenient to use the specific thickness when working with the vertically continuous equations, whereas the finite layer thickness,  $h$ , is more suitable for the layer integrated equations.

### Perfect fluid limit

When  $w^{(b)} \neq 0$ , the three terms in the thickness equation (14.9) are coupled. Likewise, the three terms in the specific thickness equation (14.16) are coupled. We discussed this coupling in Section 13.6 as part of a broader study of the vertical velocity and the dia-surface velocity. When considering perfect fluids, we have  $w^{(b)} = h \dot{b} = 0$  since buoyancy remains materially constant. In this case the layer thickness is altered only through horizontal rearrangements of volume within a layer according to the perfect fluid thickness equation

$$\left[ \frac{\partial h}{\partial t} \right]_b + \nabla_{hb} \cdot (h \mathbf{u}) = 0. \quad (14.19)$$

As further discussed in Section 14.2.7, the perfect fluid limit brings the discrete isopycnal model into accord with the immiscible stacked shallow water model from Chapters 1 and 2.

### 14.2.4 Summary of the isopycnal ocean equations

Bringing the pieces together leads to the isopycnal version of the hydrostatic Boussinesq equations

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_{\text{hb}}) \mathbf{u} + (w^{(b)} \partial_z) \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_{\text{hb}} M + \mathbf{F}^h \quad \text{momentum} \quad (14.20a)$$

$$\frac{\partial M}{\partial b} = -z \quad \text{hydrostatic} \quad (14.20b)$$

$$\left[ \frac{\partial h}{\partial t} \right]_b + \nabla_{\text{hb}} \cdot (h \mathbf{u}) + \Delta_b w^{(b)} = 0 \quad \text{thickness} \quad (14.20c)$$

$$\left[ \frac{\partial (h C)}{\partial t} \right]_b + \nabla_{\text{hb}} \cdot (h C \mathbf{u} + h \mathbf{J}^h) + \Delta_b (C w^{(b)} + J^{(b)}) = 0 \quad \text{tracer}, \quad (14.20d)$$

where the tracer equation includes possible subgrid scale flux contributions in addition to advective transport. Notice how the advective transport is two-dimensional in the perfect fluid limit where  $\dot{b} = 0$ , in which case layer-integrated scalar properties, such as volume and tracer content, are constant within buoyancy layers. Also note that geostrophic balance in the horizontal momentum equation (14.20a) gives

$$f \hat{\mathbf{z}} \times \mathbf{u}_g = -\nabla_{\text{hb}} M \implies f u_g = -\left[ \frac{\partial M}{\partial y} \right]_b \quad \text{and} \quad f v_g = \left[ \frac{\partial M}{\partial x} \right]_b. \quad (14.21)$$

Hence, the Montgomery potential is the streamfunction for geostrophic flow as represented using buoyancy coordinates.

### 14.2.5 Thickness weighted velocity equation

As in our discussion of the stacked shallow water model in Chapters 1 and 2, we can write the velocity equation (14.20a) in its thickness weighted form, with this form suited to studying momentum balances and pressure form stresses. The manipulations are directly analogous to the shallow water case, whereby we multiply equation (14.20a) by the thickness,  $h$ , and multiply the thickness equation (14.20c) by the horizontal velocity,  $\mathbf{u}$ , and then summing to find

$$\left[ \frac{\partial (h \mathbf{u})}{\partial t} \right]_b + \nabla_{\text{hb}} \cdot (h \mathbf{u} \otimes \mathbf{u}) + \Delta_b (w^{(b)} \mathbf{u}) + f \hat{\mathbf{z}} \times (h \mathbf{u}) = -h \nabla_{\text{hb}} M + h \mathbf{F}^h. \quad (14.22)$$

For the diapycnal transfer term, we made use of the operator identity

$$\lim_{\delta b \rightarrow 0} \Delta_b = h \partial_z \quad (14.23)$$

from equation (14.15).

### 14.2.6 Vector-invariant horizontal velocity equation

It is common for isopycnal models to make use of the vector-invariant form of the momentum equation derived in Section 13.13.4. Introducing the isopycnal version of the relative vorticity,

$$\hat{\mathbf{z}} \tilde{\zeta} \equiv \nabla_{\text{hb}} \times \mathbf{u} = \left[ \frac{\partial v}{\partial x} \right]_b - \left[ \frac{\partial u}{\partial y} \right]_b, \quad (14.24)$$

renders the vector-invariant horizontal velocity equation

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_b + w^{(b)} \partial_z \mathbf{u} + \tilde{\zeta}_a \hat{z} \times \mathbf{u} = -\nabla_{hb} \mathcal{B} + \mathbf{F}^h, \quad (14.25)$$

where

$$\mathcal{B} = M + \mathbf{u} \cdot \mathbf{u}/2 = \varphi - b z + \mathbf{u} \cdot \mathbf{u}/2 \quad (14.26)$$

is the Bernoulli potential for a hydrostatic Boussinesq fluid, and

$$\tilde{\zeta}_a = \tilde{\zeta} + f \quad (14.27)$$

is the vertical component to the absolute vorticity using isopycnal coordinates. Note that we can further introduce the isopycnal potential vorticity (Section 14.3.2)

$$h Q = \tilde{\zeta}_a \quad (14.28)$$

to bring the horizontal velocity equation to the form

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_b + w^{(b)} \partial_z \mathbf{u} + Q \hat{z} \times (h \mathbf{u}) = -\nabla_{hb} \mathcal{B} + \mathbf{F}^h. \quad (14.29)$$

This form is commonly used as the starting point for certain theoretical analyses, particularly when considering the perfect fluid limit in which  $w^{(b)} = 0$ .

### 14.2.7 Connection to the stacked shallow water equations

We can make use of the material time derivative operator (14.6b) to write the material form of the perfect fluid equations (14.20a)-(14.20c)

$$\frac{D\mathbf{u}}{Dt} + f \hat{z} \times \mathbf{u} = -\nabla_{hb} M \quad (14.30a)$$

$$\frac{\partial M}{\partial b} = -z \quad (14.30b)$$

$$\frac{Dh}{Dt} + h \nabla_{hb} \cdot \mathbf{u} = 0. \quad (14.30c)$$

These isopycnal equations are isomorphic to those for a single layer of perfect shallow water fluid (Chapters 1 and 2). This isomorphism allows us to derive the vorticity and potential vorticity equations in Section 14.3 by directly following the shallow water manipulations.

### 14.2.8 Diapycnal transfer

At ocean boundaries, the diapycnal term,  $w^{(b)}$ , accounts for the transfer of matter across the ocean boundaries via precipitation, evaporation, ice melt/form, and river runoff. Notably, this matter transfer also generally gives rise to a transfer of trace matter (tracers), heat (evaporation and precipitation carry a heat content), and momentum (precipitation generally has nonzero momentum). In the ocean interior,  $w^{(b)}$  affects the transfer of volume, tracer, and momentum between layers as induced by irreversible processes such as mixing.

### 14.2.9 Momentum transfer

Pressure **form stress** mechanically couples isopycnal layers even in the absence of diapycnal matter transfer. We discussed the physics of form stress for the shallow water system in Chapter 2, and devoted a full chapter in VOLUME 2 to this topic. Furthermore, there are a suite of unresolved processes giving rise to lateral and vertical stresses. Typical ocean model treatments incorporate a turbulent friction in the ocean interior, with lateral stresses acting within a layer and diapycnal stresses acting across isopycnal layer interfaces. A bottom drag is typically applied at the ocean bottom and a turbulent stress applied at the ocean surface. Details for the boundary stresses involve the physics of boundary layer turbulence, which is a topic outside of our scope.

### 14.2.10 Vanished isopycnal layers

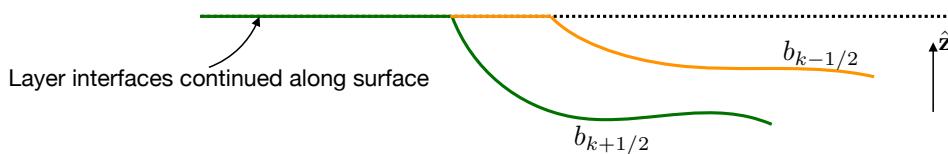


FIGURE 14.2: Schematic of isopycnal layer interfaces,  $b_{k-1/2}$  and  $b_{k+1/2}$ , that intersect the surface boundary and are analytically continued along the boundary in zero thickness regions. This construct provides a conceptual and practical method for handling layers that intersect boundaries, and thus accounting for boundary transport of matter across layer interfaces. A similar construct is used in formulating available potential energy in the presence of boundaries in VOLUME 2, and for processes leading to surface water mass transformation in Section 20.4.6.

Isopycnal layers have a transient existence at any particular horizontal position since a layer can incrop at the ocean bottom and outcrop at the ocean surface (see Figure 14.2). The seasonal cycle of warming and cooling is a canonical example of layer outcropping at the surface ocean. A formulation expedient to handle the creation and destruction of layers is to assume that all layers exist everywhere horizontally across the ocean domain, but to allow for zero layer thickness at a horizontal point where it does not exist. To admit this feature in a discrete model requires a careful realization of L'Hôpital's rule of differential calculus, thus ensuring the discrete model conserves properties in the presence of layers that can appear and disappear at any particular point in the domain.

## 14.3 Potential vorticity using isopycnal coordinates

In Chapter 6, we consider the absolute vorticity in a Boussinesq and hydrostatic ocean, with a simplified **equation of state**. This vorticity, when projected into the direction normal to constant buoyancy surface,  $\omega_a \cdot \nabla b$ , is not affected by **baroclinicity**; i.e., the projection annihilates the baroclinicity vector. From that property we conclude that  $\omega_a \cdot \nabla b$  is the potential vorticity for the Boussinesq hydrostatic fluid.

For a Boussinesq hydrostatic fluid, isopycnal coordinates build in the above property of buoyancy surfaces. Indeed, as shown in Section 13.15.3, the vertical component to baroclinicity vanishes for any generalized vertical coordinate representation of a Boussinesq fluid. Hence, buoyancy coordinates are not special from this perspective. Instead, they are special since in the case of a perfect fluid, the buoyancy based potential vorticity is materially invariant. In

Section 13.15, we derived the potential vorticity equation for a hydrostatic fluid represented with generalized vertical coordinates. We could choose to specialize that result to the current case of isopycnal coordinates. We instead choose to step through the maths via a series of exercises, thus offering further experience performing manipulations with isopycnal coordinates and vorticity.

### 14.3.1 The vorticity equation

In Exercise 14.1 we perform the few steps needed to derive the isopycnal vorticity equation

$$\left[ \frac{\partial \tilde{\zeta}_a}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_{hb}) \tilde{\zeta}_a = -\tilde{\zeta}_a \nabla_{hb} \cdot \mathbf{u} \quad (14.31)$$

where

$$\tilde{\zeta}_a = f + \hat{z} \cdot (\nabla_{hb} \times \mathbf{u}) = f + \tilde{\zeta} \quad (14.32)$$

is the absolute vorticity, written as the planetary vorticity plus the isopycnal relative vorticity. The left hand side of equation (14.31) is the material time derivative of absolute vorticity (see equation (14.6b)), so that we can write

$$\frac{D\tilde{\zeta}_a}{Dt} = -\tilde{\zeta}_a \nabla_{hb} \cdot \mathbf{u}. \quad (14.33)$$

As advertised above, there is no baroclinicity vector on the right hand side of this vorticity equation. Rather, the material time evolution of isopycnal absolute vorticity is only affected by the isopycnal convergence of the horizontal flow.

### 14.3.2 Derivation of the potential vorticity equation

In Exercise 14.2 we step through the few lines of maths to derive the potential vorticity equation

$$\frac{DQ}{Dt} = 0, \quad (14.34)$$

where the isopycnal potential vorticity is

$$Q = \frac{\tilde{\zeta}_a}{h} = \frac{f + \tilde{\zeta}}{h}. \quad (14.35)$$

Expanding the material time derivative into its components according to equation (14.6b), and making use of the perfect fluid form of the thickness equation leads to the flux-form potential vorticity equation

$$\left[ \frac{\partial (h Q)}{\partial t} \right]_b + \nabla_{hb} \cdot (h Q \mathbf{u}) = 0. \quad (14.36)$$

When formulating the vertically continuous equations rather than finite thickness layered equations, it is more convenient to make use of the specific thickness,  $h$ , rather than the layer thickness,  $h$  (see Section 14.2.3). In this case we are motivated to define the potential vorticity as

$$Q = \frac{\tilde{\zeta}_a}{h} = \frac{f + \tilde{\zeta}}{h} = (f + \tilde{\zeta}) N^2. \quad (14.37)$$

The corresponding potential vorticity equation is identical to equation (14.36), only now with  $h$  replaced by  $\mathbf{h}$ . Exercise 14.3 reveals that it is the potential vorticity (14.37) that appears when performing a coordinate transformation from the geopotential form,  $Q = (\boldsymbol{\omega}^{\text{hy}} + f \hat{z}) \cdot \nabla b$ , into its isopycnal form.

### 14.3.3 Potential vorticity with irreversible processes

We now include friction and/or boundary stresses in the momentum equation and irreversible processes in the buoyancy equation (e.g., diffusion, heating). The manipulations are straightforward yet require a bit more work than for the perfect fluid case.

As summarized in Section 14.2.4, the equations of motion with irreversible buoyancy processes as well as friction, as written using isopycnal (or buoyancy) vertical coordinates, take the form

$$\left[ \frac{\partial \mathbf{u}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_{\text{hb}}) \mathbf{u} + \dot{b} \frac{\partial \mathbf{u}}{\partial b} + \mathbf{f} \times \mathbf{u} = -\nabla_{\text{hb}} M + \mathbf{F}^{\text{h}} \quad (14.38a)$$

$$\frac{\partial M}{\partial b} = -z \quad (14.38b)$$

$$\left[ \frac{\partial \mathbf{h}}{\partial t} \right]_b + \nabla_{\text{hb}} \cdot (\mathbf{h} \mathbf{u}) = -\frac{\partial (\mathbf{h} \dot{b})}{\partial b} \quad (14.38c)$$

$$\frac{D b}{D t} = \dot{b}. \quad (14.38d)$$

Note that in this section choose to write the dia-surface transport operator in the form

$$w^{(\dot{b})} \partial_z = \dot{b} \partial_b. \quad (14.39)$$

We can make use of the material time derivative operator (14.6b) to write the material form of the equations

$$\frac{D \mathbf{u}}{D t} + \mathbf{f} \times \mathbf{u} = -\nabla_{\text{hb}} M + \mathbf{F}^{\text{h}} \quad (14.40a)$$

$$\frac{\partial M}{\partial b} = -z \quad (14.40b)$$

$$\frac{D \mathbf{h}}{D t} + \mathbf{h} \nabla_{\text{hb}} \cdot \mathbf{u} = -\mathbf{h} \frac{\partial \dot{b}}{\partial b} \quad (14.40c)$$

$$\frac{D b}{D t} = \dot{b}. \quad (14.40d)$$

### Curl of the velocity equation

Start taking the curl,  $\nabla_{\text{hb}} \times$ , of the velocity equation (14.38a), thus leading to the isopycnal vorticity equation

$$\left[ \frac{\partial \tilde{\zeta}_{\text{a}}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_{\text{hb}}) \tilde{\zeta}_{\text{a}} + \dot{b} \left[ \frac{\partial \tilde{\zeta}_{\text{a}}}{\partial b} \right] = -\tilde{\zeta}_{\text{a}} \nabla_{\text{hb}} \cdot \mathbf{u} + \hat{z} \cdot \left[ \frac{\partial \mathbf{u}}{\partial b} \times \nabla_{\text{hb}} \dot{b} + \nabla_{\text{hb}} \times \mathbf{F}^{\text{h}} \right]. \quad (14.41)$$

The left hand side of equation (14.41) is the material time derivative of absolute vorticity (see equation (14.6b)), so that

$$\frac{D\tilde{\zeta}_a}{Dt} = -\tilde{\zeta}_a \nabla_{hb} \cdot \mathbf{u} + \hat{z} \cdot \left[ \frac{\partial \mathbf{u}}{\partial b} \times \nabla_{hb} \dot{b} + \nabla_{hb} \times \mathbf{F}^h \right]. \quad (14.42)$$

Now make use of the thickness equation in the material form (14.40c) to eliminate the convergence,  $-\nabla_{hb} \cdot \mathbf{u}$ , on the right hand side, thus leading to

$$\frac{D\tilde{\zeta}_a}{Dt} - \frac{\tilde{\zeta}_a}{h} \left[ \frac{Dh}{Dt} - h \frac{\partial \dot{b}}{\partial b} \right] = \hat{z} \cdot \left[ \frac{\partial \mathbf{u}}{\partial b} \times \nabla_{hb} \dot{b} + \nabla_{hb} \times \mathbf{F}^h \right]. \quad (14.43)$$

Introducing the isopycnal potential vorticity

$$Q = \frac{\tilde{\zeta}_a}{h} = \frac{\tilde{\zeta} + f}{h} \quad (14.44)$$

leads to

$$h \frac{DQ}{Dt} = \zeta_a \frac{\partial \dot{b}}{\partial b} + \hat{z} \cdot \left[ \frac{\partial \mathbf{u}}{\partial b} \times \nabla_{hb} \dot{b} + \nabla_{hb} \times \mathbf{F}^h \right]. \quad (14.45)$$

### Massaging the irreversible terms

The terms associated with material time changes to the buoyancy can be written

$$\zeta_a \frac{\partial \dot{b}}{\partial b} + \hat{z} \cdot \left[ \frac{\partial \mathbf{u}}{\partial b} \times \nabla_{hb} \dot{b} \right] = \zeta_a \frac{\partial \dot{b}}{\partial b} + \dot{b} \frac{\partial \tilde{\zeta}}{\partial b} - \hat{z} \cdot \left[ \nabla_{hb} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \quad (14.46a)$$

$$= \zeta_a \frac{\partial \dot{b}}{\partial b} + \dot{b} \frac{\partial \tilde{\zeta}_a}{\partial b} - \hat{z} \cdot \left[ \nabla_{hb} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \quad (14.46b)$$

$$= \frac{\partial(\zeta_a \dot{b})}{\partial b} - \hat{z} \cdot \left[ \nabla_{hb} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \quad (14.46c)$$

$$= \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} + \nabla_{hb} \cdot \left[ \hat{z} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right], \quad (14.46d)$$

where the second equality follows since the Coriolis parameter is independent of the buoyancy.

### The potential vorticity equation

The potential vorticity equation takes the material form

$$h \left[ \frac{DQ}{Dt} \right] = \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} + \nabla_{hb} \cdot \left[ \hat{z} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} - \hat{z} \times \mathbf{F}^h \right]. \quad (14.47)$$

Expanding the material time derivative into its components (14.6b), and making use of the thickness equation (14.38c), leads to the flux-form equation

$$\left[ \frac{\partial(hQ)}{\partial t} \right]_b + \nabla_{hb} \cdot (hQ \mathbf{u}) + \frac{\partial(hQ \dot{b})}{\partial b} = \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} + \nabla_{hb} \cdot \left[ \hat{z} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} - \hat{z} \times \mathbf{F}^h \right]. \quad (14.48)$$

Since  $\mathbf{h} Q = \tilde{\zeta}_a$ , the  $\partial_b$  terms cancel, thus leaving the flux-form potential vorticity equation

$$\left[ \frac{\partial(\mathbf{h} Q)}{\partial t} \right]_b = -\nabla_{\mathbf{h} b} \cdot \left[ \mathbf{h} Q \mathbf{u} - \hat{\mathbf{z}} \times \dot{b} \frac{\partial \mathbf{u}}{\partial b} + \hat{\mathbf{z}} \times \mathbf{F}^h \right]. \quad (14.49)$$

#### 14.3.4 Comments

The flux-form potential vorticity equation (14.49) manifests the impermeability theorem of Chapter 8, since the right hand side is the isopycnal convergence of a flux.

## 14.4 Thickness weighted shallow water

There are a variety of mathematical formalisms used to frame the study of how linear waves, nonlinear waves, eddies, and fully developed turbulence interact with a mean flow. A distinctly geophysical element enters these studies through the primary role of vertical stratification arising from gravitation, with stratification particularly important for large scales flows where motions are approximately hydrostatic. A further specialization to the ocean context arises since there are few regions where zonal averages apply, which contrasts to the atmospheric case. The thickness weighted averaging (TWA) method has emerged as an effective formalism for stratified flows, with particular use for studies of geostrophic eddies and their parameterization. In the remainder of this chapter, we develop the TWA equations for the perfect fluid stacked shallow water model. Our focus concerns the derivation of the thickness weighted averaged equations as well as their physical interpretation.

### 14.4.1 Stacked shallow water versus continuous isopycnal

The perfect fluid stacked shallow water model exposes key facets of stratified geophysical flows without requiring the full mathematical toolkit required for continuous generalized vertical coordinates. The core simplification arises by assuming that horizontal motion has no vertical dependence within each shallow water layer, which then means that vertical motion as well as the hydrostatic pressure are linear functions of vertical position within each layer. That is, the shallow water fluid moves as extensible vertical columns (Chapter 1). It follows that horizontal pressure gradients do not need to be projected along the slope of the layer since they are vertically constant within a layer. In contrast, this projection is needed for a continuously stratified fluid described by generalized vertical coordinates, as illustrated in Figure 12.5. Hence, the shallow water equations for momentum, thickness, and tracers retain their use of Cartesian coordinates even though the layer interfaces undulate and are thus not generally horizontal. This mathematical feature of shallow water fluids aids in our pedagogical development of the TWA method.

[Young \(2012\)](#) offers an elegant application of thickness weighted averaging to the continuously stratified Boussinesq hydrostatic fluid, with his paper the culmination of many years of prior work. A natural step for our development is to work through [Young \(2012\)](#) since we have the full tensor toolkit needed for that purpose. However, we choose instead to specialize [Young \(2012\)](#) to the case of vertically discrete shallow water equations. In this manner, we minimize the required mathematical apparatus while exposing the key physical concepts. We also lay the mathematical physics foundation for the study of flow in stacked shallow water models (e.g., [Marques et al. \(2022\)](#), [Loose et al. \(2022\)](#), [Loose et al. \(2023\)](#), [Jansen et al. \(2024\)](#)). Digesting this presentation, and then coupling to the generalized vertical coordinate tensor

analysis developed in this part of the book, prepares one for [Young \(2012\)](#) as well as the slightly more mathematical treatment of [Maddison and Marshall \(2013\)](#).

#### 14.4.2 Concering the gradient operator

When studying the shallow water equations, most fields are a continuous function of  $(x, y, t)$  and labeled with a discrete vertical index,  $k$ . As such, the gradient operator acting on one of these fields only has a horizontal component, so that we can equally write  $\nabla$  or  $\nabla_h$ . However, the layer pressure,  $p_k$ , is a linear function of depth within the layer, in which case we must distinguish  $\nabla_h p_k$  from  $\nabla p_k$ . To help reduce confusion, particularly when in the midst of a detailed derivation, we write  $\nabla_h$  with its extra subscript to emphasize that the gradient operator is horizontal. This extra notation is not needed in many cases, but it proves quite useful in those few places where it is needed.

#### 14.4.3 The unaveraged thickness weighted equations

The thickness weighted averaging formalism starts from flux-form evolution equations rather than advective form equations. We thus focus on the thickness equation (14.50a), the thickness weighted tracer equation (14.50b), and the thickness weighted velocity equation (14.50c) (also called the momentum equation)

$$\frac{\partial h_k}{\partial t} + \nabla_h \cdot (h_k \mathbf{u}_k) = 0 \quad (14.50a)$$

$$\frac{\partial(h_k C)}{\partial t} + \nabla_h \cdot (h_k \mathbf{u}_k C) = 0 \quad (14.50b)$$

$$\frac{\partial(h_k \mathbf{u}_k)}{\partial t} + \nabla_h \cdot [h_k \mathbf{u}_k \otimes \mathbf{u}_k] + f \hat{z} \times (h_k \mathbf{u}_k) = -(h_k / \rho_{\text{ref}}) \nabla_h p_k. \quad (14.50c)$$

The density,  $\rho_{\text{ref}}$ , appearing in the momentum equation (14.50c) is the Boussinesq reference density, often chosen as the density in the uppermost layer,

$$\rho_{\text{ref}} = \rho_1. \quad (14.51)$$

For the analysis of thickness weighted averaging, it proves useful to move seamlessly between the thickness weighted pressure gradient body force and its equivalent [contact force](#) version studied in [VOLUME 2](#). The contact force version of the momentum equation reveals the pressure form stresses acting on the upper and lower interfaces of a shallow water layer. It also brings stresses (kinetic stresses and pressure stresses) together into the divergence of a momentum flux. As such, this formulation follows that of Cauchy as discussed in [VOLUME 2](#). The eddy correlation portion of the momentum flux is known as the [Eliassen-Palm](#) flux.<sup>1</sup>

When the dust settles, the [TWA](#) equations are isomorphic to the unaveraged thickness weighted equations (14.50a)-(14.50c), yet with the addition of momentum flux convergences to the right hand sides that arise from subgrid momentum eddy correlations. These eddy fluxes are connected to the potential vorticity fluxes, with the connection known as the [Taylor-Bretherton identity](#). The isomorphism between unaveraged thickness and averaged thickness provides some motivation to favor the [TWA](#) approach. Namely, the properties of the unaveraged equations are directly reflected in the [TWA](#) equations. It also provides a suitable framework for

---

<sup>1</sup>See [Bühler \(2014b\)](#) for a historical perspective on the Eliassen-Palm flux, which was introduced by [Eliassen and Palm \(1960\)](#) in their study of stationary mountain waves.

parameterizing the subgrid correlations within the context of flux-form conservation laws. Even so, any formalism for an eddy and mean decomposition is subjective since the mean flow and eddying fluctuations are defined by the analyst rather than prescribed by the physics. Hence, arguments concerning what is a preferable framework are subject to the needs of the analyst and yet they have no physically objective foundation.

## 14.5 Thickness transport by the bolus velocity

Prior to diving into the formalism of thickness weighted averaging, we study the eddy-induced volume transport (more precisely, thickness transport) realized by linear waves within a layer of shallow water fluid. This discussion provides a specific example of the thickness transport by the *bolus velocity*, with further discussion offered in Sections 14.7.1 and 17.5.9. Much of our intuition for bolus transport is based on the following relatively simple example of Stokes drift.<sup>2</sup>

Part of the motivation for TWA is that we do not need to compute the bolus velocity. Even so, understanding the basic physics of the bolus velocity renders useful insights into how eddies, even eddies as simple as linear waves, can provide a rectified transport of material and thermal properties.

### 14.5.1 Rectified effects

Rectification is the conversion of a fluctuating motion into motion in a particular direction. For example, the transformation of an alternating electrical current into a direct electrical current occurs through a rectifier. More generally, rectification arises from the breaking of a symmetry typically through a nonlinear mechanism. The primary example in fluid mechanics is Stokes drift. Stokes drift arises when linear waves have an amplitude that is a function of space, with this spatial dependence giving rise to net particle transport (the Stokes drift) in a preferred direction. Another example concerns the turbulent Stokes drift arising from nonlinear geostrophic waves and eddies in the ocean and atmosphere that lead to a net transport of buoyancy. The meridional transport of buoyancy by eddies in a channel provides the canonical geophysical example of eddy induced transport.

### 14.5.2 An undulating shallow water fluid layer

Figure 14.3 shows a layer of constant density shallow water fluid within a perfect fluid stacked shallow water model in a non-rotating reference frame. Since the layers are immiscible, the total volume of fluid within this layer remains constant. In its unperturbed state with flat layer interfaces, the meridional velocity in the fluid layer is zero and the thickness is a constant,  $h_0$ . When perturbed, the thickness is written

$$h(y, t) = h_0 + h'(y, t), \quad (14.52)$$

where we assume the perturbation only depends on  $(y, t)$  for simplicity. The layer thickness changes in time according to the convergence of the advective transport of thickness as found by the thickness equation (14.50a)

$$\frac{\partial h}{\partial t} = -\nabla_h \cdot (h \mathbf{u}), \quad (14.53)$$

---

<sup>2</sup>This example is based Section 2 of [Lee et al. \(1997\)](#).

where the convergence is computed within the layer and we drop the  $k$  layer index for brevity. As seen by Figure 14.3, undulations of the layer thickness at a point arise from the convergence of thickness advected to that point. Further assuming that there is no zonal dependence ( $\partial_x = 0$ ) leads to the one-dimensional thickness equation

$$\frac{\partial h}{\partial t} = -\frac{\partial (h v)}{\partial y}. \quad (14.54)$$

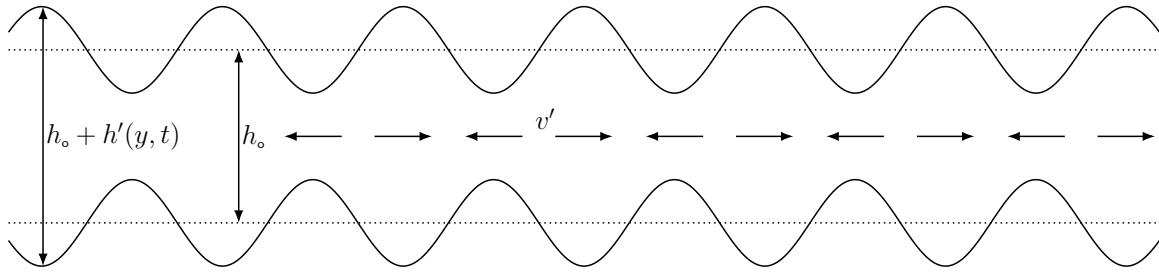


FIGURE 14.3: A single layer of constant density shallow water fluid, with resting thickness  $h = h_o$  and instantaneous thickness  $h = h_o + h'(y, t)$ . Associated with the sinusoidal undulations in thickness are fluctuations in the meridional velocity,  $v' = v_o \sin(k y - \omega t)$ , depicted here by the alternating horizontal vectors within the layer. Since the velocity and thickness are in phase as per equation (14.65), the velocity fluctuates to the right ( $v' > 0$ ) under a positive thickness anomaly ( $h' > 0$ ) and to the left ( $v' < 0$ ) with a negative anomaly ( $h' < 0$ ). The nonzero correlation between  $h'$  and  $v'$  leads to a Stokes drift, with  $\overline{v' h' / h}$  referred to as the bolus velocity. The name *bolus velocity* (originating from Rhines (1982)) is motivated by the image of a bolus of matter that moves through a tube via periodic undulations, akin to the peristaltic contractions and expansions that move food through the digestive system.

### 14.5.3 Stokes drift

Consider a monochromatic wave perturbation in the meridional velocity that propagates in the meridional direction

$$v'(y, t) = v_o \sin(\kappa y - \omega t), \quad (14.55)$$

where  $\kappa$  is a constant wave number,  $\omega > 0$  is a constant angular frequency, and  $v_o$  is the amplitude of the fluid particle velocity. This longitudinal wave is depicted in Figure 14.3. We now follow the general formalism developed in VOLUME 1 (see also Section 17.3.4) to derive an expression for the Stokes drift associated with this wave.

We are only concerned with the meridional component of the velocity, so the fluid particle trajectory equation is given by

$$\frac{dY}{dt} = v_o \sin(\kappa Y - \omega t), \quad (14.56)$$

where  $Y = Y(Y_o, t)$  is the meridional trajectory with initial position,  $Y_o$ . We can write 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)

$$\frac{dY}{dt} - v(y, t) = v_o^2 \kappa \cos(\kappa y - \omega t) \int_0^t \sin(\kappa y - \omega t') dt' \quad (14.57a)$$

$$= \frac{v_o^2 \kappa}{\omega} [\cos^2(\kappa y - \omega t) - \cos(\kappa y - \omega t) \cos(\kappa y)]. \quad (14.57b)$$

Time averaging over a single wave period,

$$T = 2\pi/\omega, \quad (14.58)$$

leads to the Stokes drift as per the general expression in VOLUME 2

$$V_{\text{stokes}} = \frac{v_o^2 \kappa}{2\omega}. \quad (14.59)$$

Introducing the phase speed for the monochromatic wave,

$$c = \omega/\kappa, \quad (14.60)$$

allows us to write the Stokes drift as

$$V_{\text{stokes}} = \frac{v_o^2}{2c}. \quad (14.61)$$

Notice how the Stokes drift becomes small when the phase speed is large. We expect this dependence since Stokes drift arises from fluid particles feeling the waves, so that for relatively fast waves the particles have little time for sampling the wave thus leading to a decrease in the Stokes drift. Correspondingly, there is only a small difference between the Eulerian and Lagrangian velocities with fast phase speeds, whereas the converse holds for slow phase speeds where Eulerian and Lagrangian velocities have a relatively large difference.<sup>3</sup>

#### 14.5.4 Linearized thickness perturbations

The velocity and thickness are written in terms of their rest state plus a perturbation due to the wave

$$h = h_o + h' \quad \text{and} \quad v = v', \quad (14.62)$$

where the velocity vanishes when the wave is absent. The thickness equation (14.54) thus takes the form

$$\frac{\partial h'}{\partial t} + h_o \frac{\partial v'}{\partial y} + v' \frac{\partial h'}{\partial y} = 0. \quad (14.63)$$

Linearizing this equation, and using the wave perturbation (14.55), leads to

$$\frac{\partial h'}{\partial t} + h_o v_o \kappa \cos(\kappa y - \omega t) = 0. \quad (14.64)$$

Time integrating this equation, and making use of the velocity perturbation in the form of equation (14.55), renders the thickness perturbation

$$h' = h_o v'/c. \quad (14.65)$$

Hence, to leading order, the thickness perturbation is directly proportional to and in phase with the velocity perturbation. We depict this in-phase relation in Figure 14.3, whereby  $v' > 0$  where  $h' > 0$  (where the layer bulges) and  $v' < 0$  where  $h' < 0$  (where the layer thins).

---

<sup>3</sup>In the case of relatively slow phase speeds, we really should keep more terms in an asymptotic expansion than those carried here.

### 14.5.5 Correlation between thickness and velocity

Over a single wave period,  $T = 2\pi/\omega$ , the temporal correlation between the linear thickness perturbation and velocity perturbation is given by

$$\overline{h' v'} = \frac{1}{T} \int_0^T h' v' dt \quad (14.66a)$$

$$= \frac{h_o}{cT} \int_0^T v' v' dt \quad (14.66b)$$

$$= \frac{v_o^2 h_o}{cT} \int_0^T \sin^2(\kappa y - \omega t) dt \quad (14.66c)$$

$$= \frac{v_o^2 h_o}{2c} \quad (14.66d)$$

$$= h_o V_{\text{stokes}}, \quad (14.66e)$$

where we introduced the Stokes drift (14.61) to reach the final equality. A nonzero correlation,  $\overline{h' v'} \neq 0$ , means that the thickness has a nonzero tendency when averaged over a wave period.

The nonzero correlation in equation (14.66e) induces a thickness transport from the one-dimensional linear longitudinal waves. This transport arises from the Stokes drift induced by the waves; without Stokes drift there is no eddy thickness transport. This behavior exemplifies the case for more general waves and nonlinear eddies moving through fluid layers. For the general case, a nonzero bolus velocity (Section 14.7.1), as determined by velocity-thickness correlations, induces an eddy thickness transport. We see that for the one-dimensional linear longitudinal wave example, the bolus velocity is the Stokes velocity.

### 14.5.6 Is the bolus velocity needed for the mean-field equations?

The bolus transport is of fundamental importance for how we think about eddy induced Stokes transport from shallow water waves. More general fluctuations, such as those from turbulent geostrophic eddies, require a parameterization to determine the thickness transport. We consider such in VOLUME 2 when studying geostrophic eddies in a zonally reentrant channel. As we see in the remainder of the current chapter, the allure of the thickness weighted averaging formalism is that it dispenses with the need to parameterize the bolus velocity. Instead, the TWA equations absorb the bolus transport into the residual mean advection operator. Operationally, the TWA exposes the eddy correlation terms only in the momentum equation, leaving the TWA thickness and TWA tracer equations in a form directly parallel to the unaveraged equations.

## 14.6 A summary of averaging operators

There are many averaging operators used in fluid mechanics, such as the wave phase average from Section 14.5, which is useful when the flow is dominantly linear waves; a long time average (formally an infinitely long time average), which is commonly used for climate studies; a space average, which is appropriate when the spatial sampling is coarse; general space filters or kernels, which are commonly used in large eddy simulations; and ensemble averages, which are generally assumed in traditional studies of turbulence. In the following, we denote the averaging or mean operator by an overbar

$$\text{average}(\Phi) = \overline{\Phi}, \quad (14.67)$$

where  $\Phi$  is any field such as velocity, thickness, or tracer concentration. Deviations (also called fluctuations) from the mean are denoted by a prime so that the full field is decomposed into a mean and eddy term according to

$$\Phi = \bar{\Phi} + \Phi'. \quad (14.68)$$

A fluctuation might be a linear wave feature moving relative to a static background flow, such as considered in Section 14.5. For nonlinear flows, a fluctuation can arise from coherent eddying flow features, such as vortices commonly generated from meandering jet-like flows (e.g., Gulf Stream rings or mid-latitude atmospheric storms spawned from the jet-stream). Furthermore, a fluctuation might represent the chaotic soup of whirls and eddies that comprise the background flow surrounding coherent vortices. In either of these cases, the nonlinear nature of the fluid mechanical equations means that any decomposition into eddy and mean leads to nonlinear eddy correlations, with such correlations the source of both the richness and complexity of geophysical fluid flows.

### 14.6.1 Reynolds average

A *Reynolds average* is an operator that annihilates its corresponding fluctuating quantity, which then means that the average of an average is the identity operator

$$\bar{\Phi}' = 0 \iff \bar{\bar{\Phi}} = \bar{\Phi}, \quad (14.69)$$

which in turn means that

$$\overline{\Phi \Psi} = \overline{(\bar{\Phi} + \Phi')} \bar{\Psi} = \bar{\Phi} \bar{\Psi}. \quad (14.70)$$

Reynolds averages are commonly used when deriving mean field equations. Even so, the assumptions of Reynolds averaging operators are not satisfied by many operators in practice. Extra technical issues arise when averaging operators do not satisfy the properties of a Reynolds average, with these issues beyond our aims in the present chapter. Hence, in this chapter we retain the Reynolds average assumption (14.69) for the averaging operator.

### 14.6.2 Ensemble average

A further assumption we make is that the average operator commutes with space and time derivatives as well as integrals. This assumption does not strictly hold if the operator is a space and/or time average operator, at least not without a bit of work. However, this assumption holds for ensemble averages. An *ensemble mean* is computed over an infinite number of realizations of the fluid flow, with approximations to this average afforded by finite sized ensembles.

Ensemble averages are typically assumed in traditional fluid turbulence studies. However, they are not always very practical nor are they the obvious choice when targeting a framework for parameterization. Even so, we prefer ensemble averages for this chapter in order to dispense with concerns about commutation of the averaging operator with derivative and integral operators. We also make use of ensemble averaging for our discussion of tracer kinematics in Chapter 17.

### 14.6.3 The algebra of thickness weighted averages

The thickness weighted average of a field is defined as the ensemble average of the thickness weighted field, and then divided by the averaged thickness

$$\widehat{\Phi} \equiv \frac{\overline{h\Phi}}{\overline{h}} \iff \overline{h}\widehat{\Phi} = \overline{h\Phi}, \quad (14.71)$$

with widehats adorning a thickness weighted average. Deviations from the thickness weighted average are denoted with two primes so that the unaveraged field is decomposed into its average plus fluctuation

$$\Phi = \widehat{\Phi} + \Phi''. \quad (14.72)$$

Since the overline average from Section 14.6.1 satisfies the Reynolds averaging assumption, so too does the thickness weighted average

$$\Phi = \widehat{\Phi} + \Phi'' \implies \widehat{\Phi''} = \frac{\overline{h\Phi''}}{\overline{h}} = 0 \implies \widehat{\Phi} = \widehat{\Phi}. \quad (14.73)$$

We are thus able to derive the following related identities

$$\Phi\Psi = (\widehat{\Phi} + \Phi'')(\widehat{\Phi} + \Phi'') \implies \widehat{\Phi\Psi} = \widehat{\Phi}\widehat{\Psi} + \widehat{\Phi''}\widehat{\Psi''} \implies \overline{h\Phi\Psi} = \overline{h}\widehat{\Phi\Psi}. \quad (14.74)$$

We sometimes need to consider mixed averages and primes, such as for

$$\overline{h}\widehat{\Phi} = \overline{h\Phi} = \overline{h}\widehat{\Phi}, \quad (14.75)$$

in which case<sup>4</sup>

$$\overline{h\Phi''} = \overline{h}(\overline{\Phi} - \widehat{\Phi}) = \overline{h\Phi} - \overline{h'}\overline{\Phi'} \neq 0. \quad (14.76)$$

Hence, the ensemble average of a fluctuation,  $\Phi''$  (which is computed relative to the thickness weighted mean), is generally nonzero. Furthermore, we sometimes find it useful to write the ensemble mean correlation between thickness and a field according to

$$\overline{h}\widehat{\Phi'} = \overline{h\Phi'} = \overline{h'}\overline{\Phi'}, \quad (14.77)$$

with the second equality following since

$$\overline{h\Phi'} = \overline{h\overline{\Phi'}} = 0. \quad (14.78)$$

The identity (14.77) allows us to write equation (14.76) as

$$\overline{h\Phi''} = \overline{h}(\overline{\Phi} - \widehat{\Phi'}). \quad (14.79)$$

A similar identity holds according to the following manipulations

$$\overline{h'\Phi'} = \overline{h\Phi'} = \overline{h(\Phi - \overline{\Phi})} = \overline{h\Phi} - \overline{h}\overline{\Phi} = \overline{h}(\widehat{\Phi} - \overline{\Phi}), \quad (14.80)$$

so that

$$\widehat{\Phi'} = \widehat{\Phi} - \overline{\Phi} = \frac{\overline{h'\Phi'}}{\overline{h}}. \quad (14.81)$$

---

<sup>4</sup>Footnote #4 in [Young \(2012\)](#) is missing the  $\overline{h\Phi}$  term appearing in equation (14.76).

Derivative operators *do not* commute with the thickness weighted average, so that, for example,

$$\partial_x \hat{u} \neq \widehat{\partial_x u}. \quad (14.82)$$

Hence, when deriving differential equations for thickness weighted fields, we first derive equations for the unaveraged thickness weighted quantities, and only thereafter do we apply the ensemble mean operator.

#### 14.6.4 Comments

For the most part, we follow the notation of [Young \(2012\)](#). Nonetheless, we caution that notational clutter and distinct conventions can present a nontrivial barrier to reading the **TWA** literature. Indeed, for our purposes with the stacked shallow water model, there is one additional piece of notation concerning the discrete layer indices. Fortunately, much of the discrete layer notation can be streamlined by exposing just the half-integer indices for fields situated at layer interfaces, along with the layer density.

### 14.7 TWA equations for thickness and tracer

In this section we derive the **TWA** versions of the thickness equation (14.50a) and the tracer equation (14.50b). The derivations involve straightforward applications of the **TWA** averaging properties (14.73) and (14.74).

#### 14.7.1 TWA thickness equation

Taking the ensemble average of the thickness equation (14.50a) renders

$$\partial_t \bar{h} + \nabla_h \cdot \bar{h} \bar{u} = 0, \quad (14.83)$$

where we dropped the layer index,  $k$ , to reduce notational clutter.<sup>5</sup> Introducing the thickness weighted average according to equation (14.71) brings the thickness equation to the form

$$\partial_t \bar{h} + \nabla_h \cdot (\bar{h} \hat{u}) = 0. \quad (14.84)$$

Consequently, the mean layer thickness,  $\bar{h}$ , evolves at a point in space according to the convergence of the thickness flux,  $-\nabla_h \cdot (\bar{h} \hat{u})$ , with the flux determined by the thickness weighted velocity,  $\hat{u}$ .

We find it useful to introduce the material time derivative operator defined with the thickness weighted velocity

$$\frac{D^\#}{Dt} = \frac{\partial}{\partial t} + \hat{u} \cdot \nabla_h = \frac{\partial}{\partial t} + \hat{u} \partial_x + \hat{v} \partial_y, \quad (14.85)$$

so that the flux-form thickness equation (14.84) can be written in the material time derivative or advective form

$$\frac{D^\# \bar{h}}{Dt} = -\bar{h} \nabla_h \cdot \hat{u}. \quad (14.86)$$

<sup>5</sup>The thickness and tracer equations do not couple to other layers, and as such we can drop the layer index,  $k = 1, N$ , when analyzing these equations. The momentum equation, in contrast, is coupled through pressure form stresses acting at the layer interfaces.

The  $D^\sharp/Dt$  notation is based on that used by [Young \(2012\)](#). The alternative,  $\widehat{D}/Dt$ , is less suitable since  $\overline{h(D/Dt)} \neq \overline{(D/Dt)h}$ . In brief, an object adorned with a sharp symbol is consistent with thickness weighted averaging but is itself not the direct result of a thickness weighted average. In the following, we find it useful to also introduce the vertical velocity,  $w^\sharp$ , in equation (14.92), and the potential vorticity,  $\Pi^\sharp$ , in equation (14.143).

The isomorphism between the **TWA** thickness equation (14.84) with the unaveraged thickness equation (14.50a) illustrates a distinct advantage of using the thickness weighted velocity,  $\widehat{\mathbf{u}}$ . Even so, for some purposes it is useful to unpack the thickness weighted velocity into its two components

$$\widehat{\mathbf{u}} = \overline{\mathbf{u}} + \frac{\overline{h' \mathbf{u}'}}{\overline{h}} \equiv \overline{\mathbf{u}} + \mathbf{u}^{\text{bolus}}, \quad (14.87)$$

with the bolus velocity defined by

$$\mathbf{u}^{\text{bolus}} = \widehat{\mathbf{u}'} = \frac{\overline{h' \mathbf{u}'}}{\overline{h}}, \quad (14.88)$$

where we made use of the identity (14.77).

We discussed the bolus velocity in Section 14.5 and see it again in Section 17.5.9 when developing the ensemble mean tracer equation in isopycnal coordinates. However, as per our discussion in Section 14.5.6, we do not need to know the bolus velocity if we write the averaged tracer and momentum equations in terms of the thickness weighted velocity,  $\widehat{\mathbf{u}}$ .

### 14.7.2 Tracer equation

Taking the ensemble average of the tracer concentration equation (14.50b) for a shallow water fluid layer renders

$$\partial_t(\overline{hC}) + \nabla_h \cdot \overline{hC \mathbf{u}} = 0. \quad (14.89)$$

Making use of the thickness weighted averages from Section 14.6.3 allows us to write

$$\overline{hC} = \overline{h}\widehat{C} \quad \text{and} \quad \overline{hC \mathbf{u}} = \overline{h}(\widehat{C}\widehat{\mathbf{u}} + \widehat{C''\mathbf{u}''}), \quad (14.90)$$

thus yielding the **TWA** tracer equation

$$\partial_t(\overline{h}\widehat{C}) + \nabla_h \cdot (\overline{h}\widehat{C}\widehat{\mathbf{u}}) = -\nabla_h \cdot (\overline{h}\widehat{C''\mathbf{u}''}). \quad (14.91)$$

The right hand side is the convergence of the thickness weighted eddy tracer flux. As seen in Section 17.6, the isopycnal form of the tracer equation is identical to that given here for a shallow water layer. In that discussion we present methods commonly used to parameterize the eddy flux convergence.

### 14.7.3 Defining $w^\sharp$ via continuity with $\widehat{\mathbf{u}}$

We generally have no need for the vertical velocity when working with the perfect fluid stacked shallow water model. Nonetheless, it is interesting to define a vertical velocity component,  $w^\sharp$ , satisfying the continuity equation

$$\nabla_h \cdot \widehat{\mathbf{u}} + \partial_z w^\sharp = 0. \quad (14.92)$$

As for the unaveraged vertical velocity component discussed in Chapter 1,  $w^\sharp$  is a linear function of  $z$  within the ensemble mean shallow water layers. Note that  $w^\sharp$  is not a thickness weighted velocity. Rather, it is the vertical velocity that is compatible, through the continuity equation, with the thickness weighted horizontal velocity. A vertical velocity is needed for the continuously stratified Boussinesq fluid, and it is defined as done here for the shallow water.<sup>6</sup>

## 14.8 Horizontal momentum equation

Taking the ensemble mean of the horizontal momentum equation (14.50c) renders

$$\partial_t(\bar{h}\bar{\mathbf{u}}) + \nabla_h \cdot [\bar{h}\bar{\mathbf{u}} \otimes \bar{\mathbf{u}}] + f\hat{\mathbf{z}} \times (\bar{h}\bar{\mathbf{u}}) = -\bar{h}\nabla_h p/\rho_{\text{ref}}, \quad (14.93)$$

where we dropped the layer interface lable,  $k$ , for brevity. Again, we make use of the thickness weighted averages from Section 14.6.3 to write

$$\bar{h}\bar{\mathbf{u}} = \bar{h}\hat{\mathbf{u}} \quad (14.94a)$$

$$\bar{h}\bar{\mathbf{u}} \otimes \bar{\mathbf{u}} = \bar{h}(\hat{\mathbf{u}} \otimes \hat{\mathbf{u}} + \widehat{\mathbf{u}'' \otimes \mathbf{u}''}), \quad (14.94b)$$

so that equation (14.93) becomes

$$\partial_t(\bar{h}\hat{\mathbf{u}}) + \nabla_h \cdot [\bar{h}\hat{\mathbf{u}} \otimes \hat{\mathbf{u}}] + f\hat{\mathbf{z}} \times (\bar{h}\hat{\mathbf{u}}) = -\nabla_h \cdot [\bar{h}\widehat{\mathbf{u}'' \otimes \mathbf{u}''}] - \bar{h}\nabla_h p/\rho_{\text{ref}}. \quad (14.95)$$

The first term on the right hand side is similar to the eddy tracer flux convergence appearing in the TWA tracer equation (14.91). In contrast, the thickness weighted pressure gradient is fundamentally distinct from anything appearing in the tracer equation. Much in the remainder of this section is devoted to developing a physical and mathematical understanding of  $\bar{h}\nabla_h p$ .

### 14.8.1 Kinetic stress and Reynolds stress

We now introduce the shallow water kinetic stress tensor

$$\mathbb{T}^{\text{sw kinetic}} = -\rho_{\text{ref}}\mathbf{u} \otimes \mathbf{u}. \quad (14.96)$$

The kinetic stress arises from motion of the fluid, with the divergence,

$$\nabla_h \cdot (h\mathbb{T}^{\text{sw kinetic}}) = -\rho_{\text{ref}}\nabla_h \cdot (\mathbf{u} \otimes \mathbf{u}), \quad (14.97)$$

contributing to changes in the momentum of a shallow water fluid column. Decomposing the velocity into the TWA velocity and fluctuation leads to the ensemble mean of the thickness weighted kinetic stress

$$\bar{h}\bar{\mathbb{T}}^{\text{sw kinetic}}(\mathbf{u}) = -\rho_{\text{ref}}\bar{h}\bar{\mathbf{u}} \otimes \bar{\mathbf{u}} \quad (14.98a)$$

$$= -\rho_{\text{ref}}\bar{h}[\hat{\mathbf{u}} \otimes \hat{\mathbf{u}} + \widehat{\mathbf{u}'' \otimes \mathbf{u}''}] \quad (14.98b)$$

$$= \bar{h}\bar{\mathbb{T}}^{\text{sw kinetic}}(\hat{\mathbf{u}}) + \bar{h}\bar{\mathbb{T}}^{\text{sw Reynolds}}, \quad (14.98c)$$

where the eddy correlation is known as the Reynolds stress tensor. The divergence of the thickness weighted Reynolds stress provides a rectified effect onto the mean flow.s

<sup>6</sup>See equation (73) in [Young \(2012\)](#).

### 14.8.2 Thickness and pressure gradient correlation

To expose the physics of the ensemble mean of the thickness weighted pressure gradient,  $\bar{h} \widehat{\nabla_h p}$ , we make use of the variety of identities that hold between layer and interface quantities, as well as between ensemble means and thickness weighted means. As a start, we write

$$\bar{h} \widehat{\nabla_h p} = \bar{h} \widehat{\nabla_h p} \quad \text{equation (14.71) defining the TWA} \quad (14.99a)$$

$$= \bar{h} \nabla_h \bar{p} + \bar{h}' \widehat{\nabla_h p'} \quad \text{expanding the ensemble mean} \quad (14.99b)$$

$$= \bar{h} (\nabla_h \bar{p} + \widehat{\nabla_h p'}) \quad \text{equation (14.77).} \quad (14.99c)$$

The eddy term is the correlation between layer thickness fluctuations and horizontal pressure gradient fluctuations

$$\bar{h} \widehat{\nabla_h p'} = \bar{h}' \widehat{\nabla_h p'}, \quad (14.100)$$

which can be written in terms of the eddy geostrophic velocity

$$\bar{h} \widehat{\nabla_h p'} = -\rho_{\text{ref}} f \hat{z} \times \bar{h}' \mathbf{u}'_g = -\rho_{\text{ref}} f \hat{z} \times \bar{h} \widehat{\mathbf{u}'_g}. \quad (14.101)$$

For the special case of geostrophic flows, the bolus velocity (14.88) equals to  $\widehat{\mathbf{u}'_g}$  so that we write

$$\bar{h} \widehat{\nabla_h p'} = -\rho_{\text{ref}} f \hat{z} \times \bar{h} \mathbf{u}^{\text{bolus}}. \quad (14.102)$$

### 14.8.3 Unpacking the thickness and pressure gradient correlation

In this subsection we unpack the correlation between eddy thickness and eddy pressure gradient as given by

$$\bar{h}_k \widehat{\nabla_h p'_k} = \bar{h}'_k \widehat{\nabla_h p_k} \quad (14.103)$$

that appears in equation (14.100). Here we expose the layer index,  $k$ , since the thickness and pressure gradient correlation will be related to interface quantities (with indices  $k \pm 1/2$ ) in the following. In words, the following manipulations proceed by writing the pressure force as a contact force rather than a body force, which exposes the eddy interfacial form stress acting at the upper and lower boundary of the layer, with the vertical divergence of this horizontal stress providing a vertical transfer of horizontal momentum between adjacent layers. In addition, we find a term arising from the horizontal gradient in the layer depth integrated pressure or, alternatively, the layer potential energy. Filling in the mathematical details requires the development given in VOLUME 2, in which we expose two equivalent expressions for the contact pressure force.

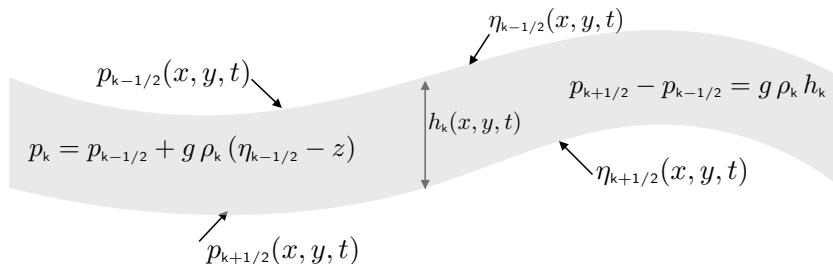


FIGURE 14.4: Schematic of the interface vertical positions,  $\eta_{k+1/2}(x, y, t)$  and  $\eta_{k-1/2}(x, y, t)$ , the interface pressures,  $p_{k+1/2}(x, y, t)$  and  $p_{k-1/2}(x, y, t)$ , and the layer pressure,  $p_k(x, y, z, t)$ , for a shallow water layer. Since  $p_k(x, y, z, t)$  is a linear function of  $z$  within the layer, its horizontal gradient is independent of  $z$  within the layer.

As part of the following derivation we make use of relations for pressure within a layer and at an interface

$$p_k = p_{k-1/2} + g \rho_k (\eta_{k-1/2} - z) \quad (14.104a)$$

$$p_{k+1/2} - p_{k-1/2} = g \rho_k h_k = -g \rho_k (\eta_{k+1/2} - \eta_{k-1/2}) \quad (14.104b)$$

$$p_{1/2} = p_a, \quad (14.104c)$$

with  $p_a$  the applied (or atmospheric) pressure at the ocean surface. In addition,  $z = \eta_{k+1/2}(x, y, t)$  is the vertical position of the lower interface, and  $z = \eta_{k-1/2}(x, y, t)$  is the vertical position of the upper interface, with these fields illustrated in Figure 14.4. Since we are working with a shallow water fluid layer, the pressure at a position within the layer,  $p_k(x, y, z, t)$ , is a linear function of vertical position within the layer (equation (14.104a)) so that its horizontal gradient is independent of vertical position within the layer,

$$\nabla_h p_k = \nabla_h p_{k-1/2} + g \rho_k \nabla_h \eta_{k-1/2}. \quad (14.105)$$

### Interfacial form stress plus gradient of layer depth integrated pressure

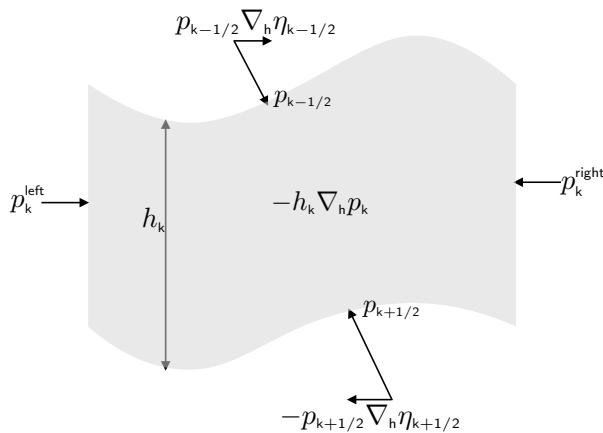


FIGURE 14.5: Illustrating the decomposition (14.106) for the thickness weighted pressure gradient acting within a shallow water layer. This volume integrated horizontal pressure body force can be decomposed into the area integrated pressure contact forces acting on the vertical sides, plus the area integrated pressure form stresses acting on the top and bottom interfaces. The pressure form stresses are drawn here as the horizontal components to the pressure contact stresses.

The first expression for thickness weighted pressure gradient is given by

$$-h_k \nabla_h p_k = -\nabla_h P_k + \mathbf{F}_k^{\text{form}}, \quad (14.106)$$

which is illustrated in Figure 14.5. In this equation we introduced the pressure that is vertically integrated over layer-k

$$P_k = \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} p_k dz = h_k (g \rho_k h_k / 2 + p_{k-1/2}), \quad (14.107)$$

with its negative gradient,<sup>7</sup>

$$-\nabla_h P_k = -(h_k \nabla_h p_{k+1/2} + p_{k-1/2} \nabla_h h_k), \quad (14.108)$$

<sup>7</sup>See Exercise 14.5 for the derivation of equation (14.108).

leading to a horizontal acceleration due to imbalances in the area integrated pressure acting along the vertical sides of the shallow water region. The second stress in equation (14.106) is the pressure form stress acting on the upper and lower layer interfaces

$$\mathbf{F}_k^{\text{form}} = p_{k-1/2} \nabla_h \eta_{k-1/2} - p_{k+1/2} \nabla_h \eta_{k+1/2} \equiv \delta_k (p_{k-1/2} \nabla_h \eta_{k-1/2}), \quad (14.109)$$

where

$$\delta_k \Phi_{k-1/2} \equiv \Phi_{k-1/2} - \Phi_{k+1/2} = -(\Phi_{k+1/2} - \Phi_{k-1/2}) \quad (14.110)$$

is a discrete vertical difference operator acting on interface properties. The use of a backward difference operator is motivated since  $k$  increases down whereas  $\hat{z}$  points up. Additionally, we define the difference operator to only act on fields defined at the layer interface, with layer fields commuting with this operator so that, for example,

$$\delta_k (h_k \eta_{k-1/2}) = h_k \delta_k (\eta_{k-1/2}). \quad (14.111)$$

This convention helps produce a tidy form for the Eliassen-Palm flux in Sections 14.8.4 and 14.8.5.

Making use of the depth integrated pressure and form stress as given by equation (14.106) allows us to write the ensemble mean thickness weighted horizontal pressure gradient

$$-\overline{h_k \nabla_h p_k} = -\nabla_h \overline{P_k} + \delta_k [\overline{p_{k-1/2} \nabla_h \eta_{k-1/2}}]. \quad (14.112)$$

Evidently, the thickness weighted horizontal pressure gradient (left hand side) has been decomposed into a horizontal gradient of the layer depth integrated pressure (first right hand side term), plus the discrete vertical divergence of the form stresses that act on the top and bottom layer interfaces labelled by  $k \pm 1/2$ .

Following equation (14.107), we write the negative gradient of the ensemble mean layer integrated pressure

$$-\nabla_h \overline{P_k} = -\nabla_h [\overline{h_k} (\overline{p_{k-1/2}} + g \rho_k \overline{h_k}/2)] - \nabla_h [\overline{h'_k} (\overline{p'_{k-1/2}} + g \rho_k \overline{h'_k}/2)], \quad (14.113)$$

and the ensemble mean of the form stress vertical divergence

$$\overline{\mathbf{F}_k^{\text{form}}} = \delta_k [\overline{p_{k-1/2} \nabla_h \eta_{k-1/2}}] = \delta_k \left[ \overline{p_{k-1/2} \nabla_h \eta_{k-1/2}} + \overline{p'_{k-1/2} \nabla_h \eta'_{k-1/2}} \right]. \quad (14.114)$$

We are thus led to the following decomposition of the eddy contribution to the thickness weighted pressure gradient

$$-\overline{h'_k \nabla_h p'_k} = -\nabla_h [\overline{h'_k} (\overline{p'_{k-1/2}} + g \rho_k \overline{h'_k}/2)] + \delta_k \left[ \overline{p'_{k-1/2} \nabla_h \eta'_{k-1/2}} \right]. \quad (14.115)$$

The right hand side is much less mathematically compact than the left hand side. The payoff for such verbosity is that the right hand side provides an alternative physical interpretation for the thickness weighted pressure gradient, which we further advance in the following. For orientation, in Figure 14.6 we illustrate the deviations of the interface positions relative to the ensemble mean.

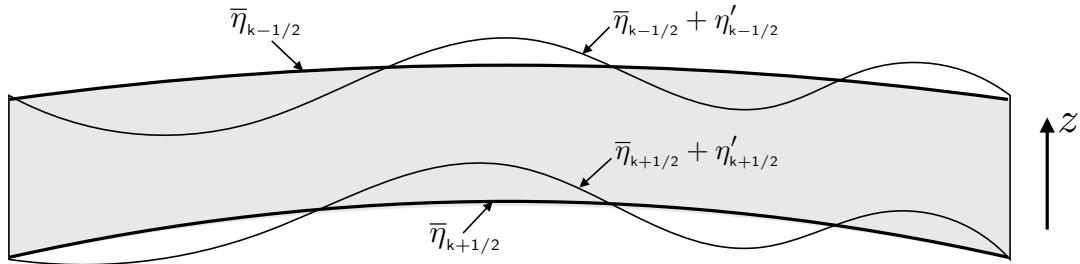


FIGURE 14.6: Schematic of the interface positions for a shallow water layer. The ensemble mean layer interfaces have vertical positions given by  $z = \bar{\eta}_{k\pm 1/2}$ , whereas the fluctuating interfaces are located at  $z = \bar{\eta}_{k\pm 1/2} + \eta'_{k\pm 1/2}$ . As depicted here, the ensemble mean interface positions are not generally horizontal.

### Dual interfacial pressure form stress plus gradient of layer potential energy

An alternative formulation uses the dual form stress and potential energy, in which case we write the thickness weighted horizontal pressure gradient as

$$-h_k \nabla_h p_k = -\nabla_h \mathcal{P}_k + \mathbf{F}_k^{\text{dual}}. \quad (14.116)$$

In this equation we introduced the layer gravitational potential energy per area

$$\mathcal{P}_k = g \rho_k \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} z \, dz = (g \rho_k / 2) (\eta_{k-1/2}^2 - \eta_{k+1/2}^2) = (g \rho_k / 2) \delta_k (\eta_{k-1/2}^2), \quad (14.117)$$

and the **dual form stress** from VOLUME 2.<sup>8</sup>

$$\mathbf{F}_k^{\text{dual}} = -\delta_k (\eta_{k-1/2} \nabla_h p_{k-1/2}) = \mathbf{F}_k^{\text{form}} - \nabla_h [\delta_k (\eta_{k-1/2} p_{k-1/2})]. \quad (14.118)$$

Since they differ by a gradient, the form stress and dual form stress have identical curls and so they contribute the same interfacial pressure torque as part of the layer vorticity evolution

$$-\nabla_h \times (h_k \nabla_h p_k) = \nabla_h \times \mathbf{F}_k^{\text{dual}} = \nabla_h \times \mathbf{F}_k^{\text{form}}. \quad (14.119)$$

However, for the momentum budget it is crucial to note that the form stress and dual form stress are distinct.

Making use of the potential energy and dual form stress as given by equation (14.116) allows us to write the ensemble mean thickness weighted pressure gradient as

$$-\overline{h_k \nabla_h p_k} = -\nabla_h \overline{\mathcal{P}_k} - \delta_k [\overline{\eta_{k-1/2} \nabla_h p_{k-1/2}}], \quad (14.120)$$

where we decomposed the potential energy gradient as

$$-\nabla_h \overline{\mathcal{P}_k} = -(g \rho_k / 2) \delta_k \left[ \nabla_h (\overline{\eta_{k-1/2}})^2 + \nabla_h (\overline{\eta'_{k-1/2}})^2 \right], \quad (14.121)$$

and the vertical convergence of the dual form stress is

$$-\delta_k [\overline{\eta_{k-1/2} \nabla_h p_{k-1/2}}] = -\delta_k \left[ \overline{\eta_{k-1/2}'} \nabla_h \overline{p_{k-1/2}} + \overline{\eta_{k-1/2}' \nabla_h p_{k-1/2}'} \right]. \quad (14.122)$$

<sup>8</sup>The form stress and dual form stress differ by a gradient, so that they are not equal. Indeed, the form stress is a force per area, whereas the dual form stress is not, even though it has the dimensions of a stress. It is unfortunate that the literature often ignores the difference between dual form stress and form stress.

We are thus led to decompose the thickness weighted pressure gradient correlation as

$$-\bar{h}' \nabla_h p'_k = -(g \rho_k / 2) \nabla_h [\delta_k (\eta'_{k-1/2})^2] - \delta_k [\eta'_{k-1/2} \nabla_h p'_{k-1/2}]. \quad (14.123)$$

Again, the first term on the right hand side arises from the eddy potential energy and the second term from the dual eddy form stress.

#### 14.8.4 Zonal and meridional Eliassen-Palm fluxes: version I

Making use of the thickness and pressure gradient correlation in the form of equation (14.115) (the version with the form stress) leads to the thickness weighted momentum equation

$$\begin{aligned} \partial_t (\bar{h} \hat{\mathbf{u}}) + \nabla_h \cdot (\bar{h} \hat{\mathbf{u}} \otimes \hat{\mathbf{u}}) + f \hat{\mathbf{z}} \times \bar{h} \hat{\mathbf{u}} + \bar{h} \nabla_h \bar{p} / \rho_{\text{ref}} \\ = -\nabla_h \cdot (\bar{h} \widehat{\mathbf{u}'' \otimes \mathbf{u}''}) - \rho_{\text{ref}}^{-1} \nabla_h [\bar{h}'_k (p'_{k-1/2} + g \rho_k h'_k / 2)] + \rho_{\text{ref}}^{-1} \delta_k \overline{(p' \nabla_h \eta')_{k-1/2}}, \end{aligned} \quad (14.124)$$

where we only exposed the interface indices to reduce notational clutter, and where we introduced the shorthand for the eddy form stress at the  $k - 1/2$  interface

$$(p' \nabla_h \eta')_{k-1/2} = p'_{k-1/2} \nabla_h \eta'_{k-1/2}. \quad (14.125)$$

The subgrid scale correlations on the right hand side of equation (14.124) can be organized into the divergence of two tensors

$$\begin{aligned} & \nabla_h \cdot (\bar{h} \widehat{\mathbf{u}'' \otimes \mathbf{u}''}) + \rho_{\text{ref}}^{-1} \nabla_h [\bar{h}'_k (p'_{k-1/2} + g \rho_k h'_k / 2)] - (\bar{h} \rho_{\text{ref}})^{-1} \delta_k \left[ \bar{h} \overline{(p' \nabla_h \eta')_{k-1/2}} \right] \\ &= [\partial_x \quad \partial_y \quad (1/\bar{h}) \delta_k] \begin{bmatrix} \bar{h} \widehat{u'' u''} & \bar{h} \widehat{u'' v''} & 0 \\ \bar{h} \widehat{u'' v''} & \bar{h} \widehat{v'' v''} & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &+ \rho_{\text{ref}}^{-1} [\partial_x \quad \partial_y \quad (1/\bar{h}) \delta_k] \begin{bmatrix} \bar{h}'_k (p'_{k-1/2} + g \rho_k h'_k / 2) & 0 & 0 \\ 0 & \bar{h}'_k (p'_{k-1/2} + g \rho_k h'_k / 2) & 0 \\ -\bar{h} \overline{(p' \partial_x \eta')_{k-1/2}} & -\bar{h} \overline{(p' \partial_y \eta')_{k-1/2}} & 0 \end{bmatrix}, \end{aligned} \quad (14.126)$$

where we recall from equation (14.111) that the vertical difference operator,  $\delta_k$ , only acts on layer interface fields so that  $\bar{h}_k$  commutes with  $\delta_k$ . The first tensor in equation (14.126) arises from Reynolds stresses and the second tensor arises from eddy pressures, including the eddy form stress in the third row.

When combined, the columns of the tensors appearing in equation (14.126) are the thickness weighted shallow water *Eliassen-Palm fluxes* for the zonal (column 1) and meridional (column 2) momentum equation

$$\mathbf{E}^{(\text{uEP})} = \left[ \bar{h} \widehat{u'' u''} + \rho_{\text{ref}}^{-1} \bar{h}'_k (p'_{k-1/2} + g \rho_k h'_k / 2) \right] \hat{\mathbf{x}} + \bar{h} \widehat{u'' v''} \hat{\mathbf{y}} - (\bar{h} / \rho_{\text{ref}}) \overline{(p' \partial_x \eta')_{k-1/2}} \hat{\mathbf{z}} \quad (14.127a)$$

$$\mathbf{E}^{(\text{vEP})} = \bar{h} \widehat{u'' v''} \hat{\mathbf{x}} + \left[ \bar{h} \widehat{v'' v''} + \rho_{\text{ref}}^{-1} \bar{h}'_k (p'_{k-1/2} + g \rho_k h'_k / 2) \right] \hat{\mathbf{y}} - (\bar{h} / \rho_{\text{ref}}) \overline{(p' \partial_y \eta')_{k-1/2}} \hat{\mathbf{z}}. \quad (14.127b)$$

The thickness-weighted Eliassen-Palm flux has physical dimensions of thickness times velocity squared. We encounter the unaveraged version of the Eliassen-Palm flux when studying the shallow water momentum equation in Chapter 2. *Maddison and Marshall* (2013) included the third column of zeros in equation (14.126) to emphasize that the Eliassen-Palm fluxes are the

first and second columns to the *Eliassen-Palm flux tensor*. They illustrated the utility of this perspective by considering gauge transformations that result in non-zero elements in the third column.

The Eliassen-Palm fluxes are second order in eddy amplitude; i.e., they are quadratic in eddy fluctuations. Furthermore, they bring together the Reynolds stress and eddy pressure terms, including the eddy form stress. The convergence of the Eliassen-Palm fluxes provides an eddy rectified acceleration to the thickness weighted velocity. To explicitly see this forcing, we write the components to the mean field momentum equation (14.124) as<sup>9</sup>

$$\partial_t(\bar{h}\hat{\mathbf{u}}) + \nabla_h \cdot (\bar{h}\hat{\mathbf{u}}\hat{\mathbf{u}}) - f\bar{h}\hat{v} + \bar{h}\partial_x\bar{p}/\rho = -(\nabla_h + \hat{z}\bar{h}^{-1}\delta_k) \cdot \mathbf{E}^{(uEP)} \quad (14.128a)$$

$$\partial_t(\bar{h}\hat{v}) + \nabla_h \cdot (\bar{h}\hat{\mathbf{u}}\hat{v}) + f\bar{h}\hat{u} + \bar{h}\partial_y\bar{p}/\rho = -(\nabla_h + \hat{z}\bar{h}^{-1}\delta_k) \cdot \mathbf{E}^{(vEP)}. \quad (14.128b)$$

Equations (14.128a) and (14.128b) are isomorphic to the unaveraged horizontal momentum equation (14.50c), yet with the addition of the convergence of the Eliassen-Palm flux on the right hand side that encapsulates rectified effects from eddies. They can be written using the material time derivative (14.85)

$$\frac{D^\# \hat{\mathbf{u}}}{Dt} - f\hat{v} + \partial_x\bar{p}/\rho = -\bar{h}^{-1}(\nabla_h + \hat{z}\bar{h}^{-1}\delta_k) \cdot \mathbf{E}^{(uEP)} \quad (14.129a)$$

$$\frac{D^\# \hat{v}}{Dt} + f\hat{u} + \partial_y\bar{p}/\rho = -\bar{h}^{-1}(\nabla_h + \hat{z}\bar{h}^{-1}\delta_k) \cdot \mathbf{E}^{(vEP)}. \quad (14.129b)$$

We emphasize that these equations only make use of the thickness weighted velocity,  $\hat{\mathbf{u}}$ , as do the averaged thickness equation (14.84) and averaged tracer equation (14.91). We advertised this point near the start of this chapter, noting that it facilitates the practical use of the TWA equations for numerical simulations. We further this correspondence in Section 14.9 by showing that the vorticity and potential vorticity equations also make use only of  $\hat{\mathbf{u}}$ .

#### 14.8.5 Zonal and meridional Eliassen-Palm fluxes: version II

We here follow the development in Section 14.8.4, only now making use of the thickness and pressure gradient correlation in the form of equation (14.123) (the version with the dual form stress). Our presentation is terse since there are few differences from Section 14.8.4. We start from the thickness weighted momentum equation

$$\begin{aligned} \partial_t(\bar{h}\hat{\mathbf{u}}) + \nabla_h \cdot (\bar{h}\hat{\mathbf{u}} \otimes \hat{\mathbf{u}}) + f\hat{z} \times \bar{h}\hat{\mathbf{u}} + \bar{h}\nabla_h\bar{p}/\rho_{ref} \\ = -\nabla_h \cdot (\widehat{\bar{h}\mathbf{u}'' \otimes \mathbf{u}''}) - (g\rho_k/2\rho_{ref})\nabla_h[\delta_k\overline{(\eta'_{k-1/2})^2}] - \delta_k[\overline{(\eta'\nabla_h p')_{k-1/2}}/\rho_{ref}]. \end{aligned} \quad (14.130)$$

The subgrid scale correlations on the right hand side can be organized into the divergence of two tensors

$$\begin{aligned} & \nabla_h \cdot (\widehat{\bar{h}\mathbf{u}'' \otimes \mathbf{u}''}) + (g\rho_k/2\rho_{ref})\nabla_h[\delta_k\overline{(\eta'_{k-1/2})^2}] + \delta_k[\overline{(\eta'\nabla_h p')_{k-1/2}}/\rho_{ref}] \\ &= [\partial_x \quad \partial_y \quad \bar{h}^{-1} \delta_k] \begin{bmatrix} \widehat{\bar{h}\mathbf{u}'' \mathbf{u}''} & \widehat{\bar{h}\mathbf{u}'' \mathbf{v}''} & 0 \\ \widehat{\bar{h}\mathbf{v}'' \mathbf{u}''} & \widehat{\bar{h}\mathbf{v}'' \mathbf{v}''} & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

<sup>9</sup>Recall from equation (14.111) that the operator  $\delta_k$  only acts on interface fields, so that  $\bar{h}^{-1}\delta_k(\bar{h}) = \delta_k$ . This convention allows us to combine the horizontal components to the Eliassen-Palm flux with the vertical component, as written in equations (14.132a) and (14.132b).

$$+ \begin{bmatrix} \partial_x & \partial_y & \bar{h}^{-1} \delta_k \end{bmatrix} \begin{bmatrix} (g \rho_k / 2 \rho_{\text{ref}}) [\delta_k \overline{(\eta'_{k-1/2})^2}] & 0 & 0 \\ 0 & (g \rho_k / 2 \rho_{\text{ref}}) [\delta_k \overline{(\eta'_{k-1/2})^2}] & 0 \\ (\bar{h} / \rho_{\text{ref}}) \overline{(\eta' \partial_x p')_{k-1/2}} & (\bar{h} / \rho_{\text{ref}}) \overline{(\eta' \partial_y p')_{k-1/2}} & 0 \end{bmatrix}. \quad (14.131)$$

The first tensor arises from Reynolds stresses and the second arises from eddy potential energy and dual eddy form stresses. When combined, the columns are the thickness weighted *Eliassen-Palm fluxes* for the zonal (column 1) and meridional (column 2), here making use of the dual form stress

$$\mathbf{E}_{\text{dual}}^{(uEP)} = \left[ \bar{h} \widehat{u'' u''} + (g \rho_k / 2 \rho_{\text{ref}}) [\delta_k \overline{(\eta'_{k-1/2})^2}] \right] \hat{\mathbf{x}} + \bar{h} \widehat{u'' v''} \hat{\mathbf{y}} + \bar{h} \overline{(\eta' \partial_x p')_{k-1/2}} / \rho_{\text{ref}} \hat{\mathbf{z}} \quad (14.132a)$$

$$\mathbf{E}_{\text{dual}}^{(vEP)} = \bar{h} \widehat{u'' v''} \hat{\mathbf{x}} + \left[ \bar{h} \widehat{v'' v''} + (g \rho_k / 2 \rho_{\text{ref}}) [\delta_k \overline{(\eta'_{k-1/2})^2}] \right] \hat{\mathbf{y}} + \bar{h} \overline{(\eta' \partial_y p')_{k-1/2}} / \rho_{\text{ref}} \hat{\mathbf{z}}. \quad (14.132b)$$

The convergence of the Eliassen-Palm fluxes provides an eddy rectified acceleration on the thickness weighted velocity

$$\partial_t (\bar{h} \hat{u}) + \nabla_h \cdot (\bar{h} \hat{\mathbf{u}} \hat{u}) - f \bar{h} \hat{v} + \bar{h} \partial_x \bar{p} / \rho = -(\nabla_h + \hat{\mathbf{z}} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}_{\text{dual}}^{(uEP)} \quad (14.133a)$$

$$\partial_t (\bar{h} \hat{v}) + \nabla_h \cdot (\bar{h} \hat{\mathbf{u}} \hat{v}) + f \bar{h} \hat{u} + \bar{h} \partial_y \bar{p} / \rho = -(\nabla_h + \hat{\mathbf{z}} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}_{\text{dual}}^{(vEP)}. \quad (14.133b)$$

### 14.8.6 Interfacial stresses from geostrophic eddies

In VOLUME 2 we study the rectified effects from geostrophic eddies in a zonally re-entrant channel for a continuously stratified fluid, where we show that the zonal mean of isopycnal eddy form stresses is equivalent to the meridional eddy flux of buoyancy. We consider similar questions within the context of the TWA shallow water fluid, here focusing on the interfacial transfer of momentum due to eddy dual form stresses as given by the vertical vectors

$$\rho_{\text{ref}} \left[ \mathbf{E}_{\text{dual}}^{(uEP)} \right]_{\text{interface}} = \hat{\mathbf{z}} \bar{h} \overline{(\eta' \partial_x p')_{k-1/2}} \quad (14.134a)$$

$$\rho_{\text{ref}} \left[ \mathbf{E}_{\text{dual}}^{(vEP)} \right]_{\text{interface}} = \hat{\mathbf{z}} \bar{h} \overline{(\eta' \partial_y p')_{k-1/2}}. \quad (14.134b)$$

Let us now write the interface pressure gradient fluctuation as

$$\nabla_h p_{k-1/2} = \nabla_h p_k - g \rho_k \nabla_h \eta_{k-1/2} \quad (14.135)$$

so that

$$\overline{(\eta' \nabla_h p')_{k-1/2}} = \overline{\eta'_{k-1/2} \nabla_h p'_k} + g \rho_k \overline{(\eta' \nabla_h \eta')_{k-1/2}}. \quad (14.136)$$

We now assume that the fluctuations are in **geostrophic balance**. Doing so allows us to introduce the layer geostrophic velocity corresponding to the gradient of the layer pressure fluctuations

$$\partial_x p'_k = \rho_{\text{ref}} f v'_k \quad \text{and} \quad \partial_y p'_k = -\rho_{\text{ref}} f u'_k. \quad (14.137)$$

The dual form stress portion of the Eliassen-Palm fluxes now takes the form

$$\rho_{\text{ref}} \left[ \mathbf{E}_{\text{dual}}^{(uEP)} \right]_{\text{interface}} = \hat{\mathbf{z}} \bar{h} \left[ f \rho_{\text{ref}} \overline{\eta'_{k-1/2} v'_k} + g \rho_k \overline{(\eta' \partial_x \eta')_{k-1/2}} \right] \quad (14.138a)$$

$$\rho_{\text{ref}} \left[ \mathbf{E}_{\text{dual}}^{(vEP)} \right]_{\text{interface}} = \hat{\mathbf{z}} \bar{h} \left[ -f \rho_{\text{ref}} \overline{\eta'_{k-1/2} u'_k} + g \rho_k \overline{(\eta' \partial_y \eta')_{k-1/2}} \right]. \quad (14.138b)$$

The  $\overline{\eta'_{k-1/2} \mathbf{u}'_k}$  term is an eddy transport of the area between  $z = \bar{\eta}_{k-1/2}$  and  $z = \eta'_{k-1/2}$  (see Figure 14.6). We studied the same transport for the continuously stratified fluid in VOLUME 2. In that discussion, we found that the interface fluctuations,  $\eta'$ , can be related to the buoyancy fluctuations,  $b'$ , in which case  $\overline{\eta'_{k-1/2} \mathbf{u}'_k}$  is proportional to the eddy buoyancy flux for the layer.

### 14.8.7 Comments

*Greatbatch and Lamb* (1990) and *Greatbatch* (1998) pursue a similar analysis for the purpose of framing the mesoscale eddy parameterization problem. They focus on the interfacial form stress contribution since, for geostrophic eddies, it dominates over the other terms in the Eliassen-Palm fluxes (14.132a) and (14.132b).

## 14.9 Vorticity and potential vorticity

We follow the procedure from our discussion of shallow water vorticity in Chapter 5 to derive the vorticity and potential vorticity for the thickness weighted shallow water equations. In the process, we connect the eddy flux of potential vorticity to the Eliassen-Palm fluxes (14.127a) and (14.127b). Note that the same manipulations also hold for the dual Eliassen-Palm fluxes (14.132a) and (14.132b).

### 14.9.1 Derivation

We make use of the vector identities from our study of the shallow water equation in Section 1.5 to bring the material evolution equations (14.129a) and (14.129b) into their equivalent vector invariant forms<sup>10</sup>

$$\partial_t \hat{u} - (f + \hat{\zeta}) \hat{v} = -\partial_x (\bar{p}/\rho + \hat{u} \cdot \hat{u}/2) - \bar{h}^{-1} (\nabla_h + \hat{z} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}^{(uEP)} \quad (14.139a)$$

$$\partial_t \hat{v} + (f + \hat{\zeta}) \hat{u} = -\partial_y (\bar{p}/\rho + \hat{u} \cdot \hat{u}/2) - \bar{h}^{-1} (\nabla_h + \hat{z} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}^{(vEP)}, \quad (14.139b)$$

where we introduced the relative vorticity of the thickness weighted horizontal velocity

$$\hat{\zeta} = \partial_x \hat{v} - \partial_y \hat{u}. \quad (14.140)$$

Taking  $\partial_x$  of the meridional equation (14.139b) and subtracting  $\partial_y$  of the zonal equation (14.139a) renders the evolution equation for absolute vorticity,  $\hat{\zeta}_a = \hat{\zeta} + f$ ,

$$\frac{D^\# \hat{\zeta}_a}{Dt} + \hat{\zeta}_a \nabla_h \cdot \hat{\mathbf{u}} = \partial_y [\bar{h}^{-1} (\nabla_h + \hat{z} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}^{(uEP)}] - \partial_x [\bar{h}^{-1} (\nabla_h + \hat{z} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}^{(vEP)}]. \quad (14.141)$$

Making use of the thickness equation (14.86) to replace  $\nabla_h \cdot \hat{\mathbf{u}}$  leads to the potential vorticity equation

$$\bar{h} \frac{D^\# \Pi^\#}{Dt} = -\nabla_h \cdot \mathbf{F}^\# \quad (14.142)$$

where

$$\Pi^\# = \frac{f + \partial_x \hat{v} - \partial_y \hat{u}}{\bar{h}} = \frac{f + \hat{\zeta}}{\bar{h}} \quad (14.143)$$

<sup>10</sup>In Section D.6 of *Griffies et al.* (2020), the authors state “In contrast to the flux-form momentum equation, the vector-invariant velocity equation does not admit a finite volume formulation.” That statement is incorrect, with equations (14.139a) and (14.139b) the finite volume vector-invariant velocity equation.

is the potential vorticity defined with the thickness weighted velocity and ensemble mean thickness. The corresponding eddy potential vorticity flux is a horizontal vector that is written in terms of the divergence of the Eliassen-Palm fluxes

$$\mathbf{F}^\sharp = \hat{\mathbf{x}} [\bar{h}^{-1} (\nabla_h + \hat{\mathbf{z}} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}^{(vEP)}] - \hat{\mathbf{y}} [\bar{h}^{-1} (\nabla_h + \hat{\mathbf{z}} \bar{h}^{-1} \delta_k) \cdot \mathbf{E}^{(uEP)}] + \hat{\mathbf{z}} \times \nabla_h \Upsilon, \quad (14.144)$$

where  $\Upsilon$  is an arbitrary gauge function.<sup>11</sup> This equation connects the potential vorticity flux to the Eliassen-Palm fluxes and it is known as the *Taylor-Bretherton identity*. Remarkably, the potential vorticity flux also provides the eddy forcing to the thickness weighted velocity equation

$$\partial_t \hat{\mathbf{u}} + (f + \hat{\zeta}) \hat{\mathbf{z}} \times \hat{\mathbf{u}} + \nabla_h (\bar{p}/\rho + \hat{\mathbf{u}} \cdot \hat{\mathbf{u}}/2) = -\hat{\mathbf{z}} \times (\mathbf{F}^\sharp - \hat{\mathbf{z}} \times \nabla_h \Upsilon), \quad (14.145)$$

which can also be written

$$\partial_t \hat{\mathbf{u}} + \hat{\mathbf{z}} \times (\bar{h} \hat{\mathbf{u}} \Pi^\sharp + \mathbf{F}^\sharp - \hat{\mathbf{z}} \times \nabla_h \Upsilon) + \nabla_h (\bar{p}/\rho + \hat{\mathbf{u}} \cdot \hat{\mathbf{u}}/2) = 0, \quad (14.146)$$

where  $\bar{h} \hat{\mathbf{u}} \Pi^\sharp + \mathbf{F}^\sharp - \hat{\mathbf{z}} \times \nabla_h \Upsilon$  is the net (mean plus eddy plus gauge) potential vorticity flux.

### 14.9.2 Concerning the mean field potential vorticity

We emphasize that the mean field potential vorticity arising from our development is  $\Pi^\sharp$ , which is defined by equation (14.143) using the thickness weighted velocity,  $\hat{\mathbf{u}}$  for the relative vorticity. This potential vorticity is distinct from the thickness weighted average potential vorticity

$$\hat{\Pi} = \frac{\bar{\Pi} \bar{h}}{\bar{h}} = \frac{f + \bar{\zeta}}{\bar{h}} = \frac{f + \partial_x \bar{v} - \partial_y \bar{u}}{\bar{h}}, \quad (14.147)$$

which is the mean field potential vorticity considered by [Greatbatch \(1998\)](#) and [Peterson and Greatbatch \(2001\)](#). The two forms of potential vorticity differ by the potential vorticity of the bolus velocity

$$\Pi^\sharp - \hat{\Pi} = \frac{(f + \hat{\zeta}) - (f + \bar{\zeta})}{\bar{h}} = \frac{\hat{\mathbf{z}} \cdot [\nabla_h \times (\hat{\mathbf{u}} - \bar{\mathbf{u}})]}{\bar{h}} = \frac{\hat{\mathbf{z}} \cdot (\nabla_h \times \hat{\mathbf{u}}')}{\bar{h}} = \frac{\hat{\mathbf{z}} \cdot (\nabla_h \times \mathbf{u}^{\text{bolus}})}{\bar{h}}, \quad (14.148)$$

where the penultimate equality made use of equation (14.81) for  $\hat{\mathbf{u}}'$ , and the final equality introduced the bolus velocity according to equation (14.88). Use of  $\Pi^\sharp$  allows us to develop a potential vorticity conservation statement solely in terms of  $\hat{\mathbf{u}}$ , whereas the use of  $\hat{\Pi}$  by [Greatbatch \(1998\)](#) and [Peterson and Greatbatch \(2001\)](#) requires both  $\bar{\mathbf{u}}$  and  $\hat{\mathbf{u}}$ .

### 14.9.3 Comments

As in [Young \(2012\)](#), and as advertised in Section 14.4, we have developed the full set of mechanical equations for the TWA shallow water solely in terms of the thickness weighted velocity,  $\hat{\mathbf{u}}$ . This development includes the thickness equation (14.84), the tracer equation (14.91), the velocity equation (14.95) and the potential vorticity equation (14.142). There is no need for the ensemble mean velocity,  $\bar{\mathbf{u}}$ , and thus no need to parameterize the bolus velocity.

<sup>11</sup>Equation (129) in [Young \(2012\)](#) should have a gauge function on its right hand side, which follows from his footnote #3.

## 14.10 Vorticity fluxes for non-divergent barotropic flow

In our study of the horizontally non-divergent barotropic flow in Chapter 4, we encountered the mechanics of a two dimensional fluid whose horizontal flow is non-divergent. As for the shallow water, the non-divergent flow moves as vertical columns. However, since the horizontal flow is non-divergent, each column is rigid so there is no stretching or squeezing of columns. Correspondingly, there are no form stresses acting on these columns since the top and bottom interfaces are flat. We here briefly specialize the shallow water vorticity analysis in this section to consider rigid columnar motion.

For rigid fluid columns, the thickness weighted average reduces to just the ensemble mean since all layer thicknesses are fixed constants. Correspondingly, there are no form stresses acting at the layer interfaces since the layer interfaces are horizontal. Hence, the Eliassen-Palm fluxes (14.127a) and (14.127b) reduce to just their Reynolds stress contributions

$$h^{-1} \mathbf{E}^{(uEP)} = \overline{u' u'} \hat{\mathbf{x}} + \overline{u' v'} \hat{\mathbf{y}} \quad (14.149a)$$

$$h^{-1} \mathbf{E}^{(vEP)} = \overline{u' v'} \hat{\mathbf{x}} + \overline{v' v'} \hat{\mathbf{y}}. \quad (14.149b)$$

The corresponding eddy potential vorticity flux (14.144), absent the gauge term, is

$$\mathbf{F}^\sharp = \hat{\mathbf{x}} \nabla_h \cdot [\overline{u' v'} \hat{\mathbf{x}} + \overline{v' v'} \hat{\mathbf{y}}] - \hat{\mathbf{y}} \nabla_h \cdot [\overline{u' u'} \hat{\mathbf{x}} + \overline{u' v'} \hat{\mathbf{y}}] \quad (14.150a)$$

$$= \hat{\mathbf{x}} [\partial_x (\overline{u' v'}) + \partial_y (\overline{v' v'})] - \hat{\mathbf{y}} [\partial_x (\overline{u' u'}) + \partial_y (\overline{u' v'})]. \quad (14.150b)$$

Does the eddy potential vorticity flux (14.150b) agree, to within a gauge function, with the eddy flux resulting from a direct decomposition into eddy and mean within a two dimensional non-divergent model? To address this question, note that from Chapter 4, the advective flux of potential vorticity for the two dimensional non-divergent flow is given by

$$\mathbf{u} q = \mathbf{u} f + \nabla_h \cdot (\hat{\mathbf{z}} \times \mathcal{E}), \quad (14.151)$$

where  $\mathcal{E}$  is the trace-free anisotropic portion of the kinetic stress tensor

$$\mathcal{E} = \begin{bmatrix} -(u^2 - v^2)/2 & -uv \\ -uv & (u^2 - v^2)/2 \end{bmatrix}. \quad (14.152)$$

The mean of the potential vorticity flux is (14.151) is given by

$$\overline{\mathbf{u} q} = \overline{\mathbf{u} f} + \overline{\mathbf{u}' q'}, \quad (14.153)$$

where the flux computed from the mean fields is

$$\overline{\mathbf{u} q} = \overline{\mathbf{u}} (f + \bar{\zeta}), \quad (14.154)$$

whereas the eddy potential vorticity flux is

$$\overline{\mathbf{u}' q'} = \mathbf{F}^\sharp - \hat{\mathbf{z}} \times \nabla_h (\overline{\mathbf{u}' \cdot \mathbf{u}'})/2, \quad (14.155)$$

with the missing steps for this derivation provided in Exercise (14.6). Hence,  $\overline{\mathbf{u}' q'}$  agrees with  $\mathbf{F}^\sharp$  in equation (14.150b) to within a gauge function given by the rotated gradient of the eddy

kinetic energy per mass, so that their divergences are equal

$$\nabla_h \cdot \mathbf{F}^\sharp = \nabla_h \cdot \overline{\mathbf{u}' q'}. \quad (14.156)$$

That is, when diagnosing contributions to the potential vorticity flux, the gauge term,  $-\hat{\mathbf{z}} \times \nabla_h(\overline{\mathbf{u}' \cdot \mathbf{u}'})/2$ , plays no role in forcing potential vorticity.



## 14.11 Exercises

### EXERCISE 14.1: DERIVATION OF THE VORTICITY EQUATION

Derive the isopycnal version of the vorticity equation

$$\left[ \frac{\partial \tilde{\zeta}_a}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_{hb}) \tilde{\zeta}_a = -\tilde{\zeta}_a \nabla_{hb} \cdot \mathbf{u}, \quad (14.157)$$

with this equation used in Section 14.3.1 for deriving the potential vorticity equation.

### EXERCISE 14.2: DERIVATION OF THE POTENTIAL VORTICITY EQUATION

Derive the potential vorticity equation,

$$\frac{DQ}{Dt} = 0, \quad (14.158)$$

discussed in Section 14.3.2.

### EXERCISE 14.3: COORDINATE TRANSFORMATION OF THE POTENTIAL VORTICITY

Show that a coordinate transformation from geopotential coordinates to isopycnal coordinates bring the potential vorticity,  $Q = (\boldsymbol{\omega}^{hy} + f \hat{\mathbf{z}}) \cdot \nabla_h b$ , into the isopycnal coordinate form of equation (14.37).

### EXERCISE 14.4: POTENTIAL VORTICITY EQUATION WITH IRREVERSIBLE PROCESSES

In Section 14.3.3 we derived the potential vorticity equation in the presence of irreversible processes. We here consider an alternative derivation that is directly analogous to Exercise 14.3. Namely, start from the discussion in Chapter 7 for the material evolution of potential vorticity in a hydrostatic and Boussinesq fluid

$$\frac{DQ}{Dt} = \nabla_h \cdot \left[ (f \hat{\mathbf{z}} + \boldsymbol{\omega}_{hy}) \dot{b} + b \nabla_h \times \mathbf{F}^h \right], \quad (14.159)$$

where

$$Q = \boldsymbol{\omega}_a^{hy} \cdot \nabla_h b = \boldsymbol{\omega}^{hy} \cdot \nabla_h b + f \frac{\partial b}{\partial z} \quad \text{and} \quad \boldsymbol{\omega}_{hy} = -\hat{\mathbf{x}} \frac{\partial v}{\partial z} + \hat{\mathbf{y}} \frac{\partial u}{\partial z} + \hat{\mathbf{z}} \left[ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right]. \quad (14.160)$$

Perform a coordinate transformation from this geopotential form into its isopycnal form, thus confirming the potential vorticity equation (14.47) that we derived from the isopycnal version of the equations of motion. Hint: this exercise is a bit tedious but straightforward.

### EXERCISE 14.5: DERIVING THE EXPRESSION FOR $\nabla_h P_k$

Fill in the details for the derivation of equation (14.108) for  $\nabla_h P_k$ . Hint: make use of the identities summarized in Figure 14.4.

## EXERCISE 14.6: VORTICITY FLUXES FOR NON-DIVERGENT BAROTROPIC FLOW

Fill in the missing steps needed to derive equation (14.155) for  $\overline{u' q'}$  in the two dimensional non-divergent barotropic flow.





# **Part V**

## **Scalar fields**

In this part of the book, we study the physics and maths of **passive tracers**, **conservative tracers**,<sup>12</sup> as well as density and **Archimedean buoyancy**. Tensorially, these fields are scalars and so they provide a number (e.g., temperature, humidity, mass density) throughout the continuum fluid. This study of **scalar mechanics** complements that of momentum, vorticity, and energy considered in other parts of this book, with each scalar offering information about the mechanics of fluid motion. Much of the material is relevant to both the atmosphere and ocean, though specialized topics are motivated from ocean applications.

Although the physics and maths of scalar fields are simpler than that describing momentum, vorticity, and energy, there is a remarkable richness to the study. We only touch upon a few of the many topics, aiming to provide a theoretical platform for further study by the interested reader. Here is a synopsis of the chapters in this part of the book.

- In Chapter 15 we focus on tracer diffusion in the absence of advection. There is a long and rich history of research into diffusive (or conductive) processes across science and engineering, with books such as *Crank (1956)* and *Carslaw and Jaeger (1959)* offering a wealth of theoretical results and mathematical methods.
- In Chapter 16 we consider advection along with diffusion in affecting the evolution of tracer concentration. Advection results through viewing fluid flow from within the Eulerian reference frame, rather than the Lagrangian material frame. When acting alone on a conservative tracer, advection affects a reversible stirring of tracer concentration that can increase tracer gradients. When diffusion is included along with advection, reversibility is lost and tracer gradients increase or decrease depending on the relative dominance of advection or diffusion.
- In Chapter 17 we introduce notions of wave-mean flow interactions that give rise to eddy-induced advection (or skew diffusion) as well as diffusion. This chapter, which mostly focuses on kinematic properties and is restricted to tracers, makes use of both geopotential coordinates as well as isopycnal coordinate equations from Chapter 14. Doing so provides examples of the dual roles these two vertical coordinate choices fill for describing turbulent geophysical flows.
- In Chapter 18 we study elements of tracer parameterizations used for coarse resolution models of the ocean circulation. We particularly focus on a variety of mathematical properties of the parameterizations, and unpack the physics embodied within the mathematics. This chapter exposes a handful of questions at the leading edge of research.
- In Chapter 19 we consider ocean density and the budget for global sea level. This study prompts us to dive into the niceties of the enthalpy (heat), salt, and mass budgets for the ocean. These budgets are central to climate science since the ocean is the dominant sink of the anthropogenically induced increase in planetary enthalpy, and with this increase affecting a global rise in sea level.
- In Chapter 20 we present fundamental elements of water mass transformation analysis, which offers a view on ocean circulation that complements those available from Eulerian and Lagrangian kinematics. Many of the methods of ocean water mass analysis are relevant to atmospheric analyses as well, though this chapter is written from an ocean

---

<sup>12</sup>Conservative tracers evolve only via the convergence of advective and diffusive fluxes within the fluid interior, along with boundary conditions. Conservative tracers have no interior sources or sinks, so the net content of a conservative tracer over any finite volume domain is affected only through transport across boundaries.

perspective. Furthermore, this chapter is arguably the toughest in this part of the book, with progress in understanding water mass transformation theory sometimes taking years to ponder the concepts and apply the methods.

#### MATHEMATICS AND KINEMATICS IN THIS PART

The mathematics in this part of the book rely mostly on the Cartesian tensor analysis and vector calculus from VOLUME 1, though with some exceptions such as the generalized Laplacian operator encountered with the diffusion equation. Additionally, we start this part with a chapter diving into the fundamentals of parabolic partial differential equations and Green's functions. This material again relies mostly on Cartesian vector calculus, though it does provide some heavy lifting particularly in the Green's function method. For water mass analysis, we make use of rudimentary ideas from generalized vertical coordinates studied earlier in Part IV of this volume. In regards to fluid kinematics, we mostly stay within an Eulerian perspective, though at times we touch upon a bit of Lagrangian methods such as when introducing the generalized Lagrangian mean in the context of wave-mean flow interactions in Chapter 17.



## Chapter 15

### DIFFUSIVE PROCESSES

In this chapter we study the physics of diffusion with a focus on how diffusion affects properties of geophysical fluids. We start with a brief discussion of molecular diffusion, which is the most basic form of diffusion affecting matter concentrations, temperature, momentum, and other properties. The continuum approximation proposes that a macroscopic description of fluid motion does not require direct information about the motion of individual molecules. Nonetheless, random molecular motion and properties of the constituent molecules impact on fluid properties through molecular diffusion. The presence of molecular diffusion signals a system that is not in thermodynamic equilibrium, with diffusion affecting an irreversible exchange of properties between fluid elements.

Brownian motion refers to the transport of relatively large pieces of matter, such as dust and pollen, by the random effects from molecular impulses. Although not central to the property distributions in geophysical fluids, the description of Brownian motion introduces some generic physical and mathematical concepts that appear throughout the physics of diffusion and irreversible processes, thus prompting its study in this chapter. In particular, it presages the ideas of turbulent diffusion introduced by [Taylor \(1922\)](#). A successful mathematical physics description of Brownian motion originates from [Einstein \(1905a\)](#), whose statistical formulation planted the intellectual seeds for the study of stochastic physics.<sup>1</sup> His description offers a particle-level mechanism behind diffusive spreading of relatively large particles, whereas molecular diffusion is concerned with how matter, heat, and momentum diffuse through media. They both rely on molecular chaos, but are concerned with the transport of distinct material properties.

Turbulent flows provide a stirring of fluids that affect the macroscopic mixing rates found in geophysical fluids. Turbulent stirring acts to enhance the magnitude of property gradients, thus offering the means to increase the efficiency of molecular diffusion that acts to irreversibly mix. We only briefly discuss the notions of such turbulent diffusion by summarizing the ideas formulated by [Taylor \(1922\)](#).

---

<sup>1</sup>The year 1905 was Einstein's *Annus mirabilis*. During that year, he published his Brownian motion paper ([Einstein, 1905a](#)), which provided direct evidence for the atomic nature of matter; his special relativity paper ([Einstein, 1905c](#)), which modified how we think about space and time; and his photoelectric effect paper ([Einstein, 1905b](#)), which provided evidence for the quantum nature of light.

---

## CHAPTER GUIDE

This concepts detailed in this chapter are basic to how we consider diffusion at a fundamental as well as phenomenological level. Chapter 16 is a direct descendant of the current chapter, where we study the combined effects of advection plus diffusion. Exposure to basic properties of [parabolic partial differential equations](#) from VOLUME 1 can be useful for a grounding in the mathematics of diffusion.

---

<b>15.1</b>	<b>Molecular diffusion</b>	<b>513</b>
15.1.1	Kinetic theory and phenomenological laws	513
15.1.2	Comments and further study	513
<b>15.2</b>	<b>Brownian motion</b>	<b>514</b>
15.2.1	Einstein's formulation as a Markov process	514
15.2.2	The Chapman-Kolmogorov equation	515
15.2.3	Diffusion in the continuum limit	515
15.2.4	Moments of the particle density	516
15.2.5	Langevin equation	518
15.2.6	Dissipation time scale	519
15.2.7	Ensemble statistics from the Langevin equation	519
15.2.8	Further study	521
<b>15.3</b>	<b>Turbulent diffusion</b>	<b>521</b>
15.3.1	Introducing Taylor's Lagrangian description	521
15.3.2	Comments and further study	523
<b>15.4</b>	<b>Phenomenological diffusion laws</b>	<b>523</b>
15.4.1	Fick's law for matter diffusion	523
15.4.2	Fourier's law of conduction	525
15.4.3	Newtonian frictional stress and momentum diffusion	526
15.4.4	Comments	526
<b>15.5</b>	<b>The scale selectivity of Laplacian diffusion</b>	<b>527</b>
<b>15.6</b>	<b>Gaussian concentration generated by a Dirac source</b>	<b>528</b>
<b>15.7</b>	<b>Mathematical interlude: tensor analysis tools</b>	<b>529</b>
15.7.1	Concerning the upright versus slanted notation	529
15.7.2	Metric tensor allows us to measure distance	530
15.7.3	Raising and lowering tensor indices via the metric tensor	530
15.7.4	Divergence of a vector and the divergence theorem	531
15.7.5	Example tracer fluxes	532
15.7.6	Comments about the tensor tools	533
<b>15.8</b>	<b>Further properties of tracer diffusion</b>	<b>533</b>
15.8.1	Sample diffusion tensors	534
15.8.2	Diffusion of tracer concentration powers	534
15.8.3	Moments of tracer concentration	535
<b>15.9</b>	<b>Connecting tracer dissipation to the diffusion operator</b>	<b>537</b>
15.9.1	Fréchet derivative of the diffusion dissipation functional	537
15.9.2	Connection to the diffusion operator	539
15.9.3	Why we need to assume $K^{mn}$ is independent of $C$	539
15.9.4	Relation to Hamilton's principle	539
<b>15.10</b>	<b>Exercises</b>	<b>539</b>

---

## 15.1 Molecular diffusion

Consider a fluid comprised of a single matter constituent, such as a lake of pure H<sub>2</sub>O. As discussed in the kinematics of VOLUME 1, for a macroscopic description of this single-component fluid, a constant mass fluid element is identical to a constant mass material fluid parcel. Hence, there is no mixing of matter since there is just a single matter component. Now place a **passive tracer** (e.g., a dye tracer) into a corner of the lake so that the lake is comprised of two material components (H<sub>2</sub>O and dye). Even in the absence of ambient macroscopic fluid motion, the random motion of water and dye molecules produces an exchange of matter constituents between fluid elements. Consequently, the dye spreads outward from its initial position; i.e., it diffuses into the surrounding water.

### 15.1.1 Kinetic theory and phenomenological laws

We introduced the notion of matter exchange between fluid elements when discussing the kinematics of tracers in VOLUME 1. In the present context, matter exchange occurs through the random motion of molecules acting in the presence of a matter concentration gradient. Even though the continuum approximation has removed all explicit concern for details of molecular motion, we confront the underlying molecular nature of matter since molecular motions have a measurable impact on macroscopic fluid properties. The transport of matter, heat, and momentum by random molecular motions is known as **molecular diffusion**. A statistical description of molecular diffusion is available for certain ideal-like gases using methods from kinetic theory largely developed in the 19th century. In that theory we are concerned with statistically describing the scattering of molecules off one another, and quantifying how such scattering affects a transport of matter, momentum, and heat.

A kinetic theory description of diffusive transport mathematically supports the phenomenological downgradient diffusive laws examined in Section 15.4, such as those for heat (**Fourier's law of conduction**), for matter (**Fick's law of diffusion**), and momentum (**Newton's law of viscous friction**), with these laws developed prior to kinetic theory. We extend the phenomenological laws to diffusion in liquids, though note that the associated kinetic theory is far less developed than for gases. We also note that for stratified fluids with multiple constituents, one encounters matter diffusion in the presence of a temperature gradient (**Soret effect**), and temperature diffusion in the presence of gradients of pressure and/or matter (**Dufour effect**).<sup>2</sup> However, geophysical fluids are nearly always dominated by turbulence, which renders a turbulent diffusion that swamps the effects from cross-diffusion. We thus ignore cross-diffusion in this chapter.

### 15.1.2 Comments and further study

Molecular diffusion as formulated by kinetic theory for ideal gases is examined in [Reif \(1965\)](#) and [Huang \(1987\)](#), as well as other statistical physics books. Molecular diffusion sets the ultimate dissipation scale and irreversibility of scalar variance. It is crucial microscopically as well as in boundary layers. Even so, it is too weak to explain large-scale mixing in the ocean and atmosphere without the aid of turbulent stirring to affect a downscale cascade of variance to the microscale.

---

<sup>2</sup>We study cross-diffusion in VOLUME 2 when examining energy and entropy flows in fluids.

## 15.2 Brownian motion

Brownian motion consists of the random walk motion of a macroscopically small but microscopically large particle (e.g., a pollen or dust particle) in a liquid. The motion of the Brownian particle arises from the incessant and random impacts on the particle by molecules from the surrounding liquid. A typical Brownian particle has size on the order of nanometers ( $10^{-9}$  m) to a few microns ( $10^{-6}$  m), which compares to the roughly  $10^{-10}$  m for the mean free path of molecular motion. Smaller particles exhibit more vigorous motion, whereas larger particles have less motion due to their increased inertia. Evidently, size, and corresponding mass, is key since particles must be microscopically huge, so to experience significant molecular impulses, yet macroscopically tiny, so to exhibit motion from the molecular impulses.

### 15.2.1 Einstein's formulation as a Markov process

Einstein (1905a) advanced a theory of Brownian motion, making a key assumption that the motion arises from a series of discrete time random impulses. He assumed that the impulses are separated by a time increment,  $\tau$ , that is tiny from the perspective of a macroscopic (human) observer but huge relative to the time scale of molecular motion. We return in Section 15.2.7 to provide an estimate for  $\tau$ , with those results supporting the assumption that it is tiny macroscopically.

In this formulation, Einstein further assumed the impulses on the Brownian particle at times  $\tau, 2\tau, 3\tau \dots$  are statistically uncorrelated. In today's language, we say that the impulses provide a white noise forcing, and that the resulting position of the particle manifests a [Markov processes](#). A Markov process has a future depending on the present but independent of the past. In space, Brownian motion is generally two dimensional, as per the motion of a neutrally buoyant dust particle in a petri dish, or three-dimensional, as per a smoke/dust particle floating in a quiet room. For simplicity, and following Einstein, we here consider only a single space dimension. We can interpret the one-dimensional position as the projection of a two or three dimensional motion along the real line.

Consider  $N$  identical Brownian particles that do not interact with one another. During a time interval,  $\tau$ , a particle moves a distance,  $\ell$ , which is continuous (i.e.,  $\ell$  is not discretized). Let  $p(\ell|\tau)$   $\delta\ell$  be the probability that a particle moves a distance between  $\ell$  and  $\ell + \delta\ell$  during one of the  $\tau$  intervals, thus measuring the probability of the particle transitioning to a new position. We write this transition probability density function as  $p(\ell|\tau)$ , with  $\tau$  a parameter that reminds us of the discrete time interval between the impulsive movements of the Brownian particle. The transition probability density is normalized over the spatial range of the motion,

$$\int_{-\infty}^{\infty} p(\ell|\tau) d\ell = 1, \quad (15.1)$$

and the random walk displacements are assumed to be unbiased, in which case the probability density is symmetric

$$p(\ell|\tau) = p(-\ell|\tau). \quad (15.2)$$

In a liquid with  $N$  Brownian particles, and ignoring any interactions between the particles, the number of particles that moves a distance within the range  $[\ell, \ell + \delta\ell]$  is written

$$\delta N = N p(\ell|\tau) \delta\ell. \quad (15.3)$$

We are interested in the case of many (effectively a continuum) of non-interacting Brownian particles that move through the fluid via the effects from molecular impulses. To connect to the notion of a tracer concentration,  $C$  (mass of tracer per mass of fluid element), we introduce the mass density,  $\varphi$ , of Brownian particles,

$$\varphi(x, t) = \rho(x, t) C(x, t), \quad (15.4)$$

with  $\rho$  the mass density of the fluid. Evidently,  $C$  is the mass concentration of Brownian particles, whereas  $\varphi$  is the mass density of Brownian particles. It follows that the total mass of Brownian particles is

$$M = \int_{-\infty}^{\infty} \rho C dx = \int_{-\infty}^{\infty} \varphi dx, \quad (15.5)$$

which we assume to be constant so that

$$\frac{dM}{dt} = \frac{d}{dt} \int_{-\infty}^{\infty} \varphi dx = \int_{-\infty}^{\infty} \frac{\partial \varphi}{\partial t} dx = 0. \quad (15.6)$$

### 15.2.2 The Chapman-Kolmogorov equation

Einstein introduced the following equation, now referred to as the one-step Chapman-Kolmogorov equation (also the discrete-time master equation), that determines the value of  $\varphi$  after a discrete time step

$$\varphi(x, t + \tau) = \int_{-\infty}^{\infty} p(\ell|\tau) \varphi(x - \ell, t) d\ell. \quad (15.7)$$

Reading this equation from right to left, we start with the particle density,  $\varphi(x - \ell, t)$ , at point  $x - \ell$  and time  $t$ . Convoluting this density with the probability,  $p(\ell|\tau) d\ell$ , to transition a distance  $\ell$  over a single time step,  $\tau$ , we arrive at the density,  $\varphi(x, t + \tau)$ .

We can verify that the Chapman-Kolmogorov equation (15.7) is consistent with mass conservation by integrating over the full domain

$$\int_{-\infty}^{\infty} \varphi(x, t + \tau) dx = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} p(\ell|\tau) \varphi(x - \ell, t) d\ell \right] dx. \quad (15.8)$$

The left hand side equals to the total mass,  $M$ , since the total mass remains constant in time as per equation (15.6). For the right hand side we swap the  $x$  and  $\ell$  integrals to find<sup>3</sup>

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} p(\ell|\tau) \varphi(x - \ell, t) d\ell \right] dx = \int_{-\infty}^{\infty} p(\ell|\tau) \left[ \int_{-\infty}^{\infty} \varphi(x - \ell, t) dx \right] d\ell. \quad (15.9)$$

The  $x$  integral yields  $M$ , which then leaves the  $\ell$  integral, which equals to unity due to the normalization condition (15.1).

### 15.2.3 Diffusion in the continuum limit

The Chapman-Kolmogorov equation (15.7) is an exact expression for the particle density at a new time step,  $t + \tau$ , given the density at time  $t$  and the probability for a displacement. For macroscopic purposes, we are most interested in the limit where the time step,  $\tau$ , between impulses is small macroscopically, in which case we take a Taylor series on the left hand side of

<sup>3</sup>We assume  $p(\ell|\tau)$  has a finite second moment and the particle density,  $\varphi(x, t)$ , decays fast enough at  $x = \pm\infty$  that allows us to perform a Taylor expansion, swap integrals, and drop boundary terms.

the Chapman-Kolmogorov equation (15.7) to find

$$\varphi(x, t + \tau) \approx \varphi(x, t) + \tau \partial_t \varphi(x, t). \quad (15.10)$$

Likewise, if the probability for a displacement is assumed to be concentrated around zero, meaning that the individual displacements,  $\ell$ , are macroscopically small, then we can expand the concentration inside the integral on the right hand side of the Chapman-Kolmogorov equation (15.7) to find

$$\int_{-\infty}^{\infty} p(\ell|\tau) \varphi(x - \ell, t) d\ell \approx \int_{-\infty}^{\infty} p(\ell|\tau) [\varphi(x, t) - \ell \partial_x \varphi(x, t) + \frac{1}{2} \ell^2 \partial_{xx} \varphi(x, t)] d\ell. \quad (15.11)$$

The integral is over the displacement,  $\ell$ , so that the density,  $\varphi(x, t)$  and its derivatives can be pulled outside of the integral. Making use of the normalization condition (15.1) and symmetry condition (15.2) leads to

$$\int_{-\infty}^{\infty} p(\ell|\tau) [\varphi(x, t) - \ell \partial_x \varphi(x, t) + \frac{1}{2} \ell^2 \partial_{xx} \varphi(x, t)] d\ell = \varphi(x, t) + \frac{1}{2} \partial_{xx} \varphi(x, t) \int_{-\infty}^{\infty} \ell^2 p(\ell) d\ell. \quad (15.12)$$

Bringing both sides together leads us to conclude that, in the macroscopic limit the particle density satisfies the diffusion equation

$$\partial_t \varphi = \kappa \partial_{xx} \varphi, \quad (15.13)$$

where we introduced the kinematic diffusivity

$$\kappa = \langle \ell^2 \rangle / (2\tau) \quad \text{with} \quad \langle \ell^2 \rangle = \int_{-\infty}^{\infty} \ell^2 p(\ell|\tau) d\ell. \quad (15.14)$$

The diffusivity is seen to be directly proportional to the mean-square particle displacement,  $\langle \ell^2 \rangle$ , and inversely proportional to the time between impulses,  $\tau$ .

#### 15.2.4 Moments of the particle density

Introduce the Brownian particle density,  $\varphi$ , and define position moments with respect to the particle density

$$\overline{x^n} = M^{-1} \int_{-\infty}^{\infty} x^n \varphi(x, t) dx. \quad (15.15)$$

Moments based on the Brownian particle density are equivalent to moments computed from an ensemble of Brownian particles as considered in Section 15.2.5 as part of Langevin's formulation of Brownian motion. However, the Brownian moments (15.15) are distinct from moments defined with respect to the probability density,  $p(\ell)$ . So although the zeroth moment of both  $p(\ell)$  and  $\varphi(x, t)$  are unity

$$1 = \overline{x^0} = M^{-1} \int_{-\infty}^{\infty} \varphi(x, t) dx \quad \text{and} \quad 1 = \int_{-\infty}^{\infty} p(\ell|\tau) d\ell, \quad (15.16)$$

their higher moments are distinct, as seen in the following.

**First moment of the particle density is static**

The first moment of the transition probability vanishes since  $p(\ell) = p(-\ell)$

$$\langle \ell \rangle = \int_{-\infty}^{\infty} \ell p(\ell|\tau) d\ell = 0. \quad (15.17)$$

In contrast, the first moment of the particle density, which we refer to as the particle center of mass, is generally nonzero,

$$\bar{x} = M^{-1} \int_{-\infty}^{\infty} x \varphi(x, t) dx. \quad (15.18)$$

However, we can show that this position is static for Brownian motion by multiplying the Chapman-Kolmogorov equation (15.7) by  $x$  and then integrating

$$\int_{-\infty}^{\infty} x \varphi(x, t + \tau) dx = \int_{-\infty}^{\infty} x \left[ \int_{-\infty}^{\infty} p(\ell|\tau) \varphi(x - \ell, t) d\ell \right] dx. \quad (15.19)$$

For the right hand side we change variables to  $y = x - \ell$  to find

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} p(\ell|\tau) (y + \ell) \varphi(y, t) d\ell \right] dy = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} p(\ell|\tau) y \varphi(y, t) d\ell \right] dy, \quad (15.20)$$

where we made use  $\langle \ell \rangle = 0$ . Making use of the normalization of  $p(\ell)$  (so that  $\langle \ell^0 \rangle = 1$ ), and then renaming  $y$  to  $x$  leads to

$$\int_{-\infty}^{\infty} x \varphi(x, t + \tau) dx = \int_{-\infty}^{\infty} x \varphi(x, t) dx, \quad (15.21)$$

which means that  $\bar{x}$  is independent of time

$$\frac{d\bar{x}}{dt} = 0. \quad (15.22)$$

We thus find that the Brownian process leaves the center of mass of the particle distribution unchanged. This result follows since there is an equal probability for a Brownian particle to be displaced in either direction, since  $p(\ell|\tau) = p(-\ell|\tau)$ . We can readily verify that for the diffusion equation description, we also find  $\bar{x}$  is static. To do so, multiply the diffusion equation (15.13) by  $x$  and integrate over the real line. Integration by parts, and assuming the particle density falls off sufficiently fast at  $\pm\infty$ , then leads to a time independent  $\bar{x}$ .

**Second moment of the particle density evolves with unit power of time**

Next consider the mean squared spread of the particle density

$$\overline{x^2} = M^{-1} \int_{-\infty}^{\infty} x^2 \varphi(x, t) dx. \quad (15.23)$$

We can determine its time evolution by multiplying the Chapman-Kolmogorov equation (15.7) by  $x^2$  and integrating over space

$$\int_{-\infty}^{\infty} x^2 \varphi(x, t + \tau) dx = \int_{-\infty}^{\infty} x^2 \left[ \int_{-\infty}^{\infty} p(\ell|\tau) \varphi(x - \ell, t) d\ell \right] dx. \quad (15.24)$$

Changing variables to  $y = x - \ell$  and making use of normalization and symmetry of  $p(\ell)$  leads to

$$\int_{-\infty}^{\infty} x^2 \varphi(x, t + \tau) dx = M \langle \ell^2 \rangle + \int_{-\infty}^{\infty} y^2 \varphi(y, t) dy, \quad (15.25)$$

which then renders

$$\int_{-\infty}^{\infty} x^2 [\varphi(x, t + \tau) - \varphi(x, t)] dx = M \langle \ell^2 \rangle = 2 \tau \kappa M, \quad (15.26)$$

where the final equality introduced the Brownian diffusivity (15.14). Taking the continuous time limit with small  $\tau$  leads to

$$\frac{d\bar{x}^2}{dt} = 2 \kappa = \langle \ell^2 \rangle / \tau. \quad (15.27)$$

Evidently, the mean-squared spread of a cloud of Brownian particles grows as the first power of time. It is notable that there are processes that display fractional time dependencies, with such processes referred to as **anomalous diffusion**.

### 15.2.5 Langevin equation

[Langevin \(1908\)<sup>4</sup>](#) provided an alternative formulation of Brownian motion, with his approach initiating the study of stochastic differential equations. In his formulation, Langevin assumed the Brownian particle experiences two forces. The first force arises from the Stokes' drag acting on a spherical body moving through a viscous fluid and with velocity,  $V$

$$F_{\text{drag}} = -\alpha V \quad \text{with} \quad \alpha = 6 \pi \mu R. \quad (15.28)$$

In this equation,  $R$  is the radius of the Brownian particle, and  $\mu$  is the molecular dynamic viscosity of the liquid (dimensions of  $M L^{-1} T$ ). The drag parameter,  $\alpha$ , has dimensions of  $M T^{-1}$ . The second force,  $\mathcal{N}$ , is a white noise force that arises from the random and uncorrelated impulses imparted by the liquid molecules. The resulting **Langevin equation** for the position of the Brownian particle is given by

$$m \ddot{X} = -\alpha \dot{X} + \mathcal{N}. \quad (15.29)$$

Note that since we are here focused on particle trajectories, we use the capital,  $X$ , to denote the position in space of a Brownian particle, and write

$$V = \dot{X} \quad (15.30)$$

for its velocity. It is notable that Langevin's equation is identical to the response function equation studied in VOLUME 1, only here we have introduced a random forcing via the noise term,  $\mathcal{N}$ .

---

<sup>4</sup>For an English translation, see [Lemons and Gythiel \(1997\)](#).

### 15.2.6 Dissipation time scale

We derive the evolution equation for kinetic energy of the Brownian particle by multiplying the Langevin equation (15.29) by  $V = \dot{X}$

$$m \frac{dV^2}{dt} = -\alpha V^2 + V \mathcal{N}. \quad (15.31)$$

As expected, the viscous term acts to dissipate the ensemble averaged kinetic energy, whereas the noise acts to increase or decrease the kinetic energy. We now estimate the dissipation time scale,

$$\tau_d = m/\alpha, \quad (15.32)$$

which we assume corresponds to the time interval,  $\tau$ , between the discrete impulses assumed by Einstein in Section 15.2.1. To do so requires details of the Brownian particle, which we assume to be spherical and with density,  $\rho_p$ , so that

$$\tau_d = \frac{\frac{4}{3}\pi R^3 \rho_p}{6\pi\mu R} = \frac{2\rho_p R^2}{9\mu} = \frac{2\rho_p R^2}{9\rho\nu}, \quad (15.33)$$

where the final equality introduced the kinematic viscosity,  $\nu$ , through

$$\mu = \rho\nu. \quad (15.34)$$

For the Brownian particle density and size, assume a pollen particle and make use of its hydrated density (i.e., density of a saturated particle)

$$\rho_p = 1.1\rho_{\text{water}} \quad \text{and} \quad L_{\text{brownian}} = 10^{-5} \text{ m}. \quad (15.35)$$

Note that the approximate size of a water molecule is

$$L_{\text{water}} = 10^{-10} \text{ m}, \quad (15.36)$$

so that the pollen particle is indeed far larger than water molecules, even though it is tiny macroscopically. The kinematic molecular viscosity of water at room temperature is around  $10^{-6} \text{ m}^2 \text{ s}^{-1}$ , and its density is  $\rho = 10^3 \text{ kg m}^{-3}$ . Taking the radius of the pollen equal to  $L_{\text{brownian}}$ , we have

$$\tau_d = \frac{2\rho_p L_{\text{brownian}}^2}{9\rho_{\text{water}}\nu} \approx 2.4 \times 10^{-5} \text{ s}. \quad (15.37)$$

This is a tiny time interval from a macroscopic perspective, which supports the use of Einstein's continuum limit of the Chapman-Kolmogorov equation in Section 15.2.3.

### 15.2.7 Ensemble statistics from the Langevin equation

The trajectory,  $X(t)$ , for a single Brownian particle is stochastic (i.e., random), with such randomness induced by the white noise forcing,  $\mathcal{N}$ , appearing in the Langevin equation (15.29). We are thus interested in developing the statistics arising from an ensemble of many Brownian particles. As noted in Section 15.2.4, statistics based on an ensemble of Brownian particles is the same, in the continuum limit, as when examining statistics based on the particle density function,  $\varphi(x, t)$ . Hence, we use the overline here for the ensemble statistics developed from the Langevin equation.

### Evolution of the ensemble mean kinetic energy

The evolution equation for the ensemble mean kinetic energy of the Brownian particle is found by taking the ensemble average of the kinetic energy equation (15.31)

$$m \frac{d\overline{V^2}}{dt} = -\alpha \overline{V^2} + \overline{V \mathcal{N}}. \quad (15.38)$$

Note that we assumed the ensemble average commutes with the time derivative. As expected, the viscous term acts to dissipate the ensemble averaged kinetic energy, whereas correlations between the velocity and the noise acts to increase or decrease the kinetic energy.

### Evolution of ensemble mean squared spread

Rather than kinetic energy, consider now an expression for the ensemble mean squared spread,  $\overline{X^2}$ , whose evolution equation is derived by multiplying the Langevin equation (15.29) by  $X$  and rearranging

$$m \frac{d}{dt} \left[ \frac{dX^2}{dt} + \alpha X^2 \right] = 2m \dot{X}^2 + 2X \mathcal{N}. \quad (15.39)$$

We now take an ensemble average and make two key assumptions. First, assume that the Brownian particle is in thermal equilibrium with the surrounding liquid. This assumption allows us to use the equipartition theorem from classical statistical mechanics to equate the ensemble average kinetic energy to the thermal energy<sup>5</sup>

$$\frac{1}{2} \overline{m \dot{X}^2} = \frac{1}{2} k_B T, \quad (15.40)$$

where  $T$  is the absolute temperature (in Kelvin) of the liquid, and  $k_B$  is Boltzmann's constant. The second assumption sets to zero the correlation between the particle position and the noise forcing

$$\overline{X \mathcal{N}} = 0. \quad (15.41)$$

This assumption means that the noise and Brownian particle positions are statistically independent and so do not somehow conspire to produce a nonzero correlation.

The above two assumptions in equation (15.39) yield the evolution equation for the ensemble mean squared distance

$$m \frac{d}{dt} \left[ \frac{d\overline{X^2}}{dt} + \alpha \overline{X^2} \right] = 2k_B T. \quad (15.42)$$

We solve this differential equation by first deriving the homogeneous solution (with zero on the right hand side), which is given by

$$e^{-t/\tau_d} \overline{X^2(0)}, \quad (15.43)$$

where  $\overline{X^2(0)}$  is the mean squared spread at the initial time. Adding in the particular solution leads to

$$\overline{X^2}(t) = e^{-\alpha t/\tau_d} \overline{X^2(0)} + (2k_B T / \alpha) t. \quad (15.44)$$

Since  $\tau_d$  is quite tiny macroscopically, we can ignore the decaying exponential term in

---

<sup>5</sup>The equipartition theorem (15.40) was also made by [Einstein \(1905a\)](#) in his formulation, which connects the diffusivity of the Brownian particle motion,  $\kappa$ , to the fluid viscosity. We see this connection in the following approach as well.

equation (15.44), so that the mean squared spread evolves approximately as

$$\overline{X^2} \approx (2 k_B T / \alpha) t \implies \frac{d\overline{X^2}}{dt} \approx 2 k_B T / \alpha. \quad (15.45)$$

Recalling our approach following Einstein earlier in this section, in particular equation (15.27), we identify the diffusivity acting to spread the Brownian particle distribution

$$\kappa = \frac{2 k_B T}{\alpha} = \frac{k_B T}{3\pi\rho\nu R} = \frac{R^g T}{3\pi\rho\nu R A^v}, \quad (15.46)$$

where the final equality introduced the universal gas constant,  $R^g$ , and Avogadro's number,  $A^v$ ,

$$R^g = 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1} \quad \text{and} \quad A^v = 6.0222 \times 10^{23} \text{ mole}^{-1}. \quad (15.47)$$

This same result was also derived by [Einstein \(1905a\)](#), and it was used by [Perrin \(1909\)](#) to estimate Avogadro's number using macroscopically measurable properties.

### 15.2.8 Further study

A summary of Brownian motion as formulated by [Einstein \(1905a\)](#) and [Langevin \(1908\)](#) can be found in Section 1.2 of [Gardiner \(1985\)](#). Section 3 of [Young \(1999\)](#) also discusses Einstein's approach, and then introduces anomalous diffusion, such as occurs with interrupted random walk processes. The concepts and methods promoted by [Einstein \(1905a\)](#) and [Langevin \(1908\)](#) formed the seeds for the study of stochastic processes during the 20th century, with [Gardiner \(1985\)](#) providing a lucid presentation. We return to the evolution of the tracer moments in Exercise 16.3 as part of our study of the advection-diffusion equation.

## 15.3 Turbulent diffusion

Turbulent flow is the norm for geophysical fluids, with turbulence greatly enhancing the efficiency of property mixing relative to the case without flow.<sup>6</sup> Namely, the stirring of fluid properties by turbulent flows acts to stretch and fold contours/surfaces of constant properties, thus increasing property gradients and so increasing the action of molecular diffusion. We refer to this turbulent enhanced mixing as [turbulent diffusion](#).

Turbulent diffusion is not concerned with molecular properties of the fluid. Rather, the properties of turbulent diffusion (e.g., its efficiency) depend on the nature of the turbulent motion of fluid elements. Hence, each type of turbulent motion gives rise to a distinct form of turbulent diffusion. For example, turbulent diffusion associated with a turbulent field of breaking internal gravity waves is distinct from turbulent diffusion from geostrophic eddies. In this way, turbulent diffusion sits within the realm of continuum mechanics, whereas molecular diffusion is a subject for kinetic theory. Furthermore, turbulent diffusion concerns fluid motions that are continuous in time, which contrasts to the discrete time impulses assumed by Einstein in his formulation of Brownian motion in Section 15.2.

### 15.3.1 Introducing Taylor's Lagrangian description

[Taylor \(1922\)](#) described the statistical properties of turbulent flows by focusing on an ensemble

---

<sup>6</sup>Turbulent flow is characterized by a quasi-random fluid motion that acts to transport fluid elements and their properties.

of fluid particles and thus deriving statistics for the ensemble. The mathematical apparatus is very similar to that encountered in the trajectory analysis of Brownian particles in Section 15.2, with the analogy allowing us to infer that turbulent transport has a diffusive quality at its core. Indeed, the many stochastic theories of turbulent diffusion make use of the Langevin equation originally formulated for Brownian motion.<sup>7</sup>

Following our study of fluid kinematics in VOLUME 1, recall that the Cartesian position and velocity of a typical fluid particle is written

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{V}(t) \implies \mathbf{X}(t) = \mathbf{X}(t_0) + \int_{t_0}^t \mathbf{V}(t') dt', \quad (15.48)$$

where  $t_0$  is some arbitrary initial time. Now take the ensemble mean and, just as for the Brownian motion analysis, assume the time derivative commutes with the ensemble mean to yield

$$\overline{\mathbf{X}(t)} = \overline{\mathbf{X}(0)} + \int_{t_0}^t \overline{\mathbf{V}(t')} dt'. \quad (15.49)$$

Assuming the velocity field has stationary and unbiased statistics, and there is no drift velocity, then we are led to conclude that the ensemble mean position remains unchanged from its initial position

$$\overline{\mathbf{X}(t)} = \overline{\mathbf{X}(0)}. \quad (15.50)$$

For simplicity, and without loss of generality, we set

$$\overline{\mathbf{X}(0)} = 0. \quad (15.51)$$

Now multiply the trajectory equation (15.48) by  $\mathbf{X}(t)$  to render an equation for the evolution of the squared position

$$\frac{d\mathbf{X}(t)^2}{dt} = \mathbf{V}(t) \cdot \mathbf{X}(t) = \mathbf{V}(t) \cdot \int_{t_0}^t \mathbf{V}(t') dt' = \int_{t_0}^t \mathbf{V}(t) \cdot \mathbf{V}(t') dt', \quad (15.52)$$

where the second equality follows from the trajectory equation (15.48). Taking the ensemble mean yields

$$\frac{1}{2} \frac{d\overline{\mathbf{X}(t)^2}}{dt} = \int_{t_0}^t \overline{\mathbf{V}(t) \cdot \mathbf{V}(t')} dt' \equiv \int_{t_0}^t C(t, t') dt', \quad (15.53)$$

where we introduced the correlation function between the velocity of fluid particles<sup>8</sup>

$$C(t, t') = \overline{\mathbf{V}(t) \cdot \mathbf{V}(t')}. \quad (15.54)$$

Equation (15.53) describes the ensemble mean dispersion of a fluid particle trajectories relative to a center of mass position which, as noted above, we assume to vanish,  $\overline{\mathbf{X}(0)} = 0$ . More specifically, it says that the time derivative of the ensemble mean of the squared fluid particle trajectory equals to the time integral of the time correlation between the fluid particle's velocity.

Assuming the correlation (15.54) is localized in time, so that it has a finite limit as  $t \rightarrow \infty$ ,

<sup>7</sup>See [McComb \(1990\)](#) for an in-depth study of stochastic theories of turbulence.

<sup>8</sup>The correlation,  $C(t, t')$ , in equation (15.54) is a one-point correlation, since it is computed between the velocity of a single fluid particle and so it is computed at a single point in space. Multiple point correlations are also of interest in fluid turbulence theories, with Section 13.3 of [Vallis \(2017\)](#) providing an introduction to two-point or two-particle correlations.

yields

$$\frac{d\overline{\mathbf{X}(t)^2}}{dt} = 2 \int_{t_0}^{\infty} C(t, t') dt' \equiv 2 D, \quad (15.55)$$

where we introduced the Lagrangian diffusivity,

$$D \equiv \int_{t_0}^{\infty} C(t, t') dt'. \quad (15.56)$$

We thus find that the mean-squared spread of the ensemble trajectories grows as

$$\overline{\mathbf{X}(t)^2} = 2 D t, \quad (15.57)$$

which is identical to the diffusive behavior of a Brownian particle found in Section 15.2. It is this connection that prompts us to conclude that the spread of fluid particle trajectories by turbulent flows is directly akin to a Brownian/diffusive process.

### 15.3.2 Comments and further study

The conceptual and technical advances from [Taylor \(1922\)](#) remain part of the null hypothesis for how turbulent flow affects mixing in geophysical fluids. However, there is a nontrivial level of work required to make use of these insights for practical parameterizations, with that work remaining at the forefront of research in ocean and atmospheric physics. We do not pursue this analysis in this book, deferring to the various chapters in [McComb \(1990\)](#) and [Vallis \(2017\)](#) (for example) for in-depth analyses of turbulent diffusion. Even so, we do pursue a phenomenological perspective on parameterized turbulent mixing in the remainder of this chapter as well as in Chapter 18.

An introduction to turbulent diffusion can be found in [Young \(1999\)](#). More detailed treatments are given in [Csanady \(1973\)](#), who focuses on turbulent diffusion in the environment (e.g., for the study of pollution dispersal), and in Chapter 13 of [Vallis \(2017\)](#), who focuses on geophysical flows. Further elements can be found in Section 1.5 of [Kundu et al. \(2016\)](#).

## 15.4 Phenomenological diffusion laws

In this section we summarize a suite of phenomenological laws that result in a diffusion equation. These laws were formulated prior to the ideas from kinetic theory, Brownian motion, and turbulent diffusion. They remain a very useful null hypothesis for turbulent diffusive transport in geophysical fluids, with the diffusivities set according to their turbulent eddy values rather than molecular values.

### 15.4.1 Fick's law for matter diffusion

Consider a fluid with a non-uniform tracer concentration such as that drawn for a one-dimensional case in Figure 15.1. Random motion, due either to molecular motion or turbulent fluctuations, transfers tracer across an arbitrary point, line, or plane. Random motion preferentially moves tracer from regions of high concentration to regions of low concentration, thus smoothing gradients.

To a good approximation, the mass flux (mass per time per cross-sectional area) of a material tracer is linearly proportional to the concentration gradient, and thus can be written

in the form

$$\mathbf{J} = -\kappa_c \rho \nabla C. \quad (15.58)$$

In this equation, we introduced the positive proportionality factor,  $\kappa_c > 0$ , known as the kinematic diffusivity, whereas the product  $\kappa_c \rho$  is known as the dynamic diffusivity:

$$\kappa_c \quad \text{kinematic diffusivity with SI units m s}^{-2} \quad (15.59)$$

$$\rho \kappa_c \quad \text{dynamic diffusivity with SI units kg m}^{-2} \text{ s}^{-2}. \quad (15.60)$$

The kinematic diffusivity has dimensions of squared length per time and it sets the efficiency of diffusion when acting within a fluid with a tracer gradients. The diffusive flux (15.58) is known as Fick's law of diffusion, and it is commonly used in geophysical fluid mechanics to represent the mixing of matter through diffusion. The minus sign in the diffusive flux arises since the flux is directed down the concentration gradient. When considering molecular diffusion, we distinguish diffusivities according to their respective tracers since they generally differ. In contrast, turbulent diffusivities are commonly assumed to be independent of tracer since they are determined by the flow, in which case we write the generic,  $\kappa$ .

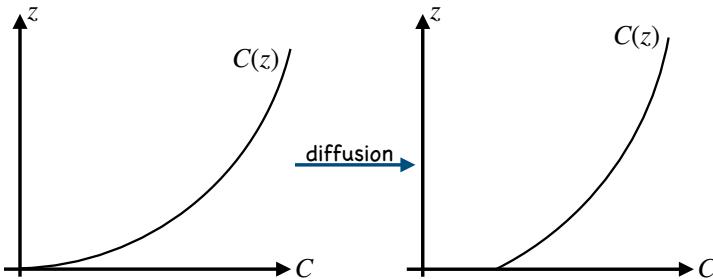


FIGURE 15.1: A graph illustrating a material tracer concentration,  $C$  as a function of the space coordinate  $z$ , with the left panel showing the concentration at an earlier time than the right panel. Across any arbitrary point, matter is transported through random motions, with this transport generally reducing the magnitude of the concentration gradient. This reduction of concentration gradient is a direct result of the downgradient (i.e., down the concentration gradient) orientation of the diffusive transport. For example, where the concentration is relatively high, random motion mixes this high concentration with adjacent lower concentration, thus acting to lower the concentration in the originally high concentration region and raise the concentration in the originally low concentration region. In this particular example,  $\partial C / \partial z > 0$ , so that random fluid motions (either molecular or turbulent) lead to a diffusive flux directed in the  $-\hat{z}$  direction; i.e., downward. This downward flux brings high concentration fluid into the lower regions and low concentration fluid into higher regions. The concentration is vertically uniform if allowed to equilibrate under the action of diffusion.

The kinematic diffusivity has physical dimensions equal to the product of a length and a speed. For molecular diffusion, the kinematic diffusivity is proportional to the mean free path,  $L_{\text{mfp}}$  and the root-mean-square molecular speed,  $v_{\text{rms}}$ , both of which are described in the chapter in VOLUME 1 describing kinetic theory and the continuum approximation. Both  $L_{\text{mfp}}$  and  $v_{\text{rms}}$  are functions of the molecules comprising the matter. For air, the mean free path is roughly  $2 \times 10^{-7}$  m and the root-mean-square speed is 500 m s<sup>-1</sup>, so that  $L_{\text{mfp}} v_{\text{rms}} \approx 10^{-4}$  m<sup>2</sup> s<sup>-1</sup>. The precise value for the molecular diffusivity depends on the molecular properties of the matter diffusing through the fluid; e.g., molecular size and speed.

For turbulent diffusion, Prandtl suggested that the characteristic length and velocity scales are determined by properties of the turbulent flow, not by the molecular properties of the fluid or the tracer.<sup>9</sup> The turbulent length scale (also called the mixing length) is generally much

<sup>9</sup> Prandtl (1925) is the original paper, with an English translation given by Prandtl (1949) and further historical context given by Bradshaw (1974).

larger than the molecular mean free path, whereas the [turbulent velocity scale](#) is much smaller than molecular speeds. Determination of turbulent length and velocity scales is subject to large uncertainties given the multiple regimes of turbulence exhibited by geophysical flows. As a result, tracer transport by turbulent flows remains an active topic of research.

In regions where the diffusive flux converges, there is a net transport of matter that leads to the reduction of the tracer concentration gradient as determined by the convergence of the diffusive flux

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J} = \nabla \cdot (\kappa_c \rho \nabla C). \quad (15.61)$$

That is, the concentration increases in regions where the diffusive flux,  $\mathbf{J}$ , converges, and decreases where the flux diverges. Expanding the divergence operator leads to

$$\rho \frac{DC}{Dt} = \nabla(\kappa_c \rho) \cdot \nabla C + \kappa_c \rho \nabla^2 C. \quad (15.62)$$

The first term is nonzero in regions where the [dynamic diffusivity](#),  $\kappa_c \rho$ , spatially varies. This term vanishes for molecular diffusion, in which case the diffusivity is a spatial constant. However, for turbulent diffusion this term can be quite important given the potential for strong flow dependence to the diffusivity. Indeed, there are cases in which this spatial dependence can enhance tracer gradients, overcoming the effects from the curvature term. We consider an example in Exercise 15.6 known as the [Phillip's layering instability](#).

The second term in the diffusion equation (15.62) is proportional to the Laplacian of the tracer concentration, which provides a measure of the curvature in the tracer field. Hence, this term vanishes when the tracer concentration is a constant or a linear function of space, whereas it is nonzero for tracers having less trivial spatial structure. As discussed in Section 15.5, this term provides a scale selectivity to the diffusion operator, thus resulting in a preferential dampening of small scale features relative to large scale features.

## 15.4.2 Fourier's law of conduction

In the same way that matter concentration gradients lead to diffusion by random motions, temperature gradients lead to diffusion of heat. The corresponding phenomenological relation is known as [Fourier's law of conduction](#), with the diffusive (or conductive) flux given by

$$\mathbf{J} = -\kappa_T \rho \nabla T, \quad (15.63)$$

where  $\kappa_T > 0$  is the kinematic thermal diffusivity. As for the matter diffusivity, the molecular thermal diffusivity can be expressed in terms of fundamental properties of the fluid, and it is different from the diffusivity for matter concentration diffusion. In general, molecular processes diffuse matter concentrations slower than temperature, so that the molecular matter diffusivity is smaller than the molecular thermal diffusivity. The reason for the difference is that the diffusion of material concentration requires the movement of matter (molecules), whereas thermal diffusion occurs through the exchange of thermal energy between molecules, and that exchange does not require the motion of matter. For turbulent transport, however, the turbulent thermal diffusivity is roughly the same as the turbulent matter diffusivity. The reason is that the turbulent diffusion of both matter and heat are mediated by the same turbulent fluctuations of fluid elements.

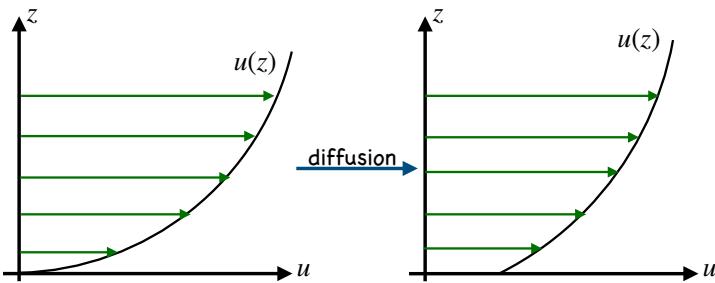


FIGURE 15.2: A graph illustrating the zonal velocity,  $u$ , as a function of the space coordinate  $z$ , with the left panel showing the velocity at an earlier time than the right panel. Across any arbitrary point, transport of momentum through random motions generally reduces the magnitude of the velocity gradient; i.e., the diffusive transport leads to a viscous stress that acts to reduce the velocity shear.

### 15.4.3 Newtonian frictional stress and momentum diffusion

In the same way that gradients in matter concentration and temperature lead to diffusion by random molecular and turbulent motions, the momentum of fluid elements is exchanged through diffusion in the presence of viscosity. The corresponding phenomenological relation is known as [Newton's law of viscous friction](#). As momentum is a vector, a general treatment of momentum transport through irreversible viscous processes involves a second order stress tensor and a fourth order viscosity tensor. For the specific case shown in Figure 15.2, with shear (i.e., nonzero velocity gradient) in a single direction, Newtonian frictional stress takes the form

$$\tau = \rho \mu \partial_z u, \quad (15.64)$$

where  $\mu > 0$  is the kinematic viscosity. Note the absence of a minus sign, in contrast to diffusive fluxes of scalars. The sign difference arises since it is the divergence of the [stress tensor](#) that leads to contact forces on the fluid, whereas it is the convergence of [diffusive fluxes](#) that leads to diffusion of matter and heat. We consider these general properties of the stress tensor in [VOLUME 2](#).

For geophysical fluid mechanics, we are most generally interested in the molecular viscosity of water and air. Quite generally, the dynamic viscosity of water ( $\rho \mu$ ) is about  $10^2$  times larger than that for air. But since the density of water is about  $10^3$  times larger than air, the kinematic viscosity of air is roughly 10 times greater than that of water.

The molecular kinematic viscosity can be expressed in terms of fundamental properties of the fluid, and it is different from the molecular matter diffusivity and molecular thermal diffusivity. For some turbulent processes, the turbulent viscosity,  $\mu$ , is proportional to the turbulent diffusivity,  $\kappa$ , of scalar fields (e.g., temperature, salinity, humidity). In general, the non-dimensional ratio of the viscosity to the diffusivity is known as the [Prandtl number](#)

$$\text{Pr} = \mu/\kappa. \quad (15.65)$$

Theories for the turbulent Prandtl number are largely empirical, with first principles arguments elusive.

### 15.4.4 Comments

As noted on page 4 of [Csanady \(1973\)](#), results from both kinetic theory of gases and Brownian motion suggest that the distance over which a typical mixing “event” occurs is small relative to the macroscopic scales of motion of concern for continuum mechanics. Consequently, we

are justified in using the diffusive flux expression (15.58) arising from Fick's law of diffusion as considered in Section 15.4.1.

## 15.5 The scale selectivity of Laplacian diffusion

Let us focus on the Laplacian term appearing in the tracer equation (15.62) to establish some properties characteristic of Laplacian diffusion. Start by considering a tracer concentration whose spatial structure is given by two Fourier modes,

$$C(\mathbf{x}) = C_{\mathbf{p}} \sin(\mathbf{p} \cdot \mathbf{x}) + C_{\mathbf{q}} \sin(\mathbf{q} \cdot \mathbf{x}), \quad (15.66)$$

where  $\mathbf{p}$  and  $\mathbf{q}$  are specified wavevectors and  $C_{\mathbf{p}}, C_{\mathbf{q}}$  are their corresponding amplitudes. In this case the Laplacian of the tracer is given by

$$\nabla^2 C = -[|\mathbf{p}|^2 C_{\mathbf{p}} \sin(\mathbf{p} \cdot \mathbf{x}) + |\mathbf{q}|^2 C_{\mathbf{q}} \sin(\mathbf{q} \cdot \mathbf{x})]. \quad (15.67)$$

Consequently, the Laplacian diffusion operator acts preferentially on waves of smaller wavelength (and larger wavenumber). For example, assume  $|\mathbf{p}| \ll |\mathbf{q}|$ , in which case the  $\mathbf{q}$ -mode is more rapidly damped towards zero than the  $\mathbf{p}$ -mode.<sup>10</sup> For this reason we say that Laplacian diffusion is **scale selective**. Note that zero is the wave averaged concentration for each Fourier mode. We thus see that diffusion acts to dampen each mode towards its average. Scale selectivity results geometrically from a property of the Laplacian operator as a measure of curvature. Tracer features with large curvature have a larger magnitude for their Laplacian, and as such they are damped more rapidly than tracer features with relatively small curvature.

As a second means to understand properties of Laplacian diffusion, consider a Taylor series for the tracer concentration computed relative to an arbitrarily chosen origin,

$$C(\mathbf{x}) = C(0) + x^m \partial_m C|_{\mathbf{x}=0} + (1/2) x^m x^n \partial_n \partial_m C|_{\mathbf{x}=0} + \dots \quad (15.68)$$

Now compute the average of this tracer concentration over a cube centered at the origin with sides  $L$  and volume  $L^3$ , and furthermore make use of the identities

$$\int_{-L/2}^{L/2} \int_{-L/2}^{L/2} \int_{-L/2}^{L/2} x^m dx dy dz = 0 \quad \forall m \quad (15.69a)$$

$$\int_{-L/2}^{L/2} \int_{-L/2}^{L/2} \int_{-L/2}^{L/2} x^m x^n dx dy dz = \delta^{mn} (L^5 / 12). \quad (15.69b)$$

We thus find that the volume averaged tracer concentration,  $\langle C \rangle$ , deviates from the concentration at the origin by a term proportional to the Laplacian of the concentration evaluated at the origin

$$\langle C \rangle - C(0) = (L^2 / 24) \nabla^2 C|_{\mathbf{x}=0} \implies \partial_t C|_{\mathbf{x}=0} = -(24 \kappa_c \rho / L^2) [C(0) - \langle C \rangle], \quad (15.70)$$

where we made use of the Laplacian portion of the diffusion equation (15.62) for the second expression. We can perform this calculation for any point taken as the origin. So this result says that Laplacian diffusion generally provides a tendency to damp the tracer concentration at a point towards the average tracer concentration in the region surrounding that point. For

---

<sup>10</sup>In VOLUME 4, we study this process of scale selectivity in the context of Fourier analysis.

example, consider the case where the averaged tracer concentration has no time dependence, as occurs in a region with zero boundary fluxes of tracer. If we are at a point in the region where the concentration is less than the average concentration,  $C < \langle C \rangle$ , then diffusion provides a positive tendency to increase  $C$  towards  $\langle C \rangle$ , and vice versa if  $C > \langle C \rangle$ . These results offer another expression of what we found in studying Laplacian diffusion on Fourier modes. In that case, the Laplacian operator, as revealed through equation (15.67), damps each mode towards its average, which is zero.

## 15.6 Gaussian concentration generated by a Dirac source

Consider a one-dimensional tracer concentration in an unbounded domain whose initial ( $t = 0$ ) value vanishes everywhere except at the origin, where it is given by a [Dirac delta](#)

$$C(x, t = 0) = Q \delta(x), \quad (15.71)$$

where  $\delta(x)$  is the Dirac delta. The Dirac delta has dimensions of inverse length, so that the constant,  $Q$ , has dimensions of  $[C] L$ . Integrating over any region containing the origin reveals that  $Q$  is the domain integrated tracer concentration at the initial time,

$$\int_{-\infty}^{\infty} C(x, t = 0) dx = Q. \quad (15.72)$$

We are ensured that this integral holds for all time if the domain has no boundary fluxes of tracer nor any interior tracer sources.

Assume now that the tracer concentration evolves according to the one-dimensional (one space dimension) diffusion equation with a constant diffusivity,  $\kappa > 0$ , and in a fluid with a constant density. In the absence of spatial boundaries (i.e., diffusion occurs on the real line,  $\mathbb{R}^1$ ), the concentration is proportional to the causal free space Green's function given by the Gaussian

$$C(x, t) = \frac{Q}{(4 \pi \kappa t)^{1/2}} e^{-x^2/(4 \kappa t)}, \quad (15.73)$$

which indeed satisfies (for any time,  $t$ )

$$Q = \int_{-\infty}^{\infty} C(x, t) dx, \quad (15.74)$$

as required by tracer conservation (15.72) for the infinite domain. We illustrate the Gaussian tracer concentration (15.73) in Figure 15.3. The variance of the tracer distribution is given by

$$Q^{-1} \int_{-\infty}^{\infty} C x^2 dx = 2 \kappa t, \quad (15.75)$$

so that the standard deviation grows according to  $\sqrt{2 \kappa t}$ . As seen in earlier sections, this square root time dependence is characteristic of diffusive processes, which is distinct from the time dependence of ballistic processes, which have a power of  $t$  behavior.<sup>11</sup>

---

<sup>11</sup>The simplest ballistic process arises for particles moving with a constant velocity,  $v$ , so that the position is  $x(t) = v t$ .

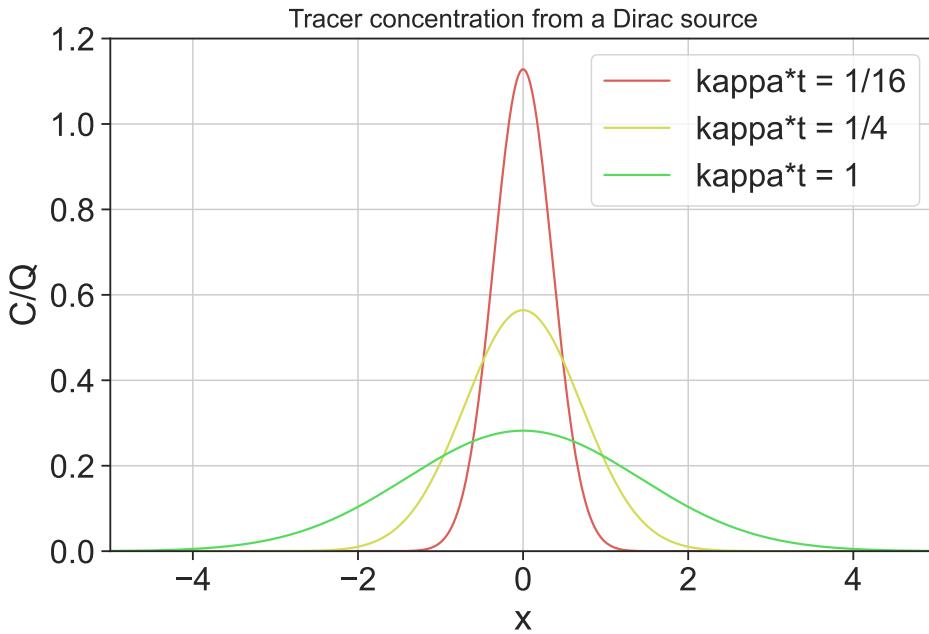


FIGURE 15.3: Illustrating the tracer concentration resulting from a Dirac source at the origin, as given by the Gaussian Green's function (15.73) for three values of  $\kappa t$ . According to equation (15.75), the standard deviation is given by  $\sqrt{2\kappa t}$ , with the standard deviation measuring the spread of the distribution.

## 15.7 Mathematical interlude: tensor analysis tools

As derived in VOLUME 1, the equation for a conservative tracer<sup>12</sup> takes on the form

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}, \quad (15.76)$$

where  $C$  is the tracer concentration scalar, and  $\mathbf{J}$  is a tracer flux vector that embodies molecular diffusion as well as subgrid scale advection and subgrid scale diffusion (Chapter 18). Advective transport from the resolved flow,  $\mathbf{v}$ , appears when transforming to an Eulerian or laboratory reference frame, in which case the tracer equation takes on the equivalent forms

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J} \iff \partial_t(\rho C) + \nabla \cdot (\mathbf{v} \rho C + \mathbf{J}) = 0, \quad (15.77)$$

with  $\mathbf{v} \rho C$  the advective flux. In this chapter we assume  $\mathbf{J}$  arises just from diffusion, and we furthermore ignore any flow (i.e.,  $\mathbf{v} = 0$ ) so that advection is absent.

### 15.7.1 Concerning the upright versus slanted notation

The two forms of the tracer in equation (15.77) are written as a tensor equation, prompting the upright  $C$ ,  $\mathbf{v}$ ,  $\mathbf{J}$ , and  $\nabla$ , which follows the notation convention from the tensor analysis in VOLUME 1. Hence, these equations are valid in any coordinate system. When represented in a particular coordinate system, such as Cartesian, then we write the slanted  $C$ ,  $\mathbf{v}$ ,  $\mathbf{J}$ , and  $\nabla$ , which are particular coordinate representations of  $C$ ,  $\mathbf{v}$ ,  $\mathbf{J}$ , and  $\nabla$ .

The upright versus slanted notation is fundamental conceptually, since it is important to

---

<sup>12</sup>Conservative tracers evolve without sources or sinks, and so their material time evolution is only affected by the convergence of a flux.

appreciate that tensors are geometric objects that are not subject to the whims of a particular coordinate choice. Correspondingly, physically robust differential and integral equations are coordinate invariant. Even so, the upright-slanted notation can be softly adhered to without much cause for concern, so long as we are careful to write the coordinate equations using rules of tensor analysis. In that case, the coordinate equations are unaltered in form when changing coordinates; i.e., they are tensor equations. Developing a practical and conceptual understanding of what careful means in this context requires the tensor analysis material presented in VOLUME 1. In the remainder of this section, we summarize salient points from those chapters that are of particular use in the study of tracer diffusion.

### 15.7.2 Metric tensor allows us to measure distance

The **metric** tensor,  $g$ , is a symmetric and positive-definite (i.e., all eigenvalues are positive) second order tensor that is used to measure the distance between points in space. The **Kronecker delta** is the Cartesian coordinate representation of the metric tensor for Euclidean space. In this case,  $g_{mn} = \delta_{mn}$ , where  $\delta_{mn}$  is unity when  $m = n$  and zero otherwise. In this book, we are only concerned with fluid motion through a background Euclidean space. Even so, we find many occasions to use non-Cartesian coordinates and/or to consider flow constrained to non-Euclidean surfaces that are embedded in Euclidean space (e.g., a spherical planet or an isopycnal surface). Example non-Cartesian coordinates of use for geophysical fluids include the spherical coordinates and cylindrical-polar coordinates detailed in VOLUME 1, the **generalized vertical coordinate** in the present volume, and the Lagrangian or material coordinates of VOLUME 1. For these cases, and others, a coordinate representation of the metric tensor is distinct from the Kronecker delta.

We often have need to work with the inverse metric tensor,  $g^{-1}$ , which we know exists since the metric is a symmetric and positive-definite tensor. To reduce notational clutter, we write  $g^{mn}$  for the coordinate representation of the inverse metric, rather than the more clumsy  $(g^{-1})^{mn}$ . By definition of the inverse metric, we have the following identity holding for any coordinate choice

$$\delta^m{}_n = g^{mp} g_{pn}, \quad (15.78)$$

where  $\delta^m{}_n$  is an expression of the identity tensor, which is coordinate invariant, meaning that it has the same numerical values for all coordinates. Notably, for Cartesian tensors, where  $g_{mn} = \delta_{mn}$ , the identity (15.78) reduces to the identity

$$\delta^m{}_n = \delta^{mp} \delta_{pn}. \quad (15.79)$$

Since  $\delta^{mn} = \delta_{mn}$  (inverse of the Kronecker delta is the same), equation (15.79) means that numerically we have

$$\delta^{mn} = \delta_{mn} = \delta^m{}_n. \quad (15.80)$$

Identities (15.79) and (15.80) are indicative of the relative simplicity of Cartesian tensor analysis, in which there is no distinction between index placement so that there is no quantitative need to keep track of upstairs versus downstairs indices. Yet there is a distinction for general tensors, and so we must be careful to use the metric tensor to raise and lower indices, as now discussed.

### 15.7.3 Raising and lowering tensor indices via the metric tensor

Besides measuring distance between points in space, the metric tensor provides the operational means to raise and lower indices that appear on the representations of tensors. For example,

consider the second order diffusion tensor,  $\mathbf{K}$ , with a particular coordinate representation  $K^m{}_n$ . The  $K^m{}_n$  representation is sometimes referred to as the (1, 1) **natural representation**, and it is related to its totally contravariant (2, 0) **sharp representation**,  $K^{mn}$ , through contraction with the metric tensor,<sup>13</sup>

$$K^m{}_n = \mathfrak{g}_{pn} K^{mp}. \quad (15.81)$$

It follows that to relate the **flat representation** (the (0, 2) representation) to the **sharp representation** requires two contractions with the metric tensor,

$$K_{mn} = \mathfrak{g}_{mq} \mathfrak{g}_{pn} K^{qp}. \quad (15.82)$$

Since the metric is symmetric,  $\mathfrak{g}_{mn} = \mathfrak{g}_{nm}$ , there is no need to be concerned with the ordering of its indices.

#### 15.7.4 Divergence of a vector and the divergence theorem

As seen from the tracer equation (15.77), the convergence of the tracer flux drives the time evolution of conservative tracers. It is thus important to know how the divergence is expressed as a tensor equation. In the general tensor analysis chapter in VOLUME 1, we derived the following coordinate invariant expression for the divergence of a vector

$$\boldsymbol{\nabla} \cdot \mathbf{J} = \nabla_m J^m = \frac{1}{\sqrt{\det(\mathfrak{g}_{mn})}} \partial_m [\sqrt{\det(\mathfrak{g}_{mn})} J^m], \quad (15.83)$$

where  $\nabla_m$  are components to the **covariant derivative**. Equation (15.85) is a convenient result since it only requires partial derivatives in the chosen coordinate system, with all the coordinate dependent properties summarized by the square root of the metric determinant,  $\sqrt{\det(\mathfrak{g}_{mn})}$ . Since  $\sqrt{\det(\mathfrak{g}_{mn})}$  appears in many places within this chapter, we find it useful to introduce the shorthand

$$\} \equiv \sqrt{\det(\mathfrak{g}_{mn})}, \quad (15.84)$$

in which the covariant divergence (15.85) is written in the more tidy manner

$$\boldsymbol{\nabla} \cdot \mathbf{J} = \}^{-1} \partial_m (\} J^m). \quad (15.85)$$

For Cartesian coordinates,  $\} = \sqrt{\det(\mathfrak{g}_{mn})} = 1$ , in which case the divergence in equation (15.85) reduces to its familiar Cartesian form. Yet other coordinates have a nonzero  $\}$ , which accounts for the squeezing and expansion of the coordinate surfaces that affect the divergence. We display the divergence in spherical coordinates and cylindrical-polar coordinates in VOLUME 1, and for **generalized vertical coordinate** in Sections 12.14 and 12.15.

The  $1/\}$  factor appearing in the covariant divergence (15.85) is convenient since it cancels the same factor appearing in the coordinate representation of the volume element. As seen in the general tensor analysis chapter in VOLUME 1, this cancellation greatly simplifies the

---

<sup>13</sup>In the tensor algebra chapter of VOLUME 1, we introduced the **musical nomenclature** sometimes used for the representations of second order tensors. The **natural representation** of a second order tensor occurs with one tensor index upstairs and the other downstairs. The natural representation is sometimes denoted by (1, 1), to indicate the number of indices up and down. The **sharp representation**, or (2, 0) representation, is when the tensor is represented with both indices upstairs. Finally, the **flat representation** or (0, 2) representation is where both indices are downstairs.

divergence theorem, which takes on the form

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{J} dV = \int_{\mathcal{R}} \nabla_m J^m dV = \int_{\mathcal{R}} \partial_m (\} J^m) d^3\xi = \oint_{\partial\mathcal{R}} J^m \hat{n}_m dS, \quad (15.86)$$

where  $\hat{\mathbf{n}}$  is the outward normal, and  $d^3\xi = d\xi^1 d\xi^2 d\xi^3$  is the coordinate volume element.

### 15.7.5 Example tracer fluxes

We here briefly consider example tracer fluxes that are studied later in this chapter or in Chapter 16. We start from their form written in Cartesian coordinates and then transform to general coordinates.

#### Advection tracer flux

The advective tracer flux, as represented using Cartesian coordinates (denoted by  $\xi^m$ ), is written as

$$J^m = \rho v^m C, \quad (15.87)$$

where  $\rho$ ,  $C$ , and  $v^m$  are functions that represent the density, tracer concentration, and velocity using Cartesian coordinates as independent variables. To transform the advective tracer flux to another set of coordinates,  $\xi^{\bar{m}}$ , requires the transformation matrix,  $\Lambda^{\bar{m}}_m$  (see VOLUME 1), where

$$J^{\bar{m}} = \bar{\rho} \bar{C} \Lambda^{\bar{m}}_m v^m = \bar{\rho} \bar{C} v^{\bar{m}}. \quad (15.88)$$

In this equation we wrote  $\bar{\rho}$ ,  $\bar{C}$ , and  $v^{\bar{m}}$  for the functions representing the density, tracer concentration, and velocity with  $\xi^{\bar{m}}$  as the independent variables. Furthermore, the transformation matrix,  $\Lambda^{\bar{m}}_m$ , is built from the partial derivatives of the two sets of coordinates

$$\Lambda^{\bar{m}}_m = \partial \xi^{\bar{m}} / \partial \xi^m. \quad (15.89)$$

For nomenclature brevity in the following, we do not write  $\bar{\rho}$  and  $\bar{C}$  for the scalar fields, thus relying on the functional dependence implicit in the coordinate choice.

#### Isotropic diffusive tracer flux and the Laplace-Beltrami operator

The isotropic diffusive tracer flux has the following representation using Cartesian coordinates

$$J^m = -\rho \kappa \delta^{mn} \partial_n C, \quad (15.90)$$

with  $\kappa > 0$  the diffusivity scalar. We generalize the isotropic diffusive flux to arbitrary coordinates,  $\xi^{\bar{m}}$ , by introducing the inverse metric tensor represented using the  $\xi^{\bar{m}}$  coordinates<sup>14</sup>

$$J^{\bar{m}} = -\rho \kappa \bar{g}^{\bar{m}\bar{n}} \partial_{\bar{n}} C. \quad (15.91)$$

Making use of equation (15.85) leads to the flux convergence

$$-\nabla \cdot \mathbf{J} = \}^{-1} \partial_{\bar{m}} (\rho \kappa \} \bar{g}^{\bar{m}\bar{n}} \partial_{\bar{n}} C). \quad (15.92)$$

---

<sup>14</sup>The covariant derivative acting on a scalar field is just the partial derivative. So we could just as well have written  $J^{\bar{m}} = -\rho \bar{g}^{\bar{m}\bar{n}} \nabla_{\bar{n}} C$  for equation (15.91).

If the product,  $\rho \kappa$ , is a constant in space then the resulting flux convergence exposes the Laplace-Beltrami operator acting on the tracer concentration

$$-\nabla \cdot \mathbf{J} = \rho \kappa \underbrace{\left\{^{-1} \partial_{\bar{m}} (\{} g^{\bar{m}\bar{n}} \partial_{\bar{n}} C \})}_{\text{Laplace-Beltrami acting on } C} \equiv \rho \kappa \bar{\nabla}^2 C. \quad (15.93)$$

### Diffusive tracer flux with an anisotropic diffusion tensor

With a general diffusion tensor (whose form is discussed more in later sections), the Cartesian expression for the diffusive flux is given by

$$J^m = -\rho K^{mn} \partial_n C. \quad (15.94)$$

The tracer flux has a corresponding expression using arbitrary coordinates

$$J^{\bar{m}} = -\rho K^{\bar{m}\bar{n}} \partial_{\bar{n}} C, \quad (15.95)$$

which is identical in form to the Cartesian expression (15.94). We make use of the transformation matrix and its inverse to write

$$\partial_{\bar{n}} = \partial_n \Lambda^n_{\bar{n}} \quad \text{and} \quad K^{\bar{m}\bar{n}} = \Lambda^{\bar{m}}_m \Lambda^{\bar{n}}_n K^{mn}. \quad (15.96)$$

Note that we actually only need to perform a single transformation since the contraction between the diffusion tensor and the derivative of the tracer concentration is coordinate invariant

$$J^{\bar{m}} = -\rho K^{\bar{m}\bar{n}} \partial_{\bar{n}} C = -\rho K^{\bar{m}n} \partial_n C = -\rho \Lambda^{\bar{m}}_m K^{mn} \partial_n C = \Lambda^{\bar{m}}_m J^m. \quad (15.97)$$

### 15.7.6 Comments about the tensor tools

Most of this chapter is unconcerned with the niceties of general tensor analysis. Part of the reason is that Cartesian coordinates are sufficient to exemplify the key maths and physics ideas. Even so, we are mindful to use rudimentary tensor notation, thus allowing for the equations derived with Cartesian coordinates to be valid tensor equations that hold for all coordinates. Given the growing suite of coordinates used in geophysical fluid studies, this extra bit of formalism has a nontrivial payoff.

## 15.8 Further properties of tracer diffusion

We examined a variety of mathematical properties of the diffusion equation in VOLUME 1, as part of our study of [parabolic partial differential equations](#). We also examined the diffusion equation as part of our analysis of the [Green's function](#) method of solutions for [passive tracers](#). Here, we explore further mathematical properties of diffusion. In particular, we here allow for distinct behavior of the diffusive fluxes in the different directions. Such distinctions are relevant especially for the turbulent diffusivity arising in stratified fluids, where turbulent mixing across stratification surfaces is suppressed relative to turbulent mixing parallel to these surfaces<sup>15</sup> For this purpose we make use of the second order positive definite and symmetric [diffusion tensor](#),

---

<sup>15</sup>We pursue this idea far more thoroughly in Chapter 18 when describing closures for the tracer equation in the presence of turbulence.

$\mathbf{K} = \mathbf{K}^T$ , with the resulting downgradient diffusive tracer flux given by

$$J^m = -\rho K^{mn} \partial_n C. \quad (15.98)$$

### 15.8.1 Sample diffusion tensors

For the case of molecular diffusion considered in equation (15.58), the diffusion tensor is generally assumed to be isotropic. Written in Cartesian coordinates, the isotropic diffusion tensor takes the form

$$K^{mn} = \kappa \delta^{mn}, \quad (15.99)$$

whereas with general coordinates it is

$$K^{\overline{mn}} = \kappa g^{\overline{mn}}. \quad (15.100)$$

It is notable that the natural or  $(1, 1)$  form for the molecular diffusion tensor is numerically identical across coordinates, in that

$$K^m{}_p = \kappa \delta^m{}_p \quad \text{and} \quad K^{\overline{m}}{}_{\overline{p}} = \kappa g_{\overline{np}} g^{\overline{mn}} = \kappa \delta^{\overline{m}}{}_{\overline{p}}, \quad (15.101)$$

where we made use of the identity (15.78) satisfied by the metric and its inverse.

To parameterize flows that are turbulent and larger than the microscale (e.g., ocean mesoscale turbulence), it is common to rotate the diffusive fluxes to be along surfaces of constant scalar field,  $\gamma(\mathbf{x}, t)$ , in which case the diffusion tensor is

$$K^{mn} = \kappa (\delta^{mn} - \hat{\gamma}^m \hat{\gamma}^n), \quad (15.102)$$

where

$$\hat{\gamma}_n = \frac{\partial_n \gamma}{|\nabla \gamma|} \quad (15.103)$$

is the surface normal direction. The general coordinate representation of this anisotropic diffusion tensor is

$$K^{\overline{mn}} = \kappa (g^{\overline{mn}} - \hat{\gamma}^{\overline{m}} \hat{\gamma}^{\overline{n}}). \quad (15.104)$$

One choice for the orientation direction is to set  $\hat{\gamma} = \hat{z}$ , in which case the diffusion tensor orients the tracer fluxes along surfaces of constant geopotential to thus realize horizontal diffusion. Another choice, motivated from the physics of ocean mesoscale eddy transport, sets  $\gamma$  equal to a measure of the local buoyancy. In this case we have the neutral diffusion process studied in Section 18.4.

### 15.8.2 Diffusion of tracer concentration powers

In Exercise 15.3, we assume that the tracer concentration satisfies the tracer equation

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}, \quad (15.105)$$

for a diffusive flux,  $\mathbf{J}$ , and then show that for any  $\Gamma \geq 1$  that

$$\rho \frac{DC^\Gamma}{Dt} = -\nabla \cdot \mathbf{J}(C^\Gamma) + \Gamma(\Gamma - 1) C^{\Gamma-2} \mathbf{J} \cdot \nabla C. \quad (15.106)$$

The first term in equation (15.106) is the convergence of the diffusive flux defined in terms of  $C^\Gamma$ . This term acts to diffuse  $C^\Gamma$  just like diffusion acts on  $C$ . The second term in equation (15.106) is negative since the diffusion tensor is symmetric and positive-definite so that

$$\mathbf{J} \cdot \nabla C = -\rho K^{mn} \partial_m C \partial_n C < 0. \quad (15.107)$$

That is, the diffusive flux, by construction, is oriented down the tracer concentration gradient. Consequently, the second term in equation (15.106) always acts to reduce the magnitude of  $C^\Gamma$  towards zero.

### 15.8.3 Moments of tracer concentration

Next we consider the evolution of domain integrated tracer concentration and its powers. To focus on impacts just from diffusion, we assume the boundaries are insulating (i.e., zero normal boundary flux) so that  $\mathbf{J} \cdot \hat{\mathbf{n}} = 0$  with  $\hat{\mathbf{n}}$  the outward normal at the boundary. We also assume the total fluid mass in the domain remains fixed

$$M = \int \rho dV \quad \text{with} \quad \frac{dM}{dt} = 0. \quad (15.108)$$

We can thus treat the domain as material given that there is no exchange of mass or tracer across the boundaries. These assumptions allow us to focus just on the effects from tracer diffusion.

#### Domain average tracer concentration

The domain averaged tracer concentration is defined by

$$\bar{C} = \frac{\int C \rho dV}{M}, \quad (15.109)$$

and it follows that its time derivative vanishes since

$$M \frac{d\bar{C}}{dt} = \frac{d}{dt} \int C \rho dV = \int \frac{DC}{Dt} \rho dV = - \int \nabla \cdot \mathbf{J} dV = - \oint \mathbf{J} \cdot \hat{\mathbf{n}} dS = 0, \quad (15.110)$$

where  $\mathbf{J} \cdot \hat{\mathbf{n}} = 0$  since we are assuming an insulating boundary. Also note that we brought the time derivative inside the integral as a material derivative since the region is itself material, thus allowing us to make use of [Reynolds transport theorem](#) from VOLUME 1. The result (15.110) follows since there is no change in the total mass of fluid nor is there any exchange of tracer across the boundaries. Hence, the domain averaged tracer concentration remains fixed in time.

#### Tracer variance within the domain

The variance of the tracer concentration is defined by

$$\text{var}(C) \equiv \frac{\int (C - \bar{C})^2 \rho dV}{M} = \bar{C^2} - \bar{C}^2 \geq 0. \quad (15.111)$$

The tracer variance measures the deviation of the tracer concentration relative to the domain averaged concentration. Since the domain average remains fixed in time, the time change of

the variance is given by

$$\frac{d[\text{var}(C)]}{dt} = \frac{d\bar{C}^2}{dt}. \quad (15.112)$$

Thus, it is common to refer to  $\bar{C}^2$  as the tracer variance, although strictly speaking only time derivatives of  $\bar{C}^2$  and  $\text{var}(C)$  are equal as per equation (15.112). Performing the time derivative, and again noting that the domain is material so that we can use Reynolds transport theorem, renders

$$M \frac{d\bar{C}^2}{dt} = \frac{d}{dt} \int C^2 \rho dV = 2 \int C \frac{DC}{Dt} \rho dV = -2 \int C \nabla \cdot \mathbf{J} dV = 2 \int \nabla C \cdot \mathbf{J} dV. \quad (15.113)$$

The final equality again made use of the insulating boundary condition,  $\mathbf{J} \cdot \hat{\mathbf{n}} = 0$ . The time change in the tracer variance is thus determined by the integral of the projection of the tracer flux onto the tracer gradient. We already saw from equation (15.107) that diffusive fluxes are oriented down the tracer gradient. Consequently, diffusion of the tracer concentration results in a reduction in tracer variance

$$\frac{d[\text{var}(C)]}{dt} = \frac{d\bar{C}^2}{dt} \leq 0. \quad (15.114)$$

This result further supports our common experience where diffusion removes differences (i.e., gradients) within the tracer field.

### Turbulent stirring enhances the efficiency of mixing from molecular diffusion

Let us expose the diffusion tensor in equation (15.113) to have

$$M \frac{d[\text{var}(C)]}{dt} = -2 \int \nabla C \cdot \mathbf{K} \cdot \nabla C dV, \quad (15.115)$$

which takes on the particularly simple form for molecular diffusion

$$M \frac{d[\text{var}(C)]}{dt} = -2 \kappa \int |\nabla C|^2 dV, \quad (15.116)$$

These expressions highlight a key point that we return to in Chapter 16 as part of our study of advection plus diffusion. In that chapter, we show that advection has no direct impact on tracer variance or any other tracer moments. However, and crucially, advection affects changes to the tracer gradient. Indeed, turbulent flows generally increase the magnitude of the gradient via a stirring process that stretches and folds tracer contours and surfaces. Consequently, as seen from equation (15.116), turbulent stirring increases the impacts from molecular diffusion by enhancing the magnitude of tracer gradients that diffusion can act upon. This mechanism is the fundamental reason that mixing in turbulent flows is far more efficient than mixing in quiescent flows.

### Diffusion of arbitrary tracer moments

Proceeding as before, and dropping boundary contributions since the domain is material and insulating, the identity (15.106) shows that the time derivative of an arbitrary tracer moment is given by

$$\frac{d\bar{C}^\Gamma}{dt} = \Gamma(\Gamma - 1) \int C^{\Gamma-2} \nabla C \cdot \mathbf{J} dV \leq 0. \quad (15.117)$$

For  $\Gamma = 0$  we have an expression of mass conservation for the domain, whereas  $\Gamma = 1$  is an expression of tracer conservation. The case of  $\Gamma = 2$  yields the tracer variance result (15.114). The result for higher powers also holds. Hence, we conclude that the downgradient orientation of diffusive tracer fluxes acts to dissipate all powers of tracer concentration when integrated globally; i.e., all tracer moments are dissipated by diffusion.

## 15.9 Connecting tracer dissipation to the diffusion operator

We here take an excursion into linear operator theory. In particular, we make a connection between the diffusion operator with natural boundary conditions (defined below) and the functional derivative of the global tracer dissipation functional. This connection holds so long as the diffusion operator is self-adjoint and linear, as it is when diffusing passive tracers and with natural boundary conditions. The connection between a linear self-adjoint operator and a functional is developed in such books as *Courant and Hilbert* (1953, 1962). In the simplest case, the Laplacian of the tracer,  $\nabla^2 C$ , is equal to the functional derivative,

$$\nabla^2 C = \frac{\delta \mathcal{F}}{\delta C}, \quad (15.118)$$

where<sup>16</sup>

$$\mathcal{F} \equiv -\frac{1}{2} \int |\nabla C|^2 \rho d^3x \quad (15.119)$$

is the associated functional. In the following, we prove this result for a general diffusion tensor acting on an arbitrary tracer concentration,  $C$ , with the proof holding so long as the diffusion tensor is not a function of the tracer concentration. Besides offering an interesting theoretical tidbit, this result provides a suitable framework for developing numerical methods for discretizing the diffusion operator, with examples provided by *Griffies et al.* (1998) and Chapter 16 of *Griffies* (2004).

### 15.9.1 Fréchet derivative of the diffusion dissipation functional

Define the diffusion dissipation functional

$$\mathcal{F} = \int \mathcal{L} d^3x, \quad (15.120)$$

where the integrand is the negative semi-definite quadratic form

$$2\mathcal{L} = \mathbf{J} \cdot \nabla C = -\rho K^{mn} \partial_m C \partial_n C \leq 0. \quad (15.121)$$

The goal is to relate the diffusion operator, given by the convergence of the diffusion flux,  $-\nabla \cdot \mathbf{J}$ , to the functional derivative of  $\mathcal{F}$ , with the derivative taken with respect to the tracer concentration,  $C$ . We compute the functional derivative using variational calculus technology from VOLUME 1, as well as in our use of Hamilton's principle in VOLUME 4.

For that purpose, consider a functional variation to the tracer concentration,  $\delta C$ , and insert

---

<sup>16</sup>In this section we write the integration volume element as  $d^3x = dV$ . Motivation for this notation will become apparent at the point of equation (15.127).

it into the dissipation functional

$$\delta\mathcal{F} = \int \left[ \delta C \frac{\delta\mathcal{L}}{\delta C} + \delta(\partial_m C) \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right] d^3x. \quad (15.122)$$

As discussed in VOLUME 1, functional variations are perturbations to the form of the function, in which case

$$C \rightarrow C + \delta C \quad \text{with} \quad |\delta C| \ll |C|. \quad (15.123)$$

Notably,  $\delta C$  is itself a function of space and time,  $\delta C(\mathbf{x}, t)$ , but it is assumed to have much smaller magnitude than the concentration,  $C(\mathbf{x}, t)$ . Additionally, the functional variation,  $\delta C$ , has no affect on the space-time points so that the variational operator,  $\delta$ , commutes with space and time derivatives and integrals. Integration by parts on the second term in equation (15.122) leads to

$$\delta\mathcal{F} = \int \left[ \delta C \frac{\delta\mathcal{L}}{\delta C} + \partial_m \left( \delta C \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) - \delta C \partial_m \left( \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) \right] d^3x. \quad (15.124)$$

The middle term is a total derivative that integrates to a boundary contribution and the associated [natural boundary condition](#)

$$\hat{\mathbf{n}} \cdot \frac{\delta\mathcal{L}}{\delta \nabla C} = \hat{\mathbf{n}} \cdot \mathbf{J} = \text{boundary flux}, \quad (15.125)$$

with  $\hat{\mathbf{n}}$  the boundary outward normal. This natural boundary condition is the [Neumann boundary condition](#).

To focus on the connection between the diffusion operator and the diffusion dissipation functional, we ignore boundary fluxes so that the functional variation is given by

$$\delta\mathcal{F} = \int \delta C \left[ \frac{\delta\mathcal{L}}{\delta C} - \partial_m \left( \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) \right] d^3x. \quad (15.126)$$

Consequently, the [functional derivative](#) (also known as the Fréchet derivative) is given by

$$(d^3y)^{-1} \frac{\delta\mathcal{F}}{\delta C(\mathbf{y})} = \frac{\delta\mathcal{L}}{\delta C} - \partial_m \left[ \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right], \quad (15.127)$$

where  $d^3y$  is the volume element at the field point,  $\mathbf{y}$ . To reach the last step required the identity

$$\frac{\delta C(\mathbf{x})}{\delta C(\mathbf{y})} = d^3y \delta(\mathbf{x} - \mathbf{y}), \quad (15.128)$$

where  $\delta(\mathbf{x} - \mathbf{y})$  is the [Dirac delta](#)<sup>17</sup> satisfying

$$\int \delta(\mathbf{x} - \mathbf{y}) d^3y = 1, \quad (15.129)$$

so long as the integration domain includes the singular point  $\mathbf{x} = \mathbf{y}$ . Note that the Dirac delta has dimensions of inverse volume, which necessitates the appearance of the volume factor,  $d^3y$ , on the right hand side of equation (15.128).<sup>18</sup>

<sup>17</sup>Note the unfortunate, though nearly universal, double meaning for the  $\delta$  symbol: one referring to the variation operator and one referring to the Dirac delta.

<sup>18</sup>Many treatments of functional derivatives in mathematics texts ignore the volume factor,  $d^3y$ , in equation

### 15.9.2 Connection to the diffusion operator

Reintroducing the specific form of the diffusion integrand  $2\mathcal{L} = -\rho K^{mn} \partial_m C \partial_n C$  leads to

$$\frac{\delta \mathcal{F}}{\delta C(\mathbf{y})} = -\partial_m \left[ \frac{\delta \mathcal{L}}{\delta (\partial_m C)} \right] d^3 y = \partial_m (\rho K^{mn} \partial_n C) d^3 y. \quad (15.130)$$

The second equality identifies the diffusion operator, thus revealing the connection between the dissipation functional, the diffusion fluxes, and the diffusion operator

$$\frac{\delta \mathcal{F}}{\delta C(\mathbf{y})} = -(\nabla \cdot \mathbf{J}) d^3 y. \quad (15.131)$$

### 15.9.3 Why we need to assume $K^{mn}$ is independent of $C$

There are many geophysical applications in which the diffusion tensor is a function of the tracer concentration, in which case the diffusion equation is no longer a linear differential equation. For example, the neutral diffusion of Section 18.4 makes use of a diffusion tensor that is a function of temperature and salinity gradients. In this case the functional derivative in terms of temperature or salinity appearing in equation (15.130) becomes

$$2 \frac{\delta \mathcal{L}}{\delta (\partial_m C)} = -2 \rho K^{mn} \partial_n C - \rho \partial_m C \partial_n C \frac{\delta K^{mn}}{\delta (\partial_m C)}. \quad (15.132)$$

The specific form of the term  $\delta K^{mn}/\delta(\partial_m C)$  depends on details of the diffusion tensor. Hence, the general results derived above for the linear diffusion equation no longer hold for this nonlinear diffusion equation. We have more to say about nonlinear advection-diffusion in Section 16.7.

### 15.9.4 Relation to Hamilton's principle

We make use of functional derivatives when using Hamilton's principle for non-dissipative continuous systems in VOLUME 4. For those systems, the Euler-Lagrange equation equations of motion result from setting the functional derivative of the action to zero, which is the mathematical statement of Hamilton's principle. In contrast, we here showed that the functional derivative of the tracer dissipation equals to the diffusion operator. In fact, the construction in this section suggests that linear self-adjoint operators, such as generalized Laplacian operators, can generally be expressed as the functional derivative of its corresponding functional. Chapters 16 and 19 of Griffies (2004) provide further examples, with applications to numerical methods. Further mathematical details can be found in such books as Courant and Hilbert (1953, 1962).



## 15.10 Exercises

### EXERCISE 15.1: VERTICAL DIFFUSION OF TEMPERATURE IN THE OCEAN (Vallis, 2017)

There is a natural time scale associated with diffusive transport. This time scale can be found

---

(15.128). Yet for physical applications it is necessary to maintain dimensional consistency, with the volume factor required for that reason. The volume factor also appears when using functional methods to derive numerical discretizations, with examples provided by Griffies et al. (1998), Griffies and Hallberg (2000), and Griffies (2004).

from scaling the diffusion equation, which reveals that it takes the form

$$\tau_{\text{diffusion}} = \Delta^2 / \kappa, \quad (15.133)$$

where  $\Delta$  is the length scale and  $\kappa$  is the kinematic diffusivity (dimensions of squared length per time). We now make use of this time scale to consider the diffusion of temperature in the ocean, with diffusion due solely to molecular processes.

Using the observed value of molecular diffusivity of temperature in water (look it up), estimate the time for a temperature anomaly to mix from the top of the ocean to the bottom, assuming vertical diffusion through the molecular diffusivity is the only means for mixing. This time scale follows from the one-dimensional diffusion equation and is determined by the diffusivity and the depth of the ocean. Comment on whether you think the real ocean has reached equilibrium after the last ice age (which ended about 12Kyr ago).

#### EXERCISE 15.2: SOLUTION TO ONE-DIMENSIONAL DIFFUSION EQUATION

Consider a one-dimensional diffusion equation

$$\partial_t C = \kappa \partial_{zz} C, \quad (15.134)$$

where  $C$  is a tracer concentration (e.g., temperature or salinity),  $\kappa$  is a constant kinematic diffusivity, and  $z$  is the vertical coordinate. Assume the domain has fixed boundaries at  $z = 0$  and  $z = H$ .

- (a) Assume there is a zero flux of tracer at the two boundaries. Mathematically express this no-flux boundary condition.
- (b) Assume that the initial tracer concentration is confined to an area near the center of the domain. Use dimensional analysis to estimate the time scale for the concentration to homogenize throughout the domain.
- (c) Consider the initial-boundary value problem

$$\partial_t C = \kappa \partial_{zz} C, \quad (15.135a)$$

$$\text{no-flux boundary condition from part (b)} \quad (15.135b)$$

$$C(z, t = 0) = C_0 \cos(kz), \quad (15.135c)$$

where  $C_0$  is a constant. What values for the wave-number,  $k$ , satisfy the no-flux boundary condition?

- (d) Solve the diffusion equation analytically for the given initial condition. Hint: consult your favorite partial differential equation book to learn how to solve this linear 1+1 dimensional diffusion equation.
- (e) Explain how the analytical answer you obtained is consistent with the dimensional analysis answer from part (b).

#### EXERCISE 15.3: DIFFUSION OF TRACER CONCENTRATION POWERS

Derive equation (15.106) for the diffusion of tracer concentration powers.

#### EXERCISE 15.4: LOCALLY DISSIPATIVE PROPERTIES OF DIFFUSION

This exercise explores the dissipative property of diffusion when acting on a tracer extrema.

- (a) ONE-DIMENSIONAL DIFFUSION

Consider the diffusion equation in one spatial dimension, and assume a Boussinesq ocean

in which case the density factors are all constant and so can be dropped

$$\partial_t C = \partial_z (\kappa \partial_z C) = \partial_z \kappa \partial_z C + \kappa \partial_{zz} C, \quad (15.136)$$

where  $\kappa(z, t)$  is an **eddy diffusivity** (also *turbulent diffusivity*). The eddy diffusivity is assumed to be a function of  $(z, t)$ , with the spatial dependence determined by the flow. Show that a tracer extrema,  $C^*$ , evolves under diffusion according to

$$\partial_t C^* = \kappa \partial_{zz} C^*. \quad (15.137)$$

So what does diffusion do to a local maxima (e.g., a local hot region) in the tracer field? What about a minima (e.g., a local cold region)? To answer this question, discuss the mathematical equation satisfied by the tracer extrema.

(b) THREE-DIMENSIONAL DIFFUSION

Generalize the above one dimensional result to three dimensions, whereby the diffusivity  $\kappa$  becomes a symmetric positive-definite diffusion *tensor*, in which case

$$\partial_t C = \partial_m (K^{mn} \partial_n C). \quad (15.138)$$

Now consider an extrema in the tracer field, which is defined by

$$\partial_n C^* = 0 \quad \forall n = 1, 2, 3. \quad (15.139)$$

Prove that three dimensional diffusion acts to *dissipate* an extrema. Hint: recall some linear algebra properties of a symmetric positive-definite matrix. In particular, note that a symmetric positive-definite matrix has positive eigenvalues.

**EXERCISE 15.5: DIFFUSION INCREASES INFORMATION ENTROPY OF A TRACER**

Diffusion is an irreversible process. Here we illustrate this property by considering the **information entropy** associated with a non-negative tracer concentration<sup>19</sup>

$$\mathcal{S}_C \equiv - \int (C \ln C) \rho dV, \quad (15.140)$$

where  $C \in [0, 1]$  is the non-dimensional material concentration of a tracer. Show that

$$\frac{d\mathcal{S}_C}{dt} \geq 0 \quad (15.141)$$

over a material region with  $C > 0$  and with downgradient diffusion,  $\mathbf{J} \cdot \nabla C < 0$ . Consequently, diffusion always increases the information entropy. Hint: follow the discussion of tracer moments in Section 15.8.3. This result is related to the **H-theorem** proven by [Boltzmann \(1966\)](#) for the kinetic theory of gases.

**EXERCISE 15.6: PHILLIPS LAYERING INSTABILITY**

This exercise is based on the discussion in Section 12.2 of [Smyth and Carpenter \(2019\)](#), in which we consider an oceanographically relevant example of a turbulent diffusivity that is a function of vertical buoyancy stratification. Under certain circumstances, the flow dependent diffusivity can enhance, rather than reduce, vertical gradients in the buoyancy, with the associated **Phillip's layering instability** leading to layering. We here only work through the basic mathematical

<sup>19</sup>Information entropy is used in statistical physics as a measure of the order/disorder of a probability distribution. We here apply these notions to measure the information entropy of a tracer concentration.

formulation, leaving the interested reader to consult [Smyth and Carpenter \(2019\)](#) for more details.

Consider a buoyancy field that is a function of vertical position and time,  $b(z, t)$ , and let the squared buoyancy frequency be given by the vertical derivative of the buoyancy

$$N^2 = \partial_z b. \quad (15.142)$$

If buoyancy is affected only by vertical diffusion, then its evolution equation is the one-dimensional vertical diffusion equation

$$\partial_t b = \partial_z (\kappa N^2), \quad (15.143)$$

where  $\kappa > 0$  is the vertical diffusivity for buoyancy. Correspondingly, a vertical derivative of the buoyancy equation leads to the evolution equation for the squared buoyancy frequency

$$\partial_t N^2 = \partial_{zz} (\kappa N^2). \quad (15.144)$$

Assume the diffusivity has the following functional dependence

$$\kappa = \kappa(N^2), \quad (15.145)$$

so that it is a function of the squared buoyancy frequency. A physically relevant choice has the diffusivity get smaller as the stratification increases, so that

$$\frac{d\kappa}{dN^2} < 0. \quad (15.146)$$

Now consider the case of a squared buoyancy frequency that is a small deviation relative to a constant background value

$$N^2(z, t) = N_0^2 + \epsilon N_1^2(z, t), \quad (15.147)$$

where  $\epsilon$  is a small non-dimensional number. Derive the condition whereby, to first order in  $\epsilon$ , we have  $N_1^2$  growing in the presence of downgradient diffusion rather than decaying. That is, what is the condition satisfied by  $d\kappa/dN^2$ ,  $N_0^2$ , and  $\kappa$  that renders an unstable diffusion equation, whereby  $\kappa > 0$  leads to an increase in  $N^2$  rather than a decrease?



## Chapter 16

### TRACER ADVECTION AND DIFFUSION

In this chapter we study tracer advection and diffusion, building on the study of tracer diffusion in Chapter 15 and the mathematics of parabolic partial differential equations in VOLUME 1. We focus on the particular case of conservative tracers, which are tracers whose evolution is only affected by advection and diffusion within the fluid interior, along with either a Neumann boundary condition or a Dirichlet boundary condition. The adjective “conservative” refers to the property that such tracers evolve only through the convergence of a tracer flux vector, and so the net tracer content is altered only through transport across boundaries. That is, conservative tracers have no interior sources or sinks, thus making their budgets simpler than other tracers, such as chemical and biogeochemical tracers, that are also affected by sources and sinks.

Example geophysical tracers that are nearly conservative include salinity in the ocean and humidity in the atmosphere. Both of these tracers are material tracers, which measure the ratio of the mass of a matter substance within a fluid element to the mass of the fluid element. Hence, we treat material tracers as non-dimensional scalar fields whose concentrations range from zero to unity. In VOLUME 1 we derived their mass budget equation (also referred to as a continuity equation) when studying the kinematics of material tracers. Conservative Temperature,  $\Theta$ , as defined VOLUME 2, is a nearly conservative thermodynamical tracer that provides a measure of the potential enthalpy in a fluid element.<sup>1</sup> Conservative Temperature is typically measured in K in the atmosphere and °C in the ocean. Finally, there are many applications of conceptual passive tracers, with such tracers assumed to have zero impact on the flow. Passive tracers are versatile theoretical tools allowing us to probe aspects of the flow, including pathways and time scales ([Haine et al., 2025](#)). It is notable that passive tracers afford a Green’s function solution, which proves quite useful in both theory and applications.

#### CHAPTER GUIDE

This chapter follows directly from our study of the maths of tracer diffusion in VOLUME 1, as well as the physics of diffusive processes in Chapter 15. Results for the Boussinesq ocean from VOLUME 2 are found merely by setting the density factor,  $\rho$ , to the Boussinesq reference density,  $\rho_0$ , where it appears in the budget equations of this chapter. We generally assume Cartesian coordinates in this chapter. Even so, the equations are written in a tensorially consistent manner to allow for extension to arbitrary coordinates.

<sup>1</sup>By convention, Conservative Temperature uses capital letters for its name.

---

<b>16.1</b>	<b>Introduction to advection and diffusion . . . . .</b>	<b>545</b>
16.1.1	Eckart's conceptual description of stirring and mixing . . . . .	545
16.1.2	A kinematic model . . . . .	546
16.1.3	Further study . . . . .	547
<b>16.2</b>	<b>Perfect fluid tracer advection . . . . .</b>	<b>548</b>
16.2.1	The advection equation . . . . .	548
16.2.2	Eulerian time tendencies from advection . . . . .	549
16.2.3	Impermeability property of tracer isosurfaces . . . . .	549
<b>16.3</b>	<b>Worked example: hyperbolic flow . . . . .</b>	<b>549</b>
16.3.1	Solution via the fluid particle trajectories . . . . .	550
16.3.2	Further study . . . . .	551
<b>16.4</b>	<b>Mathematical properties of tracer advection . . . . .</b>	<b>551</b>
16.4.1	Material constancy of $C^F$ . . . . .	551
16.4.2	Evolution of squared tracer gradient . . . . .	551
16.4.3	Eddy-induced and residual mean . . . . .	552
16.4.4	Advection tracer fluxes and skew tracer fluxes . . . . .	553
16.4.5	Skew diffusion . . . . .	554
16.4.6	A comment about skew fluxes and Lagrangian kinematics . . . . .	555
16.4.7	Further reading . . . . .	555
<b>16.5</b>	<b>Advection and skewson . . . . .</b>	<b>555</b>
16.5.1	Choosing a gauge . . . . .	555
16.5.2	Vertical gauge . . . . .	556
16.5.3	Boundary conditions . . . . .	557
<b>16.6</b>	<b>Finite volume budgets with eddy velocities . . . . .</b>	<b>558</b>
16.6.1	Advective flux formulation . . . . .	559
16.6.2	Skew flux formulation . . . . .	560
16.6.3	Domain with a tracer boundary . . . . .	560
16.6.4	Budget for a region with interior sides . . . . .	563
16.6.5	Budget for a perfect fluid in a region with interior sides . . . . .	565
<b>16.7</b>	<b>Active tracers and dia-surface flow . . . . .</b>	<b>567</b>
16.7.1	Adiabatic flow . . . . .	568
16.7.2	Diabatic processes generating dia- $\Theta$ transport . . . . .	568
<b>16.8</b>	<b>Tracer homogenization inside closed tracer contours . . . . .</b>	<b>569</b>
16.8.1	Proof of the theorem . . . . .	569
16.8.2	Comments . . . . .	572
16.8.3	Further study . . . . .	572
<b>16.9</b>	<b>Green's function method for passive tracers . . . . .</b>	<b>572</b>
16.9.1	Concerning time dependent domain boundaries . . . . .	573
16.9.2	Passive tracer boundary conditions . . . . .	573
16.9.3	Advection-diffusion initial-boundary value problem . . . . .	575
16.9.4	The Green's function problem . . . . .	575
16.9.5	Adjoint Green's function problem . . . . .	576
16.9.6	Reciprocity condition . . . . .	576
16.9.7	Composition property . . . . .	579
16.9.8	Integral expression for the tracer concentration . . . . .	581
16.9.9	Properties of the Green's function solution . . . . .	582
16.9.10	Boundary propagator . . . . .	583
16.9.11	Comments . . . . .	585
<b>16.10</b>	<b>Exercises . . . . .</b>	<b>585</b>

---

## 16.1 Introduction to advection and diffusion

In this chapter we consider the equation describing the evolution of a [conservative tracer](#)

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J} \iff \partial_t(\rho C) + \nabla \cdot (\mathbf{v} \rho C + \mathbf{J}) = 0, \quad (16.1)$$

where the flow is nonzero,  $\mathbf{v} \neq 0$ , so that advection contributes to tracer evolution in addition to subgrid scale fluxes, such as [diffusion](#), captured by the tracer flux  $\mathbf{J}$ . For the first part of this chapter, we focus on the effects from advection alone, in which  $\mathbf{J} = 0$  thus reducing equation (16.1) to [advection equation](#)

$$\rho \frac{DC}{Dt} = \partial_t(\rho C) + \nabla \cdot (\mathbf{v} \rho C) = 0. \quad (16.2)$$

In this case, convergences of the advective tracer flux render a reversible [stirring](#) that stretches fluid elements. This stirring, particularly in the presence of turbulent flows, can increase the magnitude of tracer concentration gradients (see Section 16.4.2), and it does so while maintaining, for each fluid element, a fixed mass for all matter constituents and fixed specific entropy.<sup>2</sup>

When [diffusion](#) is enabled, as seen in Chapter 15, the fluid experiences an irreversible exchange that causes [mixing](#) of properties between fluid elements. Correspondingly, diffusion reduces the magnitude of property gradients between fluid elements. When acting together, advection is no longer a pure stirring and diffusion is no longer a pure mixing. Indeed, in the steady state, advection and diffusion exactly balance. [Eckart \(1948\)](#) articulated what has become the standard conceptual paradigm for stirring and mixing in geophysical fluids, with elements of that paradigm reflected in this chapter.

### 16.1.1 Eckart's conceptual description of stirring and mixing

[Eckart \(1948\)](#) provided a conceptual description of transport in fluids that describes the combined roles of advection plus molecular diffusion in affecting the mixing of fluid properties. Quoting from his introduction, he proposes the following three phases.

It is useful to consider a trivial experiment by way of introduction: the mixing of coffee and cream. Three more or less distinct stages can be observed.

1. The initial stage, in which rather large volumes of cream and coffee are distinctly visible; there are sharp gradients at the interfaces between the volumes, but elsewhere the gradient is practically zero. Averaged over the entire volume, the gradient is small. If motion of the liquid is avoided, this state persists for a considerable time.
2. The intermediate stage, after motion has been induced by stirring the liquids; the masses of cream and coffee are distorted, with a rapid increase in the extent of the interfacial regions having high concentration gradients. The average value of the gradient is correspondingly increased.
3. The final stage, in which the gradients disappear, apparently quite suddenly and spontaneously, with the liquid becoming homogeneous.

---

<sup>2</sup>Recall from our study of thermodynamics in VOLUME 2 that specific entropy remains materially constant on fluid elements in the absence of mixing or [diabatic](#) sources.

It is a reasonable working hypothesis to assume that these three stages (or at least the second and third) also occur in the ocean and atmosphere when concentration or temperature differences arise.

### 16.1.2 A kinematic model

Consider the following kinematic advection-diffusion model to illustrate Eckart's three stages of stirring and mixing

$$\partial_t C + u \partial_x C = \kappa (\partial_{xx} + \partial_{yy}) C, \quad (16.3)$$

with tracer concentration,  $C = C(x, y, t)$ , and the domain a zonally periodic channels with free-slip walls at the northern boundaries. The prescribed zonal and non-divergent flow is static and a quadratic function of meridional position

$$u(y) = u_\circ (y/L)^2, \quad (16.4)$$

with  $u_\circ$  an arbitrary velocity scale, and  $y \in [0, L]$  the meridional domain. We non-dimensionalize the advection-diffusion equation (16.3) using the following non-dimensional variables

$$(\bar{x}, \bar{y}) = (x, y)/L \quad \text{and} \quad \bar{t} = (u_\circ/L) t, \quad (16.5)$$

which renders

$$\partial_{\bar{t}} C + \bar{y}^2 \partial_{\bar{x}} C = \text{Pe}^{-1} (\partial_{\bar{xx}} + \partial_{\bar{yy}}) C, \quad (16.6)$$

where we introduced the **Peclet number**,

$$\text{Pe} = u_\circ L / \kappa. \quad (16.7)$$

The Peclet number is sole non-dimensional number for this system and it measures the strength of the advection versus the diffusion. Note that we chose to non-dimensionalize time according to an advective time scale,  $L/u_\circ$ , which is the natural time scale for this system.

Figure 16.1 illustrates numerical solutions for two cases, one without diffusion and one with diffusion. As time moves ahead, the meridional sheared zonal flow causes flow next to the northern boundary to wrap around the periodic channel, whereas flow near the southern boundary is quiescent. Such differential flow acts to progressively increase the magnitude of the tracer gradient. Absent diffusion, this tracer gradient magnitude grows unbounded. In contrast, and as described by [Eckart \(1948\)](#), once the tracer gradient magnitudes are sufficiently large, diffusion damps the tracer gradients toward zero magnitude, which is precisely what we see in Figure 16.1 for the case with diffusion.

To be a bit more quantitative requires us to anticipate results derived later in this chapter. In Section 16.4.1 we show that advection leaves all powers of the tracer concentration unchanged, so that advection leaves the domain integrated tracer variance unchanged.<sup>4</sup> In contrast, from Section 15.8.2 we showed that diffusion always reduces the magnitude of tracer powers, which means that it decreases the tracer variance. The key role for advection appears when characterizing the evolution of  $|\nabla C|^2$ , in which equation (16.32) shows that

$$\frac{1}{2} \frac{\text{D}|\nabla C|^2}{\text{Dt}} = -\nabla C \cdot \mathbf{S} \cdot \nabla C. \quad (16.8)$$

<sup>4</sup>Numerical discretization errors in Figure 16.1 cause a slight decrease in the tracer variance, but the decrease is negligible for our purposes.

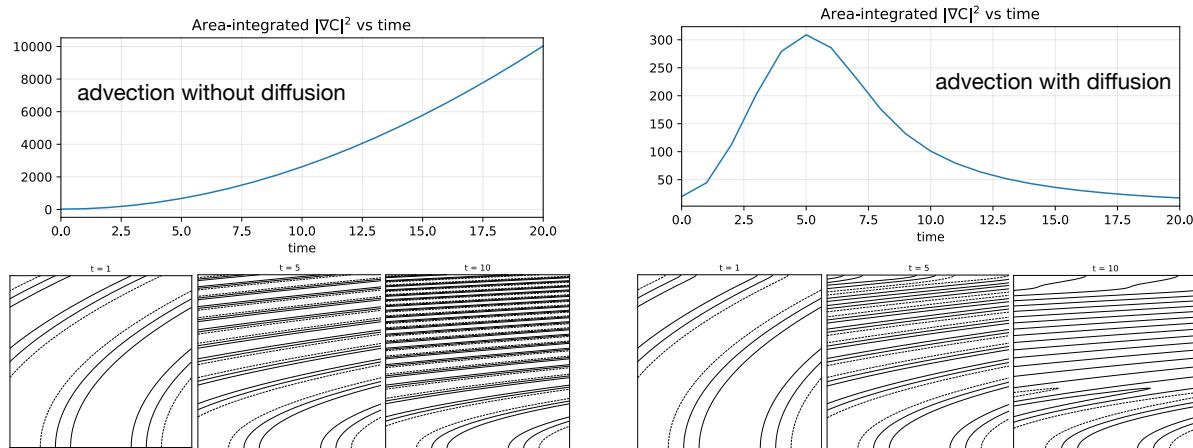


FIGURE 16.1: Illustrating Eckart’s stirring and mixing mechanism for a non-dimensional meridionally sheared zonal flow,  $\bar{u} = \bar{y}^2$ , in a zonally periodic and square channel with solid walls at the north and southern boundaries. The tracer satisfies homogeneous Neumann boundary conditions at the walls, so that the area integrated tracer concentration remains constant in time. The initial tracer profile is given by  $C(x, y, 0) = \cos(2\pi x/L)$ , in which the tracer contours are aligned parallel to the  $y$ -axis.<sup>3</sup> The left panels show three snapshots of the tracer field ( $t = 1, 5, 10$ ) for the case of advection without diffusion. Here, the area integrated tracer variance remains unchanged, and yet the squared tracer gradient grows monotonically as the tracer contours get progressively wrapped around the periodic channel so to introduce increasingly fine scale gradients. The right panel shows the case with a diffusivity, measured by  $\text{Pe}^{-1} = 10^{-4}$ , in which case  $\int |\nabla C|^2 dA$  peaks around non-dimensional time 5 (the non-dimensional time step is  $1/10$ ), and then drops thereafter. Evidently, once the tracer gradient reaches large enough magnitude, the mixing from Laplacian diffusion can efficiently act to mix the concentration. Note that the contours on each panel extend over the range  $\pm 1$  so that they are all directly comparable. However, the vertical range for the top panel is different, with the case on the right with diffusion having a much smaller vertical range than the left pure advection case.

For the prescribed zonal flow in Figure 16.1, we have the tracer gradient source term given by

$$-\nabla C \cdot \mathbf{S} \cdot \nabla C = -2\bar{y}\partial_{\bar{x}}C\partial_{\bar{y}}C. \quad (16.9)$$

The non-dimensional meridional position is positive,  $\bar{y} > 0$ , for this domain, and it weights properties in the north more than in the south. Evidently, a sufficient condition for growth of the squared tracer gradient is that the slope of a tracer contour line in the  $\bar{x}$ - $\bar{y}$  plane is positive

$$\text{slope} = \left[ \frac{\partial \bar{y}}{\partial \bar{x}} \right]_C = -\frac{\partial_{\bar{x}}C}{\partial_{\bar{y}}C}. \quad (16.10)$$

This slope is indeed positive throughout the simulation with pure advection. However, in the presence of diffusion, the source becomes negative around time 5 (see Figure 16.1), which then acts to reduce the tracer gradients. Apparently, the diffusion, primarily that acting in the far north, wipes out tracer gradients sufficiently to change the sign of the source.

### 16.1.3 Further study

Eckart’s conceptual description of stirring and mixing pervades our understanding of mixing within geophysical fluids, where turbulence is ubiquitous. [Young \(1999\)](#), [Müller and Garrett \(2002\)](#), and [Garrett \(2006\)](#) provide updated discussions drawing mostly on ocean examples though with relevance to all geophysical fluids.

The kinematic model in Figure 16.1 is inspired by a similar model considered by [Young \(1999\)](#), who used a circular domain rather than a zonally periodic channel. We chose the zonally

periodic channel in order to avoid the polar coordinate singularity arising at the center of the circle. Such coordinate singularities are a pesky feature of any numerical discretization, particularly for the Laplacian operator that takes on the polar coordinate form (see VOLUME 1)

$$\nabla^2 C = r^{-2} \partial_{\varphi\varphi} C + r^{-1} \partial_r(r \partial_r C). \quad (16.11)$$

## 16.2 Perfect fluid tracer advection

A **perfect fluid** is comprised of material fluid elements whose matter content and thermodynamic properties remain fixed. From the discussion of **molecular diffusion** in Chapter 15, we know that a perfect fluid can at most consist of a single matter constituent and uniform thermodynamic properties. The reason is that in the presence of multiple constituents with non-uniform concentrations, molecular motions irreversibly exchange matter and thermodynamic properties (e.g., temperature, specific entropy) among fluid elements. This exchange, or **mixing**, breaks the assumption of a perfect fluid. Nonetheless, we find many occasions to ignore molecular diffusion when focusing on macroscopic motions of the continuum fluid. Such is the case when considering the advection equation in the absence of mixing.

### 16.2.1 The advection equation

In the absence of mixing or other irreversible processes, the matter content of a fluid element remains fixed as the element moves within the fluid environment. Since the total mass of the element is also constant, then the tracer concentration remains constant and thus satisfies the reversible (source-free) **advection equation**

$$\frac{DC}{Dt} = (\partial_t + \mathbf{v} \cdot \nabla) C = 0. \quad (16.12)$$

The first equality relates the material time derivative to the Eulerian time derivative plus advective transport, with  $\mathbf{v}$  the **barycentric velocity** of a fluid element. We can convert the material form of the advection equation (16.12) into a **flux-form conservation law** or **tracer continuity equation** by combining with the mass continuity equation

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (16.13)$$

which yields

$$\partial_t(\rho C) + \nabla \cdot (\rho C \mathbf{v}) = 0. \quad (16.14)$$

The material form of the advection equation states that tracer concentration is a **material invariant** in the absence of sources or mixing. Hence, a general solution to the advection equation is

$$C(\mathbf{x}, t) = C[\mathbf{X}(0)], \quad (16.15)$$

where  $\mathbf{X}(0)$  is the initial position of a fluid element that is at the position  $\mathbf{x}$  at time  $t$ . If we know the trajectories for all fluid elements and their initial tracer concentration, then we know the tracer concentration for all space and time. For those cases where trajectories are unknown (which is the normal case), it is useful to make use of the Eulerian form of the advection equation to deduce the evolution of tracer concentration.

### 16.2.2 Eulerian time tendencies from advection

At a point in the fluid, the advection equation (16.12) leads to the Eulerian time tendency for tracer concentration

$$\partial_t C = -\mathbf{v} \cdot \nabla C. \quad (16.16)$$

Geometrically, the tendency arises from the projection of the fluid velocity onto the normal to surfaces of constant tracer concentration (isosurfaces). The concentration remains fixed in time (steady) at points where the velocity is parallel to tracer surfaces.

From the flux-form advection equation (16.14), the density-weighted tracer concentration (the tracer mass per volume) has an Eulerian time tendency given by the convergence of the advective flux

$$\partial_t (\rho C) = -\nabla \cdot (\rho C \mathbf{v}). \quad (16.17)$$

The tendency vanishes at a point if there is no convergence of tracer mass towards the point.

### 16.2.3 Impermeability property of tracer isosurfaces

We offer a geometric interpretation of the advection equation

$$(\partial_t + \mathbf{v} \cdot \nabla) C = 0, \quad (16.18)$$

following the discussion of dia-surface transport in Section 13.4. For this purpose, introduce the unit normal on a tracer isosurface

$$\hat{\mathbf{n}} = \frac{\nabla C}{|\nabla C|} \quad (16.19)$$

and the normal projection for the velocity of a point on that surface

$$\mathbf{v}^{(C)} \cdot \hat{\mathbf{n}} = -\frac{\partial_t C}{|\nabla C|}. \quad (16.20)$$

The advection equation (16.18) thus can be written as an impermeability condition for a tracer isosurface

$$\rho (\mathbf{v} - \mathbf{v}^{(C)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{on } C \text{ isosurfaces.} \quad (16.21)$$

We encountered this condition in VOLUME 1 when studying the kinematics of a moving material surface. Hence, in the absence of mixing, tracer isosurfaces are indeed material surfaces since they allow no fluid elements, moving with the fluid velocity  $\mathbf{v}$ , to cross them. This is an important kinematic result that is extended in Section 16.6.5 to include effects from an eddy induced velocity.

## 16.3 Worked example: hyperbolic flow

Consider the following two-dimensional flow in a constant density fluid

$$\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi \quad \text{with} \quad \psi = -\alpha x y \implies \mathbf{u} = \alpha (\hat{\mathbf{x}} x - \hat{\mathbf{y}} y), \quad (16.22)$$

where  $\alpha$  is a constant with dimensions of inverse time. This horizontally non-divergent hyperbolic flow stretches tracer contours along the  $x$ -axis and compresses along the  $y$ -axis. We depict this flow in Figure 16.2, which we studied as part of fluid kinematics in VOLUME 1. This flow is

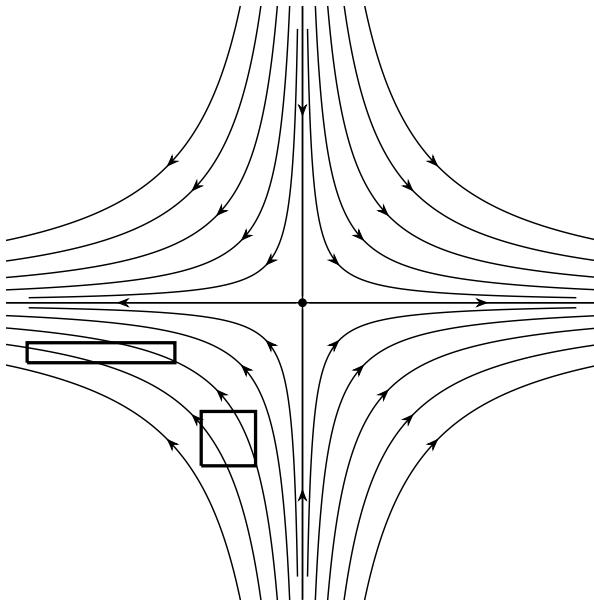


FIGURE 16.2: An ensemble of hyperbolic fluid particle trajectories defined from the horizontally non-divergent velocity field,  $\mathbf{u} = \alpha(\hat{\mathbf{x}}x - \hat{\mathbf{y}}y)$  and streamfunction,  $\psi = -\alpha xy$ . The flow stretches along the  $x$  direction and compresses along the  $y$  direction, as depicted by the material square region evolving to a rectangular region.

simple enough to allow for an exact solution to the fluid particle trajectories, which then means we can provide an exact solution to the tracer advection equation

$$(\partial_t + \alpha x \partial_x - \alpha y \partial_y)C = 0. \quad (16.23)$$

### 16.3.1 Solution via the fluid particle trajectories

A fluid particle trajectory,  $\mathbf{X}(t)$ , is found through time integrating

$$\frac{dX}{dt} = \alpha X \implies X(t) = X(0) e^{\alpha t} \quad (16.24a)$$

$$\frac{dY}{dt} = -\alpha Y \implies Y(t) = Y(0) e^{-\alpha t}, \quad (16.24b)$$

where  $\hat{\mathbf{x}}X(0) + \hat{\mathbf{y}}Y(0)$  is the initial position for the fluid particle. Knowing the fluid particle trajectory allows us to determine the tracer concentration at any time,  $t$ , given the initial concentration,  $C(x, y, 0)$ . We do so by noting that at time  $t$ , the trajectory that passes through the point  $(x, y)$  started at at time  $t = 0$  at the point

$$\hat{\mathbf{x}}X(0) + \hat{\mathbf{y}}Y(0) = \hat{\mathbf{x}}x e^{-\alpha t} + \hat{\mathbf{y}}y e^{\alpha t}. \quad (16.25)$$

Consequently, the concentration is

$$C(x, y, t) = C(x e^{-\alpha t}, y e^{\alpha t}, 0). \quad (16.26)$$

Referring to Figure 16.2, we know the concentration in the rectangular box region at time  $t$ , if we know the concentration in the square box at some earlier time.

### 16.3.2 Further study

This example is taken from Section 4.3 of [Young \(1999\)](#). Furthermore, we provide details for the Lagrangian coordinate representation in VOLUME 1, including an expression for the diffusion operator in Lagrangian coordinates.

## 16.4 Mathematical properties of tracer advection

We now explore various mathematical properties of the advection equation. For that purpose, recall the mass continuity equation (16.13) and flux-form tracer advection equation (16.17)

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (16.27a)$$

$$\partial_t(\rho C) + \nabla \cdot (\rho C \mathbf{v}) = 0. \quad (16.27b)$$

These equations manifest compatibility whereby the tracer equation (16.27b) reduces to the mass continuity equation (16.27a) if the tracer concentration is spatially uniform.

### 16.4.1 Material constancy of $C^\Gamma$

$C^\Gamma$ , for any constant  $\Gamma$ , is materially constant due to the material constancy of  $C$ . We show this property mathematically by noting that the chain rule holds for a material time derivative, so that

$$\frac{DC^\Gamma}{Dt} = \Gamma C^{\Gamma-1} \frac{DC}{Dt} = 0. \quad (16.28)$$

Likewise, making use of the Eulerian form yields

$$\partial_t C^\Gamma + \mathbf{v} \cdot \nabla C^\Gamma = \Gamma C^{\Gamma-1} [\partial_t C + \mathbf{v} \cdot \nabla C] = 0. \quad (16.29)$$

We conclude that advection, in the absence of diffusion, serves to reversibly transport the tracer concentration without altering any of its powers. Correspondingly, all tracer moments are untouched by advection.

### 16.4.2 Evolution of squared tracer gradient

As noted at the start of Section 16.1, some flows can enhance the magnitude of the tracer concentration gradient,  $|\nabla C|^2 = \partial_m C \delta^{mn} \partial_n C$ . To mathematically manifest that process, consider the following equation for the evolution of the squared tracer gradient under the effects of advection

$$(1/2) \partial_t (\partial_m C \delta^{mn} \partial_n C) = \partial_m C \delta^{mn} \partial_n \partial_t C \quad (16.30a)$$

$$= -\partial_m C \delta^{mn} \partial_n (v^k \partial_k C) \quad (16.30b)$$

$$= -\partial_m C \delta^{mn} (v^k \partial_k \partial_n C + \partial_n v^k \partial_k C) \quad (16.30c)$$

$$= -(1/2) v^k \partial_k (\partial_m C \delta^{mn} \partial_n C) - \partial_m C \delta^{mn} G^k{}_n \partial_k C \quad (16.30d)$$

$$= -(1/2) (\mathbf{v} \cdot \nabla) |\nabla C|^2 - \partial_m C \delta^{mn} S^k{}_n \partial_k C \quad (16.30e)$$

$$= -(1/2) (\mathbf{v} \cdot \nabla) |\nabla C|^2 - \nabla C \cdot \mathbf{S} \cdot \nabla C. \quad (16.30f)$$

The second equality made use of the advection equation,  $\partial_t C = -v^k \partial_k C$ , and then we introduced the velocity gradient tensor,  $\mathbf{G}$ , from our study of flow kinematics in VOLUME 1, as well as its

symmetric component, the strain rate tensor,  $\mathbf{S}$ ,

$$G^m{}_n = \partial_n v^m \quad \text{and} \quad S^m{}_n = (G^m{}_n + G_n{}^m)/2. \quad (16.31)$$

We are thus led to the material evolution equation

$$\frac{1}{2} \frac{D|\nabla C|^2}{Dt} = -\nabla C \cdot \mathbf{S} \cdot \nabla C. \quad (16.32)$$

The strain rate tensor is symmetric and yet it is not positive-definite. Hence, we can have either growth or decay of the squared tracer gradient depending on details of the velocity gradient and tracer gradient.

Much of the stirring by turbulent geophysical fluid flows occurs on quasi-horizontal surfaces. Following [Okubo \(1970\)](#); [Weiss \(1991\)](#); [Lapeyre et al. \(1999\)](#); [Balwada et al. \(2021\)](#), we find it useful to examine the evolution of  $|\nabla C|^2$  in a two-dimensional flow. In this case it is relatively straightforward to diagonalize the strain rate tensor.<sup>5</sup> The two-dimensional rate of strain tensor has the following eigenvalues

$$\lambda = (\Delta \pm \sigma)/2 \quad \text{with} \quad \Delta = \nabla \cdot \mathbf{u} \quad \text{and} \quad \sigma^2 = (\partial_x v + \partial_y u)^2 + (\partial_x u - \partial_y v)^2, \quad (16.33)$$

with  $\Delta$  the horizontal divergence and  $\sigma$  the magnitude of the strain rate. By rotating to directions defined by the eigenvectors, we simplify the source in the evolution equation (16.32) to read

$$\frac{D|\nabla C|^2}{Dt} = -(\partial_{\bar{x}} C)^2 (\Delta + \sigma) - (\partial_{\bar{y}} C)^2 (\Delta - \sigma) = -\Delta |\nabla C|^2 - \sigma [(\partial_{\bar{x}} C)^2 - (\partial_{\bar{y}} C)^2], \quad (16.34)$$

where  $\bar{x}$  and  $\bar{y}$  are the orthogonal coordinates aligned with the directions defined by the eigenvectors. Evidently, any horizontal flow convergence,  $\Delta < 0$ , leads to  $|\nabla C|^2$  growth, whereas the strain magnitude,  $\sigma \geq 0$ , leads to growth or decay depending on the relative magnitude of the tracer gradient along the eigendirections.

### 16.4.3 Eddy-induced and residual mean

The mass density time tendency

$$\partial_t \rho = -\nabla \cdot (\mathbf{v} \rho) \quad (16.35)$$

remains unchanged if the advective mass flux,  $\rho \mathbf{v}$  (dimensions of mass per time per area), is modified by the addition of a total curl

$$\rho \mathbf{v} \rightarrow \rho \mathbf{v}^\dagger = \rho \mathbf{v} + \nabla \times (\rho \Psi^*). \quad (16.36)$$

The arbitrariness manifest in equation (16.36) is known as a [gauge symmetry](#). The additional mass flux,  $\nabla \times (\rho \Psi^*)$ , leads to no accumulation of mass at a point since it has zero divergence. In the case of a Boussinesq ocean with  $\rho$  set to a constant  $\rho_0$ , the divergent-free velocity  $\nabla \times \Psi^*$  leads to zero accumulation of volume at a point.

The non-divergent mass flux

$$\rho \mathbf{v}^* \equiv \nabla \times (\rho \Psi^*) \quad (16.37)$$

often arises when we decompose the mass flux into a mean and non-divergent eddy fluctuations.

---

<sup>5</sup>The strain rate tensor is symmetric so it has real and orthogonal eigenvectors. The three dimensional case is algebraically more tedious than the two dimensional case considered here.

In that context, we use the following terminology:

$$\mathbf{v} = \text{Eulerian mean velocity} \quad (16.38a)$$

$$\rho \mathbf{v} = \text{Eulerian mean mass flux} \quad (16.38b)$$

$$\mathbf{v}^* = \text{eddy-induced velocity} \quad (16.38c)$$

$$\rho \Psi^* = \text{eddy-induced mass streamfunction} \quad (16.38d)$$

$$\rho \mathbf{v}^* = \nabla \times (\rho \Psi^*) = \text{eddy-induced mass flux} \quad (16.38e)$$

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^* = \text{residual mean velocity} \quad (16.38f)$$

$$\rho \mathbf{v}^\dagger = \rho(\mathbf{v} + \mathbf{v}^*) = \text{residual mean mass flux.} \quad (16.38g)$$

The name “residual mean” is motivated since the **residual mean velocity**,  $\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*$ , is often smaller than either term individually. That is, the eddy contribution often compensates for the mean, with sum of the mean and eddy representing a residual. We study particular forms of the parameterized eddy induced velocity in Chapter 18.

#### 16.4.4 Advection tracer fluxes and skew tracer fluxes

Following from the previous discussion, we consider the advection equation with the advective tracer transport determined by the residual mean velocity

$$\partial_t(\rho C) + \nabla \cdot (\rho C \mathbf{v}^\dagger) = 0. \quad (16.39)$$

Given the form (16.37) for the eddy mass flux  $\rho \mathbf{v}^*$ , we can write the **advective tracer flux** as

$$\rho C \mathbf{v}^\dagger = C(\rho \mathbf{v} + \rho \mathbf{v}^*) \quad (16.40a)$$

$$= C \rho \mathbf{v} + C \nabla \times (\rho \Psi^*) \quad (16.40b)$$

$$= C \rho \mathbf{v} + \nabla \times (C \rho \Psi^*) - \nabla C \times \rho \Psi^*. \quad (16.40c)$$

It is the convergence of the tracer flux that determines the time tendency, in which the total curl plays no role

$$-\nabla \cdot (\rho C \mathbf{v}^\dagger) = -\nabla \cdot (\rho C \mathbf{v} + \rho C \mathbf{v}^*) \quad (16.41a)$$

$$= -\nabla \cdot (\rho C \mathbf{v} - \nabla C \times \rho \Psi^*). \quad (16.41b)$$

That is, the convergence of the advective tracer flux equals to the convergence of the **skew tracer flux**

$$\underbrace{-\nabla \cdot (\rho C \mathbf{v}^*)}_{\text{advective flux convergence}} = \underbrace{-\nabla \cdot (-\nabla C \times \rho \Psi^*)}_{\text{skew flux convergence}} \quad (16.42)$$

since the advective flux and skew flux differ by a rotational flux,

$$\mathbf{J}^{\text{adv}} = \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}, \quad (16.43)$$

where

$$\mathbf{J}^{\text{adv}} = C \rho \mathbf{v}^* \quad \text{and} \quad \mathbf{J}^{\text{skew}} = -\nabla C \times \rho \Psi^* \quad \text{and} \quad \mathbf{J}^{\text{rot}} = \nabla \times (\rho C \Psi^*). \quad (16.44)$$

Notably, the skew tracer flux is neither upgradient nor downgradient. Rather, it is oriented parallel to isosurfaces of tracer concentration (see Figure 16.3)

$$\nabla C \cdot \mathbf{J}^{\text{skew}} = \nabla C \cdot (-\nabla C \times \rho \Psi^*) = 0. \quad (16.45)$$

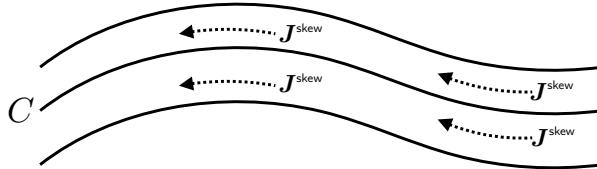


FIGURE 16.3: Skew tracer flux (dotted lines with arrows) for a tracer concentration,  $C$ , are oriented parallel to lines of constant tracer concentration (tracer isolines are the solid lines). Even though the skew fluxes are aligned with tracer contours, skew fluxes generally affect a time change to the tracer concentration through their convergence.

#### 16.4.5 Skew diffusion

Introducing tensor labels brings the skew tracer flux into the form

$$(\mathbf{J}^{\text{skew}})^m = -(\nabla C \times \rho \Psi^*)^m \quad (16.46a)$$

$$= -\epsilon^{mnp} \partial_n C \rho \Psi_p^* \quad (16.46b)$$

$$\equiv -\rho A^{mn} \partial_n C, \quad (16.46c)$$

where we defined the anti-symmetric **skew diffusion** tensor

$$A^{mn} \equiv \epsilon^{mnp} \Psi_p^* \implies \mathbf{A} = \begin{bmatrix} 0 & \Psi_3^* & -\Psi_2^* \\ -\Psi_3^* & 0 & \Psi_1^* \\ \Psi_2^* & -\Psi_1^* & 0 \end{bmatrix}. \quad (16.47)$$

Evidently, tracer advection by a non-divergent mass flux is equivalent to skew-diffusion through the action of an anti-symmetric tensor.

Skew and advective fluxes possess the following dual properties.

- **DERIVATIVE OPERATOR:** The skew flux is proportional to the vector streamfunction and the gradient of the tracer, whereas the advective flux is related to the curl of the streamfunction and the value of the tracer concentration. In effect, the fluxes swap their placement of the derivative operator. Correspondingly, the advective flux vanishes if the velocity vanishes, whereas the skew flux vanishes if the tracer gradient vanishes (just as for a diffusive flux).
- **FLUX ORIENTATION:** The orientation of the advective flux is determined by the velocity field, which is oriented according to trajectories of fluid particles. This orientation is the same regardless of the tracer. In contrast, a skew tracer flux is directed along lines of constant tracer; i.e., it is neither upgradient nor downgradient. Hence, orientation of the skew flux is directly tied to the tracer field, with each tracer yielding a generally distinct flux orientation. The distinct orientations of the advective and skew fluxes can cause confusion. We explore features of these geometric distinctions in Section 17.4 in studying eddy induced tracer fluxes.

- MATERIAL FLUX: Fluid elements carry a particular amount of trace matter so that an advective flux of a material tracer measures the passage of matter across an area per unit time (dimensions of mass per area per time). In contrast, a skew flux is not interpreted as the passage of matter across an area per time. This distinction is particularly important when deriving boundary conditions discussed in Section 16.5.3.

In Section 16.5 we pursue the above points to further reveal the dual relation between advective fluxes and skew fluxes.

#### 16.4.6 A comment about skew fluxes and Lagrangian kinematics

The advective tracer flux and skew tracer flux are very distinct vectors and we further explore the distinction in Sections 16.5 and 17.4. As detailed in each of those sections, it is a matter of convenience how one chooses to mathematically formulate the Eulerian tracer equation since the advective flux and skew flux lead to the same tracer evolution. Furthermore, the choice to formulate the tracer equation in terms of a skew flux in no way eliminates the Lagrangian fluid particle perspective. That is, we still conceive of fluid particles as moving through the fluid transporting tracer as part of this motion, regardless of how we choose to mathematically describe the consequences of that transport.

#### 16.4.7 Further reading

The uses of residual-mean transport are many and varied in the ocean and atmospheric literature. [Vallis \(2017\)](#) offers a thorough and pedagogical treatment. Skew diffusion is treated in [Moffatt \(1983\)](#), in which he raises the connection to rotating and/or magnetic fluids. [Middleton and Loder \(1989\)](#) applied these ideas to ocean gravity waves, tides, and Rossby waves. [Griffies \(1998\)](#) applied these ideas to the methods used for parameterizing tracer transport from ocean mesoscale eddies.

### 16.5 Advection and skewson

We introduced skew diffusion in Section 16.4.5 and will again encounter it in Chapters 17 and 18. Following the terminology of Section 9.2 of [Griffies \(2004\)](#), we refer to skewson as any process that leads to tracer transport via skew fluxes, with skew diffusion a particular example. There are occasions where it is conceptually and operationally more convenient to use advective fluxes, such as when considering the transport of tracers by the flow field explicitly resolved by a numerical simulation. In contrast, skew fluxes are sometimes more convenient for certain subgrid scale eddy parameterizations, such as the one discussed in Section 18.1. We here consider facets of advection and skewson for those interested in diving deeper into the mathematical physics.

#### 16.5.1 Choosing a gauge

Consider an arbitrary divergent-free mass transport

$$\nabla \cdot (\rho \mathbf{v}^*) = 0, \quad (16.48)$$

where the divergent-free constraint is satisfied by introducing a vector streamfunction

$$\rho \mathbf{v}^* = \nabla \times (\rho \Psi^*). \quad (16.49)$$

The streamfunction is arbitrary up to a gauge transformation

$$\rho \Psi' = \rho \Psi^* + \nabla(\rho \Lambda), \quad (16.50)$$

where  $\Lambda$  is a gauge function.

### Changes to the skew flux under a gauge transformation

Although the velocity is invariant up to an arbitrary gauge function, the skew flux,  $\mathbf{J}^{\text{skew}} = -\nabla C \times \rho \Psi^*$ , changes. Nonetheless, the divergence of the skew flux is invariant, as we see by noting that

$$\nabla C \times [\rho \Psi^* + \nabla(\rho \Lambda)] = \nabla C \times (\rho \Psi^*) + \nabla \times [C \nabla(\rho \Lambda)]. \quad (16.51)$$

and since  $\nabla \cdot \nabla \times [C \nabla(\rho \Lambda)] = 0$ , the flux divergence,  $\nabla \cdot \mathbf{J}^{\text{skew}}$ , remains unchanged.

### Coulomb gauge

We have some freedom in specifying the gauge function. One choice is to set  $\Lambda = 0$ . However, there are occasions in which it is useful to set the gauge function in a manner to cancel unwanted terms. The Coulomb gauge is commonly used in electrodynamics (e.g., [Jackson \(1975\)](#); [Griffiths \(1981\)](#)), which is defined by setting

$$\nabla \cdot (\rho \Psi^*) = 0. \quad (16.52)$$

Making use of the vector calculus identity (see VOLUME 1 for a proof)

$$\nabla \times (\nabla \times \mathbf{F}) = -\nabla^2 \mathbf{F} + \nabla(\nabla \cdot \mathbf{F}), \quad (16.53)$$

leads to Poisson's equation for the vector potential

$$\nabla^2(\rho \Psi^*) = -\nabla \times (\rho \mathbf{v}^*). \quad (16.54)$$

In the absence of boundaries, this equation has a Coulomb-Ampere solution comprised of the convolution of the source with the free-space Green's function<sup>6</sup>

$$\rho(\mathbf{x}, t) \Psi^*(\mathbf{x}, t) = \int \frac{\nabla \times [\rho(\mathbf{x}', t) \mathbf{v}^*(\mathbf{x}', t)]}{4 \pi |\mathbf{x} - \mathbf{x}'|} dV', \quad (16.55)$$

where  $dV'$  is the volume element for integration over the test points,  $\mathbf{x}'$ . We know of no geophysical fluid application making use of the Coulomb gauge, perhaps because it can be difficult to determine an expression for  $\rho \Psi^*$  in the presence of boundaries.

### 16.5.2 Vertical gauge

As introduced in VOLUME 1 as part of the kinematics in non-divergent flows, a gauge commonly used for eddy parameterizations (Section 18.1) sets to zero one of the three components of the vector streamfunction. This gauge choice is available since there are only two independent functional degrees of freedom available from a divergence-free vector. A common choice is the

<sup>6</sup>See the chapter in VOLUME 1 that discusses elliptic partial differential equation and their corresponding Green's function solutions.

vertical gauge in which

$$\Psi_3^* = 0. \quad (16.56)$$

To further specify the vertical gauge we invert the relations

$$\rho u^* = -\partial_z(\rho\Psi_2^*) \quad \text{and} \quad \rho v^* = \partial_z(\rho\Psi_1^*) \quad \text{and} \quad \rho w^* = \partial_x(\rho\Psi_2^*) - \partial_y(\rho\Psi_1^*), \quad (16.57)$$

to render the vector streamfunction

$$\rho \Psi^* = \hat{z} \times \int_{-H}^z \rho \mathbf{u}^* dz' = \hat{z} \times \underline{\mathbf{U}}^{(*\rho)} \quad (16.58)$$

where

$$\underline{\mathbf{U}}^{(*\rho)}(z) = \int_{\eta_b}^z \rho \mathbf{u}^* dz' \quad (16.59)$$

is the horizontal mass transport associated with  $\mathbf{u}^*$  passing between the bottom and a vertical position,  $z \geq \eta_b$ . The anti-symmetric stirring tensor for the vertical gauge is given by

$$\rho A^{mn} = \begin{pmatrix} 0 & 0 & \underline{U}^{(*\rho)} \\ 0 & 0 & \underline{V}^{(*\rho)} \\ -\underline{U}^{(*\rho)} & -\underline{V}^{(*\rho)} & 0 \end{pmatrix}, \quad (16.60)$$

and the corresponding skew, rotational, and advective fluxes are

$$\mathbf{J}^{\text{skew}} = -\underline{\mathbf{U}}^{(*\rho)} \partial_z C + \hat{z} \underline{\mathbf{U}}^{(*\rho)} \cdot \nabla_h C \quad (16.61a)$$

$$\mathbf{J}^{\text{rot}} = \partial_z(C \underline{\mathbf{U}}^{(*\rho)}) - \hat{z} \nabla_h \cdot (C \underline{\mathbf{U}}^{(*\rho)}) \quad (16.61b)$$

$$\mathbf{J}^{\text{adv}} = C(\partial_z \underline{\mathbf{U}}^{(*\rho)}) - \hat{z} C \nabla_h \cdot \underline{\mathbf{U}}^{(*\rho)}. \quad (16.61c)$$

Note that the identity  $\mathbf{J}^{\text{adv}} = \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}$  is manifest in these expressions. The horizontal components to the skew flux vanish when the tracer is uniform in the vertical, and the vertical skew flux vanishes with a horizontally uniform tracer field. These properties manifest the skewed nature of the fluxes.

### 16.5.3 Boundary conditions

We assume that all external domain boundaries are material in regards to the eddy-induced velocity,  $\mathbf{v}^*$ . Furthermore, even for moving domain boundaries, we assume that the suite of kinematic boundary conditions is based on the barycentric velocity,  $\mathbf{v}$ , introduced in VOLUME 1. Consequently,  $\mathbf{v}^*$  satisfies the no-normal flow condition even on moving boundaries

$$\hat{\mathbf{n}} \cdot \mathbf{v}^* = 0 \quad \text{external domain boundaries.} \quad (16.62)$$

As we discuss in Section 16.6.1, this boundary condition is required for the eddy-induced velocity to have zero impact on the total mass of an arbitrary tracer within the fluid domain.

Correspondingly, the advective tracer flux also satisfies a no-normal boundary condition on all external boundaries

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{adv}} = \hat{\mathbf{n}} \cdot \mathbf{v}^* \rho C = 0. \quad (16.63)$$

The corresponding boundary condition for the skew flux is found by inserting the relation

(16.43) into the advective flux boundary condition (16.63) to render

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{adv}} = \hat{\mathbf{n}} \cdot [\mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}] = 0. \quad (16.64)$$

Hence, the skew flux generally has a non-zero normal component at the solid boundaries as determined by the rotational flux

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{skew}} = -\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{rot}}. \quad (16.65)$$

Even so, there might be occasions in which  $\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{skew}} = 0$ , which is ensured so long as

$$(-\nabla C \times \rho \Psi^*) \cdot \hat{\mathbf{n}} = -(\rho \Psi^* \times \hat{\mathbf{n}}) \cdot \nabla C = 0. \quad (16.66)$$

A sufficient condition is to have  $\Psi^* \times \hat{\mathbf{n}} = 0$ , in which case the vector streamfunction is parallel to the boundary normal. An alternative sufficient condition is to have the streamfunction vanish at the boundary. Further details for boundary conditions depend on physical properties of the velocity  $\mathbf{v}^*$ . We discuss one example in Section 18.1 as prescribed by the [Gent et al. \(1995\)](#) mesoscale eddy parameterization.

## 16.6 Finite volume budgets with eddy velocities

In this section we examine how an [eddy-induced velocity](#) modifies the budgets for fluid mass and tracer mass in finite domains. We start by writing the local/differential mass and tracer budgets in the form

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}^\dagger) = 0 \quad (16.67a)$$

$$\partial_t (\rho C) + \nabla \cdot (\rho \mathbf{v}^\dagger C + \mathbf{J}^{\text{diff}}) = 0, \quad (16.67b)$$

where (see Section 16.5)

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^* \quad \text{and} \quad \nabla \cdot (\rho \mathbf{v}^*) = 0, \quad (16.68)$$

and where  $\mathbf{J}^{\text{diff}}$  is a subgrid scale flux encompassing all processes, such as diffusion and boundary conditions, that are not represented by an eddy-induced advection. Given that  $\nabla \cdot (\rho \mathbf{v}^*) = 0$  in the fluid interior and  $\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0$  along all boundaries (including moving boundaries), the mass budget (16.67a) can be written in the equivalent manners

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}^\dagger) = \partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (16.69)$$

That is, the eddy-induced velocity does not lead to any local or global sources of fluid mass. This property is central to the budget analysis in this section.

As shown in the following, the finite volume budgets for fluid mass and tracer mass also make use of the residual mean velocity,  $\mathbf{v}^\dagger$ . That result is not surprising, since the finite volume budgets are consistent with the differential budgets (16.67a) and (16.67b). Nonetheless, it is useful to expose the details as they appear in many budget analysis applications, such as the water mass and tracer mass analysis of Chapter 20. We furthermore explore how the budgets for tracer mass appear when formulated using advective fluxes versus skew fluxes. As we show, the finite volume budgets are consistent across the variety of formulations only if the eddy velocity and eddy vector streamfunction satisfy boundary conditions detailed in Section 16.5.3.

### 16.6.1 Advection flux formulation

Making use of the tracer equation (16.67b) in the Leibniz-Reynolds transport theorem (see VOLUME 1) renders the finite volume tracer mass budget for an arbitrary domain,  $\mathcal{R}$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho C dV \right] = - \oint_{\partial\mathcal{R}} \left[ \rho C (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) + \mathbf{J}^{\text{diff}} \right] \cdot \hat{\mathbf{n}} dS, \quad (16.70)$$

where  $\mathbf{v}^{(b)}$  is the velocity of a point on the domain boundary. Appearance of the residual mean velocity,  $\mathbf{v}^\dagger$ , in the finite volume budget (16.70) follows from its appearance in the local tracer budget (16.67b). We thus see that the eddy-induced velocity impacts on the tracer mass budget for an arbitrary domain. However, its impacts disappear when integrating over a closed or periodic fluid domain so long as

$$\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0 \quad \text{on all boundaries.} \quad (16.71)$$

We already encountered this boundary condition in Section 16.5.3. It holds on all boundaries, including those such as the ocean free surface that are time dependent and/or permeable. It is required if we assume the eddy-induced velocity does not modify the mass of any tracer in the full fluid domain. That assumption is generally made for eddy-induced velocities such as those associated with mesoscale and submesoscale eddies in the ocean (see Section 18.3).

Setting the tracer concentration to a constant in equation (16.70) leads to the fluid mass budget

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho dV \right] = - \oint_{\partial\mathcal{R}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS, \quad (16.72)$$

where we set the diffusive tracer flux,  $\mathbf{J}^{\text{diff}}$ , to zero since there is no diffusion of fluid mass between fluid elements.<sup>7</sup> As for the differential expression (16.67a), the mass budget for any domain is not changed by the eddy-induced velocity since

$$\nabla \cdot (\rho \mathbf{v}^*) = 0 \implies \oint_{\partial\mathcal{R}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} dS = 0, \quad (16.73)$$

so that the mass budget is given by

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho dV \right] = - \oint_{\partial\mathcal{R}} \left[ \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \right] \cdot \hat{\mathbf{n}} dS = - \oint_{\partial\mathcal{R}} [\rho (\mathbf{v} - \mathbf{v}^{(b)})] \cdot \hat{\mathbf{n}} dS \quad (16.74)$$

Hence, the eddy velocity contribution to the mass budget for any finite region vanishes, which is expected since it provides no local nor finite volume mass source to the fluid. Furthermore, one may choose to diagnose the right hand side of the mass budget in either the residual mean or Eulerian mean form. The choice is based on convenience, such as whether one has easier access to the residual mean velocity or the Eulerian mean velocity. Although the patterns of the fluxes across any particular boundary differs if  $\mathbf{v}^* \neq 0$ , the accumulation of mass within the region is identical for the two formulations.

---

<sup>7</sup>We emphasized this property when discussing the **barycentric velocity** as part of material tracer conservation in VOLUME 1.

### 16.6.2 Skew flux formulation

Now consider the perspective afforded by the skew flux formulation from Section 16.5. Here we decompose the advective tracer flux according to

$$C \rho \mathbf{v}^\dagger = C \rho \mathbf{v} - \nabla C \times \rho \Psi^* + \nabla \times (C \rho \Psi^*) = C \rho \mathbf{v} + \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}, \quad (16.75)$$

where we introduced the skew tracer flux and rotational flux arising from the eddy-induced streamfunction

$$\mathbf{J}^{\text{skew}} = -\nabla C \times (\rho \Psi^*) \quad \text{and} \quad \mathbf{J}^{\text{rot}} = \nabla \times (\rho C \Psi^*). \quad (16.76)$$

The flux-form tracer equation is thus given by

$$\partial_t(\rho C) + \nabla \cdot [\rho C \mathbf{v} + \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{diff}}] = 0, \quad (16.77)$$

where  $\nabla \cdot \mathbf{J}^{\text{rot}} = 0$  and so  $\mathbf{J}^{\text{rot}}$  does not affect the tracer budget. The corresponding finite volume tracer mass budget is

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho C dV \right] = - \oint_{\partial \mathcal{R}} [\rho C (\mathbf{v} - \mathbf{v}^{(b)}) - \nabla C \times (\rho \Psi^*) + \mathbf{J}^{\text{diff}}] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.78)$$

In this form, the contribution from the eddy induced transport is included inside the skew tracer flux rather than in the residual mean advective tracer flux. Setting  $C$  to a constant reveals the mass budget as in the second form of equation (16.74).

### 16.6.3 Domain with a tracer boundary

We now apply the previous general budget discussion to a specific domain that anticipates the more complete budget analysis provided in Section 20.9 as part of our study of water mass analysis. Here, as in Figure 16.4, we consider the fluid mass and tracer mass within an ocean region with at least one of its bounds determined by an isosurface of constant tracer concentration.

#### Advective formulation

The tracer mass budget written using the advective formulation (16.70) is given by

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho C dV \right] = \int_{\partial \Omega_{\text{surf}}(\tilde{C})} Q_m C dA - \tilde{C} \int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} - \oint_{\partial \mathcal{R}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.79)$$

For the first right hand side term we made use of the surface kinematic boundary condition from our discussion of mass budgets in VOLUME 1, with  $Q_m dA$  the mass per time crossing the surface interface and where  $dA$  is the horizontal projection of the interface area element,  $d\mathcal{S}$ . We also made use of the exterior boundary condition (16.71) for the eddy-induced velocity. For the second term, we pulled the tracer concentration outside of the boundary integral over the  $C = \tilde{C}$  interface, since the concentration is fixed at  $\tilde{C}$  on this interface.

The mass budget for this region, also formulated using advective fluxes, is given by

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho dV \right] = \int_{\partial \Omega_{\text{surf}}(\tilde{C})} Q_m dA - \int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.80)$$

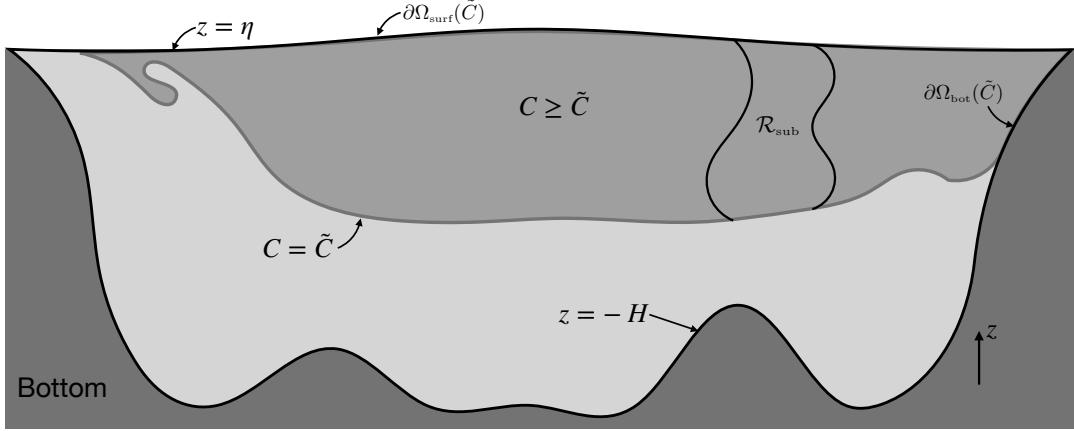


FIGURE 16.4: An ocean region with one of its boundaries set by a surface of constant tracer,  $C = \tilde{C}$ . Note that the region need not be monotonically stratified in the tracer concentration, nor does it need to be simply connected. The region is bounded at the top by  $\partial\Omega_{\text{surf}}(\tilde{C})$ , with the geometry of this surface depending on  $\tilde{C}$ . The bottom boundary is set by the tracer surface,  $C = \tilde{C}$ , as well as the solid-earth bottom,  $\partial\Omega_{\text{bot}}(\tilde{C})$ . The region can generally be multiply connected. A subregion,  $\mathcal{R}_{\text{sub}}$ , is also considered where its sides extend from the free surface to the tracer isosurface, and they are assumed to be fully within the fluid domain. We develop the tracer and fluid mass budgets for region  $\mathcal{R}_{\text{sub}}$  in Section 16.6.4, whereas the budget in the full region  $C \geq \tilde{C}$  is considered in Section 16.6.3.

Combining this budget with the tracer mass budget allows us to write

$$\frac{d}{dt} [M_C - \tilde{C} M] = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m (C - \tilde{C}) dA - \oint_{\partial\mathcal{R}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (16.81)$$

where we introduced the shorthand for the tracer mass and fluid mass in the region

$$M_C = \int_{\mathcal{R}} C \rho dV \quad \text{and} \quad M = \int_{\mathcal{R}} \rho dV. \quad (16.82)$$

In Section 20.9.2 we motivate the name *internal tracer mass* for the quantity  $M_C - \tilde{C} M$ .

### Skew flux formulation

The tracer mass budget formulated using skew tracer fluxes is generally given by equation (16.78). It takes on the following specific form for the domain in Figure 16.4

$$\begin{aligned} \frac{d}{dt} \left[ \int_{\mathcal{R}} \rho C dV \right] &= \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m C dA - \tilde{C} \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} \\ &\quad - \oint_{\partial\mathcal{R}} [-\nabla C \times (\rho \Psi^*) + \mathbf{J}^{\text{diff}}] \cdot \hat{\mathbf{n}} d\mathcal{S}, \end{aligned} \quad (16.83)$$

and the corresponding budget for the fluid mass is

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho dV \right] = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m dA - \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.84)$$

As for the advective formulation, we combine the fluid mass budget equation (16.84) with the tracer mass equation (16.83) to render a budget equation for the internal mass content

$$\frac{d}{dt} \left[ M_C - \tilde{C} M \right] = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m (C - \tilde{C}) dA - \oint_{\partial\mathcal{R}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (16.85)$$

which is identical to the advective formulation given by equation (16.81).

### Proving the budgets based on the two formulations are equivalent

The two tracer budgets, (16.79) and (16.83), must be the same since they measure changes to the tracer mass within the same region. Likewise, the two mass budgets, (16.80) and (16.84), must be the same, as are the two internal tracer mass budgets (16.81) and (16.85). We here expose the manipulations required to verify these equalities.

To prove the  $C = \tilde{C}$  terms in the tracer budget equations (16.79) and (16.83) are the same, consider the identity (16.73) applied to the region under consideration

$$0 = \oint_{\partial\mathcal{R}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\partial\Omega_{\text{bot}}(\tilde{C})} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{C=\tilde{C}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.86)$$

The surface and bottom boundary terms vanish due to the external boundary condition (16.71); namely,  $\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0$  for each point along an external fluid boundary. We are thus led to conclude that

$$\int_{C=\tilde{C}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} = 0. \quad (16.87)$$

This boundary integral means that there is no net accumulation of mass in the region due to action of the eddy velocity. Notably, we generally have  $\mathbf{v}^* \cdot \hat{\mathbf{n}} \neq 0$  at any particular point on the  $C = \tilde{C}$  surface, yet its integral over the  $C = \tilde{C}$  interface vanishes. Given the boundary integral (16.87) we are led to conclude

$$\int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.88)$$

This identity proves that the two mass budgets (16.80) and (16.84) are indeed measuring changes to the same fluid mass, even though one computes the domain boundary fluxes based on the residual mean velocity,  $\mathbf{v}^\dagger$ , whereas the other uses the Eulerian mean,  $\mathbf{v}$ .

Next we need to show that the skew flux term vanishes when integrated around the domain boundary. For the  $C = \tilde{C}$  boundary we have

$$\int_{C=\tilde{C}} [\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S} = 0, \quad (16.89)$$

which follows since  $\hat{\mathbf{n}}$  is parallel to  $\nabla C$  along this boundary. For the external boundaries, equality of the tracer mass budgets (16.79) and (16.83) is satisfied for an arbitrary tracer concentration if one of the boundary conditions discussed in Section (16.5.3) is maintained; i.e., if  $\Psi^*$  vanishes on an external boundary or if it is parallel to the boundary normal direction ( $\hat{\mathbf{n}} \times \Psi^* = 0$ ). Maintenance of either of these two boundary conditions allows us to conclude that the two budgets (16.79) and (16.83) are indeed identical.

### 16.6.4 Budget for a region with interior sides

Consider the subregion,  $\mathcal{R}_{\text{sub}}$ , shown in Figure 16.4. This region is bounded above by the free surface and below by the tracer isosurface,  $C = \tilde{C}$ , just like the region  $\mathcal{R}$  encountered in Section 16.6.3. Additionally, region  $\mathcal{R}_{\text{sub}}$  is bounded along its sides by surfaces assumed to be within the fluid interior. For much of this discussion we allow the sides to have an arbitrary shape and to move. Towards the end of this section we specialize to the case of static sides, such as relevant for a vertical section through the fluid and/or a numerical model grid cell.

#### Fluid mass budget

The fluid mass budget for the region,  $\mathcal{R}_{\text{sub}}$ , can be formulated using either the residual mean velocity or the Eulerian mean velocity

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m dA - \int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (16.90a)$$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m dA - \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.90b)$$

The two budgets are identical since the eddy velocity satisfies  $\oint \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} = 0$  for any domain, as well as  $\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0$  along any external domain boundary. Hence, as already noted in Section 16.6.1, the eddy velocity contribution to the right hand side of equation (16.90a) vanishes; it provides no net mass source to any region. We next show the same equality holds for the tracer mass budgets, with that equality requiring a bit more effort.

#### Tracer mass budget

The advective flux formulation of the tracer mass budget is given by

$$\begin{aligned} \frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] &= \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m C dA - \tilde{C} \int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} \\ &\quad - \oint_{\partial\mathcal{R}_{\text{sub}}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} C \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}, \end{aligned} \quad (16.91)$$

and the corresponding skew flux formulation is

$$\begin{aligned} \frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] &= \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m C dA - \tilde{C} \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} \\ &\quad - \oint_{\partial\mathcal{R}_{\text{sub}}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} C \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S}. \end{aligned} \quad (16.92)$$

As for the discussion in Section 16.6.3, we introduce the internal tracer mass and make use of the fluid mass budgets (16.90a) and (16.90b) to write the advective form of the internal mass budget

$$\frac{d}{dt} [M_C - \tilde{C} M] = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m (C - \tilde{C}) dA - \oint_{\partial\mathcal{R}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S}$$

$$-\int_{\text{sides}} (C - \tilde{C}) \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (16.93)$$

and the corresponding skew flux form of the same budget

$$\begin{aligned} \frac{d}{dt} [M_C - \tilde{C} M] &= \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m (C - \tilde{C}) dA - \oint_{\partial\mathcal{R}} \mathbf{J}^{\text{diff}} \cdot \hat{\mathbf{n}} d\mathcal{S} \\ &\quad - \int_{\text{sides}} (C - \tilde{C}) \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S}. \end{aligned} \quad (16.94)$$

We now examine the right hand side of the budgets (16.93) and (16.94) to show they are indeed measuring the same tracer mass budget. For that purpose, consider the skew flux integral in equation (16.94) and note that the integrand vanishes on both the top of the domain, at  $z = \eta$ , and bottom at  $C = \tilde{C}$ , thus allowing us to write

$$-\int_{\text{sides}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S} = -\oint_{\partial\mathcal{R}_{\text{sub}}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (16.95)$$

where the right hand side is an integral around the full domain boundaries. Now reintroduce the eddy induced velocity and rotational flux to have

$$-\oint_{\partial\mathcal{R}_{\text{sub}}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S} = -\oint_{\partial\mathcal{R}_{\text{sub}}} [C \rho \mathbf{v}^* - \nabla \times (C \rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.96)$$

The rotational flux has zero divergence, so that Gauss's divergence theorem means that the rotational flux vanishes when integrated along the domain boundaries

$$\oint_{\partial\mathcal{R}_{\text{sub}}} \nabla \times (C \rho \Psi^*) \cdot \hat{\mathbf{n}} d\mathcal{S} = 0. \quad (16.97)$$

The eddy advection term in equation (16.96) vanishes on the top boundary at  $z = \eta$  due to the boundary condition  $\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0$ , thus yielding

$$\oint_{\partial\mathcal{R}_{\text{sub}}} C \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\text{sides}} C \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{C=\tilde{C}} C \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (16.98a)$$

$$= \int_{\text{sides}} C \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} + \tilde{C} \int_{C=\tilde{C}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.98b)$$

Again make use of the property  $\nabla \cdot (\rho \mathbf{v}^*) = 0$  and  $\hat{\mathbf{n}} \cdot \mathbf{v}^* = 0$  at  $z = \eta$  to write

$$0 = \oint_{\partial\mathcal{R}_{\text{sub}}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} = \oint_{\text{sides}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} + \oint_{C=\tilde{C}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.99)$$

We are thus led to

$$-\int_{\text{sides}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S} = -\oint_{\partial\mathcal{R}_{\text{sub}}} C \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (16.100a)$$

$$= -\int_{\text{sides}} (C - \tilde{C}) \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.100b)$$

Making use of this result in the skew flux formulated budget equation (16.94) brings it to the advective flux form found in equation (16.93).

We are thus led to conclude that the right hand side to equation (16.94) does indeed equal to the right hand side of equation (16.93). So although the formulation of the boundary flux contributions is rather distinct between the advective flux and skew flux formulations, the resulting tracer mass budget is the same. The choice for how to formulate the budget is thus a matter of convenience.

### 16.6.5 Budget for a perfect fluid in a region with interior sides

Although contained within the formalism developed in Section 16.6.4, it is revealing to specialize those budgets to the case of zero mixing, in which  $\mathbf{J}^{\text{diff}} = 0$ , and there is zero boundary mass flux,  $Q_m = 0$ . In this case the fluid is reversibly stirred. Examining the finite region budgets for this perfect fluid allows us to further reveal the complementary treatments available from advection versus skewson.

#### Summary of the differential stirring formalism

As explored in this chapter, an Eulerian description of tracer stirring can arise from either advection or skewson. In the presence of an eddy induced velocity we consider two representations of tracer stirring, with the first being advection by the residual mean velocity,  $\mathbf{v}^\dagger$

$$\rho \frac{D^\dagger C}{Dt} = \partial_t(\rho C) + \nabla \cdot (\rho \mathbf{v}^\dagger C) = 0. \quad (16.101)$$

This formulation makes it clear that surfaces of constant  $C$  are material as defined by the residual mean velocity rather than by the Eulerian mean,  $\mathbf{v}$ . That is, tracer isosurfaces satisfy the residual mean impermeability condition

$$\rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{on } C \text{ isosurfaces}, \quad (16.102)$$

with

$$\hat{\mathbf{n}} = \frac{\nabla C}{|\nabla C|} \quad \text{and} \quad \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} = -\frac{\partial_t C}{|\nabla C|}. \quad (16.103)$$

The impermeability condition (16.102) offers a geometric interpretation of the tracer equation (16.101) following from the discussion of dia-surface transport in Section 13.4. Correspondingly, Lagrangian fluid particles moving with the residual mean velocity do not cross tracer isosurfaces even if those isosurfaces move. Furthermore, we observe that the eddy induced velocity has a nonzero projection across tracer isosurfaces

$$(\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0 \implies (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = -\mathbf{v}^* \cdot \hat{\mathbf{n}} \quad \text{on } C \text{ isosurfaces}. \quad (16.104)$$

This property of the eddy induced velocity was emphasized by [McDougall and McIntosh \(2001\)](#). It reveals that in the absence of mixing, eddy motion crossing tracer isosurfaces is exactly balanced by Eulerian motion plus surface motion, thus leaving a net zero residual mean transfer of matter across the surface. Equation (16.104) is a key kinematic property used for interpreting features of the finite volume budgets detailed below.

Our second means to represent tracer stirring makes use of advection by the Eulerian mean velocity plus skewson by the eddy induced streamfunction

$$\rho \frac{DC}{Dt} + \nabla \cdot [-\nabla C \times (\rho \Psi^*)] = \partial_t(\rho C) + \nabla \cdot [\rho \mathbf{v} C - \nabla C \times (\rho \Psi^*)] = 0. \quad (16.105)$$

In terms of the eddy streamfunction,  $\rho \Psi^*$ , the impermeability condition (16.102) takes on the form

$$[\rho \mathbf{v} + \nabla \times (\rho \Psi^*) - \rho \mathbf{v}^{(b)}] \cdot \hat{\mathbf{n}} = 0 \quad \text{on } C \text{ isosurfaces.} \quad (16.106)$$

### Budgets via residual mean advection

The mass budget formulated in terms of residual mean advection, and the corresponding residual mean advective flux formulation of the tracer mass budget, are given by

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = - \int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS - \int_{\text{sides}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS \quad (16.107a)$$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] = - \tilde{C} \int_{C=\tilde{C}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS - \int_{\text{sides}} C \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS. \quad (16.107b)$$

The residual mean impermeability condition (16.102) for the  $C = \tilde{C}$  surface renders a simplification to the fluid mass and tracer mass budgets

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = - \int_{\text{sides}} \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS \quad (16.108a)$$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] = - \int_{\text{sides}} C \rho (\mathbf{v}^\dagger - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS. \quad (16.108b)$$

Hence, in the residual mean formulation, the only fluxes that affect changes to the mass budgets are those that cross the side faces of the region.

### Budgets via Eulerian mean advection plus eddy skewson

The mass budget formulated in terms of Eulerian mean advection, and the corresponding tracer mass budget using eddy skewson, are given by

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = - \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS - \int_{\text{sides}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS \quad (16.109a)$$

$$\begin{aligned} \frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] = & - \tilde{C} \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS - \int_{\text{sides}} C \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS \\ & - \int_{\text{sides}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} dS. \end{aligned} \quad (16.109b)$$

We already saw in Section 16.6.4 how to bring the right hand side terms into the form realized by the residual mean advective approach. So there is no question concerning the equivalence of the advective and skew flux formulations for the tracer mass budget. Nonetheless, what is here clearly emphasized is that the skew flux approach requires us to account for Eulerian advective transport across the  $C = \tilde{C}$  isosurface, whereas for the advective flux approach the only flux in equation (16.108b) is that crossing the region side boundaries. Even so, as stated earlier, an Eulerian mean transport of tracer across the  $C = \tilde{C}$  isosurface *does not* correspond to material transport across this surface. The reason is that material transport is determined by the residual mean velocity,  $\mathbf{v}^\dagger$ , as per the residual mean impermeability conditions (16.104) and (16.106). So even though there is a contribution to the skew flux formulated budget from Eulerian transport across the  $C = \tilde{C}$  material surface, there remains zero net material crossing that surface.

### Zero Eulerian mean advection and static side walls

One further specialization serves to clearly emphasize the complementary nature of the advective and skew flux approaches. Here, we assume the sides of the region are static and the Eulerian mean velocity vanishes. With a zero Eulerian velocity, the residual mean impermeability condition (16.104) means that on the  $C = \tilde{C}$  isosurface, the normal component of the eddy-induced velocity is balanced by the boundary velocity as per the impermeability condition (16.110):

$$(\mathbf{v}^* - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{on } C \text{ isosurfaces and with } \mathbf{v} = 0. \quad (16.110)$$

When formulated using the residual mean advection, the fluid mass budget (16.108a) and tracer mass budget (16.108b) reduce in this case to

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = - \int_{\text{sides}} \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (16.111a)$$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] = - \int_{\text{sides}} C \rho \mathbf{v}^* \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (16.111b)$$

so that these budgets are only affected by eddy advection across the side boundaries. The corresponding mass budget written in terms of Eulerian mean advection (16.109a), and tracer mass budget written in terms of skew fluxes (16.109b), are given by

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho dV \right] = + \int_{C=\tilde{C}} \rho \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (16.112a)$$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}_{\text{sub}}} \rho C dV \right] = + \tilde{C} \int_{C=\tilde{C}} \rho \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{\text{sides}} [-\nabla C \times (\rho \Psi^*)] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (16.112b)$$

For the mass budget, (16.112a), there are no contributions to the side walls since they are static and the Eulerian velocity is assumed to vanish. The only contribution comes from the eddy term acting on the  $C = \tilde{C}$  isosurface where  $\mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} = \mathbf{v}^* \cdot \hat{\mathbf{n}}$ . For the tracer mass budget (16.112b), we also have the eddy contribution on the  $C = \tilde{C}$  isosurface, plus skew fluxes that penetrate the side walls.

The right hand sides to the fluid mass budgets (16.111a) and (16.112a), and tracer mass budgets (16.111b) and (16.112b), are remarkably distinct. Even so, they both measure the same budgets. Furthermore, in both cases the  $C = \tilde{C}$  boundary is a material boundary as defined by the residual mean velocity.

## 16.7 Active tracers and dia-surface flow

An **active tracer** impacts the fluid flow through its impacts on buoyancy, which in turn affects pressure and velocity. Hence, the advection-diffusion equation for active tracers is nonlinear since the velocity field is dependent on active tracers. We here write the advection-diffusion equation in terms of the residual mean velocity using **Conservative Temperature** as an example active tracer

$$\rho \frac{D^\dagger \Theta}{Dt} = \rho (\partial_t + \mathbf{v}^\dagger \cdot \nabla) \Theta = -\nabla \cdot \mathbf{J}^{\text{diff}}(\Theta). \quad (16.113)$$

Further nonlinearities arise when the subgrid scale diffusion tensor is itself a function of the buoyancy, as discussed at the end of Section 15.9, and/or when the parameterized eddy-induced velocity is a function of the buoyancy, as discussed in Section 18.3.

### 16.7.1 Adiabatic flow

Conservative Temperature is materially invariant (as defined by the residual mean velocity) in an adiabatic flow

$$\frac{D^\dagger \Theta}{Dt} = (\partial_t + \mathbf{v}^\dagger \cdot \nabla) \Theta = 0. \quad (16.114)$$

Furthermore, following the mass kinematics in VOLUME 1, the adiabatic residual mean flow field does not penetrate surfaces of constant Conservative Temperature ( $\Theta$ -isosurfaces are impermeable) since

$$\mathbf{v}^\dagger \cdot \nabla \Theta = -\partial_t \Theta. \quad (16.115)$$

In this case we say that residual mean advection reversibly stirs the Conservative Temperature field. This property of the residual mean velocity was also considered in the discussion of pure stirring in Section 16.6.4.

### 16.7.2 Diabatic processes generating dia- $\Theta$ transport

Conservative Temperature is not materially invariant in the presence of diabatic processes such as mixing. Correspondingly, the residual mean velocity picks up a dia-surface transport velocity,  $w^{\text{dia}}$ , that crosses the moving Conservative Temperature surface, thus making  $\Theta$  surfaces permeable to fluid flow. In turn, advective transport in the presence of mixing is not reversible. Following the kinematics from Section 13.4, we have the expression (13.26) for  $w^{\text{dia}}$  written as

$$w^{\text{dia}} \equiv \hat{\mathbf{n}} \cdot (\mathbf{v}^\dagger - \mathbf{v}^{(\Theta)}) = \frac{1}{|\nabla \Theta|} \frac{D^\dagger \Theta}{Dt} \quad (16.116)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla \Theta}{|\nabla \Theta|} \quad \text{and} \quad \mathbf{v}^{(\Theta)} \cdot \nabla \Theta = -\partial_t \Theta. \quad (16.117)$$

Rearrangement of equation (16.116) renders the kinematic identity

$$\frac{D^\dagger \Theta}{Dt} = \partial_t \Theta + \mathbf{v}^\dagger \cdot \nabla \Theta = w^{\text{dia}} |\nabla \Theta|. \quad (16.118)$$

With nonzero  $w^{\text{dia}}$ , we no longer have residual mean advection preserving properties along fluid element trajectories. There can be many physical processes contributing to a nonzero  $w^{\text{dia}}$ , most notably mixing as in the following examples.

#### Diffusion with no fluid motion

Diffusion is the canonical example of a diabatic process, with molecular diffusion leading to

$$\rho \frac{D^\dagger \Theta}{Dt} = \nabla \cdot (\kappa \rho \nabla \Theta), \quad (16.119)$$

with  $\kappa > 0$  the scalar kinematic diffusivity and the product,  $\kappa \rho$ , the dynamic diffusivity. Following the definition (16.116), we see that diffusion drives the following diabatic transport velocity

$$\rho w^{\text{dia}} = \frac{\nabla \cdot (\kappa \rho \nabla \Theta)}{|\nabla \Theta|}. \quad (16.120)$$

Consider a horizontally homogeneous Conservative Temperature field. If buoyancy is alone determined by  $\Theta$ , then there is no fluid motion since buoyancy surfaces are flat (and we assume

the eddy-induced motion is also zero). Yet in the presence of vertical diffusion and vertical stratification there is a diabatic transport since

$$\rho w^{\text{dia}} = \frac{\partial_z(\kappa \rho \partial_z \Theta)}{|\partial_z \Theta|} \neq 0. \quad (16.121)$$

In the absence of fluid flow, the dia-surface transport is determined solely by movement of the  $\Theta$  surfaces. Correspondingly,  $\Theta$  evolution is determined only by vertical diffusion since with  $\mathbf{v}^\dagger = 0$  we have

$$\partial_t \Theta = \rho^{-1} \partial_z(\kappa \rho \partial_z \Theta) = w^{\text{dia}} |\partial_z \Theta|. \quad (16.122)$$

### Steady state advective-diabatic balance

A steady state Conservative Temperature field in the presence of diabatic processes is realized when there is an exact balance between advective transport and dia-surface transport enabled by diffusion

$$\rho \mathbf{v}^\dagger \cdot \nabla \Theta = \rho w^{\text{dia}} |\nabla \Theta| = \nabla \cdot (\kappa \rho \nabla \Theta). \quad (16.123)$$

That is, maintaining static  $\Theta$ -surfaces ( $\partial_t \Theta = 0$ ) requires the residual mean advective transport to cross  $\Theta$  surfaces (left hand side) by an amount that exactly balances diabatic processes such as diffusion (right hand side).

## 16.8 Tracer homogenization inside closed tracer contours

In this section we prove a theorem involving the steady advective-diffusive balance that holds within regions on a surface that are bounded by tracer contours. For this purpose, consider the equation for a conservative tracer

$$\partial_t(\rho C) + \nabla \cdot (\rho \mathbf{v} C) = -\nabla \cdot \mathbf{J} \quad (16.124)$$

where

$$\mathbf{J} = -\rho \mathbf{K} \cdot \nabla C \quad (16.125)$$

is a downgradient diffusive flux with  $\mathbf{K}$  a symmetric positive-definite diffusion tensor. In the steady state, the divergence of the advective tracer flux balances the convergence of the diffusive flux

$$\nabla \cdot (\rho \mathbf{v} C) = -\nabla \cdot \mathbf{J}. \quad (16.126)$$

We use this identity to prove that the tracer concentration is homogeneous (i.e., a spatially constant) within a region bounded by a contour of constant  $C$ , such as shown in Figure 16.5. Evidently, in the steady state, diffusion removes all tracer variations within closed tracer contours; i.e., there are no tracer extrema within a closed tracer contour.

### 16.8.1 Proof of the theorem

The following proof follows that given in Section 3.2 of [Rhines and Young \(1982\)](#) and Section 13.5 of [Vallis \(2017\)](#), both given for flow in a Boussinesq ocean. We slightly extend their proof by working with a non-Boussinesq fluid. The proof is based on a *reductio ad absurdum* argument, whereby we first assume the tracer is not homogeneous within a closed tracer contour, and then show that this assumption leads to an inconsistency and so is wrong. Notably, if the tracer

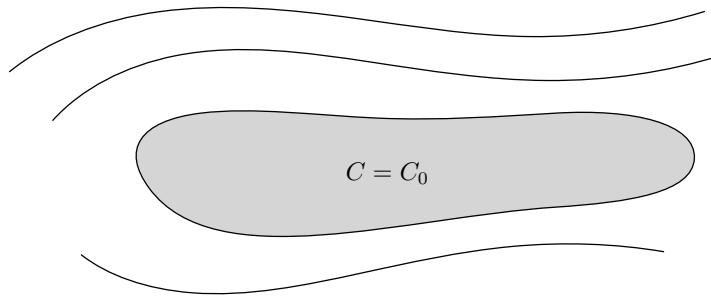


FIGURE 16.5: In a steady flow, the tracer concentration within a region bounded by a constant tracer contour is uniform. Diffusion provides the mechanism for homogenizing the tracer. Evidently, diffusion expels all variations in tracer concentration from the region. In this figure, the concentration within the closed region has constant value,  $C = C_0$ . Note that the region need not be horizontal. For example, it could be an undulating constant buoyancy surface, or it could be the surface of a sphere. The only condition is that the region is bounded by a closed tracer contour.

concentration is not homogeneous within a closed tracer contour, then there is necessarily an extrema within that contour.

To start the proof, integrate the left hand side of the steady state advection-diffusion equation (16.126) over an arbitrary simply connected surface (not necessarily a horizontal surface) and make use of the divergence theorem

$$\int_S \nabla \cdot (\rho v C) d\mathcal{S} = \oint_{\partial S} \rho v C \cdot \hat{n} dl, \quad (16.127)$$

where  $\hat{n}$  is the outward normal along the area's boundary,  $\partial S$ , and  $dl$  is the line element on the boundary. Now assume the surface is bounded by a constant tracer contour whose value is  $C = C_0$ . We can thus remove the tracer concentration from the contour integral to have

$$\oint_{\partial S} \rho v C \cdot \hat{n} dl = C_0 \oint_{\partial S} \rho v \cdot \hat{n} dl = C_0 \int_S \nabla \cdot (\rho v) d\mathcal{S}. \quad (16.128)$$

The second equality follows from the divergence theorem, only now observe that the tracer concentration is outside of the area integral. For a steady state flow, mass continuity means that density at a point is time independent, so that the density-weighted velocity has zero divergence

$$\nabla \cdot (\rho v) = 0. \quad (16.129)$$

Consequently, when integrated over a closed tracer contour we have the identity

$$\int_S \nabla \cdot (\rho v C) d\mathcal{S} = C_0 \int_S \nabla \cdot (\rho v) d\mathcal{S} = 0 \quad \text{with } S \text{ enclosed by a closed } C \text{ contour.} \quad (16.130)$$

Returning to the steady state advection-diffusion equation (16.126), the identity (16.130) then implies the analogous result for the diffusive flux

$$\int_S \nabla \cdot \mathbf{J} d\mathcal{S} = 0 \quad \text{with } S \text{ enclosed by a closed } C \text{ contour.} \quad (16.131)$$

We now provide two arguments to show that the identity (16.131) holds only if the tracer is homogeneous inside the closed contour; i.e., there are no extrema within a closed contour.

**Argument 1: Appealing to a nonzero diffusive flux in the presence of an extrema**

Consider the right hand side of the steady state advection-diffusion equation (16.126) and integrate it over a closed region

$$-\int_S \nabla \cdot \mathbf{J} dS = -\oint_{\partial S} \mathbf{J} \cdot \hat{\mathbf{n}} dl = \oint_{\partial S} (\mathbf{K} \cdot \nabla C) \cdot \hat{\mathbf{n}} \rho dl. \quad (16.132)$$

If the contour surrounds an extremum of the tracer concentration, then a downgradient diffusive flux is nonzero and has a nonzero projection in the outward normal direction. For example, consider a circular tracer contour surrounding a circular maximum, and assume isotropic diffusion. In this case the diffusive flux is radial so that it has a nonzero projection onto the outward normal. So when there is an extrema within the closed contour, the surface integral (16.132) is nonzero. But a nonzero integral of the diffusive flux around the contour contradicts equation (16.131). The inconsistency arose from assuming the existence of a tracer extrema. Dispensing with this assumption means that the tracer concentration is homogeneous inside the closed tracer contour.

**Argument 2: Considering the outward normal**

Let us again presume that the tracer is not homogeneous within the domain. In this case  $\nabla C \neq 0$  so that we can introduce the normal direction along the tracer contour,

$$\hat{\mathbf{n}} = \frac{\nabla C}{|\nabla C|}, \quad (16.133)$$

which is directed orthogonal to a constant tracer contour. Using this expression for the normal direction within the tracer flux integral leads to

$$-\int_S \nabla \cdot \mathbf{J} dS = \oint_{\partial S} \nabla C \cdot \mathbf{K} \cdot \nabla C \frac{\rho dl}{|\nabla C|}. \quad (16.134)$$

The diffusion tensor is a symmetric positive-definite second order tensor, so that the quadratic form in the integral is non-negative

$$\nabla C \cdot \mathbf{K} \cdot \nabla C = K^{mn} \partial_m C \partial_n C \geq 0. \quad (16.135)$$

Furthermore, the line element,  $dl$ , is positive, and so are  $\rho$  and  $|\nabla C|$ . Consequently, the integral around a closed tracer contour, where the normal to that contour is given by equation (16.133), is positive

$$-\int_S \nabla \cdot \mathbf{J} dS = \int_{\partial S} \partial_m C K^{mn} \partial_n C dl > 0 \quad \text{if } \hat{\mathbf{n}} = \nabla C / |\nabla C|. \quad (16.136)$$

Again, we reach a conclusion that contradicts the zero result (16.131). The zero result (16.131) is based only on the steady state assumption, whereas the inequality (16.136) is a direct result of assuming the tracer is inhomogeneous. Dropping the tracer inhomogeneity assumption is the only way to recover consistency.

### 16.8.2 Comments

Both of the above arguments lead us to conclude that the steady state tracer concentration is homogeneous within a closed tracer contour. Given enough time to reach a steady state, diffusion, even arbitrarily weak diffusion, expels all steady state tracer gradients from within regions bounded by closed tracer contours.

We presented the arguments based on integration over a surface bounded by a tracer contour. The result can be extended to integration over volumes in three dimensions, in which case steady state tracers are found to be homogeneous within closed volumes bounded by a surface of constant tracer concentration. However, such bubble-like tracer bounded regions are not common for large-scale geophysical flows. In contrast, closed tracer contours are commonly found in quasi-two dimensional flows, including flows on isopycnals. So the theorem is more readily applied to two dimensional (or quasi two dimensional) flows. We offer another proof of the tracer homogenization result in Section 20.9.1 as part of our study of tracer mass analysis.

We assumed that the tracer equation include just a symmetric diffusion tensor,  $\mathbf{K}$ , in defining the subgrid flux (16.125). However, the theorem also holds if there is an additional anti-symmetric tensor,  $\mathbf{A}$ , added to  $\mathbf{K}$ . The reason is that an anti-symmetric tensor vanishes from the quadratic form (16.135)

$$\nabla C \cdot \mathbf{A} \cdot \nabla C = A^{mn} \partial_m C \partial_n C = 0. \quad (16.137)$$

That is, the tracer skew flux,  $\mathbf{J}^{\text{skew}} = -\mathbf{A} \cdot \nabla C$ , does not cross tracer isolines:  $\nabla C \cdot \mathbf{J} = 0$ , which we already found when studying skew fluxes in Section 16.5.

### 16.8.3 Further study

A prominent application of this theorem appears when the tracer,  $C$ , is a dynamically active tracer, such as vorticity or quasi-geostrophic potential vorticity (VOLUME 3). The case of vorticity was discussed by [Batchelor \(1956\)](#), with [Rhines and Young \(1982\)](#) extending that work to the case of quasi-geostrophic potential vorticity. For the quasi-geostrophic case, contours of constant potential vorticity are known as geostrophic contours, and [Rhines and Young \(1982\)](#) used the homogenization theorem to develop a theory of ocean circulation in wind-driven gyre regions. They did so by considering potential vorticity homogenization by the mesoscale eddies that are active in regions of closed geostrophic contours, such as in mid-latitude ocean gyres.

## 16.9 Green's function method for passive tracers

A [passive tracer](#) has no impact on the fluid density nor the diffusion tensor, which means it has no impact the fluid flow. Consequently, a passive tracer provides a theoretical tool to probe the effects from advective and diffusive transport without affecting the flow. Furthermore, the advection-diffusion equation for a passive tracer is a linear partial differential equation, and as such it affords a [Green's function](#) solution. With some qualifications identified below, we here extend the Green's function method studied for the diffusion equation in VOLUME 1. There are many parallels to that discussion, so it is a very good idea to have VOLUME 1 accessible. Although our formulation is largely based on ocean applications, the Green's function method for the advection-diffusion equation is also applicable to the atmosphere.

### 16.9.1 Concerning time dependent domain boundaries

The ocean free surface is time dependent, so that the ocean domain,  $\mathcal{R}$ , is itself time dependent. Mathematically, this time dependence means that time and space integrations do not commute. In particular, space integration is generally written in the form

$$\int_{\mathcal{R}} dV = \iint \left[ \int_{\eta_b}^{\eta} dz \right] dx dy, \quad (16.138)$$

where  $z = \eta_b(x, y)$  is the static bottom and  $z = \eta(x, y, t)$  is the time dependent free surface. We must, in turn, first compute the space integration over the full domain and then do the time integration when deriving the [reciprocity condition](#) satisfied by the Green's function, as well as the integral expression for the tracer concentration in terms of the Green's function.

The free surface undulations make the vertical extent of the domain time dependent. Additionally, in an ocean with sloping sidewalls, the horizontal domain boundaries are time dependent due to the motion of the ocean up and down the sloping sides. However, allowing for a fluctuating horizontal domain extent does not introduce any fundamentally new kinematics to the analysis. The reason is that when integrating to the lateral domain boundaries, all terms vanish since the water depth vanishes at the edge of the sloping beaches. We saw this kinematic result in [VOLUME 2](#) when integrating the angular momentum budget in a channel with sloping sidewalls.

To slightly ease the analysis in this section, we assume the horizontal extent of the domain to be static. We do so by imagining a few meter high vertical seawall placed around the ocean domain edges, and by assuming a minimum depth so that there is nonzero water everywhere in the domain. These assumptions are common in ocean modeling, except in models allowing for wetting and drying of land/ocean cells. So in conclusion, we limit our analysis to time dependence of the vertical extent of the domain, with the horizontal extent static. Such limitation can be removed without much difficulty, but doing so adds nothing new fundamentally.

### 16.9.2 Passive tracer boundary conditions

In this section, we are concerned with the evolution of a smooth passive tracer concentration,  $C$ , which is the dimensionless number between zero and unity. Boundary conditions play a key role in the evolution. We here discuss the boundary conditions placed on the passive tracer along the ocean bottom, at  $z = \eta_b(x, y)$ , and the ocean free surface, at  $z = \eta(x, y, t)$ .

#### Ocean bottom

At the static and solid-earth bottom to the ocean, we consider a no-flux condition for the diffusive flux

$$\mathbf{J} \cdot \hat{\mathbf{n}} = -\rho \mathbf{K} \cdot \nabla C \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta_b. \quad (16.139)$$

The no-flux condition, along with the kinematic no-normal flow condition,  $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$ , means that there is zero flux of the passive tracer through the bottom. This assumption can be readily modified if interested in probing the role of processes next to the ocean bottom that carry boundary data into the interior. However, most applications are concerned with surface boundary data, thus motivating us to consider homogeneous Neumann bottom boundary conditions.

### Ocean free surface

At the ocean free surface we use results from our study of mass kinematics in VOLUME 1, in which we developed the boundary conditions for mass flux  $\mathcal{Q}_m$  (mass per time per area) across a permeable free surface, as well as our study of tracer mass kinematics also in VOLUME 1, where we developed the analogous boundary conditions for tracers. In particular, we make use of the following expression for the net mass flux of tracer crossing the free surface,  $\mathcal{Q}_c$ , written as the sum of an advective flux plus a non-advectional flux

$$\mathcal{Q}_c = C \mathcal{Q}_m - \mathbf{J} \cdot \hat{\mathbf{n}} = \underbrace{C \mathcal{Q}_m}_{\text{advective}} + \underbrace{\rho \mathbf{K} \cdot \nabla C \cdot \hat{\mathbf{n}}}_{\text{diffusive}} \quad \text{for } \mathbf{x} \in \partial \mathcal{R}_{\text{surf}}, \quad (16.140)$$

where  $C$  is the concentration at the free surface,  $z = \eta(x, y, t)$ , and we assumed the non-advectional flux is given by a diffusive flux. We consider the following prescribed boundary conditions.

- ROBIN CONDITIONS: Prescribing the boundary tracer mass flux,  $\mathcal{Q}_c$ , leads to a [Robin boundary condition](#), also called a mixed boundary condition

$$\mathcal{Q}_c = \text{prescribed} = C \mathcal{Q}_m + \rho \mathbf{K} \cdot \nabla C \cdot \hat{\mathbf{n}} \quad \text{for } \mathbf{x} \in \partial \mathcal{R}_{\text{surf}}. \quad (16.141)$$

This boundary condition is relevant for enthalpy and salt, with full discussion given in [Section 19.5](#). However, the Robin condition is rarely used for passive tracers along the ocean surface and so it is not further considered in this section.

- NEUMANN CONDITIONS: Prescribing the diffusive flux leads to the [Neumann boundary condition](#)

$$\rho \mathbf{K} \cdot \nabla C \cdot \hat{\mathbf{n}} = \text{prescribed} \equiv \Sigma(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in \partial \mathcal{R}_{\text{surf}}. \quad (16.142)$$

This surface ocean boundary condition is also rarely used for passive tracers, though we do examine it within the following.

- DIRICHLET CONDITIONS: Prescribing the value of the tracer concentration at the boundary leads to the [Dirichlet boundary condition](#)

$$C = \text{prescribed} \equiv \sigma(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in \partial \mathcal{R}_{\text{surf}}. \quad (16.143)$$

The Dirichlet boundary condition is the most commonly used condition for passive tracers. Hence it is our focus in this section. Note that both the Neumann and Dirichlet conditions generally involve a net transport of tracer,  $\mathcal{Q}_c \neq 0$ , across the ocean boundary.

As shown in [Section 16.9.6](#), the Neumann boundary condition in the presence of a surface mass flux is problematic due to the associated non-closed reciprocity relation satisfied by the Green's function and its adjoint. The absence of a suitable reciprocity relation makes it difficult to use the Green's function method. In contrast, the Dirichlet condition allows for a reciprocity corresponding to that found for the diffusion equation in VOLUME 1. Hence, the advection-diffusion Green's function, with a Dirichlet condition, is suitable even in the presence of a surface mass flux.<sup>8</sup>

---

<sup>8</sup>See page 2450 of [Larson \(1999\)](#) for a similar point.

### 16.9.3 Advection-diffusion initial-boundary value problem

We study the initial-boundary value problem for a smooth passive tracer concentration,  $C$ , which is the dimensionless ratio of the tracer mass to seawater mass. The tracer is affected by advection and diffusion on a spatial domain,  $\mathcal{R}$ , in the presence of a tracer source,  $\rho \Lambda$  (with dimensions of tracer mass per volume per time), with initial data available for the density and tracer concentration at time  $t = t_{\text{init}}$ . The initial-boundary value problem in the presence of Neumann or Dirichlet boundary conditions is given by

$$\partial_t(\rho C) + \nabla \cdot (\rho \mathbf{v}^\dagger C - \rho \mathbf{K} \cdot \nabla C) = \rho \Lambda \quad \mathbf{x} \in \mathcal{R}, t \geq t_{\text{init}} \quad (16.144a)$$

$$\rho C = \rho I \quad \mathbf{x} \in \mathcal{R}, t = t_{\text{init}} \quad (16.144b)$$

$$\hat{\mathbf{n}} \cdot \rho \mathbf{K} \cdot \nabla C = \rho \Sigma \quad \text{or} \quad \rho C = \rho \sigma \quad \mathbf{x} \in \partial \mathcal{R}, t \geq t_{\text{init}}. \quad (16.144c)$$

The prescribed initial condition data for the tracer concentration at time  $t = t_{\text{init}}$  is given by  $I(\mathbf{x})$ , and the initial fluid density is also prescribed at this time,  $\rho(\mathbf{x}, t_{\text{init}})$ . We consider two options for the boundary condition at  $\mathbf{x} \in \partial \mathcal{R}$ : (i) the Neumann boundary condition with a prescribed flux,  $\hat{\mathbf{n}} \cdot \rho \mathbf{K} \cdot \nabla C = \rho \Sigma$ , or the (ii) Dirichlet boundary condition with a prescribed value,  $\rho C = \rho \sigma$ . Furthermore, we assume the flow field,  $\mathbf{v}(\mathbf{x}, t)$ , the eddy-induced velocity,  $\mathbf{v}^*(\mathbf{x}, t)$ , the seawater density,  $\rho(\mathbf{x}, t)$ , and the diffusion tensor,  $\mathbf{K}(\mathbf{x}, t)$ , are known functions of space-time that are determined by solving for the dynamics, kinematics, thermodynamics, and material tracers. Finally, we assume the tracer concentration source,  $\Lambda(\mathbf{x}, t)$ , does not itself depend on the tracer concentration,  $C$ , thus ensuring linearity of the partial differential equation (16.144a).

### 16.9.4 The Green's function problem

The Green's function corresponding to the passive tracer advection-diffusion equations (16.144a)-(16.144c) satisfies the following causal boundary value problem

$$\partial_t[\rho G(\mathbf{x}, t | \mathbf{x}_o, t_o)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(\mathbf{x}, t | \mathbf{x}_o, t_o) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(\mathbf{x}, t | \mathbf{x}_o, t_o)] = \delta(\mathbf{x} - \mathbf{x}_o) \delta(t - t_o) \quad (16.145a)$$

$$G(\mathbf{x}, t < t_o | \mathbf{x}_o, t_o) = 0 \quad (16.145b)$$

$$\hat{\mathbf{n}}_{\mathbf{x}} \cdot \mathbf{K} \cdot \nabla_{\mathbf{x}} G(\mathbf{x} \in \partial \mathcal{R}, t | \mathbf{x}_o, t_o) = 0 \quad \text{or} \quad G(\mathbf{x} \in \partial \mathcal{R}, t | \mathbf{x}_o, t_o) = 0. \quad (16.145c)$$

The space-time point,  $(\mathbf{x}_o, t_o)$ , is where the Dirac delta source is located, which is within the spatial domain,  $\mathcal{R}$ , and it fires at a time after the initial time

$$\mathbf{x}_o \in \mathcal{R} \quad \text{and} \quad t_o \geq t_{\text{init}}. \quad (16.146)$$

The Green's function satisfies homogeneous boundary conditions that correspond to the boundary conditions satisfied by the passive tracer concentration in equation (16.144c). That is, if the passive tracer satisfies a Dirichlet boundary condition, then the Green's function satisfies a homogeneous Dirichlet condition, and likewise for Neumann boundary conditions. Finally, since the Dirac delta source,  $\delta(\mathbf{x} - \mathbf{x}_o) \delta(t - t_o)$ , has dimensions of inverse volume times inverse time, the Green's function has dimensions of inverse mass. We physically interpret the Green's function as the tracer concentration resulting from an impulsive tracer concentration source, divided by the mass of tracer injected by the source.

### 16.9.5 Adjoint Green's function problem

The adjoint Green's function,  $G^\ddagger$ , satisfies the adjoint problem<sup>9</sup>

$$-\partial_t[\rho G^\ddagger(\mathbf{x}, t|\mathbf{x}_o, t_o)] + \nabla_{\mathbf{x}} \cdot [-\rho \mathbf{v}^\dagger G^\ddagger(\mathbf{x}, t|\mathbf{x}_o, t_o) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(\mathbf{x}, t|\mathbf{x}_o, t_o)] = \delta(\mathbf{x} - \mathbf{x}_o) \delta(t - t_o) \quad (16.147a)$$

$$G^\ddagger(\mathbf{x}, t > t_o|\mathbf{x}_o, t_o) = 0 \quad (16.147b)$$

$$\hat{\mathbf{n}}_{\mathbf{x}} \cdot \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(\mathbf{x} \in \partial\mathcal{R}, t|\mathbf{x}_o, t_o) = 0 \text{ or } G^\ddagger(\mathbf{x} \in \partial\mathcal{R}, t|\mathbf{x}_o, t_o) = 0. \quad (16.147c)$$

Note the sign change on both the time derivative, as for the diffusion equation in VOLUME 1, as well as the advection term. The sign change on advection is expected since with time running backwards, so too does the velocity. Hence, the adjoint advection-diffusion equation is a backwards in time advection-diffusion equation.

### 16.9.6 Reciprocity condition

We here derive the **reciprocity condition** satisfied by the Green's function,  $G$ , and its adjoint,  $G^\ddagger$ . The derivation follows that in VOLUME 1 for the diffusion equation. A new feature here arises from the advection operator, and another arises from allowing the domain boundary to be time dependent as occurs at the ocean free surface. We will see that reciprocity for Neumann boundary conditions (16.145c) and (16.147c) does not “close” when there is mass transport across the ocean free surface (see equation (16.165) below). In contrast, reciprocity closes with Dirichlet boundary conditions, taking the same form as for the diffusion equation from VOLUME 1. We have more to say on this distinct behavior after its derivation.

#### Notation and setup

Consider the Green's function partial differential equation (16.145a) with a Dirac delta source,  $\delta(\mathbf{x} - \mathbf{x}_1) \delta(t - t_1)$ , along with the adjoint Green's function equation (16.147a) with a Dirac delta source,  $\delta(\mathbf{x} - \mathbf{x}_2) \delta(t - t_2)$ , where both sources are within the spatial domain and both occur later than the initial time:

$$\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{R} \quad \text{and} \quad t_{\text{init}} < t_1, t_2. \quad (16.148)$$

We follow the approach for the diffusion equation in VOLUME 1 by introducing the arbitrarily large time,  $T$ , such that

$$-T < t_1, t_2 < T. \quad (16.149)$$

As for the diffusion equation, causality conditions ensure that  $T$  drops out from the final expression. Additionally, to help ease notational clutter, we make use of the following shorthand where convenient

$$G(\mathbf{x}, t|\mathbf{x}_1, t_1) = G(1) \quad \text{and} \quad G^\ddagger(\mathbf{x}, t|\mathbf{x}_2, t_2) = G^\ddagger(2). \quad (16.150)$$

#### Cross-multiplication

Multiply the Green's function equation (16.145a) by  $G^\ddagger(2)$  and the adjoint equation (16.147a) by  $G(1)$  to find

$$G^\ddagger(2) (\partial_t[\rho G(1)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(1) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(1)]) = G^\ddagger(2) \delta(\mathbf{x} - \mathbf{x}_1) \delta(t - t_1) \quad (16.151a)$$

---

<sup>9</sup>We use the double dagger for the adjoint Green's function,  $G^\ddagger$ , to distinguish it from the single dagger used for the residual velocity,  $\mathbf{v}^\dagger$ .

$$G(1) (-\partial_t[\rho G^\ddagger(2)] + \nabla_{\mathbf{x}} \cdot [-\rho \mathbf{v}^\dagger G^\ddagger(2) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(2)]) = G(1) \delta(\mathbf{x} - \mathbf{x}_2) \delta(t - t_2). \quad (16.151b)$$

In the following, we work from the left hand side of equation (16.151a) and bring the differential operators from  $G(1)$  onto  $G^\ddagger(2)$ . The result of this movement will be equation (16.151b) plus some extra terms whose form depends on the causality condition and boundary conditions. Integration over space and time will then render the reciprocity relation.

### Self-adjointness of the generalized Laplacian operator

The generalized Laplacian operator term on the left hand side of equation (16.151a) can be written

$$\begin{aligned} & -G^\ddagger(2) \nabla_{\mathbf{x}} \cdot [\rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(1)] \\ &= \nabla_{\mathbf{x}} \cdot [-G^\ddagger(2) \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(1) + G(1) \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(2)] - G(1) \nabla_{\mathbf{x}} \cdot [\rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(2)]. \end{aligned} \quad (16.152)$$

A spatial integration of this equation over the region  $\mathcal{R}$ , and use of the homogeneous boundary conditions in equations (16.144c) or (16.147c), eliminates the divergence term to reveal

$$\begin{aligned} & \int_{\mathcal{R}} G^\ddagger(\mathbf{x}, t | \mathbf{x}_2, t_2) \nabla_{\mathbf{x}} \cdot [\rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(\mathbf{x}, t | \mathbf{x}_1, t_1)] dV \\ &= \int_{\mathcal{R}} G(\mathbf{x}, t | \mathbf{x}_1, t_1) \nabla_{\mathbf{x}} \cdot [\rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(\mathbf{x}, t | \mathbf{x}_2, t_2)] dV. \end{aligned} \quad (16.153)$$

This equality proves that the generalized Laplacian operator with a symmetric diffusion tensor is self-adjoint, which is a result encountered in our study of the diffusion equation in VOLUME 1. This result holds for either Neumann or Dirichlet boundary conditions.

### Time derivative plus advection

Next write the time derivative and advection portion of equation (16.151a) as

$$\begin{aligned} & G^\ddagger(2) \left( \partial_t[\rho G(1)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(1)] \right) \\ &= \partial_t[G^\ddagger(2) \rho G(1)] + \nabla_{\mathbf{x}} \cdot [G^\ddagger(2) \rho \mathbf{v}^\dagger G(1)] - G(1) \rho \left[ \partial_t G^\ddagger(2) + \mathbf{v}^\dagger \cdot \nabla_{\mathbf{x}} G^\ddagger(2) \right] \\ &= \partial_t[G^\ddagger(2) \rho G(1)] + \nabla_{\mathbf{x}} \cdot [G^\ddagger(2) \rho \mathbf{v}^\dagger G(1)] - G(1) \left( \partial_t(\rho G^\ddagger(2)) + \nabla \cdot [\rho \mathbf{v}^\dagger G^\ddagger(2)] \right), \end{aligned} \quad (16.154)$$

where we used mass continuity (16.67a) for the final equality. Rearrangement thus leads to

$$\begin{aligned} & G^\ddagger(2) \left( \partial_t[\rho G(1)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(1)] \right) - G(1) \left( -\partial_t[\rho G^\ddagger(2)] - \nabla \cdot [\rho \mathbf{v}^\dagger G^\ddagger(2)] \right) \\ &= \partial_t[G^\ddagger(2) \rho G(1)] + \nabla_{\mathbf{x}} \cdot [G^\ddagger(2) \rho \mathbf{v}^\dagger G(1)]. \end{aligned} \quad (16.155)$$

### Space integration

Now integrate equations (16.151a) and (16.151b) over the spatial domain,  $\mathcal{R}$ , subtract these two equations, and make use of the results (16.153) and (16.155) to reveal

$$\begin{aligned} & G^\ddagger(\mathbf{x}_1, t | \mathbf{x}_2, t_2) \delta(t - t_1) - G(\mathbf{x}_2, t | \mathbf{x}_1, t_1) \delta(t - t_2) \\ &= \int_{\mathcal{R}} \left[ \partial_t[G^\ddagger(2) \rho G(1)] + \nabla_{\mathbf{x}} \cdot [G^\ddagger(2) \rho \mathbf{v}^\dagger G(1)] \right] dV. \end{aligned} \quad (16.156)$$

The divergence term on the right hand side takes the form

$$\int_{\mathcal{R}} \nabla_{\mathbf{x}} \cdot [G^{\ddagger}(2) \rho \mathbf{v}^{\dagger} G(1)] dV = \oint_{\partial\mathcal{R}} G^{\ddagger}(2) G(1) \rho \mathbf{v}^{\dagger} \cdot \hat{\mathbf{n}}_{\mathbf{x}} d\mathcal{S} \quad \text{divergence thm (16.157a)}$$

$$= \oint_{\partial\mathcal{R}} G^{\ddagger}(2) G(1) \rho \mathbf{v} \cdot \hat{\mathbf{n}}_{\mathbf{x}} d\mathcal{S} \quad \hat{\mathbf{n}} \cdot \mathbf{v}^* = 0 \quad (16.157b)$$

$$= \int_{z=\eta} G^{\ddagger}(2) G(1) \rho \mathbf{v} \cdot \hat{\mathbf{n}}_{\mathbf{x}} d\mathcal{S} \quad \mathbf{v} \cdot \hat{\mathbf{n}} = 0 \text{ at } z = \eta_b \quad (16.157c)$$

$$= \int_{z=\eta} G^{\ddagger}(2) G(1) \rho \mathbf{v} \cdot \nabla(z - \eta) dA \quad \text{area relation} \quad (16.157d)$$

$$= \int_{z=\eta} G^{\ddagger}(2) G(1) \rho (w - \mathbf{u} \cdot \nabla \eta) dA. \quad (16.157e)$$

Equation (16.157d) made use of the area relation from the kinematics in VOLUME 1 that relates the surface area element,  $d\mathcal{S}$ , on the free surface, to its horizontal projection,  $dA$ ,

$$\hat{\mathbf{n}}_{\mathbf{x}} d\mathcal{S} = \nabla(z - \eta) dA. \quad (16.158)$$

The time derivative term in equation (16.156) takes the form

$$\int_{\mathcal{R}} \partial_t [G^{\ddagger}(2) G(1) \rho] dV = \frac{\partial}{\partial t} \left[ \int_{\mathcal{R}} G^{\ddagger}(2) G(1) \rho dV \right] - \int_{z=\eta} [G^{\ddagger}(2) G(1) \rho \partial_t \eta] dA, \quad (16.159)$$

where we made use of Leibniz's rule to bring the time derivative across the integral sign and made note of the time dependent free surface,  $z = \eta(x, y, t)$ . Combining equations (16.159) and (16.157e) leads to

$$G^{\ddagger}(\mathbf{x}_1, t | \mathbf{x}_2, t_2) \delta(t - t_1) - G(\mathbf{x}_2, t | \mathbf{x}_1, t_1) \delta(t - t_2) \quad (16.160a)$$

$$= \frac{\partial}{\partial t} \left[ \int_{\mathcal{R}} G^{\ddagger}(2) G(1) \rho dV \right] + \int_{z=\eta} G^{\ddagger}(2) G(1) [\rho (w - \mathbf{u} \cdot \nabla \eta - \partial_t \eta)] dA \quad (16.160b)$$

$$= \frac{\partial}{\partial t} \left[ \int_{\mathcal{R}} G^{\ddagger}(2) G(1) \rho dV \right] - \int_{z=\eta} G^{\ddagger}(2) G(1) Q_m dA, \quad (16.160c)$$

where the final equality follows from the surface ocean kinematic boundary condition from VOLUME 1,

$$(\partial_t + \mathbf{u} \cdot \nabla_h) \eta = w + Q_m / \rho, \quad (16.161)$$

with  $Q_m$  the mass per time per horizontal area crossing the ocean free surface.

### Time integration

We are now ready to time integrate equation (16.160c), with the left hand side leading to

$$\begin{aligned} \int_{-T}^T & \left[ G^{\ddagger}(\mathbf{x}_1, t | \mathbf{x}_2, t_2) \delta(t - t_1) - G(\mathbf{x}_2, t | \mathbf{x}_1, t_1) \delta(t - t_2) \right] dt \\ &= G^{\ddagger}(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2) - G(\mathbf{x}_2, t_2 | \mathbf{x}_1, t_1), \end{aligned} \quad (16.162)$$

which used the **sifting property** of the Dirac delta. There are two terms that appear when time integrating the time derivative on the right hand side of equation (16.160c), with each term vanishing due to the causality conditions (16.145b) and (16.147b)

$$\int_{\mathcal{R}} G^{\ddagger}(\mathbf{x}, t = T | \mathbf{x}_2, t_2) G(\mathbf{x}, t = T | \mathbf{x}_1, t_1) \rho dV = 0 \iff G^{\ddagger}(\mathbf{x}, t = T | \mathbf{x}_2, t_2) = 0 \quad (16.163)$$

$$\int_{\mathcal{R}} G^{\ddagger}(\mathbf{x}, t = t_{\text{init}} | \mathbf{x}_2, t_2) G(\mathbf{x}, t = t_{\text{init}} | \mathbf{x}_1, t_1) \rho dV = 0 \iff G(\mathbf{x}, t = t_{\text{init}} | \mathbf{x}_1, t_1) = 0. \quad (16.164)$$

We are thus left with

$$\begin{aligned} G^{\ddagger}(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2) - G(\mathbf{x}_2, t_2 | \mathbf{x}_1, t_1) \\ = - \int_{t_{\text{init}}}^{t_1} \left[ \int_{z=\eta} G^{\ddagger}(\mathbf{x}, t | \mathbf{x}_2, t_2) G(\mathbf{x}_1, t_1 | \mathbf{x}, t) Q_m dA \right] dt, \end{aligned} \quad (16.165)$$

which we refer to as a *non-closed reciprocity relation* between  $G$  and  $G^{\ddagger}$ . Note that the time limits on the integral follow from causality on the Green's function and its adjoint.

### Closed form reciprocity in special cases

There are two cases in which the relation (16.165) leads to a closed reciprocity relation:

- Zero mass flux across the ocean free surface:  $Q_m = 0$ .
- Homogeneous Dirichlet boundary conditions at the ocean free surface, in which case  $G^{\ddagger}(\mathbf{x} \in \partial\mathcal{R}, t | \mathbf{x}_o, t_o) = G(\mathbf{x} \in \partial\mathcal{R}, t | \mathbf{x}_o, t_o) = 0$ .

In either case we are led to

$$G^{\ddagger}(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2) = G(\mathbf{x}_2, t_2 | \mathbf{x}_1, t_1), \quad (16.166)$$

which is the same reciprocity condition satisfied for the diffusion equation Green's functions in VOLUME 1.

The more nuanced reciprocity for the advection-diffusion equation arises from the advective mass flux at the ocean free surface boundary. The mass flux couples the ocean with its surrounding media (e.g., the atmosphere, rivers, or cryosphere), even when using homogeneous Neumann conditions. In so doing, we are not afforded a generally closed reciprocity relation. The Dirichlet boundary condition closes the surface boundary by eliminating the direct role of the surface mass flux. Most applications of Green's function methods for passive ocean tracers make use of Dirichlet boundary conditions, in which case we are afforded a closed reciprocity relation even with a free surface open to mass transport.

### 16.9.7 Composition property

We here follow the analysis of the diffusion equation Green's function in VOLUME 1 to derive the composition property of the Green's function for the advection-diffusion equation. For this purpose, return to the cross-multiplication equations (16.151a) and (16.151b) used to derive

reciprocity, here written as

$$G^\ddagger(2) \left( \partial_t[\rho G(1)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(1) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(1)] \right) = G^\ddagger(2) \delta(\mathbf{x} - \mathbf{x}_1) \delta(t - t_1) \quad (16.167a)$$

$$G(1) \left( \partial_t[\rho G^\ddagger(2)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G^\ddagger(2) + \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(2)] \right) = -G(1) \delta(\mathbf{x} - \mathbf{x}_2) \delta(t - t_2). \quad (16.167b)$$

Adding these two equations and use of mass continuity (16.67a) brings the left hand side to

$$\text{LHS} = \partial_t[\rho G(1) G^\ddagger(2)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(1) G^\ddagger(2) + \rho G(1) \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(2) - \rho G^\ddagger(2) \mathbf{K} \cdot \nabla_{\mathbf{x}} G(1)]. \quad (16.168)$$

For the time derivative term we use Leibniz's rule to write

$$\int_{\mathcal{R}} \partial_t[\rho G(1) G^\ddagger(2)] dV = \frac{d}{dt} \int_{\mathcal{R}} \rho G(1) G^\ddagger(2) dV - \int_{z=\eta} \rho G(1) G^\ddagger(2) \partial_t \eta dA. \quad (16.169a)$$

For the advection term we follow the manipulations used for equation (16.157e) to derive

$$\int_{\mathcal{R}} \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(1) G^\ddagger(2)] dV = \int_{z=\eta} \rho G(1) G^\ddagger(2) \mathbf{v}^\dagger \cdot \hat{\mathbf{n}} dS \quad (16.170a)$$

$$= \int_{z=\eta} \rho G(1) G^\ddagger(2) (w - \mathbf{u} \cdot \nabla \eta) dA \quad (16.170b)$$

$$= \int_{z=\eta} G(1) G^\ddagger(2) (\rho \partial_t \eta - Q_m) dA. \quad (16.170c)$$

Bringing the pieces together we find that the  $\partial_t \eta$  term is eliminated, so that

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{R}} \rho(\mathbf{x}, t) G(\mathbf{x}, t | \mathbf{x}_1, t_1) G^\ddagger(\mathbf{x}, t | \mathbf{x}_2, t_2) dV &= G^\ddagger(\mathbf{x}_1, t | \mathbf{x}_2, t_2) \delta(t - t_1) \\ &\quad - G(\mathbf{x}_2, t | \mathbf{x}_1, t_1) \delta(t - t_2) - \int_{z=\eta} G(\mathbf{x}, t | \mathbf{x}_1, t_1) G^\ddagger(\mathbf{x}, t | \mathbf{x}_2, t_2) Q_m(\mathbf{x}, t) dA. \end{aligned} \quad (16.171)$$

As for the derivation of reciprocity in Section 16.9.6, we here assume either  $Q_m = 0$  or a homogeneous Dirichlet boundary condition so that

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{R}} \rho(\mathbf{x}, t) G(\mathbf{x}, t | \mathbf{x}_1, t_1) G^\ddagger(\mathbf{x}, t | \mathbf{x}_2, t_2) dV &= G^\ddagger(\mathbf{x}_1, t | \mathbf{x}_2, t_2) \delta(t - t_1) \\ &\quad - G(\mathbf{x}_2, t | \mathbf{x}_1, t_1) \delta(t - t_2). \end{aligned} \quad (16.172)$$

Equation (16.172) is identical to that satisfied by the diffusion equation Green's function, as found in VOLUME 1. As for the diffusion case, time integration of equation (16.172) requires care about how the time limits are placed relative to the two Dirac sources. However, all of these details are identical to the diffusion case from VOLUME 1, so we can simply take those results to write the composition property for the advection-diffusion Green's function

$$G(\mathbf{x}_2, t_2 | \mathbf{x}_1, t_1) = \int_{\mathcal{R}} \rho(\mathbf{x}, \tau) G(\mathbf{x}_2, t_2 | \mathbf{x}, \tau) G(\mathbf{x}, \tau | \mathbf{x}_1, t_1) dV \quad \text{if } t_1 < \tau < t_2. \quad (16.173)$$

The left hand side of this equation is the response from a Dirac source that is advected-

diffused from  $(\mathbf{x}_1, t_1)$  and measured at the space-time point  $(\mathbf{x}_2, t_2)$ . The right hand side is the composition of a Green's function feeling the source at  $(\mathbf{x}_1, t_1)$  but now sampled at an intermediate space-time position,  $(\mathbf{x}, \tau)$ , and then further advective-diffused to  $(\mathbf{x}_2, t_2)$ , with integration over all possible intermediate positions  $\mathbf{x}$ . The intermediate sampling can occur at an arbitrary intermediate time  $\tau$ , so long as  $t_1 < \tau < t_2$ . The composition property (16.173) allows us to conceive of a long-time interval Green's function as the composition of an arbitrary number of shorter time interval Green's functions.

### 16.9.8 Integral expression for the tracer concentration

We are now ready to express the passive tracer concentration,  $C$ , as a suite of integrals involving the Green's function and the known boundary and initial conditions as well as the known source function. The process for deriving this expression is identical to that used in Section 16.9.6 for reciprocity, with the following steps offered for completeness.

#### Derivation setup

The initial-boundary value problem for the passive tracer is given by

$$\partial_t(\rho C) + \nabla \cdot (\rho \mathbf{v}^\dagger C - \rho \mathbf{K} \cdot \nabla C) = \rho \Lambda \quad \mathbf{x} \in \mathcal{R}, t \geq t_{\text{init}} \quad (16.174a)$$

$$\rho C = \rho I \quad \mathbf{x} \in \mathcal{R}, t = t_{\text{init}} \quad (16.174b)$$

$$\rho C = \rho \sigma \quad \mathbf{x} \in \partial \mathcal{R}, t \geq t_{\text{init}} \quad (16.174c)$$

where we only consider the Dirichlet boundary condition to ensure a closed reciprocity relation in the presence of surface mass fluxes (Section 16.9.6). The corresponding adjoint Green's function satisfies

$$-\partial_t[\rho G^\ddagger(\mathbf{x}, t | \mathbf{x}_o, t_o)] + \nabla_{\mathbf{x}} \cdot [-\rho \mathbf{v}^\dagger G^\ddagger(\mathbf{x}, t | \mathbf{x}_o, t_o) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^\ddagger(\mathbf{x}, t | \mathbf{x}_o, t_o)] = \delta(\mathbf{x} - \mathbf{x}_o) \delta(t - t_o) \quad (16.175a)$$

$$G^\ddagger(\mathbf{x}, t > t_o | \mathbf{x}_o, t_o) = 0 \quad (16.175b)$$

$$G^\ddagger(\mathbf{x} \in \partial \mathcal{R}, t | \mathbf{x}_o, t_o) = 0, \quad (16.175c)$$

with the reciprocity condition (16.166) holding since we chose Dirichlet boundary conditions. Multiplying the adjoint Green's function equation (16.175a) by  $C(\mathbf{x}, t)$  and performing manipulations just like those for reciprocity leads to

$$\begin{aligned} -\partial_t(\rho C G^\ddagger) + \nabla_{\mathbf{x}} \cdot [G^\ddagger \rho \mathbf{K} \cdot \nabla C - C \rho \mathbf{K}(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}} G^\ddagger - C \rho \mathbf{v}^\dagger G^\ddagger] + G^\ddagger \rho \Lambda \\ = C(\mathbf{x}, t) \delta(\mathbf{x} - \mathbf{x}_o) \delta(t - t_o). \end{aligned} \quad (16.176)$$

With the homogeneous Dirichlet conditions satisfied by  $G^\ddagger$  on the spatial boundaries, a space and time integration over  $(\mathbf{x}, t)$  leads to

$$\begin{aligned} C(\mathbf{x}_o, t_o) = \int_{\mathcal{R}} G^\ddagger(\mathbf{x}, t_{\text{init}} | \mathbf{x}_o, t_o) \rho(\mathbf{x}, t_{\text{init}}) I(\mathbf{x}) dV \\ + \int_{t_{\text{init}}}^{t_o} \left[ \int_{\mathcal{R}} G^\ddagger(\mathbf{x}, t | \mathbf{x}_o, t_o) \rho(\mathbf{x}, t) \Lambda(\mathbf{x}, t) dV \right] dt \\ - \int_{t_{\text{init}}}^{t_o} \left[ \oint_{\partial \mathcal{R}} \sigma(\mathbf{x}, t) \rho(\mathbf{x}, t) \mathbf{K}(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}} G^\ddagger(\mathbf{x}, t | \mathbf{x}_o, t_o) \cdot \hat{\mathbf{n}}_{\mathbf{x}} dS \right] dt. \end{aligned} \quad (16.177)$$

Use of the reciprocity relation (16.166) allows us to write this equation in terms of the Green's function rather than the adjoint Green's function

$$\begin{aligned} C(\mathbf{x}_o, t_o) = & \int_{\mathcal{R}} G(\mathbf{x}_o, t_o | \mathbf{x}, t_{\text{init}}) \rho(\mathbf{x}, t_{\text{init}}) I(\mathbf{x}) dV \\ & + \int_{t_{\text{init}}}^{t_o} \left[ \int_{\mathcal{R}} G(\mathbf{x}_o, t_o | \mathbf{x}, t) \rho(\mathbf{x}, t) \Lambda(\mathbf{x}, t) dV \right] dt \\ & - \int_{t_{\text{init}}}^{t_o} \left[ \oint_{\partial\mathcal{R}} \sigma(\mathbf{x}, t) \rho(\mathbf{x}, t) \mathbf{K}(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}} G(\mathbf{x}_o, t_o | \mathbf{x}, t) \cdot \hat{\mathbf{n}}_{\mathbf{x}} d\mathcal{S} \right] dt. \quad (16.178) \end{aligned}$$

Finally, swapping labels  $(\mathbf{x}_o, t_o) \leftrightarrow (\mathbf{x}, t)$  renders

$$\begin{aligned} C(\mathbf{x}, t) = & \int_{\mathcal{R}} G(\mathbf{x}, t | \mathbf{x}_o, t_{\text{init}}) \rho(\mathbf{x}_o, t_{\text{init}}) I(\mathbf{x}_o) dV_0 \\ & + \int_{t_{\text{init}}}^t \left[ \int_{\mathcal{R}} G(\mathbf{x}, t | \mathbf{x}_o, t_o) \rho(\mathbf{x}_o, t_o) \Lambda(\mathbf{x}_o, t_o) dV_0 \right] dt_o \\ & - \int_{t_{\text{init}}}^t \left[ \oint_{\partial\mathcal{R}} \sigma(\mathbf{x}_o, t_o) \rho(\mathbf{x}_o, t_o) \mathbf{K}(\mathbf{x}_o, t_o) \cdot \nabla_{\mathbf{x}_o} G(\mathbf{x}, t | \mathbf{x}_o, t_o) \cdot \hat{\mathbf{n}}_{\mathbf{x}_o} d\mathcal{S}_0 \right] dt_o. \quad (16.179) \end{aligned}$$

This solution manifests causality since the tracer concentration at time  $t$  is a function only of processes occurring from  $t_{\text{init}}$  up to time  $t$ .

### 16.9.9 Properties of the Green's function solution

The integral solution (16.179) is of the same form as that for the diffusion equation in VOLUME 1. We here summarize properties of the solution, which largely follow those for the diffusion equation.

#### The role of advection and diffusion at boundaries

Explicit contributions from the advective flux are absent from the solution (16.179). Namely, there are no advective flux contributions at the surface boundary. Such contributions are missing due to the homogeneous Dirichlet boundary conditions imposed on the Green's function. For the ocean bottom, material and rigid no-flux conditions mean that  $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$  at the bottom. The presence of advection arises only through its effect on the Green's function, which is affected by both advection and diffusion.

Furthermore, notice how in the absence of diffusion (i.e.,  $\mathbf{K} = 0$ ) the Dirichlet boundary data is unable to penetrate into the ocean interior since the surface boundary integral vanishes from equation (16.179). In effect, the surface boundary becomes a material surface when there is no diffusion. That is, diffusive mixing is needed for boundary data to move into the interior. This role for diffusion again appears in our study of the surface flux condition for salt and freshwater in Section 19.5 (see also [Nurser and Griffies \(2019\)](#)).

#### Initial conditions

When sampling the tracer concentration at the initial time,  $t \downarrow t_{\text{init}}$ , causality means that all the time integrals vanish from the solution (16.179), thus leaving

$$\lim_{t \downarrow t_{\text{init}}} C(\mathbf{x}, t) = \lim_{t \downarrow t_{\text{init}}} \int_{\mathcal{R}} G(\mathbf{x}, t | \mathbf{x}_o, t_{\text{init}}) \rho(\mathbf{x}_o, t_{\text{init}}) I(\mathbf{x}_o) dV_0. \quad (16.180)$$

Self-consistency implies that the Green's function satisfies

$$\lim_{t \downarrow t_{\text{init}}} G(\mathbf{x}, t | \mathbf{x}_o, t_{\text{init}}) \rho(\mathbf{x}_o, t_{\text{init}}) = \delta(\mathbf{x} - \mathbf{x}_o) \quad \text{with } \mathbf{x}, \mathbf{x}_o \in \mathcal{R}, \quad (16.181)$$

so that

$$\lim_{t \downarrow t_{\text{init}}} \int_{\mathcal{R}} \rho(\mathbf{x}, t_{\text{init}}) G(\mathbf{x}, t | \mathbf{x}_o, t_{\text{init}}) I(\mathbf{x}_o) dV_0 = \int_{\mathcal{R}} \delta(\mathbf{x} - \mathbf{x}_o) I(\mathbf{x}_o) dV_0 = I(\mathbf{x}). \quad (16.182)$$

### Dirichlet boundary conditions

Evaluating the Dirichlet solution (16.179) on a spatial boundary,  $\mathbf{x} \in \partial\mathcal{R}$ , eliminates both the volume integrals given that the Green's function satisfies homogeneous Dirichlet boundary conditions. The tracer concentration (16.179) thus takes the form

$$C(\mathbf{x}, t) = - \int_{t_{\text{init}}}^t \left[ \oint_{\partial\mathcal{R}} \sigma(\mathbf{x}_o, t_o) \rho(\mathbf{x}_o, t_o) \mathbf{K}(\mathbf{x}_o, t_o) \cdot \nabla_{\mathbf{x}_o} G(\mathbf{x}, t | \mathbf{x}_o, t_o) \cdot \hat{\mathbf{n}}_{\mathbf{x}_o} d\mathcal{S}_0 \right] dt_o \quad \text{with } \mathbf{x} \in \partial\mathcal{R}. \quad (16.183)$$

Self-consistency with the Dirichlet boundary condition (16.144c) implies that the Green's function, when both spatial points are evaluated on the boundary, satisfies

$$\rho(\mathbf{x}_o, t_o) \mathbf{K}(\mathbf{x}_o, t_o) \cdot \nabla_{\mathbf{x}_o} G(\mathbf{x}, t | \mathbf{x}_o, t_o) \cdot \hat{\mathbf{n}}_{\mathbf{x}_o} = -\delta(t - t_o) \delta^{(2)}(\mathbf{x} - \mathbf{x}_o) \quad \text{with } \mathbf{x}, \mathbf{x}_o \in \partial\mathcal{R}, \quad (16.184)$$

so that

$$C(\mathbf{x}, t) = \int_{t_{\text{init}}}^t \left[ \oint_{\partial\mathcal{R}} \sigma(\mathbf{x}_o, t_o) \delta(t - t_o) \delta^{(2)}(\mathbf{x} - \mathbf{x}_o) d\mathcal{S}_0 \right] dt_o = \sigma(\mathbf{x}, t) \quad \text{with } \mathbf{x} \in \partial\mathcal{R}. \quad (16.185)$$

### 16.9.10 Boundary propagator

#### Defining the boundary propagator

As for the diffusion equation in VOLUME 1, we here introduce the **boundary propagator** for the advection-diffusion equation with Dirichlet boundary conditions. For this purpose, consider the special case of a passive tracer with zero interior source and with zero initial condition, thus satisfying the initial-boundary value problem

$$\partial_t(\rho C) + \nabla \cdot (\rho \mathbf{v}^\dagger C - \rho \mathbf{K} \cdot \nabla C) = 0 \quad \mathbf{x} \in \mathcal{R}, t \geq t_{\text{init}} \quad (16.186a)$$

$$\rho C = 0 \quad \mathbf{x} \in \mathcal{R}, t = t_{\text{init}} \quad (16.186b)$$

$$\rho C = \rho \sigma \quad \mathbf{x} \in \partial\mathcal{R}, t \geq t_{\text{init}}, \quad (16.186c)$$

which leads to the simplification of the Green's function solution (16.179)

$$C(\mathbf{x}, t) = - \int_{t_{\text{init}}}^t \left[ \oint_{\partial\mathcal{R}} \sigma(\mathbf{x}_o, t_o) \rho(\mathbf{x}_o, t_o) \mathbf{K}(\mathbf{x}_o, t_o) \cdot \nabla_{\mathbf{x}_o} G(\mathbf{x}, t | \mathbf{x}_o, t_o) \cdot \hat{\mathbf{n}}_{\mathbf{x}_o} d\mathcal{S}_0 \right] dt_o. \quad (16.187)$$

The tracer concentration at a point in space-time is determined by the history of the advection and diffusion that transfers boundary information to this point. To manifest this cause-effect

relation, it is useful to define the boundary propagator just as for the diffusion equation

$$G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o) \equiv -\rho(\mathbf{x}_o, t_o) \mathbf{K}(\mathbf{x}_o, t_o) \cdot \nabla_{\mathbf{x}_o} G(\mathbf{x}, t | \mathbf{x}_o, t_o) \cdot \hat{\mathbf{n}}_{\mathbf{x}_o} \quad \text{with } \mathbf{x}_o \in \partial\mathcal{R}, \quad (16.188)$$

with  $G^{\text{bp}}$  having dimensions  $L^{-2} T^{-1}$ . The boundary propagator thus brings the tracer concentration (16.187) into the rather tidy form

$$C(\mathbf{x}, t) = \int_{t_{\text{init}}}^t \left[ \oint_{\partial\mathcal{R}} \sigma(\mathbf{x}_o, t_o) G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o) dS_0 \right] dt_o. \quad (16.189)$$

### Inhomogeneous Dirichlet at the surface and homogeneous Neumann at the bottom

In applications of passive tracers to study ocean circulation, it is common to apply inhomogeneous Dirichlet boundary conditions just at the ocean surface, and homogeneous Neumann boundary conditions (no-flux) at the ocean bottom

$$\partial_t(\rho C) + \nabla \cdot (\rho \mathbf{v}^\dagger C - \rho \mathbf{K} \cdot \nabla C) = 0 \quad \mathbf{x} \in \mathcal{R}, t \geq t_{\text{init}} \quad (16.190a)$$

$$C = 0 \quad \mathbf{x} \in \mathcal{R}, t = t_{\text{init}} \quad (16.190b)$$

$$\rho C = \rho \sigma \quad \mathbf{x} \in \partial\mathcal{R}_{\text{surf}}, t \geq t_{\text{init}} \quad (16.190c)$$

$$\hat{\mathbf{n}}_{\mathbf{x}} \cdot \mathbf{K} \cdot \nabla_{\mathbf{x}} C = 0 \quad \mathbf{x} \in \partial\mathcal{R}_{\text{bot}}, t \geq t_{\text{init}}. \quad (16.190d)$$

Note that since  $\hat{\mathbf{n}} \cdot \mathbf{v} = 0$  at the solid earth ocean bottom, kinematics imposes no advective flux through the bottom,  $\hat{\mathbf{n}} \cdot \mathbf{v} C = 0$ . Since the bottom boundary conditions are homogeneous, the solution (16.187) also holds for the initial-boundary value problem (16.190a)-(16.190d). The key distinction, however, is that the Green's function now satisfies the following boundary value problem

$$\partial_t[\rho G(\mathbf{x}, t | \mathbf{x}_o, t_o)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G(\mathbf{x}, t | \mathbf{x}_o, t_o) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G(\mathbf{x}, t | \mathbf{x}_o, t_o)] = \delta(\mathbf{x} - \mathbf{x}_o) \delta(t - t_o) \quad (16.191a)$$

$$G(\mathbf{x}, t < t_o | \mathbf{x}_o, t_o) = 0 \quad (16.191b)$$

$$G(\mathbf{x}, t | \mathbf{x}_o, t_o) = 0 \quad \mathbf{x} \in \partial\mathcal{R}_{\text{surf}} \quad (16.191c)$$

$$\hat{\mathbf{n}}_{\mathbf{x}} \cdot \mathbf{K} \cdot \nabla_{\mathbf{x}} G(\mathbf{x}, t | \mathbf{x}_o, t_o) = 0 \quad \mathbf{x} \in \partial\mathcal{R}_{\text{bot}}. \quad (16.191d)$$

### Boundary value problem for the boundary propagator

Following the more detailed presentation in VOLUME 1 for the diffusion equation, we are led to the following boundary value problem satisfied by the boundary propagator

$$\partial_t[\rho G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o)] + \nabla_{\mathbf{x}} \cdot [\rho \mathbf{v}^\dagger G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o) - \rho \mathbf{K} \cdot \nabla_{\mathbf{x}} G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o)] = 0, \quad \mathbf{x} \in \mathcal{R} \quad (16.192a)$$

$$G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o) = 0, \quad \mathbf{x} \notin \partial\mathcal{R}, t \leq t_o \quad (16.192b)$$

$$G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o) = \delta(t - t_o) \delta^{(2)}(\mathbf{x} - \mathbf{x}_o), \quad \mathbf{x}, \mathbf{x}_o \in \partial\mathcal{R}. \quad (16.192c)$$

The boundary propagator acts as the mediator between boundary data,  $\sigma$ , and interior points, with the transfer of information realized through both advection and diffusion. A focus on the boundary propagator rather than the Green's function allows us to reduce the dimensionality of the problem by placing source points only on the boundary rather than both the boundary and throughout the interior. Also recall our discussion in VOLUME 1, where we argued that the boundary propagator can be considered the **impulse response function** for spatially distributed sources. Here, the mediation of the Dirac delta boundary sources is performed by advection plus diffusion, whereas in VOLUME 1 we only considered linear damping and diffusion.

### Normalization of the boundary propagator

As seen in Chapter 15, diffusion acts to smooth all structure in the tracer field. Hence, if the boundary data is a uniform constant,  $\sigma = \sigma_{\text{const}}$ , then given sufficient time the tracer concentration will equal to this constant,  $C = \sigma_{\text{const}}$ . This steady state result is independent of details for the velocity field and the diffusivity tensor, with details of advection and the diffusivity acting only to modify the transient behavior during equilibration. Assuming we wait long enough, or equivalently that the initial condition occurs infinitely far in the past, then the tracer concentration solution (16.189) leads to the normalization of the boundary propagator

$$\lim_{t_{\text{init}} \rightarrow -\infty} \int_{t_{\text{init}}}^t \left[ \oint_{\partial\mathcal{R}} G^{\text{bp}}(\mathbf{x}, t | \mathbf{x}_o, t_o) d\mathcal{S}_0 \right] dt_o = 1 \quad \text{for } \mathbf{x} \in \mathcal{R}. \quad (16.193)$$

This normalization holds for all field points,  $\mathbf{x}$ , within the region. Even though this condition was derived by assuming the special case of constant boundary data, it holds in general since the Green's function, and by extension the boundary propagator, are independent of the boundary data prescribed for the tracer concentration.

#### 16.9.11 Comments

A careful reading of this section is best considered along with the analogous, yet more detailed, treatment of the diffusion equation Green's function in VOLUME 1. Furthermore, these two presentations offer technical introductions to the review paper from [Haine et al. \(2025\)](#), who synthesize the variety of Green's function methods for studying age and ventilation time scales in geophysical flows, with particular focus on the ocean. The Green's function method is technically challenging in realistic numerical models due to the extra space-time dimensions needed to hold the Green's function,  $G(\mathbf{x}, t | \mathbf{x}_o, t_o)$ , in computer memory. As a result, realistic applications typically make approximations to reduce the dimensionality, such as by assuming steady state or by focusing on boundary propagators. Hence, there has yet to be a calculation of the full Green's function for a realistic ocean, with that calculation awaiting bigger computers and/or novel methods to side-step the nontrivial memory requirements. [Haine et al. \(2025\)](#) provide further discussion of numerical considerations.



## 16.10 Exercises

### EXERCISE 16.1: ONE-DIMENSIONAL ADVECTION

Consider the advection equation in one space dimension without boundaries

$$(\partial_t + u \partial_x) C = 0 \quad (16.194a)$$

$$C(x, z, t = 0) = C_o \cos(k x) \quad (16.194b)$$

$$u(z, t) = \alpha z \cos(\omega t). \quad (16.194c)$$

The specified zonal velocity is non-divergent, oscillatory in time, and vertically sheared

$$\partial_z u = \alpha \cos(\omega t), \quad (16.195)$$

with  $\omega$  the angular frequency of the temporal oscillations. What is the tracer concentration at times  $t > 0$ ? Hint: make use of the exact solution given by equation (16.15).

**EXERCISE 16.2: SKEW FLUX FOR OCEAN MESOSCALE EDDIES**

Consider a Boussinesq ocean description of a middle-latitude mesoscale ocean eddy respecting geostrophic balance on an f-plane. In this case, the horizontal eddy-induced velocity at the ocean surface is non-divergent

$$\mathbf{u}^* = \nabla \times \hat{\mathbf{z}}\psi. \quad (16.196)$$

In this equation, the geostrophic streamfunction is given by

$$\psi = -\hat{\mathbf{z}} g \eta / f, \quad (16.197)$$

with  $f$  the Coriolis parameter,  $g$  the gravitational acceleration, and  $\eta$  the sea level undulation associated with the eddy. Since the fluid is incompressible, the mass transport equals to the volume transport times a constant reference density,  $\rho_0$ .

- (a) Determine the skew diffusion tensor (16.47).
- (b) Determine the skew tracer flux (16.61a).

**EXERCISE 16.3: EVOLUTION OF TRACER CENTER OF MASS IN A STATIC MATERIAL DOMAIN**

The exercise introduces us to how the tracer center of mass evolves within a Boussinesq ocean. We define the tracer center of mass as

$$\langle \mathbf{x} \rangle^C = \frac{\int \mathbf{x} C \, dV}{\int C \, dV}, \quad (16.198)$$

with  $C$  the tracer concentration,  $\mathbf{x}$  the coordinate of a point in the fluid, and integration is over the full fluid domain. For example, with a spherically symmetric tracer cloud, the center of mass position is at the sphere's center. The center of mass position is not necessarily where the largest tracer concentration sits, in the same way that the center of mass of a massive object is not necessarily where the object is most dense. For example, a hollow spherical shell has its center of mass at the center of the sphere, even though there is no mass there.

For this exercise, assume the fluid is within a domain whose static boundaries are either material (no normal component to the boundary flux) or periodic. Hence, the total fluid volume and total tracer content remain constant

$$\mathcal{V} = \int dV \quad \text{and} \quad \mathcal{C} = \int C \, dV. \quad (16.199)$$

Furthermore, assume the region boundaries are static, so that the time derivative commutes with the spatial integral

$$\frac{d}{dt} \int \varphi \, dV = \int \partial_t \varphi \, dV, \quad (16.200)$$

which follows since the region boundaries are assumed to be static. Equivalently, since the region under consideration is material (no matter crosses the boundaries), we can make use of Reynolds transport theorem from VOLUME 1 to write

$$\frac{d}{dt} \int \varphi \, dV = \int \frac{D\varphi}{Dt} \, dV. \quad (16.201)$$

Finally, note that we considered the case of tracer concentration moments in Section 15.2.4 when studying the diffusion equation.

- (a) Consider a tracer concentration whose tendency at a point in space is affected only by advection

$$\frac{DC}{Dt} = 0 \implies \partial_t C + \nabla \cdot (\mathbf{v} C) = 0, \quad (16.202)$$

with  $\mathbf{v}$  a non-divergent velocity,  $\nabla \cdot \mathbf{v} = 0$ . Show that the tracer center of mass position evolves according to the tracer center of mass velocity

$$\frac{d\langle \mathbf{x} \rangle^C}{dt} = \langle \mathbf{v} \rangle^C, \quad (16.203)$$

where the tracer center of mass velocity is given by

$$\langle \mathbf{v} \rangle^C = \frac{\int \mathbf{v} C dV}{\int C dV} = \frac{1}{C} \int \mathbf{v} C dV. \quad (16.204)$$

- (b) Consider a tracer concentration whose tendency at a point in space affected only by diffusion

$$\partial_t C = \nabla \cdot (K \cdot \nabla C), \quad (16.205)$$

where  $K = K(\mathbf{x}, t) > 0$  is a kinematic diffusivity (physical dimensions of squared length per time), and which is assumed to vanish at the domain boundaries. Show that the tracer center of mass drifts up the diffusivity gradient

$$\frac{d\langle \mathbf{x} \rangle^C}{dt} = \langle \nabla K \rangle^C. \quad (16.206)$$

Hint: use the product rule and drop boundary terms.

- (c) Consider an initial tracer concentration that is a function only of latitude,

$$C(x, y, z, t = 0) = C_0(y), \quad (16.207)$$

and assume a smooth spherical domain. Assume the diffusivity,  $K$ , is a turbulent diffusivity proportional to the eddy kinetic energy of the flow, so that large diffusivity occurs in regions with large eddy activity; i.e., there is a lot of turbulent mixing where turbulence is active. Introduce an stirring from the eddies that breaks the zonal symmetry. Qualitatively discuss the process whereby this turbulent diffusive mixing causes the tracer center of mass to drift towards the turbulent region.

- (d) Generalize the result from part (b) to the case of the diffusion equation

$$\partial_t C = \nabla \cdot (\mathbf{K} \cdot \nabla C) = \partial_p (K^{pq} \partial_q C), \quad (16.208)$$

where  $\mathbf{K}$  is a second order symmetric diffusion tensor.

#### EXERCISE 16.4: EVOLUTION OF TRACER CENTER OF MASS IN MOVING REGION

Consider a finite region of fluid with fixed mass that is moving with the fluid velocity field,  $\mathcal{R}(\mathbf{v})$ . The fluid is assumed to have a tracer whose concentration is affected by an irreversible process so that

$$\frac{DC}{Dt} = \dot{C} \neq 0. \quad (16.209)$$

For example,  $\dot{C}$  may represent a diffusive process, in which case the tracer content within the region changes due to diffusion of tracer across the region boundary.

Determine the evolution equation for the tracer center of mass position

$$\langle \mathbf{x} \rangle^C = \frac{\int_{\mathcal{R}(\mathbf{v})} \mathbf{x} C \rho dV}{\int_{\mathcal{R}(\mathbf{v})} C \rho dV}. \quad (16.210)$$

Hint: the region under consideration is moving with the fluid and has constant mass. Although the region boundaries are not material, we can make use of Reynold's transport theorem from VOLUME 1 since the region has a constant mass. Consequently, we can set

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \psi \rho dV = \int_{\mathcal{R}(\mathbf{v})} \frac{D\psi}{Dt} \rho dV. \quad (16.211)$$

#### EXERCISE 16.5: STEADY TWO DIMENSIONAL ADVECTION-DIFFUSION

Consider the steady state advection-diffusion equation for a scalar field,  $Q$ , in a two dimensional non-divergent flow

$$\nabla \cdot (\mathbf{u} Q) = \nabla \cdot (\mathbf{K} \cdot \nabla Q) \quad \text{with} \quad \mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi, \quad (16.212)$$

and  $\mathbf{K}$  a diffusion tensor. Show that when evaluated along a contour of constant  $Q$  we can write

$$-(\hat{\mathbf{n}} \cdot \nabla Q) (\hat{\mathbf{t}} \cdot \nabla \psi) = \nabla \cdot (\mathbf{K} \cdot \nabla Q) \quad (16.213)$$

where  $\hat{\mathbf{t}}$  is the unit tangent along the contour and  $\hat{\mathbf{n}}$  is a unit vector pointing to the left of the tangent. Assuming  $\hat{\mathbf{n}} \cdot \nabla Q \neq 0$ , this equation takes on the form

$$\hat{\mathbf{t}} \cdot \nabla \psi = -\frac{\nabla \cdot (\mathbf{K} \cdot \nabla Q)}{(\hat{\mathbf{n}} \cdot \nabla Q)}, \quad (16.214)$$

which provides a means to integrate the streamfunction,  $\psi$ , along contours of constant  $Q$ .

If  $Q$  is the quasi-geostrophic potential vorticity (see VOLUME 3), then contours of constant  $Q$  are known as *geostrophic contours*. Within this context, [Rhines and Holland \(1979\)](#) made use of the identity (16.214) in their study of ocean circulation in the presence of eddy diffusion of potential vorticity.

Hint: write the advection operator as a Jacobian and make use of the Jacobian exercise in the vector calculus chapter in VOLUME 1.

#### EXERCISE 16.6: DISTRIBUTION OF ONE TRACER WITH RESPECT TO ANOTHER

Consider two tracers,  $\psi$  and  $B$ , that satisfy the advection-diffusion equation with the same diffusion tensor

$$\rho \frac{D\psi}{Dt} = \nabla \cdot (\rho \mathbf{K} \cdot \nabla \psi) \quad (16.215a)$$

$$\rho \frac{DB}{Dt} = \nabla \cdot (\rho \mathbf{K} \cdot \nabla B). \quad (16.215b)$$

Having access to two tracers allows us to diagnose certain properties of the flow, both in geographical/depth space as well as in the space defined by the tracers. We here study how the tracer  $B$  is distributed within layers defined by  $\psi$ , and how that distribution evolves in time. These considerations are partly motivated by the work of [Ruan and Ferrari \(2021\)](#), who assumed  $B$  to be buoyancy (with a linear equation of state). Whereas [Ruan and Ferrari \(2021\)](#) assumed a Boussinesq ocean with a constant scalar diffusivity, here we generalize to the non-Boussinesq case with a flow-dependent diffusion tensor,  $\mathbf{K}$ , which is a symmetric and positive-definite second order tensor.

(a) Derive the following identity

$$\rho \frac{D(\psi B^\Gamma)}{Dt} = \nabla \cdot (B^\Gamma \rho \mathbf{K} \cdot \nabla \psi - \psi \rho \mathbf{K} \cdot \nabla B^\Gamma) + \Gamma \psi \nabla B^{\Gamma-1} \cdot \rho \mathbf{K} \cdot \nabla B + 2\Gamma \psi B^{\Gamma-1} \rho \dot{B}, \quad (16.216)$$

where  $B^\Gamma$  is  $B$  raised to the integer power  $\Gamma$ , and where we made use of the shorthand

$$\dot{B} = \frac{DB}{Dt}. \quad (16.217)$$

Show all relevant steps in the derivation of equation (16.216). Hint: as an optional warm-up, derive the special case with  $\Gamma = 1$

$$\rho \frac{D(\psi B)}{Dt} = \nabla \cdot (B \rho \mathbf{K} \cdot \nabla \psi - \psi \rho \mathbf{K} \cdot \nabla B) + 2\psi \rho \dot{B} \quad (16.218)$$

and then the case with  $\Gamma = 2$

$$\rho \frac{D(\psi B^2)}{Dt} = \nabla \cdot (B^2 \rho \mathbf{K} \cdot \nabla \psi - \psi \rho \mathbf{K} \cdot \nabla B^2) + 2\psi \nabla B \cdot \rho \mathbf{K} \cdot \nabla B + 4\psi B \rho \dot{B}. \quad (16.219)$$

(b) Introduce the  $\psi$ -weighted mean of an arbitrary field,  $\bar{\Gamma}$ , according to

$$\bar{\Gamma} \equiv \frac{\int_{\mathcal{R}} \Gamma \psi \rho dV}{\int_{\mathcal{R}} \psi \rho dV}. \quad (16.220)$$

Furthermore, assume all boundaries to the domain are material, which means that the domain matter content is fixed in time

$$\frac{d}{dt} \int_{\mathcal{R}} \rho dV = 0 \quad \text{and} \quad \frac{d}{dt} \int_{\mathcal{R}} \psi \rho dV = 0 \quad \text{and} \quad \frac{d}{dt} \int_{\mathcal{R}} B \rho dV = 0. \quad (16.221)$$

Make use of equation (16.218) to derive the following identity

$$\frac{d\bar{B}}{dt} = 2\bar{B}, \quad (16.222)$$

and offer some discussion.

#### EXERCISE 16.7: EVOLUTION OF TRACER MOMENTS

In Section 15.8.3 we studied how tracer diffusion affects tracer moments. Here we consider the combined effects of advection and diffusion. We assume the boundaries are insulating (i.e., zero normal boundary flux of tracer) so that  $\mathbf{J} \cdot \hat{\mathbf{n}} = 0$  with  $\hat{\mathbf{n}}$  the outward normal at the boundary. We also assume there is no matter crossing the boundary, so that  $(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0$ , where  $\mathbf{v}^{(b)}$  is the velocity of a point stuck to the boundary. Correspondingly, the total fluid mass in the domain remains fixed

$$M = \int \rho dV \quad \text{with} \quad \frac{dM}{dt} = 0, \quad (16.223)$$

so that the domain is material since we assume no exchange of mass or tracer across the boundaries. These assumptions allow us to focus on the effects from tracer diffusion and advection within the domain interior.

(a) DOMAIN AVERAGED TRACER CONCENTRATION: The domain averaged tracer concentration

is defined by

$$\bar{C} = \frac{\int_{\mathcal{R}} C \rho dV}{M}. \quad (16.224)$$

Show that its time derivative vanishes.

- (b) TRACER VARIANCE WITHIN THE DOMAIN: The variance of the tracer concentration is defined by

$$\text{var}(C) \equiv \frac{\int_{\mathcal{R}} (C - \bar{C})^2 \rho dV}{M} = \bar{C^2} - \bar{C}^2 \geq 0. \quad (16.225)$$

The tracer variance measures the deviation of the tracer concentration relative to the domain averaged concentration. Since the domain average tracer concentration remains fixed in time, the time change of the variance is given by

$$\frac{d[\text{var}(C)]}{dt} = \frac{d\bar{C^2}}{dt}. \quad (16.226)$$

Thus, it is common to refer to  $\bar{C^2}$  as the tracer variance, though strictly speaking only time derivatives of  $\bar{C^2}$  and  $\text{var}(C)$  are equal as per equation (16.226). Show that

$$\frac{d[\text{var}(C)]}{dt} = \frac{d\bar{C^2}}{dt} \leq 0, \quad (16.227)$$

with this inequality determined solely by diffusion, whereas advection has no impact on the variance.

- (c) DIFFUSION OF ARBITRARY TRACER MOMENTS: Prove that

$$\frac{d\bar{C^\Gamma}}{dt} = \Gamma(\Gamma - 1) \int C^{\Gamma-2} \nabla C \cdot \mathbf{J} dV \leq 0. \quad (16.228)$$

For  $\Gamma = 0$  we have an expression of mass conservation for the domain, whereas  $\Gamma = 1$  is an expression of tracer conservation. The case of  $\Gamma = 2$  yields the tracer variance result (16.227).

Hint: This exercise reveals that tracer moments evolve solely through the effects of diffusion, whereas advection does not touch the tracer moments. The goal of this exercise is to emphasize these results by working through the details, which are largely identical to those presented in Section 15.8.3 when studying diffusion alone.

#### EXERCISE 16.8: COINCIDENCE INDEX

Consider two conservative tracers,  $A$  and  $B$ , that satisfy the advection-diffusion equation

$$\rho \frac{DA}{Dt} = \nabla \cdot (\rho \mathbf{K} \cdot \nabla A) \quad \text{and} \quad \rho \frac{DB}{Dt} = \nabla \cdot (\rho \mathbf{K} \cdot \nabla B), \quad (16.229)$$

where  $\mathbf{K}$  is the diffusion tensor. Derive the evolution equation for the product,  $C = A B$ , which [Li et al. \(2026\)](#) refer to as the coincidence index. Hint: this exercise is a simpler form of the more general case in Exercise 16.6.



## Chapter 17

### EDDY-MEAN INTERACTION FOR SCALARS

Geophysical fluid flows have multiple scales in both space and time. In the analysis of these flows, it is useful to seek a description that decomposes fluid properties into a mean component and a departure from the mean, with the departure referred to as an [eddy](#). [Rectification](#) occurs if oscillatory or turbulent eddying motions produce a mean flow through nonlinear terms in the governing equations. The mean field can be defined in many fashions with subjective choices based on particulars of the flow and the analysis goals. In turn, the chosen definition for the mean affects what we refer to as an eddy. Quite generally, eddy fluctuations take the form of transient linear waves, nonlinear and/or breaking waves, coherent structures, and/or a chaotic/turbulent soup of eddying features. In this chapter we develop a kinematic framework motivated by the analysis of scalar transport induced by small amplitude wave-like eddying features. This framework can also be used for turbulent processes and their parameterizations, though we do not pursue that application here.

We consider two related kinematic methods to decompose the flow into a mean and eddy. The first is the [generalized Lagrangian mean \(GLM\)](#), which is a hybrid Eulerian/Lagrangian method that introduces an Eulerian disturbance field to measure the position of a fluid particle relative to its mean position ([Andrews and McIntyre, 1978a,b](#); [Bühler, 2014a](#); [Gilbert and Vanneste, 2025](#)). We access a small portion of the GLM framework, focusing on the needs for describing the kinematics of eddy scalar fluxes following from [Middleton and Loder \(1989\)](#). A more thorough treatment that considers the momentum and vorticity equations is outside the scope for this chapter. The second kinematic method makes use of the [isopycnal](#) vertical coordinate, which is Eulerian in the horizontal yet Lagrangian in the vertical. Isopycnal coordinates are frequently used to describe how ocean mesoscale eddies affect stratification and tracer transport in stably stratified flows, with a thorough treatment of these coordinates presented in Part IV of this volume.

We emphasize our limited exploration of GLM, with a far more thorough presentation given in the primary literature and follow on work cited in the previous paragraph. Indeed, we could have avoided GLM altogether in this chapter given that we limit attention to small amplitude disturbances. However, the conceptual framework offered by the GLM is very useful to be exposed to since it is extensive within the literature, even for small amplitude disturbances. Furthermore, as we propose, it provides a fruitful intellectual introduction to isopycnal averaging methods, which form a central framework for parameterizing ocean mesoscale eddies. So even if one remains skeptical of the utility of GLM for their particular studies, its conceptual and technical features remain pervasive in geophysical fluid mechanics.

---

## CHAPTER GUIDE

This chapter relies on an understanding of the maths and physics of the advection-diffusion equation explored in Chapter 16. We focus on non-divergent flows with kinematics presented in VOLUME 1 as applicable to the Boussinesq ocean studied in VOLUME 2. Generalizations to non-Boussinesq flows are straightforward, with examples provided by *Griffies and Greatbatch* (2012). The kinematics of isopycnal fluid layers in a perfect fluid (Sections 17.5 and 17.6) are posed using the isopycnal vertical coordinates detailed in Chapter 14, which also develops the shallow water thickness weighted averaged equations. Our presentation of the isopycnal eddy-mean flow decomposition follows the methods developed by *McDougall and McIntosh* (2001) and summarized in Chapter 9 of *Griffies* (2004). A directly related approach is considered in Chapter 14 for the stacked shallow water equations, in which we develop the thickness weighted tracer, momentum, and vorticity budgets (see also *Young* (2012) and *Jansen et al.* (2024)).

Throughout this chapter we assume Cartesian tensors in the horizontal directions, which is sufficient for the isopycnal averaging in Sections 17.5 and 17.6. However, to extend the full GFM theory to arbitrary horizontal coordinates requires more general mathematical tools, such as those reviewed by *Gilbert and Vanneste* (2025).

---

<b>17.1</b>	<b>Loose threads</b>	<b>593</b>
<b>17.2</b>	<b>Methods for decomposing the flow</b>	<b>593</b>
17.2.1	Eulerian and Lagrangian means	593
17.2.2	Operators used to compute means	593
17.2.3	Reynolds property	594
<b>17.3</b>	<b>Generalized Lagrangian mean for scalar fields</b>	<b>595</b>
17.3.1	Motivation from the advection equation	595
17.3.2	Length scales and the small parameter	597
17.3.3	Decomposing the particle trajectory	597
17.3.4	GLM and the Stokes mean	598
17.3.5	Stokes drift	600
17.3.6	An example wave	600
17.3.7	GLM with a materially constant scalar	602
17.3.8	Further study	603
<b>17.4</b>	<b>Eddy tracer fluxes and particle displacements</b>	<b>603</b>
17.4.1	Particle displacements and eddy tracer fluxes	604
17.4.2	Decomposing into symmetric and skew symmetric fluxes	604
17.4.3	The symmetric tracer flux	605
17.4.4	Skew, advective, and rotational tracer fluxes	605
17.4.5	What does a point measurement yield?	606
17.4.6	Further study	607
<b>17.5</b>	<b>Volume transport in an isopycnal ensemble</b>	<b>607</b>
17.5.1	Specifying the vertical position	608
17.5.2	Modified mean density	608
17.5.3	Isopycnal mean and modified mean	609
17.5.4	Transformed residual mean (TRM)	610
17.5.5	Depth integrated TRM transport	611
17.5.6	Quasi-Stokes transport	612
17.5.7	Three-component TRM velocity	612
17.5.8	Volume conservation and the thickness equation	613

---

17.5.9	Ensemble mean thickness equation . . . . .	614
17.5.10	Ensemble mean kinematics in geopotential coordinates . . . . .	615
17.5.11	Approximate $\varrho$ -ensemble kinematics in geopotential coordinates . . . . .	615
17.5.12	Further study . . . . .	617
<b>17.6</b>	<b>Isopycnal mean tracer equation . . . . .</b>	<b>617</b>
17.6.1	Thickness weighted average . . . . .	617
17.6.2	Isopycnal mean thickness weighted tracer equation . . . . .	618
17.6.3	Subgrid scale tracer transport tensor . . . . .	619
17.6.4	Summary of the tracer parameterization problem . . . . .	619
17.6.5	Comments . . . . .	620
<b>17.7</b>	<b>Exercises . . . . .</b>	<b>620</b>

---

## 17.1 Loose threads

- Formulate the thickness weighted tracer variance equations as in Appendix A of *Pudig et al. (2026)*.

## 17.2 Methods for decomposing the flow

There are many methods used to decompose flow into mean and eddy components. There is no unique or objective choice, with each offering pros and cons. Furthermore, there is dependence on the coordinate choice used to represent the fields. So although physics is independent of coordinates, it is an unfortunate fact that computing a mean results in a flow decomposition that inherits particulars of the subjectively chosen coordinate representation.

### 17.2.1 Eulerian and Lagrangian means

At any point in space and time, we can decompose a field into its **Eulerian mean**,  $\bar{\chi}(\mathbf{x}, t)$ , and departures from the mean that are referred to as the eddy fluctuation,  $\chi'(\mathbf{x}, t)$ :

$$\chi(\mathbf{x}, t) = \bar{\chi}(\mathbf{x}, t) + \chi'(\mathbf{x}, t). \quad (17.1)$$

This choice is the default when working with Eulerian coordinates. It also accords with common laboratory or field methods that provide point measurements. The complement approach arises when representing fields using Lagrangian coordinates, in which we decompose a field into its **Lagrangian mean** and the departure from the mean. This choice is convenient if having knowledge of fluid particle trajectories, such as in simulations or field platforms (e.g., floats or balloons) that move with the fluid. Finally, there are hybrid approaches that we encounter in this chapter that are motivated by the generalized Lagrangian mean.

### 17.2.2 Operators used to compute means

We here summarize a variety of mean operators, whether realized using Eulerian coordinates, Lagrangian coordinates, or hybrids, with the following list non-exhaustive.

#### Time mean

If the mean operator is based on a long time average, then the mean fields are either time independent or they vary over a far longer time scale than the unaveraged fields. This is a

common operator when interest is focused on the long term mean fluid properties, such as in climate science where 30 years is often considered the minimum time average for developing climate statistics.

### Phase average

Rather than a time mean, we may choose to average over the phase (or period) of a wave. A **phase average** is particularly relevant when the fluctuating field involves quasi-linear waves. Hence, we make extensive use of phase averaging when studying waves in VOLUME 4.

### Zonal mean

The **zonal mean** refers to a spatial averaging operation computed by a line integral of a field over the longitudinal extent of the domain, and then dividing by the zonal length of the domain. The zonal mean is particularly relevant for regions that are zonally periodic, such as the atmosphere as well as the ocean in the latitudes around the Drake Passage. The resulting zonal mean field is independent of longitude.

### Coarse graining

Coarse graining refers to the process of spatial averaging over specific scales through use of a filtering operator. Coarse graining originates from renormalization methods in theoretical physics and engineering, and it provides a useful method to control the length scales over which the flow is decomposed. [Buzzicotti et al. \(2023\)](#) reviewed the method and its use with global ocean models.

### Ensemble mean

An **ensemble mean** is computed through generating an ensemble of flow realizations that differ in a controlled manner, such as through the initial conditions, and then averaging over the many (formally infinite) realizations. This method is theoretically quite convenient since space and time operators commute with the ensemble mean operator, which is not a property shared by the other operators listed above. It is for this reason we make use of the ensemble mean in this chapter in order to avoid a variety of technical details that arise with the other operators.

### 17.2.3 Reynolds property

If a mean operator satisfies the following properties then it is said to provide a **Reynolds decomposition**

$$\bar{\chi}' = 0 \quad \text{and} \quad \bar{\bar{\chi}} = \bar{\chi} \quad \text{and} \quad \bar{c\bar{\chi}} = c\bar{\chi}. \quad (17.2)$$

The first equation says that the mean of an eddy fluctuation vanishes. The second says that the mean of a mean returns the mean. The final equality says that a constant,  $c$ , commutes with the mean operator. Notably, some or all of these properties are not satisfied by certain operators used for eddy-mean decompositions. However, in the following we assume they are satisfied by our mean operator.

A Reynolds average acting on a linear equation means that both the mean and eddy quantity satisfy the linear equation. In particular, consider the non-divergence condition for a Boussinesq ocean flow,  $\nabla \cdot \mathbf{v} = 0$ . Taking the mean of this equation renders

$$\nabla \cdot \mathbf{v} = 0 \implies \nabla \cdot \bar{\mathbf{v}} = 0, \quad (17.3)$$

so that the mean velocity is non-divergent. Since both the unaveraged velocity and the mean velocity are non-divergent, it follows that the eddy velocity is also non-divergent

$$\nabla \cdot (\bar{\mathbf{v}} + \mathbf{v}') = \nabla \cdot \mathbf{v}' = 0. \quad (17.4)$$

## 17.3 Generalized Lagrangian mean for scalar fields

We here consider basic elements of **generalized Lagrangian mean (GLM)** theory. GLM is a hybrid between Lagrangian and Eulerian descriptions of fluid motions, so that it might be more appropriate to refer to it as the “hybrid Lagrangian-Eulerian mean theory”. The GLM and the Eulerian mean for a fluid property are generally distinct, with their difference referred to as the **Stokes mean**

$$\text{generalized Lagrangian mean} = \text{Eulerian mean} + \text{Stokes mean}. \quad (17.5)$$

This name is motivated from the **Stokes drift** encountered in our study of fluid kinematics in VOLUME 1, with Stokes drift referring to the difference between the Lagrangian mean velocity and Eulerian mean velocity. Note that the literature sometimes refers to the Stokes mean as the “Stokes correction”. We avoid that terminology in order to avoid the spurious notion that one type of mean operator is more correct than the other. Instead, a mean operator is subjectively chosen based on its suitability to a particular question.

### 17.3.1 Motivation from the advection equation

For much of this chapter we are concerned with a materially constant scalar field

$$\frac{D\chi}{Dt} = (\partial_t + \mathbf{v} \cdot \nabla)\chi = 0, \quad (17.6)$$

and the description of  $\chi$  when it, and the non-divergent velocity, are decomposed into a mean and eddy. The scalar field,  $\chi$ , remains constant when following trajectories of fluid particles. As described in the kinematics chapters in VOLUME 2, each fluid particle trajectory forms an integral curve of the velocity field. Furthermore, the congruence of such integral curves forms the **motion field**,  $\boldsymbol{\varphi}(\mathbf{a}, T)$ . The motion field generates a nonlinear time dependent and invertible **flow map** that continuously and smoothly reshapes the fluid continuum. We write the motion field as a function of the material space coordinate,  $\mathbf{a}$ , and the material time coordinate,  $T$ .<sup>1</sup>

#### Eulerian mean

An Eulerian mean operator considered in Section 17.2 leads to the following mean version of the advection equation (17.6), here written both in the advective form and the flux form

$$\partial_t \bar{\chi} + \bar{\mathbf{v}} \cdot \nabla \bar{\chi} = -\overline{\mathbf{v}' \cdot \nabla \chi'} \quad \text{advective form} \quad (17.7a)$$

$$\partial_t \bar{\chi} + \nabla \cdot (\bar{\mathbf{v}} \bar{\chi}) = -\nabla \cdot (\overline{\mathbf{v}' \chi'}) \quad \text{flux form.} \quad (17.7b)$$

---

<sup>1</sup>Throughout this book, we are considering all speeds (particles and waves) to be far slower than light speed. Hence, the material time coordinate equals to the Eulerian time,  $T = t$ . However, we use a distinct symbol to indicate what spatial coordinates are held fixed when taking the Eulerian time derivative,  $\partial_t$ , versus the Lagrangian time derivative,  $\partial_T$ .

The equations are equivalent since both the Eulerian mean velocity and the eddy velocity are non-divergent, as shown in Section 17.2.3 for a Reynolds decomposition. Evidently, the Eulerian mean,  $\bar{\chi}$ , is not materially constant when following trajectories defined by the Eulerian mean velocity,  $\bar{v}$ . The reason is that flow following such trajectories encounters a source term,  $-\bar{v}' \cdot \nabla \chi'$ , due to correlations between the velocity and tracer. Hence, the  $\bar{\chi}$  evolution equation is mathematically distinct from that of  $\chi$ . Furthermore, when given information only about the mean fields, then we must develop a closure for the unresolved correlation. This point touches on the largely unsolved **turbulence closure** problem pervasive in fluid mechanics.

### Lagrangian mean

An alternative approach makes use of the Lagrangian reference frame, where we sample the fluid property along a fluid particle trajectory<sup>2</sup>

$$\chi^L(\mathbf{a}, T) = \text{property } \chi \text{ following the fluid particle trajectory, } \mathbf{X}(\mathbf{a}, T). \quad (17.8)$$

In Lagrangian coordinates, the material constancy equation (17.6) becomes

$$\partial_T \chi^L(\mathbf{a}, T) = 0. \quad (17.9)$$

Consider a mean operator computed as an average over a region of material space. For example, if  $\mathbf{a}$  is the initial fluid particle position, then a mean coordinate,  $\bar{\mathbf{a}}$ , and corresponding mean field,  $\bar{\chi}^L$ , render a coarse-graining over the initial positions. Since each member of the Lagrangian mean satisfies equation (17.9), so too does the Lagrangian mean

$$\partial_T \bar{\chi}^L(\bar{\mathbf{a}}, T) = 0. \quad (17.10)$$

Notably, this equation retains the simplicity of the unaveraged version.

The elegance of the Lagrangian mean equation (17.10) hides a major limitation when applied to fluid flow. Namely, the flow map can become quite complex for turbulent flows, and even for laminar flow (e.g., chaotic advection). Such complexity makes a fully Lagrangian approach technically impractical. We thus seek something in between the Eulerian and Lagrangian approach, which leads to the generalized Lagrangian mean.

### Generalized Lagrangian mean

The **generalized Lagrangian mean (GLM)** melds elements of the Eulerian and the Lagrangian approaches. It does so by retaining the Eulerian space and time coordinates, plus by adding an additional field,  $\xi(\mathbf{x}, t)$ , that measures the position of a fluid particle relative to its mean position. Adding a new dynamical field is clearly a double-edged sword. On the positive side it provides the means to obtain Lagrangian information when using Eulerian kinematics. On the negative side we must concern ourselves with determining this new field and its evolution.

For the tracer equation, the net effect of the GLM approach is the mean scalar field that remains constant following trajectories that form integral curves of the GLM velocity

$$\partial_t \bar{\chi}^{(GLM)} + \bar{v}^{(GLM)} \cdot \nabla \bar{\chi}^{(GLM)} = 0. \quad (17.11)$$

---

<sup>2</sup>We here use the notation  $\mathbf{X}(\mathbf{a}, T)$  for a fluid particle, whereas we could alternatively make use of the motion field,  $\boldsymbol{\varphi}(\mathbf{a}, T)$ , introduced just following equation (17.6) and studied in VOLUME 1. The motion field and the particle trajectory are the same when fixing a particular particle label,  $\mathbf{a}$ . We choose the particle trajectory notion as it is somewhat more familiar.

We do not derive the GLM equation (17.11).<sup>3</sup> Even so, we motivate the GLM average from the analysis of small amplitude eddying motions, making use of this analysis as a framework for describing scalar transport.<sup>4</sup>

### 17.3.2 Length scales and the small parameter

We consider two length scales associated with an eddy. One length characterizes the size of the eddy whose length scale we write as  $\lambda$ . If the eddy is a monochromatic wave, then  $\lambda$  is its wave length. The second length scale characterizes the size of particle displacements relative to the mean position of the particle,  $|\xi|$ . In the following, we assume the particle displacements are small relative to  $\lambda$

$$|\xi| \ll \lambda \quad \text{small amplitude waves.} \quad (17.12)$$

We thus introduce the small non-dimensional ratio of length scales for the following analysis

$$\alpha = |\xi|/\lambda \ll 1. \quad (17.13)$$

### 17.3.3 Decomposing the particle trajectory

From the discussion of fluid particle trajectories given in VOLUME 1, we know that the trajectory of a fluid particle is determined by integrating the relation between the particle trajectory and the particle velocity

$$\left[ \frac{\partial \mathbf{X}(\mathbf{a}, T)}{\partial T} \right]_{\mathbf{a}} = \mathbf{v}[\mathbf{X}(\mathbf{a}, T)] \implies \mathbf{X}(\mathbf{a}, T) = \mathbf{X}(\mathbf{a}, 0) + \int_0^T \mathbf{v}[\mathbf{X}(\mathbf{a}, T')] dT'. \quad (17.14)$$

As expressed here, a trajectory describes the position of an individual fluid particle relative to a chosen origin. The congruence of all trajectories forms the integral curves of the velocity field, thus providing a Lagrangian description of the flow. The material coordinate,  $\mathbf{a}$ , distinguishes the continuum of fluid particles, thus making the trajectory a field in material space-time.

The GLM develops a hybrid Eulerian-Lagrangian method and it is motivated by linear or quasi-linear disturbances. Keeping this motivation in mind, we consider each point in space,  $\mathbf{x}$ , to be the mean position of a unique fluid particle. This identification acts to couple the Eulerian and Lagrangian kinematic descriptions. To realize this coupling, we introduce an Eulerian field,  $\xi(\mathbf{x}, t)$ , which measures the position of all fluid particles relative to their respective mean positions.<sup>5</sup> That is, at each point,  $\mathbf{x}$ , the disturbance field,  $\xi(\mathbf{x}, t)$ , measures the displacement of the particular fluid particle whose mean position is  $\mathbf{x}$ . At another point in space,  $\mathbf{y}$ , the same disturbance field measures the displacement of a distinct fluid particle whose mean position is  $\mathbf{y}$ .

As basic assumption of the method is that  $\xi(\mathbf{x}, t)$  provides a one-to-one invertible map<sup>6</sup> from the Eulerian mean spatial position of a fluid particle,  $\mathbf{x}$ , to the actual particle position,  $\mathbf{x} + \xi(\mathbf{x}, t)$ . We illustrate this decomposition in Figure 17.1. By definition, the Eulerian mean

---

<sup>3</sup>See section 10.2.2 of [Bühler \(2014a\)](#).

<sup>4</sup>It is useful to note that even if the Eulerian velocity is non-divergent, as for a Boussinesq ocean, the GLM velocity is divergent.

<sup>5</sup>We introduce a one-dimensional disturbance field when taking a Lagrangian perspective to derive the equations for acoustic waves in VOLUME 4. The GLM disturbance field,  $\xi(\mathbf{x}, t)$ , provides a three-dimensional generalization of that approach.

<sup>6</sup>Section 10.2.1 of [Bühler \(2014a\)](#) refers to  $\xi(\mathbf{x}, t)$  as the lifting map.

of the disturbance field vanishes

$$\overline{\xi(\mathbf{x}, t)} = 0, \quad (17.15)$$

which follows by considering each point in space to be the ensemble mean position of a unique fluid particle. Note that the Eulerian mean operator can be any of the operators (or others) satisfying the Reynold's decomposition property discussed in Section 17.2. Even so, we generally consider it to be an ensemble mean, both for conceptual and technical reasons. Finally, for flows that become quite nonlinear, the mapping loses its one-to-one invertible property, in which case the method breaks down.

Specification of  $\xi(\mathbf{x}, t)$  for large amplitude disturbances (i.e., nonlinear waves) requires the full machinery of GLM, which is beyond our scope. Instead, to expose the rudiments we assume small amplitude disturbances, for which the particle displacement amplitude is much smaller than the wavelength of the disturbance as given by the inequality (17.13). In this case the disturbance field is constructed by time integration of the eddy velocity field

$$\left[ \frac{\partial \xi(\mathbf{x}, t)}{\partial t} \right]_{\mathbf{x}} = \mathbf{v}'(\mathbf{x}, t) \implies \xi(\mathbf{x}, t) = \int_0^t \mathbf{v}'(\mathbf{x}, t') dt', \quad (17.16)$$

where we assume, for convenience, that  $\xi(\mathbf{x}, t=0) = 0$ . It follows that if the eddy velocity is non-divergent then so is the disturbance field

$$\nabla \cdot \mathbf{v}' = 0 \implies \nabla \cdot \xi = 0. \quad (17.17)$$

The definition (17.16) for the disturbance field is directly analogous to the particle trajectory position,  $\mathbf{X}(\mathbf{a}, T)$ , given by equation (17.14). However, there are important distinctions. Namely, the disturbance,  $\xi(\mathbf{x}, t)$ , is an Eulerian space-time field that measures the position of all fluid particles relative to their respective mean positions, with each Eulerian position,  $\mathbf{x}$ , corresponding to the mean position for a distinct fluid particle. In contrast, the particle position,  $\mathbf{X}(\mathbf{a}, T)$ , is a Lagrangian space-time field that is attached to each fluid particle and measures the position of particles relative to a chosen origin.

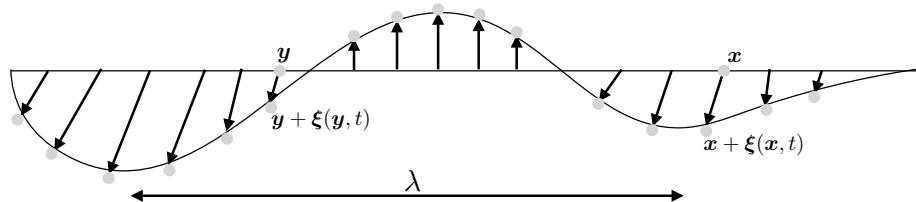


FIGURE 17.1: Illustrating the displacement of fluid particles in a wave disturbance. Each of the particle positions can be written,  $\mathbf{x}^{(\xi)}(\mathbf{x}, t) = \mathbf{x} + \xi(\mathbf{x}, t)$ , with the disturbance field,  $\xi(\mathbf{x}, t)$ , having a zero Eulerian mean,  $\bar{\xi} = 0$ . We here illustrate this role for the disturbance field for two particular fluid particles, one with mean position  $\mathbf{x}$  and the other with mean position  $\mathbf{y}$ . Small amplitude disturbances satisfy  $|\xi| \ll \lambda$ , where  $\lambda$  is the wavelength.

### 17.3.4 GLM and the Stokes mean

#### Generalized Lagrangian mean of a fluid property

The mean of a fluid property,  $\chi$ , is a function of how the property is sampled when computing the mean. For example, the mean sampled on a fluctuating fluid particle differs from the mean

sampled at the particle's mean position. Mathematically, this distinction implies that

$$\underbrace{\bar{\chi}(\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t))}_{\text{GLM}} \neq \underbrace{\bar{\chi}(\mathbf{x}, t)}_{\text{Eulerian}}, \quad (17.18)$$

where it is common to make use of the shorthand

$$\mathbf{x}^{(\xi)}(\mathbf{x}, t) \equiv \mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t) \quad (17.19)$$

for the position of the fluid particle. The mean operation,

$$\bar{\chi}^{(\text{GLM})}(\mathbf{x}, t) \equiv \bar{\chi}(\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t), t) = \bar{\chi}(\mathbf{x}^{(\xi)}, t), \quad (17.20)$$

defines the generalized Lagrangian mean for the fluid property,  $\chi$ . That is, the generalized Lagrangian mean is computed by evaluating the property,  $\chi$ , at the position of a fluid particle,  $\mathbf{x}^{(\xi)}(\mathbf{x}, t) = \mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t)$ , and then performing an Eulerian mean operation relative to the position,  $\mathbf{x}$ . In this manner we see how  $\mathbf{x}$  is both an arbitrary Eulerian field point and the mean position of a unique fluid particle,

$$\bar{\mathbf{x}}^{(\xi)} = \bar{\mathbf{x}} + \bar{\boldsymbol{\xi}}(\mathbf{x}, t) = \mathbf{x}. \quad (17.21)$$

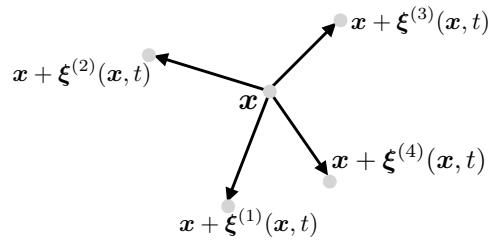


FIGURE 17.2: This figure depicts how we sample the fluid to compute the generalized Lagrangian mean of a fluid property. The central point,  $\mathbf{x}$ , is an arbitrary point in the fluid that is also the mean position of a unique fluid particle. We here interpret the Eulerian mean operation as an ensemble mean. We sample the fluid at the position of a fluid particle, and then do so again for another realization of the flow, and then again, thus building up an ensemble of flow samples, here depicting four. The ensemble mean of this sampling provides the generalized Lagrangian mean at the mean particle position,  $\mathbf{x}$ .

Since the Eulerian mean in equation (17.20) is computed with the particle disturbance as the argument to  $\chi$ , it produces a Lagrangian mean. That is, the disturbance field renders a Lagrangian mean without tracking fluid particle trajectories. Instead, we only need to keep track of the particle position relative to its mean position. The approach works so long as the disturbance field provides a one-to-one mapping between the mean particle position of all fluid particles and their instantaneous positions.

Recall from Section 17.2 that the traditional Eulerian mean is determined by evaluating  $\chi$  at a fixed Eulerian point in space

$$\bar{\chi}^{(\text{E})}(\mathbf{x}, t) \equiv \bar{\chi}(\mathbf{x}, t). \quad (17.22)$$

There is no consideration of instantaneous particle positions for this mean operation. Note that we typically drop the (E) superscript on the Eulerian mean to reduce clutter.

### Stokes mean

Following our discussion at the start of Section 17.3, we define the Stokes mean for a fluid property as the difference between the GLM and Eulerian mean

$$\bar{\chi}^{(S)}(\mathbf{x}, t) \equiv \bar{\chi}^{(GLM)}(\mathbf{x}, t) - \bar{\chi}(\mathbf{x}, t). \quad (17.23)$$

The Stokes mean arises from spatial structure in  $\chi(\mathbf{x}, t)$ , which in turn leads to differences in its mean depending on whether it is sampled on a fluid particle,  $\mathbf{x}^{(\xi)}(\mathbf{x}, t)$ , or sampled at the mean position of the fluid particle,  $\mathbf{x}$ .

We mathematically expose the origin of the Stokes mean by performing a Taylor series expansion around the mean particle position

$$\chi(\mathbf{x} + \boldsymbol{\xi}, t) = \chi(\mathbf{x}, t) + \boldsymbol{\xi} \cdot \nabla \chi(\mathbf{x}, t) + \frac{1}{2} \boldsymbol{\xi}^m \boldsymbol{\xi}^n \partial_m \partial_n \chi(\mathbf{x}, t) + \mathcal{O}(\alpha^3). \quad (17.24)$$

The non-dimensional ratio,  $\alpha = |\boldsymbol{\xi}|/\lambda \ll 1$ , was introduced in equation (17.13). It measures the ratio of the amplitude of particle displacements to the wavelength,  $\lambda$ , of fluctuations in the field  $\chi$  (see Figure 17.1). Taking the mean of equation (17.24) then leads to an expression for the Stokes mean

$$\bar{\chi}^{(S)}(\mathbf{x}, t) = \bar{\chi}^{(GLM)}(\mathbf{x}, t) - \bar{\chi}(\mathbf{x}, t) \quad (17.25a)$$

$$= \overline{\boldsymbol{\xi} \cdot \nabla \chi} + \frac{1}{2} \overline{\boldsymbol{\xi}^m \boldsymbol{\xi}^n \partial_m \partial_n \chi} + \mathcal{O}(\alpha^3). \quad (17.25b)$$

$$= \overline{\boldsymbol{\xi} \cdot \nabla \chi'} + \frac{1}{2} \overline{\boldsymbol{\xi}^m \boldsymbol{\xi}^n} \partial_m \partial_n \bar{\chi} + \mathcal{O}(\alpha^3), \quad (17.25c)$$

where we introduced the Eulerian fluctuation,

$$\chi'(\mathbf{x}, t) = \chi(\mathbf{x}, t) - \bar{\chi}(\mathbf{x}, t), \quad (17.26)$$

and all terms on the right hand side of equation (17.25c) are evaluated at  $(\mathbf{x}, t)$ . Observe that the Stokes mean (17.25c) is nonzero only starting at  $\mathcal{O}(\alpha^2)$ .

#### 17.3.5 Stokes drift

When  $\chi$  represents the velocity field, we refer to the Stokes mean as the **Stokes drift**, in which

$$\mathbf{v}^{(S)} = \overline{(\boldsymbol{\xi} \cdot \nabla) \partial_t \boldsymbol{\xi}} + \frac{1}{2} \overline{\boldsymbol{\xi}^m \boldsymbol{\xi}^n} \partial_m \partial_n \bar{\mathbf{v}} + \mathcal{O}(\alpha^3), \quad (17.27)$$

where we set  $\mathbf{v}' = \partial_t \boldsymbol{\xi}$  according to equation (17.16). Furthermore, the traditional case of Stokes drift occurs when the Eulerian mean field has a zero shear, so that<sup>7</sup>

$$\mathbf{v}^{(S)} = \overline{(\boldsymbol{\xi} \cdot \nabla) \partial_t \boldsymbol{\xi}} + \mathcal{O}(\alpha^3) \quad \text{if } \partial_n \bar{\mathbf{v}} = 0. \quad (17.28)$$

#### 17.3.6 An example wave

Consider a small amplitude wave with zero Eulerian mean,  $\bar{\mathbf{v}} = 0$ , along with the wave fields

$$\boldsymbol{\xi} = -\omega^{-1} \mathbf{U}(\mathbf{x}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (17.29a)$$

$$\mathbf{v}' = \partial_t \boldsymbol{\xi} = \mathbf{U}(\mathbf{x}) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (17.29b)$$

<sup>7</sup>We derive the Stokes' drift equation (17.28) using traditional Lagrangian methods when studying surface gravity waves in VOLUME 4.

$$\nabla v'^p = \nabla U^p \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \mathbf{k} U^p \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (17.29c)$$

$$\nabla \cdot \mathbf{v}' = (\nabla \cdot \mathbf{U}) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \mathbf{k} \cdot \mathbf{U} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (17.29d)$$

where  $\mathbf{U}$  is the velocity amplitude that is generally a function of space,  $\mathbf{k}$  is the wavevector, and  $2\pi/\omega$  is the wave period. The wave renders an oscillatory motion to fluid particles, with the disturbance field specifying the instantaneous position of fluid particles whose mean position is  $\mathbf{x}$ . The disturbance field and velocity field both have a zero mean when time integrated over a wave period

$$\frac{1}{2\pi/\omega} \int_0^{2\pi/\omega} \boldsymbol{\xi}(\mathbf{x}, t) dt = 0 \quad \text{and} \quad \frac{1}{2\pi/\omega} \int_0^{2\pi/\omega} \mathbf{v}'(\mathbf{x}, t) dt = 0. \quad (17.30)$$

To maintain a non-divergent eddy velocity at arbitrary times requires

$$\nabla \cdot \mathbf{v}' = 0 \implies \nabla \cdot \mathbf{U} = \mathbf{U} \cdot \mathbf{k} = 0. \quad (17.31)$$

As studied in VOLUME 4,  $\mathbf{U} \cdot \mathbf{k} = 0$  means that the wave is transverse, so that particle displacements arising from the wave are orthogonal to the wavevector.

### Stokes drift

Specializing to the wave velocity field (17.29b), substituting into the Stokes drift expression (17.28), and making use of an average over a wave period yields

$$\overline{(\boldsymbol{\xi} \cdot \nabla)v'^p} = \frac{U^p \mathbf{U} \cdot \mathbf{k}}{2\omega} \quad \text{and} \quad \bar{v} = 0. \quad (17.32)$$

Hence, to  $\mathcal{O}(\alpha^3)$ , the Stokes drift velocity associated with the GLM is given by

$$\mathbf{v}^{(s)} = \frac{\mathbf{U}(\mathbf{U} \cdot \mathbf{k})}{2\omega} + \mathcal{O}(\alpha^3). \quad (17.33)$$

The Stokes drift vanishes at this order of accuracy for transverse waves since  $\mathbf{U} \cdot \mathbf{k} = 0$ .

As a check on the formalism, consider a one-dimensional longitudinal wave, in which the Stokes drift is given by

$$\bar{v}^{(s)} = \frac{U^2}{2c} + \mathcal{O}(\alpha^3), \quad (17.34)$$

where  $c = \omega/k$  is the wave speed. This result agrees with that derived using Lagrangian trajectories in VOLUME 1. It is notable that the GLM displacement field offers a somewhat more streamlined method for computing Stokes drift.

### Stokes mean for an arbitrary field

The Stokes mean for an arbitrary field is given by

$$\bar{\chi}^{(s)}(\mathbf{x}, t) = -\omega^{-1} \mathbf{U} \cdot \overline{\nabla \chi' \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)} + \mathcal{O}(\alpha^3) \quad (17.35a)$$

$$= -\omega^{-1} \overline{\nabla \cdot (\mathbf{U} \chi') \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)} + \mathcal{O}(\alpha^3), \quad (17.35b)$$

where the second equality made use of the non-divergent nature of the wave field (17.31). We also assumed the Eulerian mean for this field vanishes,  $\bar{\chi} = 0$ . To third order in wave amplitude, the Stokes mean is determined by the projection of the gradient of the Eulerian fluctuation,  $\nabla \chi'$ ,

onto the wave amplitude,  $\mathbf{U}$ . For example, consider a transverse wave such as that shown in Figure 17.1. Even though the Stokes drift vanishes to order  $\mathcal{O}(\alpha^3)$ , the Stokes mean,  $\bar{\chi}^{(s)}(\mathbf{x}, t)$ , can be nonzero.

### 17.3.7 GLM with a materially constant scalar

Consider a materially constant scalar field, such as a tracer concentration in the absence of mixing and sources

$$\frac{DC}{Dt} = 0. \quad (17.36)$$

This equation means that fluid particles retain their value of  $C$ . We here derive some expressions that are fundamental to the kinematics studied in the remaining sections of this chapter, with Figure 17.3 providing a key schematic.

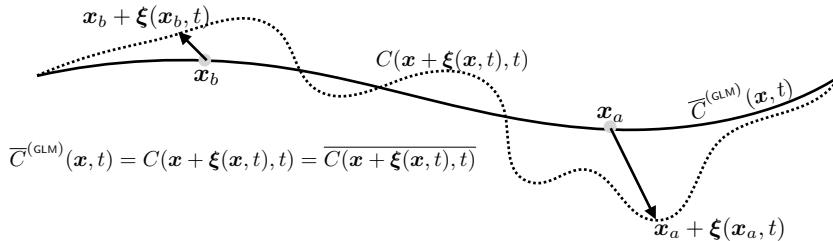


FIGURE 17.3: This figure depicts the GLM tracer concentration,  $\bar{C}^{(\text{GLM})}(\mathbf{x}, t)$ , along with one particular realization of the unaveraged tracer concentration,  $C(\mathbf{x} + \xi, t)$ , that is part of building the GLM. That is,  $\bar{C}^{(\text{GLM})}(\mathbf{x}, t) = C(\mathbf{x} + \xi, t)$ . For example,  $C$  could be the Conservative Temperature,  $\Theta$ , with interest in the 12°C isotherm, in which case  $\bar{\Theta}^{(\text{GLM})}(\mathbf{x}, t) = \Theta(\mathbf{x} + \xi, t) = 12^\circ\text{C}$ . We highlight two fluid particles labelled by  $a$  and  $b$ . The mean positions for these particles are  $\mathbf{x}_a$  and  $\mathbf{x}_b$ , and the positions for the single ensemble member are  $\mathbf{x}_a + \xi(\mathbf{x}_a, t)$  and  $\mathbf{x}_b + \xi(\mathbf{x}_b, t)$ .

### The GLM and its individual realizations

Since  $C$  is a materially invariant tracer concentration, the generalized Lagrangian mean of  $C$  equals to the value of  $C$  when sampled on fluid particles

$$\overline{C(\mathbf{x} + \xi, t)} = \bar{C}^{(\text{GLM})}(\mathbf{x}, t) = C(\mathbf{x} + \xi, t). \quad (17.37)$$

This identity contains in a lot of kinematic information that is basic to the GLM as well as the isopycnal averaging considered in Sections 17.5 and 17.6. Equation (17.37) says that when evaluated at the mean fluid particle position,  $\mathbf{x}$ , the GLM tracer concentration,  $\bar{C}^{(\text{GLM})}(\mathbf{x}, t)$ , equals to the concentration evaluated on a fluid particle,  $C(\mathbf{x} + \xi, t)$ . Hence, the GLM averaging operator has no effect on  $C(\mathbf{x} + \xi, t)$ , simply because the averaging operator follows fluid particles with constant  $C$ . Hence, the GLM for  $C$  equals the value of  $C$  on each member of the ensemble.

### Relating the particle disturbance field to Eulerian properties of $C$

There is a frequently used small amplitude consequence of the identity (17.37) involving the Eulerian fluctuation

$$C'(\mathbf{x}, t) = C(\mathbf{x}, t) - \bar{C}(\mathbf{x}, t) \quad (17.38)$$

and the Eulerian mean

$$C^{(\epsilon)}(\mathbf{x}, t) = \overline{C}(\mathbf{x}, t). \quad (17.39)$$

To derive the identity, recall the Taylor series expansion (17.24), truncated here to first order accuracy

$$C(\mathbf{x} + \boldsymbol{\xi}, t) = C(\mathbf{x}, t) + \boldsymbol{\xi} \cdot \nabla \overline{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (17.40)$$

An Eulerian mean of both sides to this equation renders

$$\overline{C(\mathbf{x} + \boldsymbol{\xi}, t)} = \overline{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2), \quad (17.41)$$

which follows since  $\overline{\boldsymbol{\xi}^p} = 0$  for each component of the displacement field. The identity (17.41) says that the GLM tracer concentration equals to the Eulerian mean tracer concentration to order  $\mathcal{O}(\alpha^2)$ . This result is consistent with the Stokes mean appearing only at order  $\mathcal{O}(\alpha^2)$ , as revealed by equation (17.25c). From equation (17.37) we know that  $C(\mathbf{x} + \boldsymbol{\xi}, t) = \overline{C(\mathbf{x} + \boldsymbol{\xi}, t)}$ . As a result, subtracting equations (17.40) and (17.41) yields<sup>8</sup>

$$C'(\mathbf{x}, t) = -\boldsymbol{\xi} \cdot \nabla \overline{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (17.42)$$

Evidently, the Eulerian eddy tracer concentration is first order in the particle disturbance. More precisely, the eddy tracer concentration equals to minus the disturbance field as projected onto the gradient of the Eulerian mean tracer concentration. Hence, the magnitude of the eddy tracer concentration is maximized for fluid particle disturbances aligned with the gradient of the Eulerian mean tracer concentration, and the eddy concentration vanishes when particle disturbances align with surfaces of constant Eulerian mean tracer. Equation (17.42) forms the starting point for our study in Section 17.4 of the kinematics of eddy tracer fluxes.

### 17.3.8 Further study

GLM was introduced in the seminal papers by [Andrews and McIntyre \(1978a,b\)](#). These papers offer a wealth of intellectual rewards after much study. GLM is also detailed in the monograph on waves and mean flows by [Böhler \(2014a\)](#). [Gilbert and Vanneste \(2025\)](#) provide an elegant and powerful mathematical framework for GLM that allows for its use on arbitrary manifolds.

## 17.4 Eddy tracer fluxes and particle displacements

As introduced in Section 17.3.1, consider the Eulerian eddy-mean decomposition for a materially constant tracer in a non-divergent flow ( $\nabla \cdot \mathbf{v} = 0$ ). The advection equation for this tracer is given by

$$\partial_t C + \nabla \cdot (\mathbf{v} C) = 0, \quad (17.43)$$

and its Eulerian mean is

$$\partial_t \overline{C} + \nabla \cdot (\overline{\mathbf{v}} \overline{C}) = -\nabla \cdot (\overline{\mathbf{v}' C'}). \quad (17.44)$$

The eddy advective flux,  $\mathbf{v}' C'$ , is the product of the eddy velocity and eddy tracer concentration. Its Eulerian mean provides a measure of the statistical correlation between  $\mathbf{v}'$  and  $C'$ , which is commonly referred to as the Eulerian mean eddy tracer flux,  $\overline{\mathbf{v}' C'}$ . Equation (17.44) says that the convergence of the Eulerian mean eddy flux provides a source to the advection equation

---

<sup>8</sup>The minus sign in equation (17.42) is crucial and yet can be easily missed in a cursory analysis.

satisfied by the Eulerian mean tracer concentration, with advection by the Eulerian mean velocity,  $\bar{\mathbf{v}}$ .

In this section we make use of the particle disturbance field of Section 17.3 to help interpret the kinematics of eddy tracer fluxes induced by small amplitude waves. As we show, the particle disturbance field affords a useful conceptual tool to frame the kinematics of eddy tracer fluxes. The kinematic expressions are accurate only for small amplitude disturbances. Even so, they provide a useful conceptual framework for interpreting eddy fluxes from large amplitude turbulent stirring.

### 17.4.1 Particle displacements and eddy tracer fluxes

Following Section 17.3, we introduce a particle disturbance vector corresponding to small amplitude eddy fluctuations

$$\partial_t \boldsymbol{\xi}(\mathbf{x}, t) = \mathbf{v}'(\mathbf{x}, t) + \mathcal{O}(\alpha^2) \quad (17.45a)$$

$$\overline{\boldsymbol{\xi}(\mathbf{x}, t)} = 0. \quad (17.45b)$$

The disturbance vector,  $\boldsymbol{\xi}(\mathbf{x}, t)$ , is an Eulerian space-time field that is defined at each spatial point. It allows us to consider each spatial point,  $\mathbf{x}$ , as the Eulerian mean position of fluid particles whose position is  $\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t)$ ; that is,

$$\overline{\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t)} = \bar{\mathbf{x}} = \mathbf{x}. \quad (17.46)$$

Following the results from Section 17.3.7, to leading order we can write the Eulerian fluctuation in terms of the particle displacement (equation (17.42))

$$C'(\mathbf{x}, t) = -\boldsymbol{\xi} \cdot \nabla \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (17.47)$$

Notice that if the particle displacement is oriented along a mean tracer iso-surface, then  $\boldsymbol{\xi} \cdot \nabla \bar{C}(\mathbf{x}, t) = 0$  and there is no tracer fluctuation,  $C' = 0$ , to order  $\mathcal{O}(\alpha^2)$ . More general eddying motions lead to a nonzero tracer fluctuation with the eddy tracer flux taking on the form

$$\mathbf{v}' C' = -\partial_t \boldsymbol{\xi} (\boldsymbol{\xi} \cdot \nabla) \bar{C} + \mathcal{O}(\alpha^3). \quad (17.48)$$

### 17.4.2 Decomposing into symmetric and skew symmetric fluxes

From equation (17.48), the m'th component of the eddy tracer flux is given by

$$v'^m C' = -[(\partial_t \xi^m) \xi^n] \partial_n \bar{C}. \quad (17.49)$$

To explore the kinematic properties of this tracer flux, decompose the second order tensor,  $(\partial_t \xi^m) \xi^n$ , into its symmetric and anti-symmetric components<sup>9</sup>

$$2(\partial_t \xi^m) \xi^n = [(\partial_t \xi^m) \xi^n + (\partial_t \xi^n) \xi^m] + [(\partial_t \xi^m) \xi^n - (\partial_t \xi^n) \xi^m] \quad (17.50a)$$

$$= \partial_t(\xi^m \xi^n) + [(\partial_t \xi^m) \xi^n - (\partial_t \xi^n) \xi^m]. \quad (17.50b)$$

<sup>9</sup>This decomposition is ubiquitous in matrix and tensor algebra, given that the symmetric and anti-symmetric components have distinct kinematic properties.

Introducing the symmetric and anti-symmetric correlation tensors

$$2K^{mn} \equiv \overline{\partial_t(\xi^m \xi^n)} \quad (17.51a)$$

$$2A^{mn} \equiv \overline{(\partial_t \xi^m) \xi^n} - \overline{(\partial_t \xi^n) \xi^m} \quad (17.51b)$$

allows us to write the mean eddy tracer flux

$$\overline{v'^m C'} = -(K^{mn} + A^{mn}) \partial_n \overline{C} \quad (17.52)$$

and the mean field tracer equation (17.44)

$$\partial_t \overline{C} + \nabla \cdot (\overline{\mathbf{v}} \overline{C}) = \partial_m [(K^{mn} + A^{mn}) \partial_n \overline{C}] = \nabla \cdot [(\mathbf{K} + \mathbf{A}) \cdot \nabla \overline{C}]. \quad (17.53)$$

The right hand side of this equation equals to the convergence of the symmetric and skew-symmetric tracer fluxes

$$\nabla \cdot [(\mathbf{K} + \mathbf{A}) \cdot \nabla \overline{C}] = -\nabla \cdot (\mathbf{F}^{\text{sym}} + \mathbf{F}^{\text{skew}}), \quad (17.54)$$

where

$$\mathbf{F}^{\text{sym}} = -\mathbf{K} \cdot \nabla \overline{C} \quad (17.55a)$$

$$\mathbf{F}^{\text{skew}} = -\mathbf{A} \cdot \nabla \overline{C} \quad (17.55b)$$

$$\overline{\mathbf{v}' C'} = \mathbf{F}^{\text{sym}} + \mathbf{F}^{\text{skew}} = -(\mathbf{K} + \mathbf{A}) \cdot \nabla \overline{C}. \quad (17.55c)$$

### 17.4.3 The symmetric tracer flux

In terms of particle displacements, the symmetric flux (17.55a) is given by

$$(F^{\text{sym}})^m = -K^{mn} \partial_n \overline{C} = -\frac{1}{2} \overline{\partial_t(\xi^m \xi^n)} \partial_n \overline{C}. \quad (17.56)$$

The symmetric tensor,  $\mathbf{K}$ , vanishes when the average is over the period of a periodic wave, in which the particle displacements undergo reversible periodic excursions (see Exercise 17.5). For waves that decay in amplitude over the averaging period, particle displacements decrease in magnitude so that  $K^{mn} < 0$ . In this case, the symmetric flux is oriented up the gradient of the Eulerian mean tracer field. In contrast, particle displacements increase in magnitude for waves that grow over the averaging period, in which case the flux is downgradient, just as for diffusion. Furthermore, growing nonlinear waves generally break and then develop into turbulence, with turbulence leading to further particle separation and tracer mixing. We thus find it sensible to parameterize turbulent motions via downgradient diffusion, with much of Chapter 18 focused on diffusive parameterizations of lateral mixing.

### 17.4.4 Skew, advective, and rotational tracer fluxes

Following our discussion in Section 16.5, we write the  $m$ 'th component to the skew flux

$$(F^{\text{skew}})^m = -A^{mn} \partial_n \overline{C} = -\epsilon^{mnp} \Psi_p \partial_n \overline{C} = -(\nabla \overline{C} \times \Psi)^m, \quad (17.57)$$

where we introduced the vector streamfunction (dimensions squared length per time)<sup>10</sup>

$$\Psi = \frac{1}{2} \partial_t \xi \times \xi = \frac{1}{2} \mathbf{v}' \times \xi. \quad (17.58)$$

As defined, the vector streamfunction is minus one half the angular momentum per mass of a fluid particle undergoing eddying motion, with the angular momentum computed relative to the mean particle position. The vector streamfunction is nonzero only if the eddy has a preferred sense of rotation (i.e., a nonzero mean angular momentum), in which case the wave field is said to be **polarized**. That is, polarization results if the eddy velocity,  $\mathbf{v}'$ , is correlated to a fluid particle displacement,  $\xi$ , in an orthogonal direction, thus giving rise to a nonzero angular momentum.

The skew flux can be written

$$\mathbf{F}^{\text{skew}} = -\nabla \bar{C} \times \Psi \quad (17.59a)$$

$$= (\nabla \times \Psi) \bar{C} - \nabla \times (\bar{C} \Psi) \quad (17.59b)$$

$$= \mathbf{U}^A \bar{C} - \nabla \times (\bar{C} \Psi) \quad (17.59c)$$

$$= \mathbf{F}^{\text{adv}} - \mathbf{F}^{\text{rot}}, \quad (17.59d)$$

so that the skew flux equals to an advective flux minus a rotational flux. We here introduced the non-divergent velocity,

$$\mathbf{U}^A = \nabla \times \Psi, \quad (17.60)$$

and the non-divergent rotational flux,

$$\mathbf{F}^{\text{rot}} = \nabla \times (\bar{C} \Psi). \quad (17.61)$$

Since  $\nabla \cdot \mathbf{F}^{\text{rot}} = 0$ , we see that the divergence of the skew flux equals to the divergence of the advective flux

$$\nabla \cdot \mathbf{F}^{\text{skew}} = \nabla \cdot (\mathbf{F}^{\text{adv}} - \mathbf{F}^{\text{rot}}) = \nabla \cdot \mathbf{F}^{\text{adv}}. \quad (17.62)$$

Consequently, the rotational flux,  $\mathbf{F}^{\text{rot}}$ , has no impact on evolution of the mean tracer concentration.

#### 17.4.5 What does a point measurement yield?

From equation (17.55c), we see that a point measurement of the correlation,  $\overline{\mathbf{v}' C'}$ , provides an estimate of the symmetric tracer flux plus the skew tracer flux

$$\overline{\mathbf{v}' C'} = \mathbf{F}^{\text{sym}} + \mathbf{F}^{\text{skew}} = -(\mathbf{K} + \mathbf{A}) \cdot \nabla \bar{C}. \quad (17.63)$$

Furthermore, for a periodic wave field, where the symmetric tensor vanishes, the correlation,  $\overline{\mathbf{v}' C'}$ , provides a direct estimate of the skew flux,  $-\nabla \bar{C} \times \Psi$ . This particular result might seem puzzling on first encounter, since one could imagine  $\overline{\mathbf{v}' C'}$  instead provides an estimate for the advective flux,  $\mathbf{U}^A \bar{C}$ . But that presumption is wrong, as indicated by the decomposition (17.63). We emphasize this point by summarizing the various relations

$$\overline{\mathbf{v}' C'} = \mathbf{F}^{\text{sym}} + \mathbf{F}^{\text{skew}} \quad (17.64a)$$

---

<sup>10</sup> Middleton and Loder (1989) and Garrett (2006) introduce a skew-diffusivity,  $\mathbf{D}$ , which is opposite in sign to the vector streamfunction:  $\Psi = -\mathbf{D}$ .

$$= -\mathbf{K} \cdot \nabla \bar{C} - \nabla \bar{C} \times \Psi \quad (17.64b)$$

$$= -\mathbf{K} \cdot \nabla \bar{C} - \nabla \times (\bar{C} \Psi) + \bar{C} \nabla \times \Psi \quad (17.64c)$$

$$= -\mathbf{K} \cdot \nabla \bar{C} - \nabla \times (\bar{C} \Psi) + \bar{C} \mathbf{U}^A \quad (17.64d)$$

$$= \mathbf{F}^{\text{sym}} - \mathbf{F}^{\text{rot}} + \mathbf{F}^{\text{adv}}. \quad (17.64e)$$

The rotational flux is generally nontrivial for polarized waves or turbulent eddies. As a result, the rotational flux provides a sizable contribution to any measurement of  $\overline{\mathbf{v}' C'}$  either from a field measurement or numerical simulation. For some purposes it can be more convenient to work directly with the skew flux rather than the advective flux.

#### 17.4.6 Further study

Much of this section follows [Plumb \(1979\)](#), [Middleton and Loder \(1989\)](#), and [Garrett \(2006\)](#), each of whom considered elements of tracer dispersion by waves and nonlinear eddies. [Middleton and Loder \(1989\)](#) work through a few oceanographically motivated examples that offer further understanding of skew fluxes. Additional treatments can be found in the review article of [Moffatt \(1983\)](#), who considers flow in a rotating reference frame as well as magneto-hydrodynamic flows.

Exercises 17.2 through 17.6 help build further insights into the kinematics introduced in this section.

## 17.5 Volume transport in an isopycnal ensemble

In this section we study the kinematics of stirring by a turbulent flow in a perfect stratified Boussinesq fluid. We move beyond the finite amplitude motions considered in the previous sections, and yet make use of the generalized Lagrangian mean kinematics to conceptually frame the formulation. As [fluid parcels](#) are stirred by turbulent and non-divergent flows, they preserve their volume while changing their shape and stretching into finer scale features. This process accords with the conceptual picture of [Eckart \(1948\)](#) discussed in Section 16.1. The turbulent motion in ocean mesoscale/baroclinic eddies, as well as nonlinear baroclinic waves in the atmosphere, offer geophysical examples of flows that lead to such stirring. We are not directly concerned with mixing in this section, instead focusing on the kinematics of eddy stirring in a stratified Boussinesq ocean.<sup>11</sup>

With a focus on stirring, we assume that each fluid particle preserves its potential density,  $\varrho$  (equivalently its specific entropy). This property motivates us to take an isopycnal coordinate approach in formulating the kinematics. Here, we track the vertical motion of potential density layer interfaces during the motion, whereas we are not concerned with following the lateral position of a fluid particle within a layer. We examine the kinematics of a  $\varrho$ -ensemble of eddying flows, whereby we focus on an ensemble of flows along an arbitrary isopycnal surface labelled by  $\varrho$ . We interpret this isopycnal approach as a vertical GLM, in which we focus on the vertical particle displacement rather than the three dimensional displacement of Section 17.3.<sup>12</sup>

We are concerned with the kinematics of parcel rearrangement in this section, with eddy correlations appearing between the specific thickness of an isopycnal fluid layer and the flow

<sup>11</sup>Although focusing on the perfect Boussinesq fluid, keep in mind that the fluid is forced to support an active and stationary turbulent flow. For example, the flow in a mechanically (wind driven) isopycnal model offers a numerical realization of this flow.

<sup>12</sup>We emphasize that the fluid particles move in three directions within the  $\varrho$ -ensemble. However, the isopycnal kinematic approach focuses just on the vertical particle displacement.

velocity within that layer. We extend the analysis to tracers in Section 17.6. We offer many details since the kinematics of isopycnal ensembles appears throughout the study of wave-mean flow interactions in geophysical fluid mechanics, with particular importance to the study of ocean mesoscale eddy transport and its parameterization. Sections 17.5.1 through 17.5.4 introduce the main ingredients of the method, whereas the following sections develop expressions for the transport of volume. Mastery of this material is not simple.

### 17.5.1 Specifying the vertical position

When studying the generalized Lagrangian mean in Section 17.3, we interpret each point in space,  $\mathbf{x}$ , as the mean position of a unique fluid particle. This interpretation acts to couple the Eulerian and Lagrangian kinematic descriptions. For purposes in this section we only track the vertical position as defined by the position of an isopycnal surface. As such, we identify each vertical position in space,  $z$ , with the mean vertical position of an isopycnal whose particular value is  $\varrho$ :

$$z = \bar{\eta}^{(\varrho)}(x, y, t) = \text{mean vertical position of isopycnal } \varrho \text{ at } (x, y, t). \quad (17.65)$$

The overline in this equation signals an ensemble mean operation as computed over the  $\varrho$ -ensemble of isopycnal surfaces with potential density,  $\varrho$ . Notably, isopycnals are generally not horizontal nor are they static (remember, the flows are forced). Hence, the mean vertical position of the  $\varrho$ -ensemble is a function of horizontal position and time.<sup>13</sup>

We write

$$z = \eta(x, y, \varrho, t) \quad (17.66)$$

for the vertical position of an isopycnal labeled by  $\varrho$  as located at a horizontal point,  $(x, y)$ , and time,  $t$ . We decompose this vertical position into its ensemble mean position,  $\bar{\eta}^{(\varrho)}(x, y, t)$ , and a vertical displacement,  $\xi(x, y, \varrho, t)$ , relative to the mean we introduce the vertical displacement field,  $\xi(x, y, \varrho, t)$ , thus writing

$$\eta(x, y, \varrho, t) = \bar{\eta}^{(\varrho)}(x, y, t) + \xi(x, y, \varrho, t) = \text{vertical position of isopycnal } \varrho. \quad (17.67)$$

Just like the vertical position,  $\eta(x, y, \varrho, t)$ , we see that the vertical displacement is a function of the chosen isopycnal,  $\varrho$ , as well as the horizontal position and time. As for the GLM displacement vector,  $\xi(\mathbf{x}, t)$ , the isopycnal vertical displacement has a zero  $\varrho$ -ensemble mean

$$\overline{\xi(x, y, \varrho, t)}^{(\varrho)} = 0 \implies \overline{\eta(x, y, \varrho, t)}^{(\varrho)} = \bar{\eta}^{(\varrho)}(x, y, t). \quad (17.68)$$

Hence, the vertical displacement field,  $\xi(x, y, \varrho, t)$ , serves to couple Eulerian kinematics to the quasi-Lagrangian kinematics of isopycnals.

### 17.5.2 Modified mean density

Figure 17.4 provides a schematic to illustrate the mean vertical position of the  $\varrho$ -ensemble of isopycnals, and the corresponding displacement field that locates the vertical position of a particular member of the  $\varrho$ -ensemble. This figure also motivates us to define a new density

<sup>13</sup>We slightly abuse notation by writing  $\varrho$  for a specific value of the density held by each  $\varrho$ -ensemble member (e.g.,  $\varrho = 1030 \text{ kg m}^{-3}$ ). We also write  $\varrho(x, y, z, t)$  for a particular ensemble member whose value at each point in space and time is  $\varrho(x, y, z, t) = \varrho$ . A more careful notation involves extra adornments to each symbol and can become rather tedious and difficult to parse. We thus rely on understanding the meaning based on the context.

field,  $\tilde{\varrho}$ , which equals to  $\varrho$  when evaluated at the  $\varrho$ -ensemble mean vertical position. Following [McDougall and McIntosh \(2001\)](#), we refer to  $\tilde{\varrho}$  as the modified mean potential density, and it is written mathematically as<sup>14</sup>

$$\tilde{\varrho}(x, y, \bar{\eta}^{(\varrho)}(x, y, t), t) = \overline{\varrho(x, y, \eta(x, y, \varrho, t), t)}^{(\varrho)} = \varrho(x, y, \eta(x, y, \varrho, t), t) = \varrho. \quad (17.69)$$

The final equality follows since each space-time functions in this equation takes on the same value of potential density (e.g.,  $\varrho = 1030 \text{ kg m}^{-3}$ ) when evaluated at their respective space-time points. The second equality follows since each member of the  $\varrho$ -ensemble has the same potential density, so that  $\varrho(x, y, \eta(x, y, \varrho, t), t)$  is invisible to the  $\varrho$ -ensemble mean operator.

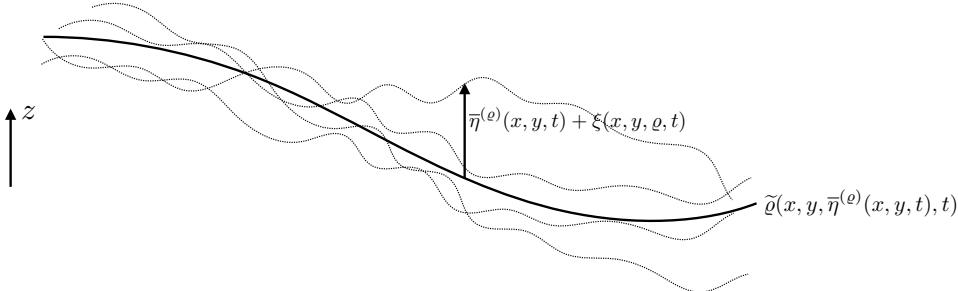


FIGURE 17.4: We here depict a few elements of the  $\varrho$ -ensemble (dotted lines; in principle there are an infinity of elements in the  $\varrho$ -ensemble) as well as the ensemble mean position of these isopycnals (dark line). The ensemble mean has a vertical position  $z = \bar{\eta}^{(\varrho)}(x, y, t)$ , whereas the vertical position of each ensemble member is  $z = \eta(x, y, \varrho, t) = \bar{\eta}^{(\varrho)}(x, y, t) + \xi(x, y, \varrho, t)$ , with a distinct displacement field,  $\xi$ , for each ensemble member. Each potential density surface within the  $\varrho$ -ensemble has the same numerical value,  $\varrho(x, y, z, t) = \varrho$  (e.g.,  $\varrho = 1030 \text{ kg m}^{-3}$ ). We define the modified mean density,  $\tilde{\varrho}$ , as the space-time field that equals to  $\varrho$  when evaluated at the mean vertical position:  $\tilde{\varrho}(x, y, \bar{\eta}^{(\varrho)}(x, y, t), t) = \varrho$ .

### 17.5.3 Isopycnal mean and modified mean

Each member of the  $\varrho$ -ensemble has the same potential density,  $\varrho$ . However, each member generally has distinct values for other fluid properties. Consider a physical scalar property,  $\chi$ , such as the temperature. It has a value at an Eulerian space point,  $\chi(x, y, z, t)$ , and a generally distinct value on a  $\varrho$  isopycnal,  $\chi(x, y, z = \eta(x, y, \varrho, t), t)$ . We may also represent this property using isopycnal coordinates, in which case we write the  $\varrho$ -ensemble mean for this property as<sup>15</sup>

$$\bar{\chi}^{(\varrho)}(x, y, \varrho, t) \equiv \varrho\text{-ensemble mean}. \quad (17.70)$$

This isopycnal mean is operationally computed just like an Eulerian mean, only now the field is represented using isopycnal coordinates, and it is sampled at the position of each member of the  $\varrho$ -ensemble that has fixed  $(x, y, \varrho, t)$ .

Following from our definition of the modified mean density,  $\tilde{\varrho}$ , we define the modified mean

<sup>14</sup>Following our connection to the GLM, we consider  $\tilde{\varrho}$  as the vertical GLM density.

<sup>15</sup>The physical scalar property,  $\chi$ , can be represented as a function of  $z$  or as a function of  $\varrho$ . These two coordinate representations require distinct mathematical functions. We could introduce distinct function names to distinguish these functions. But we choose not to in order to reduce notation. Rather, we let the distinct dependent variables signal the use of a distinct mathematical function.

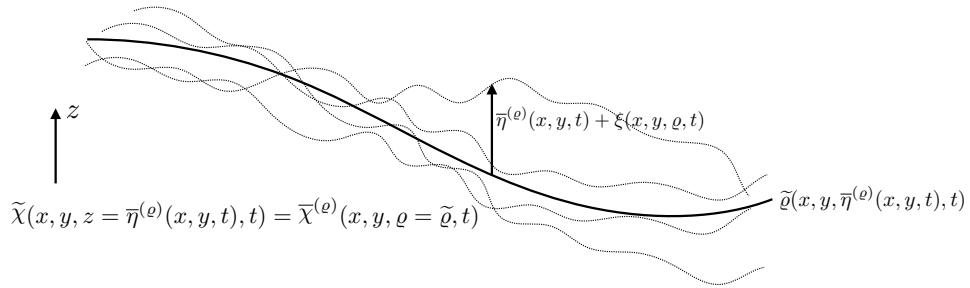


FIGURE 17.5: This figure depicts the duality relation  $\tilde{\chi}(x, y, z = \bar{\eta}^{(\varrho)}(x, y, t), t) = \bar{\chi}^{(\varrho)}(x, y, \varrho = \tilde{\varrho}, t)$  between the modified mean scalar property,  $\tilde{\chi}$ , and the  $\varrho$ -ensemble mean of that property. It expresses the objective nature of the  $\varrho$ -ensemble mean of the property, with distinct mathematical functional forms depending on whether it is represented using either  $(x, y, z, t)$  Eulerian coordinates or isopycnal coordinates,  $(x, y, \varrho, t)$ . This duality is basic to the  $\varrho$ -ensemble kinematics.

scalar field<sup>16</sup>

$$\tilde{\chi}(x, y, \bar{\eta}^{(\varrho)}(x, y, t), t) \equiv \overline{\chi(x, y, \eta(x, y, \varrho, t), t)}^{(\varrho)} = \overline{\chi(x, y, \bar{\eta}^{(\varrho)}(x, y, t) + \xi(x, y, \varrho, t), t)}^{(\varrho)}. \quad (17.71)$$

The modified mean field is a function of space and time, with the vertical position given by the  $\varrho$ -ensemble mean vertical position,  $z = \bar{\eta}^{(\varrho)}(x, y, t)$ . By construction, the modified mean, when evaluated at  $\bar{\eta}^{(\varrho)}(x, y, t)$ , equals to the  $\varrho$ -ensemble mean evaluated using the modified mean density

$$\tilde{\chi}(x, y, z = \bar{\eta}^{(\varrho)}(x, y, t), t) = \bar{\chi}^{(\varrho)}(x, y, \varrho = \tilde{\varrho}, t). \quad (17.72)$$

Evidently, as a space-time function,  $\tilde{\chi}(x, y, \bar{\eta}^{(\varrho)}(x, y, t), t)$  is numerically equal to the  $\varrho$ -ensemble mean of  $\chi$ . This identity results from the objective nature of the isopycnal averaged property,  $\chi$ , with that average represented either via  $(x, y, z)$  Eulerian coordinates or  $(x, y, \varrho)$  isopycnal coordinates. We illustrate this identity in Figure 17.5, and make use of it in the following.

#### 17.5.4 Transformed residual mean (TRM)

When working with isopycnal layers, it is very useful to employ specific thickness weighting to account for the net amount of material within a layer, or to measure the net transport in the layer.<sup>17</sup> For this purpose we introduce the **specific thickness** from Section 13.3 as given by<sup>18</sup>

$$h = \partial z / \partial \varrho = 1 / (\partial \varrho / \partial z), \quad (17.73)$$

and thus make use of thickness weighted fields,  $h \chi$ , and the corresponding thickness weighted isopycnal ensemble mean

$$\hat{\chi} = \frac{h \chi^{(\varrho)}}{h^{(\varrho)}}. \quad (17.74)$$

<sup>16</sup>We again interpret the modified mean scalar as a vertical GLM field.

<sup>17</sup>We pursue a vertical discrete version of this **thickness weighted averaging (TWA)** approach for the stacked shallow water model in Chapter 14.

<sup>18</sup>As discussed in Section 12.10.2, specific thickness is the Jacobian of transformation between geopotential coordinates,  $(x, y, z, t)$ , and isopycnal coordinates,  $(x, y, \varrho, t)$ . For stably stratified ideal fluids,  $h$  is one-signed, hence making the coordinate transformation well defined. It is also related to the buoyancy frequency through  $N^2 = -(g/\rho_0) (\partial \varrho / \partial z) = -g/(\rho_0 h)$ .

Following from the duality relation (17.72) we introduce

$$\bar{\chi}^\#(x, y, \bar{\eta}^{(\varrho)}, t) \equiv \hat{\chi}(x, y, \tilde{\varrho}, t), \quad (17.75)$$

where  $\bar{\chi}^\#$  is the transformed residual mean (TRM) evaluated at the isopycnal ensemble mean vertical position.<sup>19</sup> This is yet another important identity that is used in the following.

### 17.5.5 Depth integrated TRM transport

The horizontal TRM velocity is a particularly key field

$$\hat{\mathbf{u}}(x, y, \tilde{\varrho}, t) = \bar{\mathbf{u}}^\#(x, y, \bar{\eta}^{(\varrho)}, t), \quad (17.76)$$

where we make use of a somewhat more abbreviated notation for the functional dependencies. Following the discussion of the vertical gauge in Section 16.5.1 (see in particular equation (16.59)), we are led to define the depth integrated TRM transport as in Figure 17.6

$$\bar{\mathbf{U}}^\#(x, y, \bar{\eta}^{(\varrho)}, t) = \int_{\eta_b}^{\bar{\eta}^{(\varrho)}} \bar{\mathbf{u}}^\#(x, y, z, t) dz = \int_{\varrho(x, y, \eta_b, t)}^{\tilde{\varrho}(x, y, \bar{\eta}^{(\varrho)}, t)} \hat{\mathbf{u}}(x, y, \gamma, t) \bar{h}^{(\gamma)}(x, y, t) d\gamma, \quad (17.77)$$

with the second equality following from a change of coordinates from geopotential to isopycnal, and with  $z = \eta_b(x, y)$  the vertical position of the ocean bottom topography.

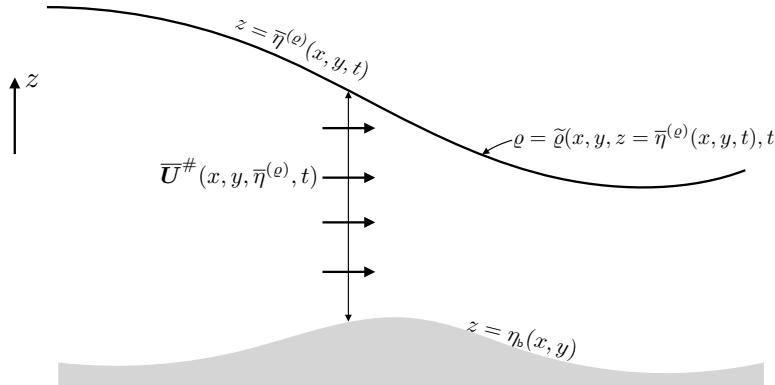


FIGURE 17.6: Depicting the horizontal TRM transport as computed from the ocean bottom at  $z = \eta_b(x, y)$ , and the vertical position  $z = \bar{\eta}^{(\varrho)}(x, y, t)$ .

We can go further with equation (17.77) by writing (and suppressing some functional dependencies for brevity)

$$\bar{\mathbf{U}}^\#(x, y, \bar{\eta}^{(\varrho)}, t) = \int_{\varrho(\eta_b)}^{\tilde{\varrho}(\bar{\eta}^{(\varrho)})} \hat{\mathbf{u}}(\gamma) \bar{h}^{(\gamma)} d\gamma \quad \text{from equation (17.77)} \quad (17.78a)$$

$$= \int_{\varrho(\eta_b)}^{\tilde{\varrho}(\bar{\eta}^{(\varrho)})} \bar{\mathbf{u}} \bar{h}^{(\gamma)} d\gamma \quad \text{from equation (17.74)} \quad (17.78b)$$

$$= \int_{\varrho(\bar{\eta}^{(\varrho)} + \xi)}^{\tilde{\varrho}(\bar{\eta}^{(\varrho)})} \bar{\mathbf{u}} \bar{h}^{(\gamma)} d\gamma \quad \text{from equation (17.69).} \quad (17.78c)$$

<sup>19</sup> McDougall and McIntosh (2001) pioneered the use of TRM. Since their averaging operator was based on a time mean, they referred to it as the temporal residual mean.

The final equality makes it clear that the TRM transport,  $\bar{\mathbf{U}}^\#$ , is the ensemble mean volume transport for fluid denser than  $\varrho(\bar{\eta}^{(\varrho)} + \xi) = \tilde{\varrho}(\bar{\eta}^{(\varrho)})$ . This transport can also be written using geopotential coordinates

$$\bar{\mathbf{U}}^\#(x, y, \bar{\eta}^{(\varrho)}, t) = \overline{\int_{\eta_b}^{\bar{\eta}^{(\varrho)} + \xi} \mathbf{u} dz}. \quad (17.79)$$

The transport from each ensemble member is determined by integrating from the ocean bottom to the vertical position,  $\bar{\eta}^{(\varrho)} + \xi$ , and then the TRM transport is determined by computing the ensemble mean for this transport.

### 17.5.6 Quasi-Stokes transport

The TRM transport (17.79) can be decomposed into an Eulerian mean plus an eddy correlation

$$\bar{\mathbf{U}}^\#(\bar{\eta}^{(\varrho)}) \equiv \bar{\mathbf{U}}(\bar{\eta}^{(\varrho)}) + \bar{\mathbf{U}}^{qs}(\bar{\eta}^{(\varrho)}), \quad (17.80)$$

where we exposed just the vertical dependency to these functions. The first term,

$$\bar{\mathbf{U}}(\bar{\eta}^{(\varrho)}) = \overline{\int_{\eta_b}^{\bar{\eta}^{(\varrho)}} \mathbf{u} dz}, \quad (17.81)$$

is the ensemble mean transport between the ocean bottom and the ensemble mean vertical position,  $\bar{\eta}^{(\varrho)}$ . We interpret this transport as an Eulerian mean since the depth ranges are fixed. In contrast, the quasi-Stokes transport

$$\bar{\mathbf{U}}^{qs}(\bar{\eta}^{(\varrho)}) \equiv \overline{\int_{\bar{\eta}^{(\varrho)}}^{\bar{\eta}^{(\varrho)} + \xi} \mathbf{u} dz} \quad (17.82)$$

measures the ensemble mean transport between the vertical position of the  $\varrho$ -ensemble mean,  $z = \bar{\eta}^{(\varrho)}$ , and that of each ensemble member,  $z = \bar{\eta}^{(\varrho)} + \xi(\varrho)$ . We refer to transport as “quasi-Stokes” given that it is the difference between an isopycnal (i.e., quasi-Lagrangian) mean and an Eulerian mean (see Section 17.3)

$$\bar{\mathbf{U}}^{qs} = \bar{\mathbf{U}}^\# - \bar{\mathbf{U}}. \quad (17.83)$$

As for the traditional Stokes drift discussed in Section 17.3.6, the quasi-Stokes transport arises from a correlation between the velocity and the undulation of the isopycnal interface.

### 17.5.7 Three-component TRM velocity

Following from the vertical gauge expression (16.58), we introduce the TRM vector streamfunction

$$\bar{\Psi}^\# = \bar{\mathbf{U}}^\# \times \hat{\mathbf{z}}, \quad (17.84)$$

and the corresponding three-dimensional non-divergent TRM velocity

$$\bar{\mathbf{v}}^\# = \nabla \times \bar{\Psi}^\#. \quad (17.85)$$

The vertical component,

$$\bar{w}^\# = \hat{\mathbf{z}} \cdot (\nabla \times \bar{\Psi}^\#), \quad (17.86)$$

has no corresponding component in an isopycnal description, which only requires the horizontal thickness weighted transport,  $\hat{\mathbf{u}}$ . However, the TRM vector streamfunction only requires the horizontal TRM transport,  $\bar{\mathbf{U}}^\#$ , so the two descriptions make use of the same number of degrees of freedom.

### 17.5.8 Volume conservation and the thickness equation

Consider two perspectives on volume conservation: one based on isopycnal coordinates and the other based on geopotential coordinates.

#### Isopycnal coordinates

In isopycnal vertical coordinates, the volume of a fluid element is written

$$\delta V = \delta x \delta y \delta z = \delta x \delta y \delta \varrho h, \quad (17.87)$$

where we introduced the specific thickness,  $h$ , from equation (17.73). Geometrically, the product

$$\delta z = |h \delta \varrho| \quad (17.88)$$

represents the vertical thickness between two infinitesimally close potential density interfaces,  $\varrho$  and  $\varrho + \delta \varrho$  (see Figure 17.7). Material conservation of both volume and potential density implies conservation of the product of specific thickness and horizontal area,  $\delta x \delta y h$ , which leads to the thickness equation (Section 14.2.3)

$$\partial_t h + \nabla_h \cdot (h \mathbf{u}) = 0, \quad (17.89)$$

with  $\mathbf{u}$  the horizontal velocity field, the time derivative is computed with  $\varrho$  held fixed, and

$$\nabla_h = \nabla_h + \mathbf{S} \partial_z \quad (17.90)$$

is the horizontal derivative operator with  $\varrho$  held fixed, and

$$\mathbf{S} = \nabla_h z \quad (17.91)$$

is the horizontal slope of the potential density surface. This equation is the continuous isopycnal version of the thickness equation appearing in vertically discrete shallow water layers.

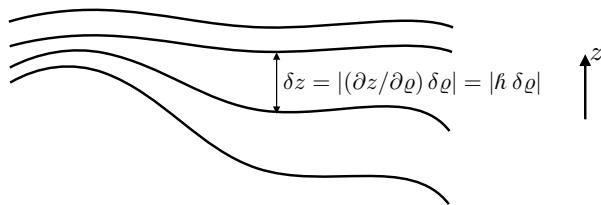


FIGURE 17.7: Illustrating the thickness between surfaces of constant generalized vertical coordinates,  $\delta z = (\partial z / \partial \sigma) \delta \sigma$ . In regions with larger magnitude for the specific thickness,  $\partial z / \partial \sigma$ , or equivalently smaller vertical stratification of the  $\sigma$  surfaces,  $\partial \sigma / \partial z$ , the layers are thicker so that isolines of constant  $\sigma$  are further apart. The converse holds where  $\partial z / \partial \sigma$  is small (equivalently  $\partial \sigma / \partial z$  is large).

### Geopotential coordinates

An Eulerian  $z$ -coordinate description of volume stirring within isopycnal layers is rendered via a combination of volume conservation,  $\nabla \cdot \mathbf{v} = 0$ , and material conservation of potential density,  $D\varrho/Dt = 0$ . When written as skewson rather than advection, the natural gauge is the vertical gauge introduced in Section 16.5.1, since this gauge only requires the same horizontal velocity field,  $\mathbf{u}$ , used with the isopycnal coordinate description. This gauge has an associated potential density skew flux,  $\mathbf{F}^{\text{skew}} = -\nabla\varrho \times \Psi$ , which leads to the evolution

$$\partial_t \varrho = \nabla \cdot (\nabla\varrho \times \Psi), \quad (17.92)$$

where all derivatives are here taken with fixed Eulerian (geopotential) coordinates,  $(x, y, z)$ , and the  $\nabla$  operator is three-dimensional.

#### 17.5.9 Ensemble mean thickness equation

Consider an ensemble of stably stratified (so that the layer specific thickness,  $\mathbf{h}$ , is single-signed and nonvanishing) perfect Boussinesq fluid parcels with the same infinitesimal volume,  $\delta V = \delta x \delta y \delta z = \delta x \delta y \mathbf{h} \delta\varrho$ , and same potential density,  $\varrho$ . Lacking any other marker, such as a tracer concentration, the ensemble members are distinguished from one another by values of their horizontal area,  $\delta A = \delta x \delta y$ , and their specific thickness,  $\mathbf{h}$ , that is, their geometric attributes. The ensemble members are assumed to be stirred by different stochastic realizations of the fluid flow. Since each flow realization alters the geometric properties of the parcels, a mean field description focuses on the mean of these geometric properties.

In isopycnal coordinates,  $(x, y, \varrho, t)$ , the thickness equation (17.89) is satisfied by each ensemble member

$$\partial_t \mathbf{h} + \nabla_{\mathbf{h}} \cdot (\mathbf{h} \mathbf{u}) = 0. \quad (17.93)$$

The ensemble mean computed over these fluid parcels, each with potential density  $\varrho$ , satisfies

$$\partial_t \bar{\mathbf{h}}^{(\varrho)} + \nabla_{\mathbf{h}} \cdot \left( \bar{\mathbf{h}}^{(\varrho)} \bar{\mathbf{u}}^{(\varrho)} + \bar{\mathbf{h}}' \bar{\mathbf{u}}'^{(\varrho)} \right) = 0, \quad (17.94)$$

where primed variables represent deviations from the isopycnal mean. It follows that the mean specific thickness,  $\bar{\mathbf{h}}^{(\varrho)}$ , of parcels with potential density,  $\varrho$ , satisfies the conservation equation

$$\partial_t \bar{\mathbf{h}}^{(\varrho)} + \nabla_{\mathbf{h}} \cdot (\bar{\mathbf{h}}^{(\varrho)} \hat{\mathbf{u}}) = 0. \quad (17.95)$$

In this equation we introduced the thickness weighted isopycnal ensemble mean horizontal velocity

$$\hat{\mathbf{u}} = \frac{\bar{\mathbf{h}} \mathbf{u}^{(\varrho)}}{\bar{\mathbf{h}}^{(\varrho)}} = \bar{\mathbf{u}}^{(\varrho)} + \frac{\bar{\mathbf{h}}' \mathbf{u}'^{(\varrho)}}{\bar{\mathbf{h}}^{(\varrho)}} \equiv \bar{\mathbf{u}}^{(\varrho)} + \mathbf{u}^{\text{bolus}}, \quad (17.96)$$

along with the isopycnal ensemble mean horizontal velocity,  $\bar{\mathbf{u}}^{(\varrho)}$ , and the horizontal **bolus velocity**,  $\mathbf{u}^{\text{bolus}}$ , originally introduced by [Rhines \(1982\)](#). The bolus velocity for an isopycnal layer corresponds to the transport

$$\bar{\mathbf{h}}^{(\varrho)} \mathbf{u}^{\text{bolus}} = \bar{\mathbf{h}}^{(\varrho)} (\hat{\mathbf{u}} - \bar{\mathbf{u}}^{(\varrho)}) = \bar{\mathbf{h}}' \bar{\mathbf{u}}'^{(\varrho)}, \quad (17.97)$$

which arises from the along-isopycnal correlations between specific thickness and horizontal velocity.

Quite conveniently, the mean conservation equation (17.95) takes the *same* mathematical form as the conservation equation (17.93) satisfied by each ensemble member. The key difference is that the isopycnal ensemble mean thickness,  $\bar{h}^{(\varrho)}$ , is stirred by the thickness weighted isopycnal ensemble mean horizontal velocity,  $\hat{\mathbf{u}}$ , whereas the thickness of each ensemble member is stirred by a randomly different realization of the horizontal velocity,  $\mathbf{u}$ . The simplicity of the mean field description (17.95) is afforded by use of the Lagrangian vertical coordinate,  $\varrho$ .

### 17.5.10 Ensemble mean kinematics in geopotential coordinates

Now consider a geopotential coordinate description of the isopycnal ensemble. For this purpose, we interpret a vertical position,  $z$ , as the ensemble mean vertical position,  $\bar{\eta}^{(\varrho)}$ . Consequently, mean fields defined at the fixed vertical position correspond to either modified mean fields when not thickness weighted (equation (17.71)), or TRM fields when thickness weighted (equation (17.75)).

#### Evolution of modified mean density

Following the skewson formulation from Section 16.5, at the ensemble mean vertical position,  $z = \bar{\eta}^{(\varrho)}(x, y, t)$ , the streamfunction,  $\bar{\Psi}^\#$ , defines an effective skew flux of the modified mean potential density given by

$$\bar{\mathbf{F}}^\# = -\nabla \tilde{\varrho} \times \bar{\Psi}^\#. \quad (17.98)$$

Using the identity  $\bar{\Psi}^\# = \bar{\mathbf{U}}^\# \times \hat{\mathbf{z}}$  (equation (17.84)) we can write this skew flux as

$$\bar{\mathbf{F}}^\# = -\bar{\mathbf{U}}^\# \partial_z \tilde{\varrho} + \hat{\mathbf{z}} \cdot \bar{\mathbf{U}}^\# \cdot \nabla_h \tilde{\varrho} \quad (17.99a)$$

$$= -(\bar{\mathbf{U}}^\# + \hat{\mathbf{z}} \cdot \mathbf{S} \cdot \bar{\mathbf{U}}^\#) \partial_z \tilde{\varrho}, \quad (17.99b)$$

where

$$\mathbf{S} = -\frac{\nabla_h \tilde{\varrho}}{\partial_z \tilde{\varrho}} \quad (17.100)$$

is the slope of the modified mean density originally introduced via equation (17.91), and  $\nabla_h = (\partial_x, \partial_y, 0)$  is the horizontal gradient operator taken with constant vertical position. The convergence of the effective skew flux leads to a stirring of the modified mean density  $\tilde{\varrho}$  at the mean vertical position,  $z = \bar{\eta}^{(\varrho)}$ ,

$$\partial_t \tilde{\varrho} = \nabla \cdot (\nabla \tilde{\varrho} \times \bar{\Psi}^\#). \quad (17.101)$$

This equation represents a geopotential coordinate specification of the evolution of the modified mean density due to stirring by the mean eddies. It corresponds directly to the evolution equation (17.92) satisfied at vertical position,  $z$ , by a single member of the  $\varrho$ -ensemble.

### 17.5.11 Approximate $\varrho$ -ensemble kinematics in geopotential coordinates

Equation (17.101) represents an exact  $z$ -coordinate description of the stirring of modified mean potential density. However, when working in geopotential coordinates, all that is available is Eulerian information. Hence, the isopycnal information used to realize this exact description can only be approximated.

### Approximating the quasi-Stokes transport

We require an approximation of the quasi-Stokes transport,  $\overline{U}^{\text{qs}}$ , defined by equation (17.82). We addressed a similar estimation in Section 17.3.4 when discussing the Stokes mean. Here, we expand the TRM transport in a Taylor series about the vertical position,  $z = \bar{\eta}^{(\varrho)}$

$$\overline{U}^\#(z) = \overline{\int_{\eta_b}^{z+\xi} \mathbf{u}(s) ds} \quad (17.102a)$$

$$= \overline{U}(z) + \overline{\mathbf{u} \xi}^{(z)} + \frac{1}{2} \overline{\partial_z \mathbf{u} \xi \xi}^{(z)} + \mathcal{O}(\alpha^3), \quad (17.102b)$$

where neglected terms are third order in deviation quantities. All ensemble means are taken at fixed vertical position (hence the  $z$  label on the overline), which accords with taking a Taylor series about the ensemble mean vertical position,  $z = \bar{\eta}^{(\varrho)}$ .

The ensemble means in equation (17.102b) are interpreted as follows. The first term is the Eulerian mean horizontal transport passing beneath the ensemble mean vertical position,  $z = \bar{\eta}^{(\varrho)}$ . The second term,  $\overline{\mathbf{u} \xi}$ , is the horizontal velocity evaluated at the ensemble mean vertical position and multiplied by the deviation,  $\xi$ , of the potential density surface from its mean vertical position, all averaged at fixed vertical position. An Eulerian split of the horizontal velocity,  $\mathbf{u}$ , into its Eulerian mean,  $\overline{\mathbf{u}}^{(z)}$ , and deviation,  $\mathbf{u}'$ , leads to the correlation

$$\overline{\mathbf{u} \xi}^{(z)} = \overline{\mathbf{u}' \xi}^{(z)}. \quad (17.103)$$

For the second order term, similar considerations lead to

$$\overline{\partial_z \mathbf{u} \xi \xi}^{(z)} \approx \partial_z \overline{\mathbf{u}^{(z)}} \overline{\xi \xi}^{(z)}, \quad (17.104)$$

where neglected terms are third order and higher. Combining these relations leads to the second order accurate expression

$$\overline{U}^\# \approx \overline{U} + \overline{\mathbf{u}' \xi}^{(z)} + \frac{1}{2} \overline{\xi \xi}^{(z)} \partial_z \overline{\mathbf{u}^{(z)}}. \quad (17.105)$$

### The disturbance field

Following the discussion in Section 17.3.7, we here determine the disturbance field,  $\xi$ , in terms of fields at constant vertical position. For this purpose, use the identity (17.69) to give

$$\tilde{\varrho}(z) = \varrho(z + \xi) \quad (17.106a)$$

$$= \varrho(z) + \partial_z \varrho(z) \xi + \frac{1}{2} \partial_{zz} \varrho(z) \xi^2 + \mathcal{O}(\alpha^3). \quad (17.106b)$$

Subtracting the Eulerian mean of equation (17.106b) from the unaveraged equation (17.106b), and noting that  $\tilde{\varrho}$  is already a mean field, leads to the second order accurate expression for the deviation

$$\xi = -\varrho'(z) / \partial_z \overline{\varrho}^{(z)} + \mathcal{O}(\alpha^2), \quad (17.107)$$

where

$$\varrho(z) = \overline{\varrho}^{(z)} + \varrho'(z). \quad (17.108)$$

To within the same order, the deviation can be written

$$\xi = -\varrho'(z) / \partial_z \tilde{\varrho}(z) + \mathcal{O}(\alpha^2). \quad (17.109)$$

### Approximate quasi-Stokes transport

Substituting the deviation (17.109) into the approximate expression (17.102b) for the TRM transport yields an approximate expression for the quasi-Stokes transport

$$\bar{U}^{\text{qs}} = -\frac{\bar{\mathbf{u}}' \bar{\varrho}'^{(z)}}{\partial_z \tilde{\varrho}} + \frac{\bar{\phi}^{(z)} \partial_z \bar{\mathbf{u}}^{(z)}}{(\partial_z \tilde{\varrho})^2} + \mathcal{O}(\alpha^3), \quad (17.110)$$

where

$$\bar{\phi}^{(z)} = \frac{1}{2} \bar{\varrho}' \bar{\varrho}'^{(z)} \quad (17.111)$$

is the mean potential density variance. [McDougall and McIntosh \(2001\)](#) noted that the [Gent et al. \(1995\)](#) scheme offers a parameterization of the two correlations on the right hand side of equation (17.110). We have more to say regarding this parameterization in Section 18.1.

Substituting the deviation (17.109) into the approximate expression (17.106b) yields, to within terms of third order, the relation

$$\tilde{\varrho} = \bar{\varrho}^{(z)} - \partial_z \left[ \frac{\bar{\phi}^{(z)}}{\partial_z \bar{\varrho}^{(z)}} \right] + \mathcal{O}(\alpha^3). \quad (17.112)$$

As for the Stokes transport, the modified mean density and Eulerian mean density, when evaluated at the same vertical position, differ by terms that are second order in eddy amplitude.

### 17.5.12 Further study

This section is largely based on presentations given by [DeSzeke and Bennett \(1993\)](#), [McIntosh and McDougall \(1996\)](#), [Kushner and Held \(1999\)](#), and [McDougall and McIntosh \(2001\)](#) as summarized in Section 9.3 of [Griffies \(2004\)](#). Many other papers have applied this formalism to a variety of analyses, with examples including [Nurser and Lee \(2004a\)](#), [Nurser and Lee \(2004b\)](#), [Young \(2012\)](#), [Wolfe \(2014\)](#), and [Jansen et al. \(2024\)](#). [Young \(2012\)](#) offers a particularly general and rewarding presentation making use of basic notions from tensor analysis.

## 17.6 Isopycnal mean tracer equation

We now include a tracer field to the perfect Boussinesq fluid and determine a mean field description for the tracer. The transport of tracer by eddies has both a reversible stirring component and an irreversible mixing component. The stirring arises from both the thickness correlation to velocity as well as the velocity correlated with tracer.

### 17.6.1 Thickness weighted average<sup>20</sup>

In equation (17.96) we introduced a specific thickness weighted average (or mean) operator, which is quite useful when considering the mean tracer equation. In general, for any field,  $\chi$ , associated with a potential density layer,  $\varrho$ , we define the decomposition into thickness weighted

---

<sup>20</sup>We also consider thickness weighted averaging for the stacked shallow water equations in Chapter 14. Many identities hold for both the continuously stratified fluid and the stacked shallow water.

average and deviation

$$\chi(\varrho) = \widehat{\chi}(\varrho) + \chi''(\varrho) \quad (17.113a)$$

$$= \frac{\overline{h\chi}^{(\varrho)}}{\overline{h}^{(\varrho)}} + \chi''. \quad (17.113b)$$

It follows by definition that the thickness weighted average of  $\chi''$  vanishes,

$$\overline{h\chi''}^{(\varrho)} = 0. \quad (17.114)$$

### 17.6.2 Isopycnal mean thickness weighted tracer equation

When attaching a tracer,  $C$ , to fluid elements, each member of a  $\varrho$ -ensemble satisfies the isopycnal tracer equation

$$\partial_t C + \mathbf{u} \cdot \nabla_{\mathbf{v}} C = 0. \quad (17.115)$$

Combining the tracer and thickness equations leads to the thickness weighted tracer equation

$$\partial_t (hC) + \nabla_{\mathbf{v}} \cdot (h\mathbf{u}C) = 0. \quad (17.116)$$

Hence, in isopycnal coordinates and in the absence of irreversible processes, the evolution of thickness weighted tracer occurs via the isopycnally oriented convergence of the two-dimensional thickness weighted horizontal advective flux,  $h\mathbf{u}C$ .

To address the problem of describing the ensemble mean tracer equation in isopycnal coordinates, decompose the tracer and velocity field into their thickness weighted average and deviation

$$\partial_t [h(\widehat{C} + C'')] + \nabla_{\mathbf{v}} \cdot [h(\widehat{\mathbf{u}} + \mathbf{u}'') (\widehat{C} + C'')] = 0. \quad (17.117)$$

Taking a mean over the  $\varrho$ -ensemble, and using equation (17.114), yield the ensemble mean thickness weighted tracer equation

$$\partial_t (\overline{h}^{(\varrho)} \widehat{C}) + \nabla_{\mathbf{v}} \cdot (\overline{h}^{(\varrho)} \widehat{C} \widehat{\mathbf{u}}) = -\nabla_{\mathbf{v}} \cdot (\overline{h} C'' \mathbf{u}''). \quad (17.118)$$

Now introduce the correlation,

$$\overline{h} C'' \mathbf{u}''^{(\varrho)} = \overline{h}^{(\varrho)} \widehat{C}'' \widehat{\mathbf{u}''}, \quad (17.119)$$

(see equation (17.113b)), and recall that the mean thickness  $\overline{h}^{(\varrho)}$  satisfies the mean thickness equation (17.95). These two points lead to the evolution equation for the mean thickness weighted tracer concentration

$$(\partial_t + \widehat{\mathbf{u}} \cdot \nabla_{\mathbf{v}}) \widehat{C} = -\frac{1}{\overline{h}^{(\varrho)}} \nabla_{\mathbf{v}} \cdot (\overline{h}^{(\varrho)} \widehat{C}'' \widehat{\mathbf{u}''}), \quad (17.120)$$

which can also be written in the flux form

$$\partial_t (\overline{h}^{(\varrho)} \widehat{C}) + \nabla_{\mathbf{v}} \cdot (\overline{h}^{(\varrho)} \widehat{\mathbf{u}} \widehat{C}) = -\nabla_{\mathbf{v}} \cdot (\overline{h}^{(\varrho)} \widehat{C}'' \widehat{\mathbf{u}''}). \quad (17.121)$$

### 17.6.3 Subgrid scale tracer transport tensor

The correlation between tracer and velocity found on the right-hand side of the mean thickness weighted tracer equation (17.120) is typically written in terms of a subgrid scale tracer transport tensor

$$\widehat{C'' \mathbf{u}''} = -\mathbb{J} \cdot \nabla_{\mathbf{r}_e} \widehat{C}. \quad (17.122)$$

This definition leads to the evolution equation

$$(\partial_t + \widehat{\mathbf{u}} \cdot \nabla_{\mathbf{r}_e}) \widehat{C} = \frac{1}{\bar{h}^{(\varrho)}} \nabla_{\mathbf{r}_e} \cdot (\bar{h}^{(\varrho)} \mathbb{J} \cdot \nabla_{\mathbf{r}_e} \widehat{C}), \quad (17.123)$$

which can also be written in the flux form

$$\partial_t (\bar{h}^{(\varrho)} \widehat{C}) + \nabla_{\mathbf{r}_e} \cdot (\bar{h}^{(\varrho)} \widehat{\mathbf{u}} \widehat{C}) = \nabla_{\mathbf{r}_e} \cdot (\bar{h}^{(\varrho)} \mathbb{J} \cdot \nabla_{\mathbf{r}_e} \widehat{C}). \quad (17.124)$$

The subgrid scale operator on the right hand side has the same general form as the diffusion operator written in isopycnal coordinates as derived in Section 12.15. However, in addition to symmetric diffusion processes, this operator includes skewed fluxes that lead to skew diffusion as discussed in Section 17.4.2. Whereas the diffusive aspect is commonly parameterized as dianeutral diffusion and neutral diffusion (Section 18.1), there is no parameterization for the skewed correlations for use in ocean models. We comment further on this situation in Section 18.3.8.

### 17.6.4 Summary of the tracer parameterization problem

Traditionally, the isopycnal parameterization problem for the evolution of the mean thickness weighted tracer requires a parameterization of the bolus velocity  $\mathbf{u}^{\text{bolus}}$ , which is related to the thickness weighted horizontal velocity via

$$\widehat{\mathbf{u}}(\varrho) = \frac{\bar{h} \mathbf{u}^{(\varrho)}}{\bar{h}^{(\varrho)}} = \bar{\mathbf{u}}^{(\varrho)} + \frac{\bar{h}' \mathbf{u}'^{(\varrho)}}{\bar{h}^{(\varrho)}} = \bar{\mathbf{u}}^{(\varrho)} + \mathbf{u}^{\text{bolus}}. \quad (17.125)$$

In addition to the bolus velocity, it is necessary to parameterize the subgrid scale tracer transport tensor

$$\widehat{C'' \mathbf{u}''} = -\mathbb{J} \cdot \nabla_{\mathbf{r}_e} \widehat{C}, \quad (17.126)$$

which generally has symmetric (diffusive) and antisymmetric (stirring) components (Section 17.4).

For a geopotential coordinate description, equation (17.75) is used to relate thickness weighted mean fields, defined as a function of  $\varrho$ , and TRM fields, defined as a function of the mean vertical position of  $\varrho$ , to write for the tracer field

$$\widehat{C}(x, y, \tilde{\varrho}, t) = \bar{C}^\#(x, y, \bar{\eta}^{(\varrho)}, t). \quad (17.127)$$

Equation (17.127), and the developed formalism, leads to the mean field tracer equation in geopotential coordinates

$$\partial_t \bar{C}^\# = \nabla \cdot (\nabla \bar{C}^\# \times \bar{\Psi}^\#) + R(\bar{C}^\#), \quad (17.128)$$

where  $R(\bar{C}^\#)$  is the geopotential coordinate form of the mixing/stirring operator on the right-

hand side of equation (17.123). Details for the transformation of the mixing/stirring operator from isopycnal to geopotential coordinates are provided in Section 12.15.

### 17.6.5 Comments

Much in this section follows from [Smith \(1999\)](#), [McDougall and McIntosh \(2001\)](#), [Young \(2012\)](#), and [Jansen et al. \(2024\)](#), each of which focused on the hydrostatic primitive equations assuming a vertically stable buoyancy stratification. The paper by [Young \(2012\)](#) is the first to formulate the ensemble mean primitive equations (continuity, tracer, momentum, vorticity, and energy equations) in a form where only the thickness weighted (residual mean) velocity appears. Hence, the formulation of [Young \(2012\)](#) eliminates the need to parameterize the [bolus velocity](#) or the [quasi-Stokes transport](#) since neither appear as separately identified terms. The paper by [Jansen et al. \(2024\)](#) further pursues the ideas from [Young \(2012\)](#) within the context of [generalized vertical coordinate](#) ocean models, and they identify some inconsistencies in how certain ocean models are implementing the eddy parameterizations. The topic of formulating the equations of motion remains an active topic of research for purposes of facilitating subgrid scale closure.



## 17.7 Exercises

### EXERCISE 17.1: STOKES DRIFT AND PSEUDOMOMENTUM

Consider a small amplitude disturbance field,  $\xi(\mathbf{x}, t)$  with an Eulerian mean velocity having zero shear,  $\partial_m \bar{v} = 0$ , so that the Stokes drift is given by equation (17.28)

$$\mathbf{v}^{(s)} = \overline{(\xi \cdot \nabla) \partial_t \xi} + \mathcal{O}(\alpha^3) = \overline{(\xi \cdot \nabla) \mathbf{v}'} + \mathcal{O}(\alpha^3) \quad \text{if } \partial_n \bar{v} = 0. \quad (17.129)$$

Show that the Stokes drift can be written as

$$\mathbf{p} = \mathbf{v}^{(s)} + \overline{\xi \times \omega'} - \frac{1}{2} \nabla \overline{\partial_t (\xi \cdot \xi)} + \mathcal{O}(\alpha^3), \quad (17.130)$$

where (using Cartesian tensors)

$$p_i = -\overline{(\partial_i \xi^j) \delta_{jk} v'^k} = -\overline{(\partial_i \xi^j) \delta_{jk} \partial_t \xi^k} \quad (17.131)$$

is known as the pseudomomentum of the disturbance.

### EXERCISE 17.2: AREA INTEGRATED TRACER FLUX

Following the GLM formalism in Section 17.4, consider the mean of the tracer flux as integrated over a static area

$$\mathcal{T} = \overline{\int_S \mathbf{v} C \cdot \hat{\mathbf{n}} dS}. \quad (17.132)$$

Extract the rotational term and use Stokes' theorem to write it as a contour integral around the boundary of the domain.

### EXERCISE 17.3: CONNECTING GLM AND STOKES DRIFT

Consider a pure wave field in which  $\bar{v} = 0$ . Connect the GLM transport to the symmetric and anti-symmetric transport tensors defined in Section 17.3.

## EXERCISE 17.4: MASSAGING THE MEAN FIELD TRACER EQUATION

Massage the mean field tracer equation (17.53) to bring it into a variety of flux-form expressions. Connect these expressions to the Stokes drift from Exercise 17.60.

## EXERCISE 17.5: LINEAR ROTATING WAVE

We illustrate some of the analysis of Section 17.4 by considering a particle displacement vector comprised of periodic and polarized motion in the horizontal plane

$$\xi(\mathbf{x}, t) = \Gamma [\hat{\mathbf{x}} \cos(\omega t) + \hat{\mathbf{y}} \sin(\omega t)] \quad (17.133a)$$

$$\partial_t \xi(\mathbf{x}, t) = \omega \Gamma [-\hat{\mathbf{x}} \sin(\omega t) + \hat{\mathbf{y}} \cos(\omega t)], \quad (17.133b)$$

where  $\Gamma > 0$  a time-independent amplitude and  $2\pi/\omega > 0$  is the period. Compute the rotational flux, skew flux, and advective flux for a tracer. Is there any evolution of the GLM tracer concentration?

## EXERCISE 17.6: SKEW FLUXES IN A SHALLOW WATER GRAVITY WAVE

In VOLUME 4 we study shallow water gravity waves. Here we compute tracer transport properties of these waves using the technology from Section 17.4, assuming the waves have small amplitude. For this purpose, consider the following  $x$ - $z$  traveling gravity waves with fluid particle velocity given by

$$\mathbf{v}' = U [\hat{\mathbf{x}} \cos(kx - \omega t) + \hat{\mathbf{z}}(z - \eta_b) |k| \sin(kx - \omega t)], \quad (17.134)$$

where  $\mathbf{k} = k \hat{\mathbf{x}}$  is the zonal wave vector,  $\omega > 0$  is the angular frequency (period =  $2\pi/\omega > 0$ ),  $z = \eta_b$  is the flat bottom, and  $U$  is the constant wave amplitude (dimensions of length per time). Compute the particle trajectories, assuming small amplitude waves. Show that the particle trajectories have a zero phase average, which means they serve as a suitable particle displacement vector for computation of wave properties and tracer transport as per Section 17.4. Compute those properties and discuss.

## EXERCISE 17.7: MEAN TRACER TRANSPORT BENEATH A DENSITY SURFACE

Making use of the formalism from Section 17.6, consider the mean horizontal tracer transport occurring beneath a particular potential density surface,  $\varrho = \tilde{\varrho}$ , which is here given by the TRM tracer concentration

$$\bar{C}^\#(\bar{\eta}^{(\varrho)}) = \overline{\int_{\eta_b}^{\bar{\eta}^{(\varrho)} + \xi} C \mathbf{u} dz}. \quad (17.135)$$

Connect this transport to the mean fields defined in Section 17.6.





## Chapter 18

### PARAMETERIZED OCEAN TRACER TRANSPORT

As discussed in Section 16.1, and following the conceptual picture of [Eckart \(1948\)](#), turbulent geophysical fluid flows affect a transfer of tracer variance to the small scales, as signaled by an increase in the magnitude of tracer gradients. This [downscale cascade](#) is facilitated by reversible stirring from balanced and unbalanced fluctuations (e.g., mesoscale eddies, submesoscale eddies, breaking gravity and lee waves, turbulent boundary layer processes). The cascade to progressively smaller scales eventually reaches the [Batchelor scale](#) (order millimetres; e.g., Section 11.5.1 of [Vallis \(2017\)](#)). At this scale, tracer gradients are sufficiently large in magnitude that molecular diffusion can readily act to dissipate tracer variance through irreversible diffusive mixing. Tracer transport at scales larger than the Batchelor scale is dominated by nearly reversible stirring, whereas transport at and below the Batchelor scale is dominated by irreversible mixing from molecular diffusion. This phenomenology provides a constraint on the form of the tracer equation to be used for coarse grained numerical models, where the model grid scale is generally much larger than the Batchelor scale.

In this chapter, we study certain of the mathematical and physical properties of parameterized advective and diffusive tracer transport. Such parameterizations aim to encapsulate key aspects of physical processes too small to observe and/or to simulate. This [subgrid scale](#) parameterization problem is far broader and deeper than available from a single chapter. We focus mostly on subgrid scale tracer advection and diffusion operators arising from ocean mesoscale eddy motions, yet even this limited focus involves far more than can be covered here. In particular, we do not discuss theories for how the eddy diffusivities are computed, which generally require studies of the momentum, energy, and vorticity budgets that are not considered here. Furthermore, we only consider parameterizations of the subgrid scale tracer flux, whose convergence provides a subgrid tendency for the coarse-grained tracer equation. Focusing on fluxes supports locality and conservation for the coarse-grained tracer equation, with these properties also shared by the uncoarsened tracer equation.

---

## CHAPTER GUIDE

We studied the physics of tracer diffusion in Chapter 15 and then advection-diffusion in Chapter 16. We also studied the kinematics of tracer transport in Chapter 17. The present chapter relies on that material, with a focus on the maths and physics of advective-diffusive parameterizations of tracer transport. We also make use of neutral directions as detailed in VOLUME 2. Mathematically, we rely on Cartesian tensor analysis from VOLUME 1.

The notation in this chapter is somewhat tedious, which arises from the many variants of tracer fluxes considered. Furthermore, there are many unanswered research questions about the suitability of certain parameterizations for ocean circulation models, particularly the anisotropic neutral diffusion discussed in Section 18.5 and the anisotropic Gent-McWilliams stirring in Section 18.6.

---

<b>18.1 Summarizing tracer transport parameterizations . . . . .</b>	<b>625</b>
18.1.1 A synopsis of ocean mixing processes . . . . .	625
18.1.2 A rough comparison . . . . .	626
18.1.3 Diffusive parameterization of fine scale mixing . . . . .	626
18.1.4 Advective-diffusive parameterization of eddy-induced transport . . . . .	627
18.1.5 Mathematical elements of eddy-induced stirring . . . . .	627
18.1.6 Dianeutral unit vector and the neutral slope . . . . .	628
<b>18.2 Expressions of small scale diffusion . . . . .</b>	<b>629</b>
18.2.1 Isotropic diffusion . . . . .	629
18.2.2 Vertical diffusion . . . . .	630
18.2.3 Dianeutral diffusion . . . . .	630
<b>18.3 Gent-McWilliams eddy-induced advection . . . . .</b>	<b>630</b>
18.3.1 Details of the parameterization . . . . .	631
18.3.2 Effects on buoyancy . . . . .	632
18.3.3 Local dissipation of available potential energy . . . . .	633
18.3.4 Connection to form stress . . . . .	635
18.3.5 Connection to isopycnal thickness diffusion . . . . .	636
18.3.6 Connection to Gent-McWilliams parameterization . . . . .	638
18.3.7 A parameterization based on a boundary value problem . . . . .	639
18.3.8 Comments . . . . .	640
<b>18.4 Neutral diffusion . . . . .</b>	<b>640</b>
18.4.1 Motivation for neutral diffusion . . . . .	640
18.4.2 Redi neutral diffusion . . . . .	641
18.4.3 Small slope neutral diffusion . . . . .	642
18.4.4 Neutral tangent plane neutral diffusion . . . . .	643
18.4.5 Neutrality condition . . . . .	644
18.4.6 Symmetry condition . . . . .	644
18.4.7 GM skewension plus small slope neutral diffusion . . . . .	644
18.4.8 Generalized vertical coordinates . . . . .	645
<b>18.5 Anisotropic neutral diffusion . . . . .</b>	<b>647</b>
18.5.1 Orthonormal triad of basis vectors . . . . .	647
18.5.2 Anisotropic neutral diffusion tensor . . . . .	648
18.5.3 Properties of the anisotropic neutral diffusive fluxes . . . . .	649
18.5.4 Small slope anisotropic neutral diffusion . . . . .	649
<b>18.6 Anisotropic Gent-McWilliams stirring . . . . .</b>	<b>651</b>

18.6.1	Streamfunction and anti-symmetric tensor . . . . .	651
18.6.2	Anisotropic Gent-McWilliams skew tracer flux . . . . .	651
18.6.3	Anisotropic GM skewson plus small slope neutral diffusion . . . . .	652
18.6.4	A parameterization based on a boundary value problem . . . . .	652
18.7	Exercises . . . . .	653

## 18.1 Summarizing tracer transport parameterizations

In this section we offer an outline of tracer transport parameterizations, starting from the small scales and moving to the mesoscale. We mostly focus on ocean applications, though similar arguments hold for the atmosphere as well.

### 18.1.1 A synopsis of ocean mixing processes

As reviewed in Chapters 15 and 16, irreversible mixing in the ocean takes place at the millimeter scale through the action of chaotic molecular motions that act to dissipate property gradients. This mixing is generally represented by downgradient molecular diffusion. The molecular diffusivity of matter (e.g., salt) in seawater is roughly  $10^{-9} \text{ m}^2 \text{ s}^{-1}$ , whereas the molecular thermal diffusivity is roughly 100 times larger (it is easier to diffuse enthalpy (heat) than matter, [Gill, 1982](#)). Reversible stirring by turbulent eddies greatly increases the magnitude of property gradients upon which molecular diffusion acts ([Eckart, 1948](#); [Nakamura, 2001](#); [Müller and Garrett, 2002](#)), thereby increasing the total amount of irreversible mixing. Motivated by the transport afforded by molecular diffusion and Brownian motion as studied in Chapter 15, and following the pioneering work on turbulent diffusion by [Taylor \(1922\)](#), it is common to parameterize mixing induced by turbulent eddy stirring as a diffusive closure with an eddy diffusivity that is far larger than molecular values. Furthermore, the eddy diffusivities are generally the same for all tracers since eddies generally act the same regardless the tracer. Double diffusive processes are the notable counter-example to this equivalence ([Schmitt, 1994](#)).

Mixing induced by eddies of length scale  $\mathcal{O}(\text{centimeters} - \text{meters})$  is associated with fine scale mixing processes such as gravitational instability, shear instability and breaking internal gravity waves ([MacKinnon et al., 2013](#)), as well as a suite of boundary layer processes ([Mellor and Yamada, 1982](#); [Large et al., 1994](#); [Kantha and Clayson, 2000](#); [Thorpe, 2005](#)). This mixing is commonly parameterized by a flow dependent isotropic eddy diffusivity. The magnitude of the eddy diffusivity is typically  $\mathcal{O}(10^{-3}-10^{-2} \text{ m}^2 \text{ s}^{-1})$  in boundary layers, and  $\mathcal{O}(10^{-5} \text{ m}^2 \text{ s}^{-1})$  in the relatively quiescent ocean interior ([Polzin et al., 1997](#); [Whalen et al., 2012](#); [Waterhouse et al., 2014](#)).

Ocean mesoscale eddies, with size  $\mathcal{O}(10 - 100) \text{ km}$  and Rossby numbers much less than unity, preferentially stir tracers along neutral directions ([McDougall, 1987a,b](#); [McDougall et al., 2014](#)). The mesoscale eddy stirring in turn induces a mixing that is parameterized by downgradient diffusion along neutral directions (Section 18.4). When feeling the geometric constraints of the surface boundary, mesoscale stirring leads to horizontal oriented mixing across outcropped density surfaces ([Treguier et al., 1997](#); [Ferrari et al., 2008](#); [Danabasoglu et al., 2008](#)). This mixing is parameterized by downgradient horizontal diffusion. The neutral and horizontal eddy diffusivities associated with mesoscale processes are typically  $\mathcal{O}(10^2 - 10^3 \text{ m}^2 \text{ s}^{-1})$  in the ocean interior and can rise to  $\mathcal{O}(10^4 \text{ m}^2 \text{ s}^{-1})$  in the ocean surface layer ([Abernathay et al., 2013](#); [Klocker and Abernathay, 2014](#); [Cole et al., 2015](#)).

### 18.1.2 A rough comparison

What process is more important for setting tracer distributions: neutral diffusion induced by mesoscale eddies or small scale isotropic diffusion induced by breaking gravity waves? Since the neutral diffusivity arises from mesoscale eddy stirring, it is many orders of magnitude larger than the isotropic diffusivity arising from fine scale mixing. However, these two eddy diffusivities act on very different tracer gradients, in which case the net effects on tracer distributions can be more comparable.

To help understand the issue, consider a scaling with a constant neutral diffusivity and a constant isotropic diffusivity. Furthermore, assume Cartesian orientation of the diffusion operators (i.e., zero neutral slope) and assume the isotropic diffusion is dominated by vertical diffusion (see Section 18.2). We are thus comparing the following two diffusion processes

$$\text{horizontal diffusion} = \kappa_{\text{horz}} \nabla_h^2 C \quad \text{and} \quad \text{vertical diffusion} = \kappa_{\text{vert}} \partial_{zz} C. \quad (18.1)$$

Now introduce a vertical length scale,  $H$ , and horizontal scale,  $L$ , over which the tracer concentration changes by the same amount  $\delta C$ . Doing so leads to the scaled diffusion operators

$$\text{horizontal diffusion} \sim (\kappa_{\text{horz}}/L^2) \delta C \quad \text{and} \quad \text{vertical diffusion} \sim (\kappa_{\text{vert}}/H^2) \delta C. \quad (18.2)$$

These operators have the same scale when

$$\kappa_{\text{vert}} = (H/L)^2 \kappa_{\text{horz}}. \quad (18.3)$$

Choosing  $L = 10^5$  m and  $H = 10^1$  m leads to

$$\kappa_{\text{vert}} = 10^{-8} \kappa_{\text{horz}}. \quad (18.4)$$

Furthermore, if  $\kappa_{\text{horz}} = 10^3$  m<sup>2</sup> s<sup>-1</sup>, then the two operators provide a similar contribution to tracer evolution if  $\kappa_{\text{vert}} = 10^{-5}$  m<sup>2</sup> s<sup>-1</sup>. This is a rather small diffusivity that is generally thought to be on the order of that afforded by the background of breaking gravity waves in the ocean interior ([MacKinnon et al., 2013](#)). This scaling is crude since the length scales are dependent on details of the flow regime, as are the eddy diffusivities. Even so, it suggests that the two diffusive processes can indeed contribute to tracer distributions by a similar amount.

### 18.1.3 Diffusive parameterization of fine scale mixing

Ignoring the cross-diffusion processes introduced in our discussion of ocean energetics in VOLUME 2 (see also [IOC et al. \(2010\)](#), Section 2.5 of [Olbers et al. \(2012\)](#), and [Graham and McDougall \(2013\)](#)), the molecular diffusion of  $\Theta$  and  $S$  lead to the material evolution equations

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot (\rho \kappa_\Theta \nabla \Theta) \quad (18.5a)$$

$$\rho \frac{DS}{Dt} = \nabla \cdot (\rho \kappa_S \nabla S), \quad (18.5b)$$

where  $\kappa_\Theta \approx 100 \kappa_S > 0$  are the molecular kinematic diffusivities for Conservative Temperature,  $\Theta$ , and salinity,  $S$ , respectively.

For a measured or simulated scale,  $\Delta$ , that is larger than the scale where gravity waves break and dissipate kinetic energy (i.e., tens to hundreds of metres), we commonly assume a diffusive parameterization for the associated irreversible tracer mixing (e.g., [MacKinnon et al.](#),

2013). Diffusion is also used to parameterize mixing from other small scale processes, such as turbulent boundary layer processes, double-diffusion, and breaking leewaves. As discussed in Section 4 of [McDougall et al. \(2014\)](#), small scale mixing generally takes place in an isotropic manner. Its parameterization thus appears just as for isotropic molecular diffusion given by equations (18.5a) and (18.5b), yet with a far larger eddy diffusivity  $\kappa_{\text{fine}} \gg \kappa_{\Theta}, \kappa_S$  that is a function of the flow

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot (\rho \kappa_{\text{fine}} \nabla \Theta) \quad (18.6a)$$

$$\rho \frac{DS}{Dt} = \nabla \cdot (\rho \kappa_{\text{fine}} \nabla S). \quad (18.6b)$$

The same eddy diffusivity is used for both  $\Theta$  and  $S$ . This assumption follows the general approach for turbulent transport parameterizations (e.g, [Vallis, 2017](#)), whereby eddies are assumed to act in the same manner on any conserved scalar tracer.

#### 18.1.4 Advection-diffusive parameterization of eddy-induced transport

Stirring from turbulent scales smaller than the grid scale is commonly parameterized by an **eddy-induced velocity**,  $\mathbf{v}^*$ . For ocean mesoscale eddies, such parameterized stirring generally follows a variant of [Gent and McWilliams \(1990\)](#) and [Gent et al. \(1995\)](#), with this stirring quite important for setting large-scale ocean tracer distributions. In addition, mixing is promoted by the direct cascade from stirring. This mixing is parameterized by **neutral diffusion**, which is distinct from that used for the small scale mixing discussed in Section 18.1.3.

Consider a second order subgrid scale eddy transport tensor,  $\mathbf{E}$ , meant to parameterize both subgrid scale eddy stirring and eddy mixing. With this tensor, the evolution of salinity and Conservative Temperature takes the form

$$\rho \frac{DS}{Dt} = \nabla \cdot (\rho \mathbf{E} \cdot \nabla S) \quad (18.7a)$$

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot (\rho \mathbf{E} \cdot \nabla \Theta). \quad (18.7b)$$

As for the fine scale diffusion equations (18.6a) and (18.6b), we here use the same transport tensor for both  $S$  and  $\Theta$  as eddies are assumed to act in the same manner on any conserved scalar tracer. As presented in Chapter 16, we decompose the second order transport tensor into the sum of its symmetric and anti-symmetric components

$$\mathbf{E} = \mathbf{K} + \mathbf{A}. \quad (18.8)$$

When the symmetric tensor,  $\mathbf{K}$ , is positive-definite, it gives rise to downgradient diffusion, whereas the **anti-symmetric tensor** (also skew-symmetric tensor),  $\mathbf{A}$ , gives rise to **skew diffusion** or eddy-induced advection affected by an **eddy-induced velocity**.

#### 18.1.5 Mathematical elements of eddy-induced stirring

As detailed in Sections 16.2, 16.4, and 17.4, the anti-symmetric tensor contributes to the parameterized transport according to

$$\nabla \cdot (\rho \mathbf{A} \cdot \nabla S) = \partial_m(\rho A^{mn} \partial_n S) \quad (18.9a)$$

$$= \partial_m(\rho A^{mn}) \partial_n S + \rho A^{mn} \partial_m \partial_n S \quad (18.9b)$$

$$= -\rho v^{*n} \partial_n S, \quad (18.9c)$$

where  $A^{mn}$  are the components to the anti-symmetric transport tensor,  $\mathbf{A}$ . Additionally, we noted that

$$\rho A^{mn} \partial_m \partial_n S = 0 \quad (18.10)$$

since  $A^{mn}$  is anti-symmetric on the indices  $m, n$  whereas  $\partial_m \partial_n S$  is symmetric. Finally, we introduced a density-weighted eddy-induced velocity

$$\rho v^{*n} = -\partial_m (\rho A^{mn}) \iff \rho \mathbf{v}^* = -\nabla \cdot (\rho \mathbf{A}). \quad (18.11)$$

Importantly,  $\rho \mathbf{v}^*$  has a zero divergence, again due to anti-symmetry of  $A^{mn}$

$$\nabla \cdot (\rho \mathbf{v}^*) = \partial_n (\rho v^{*n}) = -\partial_n \partial_m (\rho A^{mn}) = 0. \quad (18.12)$$

A zero-divergence for  $\rho \mathbf{v}^*$  means that it contributes no mass sources or sinks to the fluid.<sup>1</sup>

Transport from the parameterized anti-symmetric tensor thus provides a means to stir tracers with the eddy-induced velocity arising from unresolved eddy stirring. The mathematical form of the parameterized stirring can be either through skew-diffusion or through advection (see Section 16.4). Choosing to make use of the advection form allows us to combine the contribution from the anti-symmetric transport tensor with the resolved advection operator, thus resulting in a material transport equation making use of the residual mean velocity

$$\rho \frac{D^\dagger S}{Dt} = \nabla \cdot (\rho \mathbf{K} \cdot \nabla S) \quad (18.13a)$$

$$\rho \frac{D^\dagger \Theta}{Dt} = \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta), \quad (18.13b)$$

where the residual mean material time derivative is given by

$$\frac{D^\dagger}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^\dagger \cdot \nabla \quad (18.14)$$

and the residual mean velocity is

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*. \quad (18.15)$$

### 18.1.6 Dianeutral unit vector and the neutral slope

When considering closures for subgrid tracer mixing and stirring arising from ocean mesoscale motions, we orient the parameterized tracer fluxes according to locally defined buoyancy, as that reflects the physics of mesoscale motions.<sup>2</sup> We thus work with locally referenced Archimedean buoyancy to determine neutral directions. In particular, at each point in the fluid we orient stirring and mixing through unit directions that point in the dianeutral direction

$$\hat{\gamma} = \frac{-\alpha \nabla \Theta + \beta \nabla S}{|-\alpha \nabla \Theta + \beta \nabla S|} \quad \text{and} \quad \hat{\gamma} = \hat{\mathbf{x}} \hat{\gamma}_x + \hat{\mathbf{y}} \hat{\gamma}_y + \hat{\mathbf{z}} \hat{\gamma}_z \quad \text{and} \quad \hat{\gamma} \cdot \hat{\gamma} = 1, \quad (18.16)$$

<sup>1</sup>For a Boussinesq ocean, the density factor is replaced by the constant reference density,  $\rho$ , so that  $\nabla \cdot \mathbf{v}^* = 0$  in the Boussinesq ocean. See section 7 of Griffies and Greatbatch (2012) for more details of the Boussinesq and non-Boussinesq forms for the parameterized eddy-induced transport.

<sup>2</sup>We offer further discussion of this point at the start of Section 18.4.

with  $\hat{\gamma}$  pointing perpendicular to the neutral tangent plane in a direction towards larger density.<sup>3</sup> Furthermore, when the fluid is stably stratified in the vertical, which is common for the mesoscale and larger, then the squared buoyancy frequency is positive

$$N^2 = -g(-\alpha \partial_z \Theta + \beta \partial_z S) > 0. \quad (18.17)$$

We can thus introduce the slope of the neutral tangent plane relative to the  $(x, y)$  horizontal plane<sup>4</sup>

$$\mathbf{S} = - \begin{bmatrix} -\alpha \nabla_h \Theta + \beta \nabla_h S \\ -\alpha \partial_z \Theta + \beta \partial_z S \end{bmatrix} = \frac{g(-\alpha \nabla_h \Theta + \beta \nabla_h S)}{N^2} = \hat{\mathbf{x}} S_x + \hat{\mathbf{y}} S_y. \quad (18.18)$$

For such stably stratified fluids, the dianeutral direction can be written in terms of the neutral slope

$$\hat{\gamma} = \frac{\mathbf{S} - \hat{\mathbf{z}}}{(1 + \mathbf{S}^2)^{1/2}}. \quad (18.19)$$

In this form we see that the dianeutral direction is vertically downward when the slopes vanish (i.e., horizontal neutral directions), which accords with this direction generally pointing toward increasing density.

## 18.2 Expressions of small scale diffusion

We here follow Section 4 from [McDougall et al. \(2014\)](#) to highlight distinctions between isotropic diffusion, dianeutral diffusion,<sup>5</sup> and vertical diffusion. Although commonly considered interchangeable in the literature as parameterizations of small scale mixing, there are conceptual distinctions that we identify here. Note that the distinctions between these three diffusions are quantitatively small when neutral slopes are modest and when  $\kappa_{\text{ntr}} \gg \kappa_{\text{fine}}$ .

### 18.2.1 Isotropic diffusion

As discussed in Section 18.1.3, we generally parameterize fine scale mixing processes via an isotropic diffusion process using a diffusivity  $\kappa_{\text{fine}} > 0$ , diffusion tensor

$$\mathbf{K}^{\text{iso}} = \kappa_{\text{fine}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (18.20)$$

and corresponding diffusion flux

$$\mathbf{J}^{\text{iso}} = -\rho \mathbf{K}^{\text{iso}} \cdot \nabla C. \quad (18.21)$$

As illustrated in Figure 18.1, under the effects from isotropic diffusion, a region of tracer is diffused the same in all three directions so that, for example, a spherical tracer distribution remains spherical.

---

<sup>3</sup>Equation (4) in [McDougall et al. \(2014\)](#) makes use of the opposite convention so that their dianeutral direction points towards decreasing density. We instead follow the water mass transformation convention as in equation (20.38), so that  $\hat{\gamma}$  points in the direction of increasing density.

<sup>4</sup>Be careful to not confuse  $S$ , for salinity, with the slope components,  $S_x$  and  $S_y$ .

<sup>5</sup>Dianeutral diffusion is commonly also referred to as diapycnal diffusion, with diapycnal diffusion referring to diffusion across constant potential density surfaces. We distinguish dianeutral from diapycnal in this chapter since neutral directions are defined by locally referenced potential density, and as such neutral directions generally differ from isopycnals. Further discussion is provided in VOLUME 2 as well as [McDougall \(1987a\)](#) and [McDougall et al. \(2014\)](#).

### 18.2.2 Vertical diffusion

Because vertical density gradients are generally much larger than lateral gradients, it is common to approximate the small scale isotropic diffusion tensor with a vertical diffusion tensor

$$\mathbf{K}^{\text{vert}} = \kappa_{\text{fine}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (18.22)$$

with a corresponding vertical diffusive flux

$$\mathbf{J}^{\text{vert}} = -\rho \kappa_{\text{fine}} (\nabla C \cdot \hat{\mathbf{z}}) \hat{\mathbf{z}} = -\rho \mathbf{K}^{\text{vert}} \cdot \nabla C = -\rho \kappa_{\text{fine}} \partial_z C \hat{\mathbf{z}}. \quad (18.23)$$

In this manner, vertical mixing of a tracer patch occurs only in the vertical direction (see Figure 18.1).

### 18.2.3 Dianeutral diffusion

Dianeutral diffusion orients tracer fluxes according to the dianeutral direction (18.16)

$$\mathbf{J}^{\text{dia}} = -\rho \kappa_{\text{fine}} (\nabla C \cdot \hat{\boldsymbol{\gamma}}) \hat{\boldsymbol{\gamma}} = -\rho \mathbf{K}^{\text{dia}} \cdot \nabla C, \quad (18.24)$$

where the dianeutral diffusion tensor is given by

$$\mathbf{K}^{\text{dia}} = \kappa_{\text{fine}} \begin{bmatrix} \hat{\gamma}_x^2 & \hat{\gamma}_x \hat{\gamma}_y & \hat{\gamma}_x \hat{\gamma}_z \\ \hat{\gamma}_x \hat{\gamma}_y & \hat{\gamma}_y^2 & \hat{\gamma}_y \hat{\gamma}_z \\ \hat{\gamma}_x \hat{\gamma}_z & \hat{\gamma}_y \hat{\gamma}_z & \hat{\gamma}_z^2 \end{bmatrix}. \quad (18.25)$$

Assuming a vertically stable stratification, we can make use of the relation (18.19) to write  $\hat{\boldsymbol{\gamma}}$  in terms of the slope,  $\mathbf{S}$ , thus rendering

$$(\nabla C \cdot \hat{\boldsymbol{\gamma}}) \hat{\boldsymbol{\gamma}} = \frac{(\mathbf{S} - \hat{\mathbf{z}}) \cdot \nabla C}{1 + \mathbf{S}^2} (\mathbf{S} - \hat{\mathbf{z}}) = \frac{1}{1 + \mathbf{S}^2} \begin{bmatrix} S_x^2 & S_x S_y & -S_x \\ S_x S_y & S_y^2 & -S_y \\ -S_x & -S_y & 1 \end{bmatrix} \begin{bmatrix} \partial_x C \\ \partial_y C \\ \partial_z C \end{bmatrix}, \quad (18.26)$$

so that the dianeutral diffusion tensor now takes on the form

$$\mathbf{K}^{\text{dia}} = \frac{\kappa_{\text{fine}}}{1 + \mathbf{S}^2} \begin{bmatrix} S_x^2 & S_x S_y & -S_x \\ S_x S_y & S_y^2 & -S_y \\ -S_x & -S_y & 1 \end{bmatrix}. \quad (18.27)$$

As illustrated in Figure 18.1, dianeutral diffusion elongates a tracer patch in the direction normal to the neutral tangent plane.

## 18.3 Gent-McWilliams eddy-induced advection

As mentioned in Section 17.6.3, there are two processes that contribute to eddy-induced advection/stirring. One involves the correlations between eddy fluctuations in the velocity and tracer fields. In Section 17.4, we considered the kinematics of correlations induced by small amplitude eddying motions, where we found that the eddy-induced motion of fluid particles leads to both a symmetric (mixing) and anti-symmetric (stirring) dispersion of tracer concentrations.

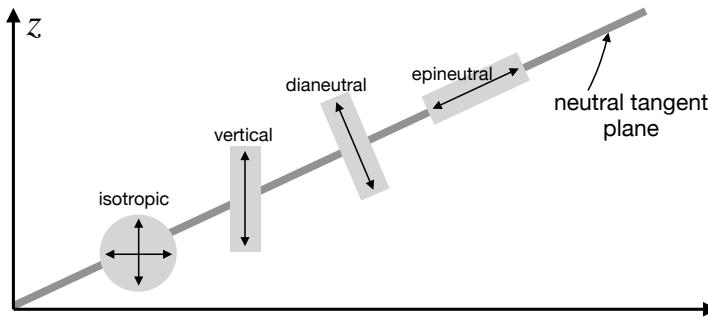


FIGURE 18.1: Illustrating the effects from various forms of diffusion on an initially spherical tracer patch. When diffused with an isotropic diffusion tensor (equation (18.20)), a spherical patch remains spherical. When diffused with a vertical diffusion tensor (equation (18.22)), a tracer patch elongates in the vertical direction. When diffused with a dianeutral diffusion tensor (equation (18.27)), a tracer patch elongates in the direction normal to the slanted neutral tangent plane. Finally, when diffused with a neutral diffusion tensor, such as the Redi tensor (18.72) or the small slope tensor (18.73), a tracer patch elongates along the neutral tangent plane; i.e., in the epineutral direction. This figure is adapted from Figure 4 of [McDougall et al. \(2014\)](#).

There is currently no method proposed for parameterizing this form of eddy-induced stirring when it arises from subgrid scale processes, thus leaving unanswered its importance to large-scale tracer distributions.

The second process leading to eddy-induced stirring arises from correlations between fluctuations in isopycnal layer thickness and horizontal velocity. As detailed in Section 17.5, this second effect leads to a movement of volume between isopycnal layers, or equivalently we can conceive of it as the quasi-Stokes transport of volume arising from transient eddy motion. This eddy-induced volume transport affects an eddy-induced tracer transport within isopycnal layers. Transient mesoscale eddies are the canonical dynamical process leading to this form of transport. For simulations that do not resolve transient mesoscale eddies, we commonly parameterize the subgrid scale stirring through variants of the [Gent et al. \(1995\)](#) scheme. Mathematical elements of this scheme are detailed in this section.

Most presentations of the [Gent et al. \(1995\)](#) scheme assume a Boussinesq ocean, with Section 7 of [Griffies and Greatbatch \(2012\)](#) an exception. We here present the non-Boussinesq form, though in places make the Boussinesq approximation since doing so simplifies the presentation without losing anything fundamental.<sup>6</sup>

### 18.3.1 Details of the parameterization

[Gent et al. \(1995\)](#) parameterize the three-dimensional non-divergent eddy-induced mass flux (recall Section 16.4.3) according to

$$\rho \mathbf{v}^* = \nabla \times (\rho \Psi^*) \quad \text{with} \quad \Psi^* = \hat{\mathbf{z}} \times \kappa_{\text{gm}} \mathbf{S}, \quad (18.28)$$

where  $\mathbf{S}$  is the neutral slope given by equation (18.18), and  $\kappa_{\text{gm}} > 0$  is a kinematic eddy diffusivity with dimensions of velocity times a length. As defined, the eddy-induced mass flux is determined by the neutral slope, with larger slopes, and thus more baroclinicity, leading to larger flux magnitudes.<sup>7</sup>

<sup>6</sup>For the Boussinesq ocean, the *in situ* density factor found throughout this section is set to the constant Boussinesq reference density,  $\rho_0$ . We study the Boussinesq ocean in VOLUME 2.

<sup>7</sup>Recalling our discussion of volume transport by eddying motion in Section 17.5, we can interpret the Gent-McWilliams parameterization (18.28) as a parameterization of the quasi-Stokes transport.

Performing the curl on the streamfunction leads to the horizontal and vertical components to the eddy-induced mass flux

$$\rho \mathbf{u}^* = -\partial_z (\kappa_{\text{gm}} \rho \mathbf{S}) \quad \text{and} \quad \rho w^* = \nabla_h \cdot (\kappa_{\text{gm}} \rho \mathbf{S}), \quad (18.29)$$

along with the skew diffusive tensor

$$\mathbf{A}^{\text{gm}} = \kappa_{\text{gm}} \begin{bmatrix} 0 & 0 & -S_x \\ 0 & 0 & -S_y \\ S_x & S_y & 0 \end{bmatrix}. \quad (18.30)$$

Following the discussion in Section 16.4.4, we identify the advective tracer flux, skew tracer flux, and rotational tracer flux

$$\mathbf{J}^{\text{adv}} = \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}} \quad (18.31)$$

where

$$\mathbf{J}^{\text{adv}} = C \rho \mathbf{v}^* = C \rho [-\partial_z (\kappa_{\text{gm}} \rho \mathbf{S}) + \hat{z} \nabla_h \cdot (\kappa_{\text{gm}} \rho \mathbf{S})] \quad (18.32\text{a})$$

$$\mathbf{J}^{\text{skew}} = -\nabla C \times \rho \Psi^* = \rho \kappa_{\text{gm}} [\mathbf{S} \partial_z C - \hat{z} (\mathbf{S} \cdot \nabla C)] \quad (18.32\text{b})$$

$$\mathbf{J}^{\text{rot}} = \nabla \times (\rho C \Psi^*). \quad (18.32\text{c})$$

### 18.3.2 Effects on buoyancy

To help understand the physics of the Gent-McWilliams parameterization (18.28), we here focus on how it affects buoyancy in a perfect Boussinesq ocean with a linear equation of state. Using potential density,  $\varrho$ , as a measure of buoyancy, we find the parameterized skew flux of potential density is given by

$$\rho_o^{-1} \mathbf{J}^{\text{skew}} = -\kappa_{\text{gm}} [\nabla_h \varrho - \hat{z} \mathbf{S}^2 \partial_z \varrho] = -\kappa_{\text{gm}} [\nabla_h \varrho + \hat{z} (\rho_o/g) (\mathbf{S} N)^2], \quad (18.33)$$

with the squared neutral slope and squared buoyancy frequency written

$$\mathbf{S}^2 = \mathbf{S} \cdot \mathbf{S} \quad \text{and} \quad N^2 = -\frac{g}{\rho_o} \frac{\partial \varrho}{\partial z}. \quad (18.34)$$

The parameterization yields a horizontal downgradient diffusive flux of potential density along with a vertical upgradient diffusive flux. Geostrophic turbulence is characterized by large Richardson numbers, so that the vertical stratification is stable ( $N^2 > 0$ ). As a result, the vertical component to the potential density skew flux is vertically downward, which corresponds to a vertically upward buoyancy skew flux. We illustrate this orientation of the skew flux components in Figure 18.2.

As we see in Section 18.3.3, this orientation of the skew flux ensures that the parameterization reduces available potential energy. Additionally, [Gent et al. \(1995\)](#) prescribe a diffusivity that vanishes at the ocean surface and ocean bottom. [McIntosh and McDougall \(1996\)](#) and [McDougall and McIntosh \(2001\)](#) present more discussion of the boundary conditions, which can be understood by considering the exact form of the quasi-Stokes transport defined by equation (17.82). Furthermore, we consider a boundary value problem approach in Section 18.3.7 that also pays particular attention to the boundary conditions.

Figure 18.3 brings elements of the parameterization together by illustrating the [Gent-McWilliams effect](#) for a meridional potential density front in the southern hemisphere. The

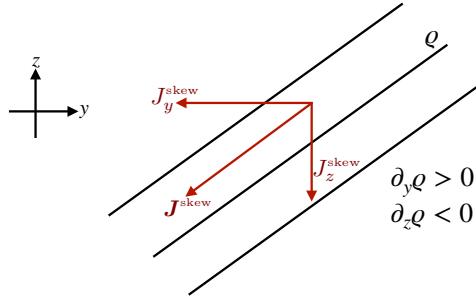


FIGURE 18.2: Orientation of the skew flux of potential density,  $\varrho$ , arising from the [Gent et al. \(1995\)](#) parameterization and as described by [Griffies \(1998\)](#). The sloped black lines are constant  $\varrho$  isosurfaces (isopycnals). The horizontal skew flux of potential density is downgradient (directed from high density to low density), whereas the vertical skew flux component is upgradient (directed from low density to high density). The net effect is a skew flux that is oriented parallel to isopycnals.

thermal wind flow is eastward, as in the Antarctic Circumpolar Current, whereas a parameterized [secondary circulation](#) acts to weaken the front, with the secondary circulation proportional to the strength of the front as measured by the isopycnal slope. That is, the Gent-McWilliams parameterization assumes that the mean effects from geostrophic eddies, whose kinetic energy is supported by the potential energy in the front, lead to a weakening of the potential density slope so that the front relaxes toward the horizontal. It is notable that on the scale of the resolved flow, the secondary circulation is [ageostrophic](#), with the geostrophic flow having a zero component in the cross-front direction.

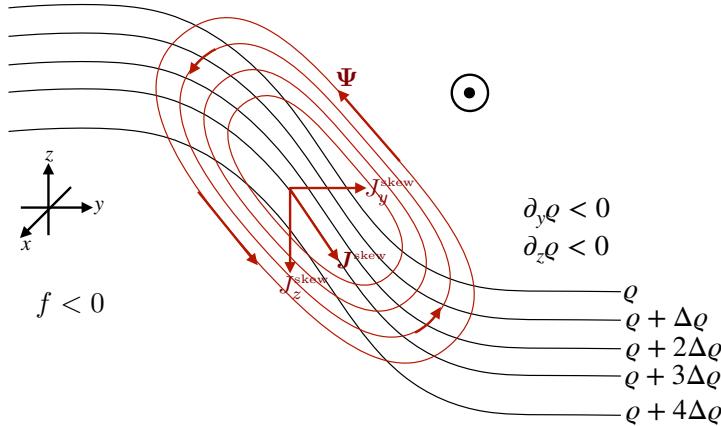


FIGURE 18.3: The [Gent-McWilliams effect](#) for a meridional potential density front in the Southern Hemisphere, where dense water rises to the south so that  $\partial_y \varrho < 0$ . The mean geostrophic thermal wind flow is eastward (out of the page), such as for the Antarctic Circumpolar Current. With a gravitationally stable stratification, whereby  $\partial_z \varrho < 0$  so that  $N^2 > 0$ , the Gent-McWilliams streamfunction given by equation (18.28) leads to a counter-clockwise [secondary circulation](#) (red isolines), which is ageostrophic. The dual perspective offered by the Gent-McWilliams skew flux for potential density, given by equation (18.33), reveals a northward (downgradient) meridional skew flux component along with a downward (upgradient) vertical component (red vectors). The result from either the streamfunction perspective or the skew flux perspective is a potential density tendency that relaxes the front towards the horizontal so to reduce the isopycnal slope.

### 18.3.3 Local dissipation of available potential energy

Now consider effects from the [Gent et al. \(1995\)](#) scheme on available potential energy, continuing to assume a Boussinesq ocean with a linear equation of state. Note that the parameterization, as an advective stirring process, preserves the quantity of fluid within density layers; i.e., there

is no dianeutral transport.<sup>8</sup> Hence, the change in potential energy is identical to the change in available potential energy. Results from this discussion support the intuition from Figure 18.3 that the scheme relaxes geostrophic fronts and thus reduces potential energy.

### Skew flux approach

Let us approach the parameterization problem from the perspective of satisfying two general properties: (I) the subgrid scale operator stirs while maintaining the same amount of fluid within isopycnal layers, (II) the subgrid operator locally dissipates potential energy through a rearrangement of the potential density surfaces, with the potential energy dissipation vanishing when there is zero baroclinicity. That is, the scheme dissipates available potential energy. What is the form of the stirring operator implied by these two assumptions?

Stirring of potential density can be realized via the convergence of a skew flux oriented parallel to potential density surfaces

$$\rho_o^{-1} \mathbf{J}^{\text{skew}} = -\nabla \varrho \times \Psi^*, \quad (18.35)$$

where at this point we have yet to specify  $\Psi^*$ . To see what the local dissipation of available potential energy imposes, consider the gravitational potential energy

$$P = g \int \varrho z \, dV, \quad (18.36)$$

where, again, we assume the *in situ* density equals to the potential density as per a linear equation of state. Assuming all boundaries are material and static allows us to focus on the time tendency of potential energy associated with the unknown flux

$$\frac{dP}{dt} = g \int z \frac{\partial \varrho}{\partial t} \, dV = -\frac{g}{\rho_o} \int (z \nabla \cdot \mathbf{J}^{\text{skew}}) \, dV = -\frac{g}{\rho_o} \int (z \partial_z J^z) \, dV = \frac{g}{\rho_o} \int J^z \, dV, \quad (18.37)$$

where  $J^z$  is the vertical flux component. We drop boundary effects by assuming the subgrid scale flux vanishes on all boundaries. To provide a local available potential energy sink requires

$$J^z \leq 0, \quad (18.38)$$

where zero occurs when the isopycnals are flat. It is sufficient to construct the vertical flux component using only the potential density field itself. For a stably stratified fluid in which  $\partial_z \varrho < 0$ , the following form provides a local available potential energy sink

$$\rho_o^{-1} J^z = \kappa_{\text{gm}} \mathbf{S}^2 \partial_z \varrho = -(\kappa_{\text{gm}} \rho_o / g) (\mathbf{S} \mathbf{N})^2 \leq 0. \quad (18.39)$$

The corresponding horizontal flux is given by a downgradient diffusive flux

$$\mathbf{J}^h = -\rho_o \kappa_{\text{gm}} \nabla_h \varrho. \quad (18.40)$$

We have thus recovered the skew flux (18.33) as proposed by [Gent et al. \(1995\)](#). Note that [Aiki et al. \(2004\)](#) proceed in a similar manner yet do not assume locality of the available potential

---

<sup>8</sup>In the ocean physics literature, such transport is often referred to as “adiabatic.” That terminology is appropriate if the density surfaces are parallel to Conservative Temperature surfaces, in which case there is no transfer of temperature as part of the stirring process. However, potential density surfaces are not parallel to temperature surfaces when the ocean has salinity, in which case we can mix temperature (and thus no longer remain adiabatic) even while maintaining a constant amount of fluid within density layers.

energy sink, thus deriving a more general subgrid scale operator.

### Advective flux approach

The impacts on potential energy should be the same when representing the parameterization as an advective flux. To verify this result, return to equation (18.37) and make use of the vertical component of the advective flux rather than the skew flux

$$\frac{dP}{dt} = -g \int z \mathbf{v}^* \cdot \nabla \varrho dV \quad (18.41a)$$

$$= g \int \varrho w^* dV \quad (18.41b)$$

$$= g \int \varrho \nabla_h \cdot (\kappa_{gm} \mathbf{S}) dV \quad (18.41c)$$

$$= g \int \nabla_h \cdot (\varrho \kappa_{gm} \mathbf{S}) dV - g \int \nabla_h \varrho \cdot \kappa_{gm} \mathbf{S} dV \quad (18.41d)$$

$$= -\rho_o \int \kappa_{gm} (\mathbf{S} N)^2 dV, \quad (18.41e)$$

which is the same result as for the skew flux. Note that to reach this result we dropped the boundary contribution, which follows since the closure assumes  $\kappa_{gm} = 0$  on boundaries.

#### 18.3.4 Connection to form stress

We now connect the [Gent et al. \(1995\)](#) closure, normally implemented in the tracer equation, to vertical transfer of momentum through form stress. For this purpose, we recall the general discussion of form stress in VOLUME 2, where we identify form stress as the horizontal pressure force acting on a sloped surface, with our present concern being form stresses acting on isopycnal surfaces.<sup>9</sup> For this purpose, assume the fluid is in Boussinesq frictional planetary geostrophic balance<sup>10</sup>, whereby the horizontal momentum equation is given by

$$\rho_o f (\hat{\mathbf{z}} \times \mathbf{u}) = -\nabla_h p + \partial_z \boldsymbol{\tau}, \quad (18.42)$$

with  $\boldsymbol{\tau}$  a horizontal subgrid scale stress vector. The Coriolis acceleration balances the acceleration from horizontal pressure gradients plus a vertical transfer of horizontal stress. The horizontal stress term is generally quite small in the ocean interior, where the flow is in geostrophic balance, whereas it is large at the ocean boundaries where it arises from turbulent boundary layer processes (e.g., turbulent wind stress and bottom stresses).

To make the connection between [Gent et al. \(1995\)](#) and the vertical transfer of horizontal form stress, add  $\rho_o f (\hat{\mathbf{z}} \times \mathbf{u}^*)$  to both sides of equation (18.42) to obtain

$$\rho_o f (\hat{\mathbf{z}} \times \mathbf{u}^\dagger) = -\nabla_h p + \partial_z \boldsymbol{\tau} + \rho_o f (\hat{\mathbf{z}} \times \mathbf{u}^*), \quad (18.43)$$

where  $\mathbf{u}^\dagger = \mathbf{u} + \mathbf{u}^*$  is the horizontal residual mean velocity. This equation says that the Coriolis acceleration from the horizontal residual mean velocity balances pressure gradients,

<sup>9</sup> [Young \(2012\)](#) provides a general method for making the connection between [Gent et al. \(1995\)](#) and form stress for a continuously stratified fluid. [Loose et al. \(2023\)](#) and [Jansen et al. \(2024\)](#) provide further theoretical and numerical analysis of this approach. For our more schematic purposes, we follow the treatment in [Greatbatch and Lamb \(1990\)](#), [Gent et al. \(1995\)](#) (their Section 4), [Ferreira and Marshall \(2006\)](#) (their Section 2), and [Zhao and Vallis \(2008\)](#) (their Section 2.2).

<sup>10</sup>We discuss planetary geostrophic vorticity balances in VOLUME 3.

the vertical divergence of the horizontal frictional stresses, plus the Coriolis acceleration from the eddy-induced velocity. We further unpack the eddy Coriolis acceleration by noting that the planetary geostrophic velocity satisfies the thermal wind relation in the ocean interior, whereby

$$f \partial_z \mathbf{u} = -(g/\rho_0) \hat{\mathbf{z}} \times \nabla \rho = -\hat{\mathbf{z}} \times N^2 \mathbf{S}. \quad (18.44)$$

We can thus write the Coriolis acceleration from the eddy-induced velocity as

$$f(\hat{\mathbf{z}} \times \mathbf{u}^*) = -f[\hat{\mathbf{z}} \times \partial_z(\kappa_{gm} \mathbf{S})] \quad (18.45a)$$

$$= -\partial_z[\hat{\mathbf{z}} \times (f \kappa_{gm} \mathbf{S})] \quad (18.45b)$$

$$= \frac{\partial}{\partial z} \left[ \frac{\kappa_{gm} f^2}{N^2} \frac{\partial \mathbf{u}}{\partial z} \right] \quad (18.45c)$$

$$= \partial_z(\nu_e \partial_z \mathbf{u}), \quad (18.45d)$$

where the final equality introduced an eddy-induced vertical viscosity

$$\nu_e \equiv \kappa_{gm} (f^2/N^2). \quad (18.46)$$

Making use of this result in the planetary geostrophic equation (18.43) thus leads to

$$\rho_0 f (\hat{\mathbf{z}} \times \mathbf{u}^\dagger) = -\nabla_h p + \partial_z(\boldsymbol{\tau} + \boldsymbol{\tau}_e), \quad (18.47)$$

where

$$\rho_0^{-1} \boldsymbol{\tau}_e = \nu_e \partial_z \mathbf{u} \quad (18.48)$$

defines a horizontal stress vector arising from the thermal wind shears and due to the mesoscale eddies. Equation (18.47) says that the Coriolis acceleration from the horizontal residual mean velocity is in balance with the horizontal pressure gradient plus the vertical transfer of horizontal shears arising from both friction/wind/bottom drag plus a contribution from parameterized mesoscale eddies.

We conclude that the [Gent et al. \(1995\)](#) parameterization appears in the planetary geostrophic residual mean momentum equation as a vertical transport of horizontal stress whose magnitude is set by a viscosity,  $\nu_e = \kappa_{gm} (f/N)^2$ . Notably, this vertical eddy transfer occurs in the absence of irreversible mixing. We thus interpret it as a parameterization of the vertical transfer of pressure form stress via mesoscale eddies that act between isopycnal layers. That is, the [Gent et al. \(1995\)](#) scheme offers a means to parameterize vertical transfer of horizontal form stress arising from undulating mesoscale eddies in the ocean interior. This interpretation is more thoroughly considered in our study of geostrophy in VOLUME 2 (see also [Greatbatch and Lamb \(1990\)](#) and [Loose et al. \(2023\)](#)).

### 18.3.5 Connection to isopycnal thickness diffusion

Recall the ensemble mean thickness equation (17.95) for a Boussinesq ocean as derived in Section 17.5.9

$$\partial_t h + \nabla_h \cdot (h \hat{\mathbf{u}}) = 0, \quad (18.49)$$

where

$$\hat{\mathbf{u}} = \mathbf{u} + \mathbf{u}^{bolus} \quad (18.50)$$

is the thickness weighted transport velocity affecting evolution of the ensemble mean specific thickness,  $\mathbf{h}$ . Note that for brevity we here drop the nomenclature  $(\cdot)^{(\varrho)}$  used in Section 17.5.9.

Isopycnal correlations of horizontal velocity and layer specific thickness define the bolus velocity via

$$\mathbf{h} \mathbf{u}^{\text{bolus}} = \overline{\mathbf{h}' \mathbf{u}'^{(\varrho)}} \quad (18.51)$$

Now consider a downgradient diffusive closure for this correlation

$$\mathbf{h} \mathbf{u}^{\text{bolus}} = \overline{\mathbf{h}' \mathbf{u}'^{(\varrho)}} = -\mathbf{K}^{\text{thick}} \cdot \nabla_{\mathbf{h}} h, \quad (18.52)$$

with  $\mathbf{K}^{\text{thick}}$  a symmetric and positive-definite  $2 \times 2$  diffusion tensor. The mean thickness equation thus takes the form of an advection-diffusion equation in isopycnal coordinates

$$\partial_t h + \nabla_{\mathbf{h}} \cdot (\mathbf{h} \mathbf{u}) = \nabla_{\mathbf{h}} \cdot (\mathbf{K}^{\text{thick}} \cdot \nabla_{\mathbf{h}} h). \quad (18.53)$$

We note one special property of the closure (18.52) revealed when considering discrete shallow water layers and assuming the thickness diffusion tensor is depth independent. Vertically summing the eddy transport from the ocean bottom up to a particular layer yields

$$\sum_{n=k_b}^{n=k} \overline{\mathbf{h}'_n \mathbf{u}'_n^{(\varrho)}} = - \sum_{n=k_b}^{n=k} \mathbf{K}^{\text{thick}} \cdot \nabla_{\mathbf{h}} h_n = -\mathbf{K}^{\text{thick}} \cdot \sum_{n=k_b}^{n=k} \nabla_{\mathbf{h}} h_n = -\mathbf{K}^{\text{thick}} \cdot \nabla_{\mathbf{h}} \eta_{k-1/2}, \quad (18.54)$$

where  $\eta_{k-1/2}$  is the upper interface of layer  $k$  (see Figure 18.4), and where  $k_b$  is the index for the layer at the ocean bottom. In this case we see that the eddy transport below an isopycnal interface is directly proportional to the slope of that interface.

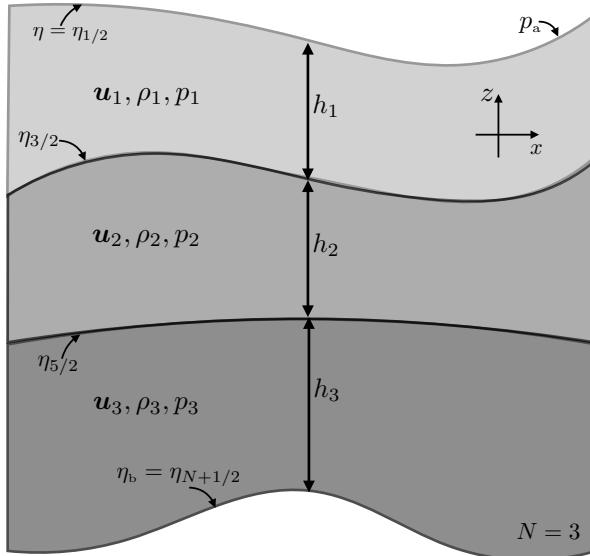


FIGURE 18.4: Three dynamically active layers of stacked shallow water fluid ( $N = 3$ ). In particular,  $\eta_{1/2}$  is the free surface,  $\eta_{1/2} = \eta$ , whereas  $\eta_{N+1/2} = \eta_b$  is the bottom interface. Hence, the total thickness of a column is  $h_1 + h_2 + h_3 = \eta_{1/2} - \eta_b$ . The “atmosphere” above the layers is assumed to apply a pressure,  $p_a$ , to the upper surface. The reduced gravity defined between each layer is given by  $g'_{k+1/2} = g(\rho_{k+1} - \rho_k)/\rho_{\text{ref}}$ . We take the reference density as  $\rho_{\text{ref}} = \rho_1$ , which results in a tidy set of layer equations. This figure is more thoroughly discussed in our study of shallow water wave models in VOLUME 4.

### 18.3.6 Connection to Gent-McWilliams parameterization

To make a direct connection between the thickness diffusion closure (18.52) and the [Gent et al. \(1995\)](#) closure discussed in Section 18.3.1, note that the specific thickness is the inverse of the vertical derivative of the potential density

$$h = (\partial_z \varrho)^{-1}. \quad (18.55)$$

Correspondingly, using the relation between the horizontal derivative operators,  $\nabla_h = \nabla + \mathbf{S} \partial_z$ , gives

$$h^{-1} \nabla_h h = -h \nabla_h (1/h) \quad \text{product rule identity} \quad (18.56a)$$

$$= -(\partial_z \varrho)^{-1} (\nabla_h + \mathbf{S} \partial_z) \partial_z \varrho \quad h = \partial z / \partial \varrho \text{ and } \nabla_h = \nabla_h + \mathbf{S} \partial_z \quad (18.56b)$$

$$= -\frac{\partial_z (\nabla_h \varrho)}{\partial_z \varrho} + \frac{\partial_{zz} \varrho \nabla_h \varrho}{(\partial_z \varrho)^2} \quad \text{rearrangement} \quad (18.56c)$$

$$= -\partial_z [\nabla_h \varrho / (\partial_z \varrho)] \quad \text{product rule identity} \quad (18.56d)$$

$$= \partial_z \mathbf{S} \quad \text{isopycnal slope } \mathbf{S} = -\nabla_h \varrho / (\partial_z \varrho). \quad (18.56e)$$

Consequently, the bolus velocity takes the form

$$\mathbf{u}^{\text{bolus}} = -h^{-1} \mathbf{K}^{\text{thick}} \cdot \nabla_h h = -\mathbf{K}^{\text{thick}} \cdot \partial_z \mathbf{S}. \quad (18.57)$$

#### The special case of depth independent diffusivity

For the special case of  $\mathbf{K}^{\text{thick}}$  that is independent of depth and proportional to the  $2 \times 2$  identity matrix, we recover the identity

$$\mathbf{u}^{\text{bolus}} = -\partial_z (\kappa_{\text{gm}} \mathbf{S}) = \mathbf{u}^*, \quad (18.58)$$

where the horizontal component of the [Gent et al. \(1995\)](#) velocity,  $\mathbf{u}^*$ , was identified from equation (18.29). Again, this identity holds only for the special case of a vertically independent diffusivity tensor proportional to the identity.

#### Further caveats

The relevance of a depth-independent diffusivity has been questioned by many authors, such as [Killworth \(1997\)](#), [Treguier et al. \(1997\)](#), [Smith and Vallis \(2002\)](#), [Smith and Marshall \(2009\)](#), and [Abernathy et al. \(2013\)](#). We conclude from these studies that a depth independent diffusivity is not the best choice for the [Gent et al. \(1995\)](#) parameterization, in which case where one places the vertical derivative is crucial.

The relation between thickness diffusion with the [Gent et al. \(1995\)](#) parameterization further breaks down near boundaries. The reason is that the eddy diffusivity vanishes next to boundaries and thus has a depth-dependence. Additionally, as noted by [Holloway \(1997\)](#) and [Griffies et al. \(2000\)](#), thickness diffusion next to solid earth boundaries leads to an increase in potential energy, with isopycnals creeping up the topographic slope. Such unphysical behavior motivates isopycnal modelers instead to use *interfacial height* diffusion to dissipate noise in the thickness field.

### 18.3.7 A parameterization based on a boundary value problem

There have been variants of the [Gent et al. \(1995\)](#) scheme proposed in the literature, such as those of [Aiki et al. \(2004\)](#) and [Ferrari et al. \(2010\)](#). As for the [Gent et al. \(1995\)](#) scheme, these alternatives dissipate available potential energy without mixing between isopycnal classes. We here briefly discuss the scheme of [Ferrari et al. \(2010\)](#), which is used by a variety of ocean climate models largely since it contains a natural means to numerically regularize the eddy-induced streamfunction in regions of weak vertical stratification. These considerations are relevant especially in ocean climate models, where weak or zero vertical stratification is inevitable and it is necessary to carefully handle such regimes.

For the [Ferrari et al. \(2010\)](#) scheme we write the parameterized eddy streamfunction as

$$\Psi^* = \hat{z} \times \Upsilon \implies \mathbf{u}^* = -\partial_z \Upsilon \quad \text{and} \quad w^* = \nabla_h \cdot \Upsilon, \quad (18.59)$$

with  $\Upsilon$  determined by solving the following vertical boundary value problem at each horizontal position<sup>11</sup>

$$(c^2 \partial_{zz} - N^2) \Upsilon = -N^2 \Upsilon^{\text{gm}} \quad \text{and} \quad \Upsilon(\eta_b) = \Upsilon(\eta) = 0, \quad (18.60)$$

where (see equation (18.28))

$$\Upsilon^{\text{gm}} = \kappa_{\text{gm}} \mathbf{S} \quad \text{and} \quad N^2 \mathbf{S} = (g/\rho_o) \nabla_h \varrho. \quad (18.61)$$

We recover the [Gent et al. \(1995\)](#) scheme when setting the squared speed to zero,  $c^2 = 0$ , in which case  $\Upsilon = \Upsilon^{\text{gm}}$ . For  $c^2 > 0$ , the second order differential operator ensures that  $\Upsilon$  smoothly and continuously transitions through regions where the vertical stratification is weak ( $N^2$  is small), and hence where  $|\mathbf{S}|$  is large. In contrast, the standard regularization approaches, with  $c^2 = 0$ , are somewhat more *ad hoc* (e.g., see Chapter 15 of [Griffies \(2004\)](#)) or very tedious (e.g., [Ferrari et al. \(2008\)](#)).

Following the discussion in Section 18.3.3, we deduce the impacts on potential energy (assuming a linear equation of state) via equation (18.37), where we make use of the vertical component of the potential density skew flux

$$\frac{1}{g} \frac{dP}{dt} = \frac{1}{\rho_o} \int J^z dV = - \int \hat{z} \cdot (\nabla \varrho \times \Psi^*) dV = - \int \nabla_h \varrho \cdot \Upsilon dV = -\frac{\rho_o}{g} \int N^2 \mathbf{S} \cdot \Upsilon dV. \quad (18.62)$$

The governing differential equation (18.60) leads to

$$\Upsilon \cdot (c^2 \partial_{zz} - N^2) \Upsilon = -(g/\rho_o) \kappa_{\text{gm}} \Upsilon \cdot \nabla_h \varrho, \quad (18.63)$$

which rearranges to

$$(g/\rho_o) \kappa_{\text{gm}} \Upsilon \cdot \nabla_h \varrho = -c^2 \partial_z (\Upsilon \cdot \partial_z \Upsilon) + c^2 (\partial_z \Upsilon \cdot \partial_z \Upsilon) + N^2 \Upsilon \cdot \Upsilon. \quad (18.64)$$

Integrating over a vertical column and making use of the homogeneous Dirichlet boundary conditions from equation (18.60) leads to

$$\frac{g}{\rho_o} \int \kappa_{\text{gm}} \Upsilon \cdot \nabla_h \varrho dz = \int (c^2 \partial_z \Upsilon \cdot \partial_z \Upsilon + N^2 \Upsilon \cdot \Upsilon) dz \geq 0. \quad (18.65)$$

This inequality means that the potential energy of a vertical column is dissipated. However,

<sup>11</sup>Note that [Ferrari et al. \(2010\)](#) used the opposite sign convention on  $\Upsilon$  from that used here.

locally at any point in the column the potential energy might increase due to the sign-indefinite term,  $-c^2 \partial_z(\mathbf{Y} \cdot \partial_z \mathbf{Y})$ . Notably, there is no *a priori* reason that mesoscale eddies dissipate potential energy locally at every point in space. Furthermore, numerical experiments documented in [Ferrari et al. \(2010\)](#) suggest that local potential energy dissipation is not necessary for a numerically stable operator. We conclude that this approach offers a suitable method for ocean climate simulations.

### 18.3.8 Comments

As noted in Section 17.6.3, there is presently no parameterization of subgrid scale stirring along neutral directions arising from the correlations between tracer and velocity fluctuations. Rather, the only parameterized subgrid scale stirring is associated with quasi-Stokes transport, with [Gent et al. \(1995\)](#) providing the canonical approach as summarized in this section. To parameterize the skew fluxes arising from tracer-velocity correlations requires one to study the polarization of the eddies giving rise to these skew fluxes, as per the discussion in Section 17.4.2 and [Middleton and Loder \(1989\)](#).

## 18.4 Neutral diffusion

Neutral diffusion, also referred to as epineutral diffusion, parameterizes the lateral mixing induced by mesoscale eddy stirring. The parameterization assumes that the neutral diffusive flux of a tracer is oriented along neutral directions, which are oriented on a neutral tangent plane. Consequently, the neutral diffusive tracer flux,  $\mathbf{J}$ , for an arbitrary tracer,  $C$ , is perpendicular to the dianeutral direction

$$\mathbf{J} \cdot \hat{\gamma} = 0 \implies \mathbf{J} \cdot (-\alpha \nabla \Theta + \beta \nabla S) = 0, \quad (18.66)$$

where  $\hat{\gamma}$  is defined by equation (18.16).

### 18.4.1 Motivation for neutral diffusion

Pioneering numerical models of the ocean general circulation, such as [Cox and Bryan \(1984\)](#), were formulated with the tracer mixing tensor oriented according to the horizontal and vertical directions, which corresponded to the orientation of the discrete grid cells. These simulations exhibited problems near strong density fronts, such as those found in western boundary currents. In such regions, the horizontally oriented tracer diffusion spuriously fluxed temperature and salinity across isopycnals, thus degrading the strength of the front and leading to, among other problems, unphysically weak meridional heat transport ([Böning et al., 1995](#)). In earlier work based on tracer measurements, [Montgomery \(1938\)](#), [Veronis \(1975\)](#), and [Rooth \(1982\)](#) suggested that ocean properties are preferentially homogenized along local potential density surfaces rather than geopotential surfaces. Such measurements motivated [Solomon \(1971\)](#) and [Redi \(1982\)](#) to propose rotating the tracer mixing tensor according to isopycnal directions, which were later generalized to neutral directions.

We offer further indirect evidence that mesoscale eddy induced diffusion is preferentially aligned along neutral directions. For that purpose, consider a diffusive flux that is not aligned with neutral directions. In this case, diffusive mixing can cause tracer distributions to cross neutral directions, thus adding to the mixing that is already parameterized from small scale mixing processes from Section 18.2. As discussed in Section 14.1.5 of [Griffies \(2004\)](#) as well as

Section 1 of [McDougall et al. \(2014\)](#), the extra mixing induced by this non-neutral orientation of the mesoscale induced diffusive fluxes is proportional to the squared slope between the proposed new direction and the neutral tangent plane. Estimates based on field measurements for interior ocean mixing constrain the magnitude of the miss-alignment to be less than  $10^{-4}$ . This number is very small, indeed it is zero within error bars of field measurements. Although the measurements are sparse, they do support the use of mesoscale eddy induced diffusive fluxes that are oriented according to neutral directions. We thus make use of this constraint in designing the diffusion tensor in the remainder of this section.<sup>12</sup>

### 18.4.2 Redi neutral diffusion

One diffusive flux satisfying the property (18.66) is given by

$$\mathbf{J}^{\text{redi}} = -\rho \kappa_{\text{ntr}} [\nabla C - \hat{\gamma} (\hat{\gamma} \cdot \nabla C)], \quad (18.67)$$

where  $\kappa_{\text{ntr}} > 0$  is the eddy neutral diffusivity (dimensions of squared length per time). In Figure 18.5 we illustrate the diffusive flux arising for a particular configuration of the neutral directions and the tracer concentration.

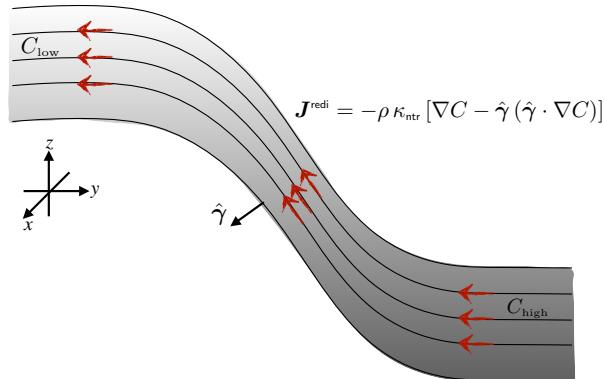


FIGURE 18.5: Schematic of the Redi neutral diffusion flux,  $\mathbf{J}^{\text{redi}} = -\rho \kappa_{\text{ntr}} [\nabla C - \hat{\gamma} (\hat{\gamma} \cdot \nabla C)]$ , from equation (18.67). The sloping lines represent surfaces whose local tangent define the neutral tangent plane with neutral directions, and with the dianeutral direction,  $\hat{\gamma}$ , perpendicular to the neutral directions. The tracer flux is aligned parallel to these surfaces. We here depict the case with higher tracer concentration on the right side so that the downgradient neutral diffusive flux is to the left and upward.

We confirm that  $\mathbf{J}^{\text{redi}}$  is oriented down the tracer gradient by noting that

$$\mathbf{J}^{\text{redi}} \cdot \nabla C = -\rho \kappa_{\text{ntr}} [|\nabla C|^2 - (\hat{\gamma} \cdot \nabla C)^2] \leq 0. \quad (18.68)$$

We can write the neutral diffusive flux (18.67) in the downgradient flux-gradient relation

$$\mathbf{J}^{\text{redi}} = -\rho \mathbf{K}^{\text{redi}} \cdot \nabla C, \quad (18.69)$$

<sup>12</sup>There have been attempts to provide a first principles argument for aligning the subgrid diffusive tracer operators along neutral directions. [McDougall et al. \(2014\)](#) provide a summary of such arguments. Currently, the strongest argument in support of neutral diffusion remains empirical.

with the neutral diffusion tensor,  $\mathbf{K}^{\text{redi}}$ , given by

$$\mathbf{K}^{\text{redi}} = \kappa_{\text{ntr}} \begin{bmatrix} 1 - \hat{\gamma}_x^2 & -\hat{\gamma}_x \hat{\gamma}_y & -\hat{\gamma}_x \hat{\gamma}_z \\ -\hat{\gamma}_x \hat{\gamma}_y & 1 - \hat{\gamma}_y^2 & -\hat{\gamma}_y \hat{\gamma}_z \\ -\hat{\gamma}_x \hat{\gamma}_z & -\hat{\gamma}_y \hat{\gamma}_z & 1 - \hat{\gamma}_z^2 \end{bmatrix} \implies (K^{\text{redi}})^{mn} = \kappa_{\text{ntr}} (\delta^{mn} - \hat{\gamma}^m \hat{\gamma}^n). \quad (18.70)$$

The corresponding neutral diffusion operator is given by the three-dimensional flux convergence

$$\mathcal{R}^{\text{redi}} = -\nabla \cdot \mathbf{J}^{\text{redi}} = \nabla \cdot (\rho \mathbf{K}^{\text{redi}} \cdot \nabla C). \quad (18.71)$$

When the neutral tangent planes are stably stratified in the vertical, so that their slopes are bounded, then the diffusion tensor takes the following form originally suggested by *Redi* (1982), which is written in terms of the neutral slope

$$\mathbf{K}^{\text{redi}} = \frac{\kappa_{\text{ntr}}}{1 + \mathbf{S}^2} \begin{bmatrix} 1 + S_y^2 & -S_x S_y & S_x \\ -S_x S_y & 1 + S_x^2 & S_y \\ S_x & S_y & \mathbf{S}^2 \end{bmatrix}. \quad (18.72)$$

### 18.4.3 Small slope neutral diffusion

Another form of the neutral diffusion flux is based on assuming a small magnitude for the slope of the neutral tangent plane relative to the horizontal, which is the case for mesoscale fronts and eddies. With this approximation, the small slope neutral diffusion tensor takes the form

$$\mathbf{K}^{\text{small}} = \kappa_{\text{ntr}} \begin{bmatrix} 1 & 0 & S_x \\ 0 & 1 & S_y \\ S_x & S_y & \mathbf{S}^2 \end{bmatrix}. \quad (18.73)$$

The corresponding small slope neutral diffusive flux is

$$\mathbf{J}^{\text{small}} = -\rho \kappa_{\text{ntr}} [\nabla_{\text{h}} + \hat{\mathbf{z}} (\mathbf{S} \cdot \nabla_{\text{h}})] C \quad (18.74)$$

where

$$\nabla_{\text{h}} = \nabla_{\text{h}} + \mathbf{S} \partial_z \quad (18.75)$$

is the horizontal derivative operator computed on the neutral tangent plane (see equation (12.85)). To show that  $\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = 0$ , we make use of the identity (18.19) so that

$$\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = \frac{\mathbf{J}^{\text{small}} \cdot \mathbf{S} - \mathbf{J}^{\text{small}} \cdot \mathbf{S}}{(1 + \mathbf{S}^2)^{1/2}} = 0. \quad (18.76)$$

Furthermore, we confirm that  $\mathbf{J}^{\text{small}}$  is oriented down the tracer gradient by noting that

$$\mathbf{J}^{\text{small}} \cdot \nabla C = -\rho \kappa_{\text{ntr}} [\nabla_{\text{h}} C \cdot \nabla_{\text{h}} C + (\mathbf{S} \cdot \nabla_{\text{h}} C) \partial_z C] \quad (18.77a)$$

$$= -\rho \kappa_{\text{ntr}} [|\nabla_{\text{h}} C|^2 + 2(\mathbf{S} \cdot \nabla_{\text{h}} C) \partial_z C + |\mathbf{S} \partial_z C|^2] \quad (18.77b)$$

$$= -\rho \kappa_{\text{ntr}} |\nabla_{\text{h}} C + \mathbf{S} \partial_z C|^2 \quad (18.77c)$$

$$= -\rho \kappa_{\text{ntr}} |\nabla_{\text{h}} C|^2 \quad (18.77d)$$

$$\leq 0. \quad (18.77e)$$

The small slope approximation was proposed by *Cox* (1987). However, his form for the

small slope neutral diffusion flux was incorrect as it did not satisfy  $\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = 0$ . The corrected form given by equation (18.74) was first written by [Gent and McWilliams \(1990\)](#). The resulting small slope neutral diffusion operator is commonly used in ocean climate models ([Griffies et al., 1998](#); [Lemarié et al., 2012](#)), which results from computing the three-dimensional convergence

$$\mathcal{R}^{\text{small}} = -\nabla \cdot \mathbf{J}^{\text{small}} = \nabla_h \cdot (\rho \kappa_{\text{ntr}} \nabla_{h\gamma} C) + \partial_z (\rho \kappa_{\text{ntr}} \mathbf{S} \cdot \nabla_{h\gamma} C). \quad (18.78)$$

In Exercise 18.5 we compute the flux from the [Cox \(1987\)](#) tensor and show why it is physically wrong.

#### 18.4.4 Neutral tangent plane neutral diffusion

A third method to compute neutral diffusion is motivated by the form of isopycnal diffusion in isopycnal layered models (e.g., see Chapter 14). Rather than isopycnal layers, we work with layers determined locally by neutral tangent planes. The neutral tangent frame makes use of projected non-orthogonal generalized vertical coordinates detailed in Chapter 12.

Following the derivations given in Section 12.15, the neutral diffusive flux in the neutral tangent frame is given by the horizontal flux

$$\mathbf{J}^{\text{ntp}} = -\rho \kappa_{\text{ntr}} \nabla_{h\gamma} C. \quad (18.79)$$

This flux is oriented down the tracer gradient as oriented along neutral directions

$$\mathbf{J}^{\text{ntp}} \cdot \nabla_{h\gamma} C = -\rho \kappa_{\text{ntr}} |\nabla_{h\gamma} C|^2, \quad (18.80)$$

which is the same as equation (18.77d) for the small slope fluxes. However, as a purely horizontal flux,  $\mathbf{J}^{\text{ntp}}$  is not oriented along neutral directions

$$\mathbf{J}^{\text{ntp}} \cdot \hat{\gamma} \neq 0. \quad (18.81)$$

Nevertheless, rather than computing the neutral diffusion operator as a horizontal convergence of this flux, the neutral tangent plane diffusion operator is computed by taking the convergence of  $\mathbf{J}^{\text{ntp}}$  along the neutral tangent plane as per equation (12.107)

$$\mathcal{R}^{\text{ntp}} = -\frac{1}{h^\gamma} [\nabla_{h\gamma} \cdot (h^\gamma \mathbf{J}^{\text{ntp}})] = \frac{1}{h^\gamma} [\nabla_{h\gamma} \cdot (h^\gamma \rho \kappa_{\text{ntr}} \nabla_{h\gamma} C)], \quad (18.82)$$

where

$$h^\gamma = \frac{\partial z}{\partial \gamma} d\gamma = - \left[ \frac{g}{\rho_0 N^2} \right] d\gamma \quad (18.83)$$

measures the thickness of a layer defined by two neutral tangent planes (see equation (12.104)).

As detailed in Section 12.15,  $\mathcal{R}^{\text{ntp}}$  is identical to the small slope neutral diffusion operator (18.78)

$$\mathcal{R}^{\text{ntp}} = \mathcal{R}^{\text{small}}. \quad (18.84)$$

In principle, it is a matter of convenience which form of the operator one uses. However, there are certain issues to consider when implementing these operators in a numerical model. Notably, a discrete realization of  $\mathcal{R}^{\text{ntp}}$  allows for a diagonal downgradient implementation of neutral diffusion, just as isopycnal diffusion in an isopycnal ocean model. In contrast, a discrete realization of either  $\mathcal{R}^{\text{redi}}$  or  $\mathcal{R}^{\text{small}}$  cannot guarantee downgradient fluxes due to the off-diagonal nature of its neutral diffusive flux components ([Griffies et al. \(1998\)](#), [Beckers et al. \(1998\)](#),

[Gnanadesikan \(1999\)](#), [Beckers et al. \(2000\)](#) [Lemarié et al. \(2012\)](#), [Shao et al. \(2020\)](#)). As a result, discrete realizations of  $\mathcal{R}^{\text{redi}}$  or  $\mathcal{R}^{\text{small}}$  can produce extrema, which are distinctly not properties of diffusion in the continuum (see Exercise 15.4). Hence, even though the continuum identity holds  $\mathcal{R}^{\text{ntp}} = \mathcal{R}^{\text{small}}$ , there are important differences that arise upon realizing these operators on a discrete lattice. [Shao et al. \(2020\)](#) provide further discussion of these points as part of their numerical realization of neutral diffusion.

### 18.4.5 Neutrality condition

Given the expression (18.16) for the dianeutral unit vector,  $\hat{\gamma}$ , it is straightforward to show that the neutral diffusive flux for Conservative Temperature and salinity satisfy the constraints

$$\nabla\Theta \cdot [-\alpha \mathbf{J}(\Theta) + \beta \mathbf{J}(S)] = 0 \quad \text{and} \quad \nabla S \cdot [-\alpha \mathbf{J}(\Theta) + \beta \mathbf{J}(S)] = 0. \quad (18.85)$$

These constraints are generally satisfied if the diffusive fluxes satisfy the balance

$$\alpha \mathbf{J}(\Theta) = \beta \mathbf{J}(S) \implies \mathbf{K} \cdot \hat{\gamma} = 0. \quad (18.86)$$

We refer to this balance as the [neutrality condition](#). It reflects the vanishing of the neutral diffusive flux of locally referenced potential density. It is maintained by the diffusive flux (18.67) of [Redi \(1982\)](#), the small slope flux (18.74) of [Gent and McWilliams \(1990\)](#), and the neutral tangent frame neutral diffusive flux (18.79). However, it is not maintained by the small slope fluxes from [Cox \(1987\)](#). Furthermore, [Griffies et al. \(1998\)](#) argued for the importance of maintaining this balance to avoid a nonlinear instability plaguing certain numerical realizations of neutral diffusion such as that from [Cox \(1987\)](#).

### 18.4.6 Symmetry condition

Since the neutral diffusion tensor is symmetric (as are all diffusion tensors), we have

$$\mathbf{J}(\Theta) \cdot \nabla S = -\kappa_{\text{ntr}} \rho K^{mn} \partial_n \Theta \partial_m S \quad (18.87a)$$

$$= -\kappa_{\text{ntr}} \rho K^{nm} \partial_n S \partial_m \Theta \quad (18.87b)$$

$$= -\kappa_{\text{ntr}} \rho K^{nm} \partial_n S \partial_m \Theta \quad (18.87c)$$

$$= \mathbf{J}(S) \cdot \nabla \Theta. \quad (18.87d)$$

This symmetry condition holds for any of the diffusion tensors introduced in this chapter. It is particularly useful in our discussion of cabbeling and thermobaricity in Section 19.3.

### 18.4.7 GM skewusion plus small slope neutral diffusion

A parameterization of mesoscale eddy stirring and mixing often appears in geopotential coordinate ocean models in the form of GM skewusion (Section 18.3.1) and small slope neutral diffusion (Section 18.4.3). The combined tracer flux takes the form

$$\rho^{-1} \mathbf{J} = -\kappa_{\text{ntr}} \nabla_h C - (\kappa_{\text{ntr}} - \kappa_{\text{gm}}) \mathbf{S} \partial_z C - \hat{\mathbf{z}} [(\kappa_{\text{ntr}} + \kappa_{\text{gm}}) \mathbf{S} \cdot \nabla_h C + \kappa_{\text{ntr}} \mathbf{S}^2 \partial_z C], \quad (18.88)$$

which can be written in terms of a subgrid scale transport tensor

$$\rho^{-1} \begin{bmatrix} J^x \\ J^y \\ J^z \end{bmatrix} = - \begin{bmatrix} \kappa_{\text{ntr}} & 0 & (\kappa_{\text{ntr}} - \kappa_{\text{gm}}) S_x \\ 0 & \kappa_{\text{ntr}} & (\kappa_{\text{ntr}} - \kappa_{\text{gm}}) S_y \\ (\kappa_{\text{ntr}} + \kappa_{\text{gm}}) S_x & (\kappa_{\text{ntr}} + \kappa_{\text{gm}}) S_y & \kappa_{\text{ntr}} \mathbf{S}^2 \end{bmatrix} \begin{bmatrix} \partial_x C \\ \partial_y C \\ \partial_z C \end{bmatrix}. \quad (18.89)$$

In the 1990s and throughout much of the 2000s, it was common to assume that  $\kappa_{\text{ntr}} = \kappa_{\text{gm}}$ , in which case the combined mixing tensor is

$$\mathbf{K}^{\text{small}} + \mathbf{A}^{\text{gm}} = \kappa_{\text{ntr}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 S_x & 2 S_y & \mathbf{S}^2 \end{bmatrix} \quad \text{if } \kappa_{\text{ntr}} = \kappa_{\text{gm}}, \quad (18.90)$$

so that the subgrid scale flux simplifies to

$$\rho^{-1} \mathbf{J} = -\kappa_{\text{ntr}} \nabla_h C - \hat{\mathbf{z}} \kappa_{\text{ntr}} (2 \mathbf{S} \cdot \nabla_h C + \mathbf{S}^2 \partial_z C) \quad \text{if } \kappa_{\text{ntr}} = \kappa_{\text{gm}}. \quad (18.91)$$

Notably, the  $2 \times 2$  horizontal mixing tensor is diagonal. Hence, the horizontal tracer flux is the same as that which arises from downgradient horizontal tracer diffusion. The simplicity of the horizontal flux component was alluring to modelers. It was furthermore argued by [Dukowicz and Smith \(1997\)](#) to be a fundamental property of mesoscale turbulence. However, as emphasized through the works of [Treguier et al. \(1997\)](#), [Ferrari et al. \(2008\)](#), [Danabasoglu et al. \(2008\)](#), and [Ferrari et al. \(2010\)](#), the boundary conditions for neutral diffusion and GM skewness are distinct, thus breaking their symmetry. Furthermore, studies such as [Smith and Marshall \(2009\)](#) and [Abernathy et al. \(2013\)](#) clearly point to the distinct vertical structure for the two diffusivities. Such distinctions are expected since the skew diffusivity and neutral diffusivity parameterize physically distinct processes: one parameterizes the quasi-Stokes transport, associated with velocity and layer thickness correlations, whereas the other parameterizes downgradient diffusion along neutral directions, associated with velocity and tracer correlations.

#### 18.4.8 Generalized vertical coordinates

Thus far we have considered neutral diffusion as realized in geopotential coordinates or using neutral tangent plane coordinates. Here, we detail the steps needed to realize neutral diffusion using [generalized vertical coordinates](#) from Part IV of this volume. This formulation is relevant for the now common use of generalized vertical coordinates for ocean modeling as reviewed by [Griffies et al. \(2020\)](#).

Start by recalling the expression (12.106) for a general diffusion operator written in terms of the generalized vertical coordinate,  $\sigma = \sigma(x, y, z, t)$

$$\mathcal{R} = -\frac{1}{h^\sigma} [\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \delta_\sigma(z_\sigma \nabla \sigma \cdot \mathbf{J})], \quad (18.92)$$

where  $\delta_\sigma \equiv d\sigma / \partial_\sigma$  is the dimensionless derivative operator, and the thickness of a  $\sigma$ -layer is

$$h^\sigma = dz = z_\sigma d\sigma = \frac{\partial z}{\partial \sigma} d\sigma. \quad (18.93)$$

Now assume the flux,  $\mathbf{J}$ , is given by equation (18.74) for small slope neutral diffusion. Trans-

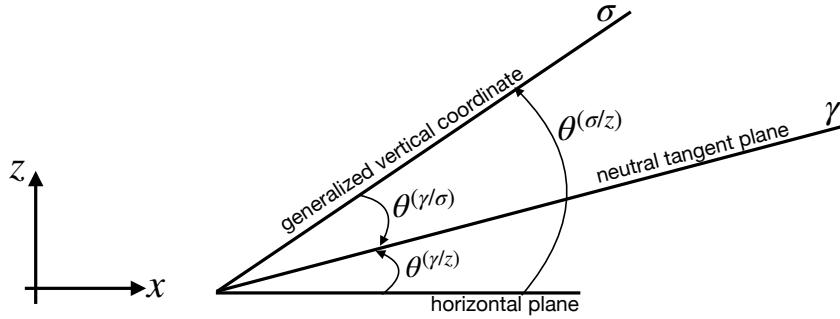


FIGURE 18.6: Slopes of a neutral tangent plane (denoted by  $\gamma$ ) relative to both the horizontal plane,  $\tan \theta^{(\gamma/z)}$ , and relative to a generalized vertical coordinate isoline,  $\tan \theta^{(\gamma/\sigma)}$ , where  $\sigma$  is a generalized vertical coordinate. We assume positive angles as measure counter-clockwise relative to the horizontal and relative to the  $\sigma$ -isoline, respectively. Hence, for this example,  $\theta^{(\gamma/z)} > 0$  yet  $\theta^{(\gamma/\sigma)} < 0$ . When extending to the two horizontal directions, the slopes generally satisfy  $\mathbf{S}^{(\sigma/z)} = \mathbf{S}^{(\gamma/z)} - \mathbf{S}^{(\gamma/\sigma)}$ , where  $|\mathbf{S}^{(\gamma/z)}| = |\tan \theta^{(\gamma/z)}|$  and  $|\mathbf{S}^{(\gamma/\sigma)}| = |\tan \theta^{(\gamma/\sigma)}|$ . Note that this relation between slope vectors also holds for arbitrary orientations of the  $\sigma$  isolines and neutral tangent planes.

forming to generalized vertical coordinates leads to the horizontal flux component

$$\mathbf{J}^{h\text{small}} = -\rho \kappa_{\text{ntr}} \nabla_h C \quad (18.94\text{a})$$

$$= -\rho \kappa_{\text{ntr}} [\nabla_h + (\nabla_h z) \partial_z] C \quad (18.94\text{b})$$

$$= -\rho \kappa_{\text{ntr}} [\nabla_{hv} + (-\nabla_{hv} z + \nabla_h z) \partial_z] C \quad (18.94\text{c})$$

$$= -\rho \kappa_{\text{ntr}} [\nabla_{hv} + (-\mathbf{S}^{(\sigma/z)} + \mathbf{S}^{(\gamma/z)}) \partial_z] C \quad (18.94\text{d})$$

$$= -\rho \kappa_{\text{ntr}} (\nabla_{hv} + \mathbf{S}^{(\gamma/\sigma)} \partial_z) C, \quad (18.94\text{e})$$

where the neutral slopes as shown in Figure 18.6 satisfy the identity

$$\mathbf{S}^{(\sigma/z)} = \mathbf{S}^{(\gamma/z)} - \mathbf{S}^{(\gamma/\sigma)}. \quad (18.95)$$

Furthermore, we made use of the identity (12.85) relating the partial derivative operators

$$\nabla_h = \nabla_h + (\nabla_h z) \partial_z \quad \text{and} \quad \nabla_{hv} = \nabla_{hv} - (\nabla_{hv} z) \partial_z. \quad (18.96)$$

The horizontal flux (18.94e) has the same form as when written using geopotential coordinates, only now with the derivative operator  $\nabla_{hv}$  and the slope  $\mathbf{S}^{(\gamma/\sigma)}$ . Correspondingly, the vertical flux component

$$\mathbf{J}^{z\text{small}} = \mathbf{J}^{h\text{small}} \cdot \mathbf{S}^{(\gamma/z)} \quad (18.97)$$

takes the form

$$z_\sigma \nabla \sigma \cdot \mathbf{J}^{\text{small}} = -\mathbf{S}^{(\sigma/z)} \cdot \mathbf{J}^{h\text{small}} + \mathbf{J}^{z\text{small}} = \mathbf{J}^{h\text{small}} \cdot \mathbf{S}^{(\gamma/\sigma)}, \quad (18.98)$$

which in turn yields the diffusion operator (18.92)

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[ \nabla_{hv} \cdot (h^\sigma \mathbf{J}^{h\text{small}}) + \delta_\sigma (\mathbf{J}^{h\text{small}} \cdot \mathbf{S}^{(\gamma/\sigma)}) \right]. \quad (18.99)$$

In the special case when  $\sigma$  is parallel to the neutral direction so that  $\mathbf{S}^{(\gamma/\sigma)} = 0$ , the diffusion operator (18.99) reduces to the neutral tangent plane version given by equation (18.82).

## 18.5 Anisotropic neutral diffusion

The neutral diffusion discussed in Section 18.4 is based on isotropic diffusion in the neutral tangent plane. That assumption has been questioned by *Smith and Gent* (2004), *Le Sommer et al.* (2011), and *Fox-Kemper et al.* (2013). We here develop some of the formalism appropriate for studying anisotropic neutral diffusion. It is notable, however, that anisotropic neutral diffusion has been rarely tested in realistic models.

### 18.5.1 Orthonormal triad of basis vectors

We make use of the following orthonormal unit vectors<sup>13</sup> as depicted in Figure 18.7

$$\hat{\mathbf{e}}_{\bar{1}} = \frac{\hat{\mathbf{d}} \times \hat{\gamma}}{|\hat{\mathbf{d}} \times \hat{\gamma}|} \quad (18.100a)$$

$$\hat{\mathbf{e}}_{\bar{2}} = \frac{\hat{\gamma} \times (\hat{\mathbf{d}} \times \hat{\gamma})}{|\hat{\mathbf{d}} \times \hat{\gamma}|} = \frac{\hat{\mathbf{d}} - (\hat{\gamma} \cdot \hat{\mathbf{d}}) \hat{\gamma}}{|\hat{\mathbf{d}} \times \hat{\gamma}|} \quad (18.100b)$$

$$\hat{\mathbf{e}}_{\bar{3}} = \hat{\gamma} \quad (18.100c)$$

where

$$\hat{\mathbf{d}} = \hat{\mathbf{x}} \hat{d}_x + \hat{\mathbf{y}} \hat{d}_y + \hat{\mathbf{z}} \hat{d}_z \quad (18.101)$$

is an arbitrary unit vector that is not parallel to  $\hat{\gamma}$ . The three unit vectors  $(\hat{\mathbf{e}}_{\bar{1}}, \hat{\mathbf{e}}_{\bar{2}}, \hat{\mathbf{e}}_{\bar{3}})$  form an orthonormal triad at each point in the fluid so that

$$\hat{\mathbf{e}}_{\bar{1}} = \hat{\mathbf{e}}_{\bar{2}} \times \hat{\mathbf{e}}_{\bar{3}} \quad \text{and} \quad \hat{\mathbf{e}}_{\bar{2}} = \hat{\mathbf{e}}_{\bar{3}} \times \hat{\mathbf{e}}_{\bar{1}} \quad \text{and} \quad \hat{\mathbf{e}}_{\bar{3}} = \hat{\mathbf{e}}_{\bar{1}} \times \hat{\mathbf{e}}_{\bar{2}}. \quad (18.102)$$

These vectors are oriented by the arbitrary direction,  $\hat{\mathbf{d}}$ , and the dianeutral direction,  $\hat{\gamma}$ . We verify that  $\hat{\mathbf{e}}_{\bar{2}}$  has unit magnitude by noting that

$$|\hat{\mathbf{d}} \times \hat{\gamma}|^2 = |\hat{\gamma} \times (\hat{\mathbf{d}} \times \hat{\gamma})|^2 = 1 - (\hat{\mathbf{d}} \cdot \hat{\gamma})^2. \quad (18.103)$$

It is also useful to verify that  $\hat{\mathbf{e}}_{\bar{3}} = \hat{\mathbf{e}}_{\bar{1}} \times \hat{\mathbf{e}}_{\bar{2}}$  through the following vector identity (see VOLUME 1)

$$(\hat{\mathbf{d}} \times \hat{\gamma}) \times [\hat{\gamma} \times (\hat{\mathbf{d}} \times \hat{\gamma})] = \hat{\gamma} |\hat{\mathbf{d}} \times \hat{\gamma}|^2. \quad (18.104)$$

The unit vectors  $\hat{\mathbf{e}}_{\bar{1}}$  and  $\hat{\mathbf{e}}_{\bar{2}}$  are both within the neutral tangent plane since they are both orthogonal to  $\hat{\gamma}$ .

The unit vector  $\hat{\mathbf{e}}_{\bar{1}}$  is orthogonal to  $\hat{\mathbf{d}}$  whereas  $\hat{\mathbf{e}}_{\bar{2}}$  is parallel to  $\hat{\mathbf{d}}$  if  $\hat{\mathbf{d}} \cdot \hat{\gamma} = 0$ . For example, *Smith and Gent* (2004) proposed setting  $\hat{\mathbf{d}}$  to be a horizontal vector set according to the local horizontal flow direction, in which case

$$\hat{\mathbf{d}} = \frac{u \hat{\mathbf{x}} + v \hat{\mathbf{y}}}{(u^2 + v^2)^{1/2}}. \quad (18.105)$$

With  $\hat{\gamma}$  nearly vertical for much of the ocean interior, then  $\hat{\mathbf{e}}_{\bar{2}}$  becomes nearly aligned with  $\hat{\mathbf{d}}$ . For these reasons we refer to  $\hat{\mathbf{e}}_{\bar{1}}$  as the across- $\hat{\mathbf{d}}$  direction and  $\hat{\mathbf{e}}_{\bar{2}}$  as the along- $\hat{\mathbf{d}}$  direction.

<sup>13</sup>The basis vectors (18.100a)-(18.100c) are more suitable for present purposes than the analogous basis vectors defined by equations (14.4)-(14.6) in *Griffies* (2004). In particular, the basis (18.100a)-(18.100c) has a sensible limit when the neutral slopes are horizontal, in which  $\hat{\gamma} = -\hat{\mathbf{z}}$ .

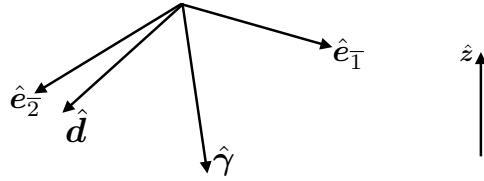


FIGURE 18.7: Depicting the orthonormal triad of basis vectors given by equations (18.100a)-(18.100c). Our convention is such that  $\hat{\gamma}$  typically points downward toward increasing density. The unit vector  $\hat{d}$  is arbitrary so long as it is not parallel to the dianeutral unit vector,  $\hat{\gamma}$ . It is horizontal when making use of the [Smith and Gent \(2004\)](#) proposal whereby  $\hat{d} = \mathbf{u}/|\mathbf{u}|$ , with  $\mathbf{u} = \hat{x} u + \hat{y} v$  the horizontal velocity vector. Since  $\hat{e}_{\bar{1}}$  is orthogonal to  $\hat{d}$ , we refer to  $\hat{e}_{\bar{1}}$  as the cross- $\hat{d}$  basis vector. Likewise, since  $\hat{e}_{\bar{2}}$  is nearly parallel to  $\hat{d}$ , especially when  $\hat{d}$  is close to horizontal and  $\hat{\gamma}$  is close to vertical (e.g., Section 18.5.4), then  $\hat{e}_{\bar{2}}$  is referred to as the along- $\hat{d}$  basis vector.

### 18.5.2 Anisotropic neutral diffusion tensor

We consider anisotropy according to the unit vectors  $\hat{e}_{\bar{1}}$  and  $\hat{e}_{\bar{2}}$ . Hence, the diffusion tensor as represented using the locally orthogonal triad  $(\hat{e}_{\bar{1}}, \hat{e}_{\bar{2}}, \hat{e}_{\bar{3}})$  is given by

$$\overline{\mathbf{K}^{\text{aniso}}} = \begin{bmatrix} \kappa_{\text{cross}} & 0 & 0 \\ 0 & \kappa_{\text{along}} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (18.106)$$

where  $\kappa_{\text{cross}} > 0$  and  $\kappa_{\text{along}} > 0$  are the generally distinct neutral diffusivities. This tensor takes on the component form

$$(\mathbf{K}^{\text{aniso}})^{\bar{m}\bar{n}} = \kappa_{\text{cross}} \hat{e}_{\bar{1}}^{\bar{m}} \hat{e}_{\bar{1}}^{\bar{n}} + \kappa_{\text{along}} \hat{e}_{\bar{2}}^{\bar{m}} \hat{e}_{\bar{2}}^{\bar{n}} \quad (18.107a)$$

$$= \kappa_{\text{cross}} (\delta^{\bar{m}\bar{n}} - \hat{e}_{\bar{2}}^{\bar{m}} \hat{e}_{\bar{2}}^{\bar{n}} - \hat{e}_{\bar{3}}^{\bar{m}} \hat{e}_{\bar{3}}^{\bar{n}}) + \kappa_{\text{along}} (\delta^{\bar{m}\bar{n}} - \hat{e}_{\bar{1}}^{\bar{m}} \hat{e}_{\bar{1}}^{\bar{n}} - \hat{e}_{\bar{3}}^{\bar{m}} \hat{e}_{\bar{3}}^{\bar{n}}), \quad (18.107b)$$

where the second expression made use of the following decomposition of the unit tensor in terms of the orthonormal basis vectors

$$\delta^{\bar{m}\bar{n}} = \hat{e}_{\bar{1}}^{\bar{m}} \hat{e}_{\bar{1}}^{\bar{n}} + \hat{e}_{\bar{2}}^{\bar{m}} \hat{e}_{\bar{2}}^{\bar{n}} + \hat{e}_{\bar{3}}^{\bar{m}} \hat{e}_{\bar{3}}^{\bar{n}}. \quad (18.108)$$

Note that  $(\mathbf{K}^{\text{aniso}})^{\bar{m}\bar{n}}$  is invariant under  $\hat{d} \rightarrow -\hat{d}$ . Likewise, it is invariant under a change in the sign of  $\hat{\gamma}$ . Furthermore, we recover the isotropic Redi diffusion tensor (18.70) by setting  $\kappa_{\text{cross}} = \kappa_{\text{along}} = \kappa_{\text{ntr}}$ , in which case

$$(K^{\text{aniso}})^{\bar{m}\bar{n}} = 2 \kappa_{\text{ntr}} (\delta^{\bar{m}\bar{n}} - \hat{e}_{\bar{3}}^{\bar{m}} \hat{e}_{\bar{3}}^{\bar{n}}) - \kappa_{\text{ntr}} (\hat{e}_{\bar{1}}^{\bar{m}} \hat{e}_{\bar{1}}^{\bar{n}} + \hat{e}_{\bar{2}}^{\bar{m}} \hat{e}_{\bar{2}}^{\bar{n}}) \quad (18.109a)$$

$$= \kappa_{\text{ntr}} (\delta^{\bar{m}\bar{n}} - \hat{e}_{\bar{3}}^{\bar{m}} \hat{e}_{\bar{3}}^{\bar{n}}) \quad (18.109b)$$

$$= (K^{\text{redi}})^{\bar{m}\bar{n}} \quad (18.109c)$$

To render a geopotential-Cartesian representation of the anisotropic diffusion tensor, we can make use of the transformation methods for Cartesian tensors developed in VOLUME 1. We do so by transforming from the locally orthogonal neutral plane coordinate system, defined by the orthonormal triad (18.100a)-(18.100c), to the geopotential-Cartesian coordinate system, defined by the Cartesian triad

$$\hat{e}_1 = \hat{x} \quad \text{and} \quad \hat{e}_2 = \hat{y} \quad \text{and} \quad \hat{e}_3 = \hat{z}. \quad (18.110)$$

Since we are working with Cartesian tensors, this transformation is a local rotation matrix,  $\mathcal{R}$ ,

so that<sup>14</sup>

$$(K^{\text{aniso}})^{mn} = \mathcal{R}^m_{\bar{m}} \mathcal{R}^n_{\bar{n}} (K^{\text{aniso}})^{\bar{m}\bar{n}} \implies \mathbf{K}^{\text{aniso}} = \mathcal{R} \mathbf{K}^{\text{aniso}} \mathcal{R}^T, \quad (18.111)$$

where the second equality made use of matrix notation with  $\mathcal{R}^T$  the transpose, and where the elements to the rotation matrix are given by the direction cosines

$$\mathcal{R} = \begin{bmatrix} \hat{\mathbf{e}}^1 \cdot \hat{\mathbf{e}}_{\bar{1}} & \hat{\mathbf{e}}^1 \cdot \hat{\mathbf{e}}_{\bar{2}} & \hat{\mathbf{e}}^1 \cdot \hat{\mathbf{e}}_{\bar{3}} \\ \hat{\mathbf{e}}^2 \cdot \hat{\mathbf{e}}_{\bar{1}} & \hat{\mathbf{e}}^2 \cdot \hat{\mathbf{e}}_{\bar{2}} & \hat{\mathbf{e}}^2 \cdot \hat{\mathbf{e}}_{\bar{3}} \\ \hat{\mathbf{e}}^3 \cdot \hat{\mathbf{e}}_{\bar{1}} & \hat{\mathbf{e}}^3 \cdot \hat{\mathbf{e}}_{\bar{2}} & \hat{\mathbf{e}}^3 \cdot \hat{\mathbf{e}}_{\bar{3}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}} \cdot \hat{\mathbf{e}}_{\bar{1}} & \hat{\mathbf{x}} \cdot \hat{\mathbf{e}}_{\bar{2}} & \hat{\mathbf{x}} \cdot \hat{\mathbf{e}}_{\bar{3}} \\ \hat{\mathbf{y}} \cdot \hat{\mathbf{e}}_{\bar{1}} & \hat{\mathbf{y}} \cdot \hat{\mathbf{e}}_{\bar{2}} & \hat{\mathbf{y}} \cdot \hat{\mathbf{e}}_{\bar{3}} \\ \hat{\mathbf{z}} \cdot \hat{\mathbf{e}}_{\bar{1}} & \hat{\mathbf{z}} \cdot \hat{\mathbf{e}}_{\bar{2}} & \hat{\mathbf{z}} \cdot \hat{\mathbf{e}}_{\bar{3}} \end{bmatrix}. \quad (18.112)$$

The machinery outlined here for the transformation is straightforward but tedious (i.e., two matrix multiplies). A more streamlined approach, also used for determining the Cartesian components to the Redi tensor (18.72), is to simply express  $(\hat{\mathbf{e}}_{\bar{1}}, \hat{\mathbf{e}}_{\bar{2}}, \hat{\mathbf{e}}_{\bar{3}})$  using geopotential-Cartesian coordinates and then plug directly into equation (18.107b).

### 18.5.3 Properties of the anisotropic neutral diffusive fluxes

We here verify some standard properties for the anisotropic neutral diffusive flux for tracers

$$\mathbf{J}^{\text{aniso}} = -\rho \mathbf{K}^{\text{aniso}} \cdot \nabla C. \quad (18.113)$$

#### Downgradient orientation within the neutral tangent plane

By construction, the flux is downgradient along the two orthogonal directions,  $\hat{\mathbf{e}}_{\bar{1}}$  and  $\hat{\mathbf{e}}_{\bar{2}}$ ,

$$\mathbf{J}^{\text{aniso}} = -\rho \kappa_{\text{cross}} \hat{\mathbf{e}}_{\bar{1}} (\hat{\mathbf{e}}_{\bar{1}} \cdot \nabla C) - \rho \kappa_{\text{along}} \hat{\mathbf{e}}_{\bar{2}} (\hat{\mathbf{e}}_{\bar{2}} \cdot \nabla C). \quad (18.114)$$

Furthermore, the flux is within the neutral tangent plane

$$\mathbf{J}^{\text{aniso}} \cdot \hat{\gamma} = 0 \quad (18.115)$$

due to orthogonality between the basis vectors

$$\hat{\mathbf{e}}_{\bar{1}} \cdot \hat{\gamma} = \hat{\mathbf{e}}_{\bar{2}} \cdot \hat{\gamma} = 0. \quad (18.116)$$

#### Neutrality condition

The neutrality condition (18.86) follows since

$$(-\alpha \nabla \Theta + \beta \nabla S) \cdot \hat{\mathbf{e}}_{\bar{1}} = (-\alpha \nabla \Theta + \beta \nabla S) \cdot \hat{\mathbf{e}}_{\bar{2}} = 0, \quad (18.117)$$

so that

$$\alpha \mathbf{J}^{\text{aniso}}(\Theta) = \beta \mathbf{J}^{\text{aniso}}(S). \quad (18.118)$$

### 18.5.4 Small slope anisotropic neutral diffusion

We now consider the special case in which the orientation direction,  $\hat{\mathbf{d}}$ , is strictly horizontal and normalized so that

$$\hat{\mathbf{d}} \cdot \hat{\mathbf{d}} = \hat{d}_x^2 + \hat{d}_y^2 = 1. \quad (18.119)$$

---

<sup>14</sup>Since we are dealing with Cartesian tensors there is no distinction between raised or lowered tensor indices in equation (18.111). Even so, we choose to follow the convention of general tensors in VOLUME 1 to help organize elements of the tensor and to facilitate use of the Einstein summation convention.

Additionally, we assume the neutral slope is small so that the neutral directions are nearly horizontal. In this case the basis vectors (18.100a)-(18.100c) take on the following form valid to  $\mathcal{O}(|\mathbf{S}|)$

$$\hat{\mathbf{e}}_1^{\text{small}} = \hat{\mathbf{d}} \times (\mathbf{S} - \hat{\mathbf{z}}) \quad (18.120\text{a})$$

$$\hat{\mathbf{e}}_2^{\text{small}} = \hat{\mathbf{d}} + \hat{\mathbf{z}} (\hat{\mathbf{d}} \cdot \mathbf{S}) \quad (18.120\text{b})$$

$$\hat{\mathbf{e}}_3^{\text{small}} = \mathbf{S} - \hat{\mathbf{z}}. \quad (18.120\text{c})$$

Note that  $\hat{\mathbf{e}}_3^{\text{small}}$  is orthogonal to  $\hat{\mathbf{e}}_1^{\text{small}}$  and  $\hat{\mathbf{e}}_2^{\text{small}}$ , however  $\hat{\mathbf{e}}_1^{\text{small}} \cdot \hat{\mathbf{e}}_2^{\text{small}}$  is  $\mathcal{O}(\mathbf{S} \cdot \mathbf{S})$ . Likewise, each of these vectors is normalized only to  $\mathcal{O}(\mathbf{S} \cdot \mathbf{S})$ .

Making use of the small slope basis vectors in the anisotropic diffusion tensor (18.107a), and expressing them in geopotential-Cartesian coordinates leads to the small slope anisotropic neutral diffusion tensor<sup>15</sup>

$$\mathbf{K}^{\text{smallaniso}} = \kappa_{\text{cross}} \begin{bmatrix} 1 & 0 & S_x \\ 0 & 1 & S_y \\ S_x & S_y & \mathbf{S} \cdot \mathbf{S} \end{bmatrix} + \Delta\kappa_{\text{ntr}} \begin{bmatrix} \hat{d}_x^2 & \hat{d}_x \hat{d}_y & (\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{d}_x \\ \hat{d}_x \hat{d}_y & \hat{d}_y^2 & (\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{d}_y \\ (\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{d}_x & (\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{d}_y & (\hat{\mathbf{d}} \cdot \mathbf{S})^2 \end{bmatrix}, \quad (18.121)$$

where

$$\Delta\kappa_{\text{ntr}} = \kappa_{\text{along}} - \kappa_{\text{cross}}. \quad (18.122)$$

As for the unapproximated anisotropic neutral diffusion tensor (18.107a), its small slope version,  $\bar{\mathbf{K}}^{\text{smallaniso}}$ , is invariant if we swap the direction  $\hat{\mathbf{d}} \rightarrow -\hat{\mathbf{d}}$ . Furthermore, in the form (18.121) we trivially see that  $\mathbf{K}^{\text{smallaniso}} = \mathbf{K}^{\text{small}}$  (equation (18.73)) in the isotropic limit where  $\kappa_{\text{cross}} = \kappa_{\text{along}} = \kappa_{\text{ntr}}$ .

The anisotropic small slope neutral diffusive flux is given by

$$\mathbf{J}^{h\text{smallaniso}} = -\rho \mathbf{K}^{\text{smallaniso}} \cdot \nabla C, \quad (18.123)$$

with horizontal and vertical components

$$\mathbf{J}^{h\text{smallaniso}} = -\rho \kappa_{\text{cross}} \nabla_h C - \rho \Delta\kappa_{\text{ntr}} \hat{\mathbf{d}} (\hat{\mathbf{d}} \cdot \nabla_h) C \quad (18.124\text{a})$$

$$J^{z\text{smallaniso}} = \mathbf{S} \cdot \mathbf{J}^{h\text{smallaniso}}, \quad (18.124\text{b})$$

where  $\nabla_h = \nabla_h + \mathbf{S} \partial_z$  is the horizontal operator as per equation (18.75). By making use of the expression (18.19),  $\hat{\gamma} = (\mathbf{S} - \hat{\mathbf{z}})(1 + \mathbf{S}^2)^{-1/2}$ , we readily find that

$$\mathbf{J}^{h\text{smallaniso}} \cdot \hat{\gamma} = 0. \quad (18.125)$$

Similarly, we can verify that the neutrality condition (Section 18.4.5) is maintained

$$\mathbf{K}^{\text{smallaniso}} \cdot \hat{\gamma} = 0 \implies \alpha \mathbf{J}^{h\text{smallaniso}}(\Theta) = \beta \mathbf{J}^{h\text{smallaniso}}(S). \quad (18.126)$$

Finally, as per the discussion in Section 18.4.4, we can evaluate the small slope anisotropic neutral diffusion operator by following the non-orthogonal neutral tangent approach rather than the three-dimensional Cartesian approach. It is the non-orthogonal neutral tangent approach that is appropriate for vertical Lagrangian ocean models such as detailed in Griffies et al. (2020) and Shao et al. (2020).

<sup>15</sup>Equation (18.121) agrees with equations (10) and (14) from Smith and Gent (2004).

## 18.6 Anisotropic Gent-McWilliams stirring

In addition to proposing the use of a small slope anisotropic neutral diffusion tensor (equation (18.121)), [Smith and Gent \(2004\)](#) proposed a complementary anisotropic version of the Gent-McWilliams stirring. We here detail the parameterization, again assuming the orientation direction,  $\hat{\mathbf{d}}$ , is horizontal

$$\hat{\mathbf{d}} = \hat{\mathbf{x}} \hat{d}_x + \hat{\mathbf{y}} \hat{d}_y, \quad (18.127)$$

just as assumed when discussing the anisotropic small slope neutral diffusion operator in Section 18.5.4.

### 18.6.1 Streamfunction and anti-symmetric tensor

The parameterized eddy-induced streamfunction is generalized from that in equation (18.28) to read

$$\Psi^* = \hat{\mathbf{z}} \times \kappa_{\text{gmcross}} \mathbf{S} + \hat{\mathbf{z}} \times (\kappa_{\text{gmcross}} - \kappa_{\text{gmalong}}) (\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{\mathbf{d}}, \quad (18.128)$$

and the corresponding anti-symmetric stirring tensor is

$$A^{\text{gmaniso}} = \kappa_{\text{gmcross}} \begin{bmatrix} 0 & 0 & -S_x \\ 0 & 0 & -S_y \\ S_x & S_y & 0 \end{bmatrix} + \Delta \kappa_{\text{gm}} (\hat{\mathbf{d}} \cdot \mathbf{S}) \begin{bmatrix} 0 & 0 & -\hat{d}_x \\ 0 & 0 & -\hat{d}_y \\ \hat{d}_x & \hat{d}_y & 0 \end{bmatrix}, \quad (18.129)$$

where

$$\Delta \kappa_{\text{gm}} = \kappa_{\text{gmalong}} - \kappa_{\text{gmcross}}. \quad (18.130)$$

As for the small slope anisotropic neutral diffusion tensor (18.121), we write the skew tensor  $A^{\text{gmaniso}}$  in equation (18.129) in a form that manifestly reduces to the isotropic Gent-McWilliams stirring tensor  $A^{\text{gm}}$  when  $\kappa_{\text{gmalong}} = \kappa_{\text{gmcross}} = \kappa_{\text{gm}}$ .

### 18.6.2 Anisotropic Gent-McWilliams skew tracer flux

The anisotropic Gent-McWilliams skew tracer flux is

$$\mathbf{J}_{\text{gm-aniso}} = -\rho A^{\text{gmaniso}} \cdot \nabla C \quad (18.131a)$$

$$= \rho \kappa_{\text{gmcross}} [\mathbf{S} \partial_z C - \hat{\mathbf{z}} (\mathbf{S} \cdot \nabla_h C)] + \rho \Delta \kappa_{\text{gm}} (\hat{\mathbf{d}} \cdot \mathbf{S}) [\hat{\mathbf{d}} \partial_z C - \hat{\mathbf{z}} (\hat{\mathbf{d}} \cdot \nabla_h C)]. \quad (18.131b)$$

When acting on locally referenced potential density,  $C = \gamma$ , the flux reduces to

$$\mathbf{J}_{\text{gm-aniso}} = \rho \kappa_{\text{gmcross}} [-\nabla_h \gamma + \hat{\mathbf{z}} \mathbf{S}^2 \partial_z \gamma] + \rho \Delta \kappa_{\text{gm}} (\hat{\mathbf{d}} \cdot \mathbf{S}) [\hat{\mathbf{d}} + \hat{\mathbf{z}} (\hat{\mathbf{d}} \cdot \mathbf{S})] \partial_z \gamma. \quad (18.132)$$

As discussed in Section 18.3.3, a negative vertical component to the potential density skew flux ensures that the available potential energy is dissipated,

$$\hat{\mathbf{z}} \cdot \mathbf{J}_{\text{gm-aniso}} = \rho [\kappa_{\text{gmcross}} \mathbf{S}^2 + \Delta \kappa_{\text{gm}} (\hat{\mathbf{d}} \cdot \mathbf{S})^2] \partial_z \gamma < 0 \implies \text{APE dissipated}. \quad (18.133)$$

Stably stratified water means that  $\partial_z \gamma < 0$ , in which case  $\hat{\mathbf{z}} \cdot \mathbf{J}_{\text{gm-aniso}} < 0$  since

$$\kappa_{\text{gmcross}} \mathbf{S}^2 + \Delta \kappa_{\text{gm}} (\hat{\mathbf{d}} \cdot \mathbf{S})^2 = \kappa_{\text{gmcross}} [\mathbf{S}^2 - (\hat{\mathbf{d}} \cdot \mathbf{S})^2] + \kappa_{\text{gmalong}} (\hat{\mathbf{d}} \cdot \mathbf{S})^2 > 0. \quad (18.134)$$

### 18.6.3 Anisotropic GM skewson plus small slope neutral diffusion

As noted in Section 18.4.7, there are strong reasons to keep the Gent-McWilliams skew flux parameterization distinct from the neutral diffusion parameterization. The central practical reason for the distinction concerns their different treatment of boundary conditions and generally distinct diffusivities. Even so, we here briefly comment on the special case where we ignore these distinctions and set the skew flux diffusivities equal to the neutral diffusivities

$$\kappa_{\text{gmcross}} = \kappa_{\text{cross}} \quad \text{and} \quad \kappa_{\text{gmalong}} = \kappa_{\text{along}}. \quad (18.135)$$

This is the approach assumed by *Smith and Gent (2004)*. With the small slope approximation to neutral diffusion, we find the combined anisotropic mixing tensor becomes

$$\begin{aligned} \mathbf{K}^{\text{smallaniso}} + \mathbf{A}^{\text{gmaniso}} = \\ \kappa_{\text{cross}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2S_x & 2S_y & \mathbf{S} \cdot \mathbf{S} \end{bmatrix} + \Delta\kappa \begin{bmatrix} \hat{d}_x^2 & \hat{d}_x \hat{d}_y & 0 \\ \hat{d}_x \hat{d}_y & \hat{d}_y^2 & 0 \\ 2(\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{d}_x & 2(\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{d}_y & (\hat{\mathbf{d}} \cdot \mathbf{S})^2 \end{bmatrix}. \end{aligned} \quad (18.136)$$

The vanishing right hand column terms simplifies the horizontal tracer fluxes computed from this tensor. Even so, this formulation is inconsistent with the theories that support a distinct treatment of the skew flux and neutral flux.

### 18.6.4 A parameterization based on a boundary value problem

We now follow the approach from Section 18.3.7 to develop a boundary value problem version of the anisotropic Gent-McWilliams stirring. For this purpose consider the vertical boundary value problem

$$(c^2 \partial_{zz} - N^2) \boldsymbol{\Upsilon} = -N^2 \boldsymbol{\Upsilon}^{\text{gmaniso}} \quad \text{and} \quad \boldsymbol{\Upsilon}(\eta_b) = \boldsymbol{\Upsilon}(\eta) = 0, \quad (18.137)$$

where (see equation (18.128))

$$\boldsymbol{\Upsilon}^{\text{gmaniso}} = \kappa_{\text{gmcross}} \mathbf{S} + \hat{\mathbf{z}} \times \Delta\kappa_{\text{gm}} (\hat{\mathbf{d}} \cdot \mathbf{S}) \hat{\mathbf{d}}. \quad (18.138)$$

As in Section 18.3.7, we deduce the impacts on potential energy (assuming a linear equation of state) via the vertical component of the potential density skew flux,

$$\frac{1}{g} \frac{dP}{dt} = \frac{1}{\rho_0} \int J^z dV = - \int \nabla_h \varrho \cdot \boldsymbol{\Upsilon} dV. \quad (18.139)$$

The governing differential equation (18.137) leads to

$$\boldsymbol{\Upsilon} \cdot (c^2 \partial_{zz} - N^2) \boldsymbol{\Upsilon} = -N^2 \boldsymbol{\Upsilon} \cdot \boldsymbol{\Upsilon}^{\text{gmaniso}} \quad (18.140)$$

which rearranges to

$$N^2 \boldsymbol{\Upsilon} \cdot \boldsymbol{\Upsilon}^{\text{gmaniso}} = -c^2 \partial_z (\boldsymbol{\Upsilon} \cdot \partial_z \boldsymbol{\Upsilon}) + c^2 \partial_z \boldsymbol{\Upsilon} \cdot \partial_z \boldsymbol{\Upsilon} + N^2 \boldsymbol{\Upsilon} \cdot \boldsymbol{\Upsilon}. \quad (18.141)$$

Integrating over a vertical column and making use of the homogeneous Dirichlet boundary conditions in equation (18.137) leads to

$$\frac{g}{\rho_0} \int \boldsymbol{\Upsilon} \cdot \nabla_h \varrho dz = - \int N^2 \Delta \kappa_{gm} (\hat{\mathbf{d}} \cdot \mathbf{S}) (\hat{\mathbf{d}} \cdot \boldsymbol{\Upsilon}) dz + \int [c^2 \partial_z \boldsymbol{\Upsilon} \cdot \partial_z \boldsymbol{\Upsilon} + N^2 \boldsymbol{\Upsilon} \cdot \boldsymbol{\Upsilon}] dz, \quad (18.142)$$

which can be rearranged into the equivalent form

$$\begin{aligned} & \frac{g}{\rho_0} \int \boldsymbol{\Upsilon} \cdot \nabla_h \varrho dz \\ &= \underbrace{\int [c^2 \partial_z \boldsymbol{\Upsilon} \cdot \partial_z \boldsymbol{\Upsilon} + N^2 (\boldsymbol{\Upsilon} \cdot \boldsymbol{\Upsilon} - (\hat{\mathbf{d}} \cdot \boldsymbol{\Upsilon})^2)] dz}_{\text{positive semi-definite}} + \underbrace{\int N^2 (\hat{\mathbf{d}} \cdot \boldsymbol{\Upsilon}) \hat{\mathbf{d}} \cdot (\boldsymbol{\Upsilon} - \mathbf{S} \Delta \kappa_{gm}) dz}_{\text{sign indefinite}}. \end{aligned} \quad (18.143)$$

The first term on the right hand side is positive semi-indefinite whereas the second term is sign indefinite. If the second term is positive, or smaller in magnitude than the first term, then the parameterization provides a column integrated sink of potential energy. Otherwise, potential energy for the column can increase. There are no existing numerical implementations of this scheme to determine its suitability for realistic ocean climate simulations.



## 18.7 Exercises

### EXERCISE 18.1: SQUARED BUOYANCY AND THE OCEAN MODEL EQUATIONS

In this exercise we develop some properties of the squared buoyancy for the hydrostatic Boussinesq equations from VOLUME 2, written here in the form

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\nabla \varphi + b \hat{\mathbf{z}} + \mathbf{F} \quad \text{velocity equation} \quad (18.144a)$$

$$\nabla \cdot \mathbf{v} = 0 \quad \text{continuity equation} \quad (18.144b)$$

$$\frac{Db}{Dt} = \dot{b} \quad \text{buoyancy equation} \quad (18.144c)$$

$$b = -\frac{g \rho'}{\rho_0} = -\frac{g(\rho - \rho_0)}{\rho_0} \quad \text{buoyancy defined} \quad (18.144d)$$

$$\varphi = \frac{p'}{\rho_0} = \frac{p - p_o(z)}{\rho_0} \quad \text{dynamic pressure defined} \quad (18.144e)$$

$$\rho = \rho_0 (1 - \alpha \Theta + \beta S) \quad \text{linear equation of state} \quad (18.144f)$$

$$\frac{dp_o}{dz} = -\rho_0 g \quad \text{background hydrostatic pressure.} \quad (18.144g)$$

Furthermore, we make use of a buoyancy flux in the form

$$\mathbf{F}^b = -\kappa \partial_z b \hat{\mathbf{z}} + \mathbf{v}^* b. \quad (18.145)$$

The first term is a downgradient vertical diffusive flux with the vertical eddy diffusivity,  $\kappa > 0$ , a function of the flow state so that

$$\kappa = \kappa(\mathbf{x}, t). \quad (18.146)$$

The second term is an advective flux, where the advective velocity,  $\mathbf{v}^* = (\mathbf{u}^*, w^*)$ , is assumed to be non-divergent

$$\nabla \cdot \mathbf{v}^* = \nabla_h \cdot \mathbf{u}^* + \partial_z w^* = 0. \quad (18.147)$$

The velocity,  $\mathbf{v}^*$ , is the eddy-induced velocity, such as that discussed in Section 18.3 from [Gent et al. \(1995\)](#).

- (a) Write the flux-form budget describing the evolution of  $b^2$ , the squared buoyancy. Write the budget equation using the residual mean velocity,  $\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*$ . Hint: start from the buoyancy equation written in the form

$$Db/Dt = -\nabla \cdot \mathbf{F}^b. \quad (18.148)$$

- (b) Discuss the impacts from vertical diffusion on the  $b^2$  budget.

#### EXERCISE 18.2: PARAMETERIZED EDDY VELOCITY AND THE OCEAN MODEL EQUATIONS

In this exercise we develop some implications of assuming a specific form for the parameterized eddy velocity for the hydrostatic Boussinesq equations from Exercise 18.1. Namely, we consider the specific form for the parameterized eddy-induced velocity proposed by [Gent et al. \(1995\)](#)

$$\mathbf{u}^* = -\partial_z(B \mathbf{S}) \quad (18.149a)$$

$$w^* = \nabla_h \cdot (B \mathbf{S}) \quad (18.149b)$$

$$\mathbf{S} = -\frac{\nabla_h b}{N^2} \quad (18.149c)$$

$$\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0 \quad \text{at all ocean boundaries.} \quad (18.149d)$$

In this expression,  $B > 0$  is an eddy diffusivity. To ensure  $\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0$  at all domain boundaries requires that  $B = 0$  along these boundaries. The horizontal vector,  $\mathbf{S} = (S_x, S_y, 0)$ , measures the slope of the buoyancy surfaces relative to the horizontal. We assume the ocean is stably stratified in the vertical, so that  $\partial b / \partial z = N^2 > 0$ .

- (a) Determine the vector streamfunction,  $\Psi^*$ , such that

$$\mathbf{v}^* = \nabla \times \Psi^*. \quad (18.150)$$

Choose the gauge with  $\hat{\mathbf{z}} \cdot \Psi^* = 0$ .

- (b) Show that

$$\int_{-H}^{\eta} \mathbf{u}^* dz = 0, \quad (18.151)$$

so that the parameterized horizontal flow has a zero depth integral.

- (c) At any chosen meridional position,  $y$ , the meridional buoyancy transport from advection (resolved plus parameterized) is computed by

$$\mathcal{B}^{(y)}(y, t) = \int_{x_1}^{x_2} dx \int_{-H}^{\eta} b(v + v^*) dz. \quad (18.152)$$

The zonal and vertical integrals are over the full zonal and vertical extent of the ocean domain. Show that the effects from  $v^*$  are to reduce the meridional gradients of buoyancy. That is, if buoyancy decreases poleward, then  $v^*$  will flux buoyancy poleward to reduce the gradient.

- (d) How does the introduction of  $\mathbf{v}^*$  to the buoyancy equation (18.144c) affect the global

integrated gravitational potential energy? Discuss.

- (e) How does the introduction of  $v^*$  to the buoyancy equation (18.144c) affect the global integrated available potential energy? Discuss.

**EXERCISE 18.3: MERIDIONAL OVERTURNING WITH *Gent et al. (1995)***

It is often of interest to compute the mass transport across a portion of the ocean. In particular, meridional-depth or meridional-potential density streamfunctions allow one to visualize and quantify the zonally integrated transport occurring in a closed basin or over the full globe. The quasi-Stokes transport provides a transport in addition to that from the resolved scale Eulerian mean transport, and the parameterization of *Gent et al. (1995)* leads to a straightforward computation of the quasi-Stokes contribution. For this purpose, write the net meridional mass transport of fluid across a basin and passing beneath a particular depth in the form (the minus sign is conventional)

$$\mathcal{T}(y, z, t) = - \int dx \int_{-H}^z \rho(v + v^*) dz. \quad (18.153)$$

Make use of the expression for the *Gent et al. (1995)* parameterized  $v^*$  and simplify this expression. Discuss its scaling for the Southern Ocean.

**EXERCISE 18.4: BIHARMONIC HORIZONTAL MIXING**

A **biharmonic operator** is commonly employed for dissipation in numerical models due to its enhanced scale selectivity. However, biharmonic operators have some unphysical properties, such as those encountered in this exercise focused on the biharmonic mixing oriented in the horizontal direction, with the flux given by

$$\mathbf{F}^{\text{bih}} = \sqrt{B} \nabla_h L, \quad (18.154)$$

where

$$L = \nabla_h \cdot (\sqrt{B} \nabla_h C) \quad (18.155)$$

is the horizontal Laplacian operator acting on the tracer and  $B > 0$  is a biharmonic mixing coefficient with units  $\text{L}^4 \text{T}^{-1}$ . The convergence of the biharmonic flux  $\mathbf{F}^{\text{bih}}$  yields the horizontal biharmonic mixing operator

$$R^{\text{bih}} = -\nabla \cdot \mathbf{F}^{\text{bih}}. \quad (18.156)$$

Examine the biharmonic operator's effects on tracer variance by computing the time tendency of the squared tracer concentration over the full domain, ignoring all boundary contributions. Discuss the result.

**EXERCISE 18.5: *Cox (1987)* SMALL SLOPE DIFFUSION**

As noted in Section 18.4.3, the neutral diffusion tensor proposed by *Cox (1987)* did not satisfy the neutrality condition of Section 18.4.5, with this incorrect tensor given by

$$\mathbf{K}^{\text{cox}} = \kappa_{\text{ntr}} \begin{bmatrix} 1 & S_x S_y & S_x \\ S_x S_y & 1 & S_y \\ S_x & S_y & S^2 \end{bmatrix}. \quad (18.157)$$

We here illustrate the basic problem with this tensor by considering the special case of neutral directions aligned with Conservative Temperature surfaces, such as occurs in an ocean with constant salinity. Show that the diffusive flux

$$\mathbf{J}^{\text{cox}}(\Theta) = -\mathbf{K}^{\text{cox}} \cdot \nabla \Theta \neq 0, \quad (18.158)$$

so that there is a non-zero flux of Conservative Temperature. Instead, a proper neutral diffusion flux leads to an identically zero flux, such as for fluxes defined by  $\mathbf{K}^{\text{redi}}$  (Section 18.4.2),  $\mathbf{K}^{\text{small}}$  (Section 18.4.3), and the neutral tangent plane method from Section 18.4.4.

**EXERCISE 18.6: *Roberts and Marshall (1998)* BIHARMONIC OPERATOR**

Assuming a Boussinesq ocean, consider the *Roberts and Marshall (1998)* divergence-free velocity

$$\mathbf{v}^* = \nabla \times \boldsymbol{\Upsilon} = \partial_z [\nabla_h^2(B \mathbf{S})] - \hat{z} \nabla_h \cdot \nabla_h^2(B \mathbf{S}), \quad (18.159)$$

with the vector streamfunction given by

$$\boldsymbol{\Upsilon} = -\hat{z} \times \nabla_h^2(B \mathbf{S}) \quad (18.160)$$

and with  $B > 0$  a biharmonic mixing coefficient. The corresponding skew tracer flux is given by

$$\mathbf{F} = -\partial_z C \nabla_h^2(B \mathbf{S}) + \hat{z} \nabla_h C \cdot \nabla_h^2(B \mathbf{S}). \quad (18.161)$$

Dropping the horizontal Laplacian  $\nabla_h^2$  and setting  $B \rightarrow -\kappa$  recovers the Laplacian scheme of *Gent et al. (1995)* as discussed in Section 18.3. *Roberts and Marshall (1998)* proposed this operator as a means to dissipate grid scale variance in the density field, yet to do so without compromising the volume of fluid within isopycnal layers. Derive an expression for the effects of this biharmonic operator on the potential energy, assuming a linear equation of state and ignoring boundary terms. Discuss the result.

**EXERCISE 18.7: DIFFUSION AND ADVECTION ACTING ON FOURIER MODES**

In this exercise we consider a single Fourier mode for the density and a single mode for the tracer concentration

$$\rho(\mathbf{x}, t) = B(t) \cos(\mathbf{p} \cdot \mathbf{x}), \quad \text{and} \quad T(\mathbf{x}, t) = A(t) \cos(\mathbf{q} \cdot \mathbf{x}) \quad (18.162)$$

where  $\mathbf{p}$  and  $\mathbf{q}$  are prescribed wave vectors. Assume a Boussinesq ocean and a density that is not a function of pressure.

- (a) Write the isotropic diffusion equation for this tracer field

$$\partial_t T = \kappa \nabla^2 T. \quad (18.163)$$

- (b) Write the neutral diffusion equation for this tracer and density

$$\partial_t T = \partial_m (K^{mn} \partial_n T), \quad (18.164)$$

where

$$K^{mn} = \kappa_{\text{nt}} (\delta^{mn} - \hat{\rho}^m \hat{\rho}^n) + \kappa_{\text{dia}} \hat{\rho}^m \hat{\rho}^n. \quad (18.165)$$

- (c) Write the eddy-induced velocity according to the Gent-McWilliams parameterization as given by equation (18.29), here written for a Boussinesq ocean as

$$\mathbf{u}^* = -\partial_z (\kappa_{\text{gm}} \mathbf{S}). \quad (18.166)$$

For simplicity, assume a constant  $\kappa_{\text{gm}}$ .

**EXERCISE 18.8: NEUTRAL TRACER FILTERS**

In this exercise we consider a tracer flux for a Boussinesq ocean in the form

$$\mathbf{F} = \mathbf{B} + \hat{z} \mathbf{S} \cdot \mathbf{B}, \quad (18.167)$$

where  $\mathbf{B}$  is an arbitrary horizontal vector built from the tracer field. Assume a linear equation of state for simplicity.

- (a) Show that  $\mathbf{F} \cdot \nabla \rho = 0$ , so that  $\mathbf{F}$  is oriented along surfaces of constant density.
- (b) Consider  $\mathbf{B} = -\kappa \nabla_h T$ , where  $\kappa > 0$  is a constant of dimensions  $L^2 T^{-1}$ . Discuss why this operator might be of use to smooth the tracer field without transferring tracer across density surfaces. Consider also the case with  $\mathbf{B} = A \nabla_h [\nabla_h \cdot (A \nabla_h T)]$ , where  $A > 0$  is a constant and with  $A^2$  having dimensions of  $L^4 T^{-1}$ .





# Chapter 19

## OCEAN DENSITY AND SEA LEVEL

As discussed in VOLUME 2, **Conservative Temperature**,  $\Theta$ , is the preferred means to measure the transport of enthalpy in the ocean, and **salinity**,  $S$ , measures the concentration of dissolved salt matter. These two scalar fields are referred to as **active tracers** since they both impact density and in turn affect pressure and ocean currents. In this chapter we study how the evolution of  $\Theta$  and  $S$  affects seawater density as well as buoyancy. As part of this study, we examine how to compute air-sea buoyancy fluxes and to compute budgets for global mean sea level.

$\Theta$  and  $S$  are **conservative tracers** so that the net changes in potential enthalpy and salt over the global ocean domain arise from net imbalances in their boundary fluxes. Likewise, ocean mass is a conserved field, with global mass changes arising from imbalances in boundary mass fluxes such as those occurring from increases in land ice melt. However, ocean volume, and hence ocean density and buoyancy, are not conserved fields. Consequently, ocean volume can change even if there is no net volume transferred to the ocean. These points have direct impact on how global mean sea level is affected by ocean processes such as mixing and heating, with the rudiments presented in this chapter.

### CHAPTER GUIDE

Our study of thermo-hydrodynamics in VOLUME 2 motivates the use of **Conservative Temperature**,  $\Theta$ , as a measure of ocean enthalpy transfer, rather than *in situ* temperature or **potential temperature**. We also make use of the ideas of parameterized turbulent mixing discussed in Chapters 15 and 18 when formulating the budget equations for  $\Theta$  and  $S$ . We use Cartesian tensors to reduce the mathematical overhead. Also note that we use subscripts on specific volume,  $\nu$ , and density,  $\rho$ , for partial derivatives with respect to  $\Theta$  and  $S$ . This is the only chapter in this book that makes use of subscript notation for partial derivatives, and we only use it for thermodynamic derivatives.

<b>19.1 Loose threads . . . . .</b>	<b>660</b>
<b>19.2 Material evolution of <i>in situ</i> density . . . . .</b>	<b>660</b>
19.2.1 Material changes to pressure . . . . .	661
19.2.2 Material changes to $\Theta$ and $S$ . . . . .	661
19.2.3 General expression for density changes . . . . .	662
19.2.4 Unpacking the subgrid contributions . . . . .	662
19.2.5 Synthesis of the density equation . . . . .	664
<b>19.3 Cabbeling and thermobaricity . . . . .</b>	<b>665</b>
19.3.1 Basic manipulations . . . . .	666
19.3.2 Cabbeling . . . . .	667

19.3.3	Thermobaricity . . . . .	667
19.3.4	Comments . . . . .	668
19.3.5	Further study . . . . .	668
<b>19.4</b>	<b>Salt and freshwater continuity equations . . . . .</b>	<b>668</b>
19.4.1	Salt and freshwater . . . . .	668
19.4.2	Continuity equations . . . . .	669
<b>19.5</b>	<b>Surface boundary conditions for <math>S</math> and <math>\Theta</math> . . . . .</b>	<b>670</b>
19.5.1	Salt and freshwater . . . . .	670
19.5.2	The non-advective salt flux boundary condition . . . . .	671
19.5.3	Conservative Temperature boundary condition . . . . .	673
19.5.4	Comments and further reading . . . . .	673
<b>19.6</b>	<b>Surface boundary fluxes of buoyancy . . . . .</b>	<b>673</b>
19.6.1	Surface boundary fluxes of enthalpy and salt . . . . .	674
19.6.2	Evolution from surface boundary fluxes . . . . .	674
19.6.3	Buoyancy tendency from surface fluxes . . . . .	675
19.6.4	Comments . . . . .	677
<b>19.7</b>	<b>Global mean sea level . . . . .</b>	<b>677</b>
19.7.1	Definitions and assumptions . . . . .	677
19.7.2	Budget for global mean sea level . . . . .	678
19.7.3	Changes due to mass input . . . . .	679
19.7.4	Steric changes due to changes in density . . . . .	679
19.7.5	Enthalpy flux imbalances implied by thermosteric sea level . . . . .	680
19.7.6	Barystatic versus steric effects from near future land ice melt . . . . .	681
<b>19.8</b>	<b>Exercises . . . . .</b>	<b>684</b>

## 19.1 Loose threads

- More on patterns of sea level and the problems with global mean of local steric not equaling global mean.

## 19.2 Material evolution of *in situ* density

Changes to the *in situ* density of seawater affects pressure forces in the ocean as well as the volume occupied by the ocean fluid (i.e., sea level). In this section we provide a mathematical framework for studying processes that affect density changes.

As discussed in VOLUME 2, we write the seawater equation of state for density as a function of **salinity**,  $S$ , and **Conservative Temperature**,  $\Theta$ , where Conservative Temperature is the potential enthalpy divided by a constant heat capacity ([McDougall \(2003\)](#); [IOC et al. \(2010\)](#)). We thus make use of the empirical equation of state for the seawater density using the functional form

$$\rho = \rho(S, \Theta, p), \quad (19.1)$$

where  $S$  is the salinity rather than the salt concentration ( $S = 1000 S$ ).

In this section we formulate the material time evolution of density as weighted by the specific volume<sup>1</sup>

$$\nu = \rho^{-1}, \quad (19.2)$$

<sup>1</sup>In other chapters we write the specific volume as  $\nu_s = 1/\rho$  to distinguish it from  $\nu$  that is used for kinematic viscosity. However, in this chapter we write  $\nu = 1/\rho$  to enable a shorthand for partial derivatives as defined by equation (19.8). We have no use kinematic viscosity in this chapter.

so that, using the chaing rule, we have

$$\frac{D \ln \rho}{Dt} = \frac{\partial \ln \rho}{\partial \Theta} \frac{D\Theta}{Dt} + \frac{\partial \ln \rho}{\partial S} \frac{DS}{Dt} + \frac{\partial \ln \rho}{\partial p} \frac{Dp}{Dt} \quad (19.3a)$$

$$= -\alpha \frac{D\Theta}{Dt} + \beta \frac{DS}{Dt} + \frac{\dot{p}}{\rho c_s^2}. \quad (19.3b)$$

In this equation we introduced the [thermal expansion coefficient](#), the [haline contraction coefficient](#), the squared speed of sound, and the vertical pseudo-velocity in pressure

$$\alpha = - \left[ \frac{\partial \ln \rho}{\partial \Theta} \right]_{p,S} \quad \beta = \left[ \frac{\partial \ln \rho}{\partial S} \right]_{p,\Theta} \quad c_s^2 = \left[ \frac{\partial p}{\partial \rho} \right]_{S,\Theta} \quad \dot{p} = \frac{Dp}{Dt}. \quad (19.4)$$

For the remainder of this section we unpack the processes contributing to the density material time evolution appearing in equation (19.3b).

### 19.2.1 Material changes to pressure

To garner some exposure to the physics of  $\dot{p}$  as it appears in equation (19.3b), consider the case of a hydrostatic fluid, where the volume per time per horizontal area of fluid crossing a surface of constant hydrostatic pressure is given by (see Section 13.4.6)

$$w^{(p)} = \frac{\partial z}{\partial p} \frac{Dp}{Dt} = -(\rho g)^{-1} \dot{p}. \quad (19.5)$$

The transport measured by  $w^{(p)}$  is the pressure-coordinate analog of the vertical velocity component,  $w = Dz/Dt$ , that arises in a geopotential coordinate representation of the vertical. That is, fluid moving into regions of increasing hydrostatic pressure ( $\dot{p} > 0$ ) represents downward movement of fluid, with  $w^{(p)} < 0$  in this case. Conversely, motion into decreasing hydrostatic pressure represents upward motion, with  $w^{(p)} > 0$ . This vertical movement generally occurs in the presence of waves, currents, and mixing; i.e., both reversible and irreversible processes give rise to vertical motion.

### 19.2.2 Material changes to $\Theta$ and $S$

We next assume the material evolution of,  $S$ , and Conservative Temperature,  $\Theta$ , are affected by the convergence of a subgrid scale flux

$$\rho \frac{D\Theta}{Dt} = -\nabla \cdot \mathbf{J}^{(\Theta)} \quad (19.6a)$$

$$\rho \frac{DS}{Dt} = -\nabla \cdot \mathbf{J}^{(S)}. \quad (19.6b)$$

The Conservative Temperature equation (19.6a) was derived in VOLUME 2, whereas the salinity equation (19.6b) follows from our study of material tracers in VOLUME 1.<sup>2</sup>

<sup>2</sup>For stratified fluids with multiple constituents, one encounters the effects of matter diffusion in the presence of a temperature gradient ([Soret effect](#)), and temperature diffusion in the presence of gradients of pressure and/or matter ([Dufour effect](#)). We introduced these cross-diffusion processes in VOLUME 2 when studying energy and entropy flows in fluids. However, geophysical fluids are nearly always dominated by turbulence, which renders a turbulent diffusion that swamps the effects from cross-diffusion. We thus ignore cross-diffusion in this chapter. We also ignore remineralization processes that can contribute to a source term in the salinity equation (19.6b). Cross-diffusion and remineralization sources are discussed in [IOC et al. \(2010\)](#).

### 19.2.3 General expression for density changes

The expressions (19.6a) and (19.6b) for material changes in  $\Theta$  and  $S$  then lead to

$$-\alpha \frac{D\Theta}{Dt} + \beta \frac{DS}{Dt} = \nu_\Theta \nabla \cdot \mathbf{J}^{(\Theta)} + \nu_S \nabla \cdot \mathbf{J}^{(S)} \quad (19.7a)$$

$$= \nabla \cdot [\nu_\Theta \mathbf{J}^{(\Theta)} + \nu_S \mathbf{J}^{(S)}] - [\mathbf{J}^{(\Theta)} \cdot \nabla \nu_\Theta + \mathbf{J}^{(S)} \cdot \nabla \nu_S] \quad (19.7b)$$

where again  $\nu = \rho^{-1}$  is the specific volume and its thermodynamic partial derivatives are written using the shorthand

$$\nu_\Theta = \frac{\partial \nu}{\partial \Theta} = \frac{\alpha}{\rho} \quad \text{and} \quad \nu_S = \frac{\partial \nu}{\partial S} = -\frac{\beta}{\rho}. \quad (19.8)$$

Bringing the above results together leads to the density equation

$$\frac{D \ln \rho}{Dt} - \frac{\dot{p}}{\rho c_s^2} = \nabla \cdot [\nu_\Theta \mathbf{J}^{(\Theta)} + \nu_S \mathbf{J}^{(S)}] - [\mathbf{J}^{(\Theta)} \cdot \nabla \nu_\Theta + \mathbf{J}^{(S)} \cdot \nabla \nu_S], \quad (19.9)$$

which has the equivalent form

$$\frac{D\rho}{Dt} - \frac{\dot{p}}{c_s^2} = \nabla \cdot [\alpha \mathbf{J}^{(\Theta)} - \beta \mathbf{J}^{(S)}] - [\mathbf{J}^{(\Theta)} \cdot \nabla \alpha - \mathbf{J}^{(S)} \cdot \nabla \beta]. \quad (19.10)$$

We brought the source term from motion across pressure surfaces (Section 19.2.1) onto the left hand side, since this term appears in the absence of subgrid processes. The first term on the right hand side represents the divergence of a buoyancy flux due to subgrid scale fluxes of Conservative Temperature and salinity. In turn, density increases in regions where the buoyancy flux diverges (e.g., reducing  $\Theta$  and increasing  $S$ ). These fluxes arise from a variety of mixing processes, some of which are surveyed in Section 18.1.1. The second term on the right hand side of equations (19.9) and (19.10) relates to local properties of the density surface. We study this source term in Section 19.3 assuming the flux arises from *neutral diffusion*. Further effects arise from an unresolved *eddy-induced velocity*, whose stirring contributes to the *residual mean velocity* as discussed in Section 18.1.4.

### 19.2.4 Unpacking the subgrid contributions

Recall from Section 18.1 that the *subgrid scale* tracer fluxes are generally written in terms of a second order eddy transport tensor,  $\mathbf{E}$ , so that

$$\mathbf{J}^{(\Theta)} = -\rho \mathbf{E} \cdot \nabla \Theta \quad \text{and} \quad \mathbf{J}^{(S)} = -\rho \mathbf{E} \cdot \nabla S. \quad (19.11)$$

Furthermore,  $\mathbf{E}$  is typically decomposed as in equation (18.8) into a symmetric downgradient diffusion tensor,  $\mathbf{K}$ , and an anti-symmetric skew diffusion (or stirring) tensor,  $\mathbf{A}$ ,

$$\mathbf{E} = \mathbf{K} + \mathbf{A}. \quad (19.12)$$

We decompose the contributions to density according to these subgrid tensors using the following manipulations

$$\frac{D\rho}{Dt} - \frac{\dot{p}}{c_s^2} = \alpha \nabla \cdot \mathbf{J}^{(\Theta)} - \beta \nabla \cdot \mathbf{J}^{(S)} \quad (19.13a)$$

$$= -\alpha \nabla \cdot (\rho \mathbf{E} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbf{E} \cdot \nabla S). \quad (19.13b)$$

Expanding the  $\Theta$  term leads to

$$\alpha \nabla \cdot \mathbf{J}^{(\Theta)} = -\alpha \nabla \cdot (\rho \mathbf{E} \cdot \nabla \Theta) \quad (19.14a)$$

$$= -\alpha \nabla \cdot (\rho \mathbf{A} \cdot \nabla \Theta) - \alpha \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta) \quad (19.14b)$$

$$= -\alpha \nabla \cdot (\rho \mathbf{A}) \cdot \nabla \Theta - \alpha \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta), \quad (19.14c)$$

$$= \alpha \mathbf{v}^* \cdot \nabla \Theta - \alpha \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta). \quad (19.14d)$$

To reach this result we made use of the identities

$$-\nabla \cdot (\rho \mathbf{A} \cdot \nabla \Theta) = -\partial_m (\rho A^{mn} \partial_n \Theta) \quad \text{expose tensor indices} \quad (19.15a)$$

$$= -\partial_m (\rho A^{mn}) \partial_n \Theta - \rho A^{mn} \partial_m \partial_n \Theta \quad \text{product rule} \quad (19.15b)$$

$$= -\partial_m (\rho A^{mn}) \partial_n \Theta \quad A^{mn} \partial_m \partial_n \Theta = 0 \quad (19.15c)$$

$$= \rho \mathbf{v}^* \cdot \nabla \Theta \quad \partial_m (\rho A^{mn}) = -\rho v^{*n}. \quad (19.15d)$$

In the final equality we introduced the density weighted eddy-induced velocity,  $\rho \mathbf{v}^*$ , defined by equation (18.11). The same manipulations for the salinity term lead to

$$\frac{D\rho}{Dt} - \frac{\dot{p}}{c_s^2} + \rho \mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S) = -\alpha \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbf{K} \cdot \nabla S). \quad (19.16)$$

We can write this expression in terms of the residual mean material time operator

$$\frac{D^\dagger}{Dt} = \partial_t + \mathbf{v}^\dagger \cdot \nabla = \frac{D}{Dt} + \mathbf{v}^* \cdot \nabla \quad (19.17)$$

through adding and subtracting  $c_s^{-2} \mathbf{v}^* \cdot \nabla p$

$$\rho \mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S) = \mathbf{v}^* \cdot (-\rho \alpha \nabla \Theta + \rho \beta \nabla S + c_s^{-2} \nabla p) - c_s^{-2} \mathbf{v}^* \cdot \nabla p = \mathbf{v}^* \cdot (\nabla \rho - c_s^{-2} \nabla p), \quad (19.18)$$

which then leads to

$$\frac{D\rho}{Dt} - \frac{\dot{p}}{c_s^2} + \rho \mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S) = \frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt}, \quad (19.19)$$

so that

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = -\alpha \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbf{K} \cdot \nabla S). \quad (19.20)$$

Transport from the symmetric tensor,  $\mathbf{K}$ , corresponds to diffusion so long as the tensor is positive definite. The diffusion operator in the residual mean evolution equation (19.20) can be written

$$\begin{aligned} & -\alpha \nabla \cdot (\rho \mathbf{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbf{K} \cdot \nabla S) \\ &= \nabla \cdot [\rho \mathbf{K} \cdot (-\alpha \nabla \Theta + \beta \nabla S)] + \rho \nabla \alpha \cdot \mathbf{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbf{K} \cdot \nabla S, \end{aligned} \quad (19.21)$$

so that the *in situ* density evolves according to

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = - \underbrace{\nabla \cdot [\rho \mathbf{K} \cdot (\alpha \nabla \Theta - \beta \nabla S)]}_{\text{conservative processes}} + \underbrace{\rho \nabla \alpha \cdot \mathbf{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbf{K} \cdot \nabla S}_{\text{sources from nonlinear EOS processes}}. \quad (19.22)$$

We now discuss the physical processes associated with the right hand side terms.

- LINEAR EQUATION OF STATE: A linear equation of state has  $\nabla \alpha = \nabla \beta = 0$ , with density independent of pressure. As a result, the evolution equation (19.22) takes the form

$$\frac{D^\dagger \rho}{Dt} = - \nabla \cdot [\rho \mathbf{K} \cdot (\alpha \nabla \Theta - \beta \nabla S)]. \quad (19.23)$$

Consequently, under the residual mean advective transport with a linear equation of state, density remains materially constant in the absence of any diffusion.

- NONLINEAR EQUATION OF STATE: A nonlinear equation of state is characterized by spatially dependent thermal expansion and haline contraction coefficients. Mixing of  $\Theta$  and  $S$  in the presence of a nonlinear equation of state generally gives rise to material evolution of *in situ* density through *cabbeling* and *thermobaricity* (McDougall, 1987b). We offer a mathematical summary of these processes in Section 19.3.

- NEUTRAL DIFFUSION:

Neutral diffusion from Section 18.4 maintains a density-compensated diffusive flux of  $\Theta$  and  $S$  so that

$$\mathbf{K}^{\text{neutral}} \cdot (\alpha \nabla \Theta - \beta \nabla S) = 0. \quad (19.24)$$

Hence, neutral diffusion leaves *in situ* density changed only via the nonlinear equation of state processes.

- ISOTROPIC SMALL SCALE DIFFUSION:

As discussed in Section 18.1.3, it is common to parameterize fine scale mixing processes using an isotropic diffusivity so that the diffusion tensor is given by

$$\mathbf{K}^{\text{iso}} = \kappa \mathbb{I}, \quad (19.25)$$

where  $\mathbb{I}$  is the unit tensor and  $\kappa > 0$  is the isotropic eddy diffusivity.

### 19.2.5 Synthesis of the density equation

In summary, the material time evolution equation for *in situ* density in the presence of subgrid scale processes takes the form

$$\begin{aligned} \frac{D\rho}{Dt} = & \underbrace{\frac{1}{c_s^2} \frac{Dp}{Dt}}_{\text{compressibility}} - \underbrace{\mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S)}_{\text{eddy-induced advection}} - \underbrace{\nabla \cdot [\rho \kappa (-\alpha \nabla \Theta + \beta \nabla S)]}_{\text{small scale diffusive mixing}} \\ & + \underbrace{\rho \nabla \alpha \cdot \mathbf{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbf{K} \cdot \nabla S}_{\text{nonlinear EOS processes from eddy mixing}}. \end{aligned} \quad (19.26)$$

We thus have the following physical processes contributing to the evolution of *in situ* density.

- ADIABATIC COMPRESSION: Material changes to pressure in the presence of a finite sound speed lead to changes in the seawater density.
- SMALL SCALE MIXING: Small scale mixing is parameterized by an isotropic diffusivity,  $\kappa > 0$ . This kinematic diffusivity is the same for all tracers, with the exception of [double diffusive processes](#) whereby material tracers (e.g., salinity, nutrients) have a diffusivity distinct from temperature ([Schmitt, 1994](#)). Given the dominance of vertical stratification over the horizontal, it is common to approximate the isotropic diffusion operator with a vertical diffusion operator (but see Section 4 of [McDougall et al. \(2014\)](#) for caveats).
- EDDY-INDUCED VELOCITY: As studied in Chapter 18, subgrid scale correlations between eddy velocity and eddy isopycnal layer thickness lead to an [eddy-induced velocity](#) which, when combined with the resolved flow velocity, leads to the residual mean material time derivative,  $D^\dagger/Dt$  acting on the tracer fields. Equivalently, eddy-induced advective transport can be realized via [skew diffusion](#).
- EDDY-INDUCED DIFFUSION: As studied in Chapter 18, subgrid scale correlations between eddy velocity and eddy tracer concentration lead to a direct cascade of  $\Theta$  and  $S$  variance to the small scales. Mixing arising from this cascade is parameterized by [neutral diffusion](#), whereby the diffusive fluxes of  $\Theta$  and  $S$  are density compensated according to the constraint (19.24).
- NONLINEAR EOS PROCESSES: Mixing of  $\Theta$  and  $S$  in the presence of a nonlinear equation of state means that *in situ* density evolves due to [cabbeling](#) and [thermobaricity](#) (Section 19.3). The dominant contributions to these processes arise from eddy induced mixing (i.e., neutral diffusion) ([McDougall, 1987b](#)), though small scale mixing also contributes.

## 19.3 Cabbeling and thermobaricity

We now return to the density equation (19.10)

$$\frac{D \ln \rho}{Dt} - \frac{\dot{p}}{\rho c_s^2} = \nabla \cdot [\nu_\Theta \mathbf{J}^{(\Theta)} + \nu_S \mathbf{J}^{(S)}] - (\mathbf{J}^{(\Theta)} \cdot \nabla \nu_\Theta + \mathbf{J}^{(S)} \cdot \nabla \nu_S), \quad (19.27)$$

and focus on  $\Theta$  and  $S$  fluxes arising just from the neutral diffusion process described in Section 18.4. The neutrality condition (18.86) is a fundamental property of neutral diffusion, and it takes the following form in terms of specific volume

$$\nu_\Theta \mathbf{J}^{(\Theta)} + \nu_S \mathbf{J}^{(S)} = 0. \quad (19.28)$$

Consequently, neutral diffusion affects density evolution only through the nonlinear equation of state source term

$$\left[ \frac{D \ln \rho}{Dt} \right]_{\text{ntrldiff}} = -\mathbf{J}^{(\Theta)} \cdot \nabla \nu_\Theta - \mathbf{J}^{(S)} \cdot \nabla \nu_S. \quad (19.29)$$

In the remainder of this section we manipulate the source term with the goal to identify the variety of physical processes associated with neutral diffusion in the presence of a nonlinear equation of state.

### 19.3.1 Basic manipulations

As a first step, eliminate the salt flux by using the neutrality condition (19.28)

$$\mathbf{J}^{(\Theta)} \cdot \nabla \nu_{\Theta} + \mathbf{J}^{(S)} \cdot \nabla \nu_S = \mathbf{J}^{(\Theta)} \cdot [\nu_S \nabla \nu_{\Theta} - \nu_{\Theta} \nabla \nu_S] / \nu_S. \quad (19.30)$$

Next, expand the gradients of the specific volume to write

$$\nabla \nu_{\Theta} = \nu_{\Theta\Theta} \nabla \Theta + \nu_{\Theta S} \nabla S + \nu_{\Theta p} \nabla p \quad \text{and} \quad \nabla \nu_S = \nu_{SS} \nabla S + \nu_{\Theta S} \nabla \Theta + \nu_{Sp} \nabla p, \quad (19.31)$$

so that

$$\begin{aligned} \nu_S \nabla \nu_{\Theta} - \nu_{\Theta} \nabla \nu_S &= \nabla \Theta (\nu_S \nu_{\Theta\Theta} - \nu_{\Theta} \nu_{\Theta S}) \\ &\quad + \nabla S (\nu_S \nu_{\Theta S} - \nu_{\Theta} \nu_{SS}) + \nabla p (\nu_S \nu_{\Theta p} - \nu_{\Theta} \nu_{Sp}). \end{aligned} \quad (19.32)$$

We again make use of the neutrality condition (19.28), as well as the symmetry condition (18.87d) to write

$$\mathbf{J}^{(\Theta)} \cdot \nabla S (\nu_S \nu_{\Theta S} - \nu_{\Theta} \nu_{SS}) = -\mathbf{J}^{(\Theta)} \cdot \nabla \Theta \left[ \nu_{\Theta} \nu_{\Theta S} - \nu_{SS} \frac{(\nu_{\Theta})^2}{\nu_S} \right]. \quad (19.33)$$

Bringing these results together leads to

$$\begin{aligned} \mathbf{J}^{(\Theta)} \cdot \nabla \nu_{\Theta} + \mathbf{J}^{(S)} \cdot \nabla \nu_S &= \mathbf{J}^{(\Theta)} \cdot \nabla p (\nu_{\Theta p} - \nu_{pS} \nu_{\Theta} / \nu_S) \\ &\quad + \mathbf{J}^{(\Theta)} \cdot \nabla \Theta [\nu_{\Theta\Theta} - 2 \nu_{\Theta S} \nu_{\Theta} / \nu_S + \nu_{SS} (\nu_{\Theta} / \nu_S)^2], \end{aligned} \quad (19.34)$$

which can be written in terms of density partial derivatives as

$$\begin{aligned} \mathbf{J}^{(\Theta)} \cdot \nabla \nu_{\Theta} + \mathbf{J}^{(S)} \cdot \nabla \nu_S &= -\rho^{-2} \mathbf{J}^{(\Theta)} \cdot \nabla p (\rho_{\Theta p} - \rho_{pS} \rho_{\Theta} / \rho_S) \\ &\quad - \rho^{-2} \mathbf{J}^{(\Theta)} \cdot \nabla \Theta [\rho_{\Theta\Theta} - 2 \rho_{\Theta S} \rho_{\Theta} / \rho_S + \rho_{SS} (\rho_{\Theta} / \rho_S)^2]. \end{aligned} \quad (19.35)$$

Following Exercise 19.1, we write the bracket terms appearing in equation (19.35) in forms consistent with those written by [McDougall \(1987b\)](#). We do so by introducing the **thermobaricity** parameter (dimensions of inverse temperature times inverse pressure)

$$\mathcal{T} = \beta \partial_p \left[ \frac{\alpha}{\beta} \right], \quad (19.36)$$

and the **cabbeling** parameter (dimensions of squared inverse temperature)

$$\mathcal{C} = \frac{\partial \alpha}{\partial \Theta} + 2 \frac{\alpha}{\beta} \frac{\partial \alpha}{\partial S} - \left( \frac{\alpha}{\beta} \right)^2 \frac{\partial \beta}{\partial S}, \quad (19.37)$$

which then leads to the tidy result

$$\mathbf{J}^{(\Theta)} \cdot \nabla \nu_{\Theta} + \mathbf{J}^{(S)} \cdot \nabla \nu_S = \rho^{-1} \mathbf{J}^{(\Theta)} \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta). \quad (19.38)$$

Finally, we are thus led to the material time evolution of *in situ* density due to neutral diffusion

$$\left[ \frac{D\rho}{Dt} \right]_{\text{ntrldiff}} = -\mathbf{J}^{(\Theta)} \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta). \quad (19.39)$$

We next discuss the terms in this equation.

### 19.3.2 Cabbeling

The intrinsic geometry of density surfaces in Conservative Temperature and salinity space render the following inequality for the cabbeling parameter

$$\mathcal{C} = \frac{\partial \alpha}{\partial \Theta} + 2 \frac{\alpha}{\beta} \frac{\partial \alpha}{\partial S} - \left[ \frac{\alpha}{\beta} \right]^2 \frac{\partial \beta}{\partial S} \geq 0. \quad (19.40)$$

With this inequality in mind, now consider the mixing of two adjacent elements of seawater. Let the fluid elements separately have distinct Conservative Temperature and/or salinity, but equal **locally referenced potential density**. For a linear equation of state, whereby density is a linear function of  $\Theta$  and  $S$ , the resulting mixed fluid element has the same density as the unmixed separate elements. However, for a nonlinear equation of state, the mixed element generally has a different density. Furthermore, the effects from downgradient diffusion of  $\Theta$ , coupled to the sign-definite nature of the cabbeling coefficient, render a mixed fluid element with a greater density than the unmixed elements. This densification upon mixing is a physical process known as **cabbeling** (*McDougall, 1987b*).

As noted above, the sign definite nature of cabbeling (i.e., cabbeling always results in denser fluid elements after mixing) is a direct result of the intrinsic geometry of density surfaces in Conservative Temperature and salinity space, combined with the downgradient orientation of the neutral diffusive flux of  $\Theta$ . As a result, the cabbeling source satisfies

$$\text{cabbeling} \equiv -\mathcal{C} \mathbf{J}^{(\Theta)} \cdot \nabla \Theta \geq 0, \quad (19.41)$$

thus providing a mathematical expression for the cabbeling source (with dimensions of density per time). An increase in the density within a column of seawater results in the reduction of the sea level due to compression of the water column.

### 19.3.3 Thermobaricity

The thermobaricity parameter

$$\mathcal{T} = \beta \partial_p (\alpha / \beta) \quad (19.42)$$

is nonzero due to pressure dependence of the ratio of the thermal expansion coefficient to the haline contraction coefficient. As both thermal and haline effects are present, the parameter  $\mathcal{T}$  is more precisely split into two terms

$$\mathcal{T} = \frac{\partial \alpha}{\partial p} - \frac{\alpha}{\beta} \frac{\partial \beta}{\partial p} = -\frac{\rho_{\Theta p}}{\rho} + \frac{\rho_{\Theta}}{\rho_S} \frac{\rho_{pS}}{\rho}. \quad (19.43)$$

**Thermobaricity** is the common name for this two-term expression, since pressure variations in the thermal expansion coefficient dominate those of the haline contraction coefficient; i.e., halobaricity is subdominant to thermobaricity. The thermal expansion coefficient generally increases as pressure increases, thus making the thermobaric parameter positive.

Since the neutral gradient of  $\Theta$  need not be oriented in a special manner relative to the neutral gradient of pressure, there is no sign-definite nature to the thermobaricity source term (with units of density per time)

$$\text{thermobaricity} \equiv -\mathcal{T} \mathbf{J}^{(\Theta)} \cdot \nabla p \quad (19.44)$$

appearing in equation (19.38). Thus, thermobaricity can either increase or decrease density, depending on details of the density and fluxes. However, as noted by [McDougall and You \(1990\)](#), thermobaricity typically increases density in much of the World Ocean.

### 19.3.4 Comments

Cabbeling and thermobaricity lead to watermass transformation and associated dianeutral transport of seawater. However, these processes are distinct from other mixing processes such as breaking gravity waves (Section 18.1). Namely, cabbeling and thermobaricity arise from the transport of  $\Theta$  and  $S$  by mesoscale eddies along neutral directions, which in turn is parameterized via neutral diffusion of these two active tracers. Transient mesoscale eddies impart a downscale cascade of tracer variance that is ultimately halted by irreversible molecular mixing. This mixing is the ultimate cause for cabbeling and thermobaricity, with the overall strength of the cabbeling and thermobaricity determined by the strength of the mesoscale stirring.

### 19.3.5 Further study

[Griffies and Greatbatch \(2012\)](#) discuss the impacts on global mean sea level from thermobaricity and cabbeling as diagnosed from an ocean model. Given that cabbeling always densifies and thermobaricity is also dominated by densification, these processes lead to a general reduction in global mean sea level. [Klocker and McDougall \(2010\)](#), [Groeskamp et al. \(2016\)](#), and [Groeskamp et al. \(2019\)](#) diagnose cabbeling and thermobaricity from observational based measurements, with [Groeskamp et al. \(2019\)](#) also offering a more robust numerical method for performing that diagnostic calculation. [McDougall \(2025\)](#) provides a review of various thermodynamic concepts used in ocean physics, including a discussion of cabbeling and thermobaricity.

## 19.4 Salt and freshwater continuity equations

We specialize the kinematics of material tracers given in VOLUME 1 to here focus on seawater, which we treat as a two component fluid comprised of salt and freshwater. We extend this discussion in Section 19.6 by studying the role of surface boundary transports of salt, enthalpy, and water, and how they affect ocean buoyancy.

### 19.4.1 Salt and freshwater

Seawater is comprised of two material tracers: freshwater plus a suite of dissolved trace “salts”. The ratio of salts is roughly constant over the World Ocean. We are thus able to make use of a single effective mass concentration known as the *salt concentration*<sup>3</sup>

$$S = \frac{\text{mass of salt}}{\text{mass of seawater}} = \frac{\text{mass of salt}}{\text{mass of freshwater} + \text{mass of salt}} \quad (19.45)$$

<sup>3</sup>We use the salt concentration,  $S$ , in this section to avoid 1/1000 factors needed if working with salinity,  $S = 1000 \text{S}$ .

to specify the amount of salt within an element of seawater. In practice oceanographers choose to work with the **salinity**,<sup>4</sup>

$$S = 1000 \text{S}, \quad (19.46)$$

which converts from typical salt concentrations of  $S = 0.035$  to a salinity of  $S = 35$ . The complement to salt concentration is the freshwater concentration or mass fraction for an element of seawater

$$F = \frac{\text{mass of freshwater}}{\text{mass of seawater}} = \frac{\text{mass of freshwater}}{\text{mass of freshwater} + \text{mass of salt}} = 1 - S. \quad (19.47)$$

Other trace matter occurs at very low concentrations so as to make seawater, in effect, a two-component fluid consisting of freshwater plus dissolved salt.<sup>5</sup> We here derive the mass budget for salt and freshwater as well as the associated kinematic boundary conditions.

## 19.4.2 Continuity equations

Following our discussion of the material tracer equation in VOLUME 1, we write the mass continuity equations for an element of seawater as

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \quad \text{seawater} \quad (19.48)$$

$$\partial_t (\rho S) + \nabla \cdot (\rho \mathbf{v} S + \mathbf{J}^{(S)}) = 0 \quad \text{salt} \quad (19.49)$$

$$\partial_t (\rho F) + \nabla \cdot (\rho \mathbf{v} F + \mathbf{J}^{(F)}) = 0 \quad \text{freshwater.} \quad (19.50)$$

Equation (19.48) is the mass continuity equation for seawater and equation (19.49) is the mass continuity equation for salt. The freshwater continuity equation (19.50) is derived by subtracting the salt equation (19.49) from the seawater equation (19.48). Hence, only two of the three continuity equations (19.48)-(19.50) are independent.

We make use of the **barycentric velocity** in the above conservation laws, where the barycentric velocity for the ocean is given by

$$\mathbf{v} = S \mathbf{v}^{(S)} + F \mathbf{v}^{(F)}. \quad (19.51)$$

The velocities  $\mathbf{v}^{(S)}$  and  $\mathbf{v}^{(F)}$  are, respectively, the molecular center of mass velocities for salt and freshwater within a fluid element, in which case

$$\partial_t S + \mathbf{v}^{(S)} \cdot \nabla S = 0 \quad \text{and} \quad \partial_t F + \mathbf{v}^{(F)} \cdot \nabla F = 0. \quad (19.52)$$

Furthermore, the fluxes  $\mathbf{J}^{(S)}$  and  $\mathbf{J}^{(F)}$  arise from the difference of the salt and freshwater velocities from the barycentric velocity

$$\mathbf{J}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v}) \quad \text{and} \quad \mathbf{J}^{(F)} = \rho F (\mathbf{v}^{(F)} - \mathbf{v}). \quad (19.53)$$

As seen in Chapter 18, these fluxes are commonly parameterized using a subgrid tensor as per

$$\mathbf{J}^{(S)} = -\rho \mathbf{E} \cdot \nabla S \quad \text{and} \quad \mathbf{J}^{(F)} = -\rho \mathbf{E} \cdot \nabla F, \quad (19.54)$$

---

<sup>4</sup>More precisely, the **salinity**,  $S$ , as defined by equation (19.46) is the *Absolute Salinity*. Absolute Salinity is distinct from the *practical salinity* determined by conductivity measurements. [IOC et al. \(2010\)](#) provides a full accounting of the theory and practice of ocean salinity.

<sup>5</sup>See [IOC et al. \(2010\)](#) for more discussion of the variations of salt concentration ratios over the ocean, as well as the impacts from biogeochemical tracers.

where  $\mathbf{E}$  is the subgrid scale (eddy) transport tensor for salt in seawater, which includes both diffusive and skew diffusive elements. We use the same eddy tensor for salt and freshwater since the eddy mixing of one is balanced by the other

$$\mathbf{J}^{(S)} = -\rho \mathbf{E} \cdot \nabla S = \rho \mathbf{E} \cdot \nabla F = -\mathbf{J}^{(F)}. \quad (19.55)$$

The advective flux of seawater is comprised of a salt flux plus a freshwater flux

$$\rho \mathbf{v} = \rho S \mathbf{v}^{(S)} + \rho F \mathbf{v}^{(F)}. \quad (19.56)$$

Conversely, the salt flux and freshwater flux can be represented as a non-advective flux plus an advective flux, where advection is defined by the barycentric velocity

$$\rho S \mathbf{v}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v}) + \rho S \mathbf{v} = \mathbf{J}^{(S)} + \rho S \mathbf{v} \quad (19.57a)$$

$$\rho F \mathbf{v}^{(F)} = \rho F (\mathbf{v}^{(F)} - \mathbf{v}) + \rho F \mathbf{v} = \mathbf{J}^{(F)} + \rho F \mathbf{v}. \quad (19.57b)$$

The non-advective fluxes,  $\mathbf{J}^{(S)}$  and  $\mathbf{J}^{(F)}$ , lead to an exchange of mass with zero net movement of mass. In contrast, the advective flux moves mass according to the barycentric velocity. Furthermore, the center of mass velocities,  $\mathbf{v}^{(S)}$  and  $\mathbf{v}^{(F)}$ , offer a conceptual framework of use to formulate the kinematic boundary conditions. Even so, they offer no new information beyond that contained in the fluxes  $\mathbf{J}^{(S)}$  and  $\mathbf{J}^{(F)}$ .

## 19.5 Surface boundary conditions for $S$ and $\Theta$

In this section we summarize the surface boundary conditions holding for the salinity and Conservative Temperature equations. This treatment complements that given in VOLUME 2 when studying boundary conditions for material tracers.

### 19.5.1 Salt and freshwater

In deriving the surface [kinematic boundary condition](#) in VOLUME 2, we made use of the [barycentric velocity](#),  $\mathbf{v}$ , for an element of seawater. We can garner further kinematic insights into the two-component ocean system by decomposing the boundary mass flux into contributions from salt and freshwater

$$Q_m = Q_S + Q_F, \quad (19.58)$$

and by introducing the center of mass velocities for salt and freshwater according to

$$-Q_m = \rho (\mathbf{v} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} \quad (19.59a)$$

$$= \rho [S \mathbf{v}^{(S)} + F \mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}] \cdot \hat{\mathbf{n}} \quad (19.59b)$$

$$= \rho [S (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) + F \mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}] \cdot \hat{\mathbf{n}} \quad (19.59c)$$

$$= S \rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} + (1 - S) \rho (\mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} \quad (19.59d)$$

$$= S \rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} + F \rho (\mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} \quad (19.59e)$$

$$\equiv -(Q_S + Q_F), \quad (19.59f)$$

where we wrote

$$S \rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -Q_S \quad (19.60a)$$

$$\mathcal{F}\rho(\mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_F. \quad (19.60b)$$

In these equations, we introduced the velocity,  $\mathbf{v}^{(\eta)}$ , of a point fixed to the free surface. We only need the projection of this velocity in the outward normal direction, which is written (see the kinematics from VOLUME 2)

$$\mathbf{v}^{(\eta)} \cdot \hat{\mathbf{n}} = \frac{\partial \eta / \partial t}{|\nabla(z - \eta)|} = \frac{\partial \eta / \partial t}{\sqrt{1 + |\nabla \eta|^2}} \implies \mathbf{v}^{(\eta)} \cdot \hat{\mathbf{n}} dS = \partial_t \eta dA, \quad (19.61)$$

where  $dS$  is the area element on the free surface and  $dA$  is its horizontal projection. Note that in many regions, there is zero flux of salt across the ocean surface, in which case the ocean surface acts as a material surface in terms of the salt velocity

$$\rho(\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{zero surface salt flux.} \quad (19.62)$$

The key exception to this boundary condition concerns sea ice, whereby salt is exchanged between liquid seawater and sea ice upon the melting or freezing of ice.

For most applications, it is preferable to make use of equation (19.57a) to eliminate the salt velocity  $\mathbf{v}^{(S)}$  in favor of the non-advection flux  $\mathbf{J}^{(S)} = \rho S(\mathbf{v}^{(S)} - \mathbf{v})$ , in which case the kinematic boundary condition (19.60a) takes the form

$$-\mathcal{Q}_S = S\rho(\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = S\rho(\mathbf{v}^{(S)} - \mathbf{v} + \mathbf{v} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = \mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} - S\mathcal{Q}_m. \quad (19.63)$$

Turning this equation around leads to the expression for the non-advection flux

$$\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = S\mathcal{Q}_m - \mathcal{Q}_S = S\mathcal{Q}_F - \mathcal{F}\mathcal{Q}_S, \quad (19.64)$$

which relates the mass transport crossing the ocean surface at  $z = \eta$  (right hand side) to the non-advection salt transport on the ocean side of the surface boundary (left hand side). A form of this equation is also given in VOLUME 2 as part of the study of tracer kinematics.

To support intuition and to check signs, consider the case with  $\mathcal{Q}_S = 0$  so that  $\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = S\mathcal{Q}_F$ . Equation (19.64) then means there is an upward non-advection flux of salt ( $\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} > 0$ ) on the ocean side of the  $z = \eta$  boundary in the presence of an input of freshwater through the ocean surface ( $S\mathcal{Q}_F > 0$ ). For the converse, let  $\mathcal{Q}_F = 0$  so that  $\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = -\mathcal{F}\mathcal{Q}_S$ . Now, there is a downward non-advection flux of salt ( $\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} < 0$ ) on the ocean side of the  $z = \eta$  boundary in the presence of salt input through the ocean surface ( $\mathcal{F}\mathcal{Q}_S > 0$ ).

## 19.5.2 The non-advection salt flux boundary condition

The above properties of boundary mass transfer result from the kinematic property of a fluid element whose mass is constant, and so the transfer of freshwater across the boundary of a fluid element is compensated by an opposite transfer of salt. The ocean boundary interface acts as a boundary for the fluid elements adjacent to the surface. Hence, to move mass across the  $z = \eta$  interface requires mass to be replenished to the surface fluid elements.

### Diffusive closure for the non-advection flux

Consider an ocean without any mixing, such as for a perfect fluid. In this case, mass arriving to the ocean surface from  $\mathcal{Q}_m > 0$  cannot be incorporated into the ambient ocean fluid, but instead will form a separate unmixed surface lens. When mass is exchanged across the ocean surface,

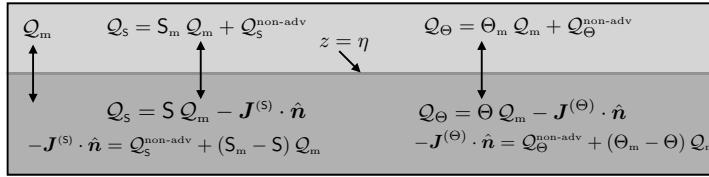


FIGURE 19.1: A schematic of an infinitesimal region of the ocean surface boundary at  $z = \eta(x, y, t)$ , with  $z < \eta$  the liquid ocean.  $Q_m dS = Q_m dA$  is the mass transport (mass per time) that crosses the  $z = \eta$  interface. We depict the case for salt concentration,  $S$ , and Conservative Temperature,  $\Theta$ , and the expressions for their net boundary fluxes. These boundary conditions are derived in Section 19.5.2 for salinity and Section 19.5.3 for Conservative Temperature.

mixing is required to incorporate the mass into the ambient ocean fluid. To determine the level of mixing, assume that  $\mathbf{J}^{(S)}$  takes the form of a diffusive flux (19.54) so that the boundary condition (19.64) becomes

$$\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = -\rho [\mathbf{K} \cdot \nabla S] \cdot \hat{\mathbf{n}} = S Q_m - Q_S = S Q_F - F Q_S, \quad (19.65)$$

where  $\mathbf{K}$  is the symmetric diffusion tensor.<sup>6</sup> This equation sets the level of diffusion on the ocean side of the surface boundary that is needed to generate the non-advectional transport. The diffusive mixing of salt and freshwater mediate the transfer of mass across the ocean surface so to incorporate that mass into the ambient ocean fluid. For example, freshwater added to the ocean ( $Q_F > 0$ ) diffuses downward as salt diffuses upward toward the surface.

### Salt dissolved within the mass transport

In the case when salt is transported across the ocean boundary, as occurs with sea ice melting and formation, it does so largely dissolved in the water that is transported. There can also be a non-advectional transport, such as via parameterized turbulent fluxes, so that the net salt flux is given by

$$Q_S = S_m Q_m + Q_S^{\text{non-adv}}. \quad (19.66)$$

If there are more sources of this transfer then a relation such as this holds for each process. We are thus led to the net salt flux

$$Q_S = -[\rho S (\mathbf{v} - \mathbf{v}^{(\eta)}) + \mathbf{J}^{(S)}] \cdot \hat{\mathbf{n}} = S Q_m - \mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = S_m Q_m + Q_S^{\text{non-adv}}. \quad (19.67)$$

which leads to the non-advectional salt flux on the ocean side of the boundary

$$-\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = Q_S^{\text{non-adv}} + (S_m - S) Q_m. \quad (19.68)$$

Figure 19.1 provides a schematic summary of the salt flux boundary condition. Furthermore, note that this boundary condition is consistent with that derived in VOLUME 1 for a general tracer.

### Treatment in observational analyses and numerical models

In ocean climate modeling applications, the salt mass flux,  $Q_S$ , typically does not affect the kinematic boundary conditions. This approximation is reasonable given that the dominant contributor to the mass flux,  $Q_m$ , is the freshwater. Even so, there remains a net salt transported

<sup>6</sup>As seen in Chapter 18, parameterized skew fluxes satisfy a no-flux condition at boundaries.

across the ocean surface in the presence of sea ice melt and formation. The above boundary conditions, in particular equations (19.67) and (19.68), remain unchanged. Furthermore, it is necessary to specify the  $z = \eta$  boundary concentration. For salt, this concentration is typically set equal to that within the ocean model surface grid cell. This choice is also common for observation-based studies.

### 19.5.3 Conservative Temperature boundary condition

Conservative Temperature, potential temperature, potential vorticity, and passive tracers each satisfy the tracer equation (19.49), with distinct tracer flux vectors,  $\mathbf{J}$ . However, they are not material tracers and so the kinematic constraints holding for salt and freshwater do not hold for these non-material tracers. We study the thermodynamic properties of Conservative Temperature in VOLUME 2, and the processes affecting its boundary fluxes in Section 19.6. Here we begin our treatment of this tracer by deriving its surface boundary condition, largely following our treatment of salinity in Section 19.5.2.

The net surface boundary flux of Conservative Temperature is written

$$\mathcal{Q}_\Theta = \Theta_m \mathcal{Q}_m + \mathcal{Q}_\Theta^{\text{non-adv}} = -[\rho \Theta (\mathbf{v} - \mathbf{v}^{(\eta)}) + \mathbf{J}^{(\Theta)}] \cdot \hat{\mathbf{n}} = \Theta \mathcal{Q}_m - \mathbf{J}^{(\Theta)} \cdot \hat{\mathbf{n}}. \quad (19.69)$$

In this equation,  $\mathcal{Q}_\Theta^{\text{non-adv}}$  arises from the non-advection enthalpy fluxes outside the ocean domain that impact on the upper ocean interface, such as from radiant and turbulent fluxes, whereas  $\Theta_m$  is the Conservative Temperature of the boundary mass flux. Rearrangement leads to the non-advection flux on the ocean side of the upper ocean boundary

$$-\mathbf{J}^{(\Theta)} \cdot \hat{\mathbf{n}} = \mathcal{Q}_\Theta^{\text{non-adv}} + (\Theta_m - \Theta) \mathcal{Q}_m, \quad (19.70)$$

where  $\Theta = \Theta(z = \eta)$  is the Conservative Temperature at the surface interface. A common assumption made for models and observational studies is to set  $\Theta_m - \Theta(z = \eta) = 0$ , in which case

$$-\mathbf{J}^{(\Theta)} \cdot \hat{\mathbf{n}} = \mathcal{Q}_\Theta^{\text{non-adv}} \quad \text{if } \Theta_m = \Theta(z = \eta). \quad (19.71)$$

Figure 19.1 provides a schematic summary of the  $\Theta$  flux boundary condition.

### 19.5.4 Comments and further reading

We make use of many results from this section when discussing surface ocean buoyancy fluxes in Section 19.6 and water mass transformation in Section 20.6. Furthermore, [Nurser and Griffies \(2019\)](#) offer further discussion of the kinematic boundary condition for salt and freshwater, with that paper motivated by questions related to water mass transformation considered in Section 20.6.

## 19.6 Surface boundary fluxes of buoyancy

As studied in VOLUME 2, **buoyancy** measures the gravitational acceleration of a fluid element relative to that of the fluid environment surrounding the element. Changes in buoyancy arise through changes in density associated with temperature and salinity changes, with buoyancy changes computed relative to a fixed pressure level. In this section we derive the equation describing changes in ocean buoyancy due to enthalpy (commonly referred to as “heat”), salt, and water fluxes crossing the ocean surface boundary. For this purpose, we expose certain

of the issues associated with coupling numerical models of the ocean, atmosphere, and land. A detailed treatment of boundary layer physics is well outside of our scope. We thus take a phenomenological perspective, developing budget equations but not diving into details of the turbulent exchange of matter and enthalpy across the ocean surface boundary. Similar considerations hold for the ocean bottom boundary, which is insulating and material except in regions of geothermal fluxes.

### 19.6.1 Surface boundary fluxes of enthalpy and salt

Broadly, the surface boundary fluxes are associated with the following physical processes.

- Turbulent processes transfer enthalpy through latent and sensible heating.
- Longwave radiation cools the upper ocean, with this radiation affected by the upper ocean skin temperature.
- Penetrative shortwave radiation is absorbed in seawater and so increases buoyancy in regions where the thermal expansion coefficient is positive.<sup>7</sup>
- All of the above transports arise from **non-advectione fluxes** in that they are not associated with a net mass transport across the ocean surface. In contrast, advective processes transfer enthalpy and salt across the ocean surface through the transfer of mass.
- Salt is transferred between the liquid ocean and sea ice when sea ice melts and forms. This transfer is proportional to the boundary flux of water and the difference in salinity between the liquid ocean and solid sea ice. There can be additional turbulent salt fluxes as well, but there is typically a negligible transfer of salt associated with precipitation, evaporation, and river runoff.

### 19.6.2 Evolution from surface boundary fluxes

We now develop finite volume budget equations for potential enthalpy (via Conservative Temperature,  $\Theta$ ), salt, and seawater mass for a grid cell region next to the ocean surface (see Figure 19.2), with a focus on contributions due to surface boundary fluxes. For that purpose, introduce the following quantities for a grid cell,

$$M = \int_{\text{cell}} \rho dV = \langle \rho \rangle V \quad V = \int_{\text{cell}} dV = A \bar{h} \quad A = \int_{\text{cell}} dA \quad (19.72\text{a})$$

$$\bar{h} A = \int_{\text{cell}} \left[ \int_{\text{cell}} dz \right] dA \quad \langle C \rangle M = \int_{\text{cell}} C \rho dV, \quad (19.72\text{b})$$

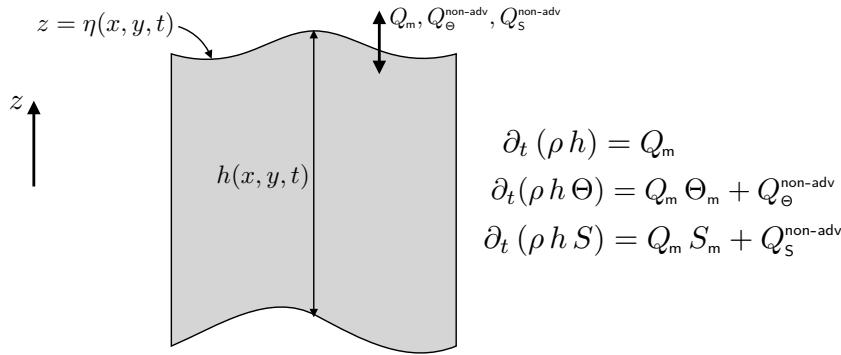
so that  $\langle \rho \rangle$  is the cell averaged density,  $\langle C \rangle$  is the cell averaged tracer concentration,  $\bar{h}$  is the cell area averaged thickness,  $V$  is the cell volume, and  $A$  is the cell horizontal area. These definitions allow us to write

$$\frac{d}{dt} \left[ \int_{\text{cell}} \rho C dV \right] = \frac{d}{dt} [\langle C \rangle M] = A \frac{d}{dt} [\bar{h} \langle C \rangle \langle \rho \rangle], \quad (19.73)$$

where the horizontal area of a cell is assumed to be constant in time and the cell is bounded by vertical side walls. The surface boundary fluxes have similar grid cell area averages.

---

<sup>7</sup>The Baltic Sea is an outlier in the World Ocean, whose fresh and cold waters often realize a negative thermal expansion so that shortwave radiation can increase density rather than reduce it.



$$\begin{aligned}\partial_t(\rho h) &= Q_m \\ \partial_t(\rho h \Theta) &= Q_m \Theta_m + Q_\Theta^{\text{non-adv}} \\ \partial_t(\rho h S) &= Q_m S_m + Q_S^{\text{non-adv}}\end{aligned}$$

FIGURE 19.2: Schematic of a finite volume region next to the ocean surface used to develop the budget equations for mass, enthalpy, and salt. To derive the surface boundary condition for buoyancy, it is sufficient to ignore all fluxes crossing interior surfaces, focusing instead on those fluxes crossing the ocean free surface at  $z = \eta(x, y, t)$ .

Focusing just on contributions from surface boundary transport leads to the budget equations

$$\partial_t(\rho h \Theta) = Q_m \Theta_m + Q_\Theta^{\text{non-adv}} \quad (19.74a)$$

$$\partial_t(\rho h S) = Q_m S_m + Q_S^{\text{non-adv}} \quad (19.74b)$$

$$\partial_t(\rho h) = Q_m, \quad (19.74c)$$

where we wrote a partial time derivative since we are holding the horizontal position fixed. Furthermore, we reduced notational clutter by dropping the angle brackets for volume average and the horizontal overline for area average. For a three-dimensional budget, the right hand side to these equations is combined with fluxes crossing interior cell boundaries, which are ignored here since we are focused just on surface fluxes. Finally, we wrote the fluxes as

$$Q_m A = Q_m \mathcal{S} \quad \text{and} \quad Q_\Theta^{\text{non-adv}} A = Q_\Theta^{\text{non-adv}} \mathcal{S} \quad \text{and} \quad Q_S^{\text{non-adv}} A = Q_S^{\text{non-adv}} \mathcal{S}, \quad (19.75)$$

where  $\mathcal{S}$  is the area on the free surface and  $A$  is the corresponding horizontal area of the grid cell.<sup>8</sup>

### 19.6.3 Buoyancy tendency from surface fluxes

For many purposes, it is of interest to quantify the impacts on ocean buoyancy arising from surface boundary fluxes. For that purpose, we here develop the budget for buoyancy in a surface model grid cell region, focusing on surface flux contributions.

Buoyancy has a local time tendency given by

$$-\frac{\rho_\Theta}{g} \frac{\partial b}{\partial t} = \rho_\Theta \frac{\partial \Theta}{\partial t} + \rho_S \frac{\partial S}{\partial t}, \quad (19.76)$$

where we introduced the shorthand

$$\rho_\Theta = \left[ \frac{\partial \rho}{\partial \Theta} \right]_{S,p} \quad \text{and} \quad \rho_S = \left[ \frac{\partial \rho}{\partial S} \right]_{\Theta,p} \quad (19.77)$$

<sup>8</sup>The surface area,  $\mathcal{S}$ , along the  $z = \eta$  boundary is not generally horizontal and it is not generally constant in time. Yet its horizontal projection,  $A$ , is time independent since we fix the horizontal positions for the vertical cell walls.

for the partial derivatives of density with respect to Conservative Temperature and salinity, respectively, each with pressure held constant. We wish to form an evolution equation for buoyancy at the ocean surface grid cell just due to the effects of surface forcing. For this purpose, multiply the temperature equation (19.74a) by  $\rho_\Theta$  and add to the salinity equation (19.74b) multiplied by  $\rho_S$

$$\rho_\Theta \left[ \frac{\partial(\rho h \Theta)}{\partial t} \right] + \rho_S \left[ \frac{\partial(\rho h S)}{\partial t} \right] = Q_m (\rho_\Theta \Theta_m + \rho_S S_m) + \rho_\Theta Q_\Theta^{\text{non-adv}} + \rho_S Q_S^{\text{non-adv}}. \quad (19.78)$$

Now use the mass budget (19.74c) and introduce the buoyancy tendency according to equation (19.76) to render an expression for the time tendency of the surface ocean buoyancy<sup>9</sup>

$$-(\rho_0/g) \rho h \left[ \frac{\partial b}{\partial t} \right]^{\text{surf}} = Q_m [\rho_\Theta (\Theta_m - \Theta) + \rho_S (S_m - S)] + \rho_\Theta Q_\Theta^{\text{non-adv}} + \rho_S Q_S^{\text{non-adv}}. \quad (19.79)$$

Introducing the thermal expansion and saline contraction coefficients

$$\alpha = -\frac{1}{\rho} \left[ \frac{\partial \rho}{\partial \Theta} \right]_{S,p} \quad \text{and} \quad \beta = \frac{1}{\rho} \left[ \frac{\partial \rho}{\partial S} \right]_{\Theta,p} \quad (19.80)$$

yields

$$\left[ \frac{\partial b}{\partial t} \right]^{\text{surf}} = \frac{g}{\rho_0 h} \left( Q_m [\alpha (\Theta_m - \Theta) - \beta (S_m - S)] + \alpha Q_\Theta^{\text{non-adv}} - \beta Q_S^{\text{non-adv}} \right) \quad (19.81a)$$

$$= \frac{g}{\rho_0 h} \left( \alpha [Q_m (\Theta_m - \Theta) + Q_\Theta^{\text{non-adv}}] - \beta [Q_m (S_m - S) + Q_S^{\text{non-adv}}] \right). \quad (19.81b)$$

In regions where the thermal expansion coefficient is positive ( $\alpha > 0$ ), adding a boundary mass ( $Q_m > 0$ ) that has  $\Theta_m > \Theta$  increases the buoyancy of the surface ocean; i.e., adding relatively warm water increases surface ocean buoyancy. Likewise, where the haline contraction coefficient is positive ( $\beta > 0$ ), adding boundary mass with  $S_m < S$  increases buoyancy of the surface ocean; i.e., adding relatively freshwater increases surface ocean buoyancy. The same behavior holds for the turbulent fluxes, where  $Q_\Theta^{\text{non-adv}} > 0$  (adding turbulent thermal energy to the ocean) increases surface ocean buoyancy whereas  $Q_S^{\text{non-adv}} > 0$  (adding salt to the ocean) decreases buoyancy.

Following from our study of Green's functions for Neumann boundary conditions in VOLUME 1, we consider the limit as the thickness,  $h$ , becomes vanishingly small, in which case we introduce a Dirac delta

$$\left[ \frac{\partial b}{\partial t} \right]^{\text{surf}} = \frac{g \delta(z - \eta)}{\rho_0} \left( Q_m [\alpha (\Theta_m - \Theta) - \beta (S_m - S)] + \alpha Q_\Theta^{\text{non-adv}} - \beta Q_S^{\text{non-adv}} \right) \quad (19.82a)$$

$$= \frac{g \delta(z - \eta)}{\rho_0} \left( \alpha [Q_m (\Theta_m - \Theta) + Q_\Theta^{\text{non-adv}}] - \beta [Q_m (S_m - S) + Q_S^{\text{non-adv}}] \right). \quad (19.82b)$$

This Dirac delta sheet form for the boundary fluxes is quite useful when organizing physical processes according to whether they arise from the ocean interior versus ocean boundary. It proves particularly useful for water mass transformation analysis in Section 20.7.

---

<sup>9</sup>We provide the few missing details in Exercise 19.2.

### 19.6.4 Comments

The buoyancy flux expression (19.81b) is of use for boundary layer parameterizations, such as the KPP scheme of [Large et al. \(1994\)](#) and [Van Roekel et al. \(2018\)](#). It is furthermore used in our study of water mass transformations in Chapter 20.

## 19.7 Global mean sea level

In this section we consider some basic features of global mean sea level by making use of the mass budget of liquid seawater. This analysis highlights the distinction between the mass budget and the volume (sea level) budget. In particular, mass is an extensive property and so it satisfies a conservation budget, with the total ocean mass affected only through boundary fluxes. In contrast, volume, just like buoyancy, has interior sources and sinks so that the ocean volume can change even if there are no boundary fluxes of volume.

### 19.7.1 Definitions and assumptions

The liquid seawater mass is given by the integral

$$M = \int_{\mathcal{R}} \rho dV, \quad (19.83)$$

with our interest in a global integral over the full liquid ocean domain,  $\mathcal{R}$ . Since mass is an extensive quantity, it has a time tendency due to boundary transport, here written as

$$\frac{dM}{dt} = A \overline{Q_m}, \quad (19.84)$$

where  $A$  is the ocean surface area,

$$\overline{Q_m} = \frac{1}{A} \int Q_m dA, \quad (19.85)$$

is the area averaged surface mass flux, and we ignore any mass entering through the ocean bottom. The global volume of liquid seawater

$$V = \int_{\mathcal{R}} dV = \frac{M}{\langle \rho \rangle} \quad (19.86)$$

changes due to mass changes and changes to the global mean density,

$$\langle \rho \rangle = \frac{\int_{\mathcal{R}} \rho dV}{\int_{\mathcal{R}} dV} = \frac{M}{V}. \quad (19.87)$$

Throughout this section we assume the surface area is constant in time, thus neglecting the relatively small changes associated with volume changes along sloping beaches. We also assume a temporally constant area averaged ocean bottom depth,  $\overline{H}$ . These two assumptions result in a time tendency in ocean volume that arises just from time tendencies in the global mean sea level,  $\bar{\eta}$ . Since around the year 2000, measurements estimate that global area mean sea level has increased at a rate of

$$\left[ \frac{d\bar{\eta}}{dt} \right]_{\text{observed}} \approx 3 \text{ mm yr}^{-1}. \quad (19.88)$$

To attach further numbers to the analysis in this section we make use of the following phenomenological numbers and make a few assumptions to facilitate calculations.

- global seawater volume  $V \approx 1.3 \times 10^{18} \text{ m}^3$
- global ocean surface area  $A \approx 3.6 \times 10^{14} \text{ m}^2$
- global ocean mean density  $\langle \rho \rangle \approx 1035 \text{ kg m}^{-3}$
- specific heat capacity for seawater  $c_p \approx 3992 \text{ J kg}^{-1}\text{K}^{-1}$
- Ignore mass fluxes transported through the sea floor, which are small relative to surface mass fluxes.
- Ignore salinity and pressure effects on density, so that changes in global mean density arise just from changes in global mean Conservative Temperature.
- Assume a constant thermal expansion coefficient

$$\alpha = -\frac{1}{\rho} \left[ \frac{\partial \rho}{\partial \Theta} \right]_{S,p} \approx 2 \times 10^{-4} \text{ K}^{-1}. \quad (19.89)$$

This is not a great approximation, since the thermal expansion coefficient ranges over the ocean by a factor of 10. Nonetheless, for this section it is sufficient for computing estimates that support conceptual understanding that is not outside the bounds of error bars in measurements for global boundary enthalpy and mass fluxes.

### 19.7.2 Budget for global mean sea level

Expression (19.86) for ocean volume leads to its time derivative

$$\frac{dV}{dt} = \frac{1}{\langle \rho \rangle} \frac{dM}{dt} - \frac{M}{\langle \rho \rangle^2} \frac{d\langle \rho \rangle}{dt} \quad (19.90a)$$

$$= \frac{A \overline{Q_m}}{\langle \rho \rangle} - \frac{V}{\langle \rho \rangle} \frac{d\langle \rho \rangle}{dt}, \quad (19.90b)$$

where we used equation (19.84) to express mass changes in terms of the surface mass flux. Additionally, the ocean volume is connected to sea level via

$$V = \int dA \int_{-H}^{\eta} dz = A (\overline{H} + \bar{\eta}), \quad (19.91)$$

so that time changes in ocean volume arise from time changes in the global mean sea level

$$\frac{dV}{dt} = A \frac{d\bar{\eta}}{dt}. \quad (19.92)$$

Combining the two volume equations (19.90b) and (19.92) yields the budget equation for global mean sea level

$$\frac{d\bar{\eta}}{dt} = \frac{\overline{Q_m}}{\langle \rho \rangle} - \frac{V}{A \langle \rho \rangle} \frac{d\langle \rho \rangle}{dt}. \quad (19.93)$$

The first term arises from changes in ocean mass (increasing total mass increases sea level) whereas the second term arises from changes in global mean seawater density (increasing the mean density decreases sea level).

### 19.7.3 Changes due to mass input

To ground these formulae in phenomenology, assume that a surface mass flux gives one-half of the observed sea level rise

$$\frac{1}{2} \left[ \frac{d\bar{\eta}}{dt} \right]_{\text{observed}} = \overline{Q_m} / \langle \rho \rangle, \quad (19.94)$$

with the other half due to changes in mean density. With  $\langle \rho \rangle = 1035 \text{ kg m}^{-3}$  and  $d\bar{\eta}/dt \approx 3 \text{ mm yr}^{-1}$ , we need an area averaged mass flux across the ocean surface

$$\overline{Q_m} \approx 5 \times 10^{-8} \text{ kg m}^{-2} \text{ s}^{-1}. \quad (19.95)$$

Integrated over the global ocean area, this flux leads to a mass transport of

$$\mathcal{T} = A \overline{Q_m} \approx 1.8 \times 10^7 \text{ kg s}^{-1} \approx 0.015 \times \mathcal{T}^{\text{river}}. \quad (19.96)$$

That is, global mean sea level rises at a rate of  $1.5 \text{ mm yr}^{-1}$  if there is a net additional mass added to the ocean equal to roughly 1.5% of the net river water entering the ocean,  $\mathcal{T}^{\text{river}}$ . This additional net mass is largely coming from the melting of land-ice in the form of mountain glaciers and ice sheets.

### 19.7.4 Steric changes due to changes in density

Steric effects generally refer to properties of a substance associated with the space occupied by atoms and molecules. In the sea level context, steric effects refer to changes in sea level associated with density changes, with changes in density associated with changes in the volume occupied by seawater molecules.

Changes in global mean sea level arising from changes in the global mean density are called global **steric sea level** changes. From the sea level budget equation (19.93) we know that steric changes are written mathematically as

$$\left[ \frac{d\bar{\eta}}{dt} \right]_{\text{steric}} \equiv - \frac{V}{A \langle \rho \rangle} \frac{d\langle \rho \rangle}{dt}. \quad (19.97)$$

In Section 19.7.6 we show that global mean density has a time tendency arising primarily from the time tendency in global mean Conservative Temperature. If we assume the ocean thermal expansion is constant, then

$$\frac{1}{\langle \rho \rangle} \frac{d\langle \rho \rangle}{dt} = -\alpha \frac{d\langle \Theta \rangle}{dt}, \quad (19.98)$$

so that steric sea level changes are primarily driven by **thermosteric** effects

$$\left[ \frac{d\bar{\eta}}{dt} \right]_{\text{thermosteric}} \equiv \frac{\alpha V}{A} \frac{d\langle \Theta \rangle}{dt}, \quad (19.99)$$

with increasing water temperature, in the presence of  $\alpha > 0$ , leading to higher sea levels. With  $\alpha \approx 2 \times 10^{-4} \text{ K}^{-1}$  and  $d\bar{\eta}/dt \approx 1.5 \text{ mm yr}^{-1}$ , we have

$$\frac{d\langle \Theta \rangle}{dt} \approx 0.2 \text{ K century}^{-1}. \quad (19.100)$$

That is, a global thermosteric sea level rise of  $1.5 \text{ mm yr}^{-1}$  corresponds to a rate of increase in the global volume mean ocean temperature of roughly  $0.2 \text{ K century}^{-1}$ .

### 19.7.5 Enthalpy flux imbalances implied by thermosteric sea level

The global budget for potential enthalpy is given by

$$\frac{d}{dt} \int_{\mathcal{R}} c_p \rho \Theta dV = \int_{\partial \mathcal{R}} Q_H dA, \quad (19.101)$$

where  $Q_H$  is the boundary flux of enthalpy. Introducing the area and volume mean quantities leads to

$$c_p \frac{d(V\langle\rho\Theta\rangle)}{dt} = A \overline{Q}_H, \quad (19.102)$$

with the expanded left hand side rendering

$$\frac{V}{A} \frac{d\langle\rho\Theta\rangle}{dt} + \langle\rho\Theta\rangle \frac{d\bar{\eta}}{dt} = \overline{Q}_H / c_p. \quad (19.103)$$

Rearrangement then leads to an expression for the global mean sea level tendency

$$\frac{d\bar{\eta}}{dt} = \frac{1}{\langle\rho\Theta\rangle} \left[ \overline{Q}_H - \frac{V}{A} \frac{d\langle\rho\Theta\rangle}{dt} \right], \quad (19.104)$$

which is directly analogous to the sea level equation (19.93) based on the mass budget. A positive tendency in global mean sea level implies a positive right hand side to equation (19.104).

In Exercise 19.6 we show that any conservative tracer,  $C$ , leads to a corresponding kinematic sea level equation just like equation (19.104). These sea level equations arise kinematically through defining the volume mean according to

$$V\langle\rho\rangle = \int_{\mathcal{R}} \rho dV \quad \text{and} \quad V\langle\rho C\rangle = \int_{\mathcal{R}} \rho C dV. \quad (19.105)$$

These definitions then lead to diagnostic connections between changes in conservative scalar properties (seawater mass, tracer mass) and sea level changes.

Rather than focus on the sea level tendency, return to equation (19.102) and estimate  $V/A \approx \overline{H}$ , thus dropping the contribution from  $d\bar{\eta}/dt$ , and further estimate a Boussinesq ocean so that

$$\frac{d\langle\rho\Theta\rangle}{dt} \approx \rho_o \frac{d\langle\Theta\rangle}{dt}, \quad (19.106)$$

in which case

$$\overline{Q}_H \approx \rho_o c_p \overline{H} \frac{d\langle\Theta\rangle}{dt}. \quad (19.107)$$

Inserting the global mean temperature tendency of  $0.2 \text{ K century}^{-1}$  from equation (19.100), as estimated from the sea level tendency, then affords an estimate of the global area mean heat flux crossing the ocean boundary<sup>10</sup>

$$\overline{Q}_H \approx 1 \text{ W m}^{-2}. \quad (19.108)$$

That is, a surface ocean enthalpy flux of roughly  $1 \text{ W m}^{-2}$  (ocean area normalized) gives rise to a global mean thermosteric sea level rise of roughly  $1.5 \text{ mm yr}^{-1}$ .

An enthalpy flux of  $1 \text{ W m}^{-2}$  is small compared to, say, that crossing the surface of a typical light bulb. However,  $1 \text{ W m}^{-2}$  is comparable to that accumulating within the earth system due

<sup>10</sup>The global area mean geothermal heat flux is roughly  $0.094 \text{ W m}^{-2}$  ([Huw Davies, 2013](#)), so that most of the thermosteric sea level rise comes from surface fluxes.

to increases in greenhouse gases ([Otto et al., 2013](#)). That is,  $1 \text{ W m}^{-2}$  averaged over the global ocean, or  $0.7 \text{ W m}^{-2}$  averaged over the surface area of the planet,<sup>11</sup> is roughly the net heating associated with anthropogenic climate change. This increase in surface heating represents a nontrivial increase in the earth's energy budget that is leading to the observed climate changes and sea level rise.<sup>12</sup>

### 19.7.6 Barystatic versus steric effects from near future land ice melt

In this subsection we consider a thought experiment to expose some questions about how global mean sea level changes when melted land ice is added to the ocean. Our concern is mostly with near future sea level rise due to melting land ice, with assumptions made based on that application.<sup>13</sup>

#### Freshwater melt and sea level changes

When freshwater enters the ocean, such as from melting glaciers or 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 reviewed by [Gregory et al. \(2019\)](#). Although ocean salinity, temperature, and pressure changes upon changing the freshwater content, the net effect on global mean sea level is almost entirely barystatic. We here focus on questions related to the changes in salinity, showing how the global halosteric effect is negligible relative to the barystatic effect. But in general, we can understand why the global steric effect (due to changes in salinity, temperature, and pressure) is so tiny by recognizing that freshwater entering the ocean sees its properties change whilst the ambient seawater is itself changed. In particular, the added freshwater has its salinity increase while the ambient seawater is freshened. These compensating effects have corresponding compensating sea level changes, thus bringing the global steric effect to near zero.<sup>14</sup>

We here summarize a two-bucket thought experiment from Appendix B of [Gregory et al. \(2019\)](#) (see also the appendix to [Lowe and Gregory \(2006\)](#)), which focuses on the questions related to salinity. In this experiment, one bucket initially holds freshwater and the other initially holds seawater, with the Conservative Temperature and pressure assumed to be identical for the two buckets, with this assumption not suited to paleoclimate questions such as those considered by [Gebbie \(2020\)](#). Even so, the analysis provides a framing of how to study steric effects associated with changes in ocean mass, with [Gebbie \(2020\)](#) pursuing these questions in far more detail.

#### Formulating the change in volume

Referring to Figure 19.3, consider two buckets that initially contain water with mass  $M_n$ , volume  $V_n$ , density  $\rho_n = M_n/V_n$ , and salinity  $S_n$ , where  $n = 1, 2$  labels the two buckets. Now

<sup>11</sup>The ocean covers roughly 70% of the earth surface. This factor is commonly forgotten when quoting heat flux numbers, so it is important to note whether the number refers to global area normalized or ocean area normalized.

<sup>12</sup>In Exercise 19.3 we compare the net power input to the ocean from anthropogenic effects to the power associated with atomic bombs.

<sup>13</sup>See [Gebbie \(2020\)](#) for applications to paleoclimate studies, where the amount of melt is far larger than that assumed in this subsection. In particular, we here ignore pressure effects, whereas pressure changes are central to the paleoclimate study of [Gebbie \(2020\)](#).

<sup>14</sup>As noted in the appendices to [Lowe and Gregory \(2006\)](#) and [Gregory et al. \(2019\)](#), these compensating effects are often mistakenly ignored.

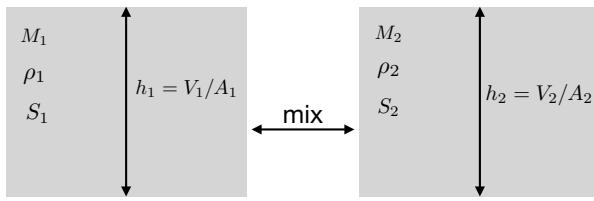


FIGURE 19.3: Schematic of the two-bucket thought experiment from [Gregory et al. \(2019\)](#), used to illustrate the tiny effects on global mean sea level from the global halosteric effect. We consider two homogeneous regions of seawater with mass  $M_n$ , volume  $V_n$ , density  $\rho_n = M_n/V_n$ , and salinity  $S_n$  (or salt concentration  $S_n = S_n/1000$ ), where  $n = 1, 2$  labels the two buckets. The water in the two buckets initially has different salinity, but they are then mixed fully, with the total mass and total salt both conserved during the mixing experiment. We ignore heat of mixing and pressure effects, which is a good approximation for buckets that correspond to upper ocean regions where melt enters, but it is not appropriate for paleoclimate questions such as considered by [Gebbie \(2020\)](#). After mixing, mass  $M_1$  of the homogenized water is put back into the first bucket and  $M_2$  is put into the second bucket. The question is how much has the volume (and hence the vertical thickness) changed upon mixing?

fully mix the water in the two buckets. In so doing, the mass,  $M$ , of seawater remains constant, as does the total mass of salt

$$M = M_1 + M_2 \quad \text{and} \quad M S = M_1 S_1 + M_2 S_2, \quad (19.109)$$

where  $S$  is the salt concentration of the homogenized fluid so that  $M S$  is the total mass of salt in the combined system. After mixing, we return a mass,  $M_1$ , of the homogenized water back in the first bucket, and a mass,  $M_2$ , into the second bucket. Our goal is to compute the change in seawater volume

$$\delta V = \delta V_1 + \delta V_2. \quad (19.110)$$

In determining this volume change, we ignore pressure changes as well as changes in enthalpy associated with the heat of mixing since we are focused here on volume changes arising from salinity changes. Both approximations are quite accurate when considering one of the buckets to be a surface ocean model grid cell and the other to be a bucket from melted land ice. However, when considering longer term paleoclimate questions, these assumptions must be removed as done by [Gebbie \(2020\)](#).

Since the mass of each bucket remains the same before and after homogenization, then the density of seawater in each bucket changes only due to the volume changes

$$\delta \rho_n = \delta(M_n/V_n) = -(M_n/V_n^2) \delta V_n \implies \delta \rho_n / \rho_n = -\delta V_n / V_n. \quad (19.111)$$

That is, the relative change in density equals to minus the relative change in volume. Now the density changes arise just from salinity changes, in which

$$\delta \rho_n / \rho_n = \beta_n \delta S_n, \quad (19.112)$$

where  $\beta_n$  is the haline contraction coefficient that measures changes in density when fixing pressure and Conservative Temperature. We are thus led to the volume change

$$\delta V = -(V_1 \delta \rho_1 / \rho_1 + V_2 \delta \rho_2 / \rho_2) = -(V_1 \beta_1 \delta S_1 + V_2 \beta_2 \delta S_2). \quad (19.113)$$

We can simplify this expression by making use of salt conservation in equation (19.109), thus

constraining salinity changes according to

$$M_1 \delta S_1 + M_2 \delta S_2 = 0 \implies \delta S_2 = -(M_1/M_2) \delta S_1, \quad (19.114)$$

in which case the volume change takes on the form

$$\delta V = -V_1 \beta_1 \delta S_1 \left[ 1 - \frac{\beta_2 \rho_1}{\beta_1 \rho_2} \right]. \quad (19.115)$$

To connect to sea level changes, assume the horizontal cross-sectional area of the two buckets is the same, so that the water column thickness differs between the two buckets by the amount

$$\delta h = -h_1 \beta_1 \delta S_1 \left[ 1 - \frac{\beta_2 \rho_1}{\beta_1 \rho_2} \right]. \quad (19.116)$$

To estimate the sign for the thickness change, consider the case with the first bucket fresh so that  $S_1 = 0$ , and assume this fresh water enters the ocean at a density less than the ambient seawater, so that  $\rho_1 < \rho_2$ . For the high latitude ocean, the haline contraction coefficient is slightly larger for the ambient seawater than for freshwater, so that  $\beta_2/\beta_1 > 1$ . We thus find that the sign for  $\delta h < 0$  depends on details of the density and haline contraction coefficient ratios.

### Numbers for the ocean

To further pursue this thought experiment, assume that the haline contraction coefficient changes only by a few percent globally (see [Roquet et al. \(2015\)](#) or Figure 1 in [Griffies et al. \(2014\)](#)), in which case we set  $\beta_1 = \beta_2 = \beta$ . In this case equation (19.116) simplifies to

$$\delta h = -h_1 \beta \delta S_1 (1 - \rho_1/\rho_2). \quad (19.117)$$

Again, assume the first bucket is initially filled with freshwater ( $S_1 = 0$ ) whereas the second bucket is initially filled with ambient seawater, with salinity  $S_2 = 1000$  S<sub>2</sub> = 25 ppt a representative value. Homogenization of the two buckets then raises salinity in the first bucket and lowers it for the second. Assume the first bucket has its salinity raised nearly to the value of the ambient seawater,<sup>15</sup> so that  $\delta S_1 = S_2$ ; let the density ratio be  $\rho_1/\rho_2 = 1000/1028$ ; and set the haline contraction coefficient to<sup>16</sup>  $\beta = 8 \times 10^{-4}$ . These values then yield a thickness change

$$\delta h/h_1 = -5 \times 10^{-4}. \quad (19.118)$$

Hence, for every meter of fresh water added to the ocean surface, the halosteric effect contributes a 0.5 mm contraction of the water thickness. That is, the total volume of homogenized water equals to the sum of the volume initially in the two separate buckets to within better than 0.05%. We conclude that the volume change is almost entirely barystatic, so that the global halosteric effect is negligible when considering global sea level changes. Furthermore, as noted above, the global halosteric changes are negative, so that we expect the salinity changes to

<sup>15</sup>Consider a numerical model simulation, in which the freshwater enters the top grid cell. During a discrete time step, the freshwater and top grid cell are completely mixed. If the grid cell mass is far larger than the freshwater volume entering that cell over a given time step, we can assume that the salinity of the top cell is only modestly lowered, whereas the salinity of the freshwater is raised to nearly that of the top cell.

<sup>16</sup>We measure the haline contraction coefficient,  $\beta = 0.8 \times 10^{-3}$ , as  $\beta = \rho^{-1} \partial \rho / \partial S = 10^{-3} \rho^{-1} \partial \rho / \partial S$ . The product  $\beta \delta S_1$  is invariant to whether we use salinity or salt concentration.

produce a slight halosteric contraction of sea level.

### What about global thermosteric sea level changes from melt?

The above derivation for the global halosteric changes can be directly transferred to the case of mixing two buckets whose water has different Conservative Temperatures but identical pressure and salinity. Conservation of salt is here replaced by conservation of potential enthalpy, so that the relative thickness change is given by

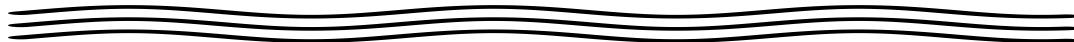
$$\delta h/h_1 = \alpha \delta\Theta_1 (1 - \rho_1/\rho_2), \quad (19.119)$$

where  $\alpha$  is the thermal expansion coefficient. Taking  $\alpha \approx 2 \times 10^{-4} \text{ K}^{-1}$ , we find that  $\delta h$  from thermal effects are on the same order as those from haline effects given by equation (19.116). That is, the contributions to the column thickness are dominated by the barystatic (mass) effects.

There is a key distinction between how the ocean is forced by thermal and haline effects. Namely, ocean thermal properties are affected by boundary radiant and turbulent enthalpy fluxes that are not necessarily associated with boundary mass fluxes. These boundary heat fluxes were studied in Section 19.7.5, where we identified the key role for ocean warming on global thermosteric sea level changes. Such changes are not negligible since they compare to the observed global [barystatic sealevel](#) changes.

### Comments

We conclude from this subsection that melting land ice, as conceived for near future climate change, alters global mean sea level almost exclusively through barystatic effects. Barystatic sea level is transferred very quickly around the World Ocean via barotropic waves. [Lorbacher et al. \(2012\)](#) present model simulations that emphasize the important distinctions between barotropic wave signals of barystatic sea level changes, versus the baroclinic signals associated with steric changes. Global steric effects from meltwater are negligible for near future climate change, though [Gebbie \(2020\)](#) noted its relative increased importance for paleoclimate studies, in which thermosteric, halosteric, and barosteric (due to pressure changes), are sizable. Furthermore, even for modern climate change, once meltwater enters the ocean, it contributes to local steric effects through impacts on temperature and salinity. These local steric effects are transported via advection and diffusion throughout the ocean, thus affecting sea level in the far field away from the melt source, yet doing so on the much longer baroclinic time scales, whereas barystatic effects are transmitted on barotropic time scales.



## 19.8 Exercises

### EXERCISE 19.1: A TIDY FORM FOR CABBELING AND THERMOBARICITY

Introduce the [thermoicity](#) parameter from equation (19.36) and the [cabbeling](#) parameter from equation (19.37) to write the bracket terms appearing in equation (19.35) in forms consistent with those written by [McDougall \(1987b\)](#), in which we have

$$\left[ \frac{D\rho}{Dt} \right]_{\text{ntrldiff}} = -\mathbf{J}^{(\Theta)} \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta), \quad (19.120)$$

thus enabling us to write the material time evolution of *in situ* density due to neutral diffusion

$$\left[ \frac{D\rho}{Dt} \right]_{\text{ntrldiff}} = -\mathbf{J}^{(\Theta)} \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta). \quad (19.121)$$

#### EXERCISE 19.2: MISSING STEPS IN BUOYANCY FLUX DERIVATION

Fill in the missing steps needed to derive equation (19.79) for the buoyancy flux crossing the surface of the ocean.

#### EXERCISE 19.3: GLOBAL WARMING COMPARED TO ATOMIC BOMBS

To gauge the magnitude of a global warming of  $1 \text{ W m}^{-2}$  distributed over the ocean surface area,  $A$ , compare that to the enthalpy flux due to blasting an atomic bomb every second ( $\Delta t = 1 \text{ s}$ ) and uniformly distributing the bomb's released energy,  $\mathcal{E}_{\text{bomb}}$ , over the ocean surface area every second. For this estimate, assume  $\mathcal{E}_{\text{bomb}} \approx 6.3 \times 10^{13} \text{ J}$ . How many bombs are needed to generate  $1 \text{ W m}^{-2}$ ?

#### EXERCISE 19.4: GLOBAL SEA LEVEL IN A BOUSSINESQ OCEAN

As seen in VOLUME 1 and VOLUME 2, the sea level for a Boussinesq ocean evolves according to the kinematic free surface equation

$$\partial_t \eta^{\text{bouss}} = -\nabla \cdot \mathbf{U} + Q_m / \rho_0, \quad (19.122)$$

where

$$\mathbf{U} = \int_{-H}^{\eta^{\text{bouss}}} \mathbf{u} dz \quad (19.123)$$

is the depth integrated horizontal velocity and  $\rho_0$  is the Boussinesq reference density. Discuss the corresponding equation for the evolution of global mean sea level in this Boussinesq ocean. What is missing?

#### EXERCISE 19.5: TO MELT OR TO EXPAND?

Following Trenberth (2009) and Griffies and Greatbatch (2012), we ask where is heating most effective at raising global sea level? Is it more effective when used to melt land ice and allow the melt water to enter the ocean? Or should the heating go directly into warming the ocean, thus impacting a thermosteric sea level rise? The answer to this question is determined by fundamental properties of seawater. Discuss the answer.

#### EXERCISE 19.6: SEA LEVEL TENDENCY AND CONSERVATIVE TRACERS

Consider an arbitrary conservative tracer,  $C$ . The time tendency for the total amount of tracer in the global ocean is given by

$$\frac{d}{dt} \int_{\mathcal{R}} \rho C dV = \int_{\partial \mathcal{R}} Q_C dA, \quad (19.124)$$

where  $Q_C$  is the boundary flux of the tracer. Show that we can write the global mean sea level tendency as

$$\frac{d\bar{\eta}}{dt} = \frac{1}{\langle \rho C \rangle} \left[ \overline{Q_C} - \frac{V}{A} \frac{d\langle \rho C \rangle}{dt} \right], \quad (19.125)$$

which expresses the tendency in terms of changes in the density weighted concentration of the conservative tracer. Discuss this equation and contrast it to equation (19.93) based on the mass budget.

#### EXERCISE 19.7: THE VARIETIES OF LOCAL SEA LEVEL TENDENCIES

In this exercise we follow [Griffies and Greatbatch \(2012\)](#) to derive some equations for the time tendency of the sea level.

- (a) Starting from the seawater mass [continuity equation](#)

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}, \quad (19.126)$$

and using the surface and bottom [kinematic boundary conditions](#), derive the following kinematic sea level equation

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho(\eta)} - \nabla \cdot \mathbf{U} - \int_{\eta_b}^{\eta} \frac{1}{\rho} \frac{D\rho}{Dt} dz, \quad (19.127)$$

where we introduced the depth integrated horizontal velocity,

$$\mathbf{U} = \int_{\eta_b}^{\eta} \mathbf{u} dz. \quad (19.128)$$

Equation (19.127) decomposes the sea level tendency into a contribution from boundary mass fluxes, the convergence of the depth integrated flow, and the depth integrated non-Boussinesq steric effect. Note that for the Boussinesq ocean, the non-Boussinesq steric effect is missing ([Griffies and Greatbatch \(2012\)](#) offered a thorough study of this term).

- (b) Consider the mass budget for a column of fluid

$$\frac{\partial}{\partial t} \int_{\eta_b}^{\eta} \rho dz = Q_m - \nabla \cdot \mathbf{U}^\rho, \quad (19.129)$$

where we introduced the depth integrated density-weighted horizontal velocity,

$$\mathbf{U}^\rho = \int_{\eta_b}^{\eta} \rho \mathbf{u} dz. \quad (19.130)$$

Using these definitions, derive the following kinematic sea level equation

$$\frac{\partial \eta}{\partial t} = \frac{Q_m - \nabla \cdot \mathbf{U}^\rho}{\bar{\rho}^z} - \frac{H + \eta}{\bar{\rho}^z} \frac{\partial \bar{\rho}^z}{\partial t}, \quad (19.131)$$

where we introduced the depth averaged density,

$$\bar{\rho}^z = \frac{1}{H + \eta} \int_{\eta_b}^{\eta} \rho dz. \quad (19.132)$$

Equation (19.131) is directly related to equation (19.127), though each term is slightly different.

- (c) Consider an approximate hydrostatic fluid, in which the ocean bottom pressure,  $p_b$ , is given by

$$p_b - p_a = g \int_{\eta_b}^{\eta} \rho dz, \quad (19.133)$$

where  $p_a$  is the applied surface pressure. Use this equation, along with equation (19.131),

to derive the following expression for the time tendency of the sea level

$$\frac{\partial \eta}{\partial t} = \frac{g^{-1} \partial_t(p_b - p_a)}{\bar{\rho}^z} - \frac{H + \eta}{\bar{\rho}^z} \frac{\partial \bar{\rho}^z}{\partial t}. \quad (19.134)$$

- (d) Consider again the approximate hydrostatic fluid that satisfies equation (19.133). Derive the following sea level equation, which is a slight modification to equation (19.134)

$$\frac{\partial \eta}{\partial t} = \frac{g^{-1} \partial_t(p_b - p_a)}{\rho(\eta)} - \frac{1}{\rho(\eta)} \int_{\eta_b}^{\eta} \frac{\partial \rho}{\partial t} dz \quad (19.135)$$

- (e) Discuss why the global area mean of the local steric sealevel change

$$A^{-1} \int \left[ \frac{\partial \eta}{\partial t} \right]_{\text{steric}} dA = -A^{-1} \int \left[ \frac{1}{\rho(\eta)} \int_{\eta_b}^{\eta} \frac{\partial \rho}{\partial t} dz \right] dA \quad (19.136)$$

does not equal to the global steric change studied in Section 19.7.4,

$$A^{-1} \int \left[ \frac{\partial \eta}{\partial t} \right]^{\text{steric}} dA \neq -\frac{V}{A \langle \rho \rangle} \frac{d \langle \rho \rangle}{dt}. \quad (19.137)$$





## Chapter 20

# WATER MASS TRANSFORMATION THEORY

In ocean physics, a water mass refers to a region of seawater characterized by a suite of physical properties. Water masses often originate through extremely large air-sea and ice-sea buoyancy fluxes at the high latitudes that form waters such as the Antarctic Bottom Water and North Atlantic Deep Water. As these waters enter the ocean interior they are transported over basin scales while they are eroded or transformed by irreversible mixing or sources. Water masses and their properties offer a conceptual means to partition or bin the fluid into distinct classes whose origin, movement, and transformation can be measured, modeled, and studied. Scalar properties generally used to classify water masses are simpler to measure than vector properties such as velocity. Hence, a water mass perspective offers the means to infer ocean circulation within the space of ocean properties without directly measuring vector fields. Such circulation inferences have been used since the early days of ocean circulation theory (e.g., [Sverdrup \(1947\)](#)).

In this chapter we develop the mathematical and physical basis for water mass transformation theory, which examines the budgets for fluid mass and tracer mass within layers or classes defined by properties such as buoyancy, Conservative Temperature, salinity, or biogeochemical tracers. Water mass transformation theory originates from the work of [Wal in \(1977\)](#) and [Wal in \(1982\)](#), and it has found many advances in both formalism and application since then (see [Groeskamp et al. \(2019\)](#) for a review). A water mass transformation perspective on circulation offers a valuable lens for describing and understanding how processes affect budgets and circulation as viewed in property space. This kinematic lens is distinct from the Eulerian and Lagrangian kinematics considered elsewhere in this book. It has proven to be particularly useful for examining questions where the irreversible transformation of properties plays a central role.

### CHAPTER GUIDE

We make use of vector calculus (VOLUME 1), and mass and tracer budget kinematics (VOLUME 2). Additionally, we use basic results from generalized vertical coordinates in Part IV of this volume. Finally, and perhaps most importantly, this chapter relies on our study of advection and diffusion from Chapters 15 and 16; aspects of parameterized tracer transport and mixing studied in Chapter 18; and buoyancy budgets in Chapter 19, including boundary conditions.

<b>20.1 Conceptual framework . . . . .</b>	<b>691</b>
20.1.1 The distinct lens of water mass configuration space . . . . .	691
20.1.2 Transformation and formation . . . . .	692

---

<b>20.2</b>	<b>Buoyancy transformation and formation . . . . .</b>	<b>693</b>
20.2.1	A three-layer thought experiment . . . . .	693
20.2.2	How processes lead to transformations . . . . .	694
<b>20.3</b>	<b>Mathematical tools . . . . .</b>	<b>696</b>
20.3.1	Fluid mass in an infinitesimal cylinder . . . . .	696
20.3.2	Fluid mass within a finite region . . . . .	697
20.3.3	Fluid mass distribution function . . . . .	697
20.3.4	Example regions . . . . .	698
20.3.5	Integrals of arbitrary functions over regions . . . . .	700
20.3.6	Moments of $\lambda$ . . . . .	701
20.3.7	Internal and external $\lambda$ -moments . . . . .	701
20.3.8	Further study . . . . .	702
<b>20.4</b>	<b>Water mass transformation across an internal <math>\lambda</math>-surface . . . . .</b>	<b>702</b>
20.4.1	Dia-surface flux and interior transformation . . . . .	702
20.4.2	Dia-surface flux from diffusion and steady state balances . . . . .	704
20.4.3	Transformation as the derivative of an integral . . . . .	704
20.4.4	Kinematic and process methods of water mass transformation . . . . .	706
20.4.5	General properties of interior transformation due to diffusion . . . . .	707
20.4.6	Interior versus boundary water mass transformation . . . . .	709
<b>20.5</b>	<b>Budget for fluid mass in a <math>\Delta\lambda</math>-layer . . . . .</b>	<b>710</b>
20.5.1	Mass transport crossing interior open boundaries . . . . .	710
20.5.2	Mass transport crossing the ocean surface . . . . .	710
20.5.3	Mass budget and water mass formation . . . . .	711
20.5.4	Mass budget and formation in terms of distribution functions . . . . .	711
<b>20.6</b>	<b>Mass budget for <math>\lambda</math> within a <math>\lambda_\infty</math>-region . . . . .</b>	<b>712</b>
20.6.1	Processes affecting the mass of $\lambda$ -stuff . . . . .	712
20.6.2	Summary of the $\lambda$ budget . . . . .	715
20.6.3	Processes leading to water mass transformation . . . . .	715
20.6.4	Water mass formation . . . . .	717
20.6.5	Interior water mass transformation from diffusion . . . . .	717
20.6.6	Surface water mass transformation . . . . .	719
<b>20.7</b>	<b>Buoyancy water mass transformation . . . . .</b>	<b>720</b>
20.7.1	Material time changes to $S$ and $\Theta$ . . . . .	720
20.7.2	Interior buoyancy water mass transformation . . . . .	721
20.7.3	Surface non-advective flux for $S$ and $\Theta$ . . . . .	721
20.7.4	Surface buoyancy water mass transformation . . . . .	721
<b>20.8</b>	<b>Tracer water mass transformation within <math>\gamma</math>-layers . . . . .</b>	<b>723</b>
20.8.1	General form of the mass budget . . . . .	723
20.8.2	Tracer processes . . . . .	724
20.8.3	Transport across an interior layer interface . . . . .	725
20.8.4	Transport across interior and surface boundaries . . . . .	726
20.8.5	The layer tracer budget . . . . .	727
20.8.6	Further study . . . . .	727
<b>20.9</b>	<b>Regions bounded by a tracer contour/surface . . . . .</b>	<b>727</b>
20.9.1	Closed region bounded by a tracer surface/contour . . . . .	727
20.9.2	Region with $C \geq \tilde{C}$ . . . . .	729
20.9.3	Comments and further study . . . . .	731
<b>20.10</b>	<b>Exercises . . . . .</b>	<b>731</b>

---

## 20.1 Conceptual framework

Water mass transformation theory is a mathematical formalism that supports the study of budgets for fluid mass and tracer mass within layers or classes defined by properties such as Archimedean buoyancy (see VOLUME 2; shortened to **buoyancy** here), **Conservative Temperature**, **salinity**, or biogeochemical tracers. The theory is concerned with how processes affect the evolution of fluid within property space and in the characterization of circulation inferred from this evolution.

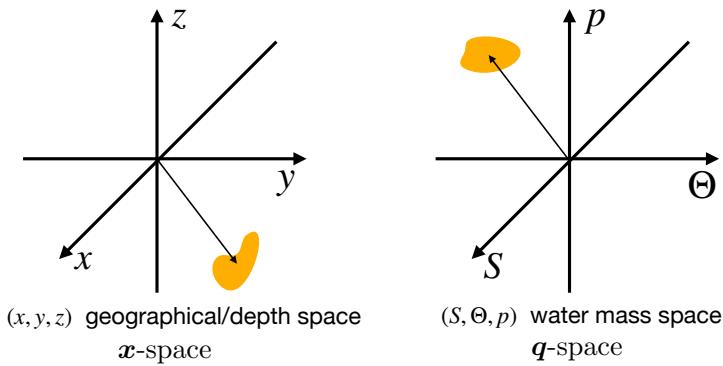


FIGURE 20.1: Left panel: a fluid element is positioned in geographical/depth  $\mathbf{x}$ -space according to its horizontal ( $x, y$ ) (longitude, latitude) position and its vertical geopotential,  $z$ . Right panel: the same fluid element is positioned in a particular water mass configuration space ( $\mathbf{q}$ -space), here defined by  $\mathbf{q} = (S, \Theta, p)$  with salinity,  $S$ , Conservative Temperature,  $\Theta$ , and pressure,  $p$ . Mapping between the two spaces is generally not 1-to-1. Namely, a point in  $\mathbf{q}$ -space can be occupied by more than one point in  $\mathbf{x}$ -space. The reason is that more than a single point in  $\mathbf{x}$ -space can have the same values for  $(S, \Theta, p)$ . Although the coordinate axes in  $\mathbf{q}$ -space are depicted here as mutually orthogonal, there is no objective means to determine angles in  $\mathbf{q}$ -space since it contains no metric. Rather,  $\mathbf{q}$ -space, just like thermodynamic state space (see the chapters on thermodynamics in VOLUME 2), forms a **differentiable manifold** that has no metric.

### 20.1.1 The distinct lens of water mass configuration space

Water mass configuration space (denoted by  $\mathbf{q}$ -space) is the space we work within to study water mass transformations.<sup>1</sup> This space has some or all of its coordinates set by properties other than geographic/depth coordinates. For example, in Figure 20.1 we present the three-dimensional  $\mathbf{q}$ -space given by  $\mathbf{q} = (S, \Theta, p)$ , where the position (or bin) for a fluid element is determined by its Conservative Temperature,  $\Theta$ , salinity,  $S$ , and pressure,  $p$ . Operationally, we fill  $\mathbf{q}$ -space by forming histograms that result in a  $\mathbf{q}$ -space distribution of the fluid properties. For example, a one-dimensional  $\mathbf{q}$ -space results from binning the ocean according to potential density, whereas retaining latitudinal information along with potential density renders a two-dimensional  $\mathbf{q}$ -space. Typically  $\mathbf{q}$ -space has three or fewer dimensions, given the three dimensionality of  $\mathbf{x}$ -space. There is no implied constraint that any of the  $\mathbf{q}$ -space coordinates are monotonic with respect to  $\mathbf{x}$ -space. Indeed, there is no presumption that points in  $\mathbf{q}$ -space maintain a 1-to-1 relation to points in  $\mathbf{x}$ -space. For example, many points in  $\mathbf{x}$ -space may fall into a single point (or bin) within  $\mathbf{q}$ -space.

The lack of 1-to-1 mapping between water mass configuration space and geographic/depth space is a fundamental kinematic distinction from the 1-to-1 relation that holds between the

<sup>1</sup>We prefer the term “configuration space” over the alternative “phase space”, since phase space in Hamiltonian dynamics specifically refers to position and momentum coordinates (see VOLUME 1). In contrast, configuration space, as used in our discussion of water masses, can be determined by almost any property or geographic position.

Eulerian reference frame and Lagrangian reference frame used for describing fluid motion (see the kinematics chapters in VOLUME 2). The lack of a 1-to-1 relation can be frustrating since circulation viewed in  $\mathbf{q}$ -space generally has incomplete  $\mathbf{x}$ -space information.<sup>2</sup> Even so, abandoning the 1-to-1 relation can be liberating since working within  $\mathbf{q}$ -space offers a framework to infer  $\mathbf{q}$ -space circulation even without measuring velocity of the fluid in geographical space. Correspondingly, ocean circulation when viewed through a water mass lens can offer understanding that complements traditional Eulerian or Lagrangian views.

Water mass configuration space generally has no metric, particularly when none of the chosen coordinates are geographical (latitude or longitude) or depth. Hence, there is generally no notion of distance or angles between points in water mass configuration space. For example, what does it mean to be orthogonal in temperature-salinity space or when studying the density-binned distribution of seawater? The absence of a metric is something we have already seen when studying thermodynamic configuration space in VOLUME 1. Mathematically, we say that both thermodynamic configuration space and water mass configuration space are [differentiable manifolds](#). Even so, one commonly sees a point in thermodynamic space depicted on a diagram with orthogonal axes (e.g., pressure-volume diagrams), or a position in water mass configuration space similarly depicted with orthogonal axes as in Figure 20.1. This depiction is arbitrary since these spaces contain none of the necessary geometric structure required to define a metric tensor. Hence, we have no objective means to determine angles or orthogonality. Depictions of property space using orthogonal axes satisfy a subjective desire for geometric structure when in fact there is none afforded to the underlying differentiable manifold.

### 20.1.2 Transformation and formation

Force imbalances lead to the motion of a fluid element through  $\mathbf{x}$ -space as captured by Newton's laws of motion. Likewise, [transformation processes](#)<sup>3</sup> render motion in water mass configuration space ( $\mathbf{q}$ -space). That is, a fluid moves through  $\mathbf{q}$ -space if it is transformed in a manner that causes water to cross surfaces of constant  $\mathbf{q}$ . It follows that the convergence (local accumulation) or divergence (local depletion) of transformation leads to the [formation](#) of water mass classes, or its negative, being the destruction of water masses. Processes leading to transformation arise from mixing, solar radiation, and chemical reactions. Correspondingly, we measure zero motion along a  $\mathbf{q}$ -space coordinate axis when the property defining that axis remains materially unchanged. For example, adiabatic and isohaline processes such as linear waves can render nontrivial motion in geographical/depth space whereas they lead to no motion in  $(S, \Theta)$  space.

Given the generally non-local and non-invertible relation between  $\mathbf{x}$ -space and  $\mathbf{q}$ -space, we find that  $\mathbf{q}$ -space is an unnatural kinematic venue to study forces and stresses acting between spatially adjacent fluid elements. Hence, the study of momentum and vorticity dynamics is more suitably handled via Eulerian or Lagrangian kinematics. Where water mass configuration space shines is by revealing the  $\mathbf{q}$ -space circulation effects from processes that affect material changes to fluid properties. For example, a water mass perspective has found use in framing key questions of primary interest in the Anthropocene, such as ocean buoyancy and its transformation through interior and boundary mixing, ocean heat uptake and transport, the hydrological cycle, steric

<sup>2</sup>Auxiliary methods such as the water tagging method of [Groeskamp et al. \(2014\)](#) can be used to recover some geographical information.

<sup>3</sup>In many parts of this book the word “transformation” refers to mathematical coordinate transformations, thus referring to a change of reference or change of description. Here, “transformation” refers to a physical process acting to change a property of a fluid element, thus acting to change one or more of the  $\mathbf{q}$ -space coordinates of a fluid element.

sea level rise, and irreversible changes to biogeochemical properties (see [Groeskamp et al. \(2019\)](#) for a review).

## 20.2 Buoyancy transformation and formation

Archimedean buoyancy is a common property used to partition or bin geophysical fluids into classes, with Archimedean buoyancy often approximated by potential density. In this section we introduce the notions of [transformation](#) and [formation](#) when partitioning the ocean according to density ( $\gamma$ ) classes that locally measure buoyancy. The ideas presented here extend to any scalar property used to bin the ocean fluid.

Quantitatively, the transformation of fluid provides a measure of the mass per time that fluid moves across an isosurface of one of the  $\mathbf{q}$ -space coordinates. Hence, water mass transformation refers to the movement of water from one  $\mathbf{q}$ -space bin to another bin within the  $\mathbf{q}$ -space mass distribution. We here follow the convention whereby the transformation is positive if water moves to a larger value of a particular  $\mathbf{q}$ -space coordinate (e.g., higher density class), and negative if water moves to a smaller value for the  $\mathbf{q}$ -space coordinate (e.g., lower density class). Water mass [formation](#) refers to the difference in transformation across the surfaces bounding a layer, so that formation measures the change in seawater mass of the layer.

The formation is the layer integrated  $\mathbf{q}$ -space convergence of transformation, with both formation and transformation having dimensions of mass per time (or volume per time when considering a [Boussinesq ocean](#)). We typically measure transformation and formation using Sverdrup units:

$$1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1} \quad \text{volume-Sverdrup} \tag{20.1a}$$

$$1 \text{ Sv} = 10^9 \text{ kg s}^{-1} \quad \text{mass-Sverdrup.} \tag{20.1b}$$

The mass-Sverdrup can be routinely used for Boussinesq fluids merely by multiplying the volume-Sverdrup by the constant Boussinesq reference density,  $\rho_0$ .

### 20.2.1 A three-layer thought experiment

To illustrate the concepts of transformation and formation, bin the World Ocean into density ( $\gamma$ ) classes so that  $\mathbf{q}$ -space is just one dimensional. In performing the binning, we lose all information about latitude, longitude, and depth, while retaining information provided by  $\gamma$ -classes. Furthermore, partition the World Ocean into just three density classes (layers or bins) that are bounded by four density interfaces:

$$\text{light density layer} = [\gamma - \delta\gamma/2, \gamma + \delta\gamma/2] \tag{20.2a}$$

$$\text{middle density layer} = [\gamma + \delta\gamma/2, \gamma + 3\delta\gamma/2] \tag{20.2b}$$

$$\text{heavy density layer} = [\gamma + 3\delta\gamma/2, \gamma + 5\delta\gamma/2], \tag{20.2c}$$

where  $\delta\gamma > 0$  is the size of the density bins. Figure 20.2 depicts a sample mass distribution; i.e., the mass census for seawater binned into these three density layers.<sup>4</sup> Now consider a physical, chemical, or biological process that results in water leaving the middle density layer

<sup>4</sup>A realistic ocean is confronted with boundary forcing that causes the maximum and minimum density to be a function of time. It is thus common to fix the lower density limit to be well below the lightest water in the ocean and the upper density limit well above the maximum density, thus ensuring that all seawater is contained within the chosen binning. We introduce such “infinity” bounds in Section 20.3.3.

and entering both the light layer and the heavy layer.<sup>5</sup> Let  $G(\sigma)$  measure the mass per time that water crosses the density interface  $\gamma = \sigma$ ; i.e.,  $G(\sigma)$  is the transformation. This particular thought experiment has the following transformations across the various layer interfaces

$$G(\sigma) = \begin{cases} 0 & \sigma = \gamma - \delta\gamma/2 \quad \text{closed boundary} \\ < 0 & \sigma = \gamma + \delta\gamma/2 \quad \text{mass moves to light density from middle density} \\ > 0 & \sigma = \gamma + 3\delta\gamma/2 \quad \text{mass moves from middle density to heavy density} \\ 0 & \sigma = \gamma + 5\delta\gamma/2 \quad \text{closed boundary.} \end{cases} \quad (20.3)$$

The difference in the transformation across the interface boundaries of a particular layer determines the formation/destruction of water in that layer. Here, the convergence of water into the light and heavy layers means that there is a positive formation of water in these two density layers. In contrast, the divergence of water from the middle density layer means there is a negative formation or a destruction of some of its water. We write these layer formations mathematically as follows

$$\text{light-formation} = -[G(\gamma + \delta\gamma/2) - G(\gamma - \delta\gamma/2)] > 0 \quad (20.4a)$$

$$\text{middle-formation} = -[G(\gamma + 3\delta\gamma/2) - G(\gamma + \delta\gamma/2)] < 0 \quad (20.4b)$$

$$\text{heavy-formation} = -[G(\gamma + 5\delta\gamma/2) - G(\gamma + 3\delta\gamma/2)] > 0. \quad (20.4c)$$

The minus sign out front emphasizes that the formation is the layer integrated convergence of the transformation.

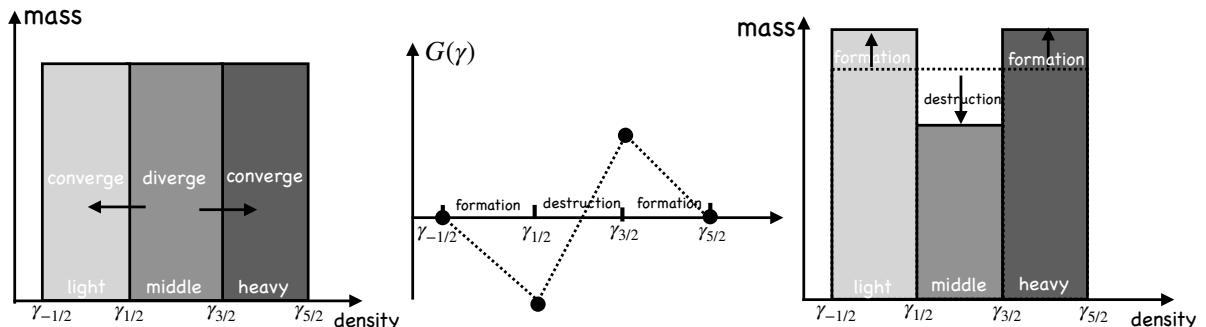


FIGURE 20.2: A sample mass distribution of the ocean binned into three density layers (light, middle, heavy) bounded by four density interfaces:  $\gamma_{n/2} = \gamma + n\delta\gamma/2$  for  $n = -1, 1, 3, 5$ . The left panel depicts an ocean state with equal mass in each layer. Some process is then imagined to cause water to diverge from the middle layer and converge to the light and heavy layers. The right panel shows the mass distribution after the water has moved, so that the middle layer has experienced a negative formation (i.e., net loss or destruction of water mass) whereas the light and heavy layers have experienced a positive formation (i.e., net mass gain). The middle panel depicts the transformation,  $G$ , which measures the mass per time moving across the layer boundaries. By convention,  $G > 0$  for water moving into a heavier density layer and  $G < 0$  for water entering a lighter density layer. The addition of more layers refines the picture (e.g., by smoothing the plot of  $G$ ) but it does not modify the basic ideas illustrated in this thought experiment.

### 20.2.2 How processes lead to transformations

Here we outline a few of the processes that affect circulation within water mass space, again focusing on buoyancy yet with easy generalization to any other property used to define a water

<sup>5</sup>Since the layer is the result of binning over the World Ocean, there can be some regions within a bin that experience processes that decrease the density, whereas other regions where density increases. In this manner, some water within the bin moves to a denser bin whereas other water moves to a lighter bin.

mass coordinate.

### Interior transformation from mixing

Mixing moves water across layer boundaries in  $q$ -space, with  $q$ -space coordinates/properties materially modified in the presence of mixing (so long as there are spatial gradients in the property). For example, in our study of tracer kinematics in VOLUME 2, we consider how mixing causes tracers to move between fluid elements even as mixing does not alter the net mass of fluid elements, with the [barycentric velocity](#) playing a central role in that theory. Hence, in the presence of mixing, seawater fluid elements retain a fixed mass, and yet the mass is redistributed among layers defined by property isosurfaces since the isosurfaces move in the presence of mixing.

### Surface mass fluxes

Precipitation, runoff, and evaporation alter the mass of the ocean. In turn, the layers where precipitation, runoff, and evaporation occur (layers that outcrop) have their mass altered through the surface mass fluxes. Additionally, if properties of the mass flux (e.g., temperature, salinity) differ from that of the ocean layer that it enters/leaves, then the properties of the ocean layer are modified upon mixing the ambient seawater with the water crossing the ocean surface boundary.

### Surface and bottom boundary transformation

Surfaces of constant buoyancy that outcrop at the ocean surface or incrop at the ocean bottom are exposed to boundary fluxes that generally modify the buoyancy of the fluid within a layer. This modification causes the layer boundaries to move so that the mass distributed within the layers can be modified if the mass moves with a velocity distinct from the buoyancy surface. A particularly striking example occurs in the upper ocean boundary layer where surface boundary fluxes lead to the seasonal migration of density outcrops. The associated lateral movement of density layers causes water to entrain and detrain from a layer since the layer boundaries generally have a velocity distinct from fluid elements. In so doing, the seasonal cycle of surface buoyancy forcing can inflate or deflate a buoyancy layer by moving the layer boundaries so that the layer entrains or detrains mass. We return to this example when discussing Figure 20.14 when studying transformations due to surface boundary buoyancy fluxes.

Penetrative shortwave radiation provides another means to modify water masses, with penetrative radiation a function of the optical properties of the fluid. This radiation provides a heat source that can penetrate into the upper tens of meters in the ocean, and can impact on the temperature and density structure of the ocean layers affected by radiation.

Layers that intersect the ocean bottom are exposed to geothermal heating as well as enhanced mixing from turbulent bottom boundary layer processes. Each of these processes affects a transformation of the buoyancy, thus modifying the buoyancy layer interfaces and mass distribution within the layers.

### Interior sources and sinks

When studying water masses defined by biogeochemical tracers (e.g., carbon, oxygen, nutrients), there are a variety of chemical reactions and biological processes that act to modify these properties. These processes generally cannot be represented mathematically as the convergence

of a flux, and so they are commonly referred to as a **non-conservative process**. They appear mathematically as a source/sink term in the tracer equation.

## 20.3 Mathematical tools

In this section we develop a suite of mathematical tools to quantify the conceptual ideas presented in Sections 20.1 and 20.2. In particular, we develop a formalism for integrating properties within a region bounded by isosurfaces of a scalar field,  $\lambda = \lambda(\mathbf{x}, t)$ . The formulation is given from a geometric perspective afforded by  $\mathbf{x}$ -space. It is also given from the complementary distributional perspective afforded by binning seawater mass according to  $\lambda$ -classes that define a one-dimensional  $\mathbf{q}$ -space (extensions to higher  $\mathbf{q}$ -space dimensions are straightforward).

In Section 20.2 we considered  $\lambda$  to be the buoyancy field,  $\lambda = \gamma$ , whereas here we assume it is a generic scalar field,  $\lambda(\mathbf{x}, t)$ . In contrast to the case of a **generalized vertical coordinate** (Part IV),  $\lambda$ -isosurfaces are free to overturn or even to be situated in spatially disconnected regions. This freedom is motivated by the behavior of most ocean scalar properties, which commonly exhibit vertically unstratified or negatively stratified profiles, particularly within boundary layers. As mentioned in Section 20.1.1, this freedom comes at the cost of losing the 1-to-1 relation between  $\mathbf{x}$ -space and  $\mathbf{q}$ -space.

### 20.3.1 Fluid mass in an infinitesimal cylinder

Consider the calculation of fluid mass within an infinitesimal  $\lambda$ -layer bounded by two isosurfaces,  $[\lambda - \delta\lambda/2, \lambda + \delta\lambda/2]$ , as depicted in Figure 20.3. The mass within a tiny cylinder extending from one interface to the other is given by the fluid density,  $\rho$ , multiplied by the volume of the cylinder,<sup>6</sup>

$$\delta M = \rho \delta V = \rho \delta h \delta S, \quad (20.5)$$

where  $\delta S$  is the cross-sectional area element and  $\delta h$  is the layer thickness. The geometric thickness,  $\delta h$ , is related to the differential  $\lambda$ -increment separating the two interfaces according to

$$\delta\lambda = \nabla\lambda \cdot \delta\mathbf{x} = |\nabla\lambda| \hat{\mathbf{n}} \cdot \delta\mathbf{x} = |\nabla\lambda| \delta h \quad \text{with} \quad \hat{\mathbf{n}} = \nabla\lambda |\nabla\lambda|^{-1}, \quad (20.6)$$

where  $\delta\mathbf{x}$  is a position vector connecting points on the two interfaces. We thus see that the layer thickness is given by

$$\delta h = \frac{\delta\lambda}{|\nabla\lambda|}, \quad (20.7)$$

which connects a geometric property of the layer,  $\delta h > 0$ , to the  $\lambda$ -increment,  $\delta\lambda > 0$ . For a given  $\lambda$ -increment, the layer thickness is smaller with more tightly packed  $\lambda$ -isosurfaces as reflected by a larger  $|\nabla\lambda|$ . Furthermore, the geometric thickness is oriented according to the normal direction,  $\hat{\mathbf{n}}$ , so that  $\delta h$  measures the distance between the  $\lambda$ -interfaces in the direction of the normal. It follows that the fluid mass within the cylinder is given by

$$\delta M = \rho \delta V = \rho \delta h \delta S = \frac{\rho \delta\lambda \delta S}{|\nabla\lambda|}. \quad (20.8)$$

---

<sup>6</sup>Recall our notational convention is as follows:  $\delta$  refers to an infinitesimal increment of a property measured within the fluid whereas  $d$  is a differential increment used for computing integrals.

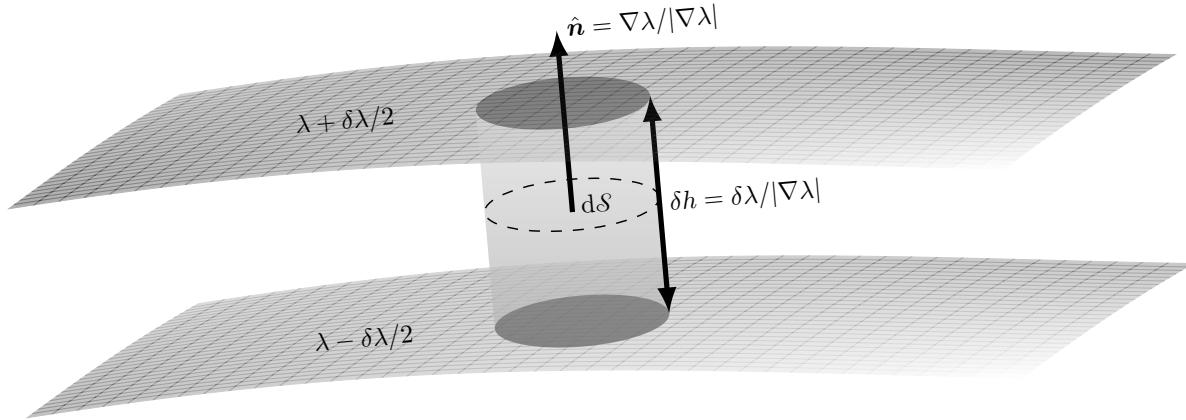


FIGURE 20.3: Depicting an infinitesimally thin  $\lambda$ -layer bounded by two interfaces  $[\lambda - \delta\lambda/2, \lambda + \delta\lambda/2]$ , with the  $\lambda$ -increment,  $\delta\lambda > 0$ . The cylinder extends between the two iso-surfaces and it has thickness,  $\delta h = \delta\lambda/|\nabla\lambda|$ , and cross-sectional area,  $\delta S$ . The cylinder is oriented according to the normal direction,  $\hat{n} = |\nabla\lambda|^{-1} \nabla\lambda$ , which points in the direction of increasing  $\lambda$ . We assume  $|\nabla\lambda| \neq 0$ , as required to define a normal direction. Indeed, if  $\nabla\lambda = 0$  then we could not perform a binning according to  $\lambda$  classes, so the  $|\nabla\lambda| \neq 0$  assumption is basic to the use of the scalar field,  $\lambda$ , for water mass analysis.

### 20.3.2 Fluid mass within a finite region

Making use of the infinitesimal cylinder mass (20.8) allows us to write the mass of fluid within the  $\lambda$ -region,  $\lambda_1 \leq \lambda \leq \lambda_2$ ,

$$M(\lambda_1, \lambda_2) \equiv \int_{\Omega(\lambda_1 \leq \lambda \leq \lambda_2)} dM = \int_{\lambda_1}^{\lambda_2} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla\lambda|} \right] d\lambda. \quad (20.9)$$

In this equation,  $\Omega(\lambda_1 \leq \lambda \leq \lambda_2)$  is the region in space bounded by the  $\lambda_1$ -interface and  $\lambda_2$ -interface, and  $\partial\Omega(\lambda)$  is the surface defined by a  $\lambda$ -isosurface. The  $\partial\Omega(\lambda)$  integral is taken over the area of the  $\lambda$ -isosurface, which is then integrated over the range,  $\lambda_1 \leq \lambda \leq \lambda_2$ , to accumulate the layer mass.

### 20.3.3 Fluid mass distribution function

The region bounded by the layer interfaces can have any shape in  $\mathbf{x}$ -space and can even be spatially disconnected. This complexity motivates us to introduce the **mass distribution function** in water mass space by integrating the mass over the surface,  $\partial\Omega(\lambda)$

$$m(\lambda) \equiv \frac{dM}{d\lambda} = \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla\lambda|}. \quad (20.10)$$

Evidently, the mass distribution function is the mass density within  $\lambda$  space; i.e., the mass per  $\lambda$ . As defined, the mass distribution function serves the analog for  $\mathbf{q}$ -space that the *in situ* density,  $\rho(\mathbf{x})$ , does for  $\mathbf{x}$ -space:

$$dM = m(\lambda) d\lambda = \text{fluid mass within the infinitesimal } \lambda\text{-layer } [\lambda - d\lambda/2, \lambda + d\lambda/2], \quad (20.11)$$

with an illustration given by Figure 20.4. In practice, the mass distribution function is constructed through binning the ocean mass according to  $\lambda$ -classes, thus producing a histogram that approximates  $m(\lambda)$ . The mass distribution function then provides the means to compute

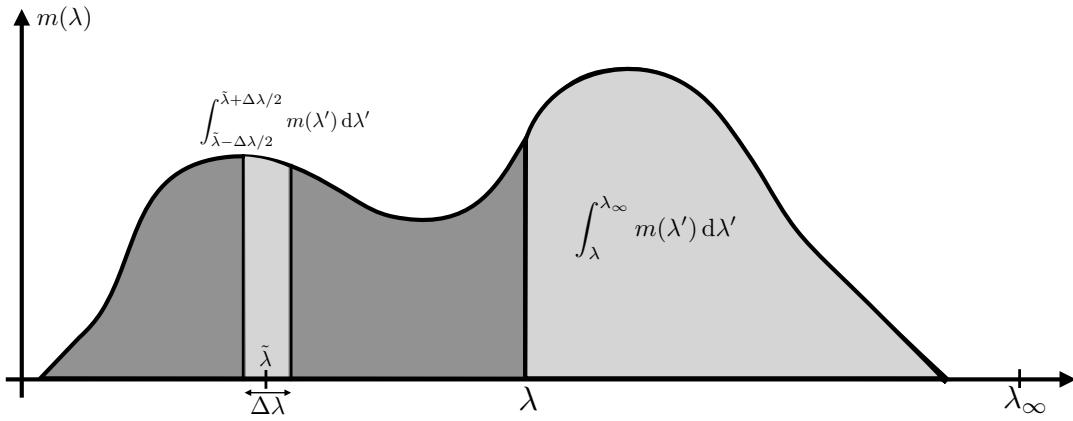


FIGURE 20.4: An example mass distribution function,  $m(\lambda) = dM/d\lambda$ , which measures the mass of fluid per  $\lambda$ -increment. Integration over a finite  $\lambda$ -region measures the fluid mass within that region. For example, the mass within a  $\Delta\lambda$ -layer is given by  $M(\tilde{\lambda} - \Delta\lambda/2, \tilde{\lambda} + \Delta\lambda/2) = \int_{\tilde{\lambda} - \Delta\lambda/2}^{\tilde{\lambda} + \Delta\lambda/2} m(\lambda') d\lambda'$  whereas the mass within the  $\lambda_\infty$ -region is  $M(\lambda, \lambda_\infty) = \int_\lambda^{\lambda_\infty} m(\lambda') d\lambda'$ , where we assume that  $\lambda_\infty$  is an arbitrary fixed value that is larger than any  $\lambda$  realized within the ocean.

the fluid mass within a finite  $\lambda$ -region as in equation (20.9)

$$M(\lambda_1, \lambda_2) \equiv \int_{\Omega(\lambda_1 \leq \lambda \leq \lambda_2)} dM = \int_{\lambda_1}^{\lambda_2} m(\lambda) d\lambda. \quad (20.12)$$

Extensions to higher dimensional  $\mathbf{q}$ -space are straightforward, in which we consider,  $m(\mathbf{q})$ .

#### 20.3.4 Example regions

To help ground the previous expressions for mass, we here consider some example regions commonly considered in water mass analysis.

**$\Delta\lambda$ -layer defined by**  $[\lambda_1, \lambda_2] = [\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$

A  $\Delta\lambda$ -layer is defined with the bounding interface values

$$\lambda_1 = \lambda - \Delta\lambda/2 \text{ and } \lambda_2 = \lambda + \Delta\lambda/2, \quad (20.13)$$

for some finite difference increment  $\Delta\lambda > 0$ . In this case the layer mass is

$$M(\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2) = \int_{\lambda - \Delta\lambda/2}^{\lambda + \Delta\lambda/2} \left[ \int_{\partial\Omega(\lambda')} \frac{\rho d\mathcal{S}}{|\nabla \lambda'|} \right] d\lambda' = \int_{\lambda - \Delta\lambda/2}^{\lambda + \Delta\lambda/2} m(\lambda') d\lambda'. \quad (20.14)$$

**$\lambda_\infty$ -region defined by**  $[\lambda_1, \lambda_2] = [\lambda, \lambda_\infty]$

A  $\lambda_\infty$ -region is defined with

$$\lambda_1 = \lambda \text{ and } \lambda_2 = \lambda_\infty, \quad (20.15)$$

where  $\lambda_\infty$  is an arbitrary fixed constant that is larger than any value of  $\lambda$  realized in the ocean. The region mass is thus given by

$$M(\lambda, \lambda_\infty) = \int_{\Omega(\lambda \leq \lambda_\infty)} dM = \int_\lambda^{\lambda_\infty} \left[ \int_{\partial\Omega(\lambda')} \frac{\rho d\mathcal{S}}{|\nabla \lambda'|} \right] d\lambda' = \int_\lambda^{\lambda_\infty} m(\lambda') d\lambda'. \quad (20.16)$$

An example  $\lambda_\infty$ -region is shown in Figure 20.5. The  $\lambda_\infty$ -region as so defined provides an expression for the differential mass increment

$$M(\lambda, \lambda_\infty) = \int_\lambda^{\lambda_\infty} m(\lambda') d\lambda' \implies dM(\lambda, \lambda_\infty) = -m(\lambda) d\lambda, \quad (20.17)$$

which follows since  $\lambda_\infty$  is a constant.

The value of the fixed constant,  $\lambda_\infty$ , is arbitrarily large, indeed it could be infinite. We can set it to an arbitrarily large and constant value since there is no contribution to the integral from regions with  $\lambda'$  outside the range realized within the ocean, merely since there is no ocean mass in that region. As an example, let  $\lambda = \Theta$ , the Conservative Temperature, in which the region  $\Theta \leq \Theta_\infty$  encompasses the ocean region where the Conservative Temperature is larger (warmer) than  $\Theta$ .

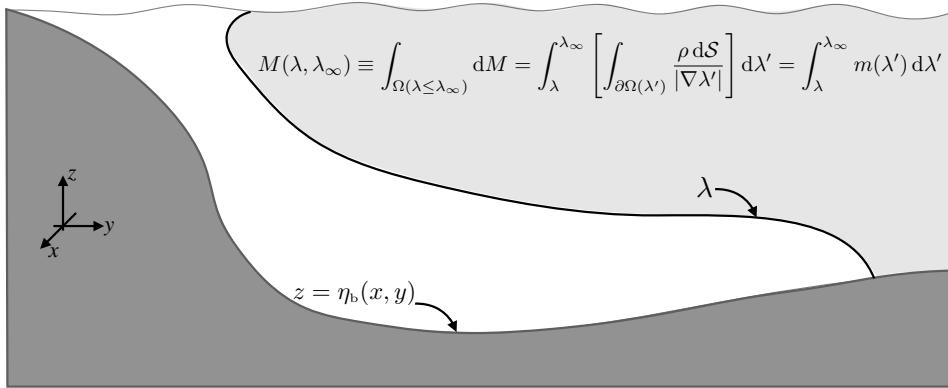


FIGURE 20.5: Depicting the mass of fluid within a  $\lambda_\infty$ -region, where  $\lambda \leq \lambda_\infty$  with  $\lambda_\infty$  an arbitrary constant that is larger than any value of  $\lambda$  in the ocean domain. An example of such a region is for  $\lambda = \Theta$ , whereby warmer waters are typically shallower and towards the equator.

#### $\lambda_{-\infty}$ -region defined by $[\lambda_1, \lambda_2] = [\lambda_{-\infty}, \lambda]$

A  $\lambda_{-\infty}$ -region is defined with

$$\lambda_1 = \lambda_{-\infty} \text{ and } \lambda = \lambda, \quad (20.18)$$

where  $\lambda_{-\infty}$  is an arbitrary constant that is smaller than any value of  $\lambda$  realized in the ocean. The region mass is thus given by

$$M(\lambda_{-\infty}, \lambda) \equiv \int_{\Omega(\lambda_{-\infty} \leq \lambda)} dM = \int_{\lambda_{-\infty}}^\lambda \left[ \int_{\partial\Omega(\lambda')} \frac{\rho dS}{|\nabla\lambda'|} \right] d\lambda' = \int_{\lambda_{-\infty}}^\lambda m(\lambda') d\lambda'. \quad (20.19)$$

This mass is the complement of that contained in the  $\lambda_\infty$ -region. The  $\lambda_{-\infty}$ -region mass implies a corresponding differential mass increment via

$$M(\lambda_{-\infty}, \lambda) = \int_{\lambda_{-\infty}}^\lambda m(\lambda') d\lambda' \implies dM(\lambda_{-\infty}, \lambda) = m(\lambda) d\lambda. \quad (20.20)$$

#### $\lambda_{\pm\infty}$ -region defined by $[\lambda_1, \lambda_2] = [\lambda_{-\infty}, \lambda_\infty]$

The full ocean is contained in the  $\lambda_{\pm\infty}$ -region

$$\lambda_1 = \lambda_{-\infty} \text{ and } \lambda_2 = \lambda_\infty, \quad (20.21)$$

so that the full ocean mass is written

$$M(\lambda_{-\infty}, \lambda_{\infty}) = \int_{\Omega(\lambda_{-\infty} \leq \lambda \leq \lambda_{\infty})} dM = \int_{\lambda_{-\infty}}^{\lambda_{\infty}} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla \lambda|} \right] d\lambda = \int_{\lambda_{-\infty}}^{\lambda_{\infty}} m(\lambda') d\lambda'. \quad (20.22)$$

### Difference of mass between two $\lambda_{\infty}$ -regions

The difference in mass between two  $\lambda_{\infty}$ -regions is given by

$$M(\lambda_1, \lambda_{\infty}) - M(\lambda_2, \lambda_{\infty}) = \int_{\lambda_1}^{\lambda_{\infty}} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla \lambda|} \right] d\lambda - \int_{\lambda_2}^{\lambda_{\infty}} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla \lambda|} \right] d\lambda. \quad (20.23)$$

The arbitrary constant,  $\lambda_{\infty}$ , drops out when taking the difference so that we are left with the mass within the intersection of the two regions

$$M(\lambda_1, \lambda_2) = M(\lambda_1, \lambda_{\infty}) - M(\lambda_2, \lambda_{\infty}) = \int_{\lambda_1}^{\lambda_2} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla \lambda|} \right] d\lambda = \int_{\lambda_1}^{\lambda_2} m(\lambda') d\lambda'. \quad (20.24)$$

### 20.3.5 Integrals of arbitrary functions over regions

We can extend the above formalism to integrals of an arbitrary function,  $F(\mathbf{x}, t)$ , over a region defined by the  $\lambda_1$  and  $\lambda_2$  interfaces

$$\mathcal{I}_F(\lambda_1, \lambda_2) \equiv \int_{\Omega(\lambda_1 \leq \lambda \leq \lambda_2)} F dM = \int_{\lambda_1}^{\lambda_2} \left[ \int_{\partial\Omega(\lambda)} \frac{F \rho dS}{|\nabla \lambda|} \right] d\lambda. \quad (20.25)$$

Performing the area integral amounts to binning the function according to  $\lambda$ -increments, in which case we define the distribution function

$$m_F(\lambda) = \int_{\partial\Omega(\lambda)} \frac{F \rho dS}{|\nabla \lambda|}, \quad (20.26)$$

so that an integral over the distribution is given by

$$\mathcal{I}_F(\lambda_1, \lambda_2) = \int_{\lambda_1}^{\lambda_2} m_F(\lambda) d\lambda. \quad (20.27)$$

In particular, consider the integral over a  $\lambda_{\infty}$ -region

$$\mathcal{I}_F(\lambda, \lambda_{\infty}) = \int_{\lambda}^{\lambda_{\infty}} \left[ \int_{\partial\Omega(\lambda')} \frac{F \rho dS}{|\nabla \lambda'|} \right] d\lambda' = \int_{\lambda}^{\lambda_{\infty}} m_F(\lambda') d\lambda', \quad (20.28)$$

which has the derivative

$$\frac{\partial \mathcal{I}_F(\lambda, \lambda_{\infty})}{\partial \lambda} = - \int_{\partial\Omega(\lambda)} \frac{F \rho dS}{|\nabla \lambda|} = -m_F(\lambda), \quad (20.29)$$

as follows from the fundamental theorem of calculus. Note how the derivative removes the arbitrary reference value,  $\lambda_{\infty}$ . Analogously, the integral over a  $\lambda_{-\infty}$ -region has the derivative

$$\frac{\partial \mathcal{I}_F(\lambda_{-\infty}, \lambda)}{\partial \lambda} = \int_{\partial\Omega(\lambda)} \frac{F \rho dS}{|\nabla \lambda|} = m_F(\lambda). \quad (20.30)$$

### 20.3.6 Moments of $\lambda$

Setting  $F = \lambda$  in the integral (20.25) renders

$$\Lambda(\lambda_1, \lambda_2) \equiv \int_{\Omega(\lambda_1 \leq \lambda \leq \lambda_2)} \lambda dM = \int_{\Omega(\lambda_1 \leq \lambda \leq \lambda_2)} \lambda \rho dV = \int_{\lambda_1}^{\lambda_2} \left[ \int_{\partial\Omega(\lambda)} \frac{\lambda \rho dS}{|\nabla \lambda|} \right] d\lambda. \quad (20.31)$$

If  $\lambda$  is a tracer concentration (tracer mass per fluid mass), then  $\Lambda(\lambda_1, \lambda_2)$  is the mass of tracer within the layer. Observe that  $\lambda$  can be pulled outside of the surface integral in equation (20.31) since  $\lambda$  is constant along  $\partial\Omega(\lambda)$ , thus rendering

$$\Lambda(\lambda_1, \lambda_2) = \int_{\lambda_1}^{\lambda_2} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla \lambda|} \right] \lambda d\lambda = \int_{\lambda_1}^{\lambda_2} m(\lambda) \lambda d\lambda. \quad (20.32)$$

We can likewise define any higher powers as

$$\Lambda^{(n)}(\lambda_1, \lambda_2) \equiv \int_{\lambda_1}^{\lambda_2} \left[ \int_{\partial\Omega(\lambda)} \frac{\rho dS}{|\nabla \lambda|} \right] \lambda^n d\lambda = \int_{\lambda_1}^{\lambda_2} m(\lambda) \lambda^n d\lambda = M(\lambda_1, \lambda_2) \langle \lambda^n \rangle. \quad (20.33)$$

The final equality introduced the mean value for the power

$$\langle \lambda^n \rangle = \frac{\int_{\lambda_1}^{\lambda_2} m(\lambda) \lambda^n d\lambda}{\int_{\lambda_1}^{\lambda_2} m(\lambda) d\lambda} \quad (20.34)$$

as defined over the  $[\lambda_1, \lambda_2]$  region. We refer to  $\langle \lambda^n \rangle$  as the  $n$ -moment of  $\lambda$ , with  $n = 1$  yielding the mean.

### 20.3.7 Internal and external $\lambda$ -moments

Now specify the region  $[\lambda_1, \lambda_2] = [\tilde{\lambda}, \lambda_\infty]$  for the moment equation (20.33). Making use of the differential mass increment,  $dM(\lambda, \lambda_\infty) = -m(\lambda) d\lambda$  as in equation (20.17) allows us to integrate the moment equation by parts

$$\Lambda^{(n)}(\tilde{\lambda}, \lambda_\infty) = \int_{\tilde{\lambda}}^{\lambda_\infty} \lambda^n m(\lambda) d\lambda \quad (20.35a)$$

$$= - \int_{\tilde{\lambda}}^{\lambda_\infty} \lambda^n dM \quad (20.35b)$$

$$= \int_{\tilde{\lambda}}^{\lambda_\infty} [-d(\lambda^n M) + n M \lambda^{n-1} d\lambda] \quad (20.35c)$$

$$= -\lambda_\infty^n M(\lambda_\infty, \lambda_\infty) + \tilde{\lambda}^n M(\tilde{\lambda}, \lambda_\infty) + n \int_{\tilde{\lambda}}^{\lambda_\infty} M(\lambda, \lambda_\infty) \lambda^{n-1} d\lambda \quad (20.35d)$$

$$= \tilde{\lambda}^n M(\tilde{\lambda}, \lambda_\infty) + n \int_{\tilde{\lambda}}^{\lambda_\infty} M(\lambda, \lambda_\infty) \lambda^{n-1} d\lambda, \quad (20.35e)$$

where the final equality follows since  $M(\lambda_\infty, \lambda_\infty) = 0$ . Making use of equation (20.33) thus leads to

$$M(\tilde{\lambda}, \lambda_\infty) \langle \lambda^n \rangle = \underbrace{M(\tilde{\lambda}, \lambda_\infty) \tilde{\lambda}^n}_{\text{external moment}} + \underbrace{n \int_{\tilde{\lambda}}^{\lambda_\infty} M(\lambda, \lambda_\infty) \lambda^{n-1} d\lambda}_{\text{internal moment}}. \quad (20.36)$$

We refer to the rightmost term as the *internal moment* since it is an integral over the region, whereas  $M(\tilde{\lambda}, \lambda_\infty) \tilde{\lambda}^n$  is the *external moment*, which is the region mass times the boundary value,  $\tilde{\lambda}^n$ . We choose the moniker “external” since the external moment increases in direct proportion to the mass crossing the layer boundaries, including the external boundaries. In Section 20.6 we develop a budget for the  $n = 1$  moment, in which the internal moment from equation (20.36) takes the form

$$M(\tilde{\lambda}, \lambda_\infty) [\langle \lambda \rangle - \tilde{\lambda}] = \int_{\tilde{\lambda}}^{\lambda_\infty} M(\lambda, \lambda_\infty) d\lambda. \quad (20.37)$$

We return to the notion of internal and external moments in Section 20.9.2.

### 20.3.8 Further study

The formulation given here in terms of mass distribution functions follows that of [Walin \(1977\)](#) and [Walin \(1982\)](#). In these two papers, Walin pioneered the formalism of water mass transformation analysis, which is sometimes referred to as [Walin analysis](#) in his honor. The concept of internal and external tracer moments follows the internal and external heat introduced by [Holmes et al. \(2019b\)](#).

## 20.4 Water mass transformation across an internal $\lambda$ -surface

We here develop the water mass transformation formalism to compute the transport of fluid crossing an interior  $\lambda$ -interface. This transport is referred to as the [water mass transformation](#) and is written as  $G(\lambda)$ . Figure 20.6 illustrates how this transformation appears in a mass budget for a  $\Delta\lambda$ -layer, with details provided in this section.

### 20.4.1 Dia-surface flux and interior transformation

The object that measures the local water mass transformation is the dia-surface flux detailed in Section 13.4.8 and illustrated in Figure 20.7. This flux is given by the following equivalent expressions

$$w^{\text{dia}} = \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\lambda)}) = \frac{\dot{\lambda}}{|\nabla \lambda|} \quad \text{with} \quad \hat{\mathbf{n}} = \frac{\nabla \lambda}{|\nabla \lambda|} \quad \text{and} \quad \dot{\lambda} = \frac{D\lambda}{Dt}, \quad (20.38)$$

and with  $w^{\text{dia}} > 0$  for water moving to regions of larger  $\lambda$ . It is computed as the projection of the relative velocity,  $(\mathbf{v} - \mathbf{v}^{(\lambda)})$ , onto the direction normal to the  $\lambda$ -surface, with the relative velocity being the difference between the fluid particle velocity,  $\mathbf{v}$ , and the velocity,  $\mathbf{v}^{(\lambda)}$ , of a point on the  $\lambda$ -surface. The velocity,  $\mathbf{v}^{(\lambda)}$ , satisfies the following kinematic constraint

$$(\partial_t + \mathbf{v}^{(\lambda)} \cdot \nabla) \lambda = 0. \quad (20.39)$$

This constraint is based on assuming that  $\mathbf{v}^{(\lambda)}$  measures the velocity of a point that is fixed to the  $\lambda$ -surface. So by construction, the dia-surface flux,  $w^{\text{dia}}$ , locally measures the flux of fluid (volume per area per time) that penetrates a  $\lambda$ -surface in the direction of increasing  $\lambda$ .

The interior water mass transformation,  $G(\lambda)$ , is the area integral of  $\rho w^{\text{dia}}$  over the full

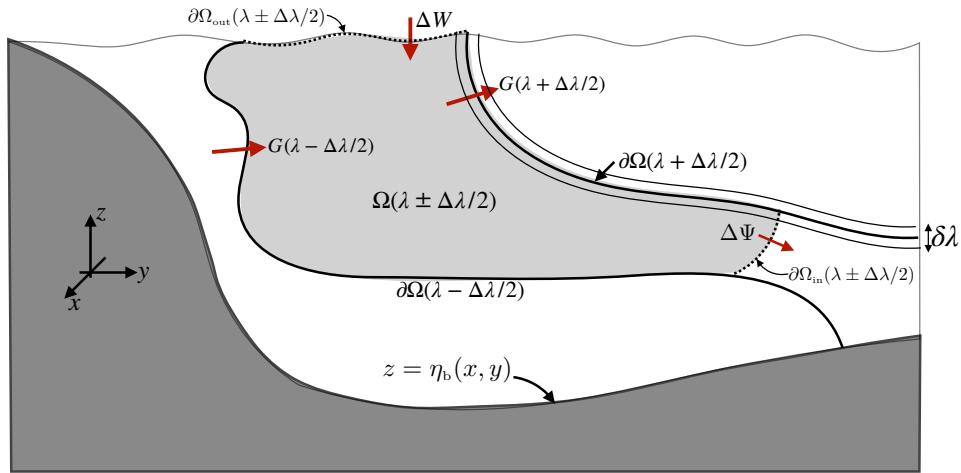


FIGURE 20.6: A layer of fluid with scalar property  $\lambda$  within the range  $[\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$  (with  $\Delta\lambda > 0$ ) and defined over a geographical domain  $\Omega(\lambda \pm \Delta\lambda/2)$ . In this example,  $\lambda$  increases to the right. The net fluid mass transport crossing the layer interfaces,  $\partial\Omega(\lambda \pm \Delta\lambda/2)$ , is  $G(\lambda \pm \Delta\lambda/2)$ , with our convention defining  $G > 0$  for water moving to regions of larger  $\lambda$ , which orients the direction of the arrows. The fluid mass crossing the layer through the geographical bounds,  $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$ , is written  $\Delta\Psi = \Psi(\lambda \pm \Delta\lambda/2)$ , with our convention setting  $\Psi(\lambda \pm \Delta\lambda/2) > 0$  for water leaving  $\Omega(\lambda \pm \Delta\lambda/2)$ . The boundary  $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$  is absent when the domain extends across a basin or the global ocean (e.g., see Figure 20.11). The mass crossing the sea surface,  $\partial\Omega_{out}(\lambda \pm \Delta\lambda/2)$ , through precipitation, evaporation, melt, and river runoff is written  $\Delta W(\lambda \pm \Delta\lambda/2)$ , with  $\Delta W(\lambda \pm \Delta\lambda/2) > 0$  for mass entering  $\Omega(\lambda \pm \Delta\lambda/2)$ . A layer interface can have an arbitrary stratification, such as the vertically non-monotonic profile depicted here for the  $\lambda + \Delta\lambda/2$  interface. Additionally, the domain  $\Omega(\lambda \pm \Delta\lambda/2)$  can generally be disconnected. The net domain boundaries are written  $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2) + \partial\Omega_{out}(\lambda \pm \Delta\lambda/2) + \partial\Omega(\lambda + \Delta\lambda/2) + \partial\Omega(\lambda - \Delta\lambda/2)$ . The infinitesimal  $\delta\lambda$  layer surrounding the  $\partial\Omega(\lambda + \Delta\lambda/2)$  interface arises as part of the method detailed in Section 20.4.3 for computing  $G(\lambda)$  according to the  $\lambda$ -derivative of a volume integral over the  $\delta\lambda$ -layer.

extent of the  $\lambda$ -surface

$$G(\lambda) \equiv \int_{\partial\Omega(\lambda)} \rho w^{\text{dia}} dS = \int_{\partial\Omega(\lambda)} \rho \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\lambda)}) dS = \int_{\partial\Omega(\lambda)} \frac{\rho \dot{\lambda}}{|\nabla \lambda|} dS, \quad (20.40)$$

where  $\partial\Omega(\lambda)$  is the surface occupied by the  $\lambda$ -interface (see Figure 20.6). Furthermore, the dimensions of  $G(\lambda)$  are mass per time

$$G(\lambda) \quad [\equiv] \quad \text{M T}^{-1}, \quad (20.41)$$

thus measuring the mass per time crossing the  $\lambda$ -interface.

Based on the definition (20.40), we see that interior water mass transformation across a  $\lambda$ -interface occurs when there is a material change,  $\dot{\lambda} \neq 0$ , in the property defining the surface. If  $\lambda$  is a conservative tracer, then interior material changes arise from the mixing of  $\lambda$ , thus driving fluid to cross the moving  $\lambda$ -interface (we consider the case of tracer diffusion in Section 20.4.2). Material changes can also arise from sources and sinks, as when considering buoyancy surfaces in the presence of a nonlinear equation of state (Chapter 19). Sources and sinks also affect biogeochemical tracers.

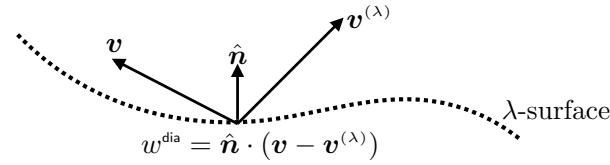


FIGURE 20.7: Schematic of the dia-surface flux,  $w^{\text{dia}} = \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\lambda)})$ , as per equation (20.38), which measures the volume of fluid per time per area that crosses the  $\lambda$ -surface. The velocity of a fluid particle is  $\mathbf{v}$ , whereas  $\mathbf{v}^{(\lambda)}$  is the velocity of a point fixed on the  $\lambda$ -surface. The normal direction,  $\hat{n} = \nabla \lambda / |\nabla \lambda|$ , points toward higher values of  $\lambda$ . If the normal projection of the fluid velocity is distinct from that of the surface velocity ( $\hat{n} \cdot \mathbf{v} \neq \hat{n} \cdot \mathbf{v}^{(\lambda)}$ ), then fluid crosses the surface and  $w^{\text{dia}}$  is nonzero. As per equation (20.39), we see that  $\mathbf{v}^{(\lambda)} \cdot \hat{n} = -|\nabla \lambda|^{-1} \partial_t \lambda$ , so that for static surfaces, in which  $\partial_t \lambda = 0$ , then  $w^{\text{dia}} = \hat{n} \cdot \mathbf{v}$ . We depict the  $\lambda$ -surface with a dotted line to emphasize that it generally has fluid crossing it. The special case of a material surface arises when each point on the surface moves with the fluid velocity,  $\mathbf{v}$ , so that  $\hat{n} \cdot \mathbf{v} = \hat{n} \cdot \mathbf{v}^{(\lambda)}$ , in which case  $w^{\text{dia}} = 0$ .

### 20.4.2 Dia-surface flux from diffusion and steady state balances

In a fluid with interior transformation arising from tracer diffusion, the dia-surface flux across the  $\lambda$ -surface is

$$\rho w^{\text{dia}} = \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\lambda)}) = \frac{\rho \dot{\lambda}}{|\nabla \lambda|} = \frac{\nabla \cdot (\rho \mathbf{K} \cdot \nabla \lambda)}{|\nabla \lambda|}, \quad (20.42)$$

where  $\mathbf{K}$  is the diffusion tensor (Section 15.8). As expected, there is more dia-surface flux crossing the  $\lambda$ -surface where there is more diffusion of  $\lambda$ .

Now assume the flow has reached a steady state so that  $\partial_t \lambda = 0$ , so  $\hat{n} \cdot \mathbf{v}^{(\lambda)} = 0$ . In this case advection of  $\lambda$  exactly balances its diffusion

$$\rho \mathbf{v} \cdot \nabla \lambda = \nabla \cdot (\rho \mathbf{K} \cdot \nabla \lambda), \quad (20.43)$$

where we used the steady mass continuity equation,  $\partial_t \rho = -\nabla \cdot (\mathbf{v} \rho) = 0$ , to write  $\nabla \cdot (\rho \mathbf{v} \lambda) = \rho \mathbf{v} \cdot \nabla \lambda$ . Hence, the dia-surface flux is given by the velocity projected onto the surface normal

$$\rho w^{\text{dia}} = \frac{\rho \dot{\lambda}}{|\nabla \lambda|} \stackrel{\text{diffusion}}{=} \frac{\nabla \cdot (\rho \mathbf{K} \cdot \nabla \lambda)}{|\nabla \lambda|} \stackrel{\text{steady}}{=} \frac{\rho \mathbf{v} \cdot \nabla \lambda}{|\nabla \lambda|} = \rho \mathbf{v} \cdot \hat{n}. \quad (20.44)$$

Notably, there is flow across the  $\lambda$  surfaces since diffusion drives  $\mathbf{v} \cdot \hat{n} \neq 0$ , thus indicating a nonzero water mass transformation. However, in a steady state this flow that crosses the surface is exactly balanced by diffusion of  $\lambda$ , as per equation (20.43). Consequently, there is no evolution of the fluid mass distribution within  $\lambda$  layers, even though there is a nonzero water mass transformation across the  $\lambda$  surfaces defining the layers. That is, with a steady state there can be nonzero water mass transformation across  $\lambda$ -surface, and yet there is no formation (or destruction) of water mass within  $\lambda$ -layers since there is no convergence (in  $\lambda$ -space) of the transformation. These are many words to express a rather simple result. Yet the meaning of these words are important to appreciate, particularly in the midst of manipulations with the water mass transformation equations arising in subsequent sections. We further pursue the case of diffusive transformation in Section 20.6.5.

### 20.4.3 Transformation as the derivative of an integral

The water mass transformation,  $G(\lambda)$ , given by equation (20.40) is an area integral over the  $\lambda$  surface. However, the integral is intractable in practice given the need to estimate an area integral over a surface that is a function of space and time. Here we derive a practical method that provides the means for estimating  $G(\lambda)$  for all numerical realizations of water

mass transformation.

For this purpose, return to the discussion from Section 20.3.5 and set  $F = \dot{\lambda}$  for the integral

$$\mathcal{I}_{\dot{\lambda}}(\lambda, \lambda_{\infty}) = \int_{\Omega(\lambda \leq \lambda_{\infty})} \dot{\lambda}' dM = \int_{\lambda}^{\lambda_{\infty}} \left[ \int_{\partial\Omega(\lambda')} \frac{\rho \dot{\lambda}'}{|\nabla \lambda'|} dS \right] d\lambda' = \int_{\lambda}^{\lambda_{\infty}} G(\lambda') d\lambda'. \quad (20.45)$$

Taking the functional derivative of this integral, and using the fundamental theorem of calculus, we arrive at an expression for the water mass transformation

$$G(\lambda) = -\frac{\partial \mathcal{I}_{\dot{\lambda}}(\lambda, \lambda_{\infty})}{\partial \lambda} \quad \text{fund. thm of calculus} \quad (20.46a)$$

$$= -\lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \left[ \int_{\lambda + \delta\lambda/2}^{\lambda_{\infty}} G(\lambda') d\lambda' - \int_{\lambda - \delta\lambda/2}^{\lambda_{\infty}} G(\lambda') d\lambda' \right] \quad \text{definition of derivative} \quad (20.46b)$$

$$= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\lambda - \delta\lambda/2}^{\lambda + \delta\lambda/2} G(\lambda') d\lambda' \quad \text{combine integral limits} \quad (20.46c)$$

$$= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} \dot{\lambda}' dM \quad \text{equation (20.45)} \quad (20.46d)$$

$$= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} \dot{\lambda}' \rho dV \quad dM = \rho dV. \quad (20.46e)$$

We conclude that interior water mass transformation,  $G(\lambda)$ , across a  $\lambda$ -surface can be computed as a volume integral over an infinitesimal layer surrounding the  $\lambda$ -surface. The volume integral requires information about the material time change,  $\dot{\lambda}$ , a weighting of the time changes according to the mass,  $dM = \rho dV$ , and a binning of the full integrand,  $\rho \dot{\lambda} dV$ , according to  $\lambda$ -class  $[\lambda - \delta\lambda/2, \lambda + \delta\lambda/2]$ .<sup>7</sup> We illustrate the identity (20.46e) in Figure 20.8, with this figure summarizing the practical method, based on binning, used in all calculations of water mass transformation.

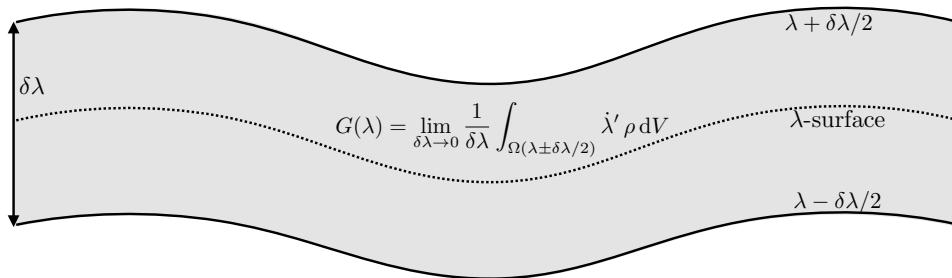


FIGURE 20.8: Illustrating equation (20.46e), in which we compute the interior water mass transformation,  $G(\lambda)$ , across a  $\lambda$ -surface as the volume integral over an infinitesimal region that surrounds the  $\lambda$ -surface. This figure summarizes the binning approach used in all calculations of interior water mass transformation, whereby processes contributing to the integrand,  $\rho \dot{\lambda} dV$ , are binned according to  $\lambda$ -classes as per the discussion in Section 20.4.4.

<sup>7</sup>The final expression in equation (20.46e) might appear to lead to a singularity given the limit  $\delta\lambda \rightarrow 0$ . However, as seen by the form in equation (20.46c), the integration volume also gets smaller as  $\delta\lambda \rightarrow 0$  so that the limit is well defined.

#### 20.4.4 Kinematic and process methods of water mass transformation

There are two complementary means to view interior water mass transformation: the **process method** and the **kinematic method**. The two methods are mathematically identical and so they offer two means to compute the same transformation. The kinematic method tells us *how* transformation happens, whereas the process method helps to understand *why* transformation happens.

##### Kinematic method

The kinematic method focuses on the kinematic means for realizing dia-surface transport, thus providing information concerning how interior transformation occurs. It does so by binning processes contributing to the right hand side of

$$\rho \dot{\lambda} = \partial_t(\rho \lambda) + \nabla \cdot (\rho \lambda \mathbf{v}), \quad (20.47)$$

which arises from the local time tendency plus advection, so that

$$G(\lambda) = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} [\partial_t(\rho \lambda) + \nabla \cdot (\rho \lambda \mathbf{v})] dV. \quad (20.48)$$

This method is operationally simpler than the process method since there are fewer terms to bin (i.e., just the local time tendency plus the divergence of the advective flux). However, it does not provide information about why there is transformation, with that information requiring us to bin tendencies arising from individual processes.

##### Process method

The **process method** focuses on physical (or biogeochemical) processes leading to movement of fluid across the  $\lambda$ -surface, thus providing information concerning *why* interior transformation occurs. It does so by binning processes contributing to the right hand side of the tracer equation

$$\rho \dot{\lambda} = -\nabla \cdot \mathbf{J} + \rho \dot{\Upsilon}, \quad (20.49)$$

where  $\mathbf{J}$  is a flux arising from non-advective processes such as diffusion, and  $\dot{\Upsilon}$  is a source/sink term (dimensions of  $\lambda$  per time) that cannot be written as the convergence of a flux. With the tracer equation (20.49) inserted into the transformation equation (20.46e), we are led to

$$G(\lambda) = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} (-\nabla \cdot \mathbf{J} + \rho \dot{\Upsilon}) dV \quad (20.50a)$$

$$= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \oint_{\partial\Omega(\lambda \pm \delta\lambda/2)} (-\mathbf{J} \cdot \hat{\mathbf{n}}) dS + \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} \rho \dot{\Upsilon} dV. \quad (20.50b)$$

The second equality made use of the divergence theorem to convert the volume integral into a surface integral, with  $\hat{\mathbf{n}}$  the outward normal to the boundary,  $\partial\Omega(\lambda \pm \delta\lambda/2)$ . For tracer sources, we assume they do not modify the fluid mass at a point so that there is no source in the fluid mass equation.

For many purposes it is useful to decompose the non-advective flux divergence into contributions from interior processes, such as ocean mixing, along with boundary fluxes

$$\nabla \cdot \mathbf{J} = \nabla \cdot (\mathbf{J}^{\text{int}} + \mathbf{J}^{\text{out}} + \mathbf{J}^{\text{bot}}). \quad (20.51)$$

By definition,

$$\mathbf{J}^{\text{int}} \cdot \hat{\mathbf{n}} = 0 \quad \text{on } \partial\Omega_{\text{out}}(\lambda \pm \delta\lambda/2) \text{ and } \partial\Omega_{\text{bot}}(\lambda \pm \delta\lambda/2), \quad (20.52)$$

whereas  $\mathbf{J}^{\text{int}} \cdot \hat{\mathbf{n}}$  is generally nonzero on interior layer boundaries. In contrast, the boundary fluxes,  $\mathbf{J}^{\text{out}} \cdot \hat{\mathbf{n}}$  and  $\mathbf{J}^{\text{bot}} \cdot \hat{\mathbf{n}}$ , are identically zero everywhere except on their respective boundaries. Correspondingly, it is convenient to bin the volume weighted convergence,  $-\nabla \cdot \mathbf{J}^{\text{int}} dV$ , according to  $\lambda$ -classes, and to likewise bin the area weighted boundary fluxes,  $\mathbf{J}^{\text{out}} \cdot \hat{\mathbf{n}} dS$  and  $\mathbf{J}^{\text{bot}} \cdot \hat{\mathbf{n}} dS$ . In this way we write the non-advection contribution to water mass transformation in the form

$$\begin{aligned} G(\lambda)_{\text{nonadv}} = & \underbrace{-\lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} \nabla \cdot \mathbf{J}^{\text{int}} dV}_{\text{interior transformation} = \text{volume integral of convergence}} \\ & -\underbrace{\lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\partial\Omega_{\text{out}}(\lambda \pm \delta\lambda/2)} \mathbf{J}^{\text{out}} \cdot \hat{\mathbf{n}} dS}_{\text{surface transformation} = \text{area integral of surface boundary fluxes}} \\ & -\underbrace{\lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\partial\Omega_{\text{bot}}(\lambda \pm \delta\lambda/2)} \mathbf{J}^{\text{bot}} \cdot \hat{\mathbf{n}} dS}_{\text{bottom transformation} = \text{area integral of bottom boundary fluxes}}. \end{aligned} \quad (20.53)$$

Again, this expression decomposes the contribution from interior processes, here represented as the volume integral of the interior flux convergence, from the surface and bottom contributions, here represented as the area integral of the boundary fluxes. This decomposition is further examined in Sections 20.6.6 and 20.7.4 where we focus on the surface contribution to water mass transformation.

Since the boundary fluxes are, by definition, zero except on the boundaries, their divergence can be written in terms of a Dirac delta

$$\nabla \cdot [\mathbf{J}^{\text{out}} + \mathbf{J}^{\text{bot}}] = \mathbf{J}^{\text{out}} \cdot \hat{\mathbf{n}} \delta(z - \eta) + \mathbf{J}^{\text{bot}} \cdot \hat{\mathbf{n}} \delta(z - \eta_b). \quad (20.54)$$

This equation's use in the transformation equation (20.50a) leads to the expression (20.53). Furthermore, as detailed in our study of Neumann boundary conditions in VOLUME 1, we are afforded the ability to introduce Dirac deltas into the boundary conditions (20.54) since the boundary conditions are Neumann (flux) conditions. We made use of this approach for the surface buoyancy fluxes in Section 19.6.3, and it proves useful both conceptually and practically in decomposing contributions to water mass transformation.

## Comments

As we saw, equality of the process method and kinematic method follows because the two provide expressions for the material time derivative. However, in the analysis of numerical model output, it can be nontrivial to realize this equivalence due to the extreme care required to diagnose the terms appearing in the scalar budget equation. See [Drake et al. \(2025\)](#) for a thorough discussion of the details as required for a particular finite volume numerical ocean model.

### 20.4.5 General properties of interior transformation due to diffusion

We here examine some general properties of water mass transformation arising just from interior processes such as diffusion.

**A global integrated constraint on  $G(\lambda)_{\text{int}}$** 

Consider the integrated water mass transformation given by equation (20.45), now integrated over the full ocean domain

$$\mathcal{I}_{\dot{\lambda}}(\lambda_{-\infty}, \lambda_{\infty}) = \int_{\lambda_{-\infty}}^{\lambda_{\infty}} G(\lambda') d\lambda' = \int_{\Omega(\lambda_{-\infty}, \lambda_{\infty})} \rho \dot{\lambda}' dV. \quad (20.55)$$

This integral vanishes for water mass transformations arising from conservative interior processes (i.e., those processes determined by the convergence of a flux)

$$[\mathcal{I}_{\dot{\lambda}}(\lambda_{-\infty}, \lambda_{\infty})]_{\text{int}} = \int_{\lambda_{-\infty}}^{\lambda_{\infty}} G(\lambda')_{\text{int}} d\lambda' \quad (20.56a)$$

$$= - \int_{\Omega(\lambda_{-\infty}, \lambda_{\infty})} \nabla \cdot \mathbf{J}^{\text{int}} dV \quad (20.56b)$$

$$= - \oint_{\partial\Omega(\lambda_{-\infty}, \lambda_{\infty})} \mathbf{J}^{\text{int}} \cdot \hat{\mathbf{n}} dS \quad (20.56c)$$

$$= 0, \quad (20.56d)$$

which follows since  $\mathbf{J}^{\text{int}} \cdot \hat{\mathbf{n}} = 0$  on the ocean boundaries. Hence, there can be no net water mass transformation across a  $\lambda$  surface arising from conservative interior processes

$$\int_{\lambda_{-\infty}}^{\lambda_{\infty}} G(\lambda')_{\text{int}} d\lambda' = 0. \quad (20.57)$$

Instead, conservative interior processes only lead to rearrangement of water within the  $\lambda$ -bins. This result follows since these interior processes conserve the total content of  $\lambda$  within the global domain. Equation (20.57) provides a constraint that is useful to verify with any numerical realization of water mass transformation analysis.

**Transformation across  $\lambda$  surfaces**

We now focus on the transformation occurring along  $\lambda$ -surfaces as in Figure 20.9, in which case

$$G(\lambda)_{\text{int}} = - \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} \nabla \cdot \mathbf{J}^{\text{int}} dV \quad (20.58a)$$

$$= - \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \left[ \int_{\Omega(\lambda + \delta\lambda/2)} \hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} dS - \int_{\Omega(\lambda - \delta\lambda/2)} \hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} dS \right]. \quad (20.58b)$$

Transformation occurs if there is an imbalance between the diffusive transport across the two bounding surfaces,  $\Omega(\lambda + \delta\lambda/2)$  and  $\Omega(\lambda - \delta\lambda/2)$ .

We examine a special case that is not so much physically motivated but is instead motivated to help unravel the meaning of the mathematical symbols. In this case we assume the ocean surface is a constant  $\lambda$  surface with  $\lambda = \lambda_{\text{top}}$  (see right panel of Figure 20.9). Along this surface we have  $\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} = 0$ , by construction since  $\mathbf{J}^{\text{int}}$  is nonzero only within the interior. Hence, the water mass transformation,  $G(\lambda_{\text{top}} - \delta\lambda/2)_{\text{int}}$ , has a contribution just from the flux crossing the  $\Omega(\lambda_{\text{top}} - \delta\lambda)$  surface

$$\lim_{\delta\lambda \rightarrow 0} G(\lambda_{\text{top}} - \delta\lambda/2)_{\text{int}} = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda_{\text{top}} - \delta\lambda)} \hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} dS. \quad (20.59)$$

Likewise, along the top surface we have

$$G(\lambda_{\text{top}})_{\text{int}} = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda_{\text{top}} - \delta\lambda/2)} \hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} dS, \quad (20.60)$$

where we set  $\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} = 0$  for the surface  $\Omega(\lambda_{\text{top}} + \delta\lambda/2)$ , since this surface exists outside of the ocean.

The results (20.59) and (20.60) make it appear that  $G(\lambda_{\text{top}} - \delta\lambda/2)_{\text{int}}$  and  $G(\lambda_{\text{top}})_{\text{int}}$  are unbounded as  $\delta\lambda \rightarrow 0$ , so long as there is a nonzero diffusive transport through  $\Omega(\lambda_{\text{top}} - \delta\lambda)$  or  $\Omega(\lambda_{\text{top}} + \delta\lambda/2)$ . However, this unbounded water mass transformation is not realized since the interior diffusive flux,  $\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} = 0$ , gets smaller in magnitude when approaching the ocean surface, and it does so in order to satisfy the no-flux surface boundary condition (20.52) satisfied by interior processes<sup>8</sup>

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{int}} = 0 \quad \text{at } z = \eta. \quad (20.61)$$

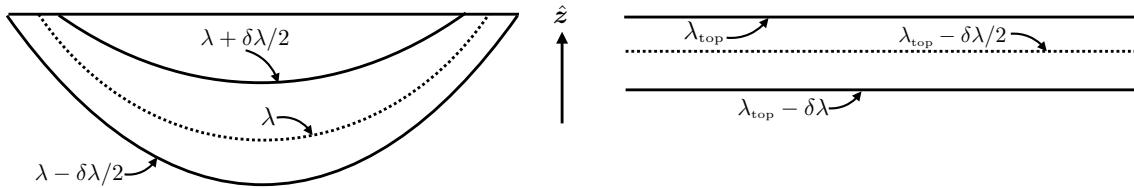


FIGURE 20.9: Example  $\lambda$  surfaces for studying interior transformation due to diffusion. The left panel shows  $\lambda$  surfaces that outcrop to the ocean surface, with the dotted surface the chosen  $\lambda$  surface across which we compute the water mass transformation,  $G(\lambda)$ . The right panel assumes the  $\lambda$  surfaces are flat and with  $\lambda = \lambda_{\text{top}}$  the value along the surface boundary.

#### 20.4.6 Interior versus boundary water mass transformation

Throughout this section we focused on transformation arising from interior processes such as diffusion. What about boundary processes? Indeed, where does interior become boundary? These questions are at the heart of boundary water mass transformation, both the ocean bottom and surface. Surface water mass transformation analysis, introduced in Sections 20.6.6 and 20.7.4, is the most common form of water mass transformation analysis in practice, given that it provides useful inferences based only on surface fluxes. Even so, it is important to note that the distinction between surface and boundary transformation is rather arbitrary for the following reasons.

Consider equation (20.46e) for the water mass transformation

$$G(\lambda) = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \int_{\Omega(\lambda \pm \delta\lambda/2)} \dot{\lambda}' \rho dV. \quad (20.62)$$

When binning the integrand, we bin fluid everywhere in the domain, including next to boundaries where boundary fluxes directly contribute to  $\dot{\lambda}$ . Indeed, these fluxes are exposed when using the divergence theorem to convert the volume integral of  $\rho \dot{\lambda} = -\nabla \cdot \mathbf{J}$  to a boundary area integral of  $\mathbf{J} \cdot \hat{\mathbf{n}}$  (e.g., see Section 20.6.1). Another means to understand the role of boundary contributions follows [Nurser et al. \(1999\)](#), in which boundary fluxes act to inflate layers (including layers

<sup>8</sup>As discussed in Sections 19.5, and 20.7.3, when water is transported across the ocean surface the diffusive flux picks up a nonzero boundary contribution. That contribution is assumed to be part of the surface transformation in equation (20.53) so that the interior diffusive flux still satisfies the no-flux boundary condition (20.61). [Nurser and Griffies \(2019\)](#) discuss these ideas for the case of salinity.

that are otherwise vanished), or to deflate outcropped layers (including layers that end up vanishing). For these reasons, it is unnecessary in principle to distinguish between interior processes and boundary processes. Even so, for practical calculations, it is common to refer to surface transformation when making use just of surface boundary fluxes, which is how we consider surface transformation in Sections 20.6.6 and 20.7.4.

## 20.5 Budget for fluid mass in a $\Delta\lambda$ -layer

In this section we construct the fluid mass budget for a  $\Delta\lambda$ -layer, making reference to Figure 20.6 for the notation. We write the budgets both in terms of extensive quantities (mass and mass transport) and intensive quantities (mass distribution and mass transport distribution). For mass, recall the definition (20.10) where layer mass is written in the equivalent manners

$$\Delta M \equiv M(\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho dV = \int_{\lambda - \Delta\lambda/2}^{\lambda + \Delta\lambda/2} m(\lambda') d\lambda'. \quad (20.63)$$

We introduce analogous expressions for mass transports in the following.

### 20.5.1 Mass transport crossing interior open boundaries

As depicted in Figure 20.6, the layer region has an open boundary that is within the interior of the fluid. The mass transport leaving the layer through this interior open boundary is written

$$\Delta\Psi \equiv \Psi(\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2) = \int_{\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (20.64)$$

where  $\mathbf{v}^{(b)}$  is the velocity for a point on the boundary and  $\hat{\mathbf{n}}$  is the outward normal along the boundary. One common example for an open interior boundary is when choosing a particular latitude, in which case  $\mathbf{v}^{(b)} = 0$  and  $\hat{\mathbf{n}} = \hat{\mathbf{y}}$  so that

$$\Psi(\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2) = \int_{\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)} \rho v dx dz. \quad (20.65)$$

In this case,  $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$  specifies the depth and longitude range for the layer at its intersection along the constant latitude boundary.

### 20.5.2 Mass transport crossing the ocean surface

The mass transport crossing the ocean free surface is written

$$W(\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2) = - \int_{\partial\Omega_{out}(\lambda \pm \Delta\lambda/2)} \rho (\mathbf{v} - \mathbf{v}^{(n)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\partial\Omega_{out}(\lambda \pm \Delta\lambda/2)} Q_m d\mathcal{S}, \quad (20.66)$$

where  $Q_m d\mathcal{S}$  is the mass transport of water crossing the free surface ( $Q_m > 0$  for water entering the ocean), and where we made use of the surface **kinematic boundary condition** to write

$$\rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} \equiv -Q_m, \quad (20.67)$$

### 20.5.3 Mass budget and water mass formation

Bringing the above pieces together leads to the layer mass budget

$$\frac{d\Delta M}{dt} = -\Delta\Psi + \Delta W - [G(\lambda + \Delta\lambda/2) - G(\lambda - \Delta\lambda/2)], \quad (20.68)$$

where for brevity we dropped  $\lambda \pm \Delta\lambda/2$  arguments for  $\Delta M$ ,  $\Delta\Psi$ , and  $\Delta W$ . Although quite simple to state, this mass budget encapsulates a tremendous amount of details about the many processes leading to transport across layer boundaries.

In addition to a layer mass budget, we find it useful to define the water mass [formation](#) for the  $\lambda \pm \Delta\lambda/2$  layer,  $\mathcal{F}\Delta\lambda$ , as

$$\mathcal{F}\Delta\lambda \equiv \underbrace{\frac{d\Delta M}{dt} + \Delta\Psi}_{\text{storage + outflow}} - \underbrace{\Delta W - \frac{[G(\lambda + \Delta\lambda/2) - G(\lambda - \Delta\lambda/2)]}{\Delta\lambda}}_{\text{formation into layer } \Omega_{(\lambda \pm \Delta\lambda/2)}} \Delta\lambda. \quad (20.69)$$

This equation defines the layer water mass formation,  $\mathcal{F}\Delta\lambda$ , is the product of the specific formation,  $\mathcal{F}$ , which is an intensive quantity (mass per time per  $\lambda$ -increment), and the layer width,  $\Delta\lambda$ ,

$$\mathcal{F}\Delta\lambda = \text{mass/time of fluid forming within the interval } [\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]. \quad (20.70)$$

The first equality in equation (20.69) defines the layer water mass formation,  $\mathcal{F}\Delta\lambda$ , as the time change of the mass within the layer (sometimes referred to as the [storage term](#)), plus the net mass leaving through the interior open boundary. Evidently, layer water mass formation is positive if the mass tendency is positive,  $(dM/dt > 0)$ . Additionally, formation is positive if mass leaves through the open boundary,  $\Psi > 0$ , which signals the transport of water to adjacent regions (still within the same layer) and so “forming” additional waters of this particular type. The second equality in equation (20.69) defines the layer water mass formation as the mass entering through the surface boundary outcrop region plus the integrated convergence of water mass transformation across interior layer interfaces.

We illustrate the water mass formation equation in Figure 20.10 for the case of a layer within the ocean interior, in which there is no contribution from  $W$ . Note that for the special case of a lower boundary set to the solid earth, then a steady state balance has

$$\Psi(\lambda_{-\infty}, \lambda) = -G(\lambda). \quad (20.71)$$

That is, the steady state transport through the inner boundary,  $\partial\Omega_{\text{in}}(\lambda_{-\infty}, \lambda)$ , extending from the bottom to the  $\lambda$ -surface is set by minus the water mass transformation,  $-G(\lambda)$ , across the  $\lambda$ -surface. Evidently, if there is no water mass transformation across the  $\lambda$  layer,  $G(\lambda) = 0$ , then a steady flow has no transport through the open boundary,  $\Psi(\lambda_{-\infty}, \lambda) = 0$ .

### 20.5.4 Mass budget and formation in terms of distribution functions

Following our use of a mass distribution function,  $m(\lambda)$ , in equation (20.63), and inspired by [Walsh \(1977\)](#) and [Walsh \(1982\)](#), we here introduce the mass distribution functions for the mass per time per  $\lambda$ -increment that is transported across the open boundary and upper ocean surface

$$\Psi(\lambda_0, \lambda) = \int_{\lambda_0}^{\lambda} \dot{m}_{\Psi}(\lambda') d\lambda' \quad \text{and} \quad W(\lambda_0, \lambda) = \int_{\lambda_0}^{\lambda} \dot{m}_W(\lambda') d\lambda', \quad (20.72)$$

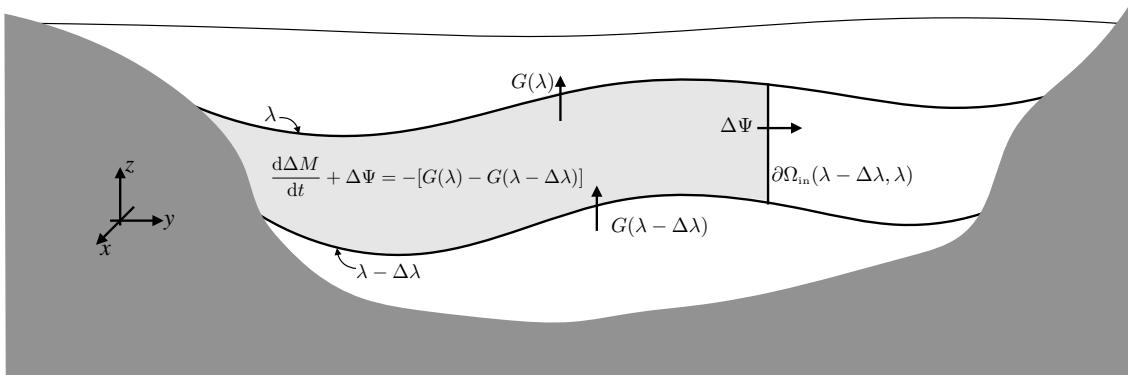


FIGURE 20.10: Water mass budget for the lightly shaded interior layer,  $[\lambda - \Delta\lambda, \lambda]$  with  $\Delta\lambda > 0$ , and with one lateral boundary set to the rock and the other set to a vertical section (e.g., a constant latitude line). Assuming there is no flux of  $\lambda$  through the solid earth bottom, then the mass within the layer changes due to water mass transformation across the lower layer boundary,  $G(\lambda - \Delta\lambda)$ , water mass transformation across the upper layer boundary,  $G(\lambda)$ , and transport crossing the inner boundary,  $\Psi(\lambda - \Delta\lambda, \lambda)$ .

so that

$$d\Psi = \dot{m}_\Psi(\lambda) d\lambda = \text{mass/time of fluid crossing } \partial\Omega_{\text{in}} \text{ within } [\lambda - d\lambda/2, \lambda + d\lambda/2] \quad (20.73a)$$

$$dW = \dot{m}_W(\lambda) d\lambda = \text{mass/time of fluid crossing } \partial\Omega_{\text{out}} \text{ within } [\lambda - d\lambda/2, \lambda + d\lambda/2]. \quad (20.73b)$$

Use of these mass distribution functions leads to the intensive version of the mass budget equation (20.68)

$$\frac{dm}{dt} = -\dot{m}_\Psi + \dot{m}_W - \frac{\partial G}{\partial \lambda}, \quad (20.74)$$

and the corresponding intensive version of the water mass formation equation (20.69)

$$\mathcal{F} = \frac{dm}{dt} + \dot{m}_\Psi = \dot{m}_W - \frac{\partial G}{\partial \lambda}. \quad (20.75)$$

## 20.6 Mass budget for $\lambda$ within a $\lambda_\infty$ -region

We build from our understanding of the fluid mass budget in Section 20.5 to develop a budget for the mass of  $\lambda$  within the  $\lambda_\infty$ -region of Section 20.3.4 and as illustrated in Figure 20.11. Part of our aim is to further develop the formalism while also offering added insights into the causes for water mass transformation,  $G(\lambda)$ .

We here choose to be specific by considering  $\lambda$  to be an intensive property such as a material tracer concentration, in which case  $\lambda \rho dV$  has dimensions of tracer mass. For non-material scalar fields, such as Conservative Temperature or buoyancy, the dimensions are modified accordingly, but the formalism is identical.

### 20.6.1 Processes affecting the mass of $\lambda$ -stuff

Our starting point is the Leibniz-Reynolds transport theorem for a scalar field, here including the possibility of scalar sources

$$\frac{d}{dt} \left[ \int_{\Omega(\lambda \leq \lambda_\infty)} \lambda \rho dV \right] = \int_{\Omega(\lambda \leq \lambda_\infty)} \rho \dot{\Upsilon} dV - \int_{\partial\Omega(\lambda \leq \lambda_\infty)} [\lambda \rho (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (20.76)$$

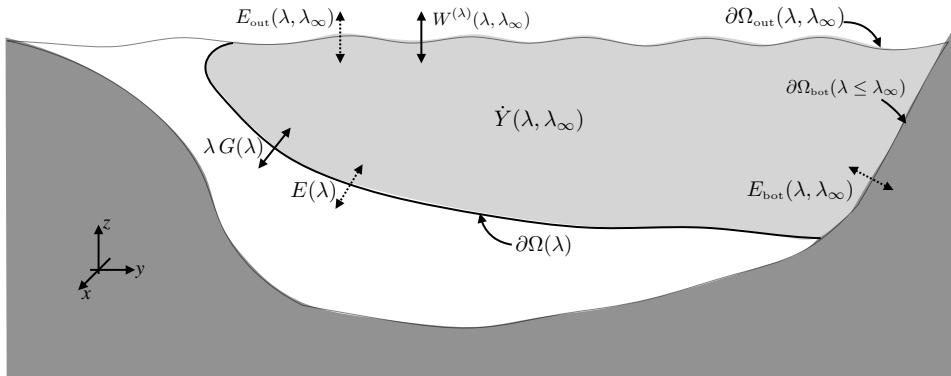


FIGURE 20.11: A  $\lambda_\infty$ -region for studying the  $\lambda$  budget, with the region bounded by the solid-earth bottom,  $\partial\Omega_{\text{bot}}(\lambda \leq \lambda_\infty)$ , the ocean surface boundary,  $\partial\Omega_{\text{out}}(\lambda, \lambda_\infty)$ , and the  $\lambda$ -interface,  $\partial\Omega(\lambda)$ . As drawn here,  $\lambda$  generally increases upward, as per Conservative Temperature, though that assumption is not necessary. In contrast to the  $\Delta\lambda$ -region shown in Figure 20.6, the  $\lambda_\infty$ -region has no interior open boundary. Along the surface boundary, the  $\lambda$  budget is affected by the non-advectional transport,  $E_{\text{out}}(\lambda, \lambda_\infty)$ , arising from processes such as diffusion, plus advective transport,  $W^{(\lambda)}(\lambda, \lambda_\infty)$ , arising from mass transported across the surface that can carry a non-zero amount of  $\lambda$ . Along the bottom, the  $\lambda$  budget is affected by non-advectional transport,  $E_{\text{bot}}(\lambda, \lambda_\infty)$ , arising from processes such as geothermal heating. There is no corresponding advective transport along the bottom since we assume there is no mass crossing the ocean bottom. Along the interior boundary,  $\partial\Omega(\lambda)$ , the budget is affected by non-advectional transport,  $E(\lambda)$ , arising from diffusion, as well as advective transport through  $\lambda G(\lambda)$ , with  $G(\lambda)$  the interior water mass transformation from Section 20.4. Finally, there is the possibility for an interior volume source,  $\dot{Y}(\lambda \leq \lambda_\infty)$ , particularly for buoyancy in the presence of a nonlinear equation of state, and for biogeochemical tracers. Note that along  $\partial\Omega(\lambda)$ , the outward normal,  $\hat{\mathbf{n}} = -\nabla\lambda/|\nabla\lambda|$ , points to regions of smaller  $\lambda$ , which is the opposite convention used to define  $w^{\text{dia}}$  in equation (20.38). We encounter this minus sign in deriving equations (20.84) and (20.104).

The left hand side of this budget equation is the time tendency for the mass of  $\lambda$ -stuff within the domain. The right hand side can be decomposed into the following processes that are depicted in Figure 20.11.

### Non-conservative sources and sinks

As noted in Section 20.4.4, the source term,  $\rho \dot{Y}$ , accounts for processes that cannot be represented as the convergence of a flux. We write its region integrated contribution using the shorthand

$$\dot{Y}(\lambda, \lambda_\infty) \equiv \int_{\Omega(\lambda \leq \lambda_\infty)} \rho \dot{Y} dV \equiv \int_{\lambda}^{\lambda_\infty} \dot{m}_Y(\lambda') d\lambda'. \quad (20.77)$$

The final equality introduced the mass distribution function for the source, in which

$$\dot{m}_Y(\lambda) d\lambda = \text{mass per time of } \lambda\text{-stuff created within } [\lambda - d\lambda/2, \lambda + d\lambda/2]. \quad (20.78)$$

### Transport from non-advectional processes

The contribution from boundary area integrated non-advectional fluxes (e.g., diffusive fluxes and turbulent boundary fluxes) appears in the term

$$-\int_{\partial\Omega(\lambda \leq \lambda_\infty)} \mathbf{J} \cdot \hat{\mathbf{n}} dS = -\int_{\partial\Omega_{\text{out}}(\lambda \leq \lambda_\infty)} \mathbf{J} \cdot \hat{\mathbf{n}} dS - \int_{\partial\Omega_{\text{bot}}(\lambda \leq \lambda_\infty)} \mathbf{J} \cdot \hat{\mathbf{n}} dS - \int_{\partial\Omega(\lambda)} \mathbf{J} \cdot \hat{\mathbf{n}} dS. \quad (20.79)$$

Recall the minus signs arise since a non-advectional flux increases the  $\lambda$  content of the region if the flux is oriented into the region, whereas  $\hat{\mathbf{n}}$  is the region outward normal. The surface,  $\partial\Omega_{\text{out}}(\lambda \leq \lambda_\infty)$ , extends along the upper ocean boundary and supports non-advectional surface

boundary fluxes. Likewise, the boundary,  $\partial\Omega_{\text{bot}}(\lambda \leq \lambda_\infty)$ , intersects the ocean bottom and generally experiences bottom boundary fluxes, such as geothermal heating affecting the enthalpy budget. Finally, the surface,  $\partial\Omega(\lambda)$ , has non-advectional fluxes that cross the  $\lambda$ -interface, with diffusive fluxes the canonical example. Note that along the surface,  $\partial\Omega(\lambda)$ , the outward normal points toward regions of smaller  $\lambda$  (see Figure 20.11), so that

$$\hat{\mathbf{n}} = -\nabla\lambda/|\nabla\lambda| \quad \text{on } \partial\Omega(\lambda). \quad (20.80)$$

The sign here is opposite to that used to define the dia-surface flux,  $w^{\text{dia}}$ , in equation (20.38), where the normal direction is defined as  $\nabla\lambda/|\nabla\lambda|$ . We encounter this minus sign in deriving equations (20.84) and (20.104) in the following.

The boundary area integrated non-advectional fluxes give rise to non-advectional transports, with these transports having dimensions of mass of  $\lambda$ -stuff per time. We find it useful to write the transports using the shorthand

$$-\int_{\partial\Omega(\lambda \leq \lambda_\infty)} \mathbf{J} \cdot \hat{\mathbf{n}} \, d\mathcal{S} = E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) + E(\lambda), \quad (20.81)$$

with a term having a positive value if it increases the  $\lambda$  mass of the region. We furthermore find it useful to introduce the mass distribution functions according to

$$E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) = \int_\lambda^{\lambda_\infty} [\dot{m}_{\text{E}}^{\text{out}}(\lambda') + \dot{m}_{\text{E}}^{\text{bot}}(\lambda')] \, d\lambda', \quad (20.82)$$

where

$$\dot{m}_{\text{E}}^{\text{out}}(\lambda) \, d\lambda = \text{mass per time of } \lambda\text{-stuff from } \partial\Omega_{\text{out}} \text{ transport in } [\lambda - d\lambda/2, \lambda + d\lambda/2] \quad (20.83a)$$

$$\dot{m}_{\text{E}}^{\text{bot}}(\lambda) \, d\lambda = \text{mass per time of } \lambda\text{-stuff from } \partial\Omega_{\text{bot}} \text{ transport in } [\lambda - d\lambda/2, \lambda + d\lambda/2]. \quad (20.83b)$$

### **$\lambda$ transported with interior mass fluxes**

We next consider the contribution to the budget equation (20.76) arising from the transport of  $\lambda$  with mass that crosses the interior interface,  $\partial\Omega(\lambda)$ , whereby

$$-\int_{\partial\Omega(\lambda)} \lambda \rho (\mathbf{v} - \mathbf{v}^{(\lambda)}) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = -\lambda \int_{\partial\Omega(\lambda)} \rho (\mathbf{v} - \mathbf{v}^{(\lambda)}) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \lambda G(\lambda), \quad (20.84)$$

To reach this result we noted that  $\lambda$  can be pulled outside of the  $\partial\Omega(\lambda)$  integral since it is constant along this surface, thus allowing for the introduction of the water mass transformation,  $G(\lambda)$ , given by equation (20.40). Furthermore, we made note of the minus sign identified in equation (20.80) for the outward normal along  $\partial\Omega(\lambda)$ .

### **Surface boundary mass fluxes**

The final term contributing to the right hand side of the  $\lambda$  budget equation (20.76) arises from the surface mass transport along the boundary,  $\partial\Omega_{\text{out}}(\lambda \leq \lambda_\infty)$ ,

$$-\int_{\partial\Omega(\lambda \leq \lambda_\infty)} \lambda \rho (\mathbf{v} - \mathbf{v}^{(\text{b})}) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \int_{\partial\Omega_{\text{out}}(\lambda \leq \lambda_\infty)} \lambda Q_m \, d\mathcal{S} \equiv W^{(\lambda)}(\lambda, \lambda_\infty). \quad (20.85)$$

To reach the first equality we followed the steps in Section 20.5.2 by using the [kinematic boundary condition](#) to introduce the surface mass transport,  $\mathcal{Q}_m dS$ . The final equality introduced a shorthand that corresponds to the  $W(\lambda, \lambda_\infty)$  from Section 20.5.2. In the following, we find it useful to introduce the mass distribution function,  $\dot{m}_w(\lambda)$ , from equation (20.73b), thus rendering

$$W(\lambda, \lambda_\infty) = \int_\lambda^{\lambda_\infty} \dot{m}_w(\lambda') d\lambda' \quad \text{and} \quad W^{(\lambda)}(\lambda, \lambda_\infty) = \int_\lambda^{\lambda_\infty} \lambda \dot{m}_w(\lambda') d\lambda'. \quad (20.86)$$

Following the discussion in Sections 19.5.2, we have not assumed a relation between  $\lambda$  along the interface,  $\partial\Omega_{\text{out}}(\lambda \leq \lambda_\infty)$ , and the concentration,  $\lambda_m$ , contained in the entering mass. We prefer to keep the discussion general for now, providing a relation only when necessary.

## 20.6.2 Summary of the $\lambda$ budget

Bringing terms together leads to the expanded version of the mass budget (20.76) for  $\lambda$ -stuff, now written as

$$\frac{d}{dt} \left[ \int_{\Omega(\lambda \leq \lambda_\infty)} \lambda \rho dV \right] = \dot{Y}(\lambda, \lambda_\infty) + E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) + E(\lambda) + W^{(\lambda)}(\lambda, \lambda_\infty) + \lambda G(\lambda), \quad (20.87)$$

which has the equivalent expression in terms of mass distribution functions

$$\begin{aligned} \frac{d}{dt} \int_\lambda^{\lambda_\infty} \lambda' m(\lambda') d\lambda' \\ = \int_\lambda^{\lambda_\infty} [\dot{m}_Y(\lambda') + \dot{m}_{\text{E}}^{\text{out}}(\lambda') + \dot{m}_{\text{E}}^{\text{bot}}(\lambda') + \lambda \dot{m}_w(\lambda')] d\lambda' + E(\lambda) + \lambda G(\lambda). \end{aligned} \quad (20.88)$$

These budget equations also include the budget for seawater mass within the  $\Omega(\lambda \leq \lambda_\infty)$  region, as seen by dropping the diffusive contribution (since diffusion does not affect the mass of a fluid element), dropping the bottom boundary terms (no seawater mass through the solid-earth bottom), and dropping the interior source term (no interior seawater mass sources), thus revealing the seawater mass budget equation within the  $\lambda_\infty$ -region

$$\frac{d}{dt} \left[ \int_{\Omega(\lambda \leq \lambda_\infty)} \rho dV \right] = W(\lambda, \lambda_\infty) + G(\lambda), \quad (20.89)$$

which takes on the following form in terms of mass distribution functions

$$\frac{d}{dt} \int_\lambda^{\lambda_\infty} m(\lambda') d\lambda' = \int_\lambda^{\lambda_\infty} \dot{m}_w(\lambda') d\lambda' + G(\lambda). \quad (20.90)$$

## 20.6.3 Processes leading to water mass transformation

We now massage the budget equations to explicitly identify processes leading to water mass transformation,  $G(\lambda)$ . For that purpose, make use of the moment equation (20.37) to write

$$\int_{\Omega(\lambda \leq \lambda_\infty)} \lambda \rho dV = M(\lambda, \lambda_\infty) \langle \lambda \rangle = M(\lambda, \lambda_\infty) \lambda + \int_\lambda^{\lambda_\infty} M(\lambda', \lambda_\infty) d\lambda', \quad (20.91)$$

which then leads to

$$\frac{d[M(\lambda, \lambda_\infty) \langle \lambda \rangle]}{dt} = \lambda \frac{dM(\lambda, \lambda_\infty)}{dt} + \int_\lambda^{\lambda_\infty} \frac{dM(\lambda', \lambda_\infty)}{dt} d\lambda'. \quad (20.92)$$

Note that there is no  $d\lambda/dt$  term since  $\lambda$  in this equation is a fixed parameter, not a field.

### Integrated water mass transformation over the $\lambda_\infty$ -region

Use of the  $\lambda$  budget equation (20.87) for the left hand side of equation (20.92), and the mass budget equation (20.89) for the right hand side, yields

$$\begin{aligned} \dot{Y}(\lambda, \lambda_\infty) + E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) + E(\lambda) + W^{(\lambda)}(\lambda, \lambda_\infty) + \lambda G(\lambda) \\ = \lambda [W(\lambda, \lambda_\infty) + G(\lambda)] + \int_\lambda^{\lambda_\infty} [W(\lambda', \lambda_\infty) + G(\lambda')] d\lambda'. \end{aligned} \quad (20.93)$$

Observe that the  $\lambda G(\lambda)$  term cancels on both sides of this equation. The three contributions from the surface boundary mass fluxes also cancel, as revealed through the following identity

$$W^{(\lambda)}(\lambda, \lambda_\infty) - \lambda W(\lambda, \lambda_\infty) = \int_\lambda^{\lambda_\infty} (\lambda' - \lambda) \dot{m}_w(\lambda') d\lambda' \quad (20.94a)$$

$$= \int_\lambda^{\lambda_\infty} \left[ \int_{\lambda'}^{\lambda_\infty} \dot{m}_w(\lambda'') d\lambda'' \right] d\lambda' \quad (20.94b)$$

$$= \int_\lambda^{\lambda_\infty} W(\lambda', \lambda_\infty) d\lambda', \quad (20.94c)$$

where the second equality follows from the double integral formula (20.164) derived in Exercise 20.5. To understand the physical reason we see no water mass transformation from surface mass fluxes, recall the discussion in Section 19.5.2. Namely, mixing and internal sources provide the only means for irreversible changes to water masses and thus to water mass transformation. In contrast, boundary mass transport contributes to transformation only if the mass participates in mixing. That is, the mass associated with boundary mass transport is incorporated into the ocean (or leaves the ocean) only in the presence of mixing. It is reassuring that the budget formalism leads to this same conclusion. Furthermore, this result is consistent with the expression (20.50b), whereby the water mass transformation is, again, determined solely in terms of the non-advectional fluxes at the region boundaries, plus the interior source term.

Cancelling the mass transport terms thus leads to the integrated water mass transformation

$$\int_\lambda^{\lambda_\infty} G(\lambda') d\lambda' = \dot{Y}(\lambda, \lambda_\infty) + E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) + E(\lambda). \quad (20.95)$$

Each term in this equation has dimensions mass of  $\lambda$ -stuff per time. This equation is an integrated version of the expression (20.50b) for the water mass transformation, here having exposed the processes contributing to transformation over the range  $\lambda \leq \lambda_\infty$ . Evidently, the accumulated effects from sources within the interior, plus non-advectional fluxes along the surface and interior boundaries, lead to an integrated interior water mass transformation.

### Water mass transformation across the $\lambda$ -interface

An expression for the water mass transformation across the  $\lambda$ -interface is derived by taking  $\partial/\partial\lambda$  of equation (20.95)

$$G(\lambda) = -\frac{\partial}{\partial\lambda} \left[ \dot{Y}(\lambda, \lambda_\infty) + E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) + E(\lambda) \right]. \quad (20.96)$$

This equation shows that water mass transformation across a  $\lambda$ -surface,  $G(\lambda)$ , is the  $\lambda$ -convergence of interior sources, boundary transport, and mixing processes. This equation takes on the following distributional form

$$G(\lambda) = \dot{m}_Y(\lambda) + \dot{m}_E^{\text{bot}}(\lambda) + \dot{m}_E^{\text{out}}(\lambda) - \frac{\partial E(\lambda)}{\partial\lambda}. \quad (20.97)$$

Again, this equation says that water mass transformation across a  $\lambda$  surface occurs if there are interior sources that modify the distribution through  $\dot{m}_Y(\lambda)$ ; diffusive processes at the boundaries through  $\dot{m}_E^{\text{bot}}(\lambda)$  and  $\dot{m}_E^{\text{out}}(\lambda)$ ; and the convergence of diffusion across the interior layer interfaces through  $-\partial E(\lambda)/\partial\lambda$ .

#### 20.6.4 Water mass formation

Recall from Section 20.5.4 that we derived equation (20.75) for water mass formation

$$\mathcal{F} = \dot{m}_W - \frac{\partial G}{\partial\lambda}. \quad (20.98)$$

Making use of equations (20.96) and (20.97) for the water mass transformation then leads to

$$\mathcal{F}(\lambda) = \dot{m}_W(\lambda) + \frac{\partial^2 E(\lambda)}{\partial\lambda^2} + \frac{\partial^2}{\partial\lambda^2} \left[ \dot{Y}(\lambda, \lambda_\infty) + E_{\text{out}}(\lambda, \lambda_\infty) + E_{\text{bot}}(\lambda, \lambda_\infty) \right] \quad (20.99a)$$

$$= \dot{m}_W(\lambda) + \frac{\partial^2 E(\lambda)}{\partial\lambda^2} - \frac{\partial [\dot{m}_Y(\lambda) + \dot{m}_E^{\text{bot}}(\lambda) + \dot{m}_E^{\text{out}}(\lambda)]}{\partial\lambda}. \quad (20.99b)$$

The water mass formation equations (20.99a) and (20.99b) are dense with information. To help unpack some of that information, we consider the example in Figure 20.12. In this example, water experiences boundary forcing that increases the lowest value of  $\lambda$  while retaining the highest value unchanged. This forcing destroys waters with low  $\lambda$  and moves the distribution towards larger values. If  $\lambda = \Theta$ , we think of this example as the effects of surface heating on water masses.

#### 20.6.5 Interior water mass transformation from diffusion

Equations (20.96) and (20.97) summarize the full suite of processes affecting water mass transformation across a chosen  $\lambda$ -surface. To help understand the interior term, follow the discussion from Section 20.4.2 by considering the interior transformation arising from diffusion

$$G_{\text{interior}}(\lambda) = -\frac{\partial E(\lambda)}{\partial\lambda} = \frac{\partial}{\partial\lambda} \left[ \int_{\partial\Omega(\lambda)} \mathbf{J} \cdot \hat{\mathbf{n}} \, d\mathcal{S} \right] = \int_{\partial\Omega(\lambda)} \rho (\mathbf{v} - \mathbf{v}^{(\lambda)}) \cdot (\nabla\lambda / |\nabla\lambda|) \, d\mathcal{S}, \quad (20.100)$$

where the final equality used the definition (20.40) of interior water mass transformation according to the integrated dia-surface flux,  $w^{\text{dia}} = \rho (\mathbf{v} - \mathbf{v}^{(\lambda)}) \cdot (\nabla\lambda / |\nabla\lambda|)$ . Inserting the

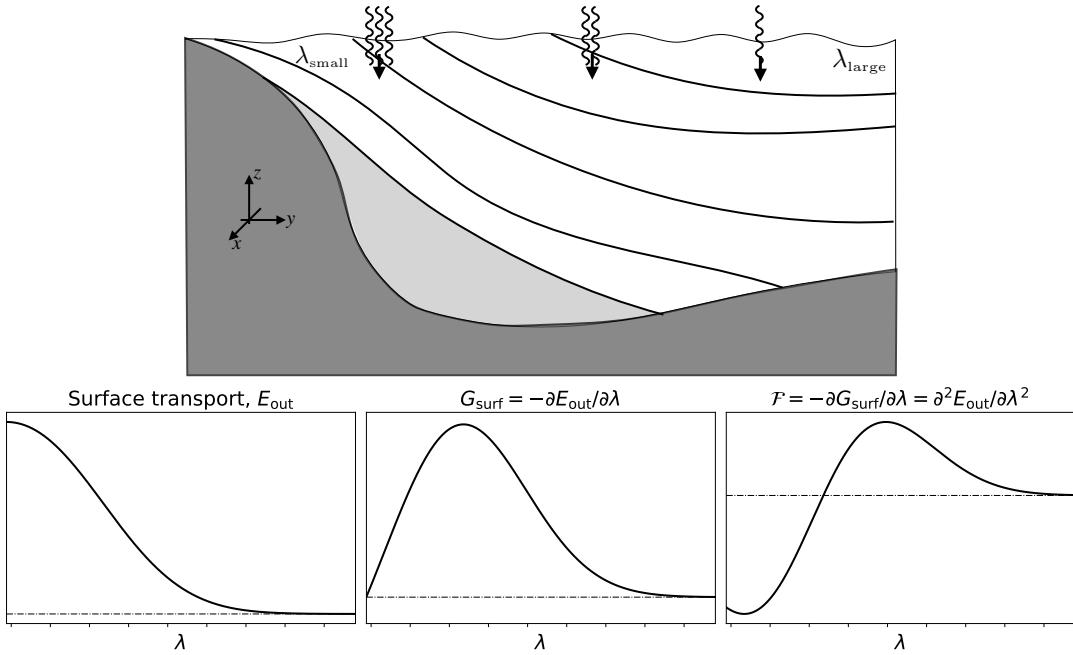


FIGURE 20.12: An example inspired from Figure 4 of Nurser et al. (1999). The top panel shows a section through the  $\lambda$ -field that outcrops with small  $\lambda$  values in the south ( $\lambda = \Theta$  in the Southern Ocean is an example). The surface transport is all directed to increase the value of  $\lambda$  throughout the section so that  $E_{\text{out}} > 0$  (e.g., surface heating), and with a stronger surface flux acting on the smaller values of  $\lambda$  (e.g., summertime warming in the high latitude Southern Ocean). There is a deep layer (the light gray layer) that does not outcrop in this section and whose values of  $\lambda$  are lower than any at the surface. The lower left panel shows the profile for the surface transport,  $E_{\text{out}}$ , binned according to  $\lambda$ .  $E_{\text{out}}$  approaches zero as  $\lambda$  approaches  $\lambda_{\max}$ , so that  $\lambda_{\max}$  remains the maximum value of  $\lambda$ . The middle lower panel shows the surface water mass transformation  $G = -\partial E_{\text{out}} / \partial \lambda$ , which is all positive, and with largest transformation at intermediate values of  $\lambda$  where the  $\lambda$ -derivative of  $E_{\text{out}}$  is largest in magnitude. The lower right panel shows the formation,  $F = -\partial G / \partial \lambda = \partial^2 E_{\text{out}} / \partial \lambda^2$ . Under this surface forcing, waters with relatively small  $\lambda$  are destroyed ( $F < 0$ ) while waters with relatively large  $\lambda$  are formed ( $F > 0$ ). The three lower panels have the same horizontal  $\lambda$  axis, with  $\lambda$  increasing to the right along the horizontal axis. The dot-dash line is the zero line for each panel. All units are arbitrary, with zero the only physically relevant number for the lower three panels.

diffusive flux,  $\mathbf{J} = -\rho \mathbf{K} \cdot \nabla \lambda$  with  $\mathbf{K}$  the diffusion tensor, renders

$$E(\lambda) = - \int_{\partial\Omega(\lambda)} \mathbf{J} \cdot \hat{\mathbf{n}} dS = \int_{\partial\Omega(\lambda)} \rho (\mathbf{K} \cdot \nabla \lambda) \cdot \hat{\mathbf{n}} dS. \quad (20.101)$$

To go a bit further, let us make a couple of simplifying assumptions. First, just like in Section 20.4.2, assume a steady state so that the divergence of the advective flux balances the convergence of the diffusive flux

$$\nabla \cdot (\rho \mathbf{v} \lambda) = \nabla \cdot (\rho \mathbf{K} \cdot \nabla \lambda), \quad (20.102)$$

in which case

$$E(\lambda) = \int_{\partial\Omega(\lambda)} \lambda \rho \mathbf{v} \cdot \hat{\mathbf{n}} dS. \quad (20.103)$$

Next, assume the only  $\lambda$ -dependence for the right hand side integral is with the explicit appearance of  $\lambda$  within the integrand. One example is if the flow is independent of  $\lambda$  and the region of interest has parallel  $\lambda$  surfaces that are equally spaced. In this case we can naively

take the functional derivative to find

$$-\frac{\partial E}{\partial \lambda} = - \int_{\partial\Omega(\lambda)} \rho \mathbf{v} \cdot \hat{\mathbf{n}} \, dS = \int_{\partial\Omega(\lambda)} \frac{\rho \mathbf{v} \cdot \nabla \lambda}{|\nabla \lambda|} \, dS. \quad (20.104)$$

This is the expected expression for a steady state version of  $G_{\text{interior}}(\lambda)$ , which results from setting  $\mathbf{v}^{(\lambda)} = 0$  in equation (20.100).

## 20.6.6 Surface water mass transformation

Contributions to transformation from surface processes in the transformation equation (20.97) are given by

$$G_{\text{out}}(\lambda) \equiv -\frac{\partial E_{\text{out}}(\lambda, \lambda_\infty)}{\partial \lambda} = \dot{m}_{\text{E}}^{\text{out}}(\lambda) = -\lim_{\delta \lambda \rightarrow 0} \frac{1}{\delta \lambda} \int_{\partial\Omega_{\text{out}}(\lambda \pm \delta \lambda / 2)} \mathbf{J} \cdot \hat{\mathbf{n}} \, dS, \quad (20.105)$$

where the final equality made use of equation (20.50b). Such surface transformation forms the focus of many studies of water mass transformation because it only requires surface boundary information, which is generally more accessible than information from interior ocean mixing processes or bottom geothermal processes. Furthermore, much of the transformation of water occurs in surface regions since this region is home to large contributions from surface boundary fluxes and associated ocean mixing. The basic equation we use is the non-advection flux equation (19.68), rewritten here for the scalar field  $\lambda$

$$-\mathbf{J} \cdot \hat{\mathbf{n}} = \mathcal{Q}_\lambda - \lambda \mathcal{Q}_m = \mathcal{Q}_\lambda^{\text{nonadv}} + (\lambda_m - \lambda) \mathcal{Q}_m. \quad (20.106)$$

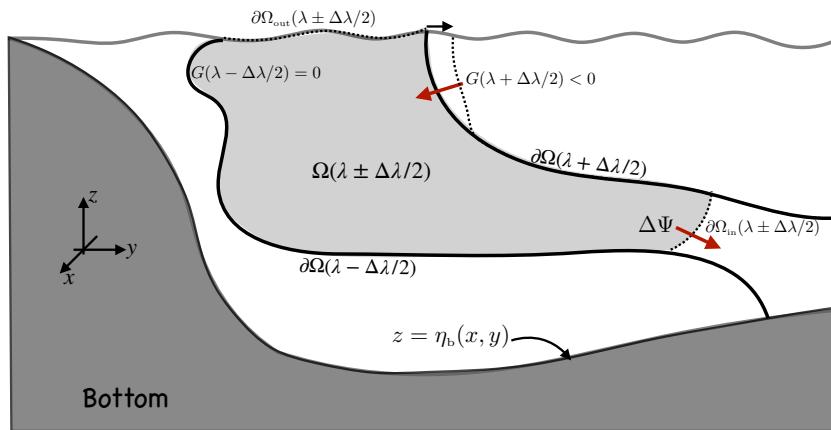


FIGURE 20.13: An example of surface transformation driven circulation oriented according to the Southern Ocean with Antarctica to the left. Here we depict a layer that is exposed to some form of air-sea interaction that causes the interface with  $\lambda + \Delta\lambda/2$  to move meridionally. For example, if  $\lambda = \Theta$  (Conservative Temperature), then a cold storm generally causes  $\Theta$  outcrops to move meridionally, with cold isotherms moving northward (black arrow at ocean surface). Movement of the near-surface portion of the  $\partial\Omega(\lambda + \Delta\lambda/2)$  interface causes fluid to entrain into the layer and thus contribute to the surface water mass transformation,  $G(\lambda + \Delta\lambda/2) < 0$  (red arrow near the surface directed to the south; remember that  $G(\lambda) > 0$  when water moves to the larger  $\lambda$  direction). In turn, the boundary,  $\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)$ , expands as the near-surface portion of the interface,  $\partial\Omega(\lambda + \Delta\lambda/2)$ , moves to the north as a result of the entrained new water (black arrow moving to the north). If there is a net convergence of water mass into the layer (as determined by the net mass crossing both layer interfaces  $\partial\Omega(\lambda + \Delta\lambda/2)$  and  $\partial\Omega(\lambda - \Delta\lambda/2)$ ), then mass accumulates within the layer  $[\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$ .

Expressions for the water mass budget are based on integrations over finite regions bounded by  $\lambda$ -surfaces. Since the budgets are formulated over layers, the mass budget offers the means

to make very general inferences about the circulation within those layers, even without a direct measurement of the flow. This is a key power of water mass transformation theory, with surface transformation playing a central role. We depict an example in Figure 20.13 where the surface outcrop of a layer is exposed to air-sea interactions that lead to a meridional movement of the interface,  $\lambda + \Delta\lambda/2$ . This movement laterally entrains mass into the layer. If there is a net convergence of mass into the layer, then the layer mass increases.

## 20.7 Buoyancy water mass transformation

In Section 20.2 we considered the transformation of water masses as defined by buoyancy classes. For that analysis we set  $\lambda = \gamma$ , where  $\gamma$  is a density-like field whose isosurfaces approximate constant buoyancy surfaces and hence [neutral directions](#). We here offer further details for such buoyancy water mass analysis. Neither buoyancy nor density are conservative scalars due to seawater's nonlinear [equation of state](#). Hence, buoyancy water mass analysis must account for the source and sink terms that add to buoyancy evolution, in addition to diffusive transport processes affecting the conservative tracers,  $S$  and  $\Theta$ .

### 20.7.1 Material time changes to $S$ and $\Theta$

The material time derivative of  $\gamma$  can be written as the sum of contributions from salinity,  $S$ , and Conservative Temperature,  $\Theta$ ,

$$\rho \dot{\gamma} = \frac{\partial \gamma}{\partial S} \rho \dot{S} + \frac{\partial \gamma}{\partial \Theta} \rho \dot{\Theta}. \quad (20.107)$$

Following the decomposition of the water mass transformation in Section 20.4.4 for a general tracer, we here write the material time derivatives in the form

$$\rho \dot{S} = -\nabla \cdot \mathbf{J}_{\text{int}}^{(S)} - \mathbf{J}_{\text{out}}^{(S)} \cdot \hat{\mathbf{n}} \delta(z - \eta) - \mathbf{J}_{\text{bot}}^{(S)} \cdot \hat{\mathbf{n}} \delta(z - \eta_b) \quad (20.108a)$$

$$\rho \dot{\Theta} = -\nabla \cdot \mathbf{J}_{\text{int}}^{(\Theta)} - \mathbf{J}_{\text{out}}^{(\Theta)} \cdot \hat{\mathbf{n}} \delta(z - \eta) - \mathbf{J}_{\text{bot}}^{(\Theta)} \cdot \hat{\mathbf{n}} \delta(z - \eta_b), \quad (20.108b)$$

where we assumed there are no interior sources of  $S$  or  $\Theta$ . The surface and bottom boundary contributions are weighted by the Dirac delta and projected into the normal direction along the two respective boundary surfaces. Following from the decomposition of water mass transformation given by equation (20.53), we are led to the following form for buoyancy transformation

$$G(\gamma) = \underbrace{- \lim_{\delta\gamma \rightarrow 0} \frac{1}{\delta\gamma} \int_{\Omega(\gamma \pm \delta\gamma/2)} \left( \frac{\partial \gamma}{\partial S} \nabla \cdot \mathbf{J}_{\text{int}}^{(S)} + \frac{\partial \gamma}{\partial \Theta} \nabla \cdot \mathbf{J}_{\text{int}}^{(\Theta)} \right) dV}_{\text{interior buoyancy transformation} = \text{volume integral of convergence}} \\ - \underbrace{\lim_{\delta\gamma \rightarrow 0} \frac{1}{\delta\gamma} \int_{\partial\Omega_{\text{out}}(\gamma \pm \delta\gamma/2)} \left( \frac{\partial \gamma}{\partial S} \mathbf{J}_{\text{out}}^{(S)} + \frac{\partial \gamma}{\partial \Theta} \mathbf{J}_{\text{out}}^{(\Theta)} \right) \cdot \hat{\mathbf{n}} d\mathcal{S}}_{\text{surface buoyancy transformation} = \text{area integral of surface boundary fluxes}} \\ - \underbrace{\lim_{\delta\gamma \rightarrow 0} \frac{1}{\delta\gamma} \int_{\partial\Omega_{\text{bot}}(\gamma \pm \delta\gamma/2)} \left( \frac{\partial \gamma}{\partial S} \mathbf{J}_{\text{bot}}^{(S)} + \frac{\partial \gamma}{\partial \Theta} \mathbf{J}_{\text{bot}}^{(\Theta)} \right) \cdot \hat{\mathbf{n}} d\mathcal{S}}_{\text{bottom buoyancy transformation} = \text{area integral of bottom boundary fluxes}}. \quad (20.109)$$

This expression is explored in the remainder of this section.

## 20.7.2 Interior buoyancy water mass transformation

Contributions from cabbeling, thermobaricity, and halobaricity (Section 19.3) arise from the interior transformation appearing in equation (20.109). Furthermore, in the special case of a linear equation of state, whereby  $\nabla(\partial\gamma/\partial\Theta) = 0$  and  $\nabla(\partial\gamma/\partial S) = 0$ , then equation (20.57) means that the global integral of the interior transformation vanishes

$$\int_{\gamma_{-\infty}}^{\gamma_{\infty}} G_{\text{int}}(\gamma) d\gamma = \int_{\mathcal{R}} \left( \frac{\partial\gamma}{\partial S} \nabla \cdot \mathbf{J}_{\text{int}}^{(S)} + \frac{\partial\gamma}{\partial\Theta} \nabla \cdot \mathbf{J}_{\text{int}}^{(\Theta)} \right) dV = 0, \quad (20.110)$$

where  $\mathcal{R}$  is the global ocean domain. By inference, we conclude that any nonzero result for this integral is a global measure of the effects from the nonlinear equation of state

$$\text{contribution from nonlinear equation of state} = \int_{\gamma_{-\infty}}^{\gamma_{\infty}} G_{\text{int}}(\gamma) d\gamma. \quad (20.111)$$

## 20.7.3 Surface non-advection flux for $S$ and $\Theta$

We review a few of the distinct characteristics of surface non-advection fluxes of  $S$  and  $\Theta$  as detailed in Section 19.5.2, here working with salinity,  $S$ , rather than salt concentration,  $S = S/1000$ .

### Non-advection salt flux

The non-advection surface boundary flux for salt is given by equation (19.68), here written as

$$-\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = Q_S^{\text{non-adv}} + (S_m - S) Q_m, \quad (20.112)$$

where  $Q_S^{\text{non-adv}}$  is a non-advection salt flux, such as might arise from parameterized turbulent transfer. For the salt concentration of water crossing the ocean surface, we generally take  $S_m = 0$  for precipitation, evaporation, and river runoff, whereas  $S_m \neq 0$  for sea ice melt and formation. Furthermore, the boundary term,  $S = S(z = \eta)$ , is commonly approximated by the bulk salt concentration in the upper few meters of the ocean.

### Non-advection flux for Conservative Temperature

For Conservative Temperature we follow the discussion in Section 19.5.3, whereby the non-advection flux is given by equation

$$-\mathbf{J}^{(\Theta)} \cdot \hat{\mathbf{n}} = Q_{\Theta}^{\text{non-adv}} + (\Theta_m - \Theta) Q_m. \quad (20.113)$$

It is common to approximate the difference  $\Theta_m - \Theta(z = \eta) = 0$ , in which case the non-advection flux is just due to turbulent and radiative heat fluxes

$$-\mathbf{J}^{(\Theta)} \cdot \hat{\mathbf{n}} = Q_{\Theta}^{\text{non-adv}} \quad \text{if } \Theta_m - \Theta(z = \eta) = 0. \quad (20.114)$$

## 20.7.4 Surface buoyancy water mass transformation

In Figure 20.14 we illustrate how a meridional gradient in the surface buoyancy loss causes entrainment into buoyancy layers. The calculation of this surface water mass transformation is

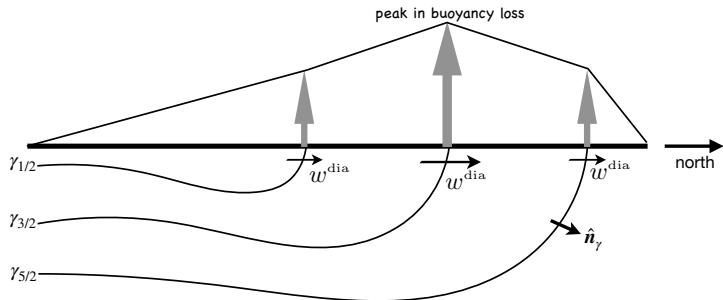


FIGURE 20.14: An example of surface water mass transformation, here illustrating the effects of transformation due to a meridional gradient in the surface buoyancy loss, with  $\gamma$  a density-like coordinate. The example is oriented for the northern hemisphere with increasing latitudes to the north/right. Buoyancy loss is denoted by the thick vertical arrows indicating the removal of buoyancy from the ocean, thus causing surface water to loose buoyancy and get more dense. The surface buoyancy loss causes  $\gamma$  interfaces to migrate to the south (denser water moves southward). This southward migration of the density outcrops causes dianeutral mass flux to move from lighter layers to denser layers, thus causing entrainment into the denser layers (black vectors pointed to the north,  $w^{\text{dia}}$ ). With a peak in the buoyancy loss at a particular latitude, more entrainment is driven into the layer to the north of the peak (water converges to the layer  $\gamma_{3/2} \leq \gamma \leq \gamma_{5/2}$ ) and less entrainment into the layer to the south (water diverges from the layer  $\gamma_{1/2} \leq \gamma \leq \gamma_{3/2}$ ).

found by inserting the surface fluxes into equation (20.109) as per Section 20.7.3 to write

$$G(\gamma)^{\text{surf}} = \lim_{\delta\gamma \rightarrow 0} \frac{1}{\delta\gamma} \int_{\partial\Omega_{\text{out}}(\gamma \pm \delta\gamma/2)} (\gamma \beta [Q_S^{\text{nonadv}} + (S_m - S) Q_m] - \gamma \alpha [Q_\Theta^{\text{non-adv}} + (\Theta_m - \Theta) Q_m]) d\mathcal{S}, \quad (20.115)$$

where we introduced the thermal expansion and saline contraction coefficients, here defined according to<sup>9</sup>

$$\alpha = -\frac{1}{\gamma} \frac{\partial \gamma}{\partial \Theta} \quad \text{and} \quad \beta = \frac{1}{\gamma} \frac{\partial \gamma}{\partial S}. \quad (20.116)$$

Recall that  $G(\gamma) > 0$  occurs when water is transformed into regions with larger  $\gamma$ . For example, net surface cooling in the presence of a positive thermal expansion coefficient ( $\alpha > 0$ ) leads to  $Q_\Theta^{\text{non-adv}} + (\Theta_m - \Theta) Q_m < 0$ . Such cooling then leads to a positive contribution to  $G(\gamma)^{\text{surf}}$  as water is transformed from light to heavy  $\gamma$ -classes. Likewise, a positive net salt transport into the upper ocean,  $Q_S^{\text{nonadv}} + (S_m - S) Q_m > 0$ , leads to a positive contribution to  $G(\gamma)^{\text{surf}}$ .

The integrand to equation (20.115) corresponds to minus the surface buoyancy flux derived in Section 19.6.3. The only difference is that we here make use of the surface element,  $d\mathcal{S}$ , and the corresponding fluxes  $Q_\Theta^{\text{non-adv}}$ ,  $Q_S^{\text{nonadv}}$ , and  $Q_m$ . However, if the ocean surface has no overturns, we can write its vertical position as  $z = \eta(x, y, t)$  and can also define the horizontal projection of the area element as

$$d\mathcal{S} = \sqrt{1 + |\nabla \eta|^2} dA. \quad (20.117)$$

In this case we can introduce the fluxes  $Q_\Theta^{\text{non-adv}}$ ,  $Q_S^{\text{non-adv}}$ , and  $Q_m$  used in Section 19.6.3 via

$$Q_\Theta^{\text{non-adv}} d\mathcal{S} = Q_\Theta^{\text{non-adv}} dA \quad (20.118a)$$

$$Q_S^{\text{nonadv}} d\mathcal{S} = Q_S^{\text{non-adv}} dA \quad (20.118b)$$

$$Q_m d\mathcal{S} = Q_m dA, \quad (20.118c)$$

<sup>9</sup>For Boussinesq oceans, we replace the factor of  $\gamma^{-1}$  with  $\rho_b^{-1}$  on the right hand side of equation (20.116) with  $\rho_b$ , the constant Boussinesq reference density.

to render

$$G(\gamma)^{\text{surf}} = \lim_{\delta\gamma \rightarrow 0} \frac{1}{\delta\gamma} \int_{\partial\Omega_{\text{out}}(\gamma \pm \delta\gamma/2)} (\gamma \beta [Q_S^{\text{non-adv}} + (S_m - S) Q_m] - \gamma \alpha [Q_\Theta^{\text{non-adv}} + (\Theta_m - \Theta) Q_m]) dA. \quad (20.119)$$

Integrating the surface transformation (20.119) over all  $\gamma$ -classes leads to the identity

$$\int_{\gamma=-\infty}^{\gamma=\infty} G(\gamma)^{\text{surf}} d\gamma = \int_{z=\eta} \left( \gamma \beta [Q_S^{\text{non-adv}} + (S_m - S) Q_m] - \gamma \alpha [Q_\Theta^{\text{non-adv}} + (\Theta_m - \Theta) Q_m] \right) dA. \quad (20.120)$$

This equality means that the diagnosed surface transformation,  $G(\gamma)^{\text{surf}}$ , which is obtained by binning surface fluxes into  $\gamma$ -classes, must properly add up to the area integrated surface fluxes as weighted by the surface value of  $\gamma$ . This equality is a useful check on the integrity of numerical binning code used to diagnose surface water mass transformation.

## 20.8 Tracer water mass transformation within $\gamma$ -layers

In Sections 20.5 and 20.6 we developed the budgets for  $\lambda$  within layers defined  $\lambda$ . Here we extend that analysis to develop budgets for a tracer concentration,  $C$ , within a region within a layer of buoyancy,  $\gamma$ , as depicted in Figure 20.15. The upper panels to this figure illustrate a tracer patch in geographic/depth  $x$ -space along with isolines of buoyancy, whereas the lower panels show the tracer distribution (histogram) binned within the buoyancy classes ( $q$ -space). If the tracer is mixed within a layer, such as via the neutral diffusion process of Section 18.4, then the tracer patch is spread laterally within the buoyancy layer and yet the distribution (lower panel) is unchanged. In contrast, if the tracer is mixed across layer interfaces (dianeutral transport) then the tracer distribution is spread within buoyancy space.

Another means to alter the tracer distribution is to modify the buoyancy field. This situation is especially common for tracer near the surface, where boundary buoyancy forcing can act to move the layers thus causing tracer to move between layers even if the tracer patch is stationary in  $x$ -space (e.g., see Figures 20.13 and 20.14). That is, if the tracer moves at a velocity that is distinct from the velocity of buoyancy surfaces, then that relative motion changes the tracer's distribution within buoyancy classes.

### 20.8.1 General form of the mass budget

As depicted in Figure 20.15 for buoyancy layers, and Figure 20.16 for generic layers, there are two general processes whereby a tracer distribution within layers can be modified: (i) the tracer can mix between layers and (ii) the layers can move relative to the tracer. These ideas transcend buoyancy and thus can be applied to any scalar field,  $\lambda$ , used to classify water masses. We quantify these two processes by writing the time change of tracer content within a  $\lambda$ -layer, which is arrived at by applying the Leibniz-Reynolds transport theorem to a  $\lambda$ -layer

$$\frac{d}{dt} \Delta M_C(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho \dot{C} dV - \oint_{\partial\Omega(\lambda \pm \Delta\lambda/2)} \rho C (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (20.121)$$

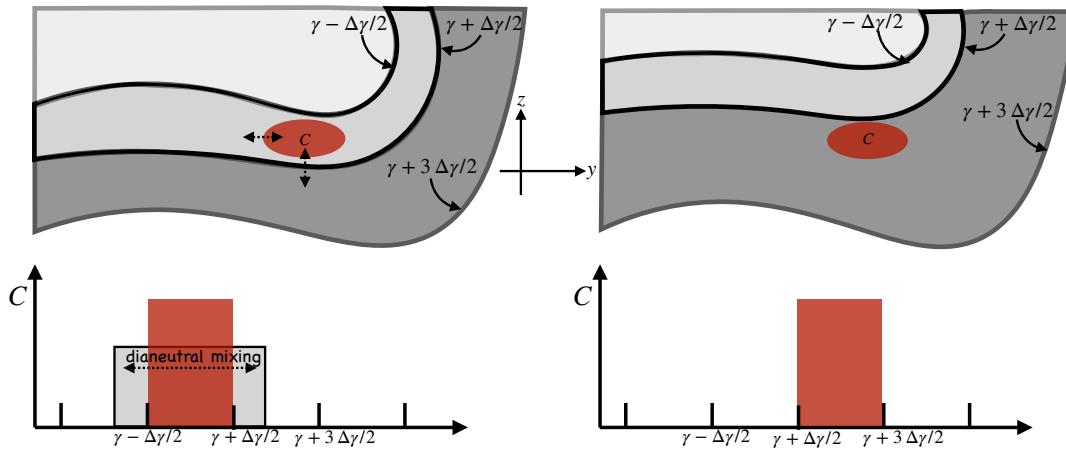


FIGURE 20.15: Depicting a tracer patch within the buoyancy layer bounded by the interface values  $[\gamma - \Delta\gamma/2, \gamma + \Delta\gamma/2]$  (left panel) and  $[\gamma + \Delta\gamma/2, \gamma + 3\Delta\gamma/2]$  (right panel). The upper panels show the tracer and buoyancy in geographic/depth  $x$ -space whereas the bottom panels show the tracer distribution (histogram) binned according to buoyancy ( $q$ -space). There are two general means to modify the distribution of tracer within the buoyancy classes. The first occurs via dianeutral mixing that spreads the tracer distribution to other buoyancy layers as depicted by the vertical arrow in the upper left panel and the horizontal arrows in the lower left panel. The lateral arrow in the upper left panel depicts neutral diffusion, which laterally spreads the tracer within a layer but does not alter the distribution across layers (see Section 18.4). The second means to alter the distribution occurs when the buoyancy surfaces move relative to the tracer. This scenario is depicted in the lower right panel whereby the tracer patch originally in buoyancy layer  $[\gamma - \Delta\gamma/2, \gamma + \Delta\gamma/2]$  now finds itself in the layer  $[\gamma + \Delta\gamma/2, \gamma + 3\Delta\gamma/2]$ . This depiction is not realistic, since motion of interior buoyancy surfaces generally occurs along with mixing of tracer patches. Nonetheless, this example emphasizes that motion of the buoyancy surfaces need not precisely coincide with motion of the tracer patch.

where we wrote the mass of tracer  $C$  within the  $\lambda$ -layer

$$\Delta M_C(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho C \, dV. \quad (20.122)$$

The volume integral on the right hand side of equation (20.121) arises from material time changes to the tracer within the layer, whereas the surface integral arises from dia-surface transport across the layer boundary.

### 20.8.2 Tracer processes

We determine the material time changes for a conservative tracer according to the convergence of a flux

$$\rho \dot{C} = \rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}. \quad (20.123)$$

Many biogeochemical tracers have additional source terms beyond the flux convergence considered here. As in Section 20.4.4, sources can be readily incorporated into the following by adding a source tendency term that acts throughout a layer.

The divergence theorem converts the convergence,  $-\nabla \cdot \mathbf{J}$ , into the area integral of fluxes over the layer boundaries, including interior layer interfaces as well as intersections with the surface and bottom boundaries. For the interior interfaces it is typically simpler in practice to bin the volume integrated material time changes within the  $\lambda$ -classes. In contrast, the surface and bottom boundary contributions are fed into the budget via Neumann boundary conditions applied to the flux  $\mathbf{J}$

$$\mathbf{J} \cdot \hat{\mathbf{n}} \, dS = \text{boundary tracer transport}. \quad (20.124)$$

Note that when there is an advective/skew diffusive component to the subgrid scale flux (Chapters 17 and 18), then it adds to the resolved advective component to render a residual mean material time operator

$$\rho \frac{D^\dagger C}{Dt} = -\nabla \cdot \mathbf{J}^{\text{non-adv}}, \quad (20.125)$$

where

$$\frac{D^\dagger}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v} + \mathbf{v}^*) \cdot \nabla, \quad (20.126)$$

with  $\mathbf{v}^*$  an eddy-induced velocity (see Section 18.1). For the purposes of water mass transformation analysis, we write

$$\dot{C} = \frac{D^\dagger C}{Dt}, \quad (20.127)$$

thus incorporating the eddy-induced advection into the kinematic expression for the material time derivative.

There are many interior and boundary processes that contribute to  $\dot{C}$  within a layer. We here write the following as a general expression for these contributions to the layer budget

$$\Delta E_c(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho \dot{C} dV = - \int_{\Omega(\lambda \pm \Delta\lambda/2)} \nabla \cdot \mathbf{J} dV, \quad (20.128)$$

which is sometimes usefully decomposed into interior and surface boundary processes

$$\Delta E_c^{\text{int}}(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho \dot{C}^{\text{int}} dV \quad (20.129a)$$

$$\Delta E_c^{\text{out}}(\lambda \pm \Delta\lambda/2) = - \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (20.129b)$$

If the region boundary intersects the ocean bottom along  $\partial\Omega_{\text{bot}}(\lambda \pm \Delta\lambda/2)$ , then there is an additional bottom boundary contribution in the form

$$\Delta E_c^{\text{bot}}(\lambda \pm \Delta\lambda/2) = - \int_{\partial\Omega_{\text{bot}}(\lambda \pm \Delta\lambda/2)} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (20.130)$$

### 20.8.3 Transport across an interior layer interface

The surface integral in the budget (20.121) involves transport across the layer interfaces, with this transport requiring motion of the interface relative to a fluid particle. The same formalism introduced earlier can be used to compute this transport. That is, we can generalize the transformation equation (20.46e) to write

$$G_c(\lambda) = \int_{\partial\Omega(\lambda)} \rho C (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \frac{\partial}{\partial \lambda} \int_{\Omega(\lambda_0 \leq \lambda)} \rho \dot{\lambda} C dV. \quad (20.131)$$

As a check, note that for the special case where the tracer concentration is a constant along the layer interface, then  $G_c(\lambda) = C G(\lambda)$ . We consider this special case in Section 20.9 when studying budgets over regions bounded by a tracer isosurface.

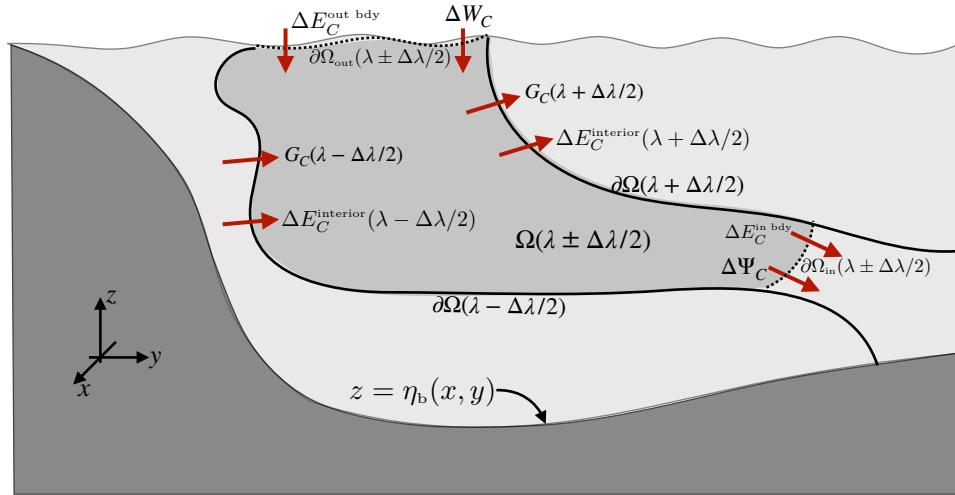


FIGURE 20.16: As for the schematic of a layer fluid mass budget depicted in Figure 20.6, we here illustrate the budget for a tracer,  $C$ , within a layer defined by the scalar property,  $\lambda$ , within the range  $[\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$  and defined over a geographical/depth domain  $\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2) + \partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2) + \partial\Omega(\lambda + \Delta\lambda/2) + \partial\Omega(\lambda - \Delta\lambda/2)$ . The budget for a tracer,  $C$ , over this layer is affected by the transport of tracer substance across the variety of layer boundaries. Transport processes include those determined by mixing and/or radiation across interior and surface boundaries,  $\Delta E_C$  (equation (20.129b)). This term has no associated transfer of fluid mass and thus is absent from the water mass budget in Figure 20.6. Tracer budgets are also affected by processes that move fluid mass across layer boundaries: water mass transformation processes giving rise to  $G_C(\lambda \pm \Delta\lambda/2)$  (equation (20.131)); transport across the surface domain boundary,  $\Delta W_C$ , arising from precipitation, evaporation, runoff, and melt (equation (20.132)); and transport within the circulation crossing an interior domain boundary,  $\Delta\Psi_C$  (equation (20.133)).

#### 20.8.4 Transport across interior and surface boundaries

We now consider the impact on layer tracer mass budgets due to boundary transport. The budget contribution from mass fluxes crossing the ocean surface boundary is determined by making use of the surface [kinematic boundary condition](#)

$$\Delta W_C = \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} \rho C (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} \, dS = \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} Q_m C_m \, dA, \quad (20.132)$$

where  $C_m$  is the tracer concentration within the mass transported across the boundary.<sup>10</sup> As a check, note that in the special case of a constant tracer concentration in the mass transported across the boundary, then  $\Delta W_C = C_m \Delta W$ , where  $\Delta W$  is the water mass transported across the ocean free surface as given by equation (20.66).

For the interior open boundary the contribution is written in the generic manner

$$\Delta\Psi_C = \int_{\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)} C \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} \, dS. \quad (20.133)$$

Again, in the special case where the tracer concentration is a constant,  $C_b$ , along the interior boundary, then  $\Delta\Psi_C = C_b \Delta\Psi$ , where  $\Delta\Psi$  is the fluid mass transport given by equation (20.64).

<sup>10</sup>Note that equation (26) in [Groeskamp et al. \(2019\)](#) incorrectly writes the integrand in equation (20.132) as  $Q_m (C_m - C)$ .

### 20.8.5 The layer tracer budget

Bringing all terms together leads to the layer tracer mass budget

$$\frac{d\Delta M_c}{dt} + \Delta\Psi_c = \Delta E_c + \Delta W_c - [G_c(\lambda + \Delta\lambda/2) - G_c(\lambda - \Delta\lambda/2)], \quad (20.134)$$

which is directly analogous to the fluid layer mass budget (20.68), with the added term  $\Delta E_c$  arising from material tracer changes. As for the fluid mass budget discussed in Section 20.5.3, the layer tracer budget (20.134) provides the framework for rather general inferences about tracer transport within  $\lambda$ -classes.

### 20.8.6 Further study

Much in this section follows the treatment given by [Groeskamp et al. \(2019\)](#), which offers specific examples of tracer mass analysis.

## 20.9 Regions bounded by a tracer contour/surface

In Section 20.8 we developed layer tracer budget equations where the scalar field,  $\lambda$ , that defines the layer is generally distinct from the tracer,  $C$ , whose budget is examined. In this section we specialize to the case where  $\lambda = C$  so that the region boundaries are determined by the tracer whose budget is under study. These budgets were introduced in Sections 20.5 and 20.6, and here we derive some further results.

As in Section 20.8, our starting point is the Leibniz-Reynolds transport theorem as applied to fluid mass and tracer mass computed over an arbitrary region,  $\mathcal{R}$ ,

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho C dV \right] = - \int_{\partial\mathcal{R}} [\rho C (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} dS \quad (20.135a)$$

$$\frac{d}{dt} \left[ \int_{\mathcal{R}} \rho dV \right] = - \int_{\partial\mathcal{R}} [\rho (\mathbf{v} - \mathbf{v}^{(b)})] \cdot \hat{\mathbf{n}} dS. \quad (20.135b)$$

The region,  $\mathcal{R}$ , is arbitrary and can in general be disconnected. Throughout this section we make use of the following shorthand notation for region-integrated quantities

$$M = \int_{\mathcal{R}} \rho dV \quad \text{region fluid mass} \quad (20.136a)$$

$$M_c = \int_{\mathcal{R}} C \rho dV \quad \text{region tracer mass} \quad (20.136b)$$

$$\langle C \rangle = \frac{1}{M} \int_{\mathcal{R}} C \rho dV = \frac{M_c}{M} \quad \text{region averaged tracer concentration.} \quad (20.136c)$$

### 20.9.1 Closed region bounded by a tracer surface/contour

Consider a closed region,  $\tilde{\mathcal{R}}$ , bounded by a surface of constant tracer concentration,  $C = \tilde{C}$ , such as depicted in Figure 20.17. The tracer budget (20.135a) for this region is given by

$$\frac{d(M \langle C \rangle)}{dt} = -\tilde{C} \int_{\partial\tilde{\mathcal{R}}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS - \int_{\partial\tilde{\mathcal{R}}} \mathbf{J} \cdot \hat{\mathbf{n}} dS, \quad (20.137)$$

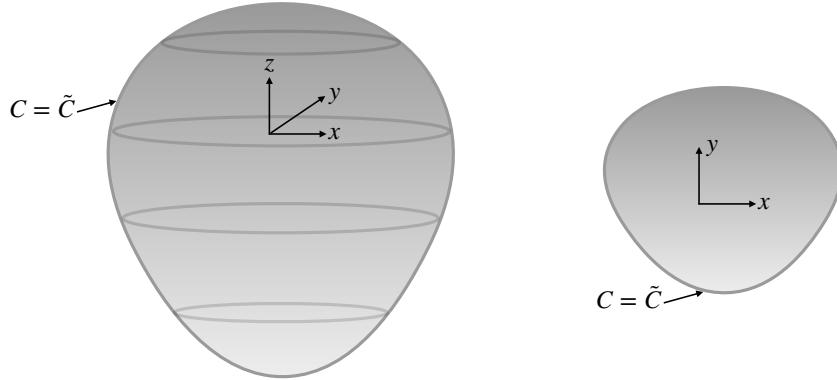


FIGURE 20.17: Left panel: a closed three-dimensional region,  $\tilde{\mathcal{R}}$ , with its boundary,  $\partial\tilde{\mathcal{R}}$ , defined by a surface of constant tracer concentration,  $C = \tilde{C}$ . Right panel: the analog closed two-dimensional region with its boundary defined by a contour of constant tracer concentration,  $C = \tilde{C}$ .

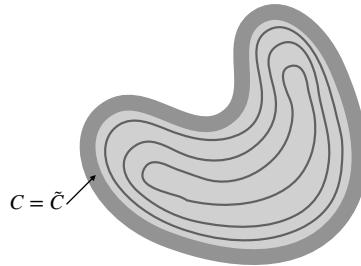


FIGURE 20.18: A two-dimensional region bounded by a finite-thick shell with constant tracer concentration,  $C = \tilde{C}$ . Inside the shell region the tracer concentration is not uniform.

where we pulled the tracer concentration outside of the surface integral since, by construction, it is constant on the boundary,  $\partial\tilde{\mathcal{R}}$ . Use of the mass budget (20.135b) then leads to the rather tidy result

$$\frac{d[M(\langle C \rangle - \tilde{C})]}{dt} = - \int_{\partial\tilde{\mathcal{R}}} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (20.138)$$

The left hand side is the time change of the mass-weighted difference between the region averaged tracer concentration,  $\langle C \rangle$ , and the value of the tracer concentration defining the region boundary,  $\tilde{C}$ . These time changes are driven by a nonzero diffusive tracer transport bringing tracer mass across the region boundary. A nonzero diffusive flux on the region boundary arises only when there is a gradient of tracer concentration across that boundary. In the special case of a zero net diffusive tracer transport across the region boundary, the budget equation (20.138) reaches a steady state whereby

$$\frac{d}{dt} \left[ M(\langle C \rangle - \tilde{C}) \right] = 0 \iff \int_{\partial\tilde{\mathcal{R}}} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S} = 0. \quad (20.139)$$

A three-dimensional region bounded by a constant tracer concentration is not commonly encountered in large-scale stratified ocean and atmospheric fluids. In contrast, we often encounter quasi-two-dimensional regions as depicted in Figure 20.18, in which one may find two-dimensional surfaces bounded by a closed contour of constant tracer concentration. For example, in many parts of the ocean and atmosphere, transport occurs predominantly along two-dimensional surfaces defined by a constant buoyancy. We may thus find closed contours of tracer concentrations along constant buoyancy surfaces.

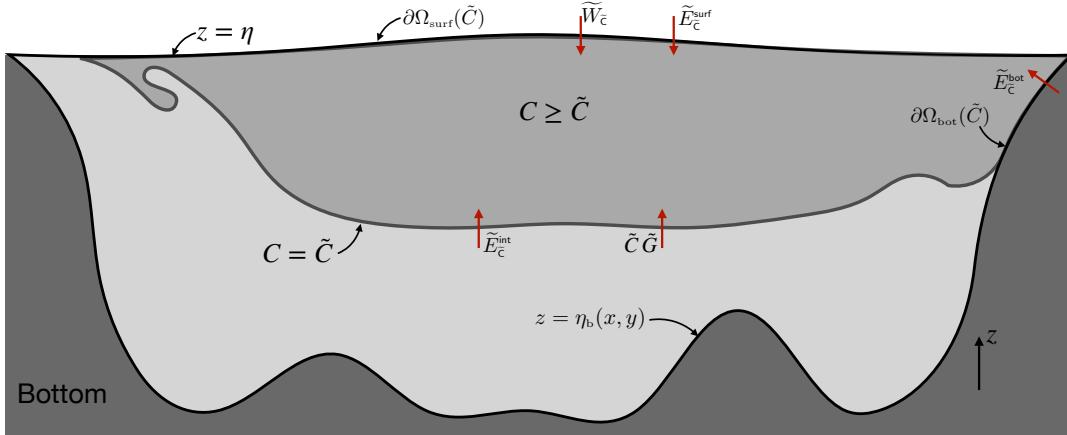


FIGURE 20.19: An ocean region where the tracer concentration is greater than a nominal value,  $C \geq \tilde{C}$ . A specific example is with  $C = \Theta_{\zeta}$ , the Conservative Temperature, in which we are concerned with the ocean with temperature greater than  $\Theta$ . Here we depict a case where the tracer concentration generally increases upward (as with  $C = \Theta$ ), and yet with vertical stratification not everywhere monotonic, such as for  $C = \Theta$  in the high latitudes where salinity effects on density stratification become dominant. Transport processes affecting the budget of  $C$  within this region arise from mixing at the interior boundary and surface boundary,  $\tilde{E}_{\tilde{C}}^{\text{int}}$  and  $\tilde{E}_{\tilde{C}}^{\text{surf}}$ , advection at the surface,  $\tilde{W}_{\tilde{C}}$ , and the tracer weighted water mass transformation across the interior layer boundary,  $\tilde{C} \tilde{G}$ . Arrows are oriented in which a positive value for the corresponding term adds tracer to the region.

To help illustrate a necessary condition to reach a steady state, consider the particular example depicted in Figure 20.18. In this figure, the tracer contour defining the region boundary is a thick shell defined by a uniform concentration  $C = \tilde{C}$ . The diffusive flux vanishes at each point within the boundary shell since the tracer concentration is uniform. Hence, the steady budget (20.139) leads to

$$(\langle C \rangle - \tilde{C}) \frac{dM}{dt} + M \frac{d\langle C \rangle}{dt} = 0. \quad (20.140)$$

If the total fluid mass within the region is constant, then the averaged tracer concentration is also constant, so that both terms in this steady budget vanish individually. Even so, this configuration does not reach a steady state at each point throughout the domain interior. The reason is that diffusion in the interior causes tracer to move from regions of high concentration to low concentration. Consequently, at any particular point within the domain there is an evolving tracer concentration. The only way for each point to reach a steady state within a region bounded by a tracer contour is for the tracer concentration to be a uniform constant throughout the region interior

$$C = \tilde{C} \quad \text{steady state tracer throughout a closed tracer region.} \quad (20.141)$$

Diffusion thus expels tracer gradients from steady state regions bounded by closed tracer contours, thus leaving a homogenous interior. We proved this same result from a different perspective in Section 16.8. It is satisfying to see this result follow from the present formalism based on Leibniz-Reynolds.

## 20.9.2 Region with $C \geq \tilde{C}$

As a second example, consider the tracer budget for a region where the tracer concentration is greater than or equal to a particular tracer value, such as depicted in Figure 20.19. In contrast

to the domain in Figure 20.16, here there is no inner boundary. To develop the fluid mass budget and the tracer mass budget, we introduce the fluid mass and tracer mass for the region with  $C \geq \tilde{C}$

$$\widetilde{M} = \int_{C \geq \tilde{C}} \rho dV \quad (20.142a)$$

$$\widetilde{M}_{\bar{C}} = \int_{C \geq \tilde{C}} C \rho dV = \widetilde{M} \langle C \rangle. \quad (20.142b)$$

We also make use of terms arising from water mass transformation across the  $\tilde{C}$  interface

$$\tilde{G} = - \int_{C=\tilde{C}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS = - \int_{C=\tilde{C}} \rho w^{\text{dia}} dS \quad (20.143a)$$

$$\tilde{G}_{\bar{C}} = - \int_{C=\tilde{C}} C \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS = - \int_{C=\tilde{C}} C \rho w^{\text{dia}} dS, \quad (20.143b)$$

as well as terms arising from ocean surface boundary mass transport

$$\widetilde{W} = - \int_{\partial\Omega_{\text{surf}}(\tilde{C})} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} Q_m dA \quad (20.144a)$$

$$\widetilde{W}_{\bar{C}} = - \int_{\partial\Omega_{\text{surf}}(\tilde{C})} C \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS = \int_{\partial\Omega_{\text{surf}}(\tilde{C})} C Q_m dA. \quad (20.144b)$$

Finally, we need terms arising from subgrid scale transport across the  $\tilde{C}$  interface, ocean surface, and ocean bottom

$$\tilde{E}_{\bar{C}}^{\text{int}} = - \int_{C=\tilde{C}} \mathbf{J} \cdot \hat{\mathbf{n}} dS \quad (20.145a)$$

$$\tilde{E}_{\bar{C}}^{\text{surf}} = - \int_{\partial\Omega_{\text{surf}}(\tilde{C})} \mathbf{J} \cdot \hat{\mathbf{n}} dS \quad (20.145b)$$

$$\tilde{E}_{\bar{C}}^{\text{bot}} = - \int_{\partial\Omega_{\text{bot}}(\tilde{C})} \mathbf{J} \cdot \hat{\mathbf{n}} dS. \quad (20.145c)$$

Recall that  $\hat{\mathbf{n}}$  is the outward normal on a boundary so that positive values for the transports (20.145a)-(20.145c) increase the tracer mass within the region. For equations (20.143a) and (20.143b), we introduced the dia-surface transport velocity according to equation (20.38) for flow across the  $C = \tilde{C}$  layer interface, with  $\hat{\mathbf{n}} = \nabla C / |\nabla C|$  for these transports. Likewise, for equations (20.144a) and (20.144b) we made use of the surface kinematic boundary condition

$$\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} dS = -Q_m dA, \quad (20.146)$$

where  $Q_m$  is the mass transport across the free surface, with  $Q_m > 0$  adding mass to the ocean, and  $dA$  is the horizontal projection of the surface area element. By inspection of Figure 20.19, the fluid mass budget and tracer mass budget for this region are given by

$$\frac{d\widetilde{M}}{dt} = \tilde{G} + \widetilde{W} \quad (20.147a)$$

$$\frac{d[\widetilde{M} \langle C \rangle]}{dt} = \tilde{C} \tilde{G} + \widetilde{W}_{\bar{C}} + \tilde{E}_{\bar{C}}^{\text{surf}} + \tilde{E}_{\bar{C}}^{\text{bot}} + \tilde{E}_{\bar{C}}^{\text{int}}, \quad (20.147b)$$

where we assumed that no mass crosses through the solid earth. Furthermore, along the  $\tilde{C}$ -boundary we pulled the tracer concentration outside of the surface integral to write  $\tilde{G}_{\bar{C}} = \tilde{C} \tilde{G}$ .

Just as we did in Section 20.9.1, the tracer budget (20.147b) can be simplified by making use of the fluid mass budget (20.147a) to eliminate the water mass transformation contribution  $\tilde{G}$ , thus rendering

$$\frac{d\tilde{M}_{\bar{C}}}{dt} = [\tilde{W}_{\bar{C}} - \tilde{W} \tilde{C}] + \tilde{E}_{\bar{C}}^{\text{surf}} + \tilde{E}_{\bar{C}}^{\text{bot}} + \tilde{E}_{\bar{C}}^{\text{int}}. \quad (20.148)$$

In this equation we introduced the *internal tracer mass* according to

$$\tilde{M}_{\bar{C}} \equiv \tilde{M} (\langle C \rangle - \tilde{C}) = \int_{C \geq \tilde{C}} (C - \tilde{C}) \rho dV. \quad (20.149)$$

For completeness we express the internal tracer mass budget (20.148) in its integral form

$$\frac{d\tilde{M}_{\bar{C}}}{dt} = \int_{\partial\Omega_{\text{out}}} [Q_m (C - \tilde{C}) dA - \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}] - \int_{\partial\Omega_{\text{bot}}} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S} - \int_{C=\tilde{C}} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (20.150)$$

### 20.9.3 Comments and further study

Elimination of the water mass transformation,  $\tilde{G}$ , from the internal tracer mass budget equations (20.148) and (20.150) offers a practical advantage since  $\tilde{G}$  can be rather noisy in applications. Furthermore, for some applications (e.g., see [Holmes et al. \(2019b\)](#)) it is sufficient to consider the simpler budget (20.148) for internal tracer mass, rather than the budget (20.147b) for the total tracer mass.



## 20.10 Exercises

### EXERCISE 20.1: TRANSFORMATION FROM DIFFUSION

We here express the equations for water mass transformation for a tracer,  $\lambda$ , satisfying

$$\rho \frac{D\lambda}{Dt} = -\nabla \cdot \mathbf{J}, \quad (20.151)$$

where  $\mathbf{J}$  is a downgradient diffusive flux

$$\mathbf{J} = -\rho \kappa \nabla \lambda, \quad (20.152)$$

with  $\kappa > 0$  a kinematic diffusivity. Furthermore, assume  $\lambda = \lambda(z, t)$  so that  $\nabla \lambda = \partial_z \lambda \hat{z}$ . Correspondingly, all  $\lambda$  surfaces are horizontal so that transformation across a  $\lambda$  surface is transformation across a constant  $z$  surface, which we write as  $z = z(\lambda)$ .

- (a) Write the expression for the interior water mass transformation,  $G(\lambda)$ , arising from downgradient vertical diffusion. Write the form as a surface integral over the horizontal surface whose vertical position is  $z = z(\lambda)$ .
- (b) Write the equivalent form for  $G(\lambda)$  as a volume integral, following the methods from Section 20.4.3, and assume the volume integral is over a layer with thickness  $\delta z$  surrounding the  $z = z(\lambda)$  horizontal surface.

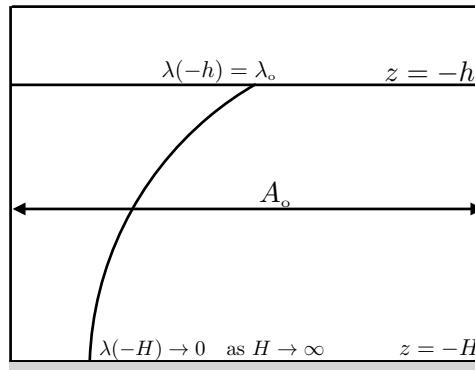


FIGURE 20.20: Depicting the setup for Exercise 20.2, in which we examine water mass transformation and formation for a tracer,  $\lambda$ , in steady state advection-diffusion balance. The tracer takes the boundary values,  $\lambda(-h) = \lambda_o$  and  $\lim_{H \rightarrow \infty} \lambda(-H) = 0$ . The domain has vertical sidewalls with horizontal area  $A_o$ , and it extends over the vertical range,  $-H < z \leq h$ , where  $H \rightarrow \infty$ .

- (c) Write the volume integral from the previous part as the difference of two surface integrals over constant  $z$  surfaces.
- (d) Assume a steady state and write the water mass transformation in terms of the vertical velocity as integrated over the  $z = z(\lambda)$  horizontal surface.

**EXERCISE 20.2: TRANSFORMATION AND FORMATION WITH STEADY ADVECTION-DIFFUSION**  
 Consider a steady vertical advection-diffusion balance for a tracer,  $0 \leq \lambda \leq \lambda_o$ , in a Boussinesq ocean:

$$w \partial_z \lambda = \kappa \partial_{zz} \lambda, \quad (20.153)$$

where  $\kappa > 0$  is a constant kinematic diffusivity and  $w = w_o > 0$  is an upwelling velocity that is constant throughout the domain. As in Figure 20.20, assume the vertical extent of the domain is  $-H < z \leq -h$ , with  $h > 0$  a finite depth such as the base of the surface boundary layer, and with  $H \gg h$  so that we can approximate  $H \rightarrow \infty$ . Furthermore, assume the domain has vertical sidewalls with a constant horizontal cross-sectional area,  $A_o$ . Let  $\lambda(z = -h) = \lambda_o$  and  $\lim_{H \rightarrow \infty} \lambda(z = -H) = 0$ . With a domain that is infinitely deep (in the limit with  $H \rightarrow \infty$ ), there is an infinite reservoir of water that feeds the constant upwelling. To maintain volume conservation, we assume water leaves the domain through  $z = -h$  at a rate equal to  $w_o$ , thus entering the surface boundary layer where flow within the boundary layer is not of interest for this exercise.<sup>11</sup>

- (a) Find the steady state solution,  $\lambda(z)$ , to equation (20.153) within the domain  $-H < z \leq -h$  in the limit that  $H \rightarrow \infty$ , and that satisfies the stated boundary conditions.
- (b) What is the downgradient diffusive flux of  $\lambda$ ? What is its direction?
- (c) What is the advective flux of  $\lambda$ ? What is its direction?
- (d) What is the dia-surface flux,  $w^{\text{dia}}(\lambda)$ ? What is its direction?
- (e) What is the water mass formation per unit  $\lambda$  increment,  $\mathcal{F} = -\partial G / \partial \lambda$ ? Hint: assume

<sup>11</sup>If we consider flow within the upper boundary layer,  $-h \leq z$ , then we might wish to set the boundary condition,  $w(z = 0) = 0$ , which is the rigid lid assumption, or to examine a nonzero transport across the ocean surface via precipitation, evaporation, or runoff. However, extending the domain to the ocean surface would make the exercise far more complicated. Namely, in that case  $w$  cannot be assumed to be a constant throughout the domain, nor can we assume  $u$  and  $v$  are zero. To avoid such complexities, and to allow us to focus on the water mass transformation equations, we restrict attention to the region below the boundary layer where we can maintain self-consistency with  $w(z) = w_o$  and  $\lambda = \lambda(z)$ .

the cross-sectional area of the domain is given by<sup>12</sup>

$$A(z) = A_\circ [\lim_{H \rightarrow \infty} \mathcal{H}(z + H) - \mathcal{H}(z)], \quad (20.154)$$

where  $\mathcal{H}$  is the Heaviside step function given by

$$\mathcal{H}(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1/2 & \text{if } z = 0 \\ 1 & \text{if } z > 0, \end{cases} \quad (20.155)$$

and whose derivative is the Dirac delta

$$d\mathcal{H}(z)/dz = \delta(z). \quad (20.156)$$

- (f) What is the net water mass formation over the domain,  $\int_0^{\lambda_\circ} \mathcal{F} d\lambda$ ?

**EXERCISE 20.3: TRANSFORMATION AND FORMATION FROM DIFFUSION IN A STAGNANT FLUID**  
In this exercise we examine how diffusion affects water mass transformation within a stagnant fluid ( $v = 0$ ) and with the tracer concentration a function of vertical position and time,  $\lambda = \lambda(z, t)$  with  $\partial_z \lambda > 0$ . Assuming a Boussinesq ocean and with a constant kinematic diffusivity,  $\kappa > 0$ , we have the diffusion equation

$$\partial_t \lambda = \kappa \partial_{zz} \lambda. \quad (20.157)$$

As in Exercises 20.1 and 20.2, all  $\lambda$  surfaces are horizontal so that transformation across a  $\lambda$  surface is transformation across a constant  $z$  surface. Furthermore, as in Exercise 20.2, assume the domain has vertical sidewalls so that the horizontal cross-sectional area is

$$A(z) = A_\circ [\lim_{H \rightarrow \infty} \mathcal{H}(z + H) - \mathcal{H}(z)], \quad (20.158)$$

where  $\mathcal{H}$  is the Heaviside step function (20.155).

- (a) Write the expression for the water mass transformation,  $G(\lambda)$ , arising from downgradient diffusion. Write it as a surface integral over the horizontal  $z = z(\lambda)$  surface.
- (b) Write the water mass transformation,  $G(\lambda)$ , for a tracer having an exponential vertical profile

$$\lambda = \lambda_\circ e^{z/d}, \quad (20.159)$$

where  $\lambda_\circ = \lambda(z = 0)$  and  $d > 0$  is a constant depth scale.

- (c) What is the velocity of a point fixed on a  $\lambda$ -surface, again assuming the exponential profile (20.159)? Hint: remember that the fluid is assumed to be stagnant.
- (d) Write the water mass formation corresponding to the exponential tracer profile (20.159). Hint: this part is where we need to make use of the area expression (20.158).

#### EXERCISE 20.4: INTEGRATION BETWEEN TWO CLOSED TRACER CONTOURS

This exercise introduces some ideas of use for determining processes affecting the transport of matter across a tracer contour. Note that in general, the tracer concentration is a function of time. However, the present suite of questions concerns the instantaneous geometry of the tracer field, so that time dependence is not considered. Elements of this exercise follow from

---

<sup>12</sup>The expression (20.154) for the horizontal area can be considered the limit of a bowl domain becoming a vertical sidewall domain. For the bowl we have  $\partial_z A \neq 0$ , which has its generalization to the hypsometry of the ocean's bottom topography. See [de Lavergne et al. \(2016\)](#) for a discussion of its role in water mass processes.

the discussion from Section 20.3.1.

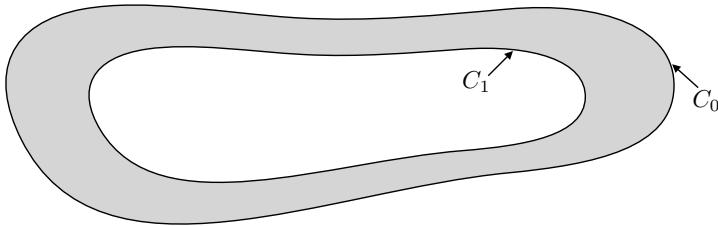


FIGURE 20.21: Illustrating the area contained between two closed tracer contours,  $C_0 \leq C(x, y, t) \leq C_1$ . Exercise 20.4 develops some mathematical expressions for integration within this area, with the resulting expressions of use for the analyses of tracer transport.

- (a) Consider a closed two-dimensional region bounded by two contours of tracer concentration,  $C_0 \leq C(x, y, t) \leq C_1$ , such as shown in Figure 20.21. Derive the following expression for the area enclosed by the two contours

$$\mathcal{A} = \int_{C_0}^{C_1} dC \oint \frac{dl}{|\nabla C|}. \quad (20.160)$$

In this expression,  $dl$  is the line element for a path taken in a counter-clockwise direction along a contour of constant  $C$ . We also assume the tracer concentration is not uniform in the region of interest so that  $\nabla C \neq 0$ .

- (b) As a corollary, show that for

$$\mathcal{A}(C) = \int_{C_0}^C dC' \oint \frac{dl}{|\nabla C'|} \quad (20.161)$$

we have the identity

$$\frac{\partial \mathcal{A}(C)}{\partial C} = \oint \frac{dl}{|\nabla C|}. \quad (20.162)$$

In words, this result means that the area between two tracer contours has a partial derivative, with respect to the tracer contour, equal to the line integral on the right hand side. The area per  $C$  is smaller in regions where the tracer gradient is larger; i.e., there is less area “concentration” in regions of strong tracer gradient.

- (c) Use the above two results to prove the following form of the Fundamental Theorem of Calculus

$$\frac{\partial}{\partial C} \left[ \int \Phi(x) d\mathcal{A} \right] = \frac{\partial}{\partial C} \left[ \int_{C_0}^C dC' \oint \frac{\Phi dl}{|\nabla C'|} \right] = \oint \frac{\Phi dl}{|\nabla C|}, \quad (20.163)$$

with  $\Phi$  an arbitrary function. This identity is analogous to that derived in Section 20.3.5. It has many useful applications such as those discussed in [Marshall et al. \(2006\)](#).

#### EXERCISE 20.5: CAUCHY’S DOUBLE INTEGRAL FORMULA

In this exercise we derive the following double-integral identity, originally due to [Cauchy \(1823\)](#),

$$\int_{t_n}^{t_{n+1}} \left[ \int_{t_n}^t \mathcal{B}(s) ds \right] dt = \int_{t_n}^{t_{n+1}} (t_{n+1} - t) \mathcal{B}(t) dt, \quad (20.164)$$

where we make reference to Figure 20.22 for the notation. To prove this identity, it is useful to

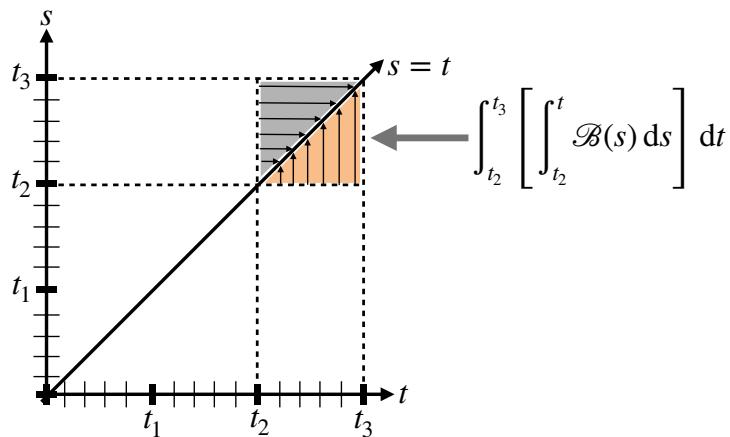


FIGURE 20.22: Gold region depicts the time integration domain used in one of the double integrals from equation (20.164) for the special case of  $n = 2$ . Note that the gray triangular region generally leads to a distinct integral.

make the substitution

$$A(s) = \int_{s_0}^s B(s') ds' \implies \mathcal{B}(s) = dA/ds, \quad (20.165)$$

and then work separately on the two sides of equation (20.164) to show that they are the same.





## **Part VI**

### **Solutions to Exercises**

We conclude each chapter of this book with a suite of exercises, and here we present solutions to the 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. Namely, the solutions presented here aim 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, exposing details helps to identify when an incorrect result follows from a physical/conceptual error (e.g., incorrect setup of the problem solution) or from 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.

## General pointers

Here are a few general pointers that can serve to motivate problem solving, either for class work or more broadly for research.

### **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?

### CLARITY EVOLVES AND HONESTY IS ESSENTIAL

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.

## Specific pointers on problem solving

### 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 VII**

### **End matter**



---

## Appendix A

### GLOSSARY OF CONCEPTS AND TERMS

**absolute velocity** The absolute velocity is the velocity as measured within an absolute or inertial reference frame. On a rotating planet, the absolute velocity,  $\mathbf{v}_a$ , is related to the rotating reference frame velocity,  $\mathbf{v}$ , through the identity,  $\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}$ , where  $\boldsymbol{\Omega}$  is the angular velocity of the planet. 231, 293

**absolute vorticity** The absolute vorticity is the sum of the relative vorticity,  $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ , plus the planetary vorticity,  $2\boldsymbol{\Omega}$ . 97, 123

**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, 539

**active tracer** An active tracer affects the flow. Consequently, the tracer equation for active tracers is nonlinear. Temperature and matter concentration are two active tracers, in which these two tracers impact on the density, with density then modifying pressure that then modifies the flow. 567, 659

**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. 14, 545, 548

**advective time scale** The advective time scale is given by  $T = L/U$ , where  $L$  is a typical length scale and  $U$  is a typical velocity scale for fluid particles. The advective time scale is commonly used for scale analysis of geophysical fluid flows. 318

**advective tracer flux** The advective tracer flux,  $\rho C \mathbf{v}$ , is the mass of tracer per area per time that transported with the fluid flow. The convergence of this flux affects a time change to the mass of tracer per volume. 553

**ageostrophic** Ageostrophic flow is any part of the flow field that is not in balance through geostrophy. In the context of Ekman mechanics, the ageostrophic flow is that flow directly affected by friction. In general, the ageostrophic flow has a nonzero horizontal divergence, which contrasts to the non-divergent horizontal flow for  $f$ -plane geostrophic flows. 312, 633

**angular momentum** Angular momentum is the moment of the linear momentum as computed around a chosen origin. For a particle of constant mass,  $m$ , the angular momentum is,  $\mathbf{L} = \mathbf{X} \times \mathbf{P} = m \mathbf{X} \times \mathbf{V}$ . 34, 80

**anomalous diffusion** Brownian motion as formulated by [Einstein \(1905a\)](#) leads to the mean squared spread of a cloud of Brownian particles evolving according to  $d\overline{x^2}/dt = 2\kappa$ , so that for a constant diffusivity the spread grows as  $\overline{x^2} = 2\kappa t$ . Anomalous diffusion arises when  $\overline{x^2} \propto t^\xi$  for  $\xi \neq 1$ . Section 3 of [Young \(1999\)](#) provides a summary of anomalous diffusion in geophysical flows. [518](#)

**anti-symmetric tensor** An anti-symmetric (also skew-symmetric) tensor equals to minus its transpose,  $\mathbf{A} = -\mathbf{A}^\top$ . [627](#)

**anticipated potential vorticity** [Sadourny and Basdevant \(1985\)](#) proposed the anticipated potential vorticity method for parameterizing subgrid scale processes. The anticipated potential vorticity method makes use of an estimate for the future value of potential vorticity in computing the advection operator, thus motivating the term “anticipated” in the method’s name. It is notable that the method is not Galilean invariant, as might be expected given the method’ attempt to anticipate the future. [160](#)

**Archimedean buoyancy** The Archimedean buoyancy is the buoyancy acting on a test fluid element computed relative to a reference buoyancy. We consider both a globally referenced density, thus leading to the globally referenced Archimedean buoyancy,  $b = -g(\rho - \rho_0)/\rho_0$ , as well as a locally referenced density, used to compute neutral directions. The Archimedean buoyancy of a test fluid element makes use of perhaps the most ancient of physical concepts in fluid mechanics. [18, 79, 238, 279, 289, 295, 508, 628, 691, 693](#)

**atmospheric form stress** Atmospheric form stress is the contribution from atmospheric pressure acting to produce a horizontal ocean acceleration that acts at the interface between the ocean and atmosphere. This form stress is equal in magnitude to the oceanic form stress, but it is oppositely directed. [50](#)

**available potential energy** The available potential energy (APE) is the difference in gravitational potential energy between an arbitrary state and a state reached by reversibly rearranging fluid to have zero baroclinicity. [27, 58, 66](#)

**axial vector** An axial vector (also referred to as a pseudo-vector) changes sign when changing the handedness of the coordinate system. Angular momentum and angular velocity are axial vectors. [90](#)

**balanced model** A balanced model in geophysical fluid mechanics is a reduced system that enforces dominant force balances diagnostically, filters fast waves, and evolves the flow on the slow, vortical manifold, usually through potential vorticity dynamics. A balanced model is obtained by imposing diagnostic balance relations (rather than prognostic equations) between key variables, such as velocity, pressure, and buoyancy, so that the flow remains close to a slow manifold determined by dominant balances. The remaining prognostic equations govern the slow, vortical evolution of the flow. Quasi-geostrophic and semi-geostrophy are two further balanced models encountered in this book. The dynamical fields in a balanced model can be diagnosed through knowledge of the potential vorticity. [375](#)

**baroclinic flow** Baroclinic flow is flow with a nonzero baroclinicity vector, so that vorticity is affected by baroclinicity. [221](#)

**baroclinic instability** Baroclinic instability is a fluid flow instability arising from the presence of baroclinicity in a thermal wind balanced flow. Baroclinic instability can be interpreted as a wave or global instability that arises from the constructive interference of vorticity waves. Baroclinic instability is formulated mathematically within quasi-geostrophy. 345

**baroclinic velocity** The baroclinic velocity refers to the horizontal velocity minus the depth averaged horizontal velocity, with the average taken over the full depth of a fluid column. This velocity is also sometimes referred to as the internal velocity. 57

**baroclinicity** Baroclinicity is the miss-alignment between density and pressure that renders a source for vorticity,  $\mathbf{B} = -\nabla \times (\rho^{-1} \nabla p) = (\nabla \rho \times \nabla p)/\rho^2$ . In a Boussinesq ocean, the baroclinicity vector arises just from horizontal gradients in the buoyancy,  $\mathbf{B} = \nabla b \times \hat{\mathbf{z}}$ . 115, 122, 213, 239, 289, 467, 479

**barotropic flow** Barotropic flow is flow with a zero baroclinicity vector, so that vorticity is not affected by baroclinicity. 221

**barotropic fluid** A barotropic fluid is one for which pressure is a function of density,  $p = p(\rho)$ , so that the baroclinicity vanishes,  $\mathbf{B} = (\nabla \rho \times \nabla p)/\rho^2 = 0$ . 115

**barotropic velocity** The barotropic velocity refers to the depth averaged horizontal velocity, averaged over the full depth of a fluid column. It is also referred to as the external velocity. 57

**barycentric velocity** The barycentric velocity is the center of mass velocity for a fluid element. The barycentric velocity plays the same role for multi-component fluids as the fluid parcel velocity does for single-component fluids. Differences between the barycentric velocity and the velocity of a specific fluid constituent can lead to the exchange of matter constituents across the boundary of the fluid element, with that exchange typically represented as diffusion. 211, 440, 460, 548, 557, 559, 669, 670, 695

**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. 681, 684, 745

**Batchelor scale** The Batchelor scale sets the smallest spatial scale of scalar (e.g., temperature or salinity) fluctuations in a flow with both turbulent advection and molecular diffusion. It is the dissipation scale for scalars, below which molecular diffusion dominates over advective stirring. 623

**Bernoulli potential** The Bernoulli potential is the sum of the mechanical energy per mass plus the enthalpy per mass,  $\mathcal{B} = \mathcal{M} + \mathcal{H} = \mathcal{K} + \Phi + \mathcal{I} + p/\rho$ . For the Boussinesq ocean, the Bernoulli potential is  $\mathcal{B} = \mathcal{K} + P^b + \varphi$ . 282, 283, 295, 478

**Bernoulli's theorem** Bernoulli's theorem is a statement of total energy conservation for steady flows. Most commonly it appears for steady perfect fluid flows, in which the Bernoulli potential is constant along streamlines. For such steady perfect fluid flows, there is an exchange between the total energy per mass,  $\mathcal{E}$ , and the injection work,  $p/\rho$ , such that their sum remains constant along the streamline. 64, 282

**beta effect ( $\beta$ -effect)** The  $\beta$ -effect arises from the meridional dependence of the Coriolis parameter:  $\beta = \partial_y f$ , thus giving rise to differential rotation. The  $\beta$ -effect supports vorticity waves known as planetary Rossby waves. 122, 169, 187, 233–235

**beta gyre** On the  $\beta$ -plane, parcels moving around a cyclonic and circularly symmetric northern hemisphere ( $f > 0$ ) monopole flow pick up anomalous relative vorticity according to the beta effect:  $D\zeta/Dt = -\beta v$ . On the western side, parcels move southward and thus pick up a positive relative vorticity anomaly ( $-\beta v > 0$ ), whereas on the eastern side the northward flow picks up a negative relative vorticity anomaly. The beta effect thus induces a westward drift of the monopole maximum, towards where the relative vorticity is increasing. Additionally, the positive anomaly on the western side of the monopole induces a secondary circulation known as a beta gyre, with this gyre rotating counter-clockwise, whereas there is an oppositely oriented beta gyre on the eastern side. The secondary circulation from the counter-rotating beta gyres induces a northward drift to the monopole. The combined effect of the westward beta induced drift and the northward drift from the beta gyres leads to a net northwestward beta drift for the monopole. 142

**beta plume** The beta plume is the steady and westward extending response of a rotating fluid to localized forcing, caused by Rossby-wave radiation on a beta plane. It represents the Green's function solution when the forcing is a Dirac delta. 146

**beta-drift ( $\beta$ -drift)** The  $\beta$ -drift accounts for the northwestward drift of an ideal vortex due to differential rotation. The  $\beta$ -drift arises from pressure effects that set up a secondary flow that induces westward drift, in addition to the northward drift from the Rossby effect. 140, 142

**beta-plane ( $\beta$ -plane)** The  $\beta$ -plane is a tangent plane approximation (tangent to the geoid) that makes use of Cartesian coordinates for studying geophysical fluid motion local to a point on the rotating planet, and using a Coriolis parameter that is a linear function of latitude. Since motion is assumed to be on a constant geopotential, the  $\beta$ -plane makes use of the effective gravitational acceleration that includes both central gravity acceleration plus the planetary centrifugal acceleration. 183, 321, 323

**biharmonic operator** The biharmonic operator is the squared Laplacian operator,  $\nabla^4 = \nabla^2 \nabla^2$ . Biharmonic operators arise in elasticity theory as well as for Stokes flow in fluid mechanics. For geophysical fluid mechanics, biharmonic operators are commonly used for numerical reasons to enhance the scale-selectivity of dissipation. 655

**Biot-Savart law** In electromagnetism, the Biot-Savart law computes the magnetic field by integrating the contributions from all current elements, with direction given by a cross product and strength decaying as inverse distance squared. In fluid mechanics, we encounter mathematically similar expressions for the velocity field written in terms of an integral of the vorticity. 127

**Bjerknes circulation theorem** Kelvin's circulation theorem applies to barotropic and perfect fluid flow, in which circulation is materially conserved. Bjerknes' circulation theorem generalizes Kelvin's result to baroclinic fluids on a rotating planet. [233](#)

**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. [xix](#), [50](#)

**bolus velocity** The bolus velocity is a horizontal velocity that arises from the correlation between layer thickness and layer velocity when performing an eddy-mean decomposition in isopycnal coordinates or in a stacked shallow water model. [614](#), [620](#)

**boundary propagator** The boundary propagator is the boundary Green's function for the diffusion and advection-diffusion problem, which incorporate time dependence along with space dependence. [583](#)

**Boussinesq ocean** The Boussinesq ocean equations serve as a useful framework to study aspects of the ocean circulation, both large-scale and small-scale. The Boussinesq ocean's prognostic velocity is non-divergent, thus representing an incompressible flow, and yet the Boussinesq ocean fluid admits density variations, as for a compressible fluid. That is, the study of a Boussinesq ocean concerns the incompressible flow of a compressible fluid, thus exemplifying the important distinction between a fluid property versus a flow property. Since the flow is non-divergent, the pressure in the Boussinesq ocean is not the thermodynamic pressure found in the compressible non-Boussinesq fluid. Rather, Boussinesq pressure serves a purely mechanical role by acting as the Lagrange multiplier to constrain the Boussinesq flow to be non-divergent. The Boussinesq ocean has an inertial mass based on a constant reference density,  $\rho_0$ , whereas the gravitational mass is based on the *in situ* density,  $\rho$ . Hence, the Boussinesq ocean does not respect the principle of equivalance, thus making it necessary to exercise special care when studying energetics of the Boussinesq ocean. [21](#), [59](#), [70](#), [93](#), [129](#), [163](#), [237](#), [254](#), [270](#), [286](#), [313](#), [353](#), [385](#), [437](#), [457](#), [468](#), [543](#), [586](#), [693](#)

**Brownian motion** Brownian motion is the process whereby a relatively large particle (e.g., piece of dust) is transported by the random motion of molecules within the fluid. [Einstein \(1905a\)](#) provided a theory for Brownian motion, which provided clear evidence for the molecular nature of matter. Brownian motion was also used as an inspiration for [Taylor \(1922\)](#) in his theory of turbulent diffusion. [511](#), [514](#), [625](#)

**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](#)

**buoyancy** Buoyancy, or more precisely the Archimedean buoyancy, is the vertical force, acting on a region,  $\mathcal{R}$ , immersed in a fluid, with the forces those static forces from gravity and hydrostatic pressure. If the density of the displaced fluid is greater than the density of the matter within the region,  $\rho^{\text{fluid}} > \rho^{\mathcal{R}}$ , then the buoyancy force is in the  $+\hat{z}$  direction,

thus leading to a rising motion of the region. The converse happens if  $\rho^{\text{fluid}} < \rho^{\mathcal{R}}$ , in which case the region experiences a negative buoyancy force so that it sinks. If  $\rho^{\text{fluid}} = \rho^{\mathcal{R}}$ , then the region experiences zero buoyancy force so that the region is neutrally buoyant and it floats. When considering buoyancy of fluid elements, then we find a net positive or negative buoyancy arises only for a fluid environment with density inhomogeneities, so that the buoyancy of a fluid element vanishes in a fluid with a homogeneous density. 673, 691

**buoyancy frequency** The buoyancy frequency,  $N$ , provides a measure of the vertical stratification of density. It also provides a measure of the angular frequency for buoyancy oscillations. We compute the squared buoyancy frequency for the ocean via  $N^2 = g(\alpha \partial_z \Theta - \beta \partial_z S)$ , whereas the atmosphere typically computes it according to the vertical derivative of the potential temperature. If  $N^2 < 0$  then the vertical column is gravitationally unstable. 21, 287, 437, 629

**Burger number** The Burger number is a non-dimensional number is the squared ratio of the deformation radius to the lateral length scale of the flow,  $\text{Bu} = (L_d/L)^2$ . 320, 401

**cabbeling** Consider the mixing of two seawater elements. Let the fluid elements separately have distinct Conservative Temperature and/or salinity, but equal locally referenced potential density. For a linear equation of state, whereby density is a linear function of  $\Theta$  and  $S$ , then the resulting mixed fluid element has the same density as the unmixed separate elements. However, for a nonlinear equation of state, the mixed element generally has a different density. Furthermore, a property of seawater is that the density of the mixed element is greater than the unmixed elements. This densification upon mixing is a physical process known as cabbeling (*McDougall*, 1987b). 664–667, 684

**Cauchy equation of motion** The Cauchy equation of motion is the expression of Newton's law of motion applied to continuous media such as a fluid. 49

**centrifugal** The centrifugal acceleration points towards the convex side of a turning trajectory; i.e., outward. Its opposing partner, the centripetal acceleration, points toward the center. The centrifugal acceleration arises from accelerated motion due to motion along a curved path. It is the centrifugal acceleration that pulls one away from the center of a merry-go-round whereas one's arms and hands provide the balancing centripetal acceleration. For the rotating Earth, gravity provides the centripetal acceleration to counteract the centrifugal acceleration of the rotating planet. 44, 130

**Chapman-Kolmogorov** The Chapman-Kolmogorov composition principle states that for a Markov process, the probability of going from a state A to a state C later in time can be found by summing over all possible intermediate states B. Namely, it is all possible paths through the middle, added up, thus expressing how multi-step transitions are built from one-step transitions. 515

**circulation** Circulation is a measure of the net rotation of the fluid flow around a closed and simply connected loop. Through Stokes' theorem, circulation equals to the area integral of the vorticity as projected normal to the area bounded by the loop. 94, 429

**compatibility** Compatibility refers to the mathematical compatibility between the equations for tracer conservation and mass conservation. Compatibility manifests by the continuity

equation for tracers reducing to the mass continuity equation when the tracer concentration is a uniform constant. Compatibility requires the non-advectional tracer fluxes to vanish when the tracer concentration is uniform, and for tracer boundary conditions to reduce to those for the fluid mass. 551

**computational modes** Computational modes refer to spurious (unphysical) spatio-temporal solutions to the discrete equations of a numerical model that have no manifestation in the continuum fluid. They generally manifest by power at the grid scale, such as through check-board patterns in space or oscillatory patterns in time. We can avoid some computational modes by designing numerical methods that do not admit them. Even so, it is not always possible to eliminate all computational modes *a priori*, thus making it necessary to introduce some form of smoothing filter in space and/or time. 206

**Conservative Temperature** Conservative Temperature,  $\Theta$ , is the potential enthalpy of a fluid element divided by a standard heat capacity. Conservative Temperature is far more conservative than potential temperature,  $\theta$ . Hence,  $\Theta$  is the preferred variable for measuring changes in heat within the ocean. 543, 567, 659, 660, 673, 691

**conservative tracer** Conservative tracers evolve via the convergence of advective and diffusive fluxes within the fluid interior, along with boundary conditions. Conservative tracers have no interior sources or sinks, so the net content of a conservative tracer over any finite volume domain is affected only through transport across boundaries. 508, 529, 543, 545, 659

**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, 50, 223, 484

**continuity equation** A continuity equation is a flux-form differential conservation law for intensive fluid properties that are typically expressed using the Eulerian kinematic perspective. Examples include the mass continuity equation for all constituents within a fluid sample,  $\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0$ , and the tracer continuity equation holding for individual matter constituents,  $\partial_t (\rho C) + \nabla \cdot (\rho \mathbf{v} C + \mathbf{J}) = 0$ . 451, 543, 548, 686

**continuum approximation** The continuum approximation assumes that mathematical limits for fluid volumes tending to zero are reached on length and time scales very large compared to molecular space and time scales. The temporal realization of the continuum approximation is based on recognizing that macroscopic motion associated with fluid flows (e.g., advection, waves, and mixing) evolves with time scales far longer than the time scales of molecular motions. Hence, from a macroscopic perspective, the continuum approximation leads us to assume that all fluid motions are continuous in both space and time. 511

**Coulomb gauge** The Coulomb gauge is commonly used in electrostatics (e.g., Jackson (1975)). For the eddy-induced velocity, with  $\rho \mathbf{v}^* = \nabla \times (\rho \Psi^*)$ , the Coulomb gauge is defined by setting  $\nabla \cdot (\rho \Psi^*) = 0$ . 556

**Coulomb-Ampere solution** The Coulomb-Ampere solution is for a non-divergent vector field in the absence of boundaries, and it is written as the convolution of the source with the free-space Green's function. 556

**covariance** Covariance means that the physical equations are form invariant under a chosen symmetry operation. The fundamental equations of geophysical fluid mechanics respect Galilean covariance, and they do so using any arbitrary coordinate system. Additionally, the Eulerian form of the equations respect coordinate invariance whereby they take on the same form for all time independent coordinates. 429

**covariant derivative** The covariant derivative is the generalization of the partial derivative operator that transforms as a  $(0, 1)$  tensor with arbitrary coordinates. 531

**curvature circle** The curvature circle is a geometric construction used to compute the radius of curvature for a curve. More precisely, the radius of curvature equals to the radius of a tangent circle (the curvature circle) that approximates, to second order accuracy, the curve at a particular point. The radius is smaller in magnitude when the curve is highly curved, the radius goes to infinity as the curve straightens. 104

**curvature vorticity** Curvature vorticity is that component of relative vorticity arising from the curvature of the trajectory of a fluid parcel. It is also sometimes referred to as orbital vorticity. 104

**deformation raduis** The deformation radius,  $L_d$ , distinguishes flows where the Coriolis force is important,  $L \geq L_d$ , from those where the Coriolis can be neglected,  $L \ll L_d$ . The deformation radius decreases toward the poles, so that rotational effects are felt by smaller scales in the high latitudes than in the tropics. 320

**dia-surface transport** The dia-surface transport refers to the transport of mass (mass per time) crossing a surface. If the surface is moving, then the dia-surface transport must take into account motion of the surface. The dia-surface transport is a fundamental kinematic object for stratified flows, and it plays a particularly central role in the fluid mechanical equations using generalized vertical coordinates. 568

**diabatic** A diabatic process in a fluid occurs with the exchange of heat between fluid elements. In this manner, physical systems experiencing diabatic processes are said to be thermally open. 545

**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

**dianeutral diffusion** Dianeutral diffusion refers to irreversible ocean mixing processes that are parameterized by diffusive fluxes that are oriented perpendicular to the local neutral direction. Dianeutral diffusion is often also called diapycnal diffusion. Furthermore, it is commonly approximated as vertical diffusion given that in much of the ocean interior has neutral directions close to horizontal. This mixing is important for ocean stratification, transformation of water mass properties, and changes to potential energy. 629

**dianeutral direction** The dianeutral direction points orthogonal to the neutral direction, with the dianeutral direction given by  $\hat{\gamma} = (-\alpha \nabla \Theta + \beta \nabla S) / | -\alpha \nabla \Theta + \beta \nabla S |$ . [628](#)

**differentiable manifold** A differentiable manifold is locally Euclidean and possess smoothness and continuous properties that allow one to perform differential calculus. A differentiable manifold is not necessarily equipped with a metric. For example, Gibbs' thermodynamic configuration space is a differentiable manifold that has no *a priori* metric structure. [691](#), [692](#)

**diffusion** Diffusion is the physical process by which a field, such as a tracer, spreads in space over time due to random motion, either from molecular chaos (molecular diffusion) or turbulent flows (turbulent diffusion). The net flux moves from regions of higher concentration to regions of lower concentration, thus moving down the concentration gradient. The mathematical equation describing diffusion is the canonical parabolic partial differential equation. [511](#), [545](#)

**diffusion dissipation functional** The diffusion dissipation functional is the domain integral of  $\mathbf{J} \cdot \nabla C / 2$ , where  $\mathbf{J}$  is the diffusive flux. The functional derivative (Fréchet derivative) of this functional equals to the diffusion operator. [537](#)

**diffusion tensor** The diffusion tensor is a second order symmetric and positive-definite tensor with dimensions of squared length per time. It is used to parameterize the downgradient eddy diffusive fluxes of tracers arising from turbulent motions. [533](#), [672](#)

**diffusive flux** The diffusion of tracer is derived by computing the convergence of the tracer diffusive flux, which is a flux that is directed down the tracer gradient and its strength is mediated by a diffusion tensor. [526](#)

**diffusivity** The diffusivity is the physical parameter that measures how fast a substance spreads out by diffusion. It sets the proportionality between the gradient of a quantity and the resulting diffusive flux. We generally work with the kinematic diffusivity, whose physical dimensions are  $L^2 T^{-1}$ . The diffusivity for molecular diffusion is a scalar, whereas for turbulent diffusion it is a second order tensor. [524](#)

**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. [112](#), [126](#), [145](#), [394](#), [528](#), [538](#), [676](#)

**Dirac delta sheet** The Dirac delta sheet refers to the ability to absorb the Neumann boundary condition into a modified interior source. By doing so, the boundary value problem has a modified interior source that is proportional to a Dirac delta (hence the term Dirac delta sheet), but it now has a homogeneous boundary condition. This reformulation of the Neumann boundary condition is commonly pursued in the geophysical fluids literature, such as when studying potential vorticity and for water mass analysis. [676](#)

**Dirichlet boundary condition** The Dirichlet boundary condition prescribes the value of a function along the boundary. [543](#), [574](#)

**divergence damping** Divergence-damping is a numerical method that acts to preferentially damp the horizontally divergent portion of the velocity field, thus acting to dissipate the divergent flow, while leaving the vorticity equation untouched. The motivation for this operator is based on noting that much of the large-scale circulation in a rotating fluid has a nontrivial absolute vorticity yet a relatively small horizontal divergence. 206

**divergence theorem** The divergence theorem, also known as Gauss's divergence theorem, provides a relation between the volume integral of the divergence of a vector to the boundary integral of the vector projected onto the outward normal:  $\int_{\mathcal{R}} \nabla \cdot \mathbf{F} dV = \oint_{\partial\mathcal{R}} \mathbf{F} \cdot \hat{\mathbf{n}} dS$ . There are many corollaries to this theorem that we use in this book, one that says for Cartesian tensors that  $\int_{\mathcal{R}} \nabla \phi dV = \oint_{\partial\mathcal{R}} \phi \hat{\mathbf{n}} dS$ , which we use in formulating the effects from pressure in a weak formulation of the equations of motion. 298, 426, 462

**double diffusive processes** Double diffusive processes refer to a class of small-scale mixing processes that occur when temperature and salinity gradients act together to produce density stratification. Because heat diffuses about 100 times faster than salt in seawater, a stratification that is stable in density can nonetheless become unstable dynamically, leading to distinctive mixing patterns such as salt fingering or diffusive convection. 625, 665

**downscale cascade** The downscale cascade refers to the transfer of a conserved (or approximately conserved) quantity, such as typically kinetic energy, enstrophy, tracer or buoyancy variance, from larger spatial scales to progressively smaller scales through nonlinear interactions, without direct forcing or dissipation at the intermediate scales. Imagine a turbulent ocean flow in which mechanical energy is injected at some large scale, for example, by winds, tides, or instabilities. The nonlinear advection terms in the equations of motion redistribute this energy across scales. A downscale cascade occurs when this redistribution is predominantly from large to small scales, eventually reaching scales where dissipation (e.g., viscosity, mixing) can act. It manifests through filamentation, front sharpening, submesoscale generation, and turbulent mixing, and is diagnosed through spectral fluxes in numerical models and observations. The downscale cascade is in contrast to an upscale cascade, where kinetic energy or other invariants flow from smaller to larger scales (e.g., barotropic kinetic energy in 2D turbulence). 623

**dual form stress** The dual form stress,  $-\eta \nabla p$ , differs by a gradient from the form stress,  $+p \nabla \eta$ . Even though both have dimensions of a pressure, the dual form stress does not act to accelerate a fluid element. When integrated over a region where one can drop the gradient term, the integrated effects from the dual form stress equal to that from the form stress. However, it is not possible to make a local identification between the form stress and the dual form stress. 497

**Dufour effect** The Dufour effect is a flux of heat that arises from matter concentration gradients and pressure gradients. 513, 661

**dynamic diffusivity** The dynamic diffusivity equals to the kinematic diffusivity times the density of the fluid. When referring to "diffusivity" in this book, we generally mean the kinematic diffusivity rather than the dynamic diffusivity. 524, 525

**dynamic pressure** The dynamic pressure is another name for the kinetic energy per mass as it appears in the vector-invariant velocity equation. Gradients in the kinetic energy per

mass contribute a dynamical pressure gradient that accelerates the fluid down the kinetic energy gradient, from regions of high kinetic energy per mass to regions of low kinetic energy per mass. 28

**dynamical pressure** The dynamical pressure is that part of the hydrostatic pressure that has a horizontal gradient, and thus contributes to horizontal motion. It is not to be confused with the dynamic pressure, which is another name for the kinetic energy per mass as it appears in the vector-invariant velocity equation. 24

**dynamical tracer** We refer to potential vorticity as a dynamical tracer since it is directly constructed from dynamical fields (velocity and buoyancy), and it provides a direct connection to the dynamics through inversion. 269

**eddy** An eddy refers to any flow feature that is a departure from a subjectively chosen mean. 591

**eddy diffusivity** An eddy diffusivity is an emergent property of the flow that aims to capture the essential features of turbulent motion to irreversible mixing of tracers. Eddy diffusivities are typically far larger than molecular diffusivities, given the far more efficient nature of turbulent transport than molecular transport. 541, 653

**eddy-induced velocity** An eddy-induced velocity is an emergent property of the flow that aims to capture the advective transport features of turbulent motion. Eddy-induced velocities are akin to Stokes' drifts, and they play an important role in parameterizations of geostrophic eddies. 558, 627, 654, 662, 665

**effective free surface height** The effective free surface height accounts for the applied surface pressure that either decreases or increases the free surface. 7

**Einstein summation convention** The Einstein summation convention says that we drop the summation symbol while repeated indices (one upstairs and one downstairs) are summed over their range. Although seemingly trivial, the summation convention proves central to a variety of tensor manipulations, and so it is important to make friends with it. 108

**Ekman boundary layer** An Ekman boundary layer arises from a balance between pressure, Coriolis, and turbulence induced friction. The role of rotation distinguishes geophysical boundary layers from engineering applications. The associated Ekman boundary layers are crucial for understanding circulation and transport in both the atmosphere and ocean. 79

**Ekman mechanics** Ekman mechanics is concerned with geophysical fluid flow affected by accelerations from horizontal pressure gradients, vertical friction, and Coriolis, with particular attention given to regions near boundaries where turbulent friction is especially large. 182, 250, 307, 353, 359, 366, 401

**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

**Eliassen-Palm flux** The Eliassen–Palm flux is a wave-activity flux whose divergence gives the eddy forcing of the zonal-mean flow in a rotating, stratified fluid. 49, 58

**elliptic partial differential equation** An elliptic partial differential equation (PDE) is a second order PDE whose coefficients satisfy certain positivity conditions that ensure it behaves like Laplace’s equation. The solutions are smooth and non-propagating, thus commonly found when studying static or equilibrium phenomena. 129, 312, 328, 336, 375, 556

**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

**ensemble mean** The ensemble mean is based on averaging over a suite of identically prepared flow realizations that differ in a controlled manner such as through the initial conditions. For many purposes this is the most analytically convenient mean operator, though it is often difficult to realize in practice. 489, 594

**epineutral** Epineutral refers to processes that are oriented parallel to the local neutral direction, which are directions on which a fluid parcel can be displaced without experiencing a buoyancy force. 631

**epineutral diffusion** Epineutral (also neutral) diffusion is based on a diffusion tensor that yields diffusive fluxes oriented along neutral directions. It is also known as neutral diffusion. It parameterizes irreversible mixing of tracers along local neutral directions. Epineutral diffusion is a refinement of isopycnal diffusion, motivated by the fact that potential density surfaces are not exactly neutral in a compressible, thermally expanding, and saline fluid like seawater. 640

**equation of state** An equation of state expresses a constraint satisfied by thermodynamic variables in thermodynamic equilibrium. In geophysical fluid mechanics, we typically refer to the equation of state as an equation specifying the mass density of a fluid element in terms of thermodynamic state properties, such as temperature, tracer concentration, and pressure. The equation of state for the atmosphere is well approximated by the ideal gas equation, whereas the ocean has a nonlinear equation of state whose coefficients are fit by empirical measurements. 479, 720

**equivalent barotropic** Equivalent barotropic flow is characterized by flow that is in the same direction for all depths, though with potentially distinct magnitudes. In stratified systems,

equivalent barotropic flows often resemble the first baroclinic mode yet with no zero crossing. 127

**equivalent barotropic depth** Equivalent barotropic depth is the single effective vertical scale that represents a vertically coherent (equivalent-barotropic) flow, as if it were a one-layer barotropic motion. 128

**Ertel potential vorticity** Ertel potential vorticity is the scalar product of absolute vorticity and the gradient of potential temperature (atmosphere) or buoyancy (ocean), and it is materially conserved in adiabatic, inviscid flow. 163, 263, 267, 289

**Euler equation** The Euler equation is the equation of motion for a perfect fluid, thus containing zero friction or other irreversible processes. We commonly make use of versions of the Euler equation in our studies, given that large-scale geophysical flows are not greatly affected by details of the molecular viscosity. 400

**Euler-Lagrange equation** The Euler-Lagrange equation is a differential equation that results from Hamilton's principle of stationary action. In classical mechanics, the Euler-Lagrange equation is the same as Newton's equation of motion. 539

**Eulerian mean** An Eulerian mean refers to any averaging operation taking place at a fixed point in space so that it is computed within the Eulerian reference frame. 593

**Eulerian reference frame** An Eulerian reference frame describes fluid motion relative to a fixed position ( $x$ -space), commonly referred to as the laboratory frame. This kinematic description measures fluid properties as the fluid streams by a fixed observer. It is not concerned with determining trajectories. Instead, Eulerian kinematics focuses on fluid properties as continuous fields that are functions of the space position,  $x$ , and time,  $t$ . 692

**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

**Exner function** The Exner function is a thermodynamic quantity that relates pressure to temperature and potential temperature, which takes on a particularly simple form when assuming an ideal gas. 214

**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, 314

**external velocity** The external velocity refers to the depth averaged horizontal velocity, averaged over the full depth of a fluid column. It is also referred to as the barotropic velocity. 57

**f-plane** The  $f$ -plane is a tangent plane approximation (tangent to the geoid) that makes use of Cartesian coordinates for studying geophysical fluid motion local to a point on the rotating planet and using a constant Coriolis parameter. Since motion is assumed to be on

a constant geopotential, the  $f$ -plane makes use of the effective gravitational acceleration that includes both central gravity acceleration plus the planetary centrifugal acceleration. 586

**fetch** The fetch refers to the region over which winds blow and thus affect the transfer of momentum to the ocean for producing surface gravity waves. 76

**Fick's law of diffusion** Fick's law of diffusion says that the diffusive flux of tracer concentration is given by a kinematic diffusivity times minus the gradient of the concentration, so that the flux is directed down the tracer concentration gradient. 513, 524, 527

**fine scale mixing** Fine-scale mixing refers to the mixing of seawater properties (e.g., heat, salt, momentum, tracers) that occurs at spatial scales just above the molecular dissipation range, typically meters to tens of meters, driven by the interaction between larger-scale internal waves or mesoscale motions and small-scale shear or stratification. It is the intermediate regime between large-scale stirring (hundreds of meters to kilometers, e.g., eddies and fronts) and microscale turbulent mixing (millimeters and below, at the Kolmogorov and Batchelor scales). 625

**flat representation** The flat representation refers to the  $(0, 2)$  representation of second order tensors. For example, the flat representation of the diffusion tensor is  $\mathbf{K} = K_{mn} \mathbf{e}^n \otimes \mathbf{e}^m$ . The flat representation follows the musical nomenclature for second order tensors, where we have sharp, flat, and natural representations. 531

**flow map** The flow map smoothly and continuously deforms the matter continuum through space as the fluid moves, and it is written as  $\boldsymbol{\varphi}(\mathbf{a}, T)$ . The flow map provides the trajectory for the fluid particle specified by the material coordinate,  $\mathbf{a}$ . So in this sense we can consider the flow map as the accumulation of all fluid particle pathlines, and with its time derivative,  $\partial_T \boldsymbol{\varphi}(\mathbf{a}, T)$ , providing the velocity of the fluid particle,  $\partial_T \boldsymbol{\varphi}(\mathbf{a}, T) = \mathbf{v}[\mathbf{x} = \boldsymbol{\varphi}(\mathbf{a}, T), t = T] = \mathbf{v}^\perp(\mathbf{a}, T)$ . The motion field is the reason there is a flow map, so that the nomenclature “motion field” and “flow map” are used interchangeably in this book since they both refer to movement of the continuum. 595

**fluid parcel** Fluid parcel are infinitesimal deformable regions of a perfect fluid that maintain a fixed matter content and fixed thermodynamic properties, so that they have fixed mass and fixed enthalpy. 607

**flux-form conservation law** A flux-form conservation law is a partial differential equation written in the form of a local time tendency of an intensive quantity that equals to (is driven by) the convergence of a flux. Such equations are generally written from an Eulerian perspective. Examples include the mass conservation equations, which are referred to as continuity equations, such as for the mass of all constituents within a fluid sample and for tracer continuity equations for individual matter constituents. 548

**form stress** Form stress is the contribution from pressure acting to give a horizontal acceleration. For a surface defined by  $z = \eta(x, y, t)$ , the form stress is given by  $p \nabla_h \eta$  when acting on the top side of the surface, and  $-p \nabla_h \eta$  on the bottom side. The name arises since the stress depends on the form, or shape, of the surface on which pressure acts. 2, 19, 24, 39, 50, 53, 190, 248, 261, 463, 465, 479

**formation** In the context of water mass analysis, fluid moves through water mass configuration space ( $\mathbf{q}$ -space) if it is transformed in a manner that causes water to cross surfaces of constant  $\mathbf{q}$ . It follows that the convergence (local accumulation) or divergence (local depletion) of transformation leads to the formation of water mass classes, or its negative, being the destruction of water masses. More precisely, formation is the layer integrated  $\mathbf{q}$ -space convergence of transformation, with both formation and transformation having dimensions of mass per time. 692–694, 711

**Fourier's law of conduction** Fourier's law expresses the conductive flux of heat as proportional to a conductivity times the gradient of temperature, with the flux directed down the temperature gradient. The time changes to the temperature are given by the convergence of the downgradient flux. 513, 525

**Fréchet derivative** The Fréchet derivative provides a generalization of an ordinary derivative. It measures how a functional (a mapping from functions to numbers) changes when the input function is varied. We make use of such derivatives when working with variational calculus. It is effectively the same as the functional derivative. 538

**free vortex** A free vortex has zero vorticity and zero circulation for all points except the origin. Yet the same points with zero vorticity and zero circulation have a constant angular momentum relative to the origin. 95

**free-space Green's function** The free-space Green's function is the Green's function defined in unbounded space. It is sometimes referred to as the fundamental solution to the partial differential equation. 126, 127, 145

**friction stress tensor** The friction stress tensor arises from relative motion (shears) within the fluid, and it is often called the deviatoric stress tensor since its trace vanishes. 34

**Froude number** The Froude number is the non-dimensional ratio of the fluid particle speed to the gravity wave speed. Froude number's larger than unity typically signal flow that is unstable to hydraulic jumps. For stratified flows, we also define the Froude number as a measure of the relative strength of vertical shears (i.e., vertical derivatives) of the horizontal velocity,  $U/H$ , versus the buoyancy stratification,  $N$ . For stratified flows, the squared Froude number equals to the inverse of the Richardson number. 316

**functional derivative** The functional derivative (effectively the same as the Fréchet derivative) provides a generalization of an ordinary derivative. It measures how a functional (a mapping from functions to numbers) changes when the input function is varied. We make use of such derivatives when working with variational calculus. 538

**Galilean boost** Galilean boost refers to a transformation in which the velocity is modified by a constant velocity. For velocity independent forces, Newtonian physics remains unchanged by a Galilean boost since since Newton's law of motion is Galilean covariant. 159

**Galilean covariance** Galilean covariance, also known as Galilean relativity, means that we can move between inertial reference frames without altering the physics, so long as the forces are velocity-independent and boundaries are absent. For such cases, all inertial reference frames are equally valid for studying physics, and the physical equations remain invariant under a Galilean transformation. 205

**gauge function** A gauge function is a physically unspecified function that is associated with a gauge symmetry. The simplest gauge function we encounter in this book is the arbitrary constant associated with streamfunctions for two-dimensional non-divergent flows. For three-dimensional non-divergent flows, gauge symmetry allows for an arbitrary gradient to be added to the vector streamfunction. We also encounter gauge functions when studying fluxes in flux-form Eulerian conservation laws. In these cases, is only the convergence of a flux that is physically relevant, so that the flux is arbitrary up to the curl of a vector function. 277, 293, 556

**gauge symmetry** A gauge symmetry arises from a redundancy in the mathematical description of a physical system. In fluid mechanics, there is a gauge symmetry associated with the ability to add a constant to the streamfunction in a horizontally non-divergent flow, with the constant not affecting the velocity field. For three dimensional non-divergent flow, the vector streamfunction is arbitrary up to the gradient of a scalar, with the details of this scalar irrelevant to the physics of the fluid flow. We also encounter gauge symmetries when defining the flux of a scalar field. Since the convergence of the flux affects time changes to the scalar, the flux is arbitrary up to the curl of a scalar. The presence of gauge symmetries affords us some freedom to choose a particular gauge to suite our subjective needs, with the choice not altering the objective physics. A standard reference for gauge symmetry, in the context of electromagnetism, is given in Section 27-4 in Volume II of *Feynman et al.* (1963). 276, 335, 552

**generalized Lagrangian mean** The generalized Lagrangian mean refers to a kinematic method used to decompose the flow into a Lagrangian mean plus a fluctuation relative to the mean. It is a hybrid Eulerian/Lagrangian method that introduces an Eulerian disturbance field to measure the position of a fluid particle relative to its mean position. 591, 593, 595, 596, 608

**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. 2, 34, 530, 531, 620, 645, 689, 696

**Gent-McWilliams effect** The general tendency for baroclinic eddies to extract potential energy from the flow and to thus relax fronts towards the horizontal. This effect from used by *Gent and McWilliams* (1990) and *Gent et al.* (1995) as the basis for the parameterization of the stirring effects from ocean geostrophic eddies. 632, 633

**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, 40, 312, 351, 500, 586

**geostrophic flow** Geostrophic flow arises from a balance between the Coriolis acceleration and the horizontal pressure gradient acceleration. 39, 44

**geostrophic streamfunction** The geostrophic streamfunction is the streamfunction for the horizontal geostrophic flow. On the  $f$ -plane, pressure divided by density and the Coriolis parameter provides a streamfunction for the geostrophic velocity:  $\mathbf{u}_g = \hat{\mathbf{z}} \times \nabla_h [p/(\rho_0 f)]$ . With steady geostrophic flow, the geostrophic streamfunction defines trajectories geostrophic contours or pathways for geostrophic flow. 178, 179, 326

**geostrophic Sverdrup balance** The geostrophic Sverdrup balance is the Sverdrup balance based on geostrophic flows. 366

**global conservation** A property is said to be globally conserved if it remains constant in time over a region in the absence of boundary fluxes. 275

**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

**Godfrey's island rule** Godfrey's island rule ([Godfrey, 1989](#)) states that in a steady, linear, large-scale ocean on a  $\beta$ -plane, the depth-integrated transport of fluid around an island is set by the integral of wind stress along a contour that encloses the island and closes to the eastern boundary. As such, the transport is largely independent of local boundary dynamics, and thus it is independent of details of the boundary currents around the island. 198

**gradient wind** Gradient wind balance refers to a steady balance in the horizontal tangent plane between the planetary Coriolis acceleration, motionally induced centrifugal acceleration, and the horizontal pressure gradient acceleration, as they act on a fluid parcel. It is commonly used to provide a diagnostic description of fluid motion with Rossby number near unity. 44, 130

**gravity waves** Gravity waves in fluids are supported by the gravitational restoring force. We encounter surface gravity waves, shallow water gravity waves, internal gravity waves, and inertia-gravity waves in this book. The dispersion relation for each of these waves includes the gravitational acceleration. 323

**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. 126, 145, 312, 375, 533, 556, 572

**H-theorem** Boltzmann's H-theorem is a foundational result in kinetic theory that explains how a gas evolves irreversibly toward thermodynamic equilibrium starting from time-reversible microscopic dynamics. The gas evolves irreversibly toward a state of maximum entropy, even though the underlying microscopic dynamics are time-reversible. 541

**haline contraction coefficient** The haline expansion coefficient is a response function that measures the relative change in density as the salinity is altered while holding the pressure and temperature. It is typically positive, so that density increases as salinity increases. 661

**Hamilton's principle** Hamilton's principle of stationary action is a variational method used to derive the equations of motion for a non-dissipative dynamical system, with the corresponding differential equations of motion known as the Euler-Lagrange equations. In classical mechanics, Hamilton's principle results in the same equations as Newton's second law. However, Hamilton's principle has widespread use beyond classical mechanics, making it part of most theories of physics. 537, 539

**harmonic function** A harmonic function has zero Laplacian:  $\nabla^2\psi = 0$ . Harmonic functions satisfy the remarkable mean value property, whereby the value of a harmonic function at a point  $\mathbf{x}$  within an open region of a domain,  $\mathcal{R}$ , equals to the average of  $\psi$  taken over the surface of a sphere within  $\mathcal{R}$  that is centered at  $\mathbf{x}$ . 93

**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 ? for more details. 394, 733

**helicity** Helicity of the fluid within a vortex tube volume is defined as the integration of the helicity density,  $\mathbf{v} \cdot \boldsymbol{\omega}$ , over the closed volume  $\int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \cdot \boldsymbol{\omega} dV$ , where the volume  $\mathcal{R}(\mathbf{v})$  is material. As reviewed by ?, helicity provides a measure of the knottedness of vortex tubes, and it plays a fundamental role in topological fluid dynamics. 260

**Helmholtz equation** the Helmholtz equation arises from assuming the monochromatic time dependence,  $\psi(\mathbf{x}, t) = e^{-i\omega t} \Psi(\mathbf{x})$ , for a solution to the wave equation,  $(\partial_{tt} - c^2 \nabla^2)\psi = 0$ . It takes the form,  $[\nabla^2 + (\omega/c)^2] \Psi = 0$ . The Helmholtz equation arise when we are uninterested in the initial value problem for waves, but instead are interested in the steady state where the wave field is fully established. 333

**Helmholtz's first theorem** Circulation around a vortex tube is the same for any position along the vortex tube. That is, the strength of a vortex tube is the same value along its length. This result is Helmholtz's first theorem. 101

**horizontal diffusion** Horizontal diffusion is based on a diffusion tensor that yields diffusive fluxes oriented along geopotential (horizontal) surfaces. 534

**hydrodynamics** A branch of fluid mechanics concerned with the flow of a homogeneous (constant density) incompressible fluid. vii

**hydrostatic approximation** 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. 345

**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](#)

**hydrostatic primitive equations** The hydrostatic primitive equations refer to the equations of motion for a fluid on a rotating planet, with the vertical momentum equation given by the hydrostatic balance. The horizontal equations are based on the traditional approximation in which the fluid thickness is much smaller than the planetary radius. The term “primitive” means that the prognostic dynamical field is the velocity rather than vorticity and divergence. [Smagorinsky \(1963\)](#) was an early proponent of the hydrostatic primitive equations for use in studying the large-scale ocean and atmospheric circulation. These equations form the basis for many general circulation models of the atmosphere and ocean.

[2](#)

**impermeability theorem** In the context of potential vorticity, the impermeability theorem is a kinematic result that says that the potential vorticity flux does not penetrate the isosurface of the scalar field used to define the potential vorticity. For example, if potential vorticity is defined by buoyancy, then the potential vorticity flux does not penetrate a buoyancy surface. This result holds even in the presence of irreversible processes, and it follows from kinematics. [88](#), [176](#), [289](#), [351](#), [483](#)

**impulse response function** The impulse response function is the response of a dynamical system to unit impulse. If the dynamical system is linear, then the impulse response function equals to the Green’s function for the initial value problem. [584](#)

**incompressible flow** Incompressible flow, also known as non-divergent flow, is defined by  $\nabla \cdot \mathbf{v} = 0$ . This flow is realized by the Boussinesq ocean, as well as fluid of constant density (and incompressible fluid). In two-dimensions, we consider the mechanics of horizontally non-divergent barotropic flow, whereby  $\nabla \cdot \mathbf{u} = 0$ . [92](#)

**incrop** An incrop is the location at the sea floor where a water property surface (e.g., an isopycnal, isotherm) meets at the ocean bottom. [27](#)

**inertial instability** Inertial instability refers to a particular form of symmetric instability in a flow in a rotating reference frame that is signaled by  $f(f + \zeta) < 0$ , so that the potential momentum of a fluid element increases in a contrary direction. This is a relatively fast and ageostrophic instability. We use the term inertial instability to refer to an instability in horizontal flows, whereas the term is sometimes used more broadly. [155](#)

**inertial mass** The inertial mass is the mass multiplying acceleration on the right hand side Newton’s law of motion,  $\mathbf{F} = m \mathbf{a}$ , and it is correspondingly used to measure kinetic energy. The principle of equivalence says that this mass equals to the gravitational used to compute the gravitational field using Newton’s law of gravity. [70](#)

**inertial oscillation** An inertial oscillation is the free horizontal motion of a fluid element on a rotating Earth that occurs when the only restoring influence is the Coriolis force, with no pressure-gradient force or friction acting. Motion is clockwise in the northern hemisphere and counter-clockwise in the southern. This motion arises from a balance between

the Coriolis acceleration of the rotating frame (planetary Coriolis) and the centrifugal acceleration arising from the particle's circular motion. Inertial oscillations appear in inertial waves, in which the only restoring force is the Coriolis force. Inertial waves are the limit of inertia-gravity waves as the stratification vanishes. Note that inertial oscillations are not geostrophic, as acceleration is leading order and the flow is unbalanced since pressure gradients are neglected. 344

**information entropy** Information entropy is used in statistical physics as a measure of the order/disorder of a probability distribution. It is proportional to the expectation value of the  $\ln p$ , where  $p$  is the probability density. It was originally introduced by *Boltzmann* (1966) in his proof of the H-theorem for kinetic theory of gases, and then by *Shannon* (1948) for information theory and *Jaynes* (1957) for statistical mechanics. 541

**integral curve** An integral curve is defined so that a given vector field is tangent to each point along the curve. For example, fluid particle trajectories are integral curves for the fluid velocity field. 595, 597

**interfacial form stress** The interfacial form stress is the contribution from the pressure stress that acts in the horizontal direction along an interface within a fluid, thus acting to produce a horizontal acceleration. 50–52

**internal 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, which we refer to as an internal scale. 314

**internal velocity** The internal velocity refers to the horizontal velocity minus the depth averaged horizontal velocity, with the average taken over the full depth of a fluid column. This velocity is also sometimes referred to as the baroclinic velocity. 57

**inverse barometer** A zero horizontal pressure gradient at the ocean surface, in the presence of an applied surface pressure, is known as an inverse barometer sea level, so that  $(\nabla_h p)_{z=\eta} = \nabla_h p_a + g \rho(\eta) \nabla_h \eta = 0$ . 7

**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

**irrotational** Irrotational flow has zero relative vorticity. 92, 99

**isopycnal** An isopycnal is a surface of constant potential density, which serves also as a surface of constant globally referenced Archimedean buoyancy. 2, 163, 591

**Kelvin's circulation theorem** Kelvin's circulation theorem says that for a perfect fluid with conservative body forces, the circulation around any closed material loop moving with the fluid is a material constant. It plays a central role in the study of circulation and

potential vorticity in fluids, including planetary fluid flows. Although the theorem only holds for restrictive cases of perfect fluids, it remains an important organizing principle for the study of geophysical fluid motion. 88

**kinematic boundary condition** The kinematic boundary conditions arise from the kinematic constraints on fluid motion when encountering a boundary. The simplest kinematic boundary condition is the no-normal flow, in which  $\mathbf{n} \cdot \mathbf{v} = 0$  for flow encountering a static and material boundary, such as the solid-earth, with  $\hat{\mathbf{n}}$  the outward normal. When the boundary is material and yet moves, then  $\mathbf{n} \cdot (\mathbf{v} - \mathbf{v}^{(b)}) = 0$ , where  $\mathbf{v}^{(b)}$  is the velocity of a point attached to the moving boundary. When the surface allows for fluid to cross, then the kinematic boundary condition is written  $\rho(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = -Q_m$ , where  $Q_m$  is the mass per time per area crossing the surface. If the boundary is the ocean surface, and the boundary is a monotonic function of vertical, then we can write the surface ocean kinematic boundary condition as  $w + \rho^{-1} Q_m = (\partial_t + \mathbf{u} \cdot \nabla) \eta$ , where  $z = \eta(x, y, t)$  is the vertical position of the ocean free surface, and  $Q_m$  is the mass per time per horizontal area of fluid crossing the boundary. 29, 53, 246, 250, 291, 306, 390, 433, 435, 459, 670, 686, 710, 715, 726, 730

**kinematic method** In water mass transformation theory, there are two complementary means to view interior water mass transformation: the *process method* and the *kinematic method*. The two methods are mathematically identical and so they offer two means to compute the same transformation. The kinematic method tells us *how* transformation happens, whereas we make use of the process to help understand *why* transformation happens. 706

**kinetic stress tensor** The kinetic stress tensor is a stress acting on an Eulerian region due to the transport of momentum across the boundary of that region,  $\mathbb{T}^{\text{kinetic}} = -\rho \mathbf{v} \otimes \mathbf{v}$ . In components, we have  $(\mathbf{v} \otimes \mathbf{v})^{ab} = v^a v^b$ . 46, 121, 247, 493

**Kronecker delta** The Kronecker delta symbol,  $\delta_{ab}$ , equals to unity if the indices  $a = b$  and zero otherwise. The Kronecker delta is the Cartesian coordinate representation of the metric for Euclidean space using Cartesian coordinates. 530

**Lagrangian mean** A Lagrangian mean refers to any averaging operation taking place at a fixed point in material space so that it is computed within the Lagrangian reference frame. 593

**Lagrangian reference frame** A Lagrangian reference frame is defined by motion of material fluid particles; i.e., it is a reference frame that is comoving with the continuum of fluid particles ( $a$ -space). The mechanical description aims to determine the continuum of trajectories, with each trajectory delineated by a continuous material coordinate that labels each fluid particle. The Lagrangian reference frame is non-inertial since fluid particles generally experience accelerations via changes to their speed and/or direction. 107, 692

**Langevin equation** The Langevin equation is a stochastic differential equation that describes the position of a Brownian particle. The two forces acting on the particle are the viscous damping from the liquid, and the random noise forcing from molecular impulses. 518

**Laplace-Beltrami operator** The Laplace-Beltrami operator is the Laplacian written using arbitrary coordinates. It is written  $\}^{-1} \partial_{\bar{m}} (\}) g^{\bar{m}\bar{n}} \partial_{\bar{n}} C$ , where  $\}$  is the square root of the determinant of the metric tensor written using the  $\xi^{\bar{m}}$  coordinates. 533

**Laplacian operator** The Laplacian operator is built from the squared gradient operator,  $\nabla \cdot \nabla = \nabla^2$ . In Cartesian coordinates we have  $\nabla^2 = \partial_{xx} + \partial_{yy} + \partial_{zz}$ , with other coordinates having specific forms. 126

**Leibniz's rule** Leibniz's rule provides the rule for differentiating integrals. 243

**Leibniz-Reynolds transport theorem** The Leibniz-Reynolds transport theorem provides the means to take the time derivative of an integrated fluid property, thus providing the basis for all finite-volume budgets within fluid mechanics. It serves to link the weak form (integral budgets) and strong form (differential budgets) of fluid mechanics. 304, 559, 712, 723, 727

**level of no motion** The level of no motion (also the depth of no motion) is an imagined level upon which there are no horizontal pressure gradients. This level is commonly used in diagnostic ocean studies where we do not have information about the absolute velocity, but instead only have information about the density field. We can then compute the geostrophic motion using thermal wind relative to the level of no motion. 15

**lid pressure** The lid pressure is the pressure applied to the surface of a rigid lid model that constrains the vertical velocity to vanish on this surface, and thus it maintains the constraint of non-divergence to the depth integrated flow. Furthermore, the lid pressure provides the force generating fluid motion in a horizontally non-divergent barotropic model. 27, 129

**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

**local steric sealevel change** The local steric sea level change refers to the depth integral of a local time derivative of the *in situ* density, thus providing a means to determine how local changes to density affect changes to sea level. There are a variety of expressions for the local steric change, with details provided by Griffies and Greatbatch (2012), Griffies et al. (2014). A common expression is given by  $(\partial\eta/\partial t)_{\text{steric}} = -\rho(\eta)^{-1} \int_{\eta_b}^{\eta} \partial\rho/\partial t dz \approx -\rho_0^{-1} \int_{\eta_b}^{\eta} \partial\rho/\partial t dz$ . 687

**locally referenced potential density** The locally referenced potential density is numerically equal to the *in situ* density at each point in the fluid. However, when computing spatial gradients, we hold the pressure fixed so that the gradient only probes changes in salinity and temperature. The locally referenced potential density is used for computing gravitational stability and neutral directions. 667

**Magnus acceleration** The Magnus acceleration appears in the vector-invariant velocity equation and has the mathematical form  $-\boldsymbol{\omega} \times \mathbf{v}$ . Since it acts only when there is both motion and vorticity, it is sometimes referred to as a *vortex force*. The Magnus acceleration is a nonlinear process that deflects a spinning fluid element in a direction perpendicular to its trajectory, in a manner analogous to the Coriolis acceleration. 29, 165

**Margules' relation** Margules' relation is an expression of thermal wind balance when realized for stacked shallow water layers. 25, 39, 42

**Markov processes** A discrete time stochastic process is Markov if its future depends on the present state but not on the past. Diffusion is the continuum limit of a discrete Markov process. 514

**mass distribution function** In the context of water mass transformation theory, the mass distribution function refers to the distribution of fluid mass as a function of water mass coordinates. 697

**material coordinate** A material coordinate (also Lagrangian coordinate) provides continuum of markers that distinguish fluid particles. These coordinates are used in the Lagrangian or material reference frame. There is a one-to-one relation between material coordinates and Eulerian coordinates, thus providing an invertible mapping between the Eulerian and Lagrangian reference frames. 597

**material invariant** The Lagrangian (material) time derivative vanishes for a property that is a material invariant. 170, 548

**material surface** Any continuous surface or interface that is impenetrable to the flow of matter or thermal energy (mechanical energy can be transferred via pressure forces). 11

**material tracers** Fluids generally contain multiple matter constituents, and we refer to such matter constituents as material tracers, with examples being salinity and freshwater in the ocean, oxygen, nitrogen, and water vapor in the atmosphere. We measure the concentration of matter within a fluid element as the ratio of the mass of matter constituent to the mass of all constituents within the sample. The tracer concentration satisfies the tracer equation. 543

**mechanical pressure** Mechanical pressure is the pressure that appears along the diagonal of the stress tensor. In this book, the mechanical pressure equals to the thermodynamical pressure that appears in the equations for equilibrium thermodynamics. 214

**mesoscale** The mesoscale in the ocean is dynamically characterized by a Rossby number less than unity, so that the flows are well described by quasi-geostrophy. For the atmosphere, the mesoscale is characterized by order unity Rossby numbers, which constitute the scales of atmospheric fronts and storms. 44

**metric** A metric tensor is a symmetric second order tensor that provides the means to measure the distance between two points in space. The Kronecker delta is the Cartesian coordinate representation of the Euclidean space metric tensor. When working with alternative coordinates (e.g., spherical, generalized vertical coordinates), the coordinate representation becomes less trivial. 530

**mixing** Mixing is the process whereby fluid elements irreversibly exchange properties. Mixing in geophysical fluids is affected by molecular diffusion, and its efficiency is enhanced through turbulent processes such as breaking waves and stirring by large scale eddies. Mixing is commonly parameterized by downgradient diffusion with an eddy diffusivity, though not all mixing processes are downgradient. 545, 548

**mixing length** The mixing length is a notion introduced by ? that estimates the length scale associated with turbulent diffusive transport. It is generally much larger than the molecular mean free path. Determination of turbulent length and velocity scales is subject

to large uncertainties and variations given the multiple regimes of turbulence exhibited by geophysical flows. 524

**modified mean** The modified mean refers to the generalized Lagrangian mean of a field, with the particle displacement tracking only the vertical position. The modified mean appears in the study of isopycnal averaging used to describe turbulent motions in a perfect stratified fluid. 609

**molecular diffusion** Diffusion is the physical process by which a field, such as a tracer, spreads in space over time due to random motion. Molecular diffusion arises from molecular chaos. The net flux moves from regions of higher concentration to regions of lower. 511, 513, 548, 625

**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

**Montgomery potential** The Montgomery potential,  $M = \varphi - bz$ , satisfies the buoyancy coordinate form of the hydrostatic balance. The Montgomery potential plays a role for isopycnal coordinates that is directly analogous to pressure in geopotential coordinates. Correspondingly, the Montgomery potential is the geostrophic streamfunction in buoyancy coordinates. 24, 474

**motion field** The motion field,  $\boldsymbol{\varphi}$ , is the mathematical object (or a “machine” using the language of *Misner et al. (1973)*) that generates a nonlinear time dependent and invertible flow map that continuously and smoothly reshapes the continuum as it moves through Euclidean space as time evolves. The motion field is the reason there is a flow map, so that the nomenclature “motion field” and “flow map” are used interchangeably in this book. Both refer to movement of the continuum as time progresses. 595

**musical nomenclature** The musical nomenclature is sometimes used for distinguishing the representation of second order tensors. The natural representation of a second order tensor occurs with one tensor index upstairs and the other downstairs. The natural representation is sometimes denoted by (1, 1), to indicate the number of indices up and down. The sharp representation, or (2, 0) representation, is when the tensor is represented with both indices upstairs. Finally, the flat representation or (0, 2) representation is where both indices are downstairs. 531

**natural boundary condition** Natural boundary conditions arise in the context of formulating the diffusion operator as the functional derivative of the dissipation functional, in which case the natural boundary condition is the Neumann (flux) boundary condition. 538

**natural coordinates** Natural coordinates (also intrinsic coordinates) refer to coordinates that are aligned along the flow direction and orthogonal to the flow. They are sometimes used as a means to kinematically decompose the flow into physically distinct components. In particular, they offer a concise means to compare the relative magnitudes of the Coriolis, pressure, and centrifugal accelerations acting on a fluid element moving horizontally 104, 131, 138

**natural representation** The natural representation refers to the  $(1, 1)$  representation of second order tensors. For example, the natural representation of the diffusion tensor is  $\mathbf{K} = K^m{}_n \mathbf{e}^n \otimes \mathbf{e}_m$ . [531](#)

**Neumann boundary condition** The Neumann boundary condition prescribes the normal derivative of a function along the boundary. For tracers, the Neumann boundary condition is often referred to as the flux boundary condition since by prescribing the normal derivative we prescribe the flux. [538](#), [543](#), [574](#)

**neutral diffusion** Neutral (also epineutral) diffusion is based on a diffusion tensor that yields diffusive fluxes oriented along neutral directions. It is also known as neutral diffusion. It parameterizes irreversible mixing of tracers along local neutral directions. Neutral diffusion is a refinement of isopycnal diffusion, motivated by the fact that potential density surfaces are not exactly neutral in a compressible, thermally expanding, and saline fluid like seawater. [534](#), [627](#), [640](#), [662](#), [664](#), [665](#)

**neutral directions** A neutral direction is a direction along which movement of a test fluid element leaves its local Archimedean buoyancy zero (i.e., where the test fluid element retains the same *in situ* density as the local environment), are directions where the fluid element remains neutrally buoyant. These directions are referred to as neutral directions. In effect, the test fluid element floats along a neutral direction. [625](#), [628](#), [640](#), [720](#)

**neutral helicity** Neutral helicity is given by  $\mathcal{H}^\gamma = \mathbf{N} \cdot (\nabla \times \mathbf{N}) = (-\alpha \nabla \Theta + \beta \nabla S) \cdot [\nabla \times (-\alpha \nabla \Theta + \beta \nabla S)]$ . Neutral surfaces are well defined only if the neutral helicity vanishes, which is not the case for a realistic seawater equation of state. [272](#)

**neutral tangent plane** A neutral tangent plane is a plane that sits tangent to a neutral direction at each point in the ocean. [640](#)

**neutrality condition** The neutrality refers to the property of neutral directions that the  $\alpha$  and  $\beta$  weighted gradients of  $\Theta$  and  $S$  are exactly balanced when aligned along a neutral direction. So when considering neutral directions from the perspective of the mixed test fluid element, the mixing-induced changes in  $\Theta$  precisely compensate mixing-induced changes in  $S$  as per the neutrality condition,  $\alpha \hat{\mathbf{t}}^\gamma \cdot \nabla \Theta = \beta \hat{\mathbf{t}}^\gamma \cdot \nabla S$ , where  $\hat{\mathbf{t}}^\gamma$  is the unit direction pointing along the neutral displacement. [644](#)

**Newton's law of viscous friction** Newton's law of viscous friction is a phenomenological constitutive relation that proposes that viscous friction within a fluid follows that for a Newtonian fluid, whereby there is a linear relationship between frictional stress and strain as realized by a constant molecular viscosity. Turbulent eddy viscosities are not constant in space and time, though they generally follow the same stress-strain mathematical form as for molecular diffusion. [513](#), [526](#)

**Newton's third law** To each action there is an equal and oppositely directed reaction. The third law holds for central forces, such as arise in Newtonian gravity and electrostatics. However, it does not hold for all forces, such as the Lorentz force acting on a moving charged particle. The strong form of Newton's third law is satisfied if the force of interaction between two particles is a central force (i.e.,  $(\mathbf{X}_{(i)} - \mathbf{X}_{(j)}) \times \mathbf{F}_{(ji)} = 0$ ) and it satisfies  $\mathbf{F}_{(ij)} = -\mathbf{F}_{(ji)}$ . The weak form of Newton's third law is satisfied if  $\mathbf{F}_{(ij)} = -\mathbf{F}_{(ji)}$  whereas the force is not central. [54](#), [78](#)

**no-slip boundary condition** The no-slip boundary condition means that fluid adheres to solid boundaries due to the role of viscosity, so that there is zero relative flow between the fluid and solid. The no-slip boundary condition leads to boundary layer formation for fluids flows adjacent to solid boundaries. The no-slip boundary condition cannot be imposed for perfect fluids, since such fluids have no viscosity. As discussed in the historical essay by [Anderson \(2005\)](#), it was the work of Prandtl in 1905 that first exposed the fundamental nature of the no-slip boundary condition, and its role in establishing boundary layers around solid bodies immersed in a fluid flow. Since the no-slip condition specifies the value of the velocity at the wall (fluid and wall have same velocity), it is often characterized mathematically as a kinematic boundary condition. However, it involves assumptions about stresses at the wall and so physically it arises as a result of dynamic processes, and so can be considered a dynamic boundary condition from this perspective. [111](#), [193](#), [260](#), [301](#)

**Noether's theorem** Noether's theorem states that for any symmetry of a physical system, there is a corresponding dynamical conservation law. For example, a classical particle system that exhibits Galilean symmetry maintains a constant linear momentum, constant mechanical energy, and constant angular momentum. A conservation law provides a dynamical constraint on the motion. The deduction of dynamical constraints is naturally arrived at using the methods of analytical mechanics, particularly through Hamilton's principle of stationary action. Noether's theorem is one of the most important foundational elements in 20th century mathematical physics. [205](#)

**non-advection flux** A non-advection flux is any flux of a scalar field that is not advective, with the canonical example being a diffusive flux. [670](#), [674](#)

**non-conservative process** A non-conservative process is a process that cannot be written as the convergence of a flux, but is instead written in the form of a source or sink in the budget equation for a property. [696](#)

**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-divergent barotropic model** The horizontally non-divergent barotropic model considers a single layer of homogeneous fluid with fixed thickness so that the horizontal flow is non-divergent,  $\nabla_h \cdot \mathbf{u} = 0$ . This model is only concerned with two-dimensional (horizontal) flow, and it focuses on vortical motion and it filters away all gravity waves. [115](#)

**non-divergent flow** Non-divergent flow is defined by  $\nabla \cdot \mathbf{v} = 0$ . This flow is realized by the Boussinesq ocean, which is the most general case of non-divergent flow. It is also realized by flow within a fluid of constant density. In two-dimensions, we consider the mechanics of horizontally non-divergent barotropic flow, whereby  $\nabla \mathbf{u} = 0$ . [157](#)

**orbital vorticity** Orbital vorticity (also known as curvature vorticity) is that component of relative vorticity arising from the curvature of the trajectory of a fluid parcel. [104](#)

**outcrop** An outcrop is the location at the ocean surface where a water property surface (e.g., an isopycnal, isotherm) meets at the surface. [27](#)

**parabolic partial differential equation** A parabolic partial differential equation a second order PDE that describes diffusive processes, in which disturbances smooth out over time but do not propagate as waves. The spatial portion of the parabolic operator is typically elliptical, so that a time independent parabolic equation reduces to an elliptic equation. 512, 533, 543

**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 canonical example, along with centrifugal and symmetric instabilities. Parcel instabilities are also referred to as local instabilities. xxi

**passive tracer** A passive tracer is a scalar field that satisfies the tracer equation, but it has zero impact on the velocity and is thus dynamically passive. Hence, a passive tracer provides a means to probe the advective-diffusive features of the flow without modifying it. 508, 513, 533, 543, 572, 673

**Peclet number** The Peclet number is a non-dimensional number that measures the strength of advective transport to diffusive transport:  $\text{Pe} = U L / \kappa$ , where  $U$  is a characteristic velocity scale,  $L$  is a length scale, and  $\kappa$  is a kinematic diffusivity. 546

**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 perfect fluid 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, 548

**permutation symbol** The permutation symbol is a three-index object that has indices that are totally anti-symmetric:  $\epsilon_{abc} = -\epsilon_{bac} = -\epsilon_{acb} = \epsilon_{cba}$ . It is the Cartesian coordinate representation of the Levi-Civita tensor,  $\varepsilon$ . 131

**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. 594

**Phillip's layering instability** The Phillip's layer instability arises from the spatial dependence of the eddy diffusivity. It explains how a vertically sheared turbulence in a stably stratified fluid can organize into horizontal layers separated by sharp density interfaces. 525, 541

**planetary Cartesian coordinates** Planetary Cartesian coordinates are Cartesian coordinates with their origin at the planet center and that rotate with the planet. 209, 230

**planetary geostrophy** Planetary geostrophy refers to a set of equations based on assuming the Rossby number is small, thus reducing the velocity equation to the diagnostic frictional geostrophic balance along with boundary forcing from winds and bottom drag. Planetary geostrophy is concerned with large-scale flow of a stably stratified laminar fluid whose vorticity is dominated by planetary rotation in the presence of planetary beta and topographic beta. It is commonly used to study the large-scale laminar ocean circulation. 144, 178, 225, 234, 241, 312, 313, 351

**planetary induction** Planetary induction is another name for the beta effect, whereby motion of a fluid within a rotating planetary frame induces relative vorticity due to the exchange with planetary vorticity. That is, planetary induction of vorticity is the change in

relative vorticity required to compensate for changes in planetary vorticity as fluid moves meridionally on the rotating Earth. 233, 234

**planetary vorticity** Planetary vorticity refers to the vorticity imparted to every fluid due to its existence in a rotating planetary reference frame. xix, 97

**Poisson's equation** Poisson's equation,  $\nabla^2 \Psi = \sigma$ , is a linear elliptic partial differential equation where the Laplacian of a scalar function,  $\Psi$ , equals to a source term,  $\sigma$ . This equation arises in many contexts within geophysical fluid mechanics. 125, 126, 556

**polarized** A wave field is polarized if it has a preferred sense of rotation. 606

**potential density** Potential density is the density of a fluid element moved to a reference pressure,  $p_R$ , while maintaining fixed specific entropy and fixed tracer concentration. For seawater, this movement is equivalent to fixed Conservative Temperature and fixed salinity, so that  $\varrho(S, \Theta) = \rho(S, \Theta, p_R)$ . 271, 294, 693

**potential enstrophy** Potential enstrophy is the domain integrated squared potential vorticity. 268

**potential enthalpy** Potential enthalpy is the enthalpy of a fluid element moved to a reference pressure,  $p_R$ , while maintaining fixed specific entropy and fixed tracer concentration:  $\mathcal{H}^{\text{pot}}(S, C) = \mathcal{H}(S, p_R, C)$ . It is used to define the Conservative Temperature,  $c_p^{\text{ref}} \Theta \equiv \mathcal{H}^{\text{pot}}(S, C) = \mathcal{H}(S, p_R, C)$ . 543, 660

**potential flow** Potential flow is characterized by an irrotational velocity ( $\nabla \times \mathbf{v} = 0$ ), which in turn means that the velocity can be expressed as the gradient of a scalar field,  $\mathbf{v} = \nabla \Psi$ . 93

**potential momentum** Potential momentum,  $\mathbf{M} = \mathbf{V}_{\text{cartesian}} + 2 \boldsymbol{\Omega} \times \mathbf{X}$ , is a constant of the motion for point particles moving along directions parallel to the geopotential (so long as the geopotential is a constant). Potential momentum is useful for the study of fronts on an  $f$ -plane, where the component of potential momentum along the front remains materially conserved for inviscid flows. 205

**potential temperature** Potential temperature,  $\theta$ , is the temperature that a fluid element reaches after performing an adiabatic and constant concentration translation to a standard pressure. For the ocean and atmosphere, the standard pressure is generally taken as the standard surface pressure. 171, 214, 659, 673

**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, 88, 123, 673

**potential vorticity inversion** Potential vorticity inversion is the mathematical procedure of recovering the balanced flow field (velocity, pressure/height, streamfunction) from a given distribution of potential vorticity, together with appropriate boundary conditions. Inversion requires solution to an elliptic partial differential equation. 312, 333

**potential vorticity pancake** Rather than being a property carried by each fluid particle, the pancake potential vorticity is carried by a finite fluid region and it is materially invariant for perfect fluid flows, even for a realistic equation of state. The concept was introduced by [Kooloth et al. \(2022\)](#). 284

**potential vorticity substance** Potential vorticity substance is the conceptual interpretation of potential vorticity as a materially conserved quantity that is transported and rearranged by the flow, with changes arising only from diabatic or frictional processes. 276, 289

**Prandtl number** The Prandtl number is the ratio of the kinematic viscosity to the kinematic diffusivity. 526

**Prandtl ratio** The Prandtl ratio is the ratio of the Coriolis parameter to the buoyancy frequency,  $f/N$ . It measures the relative importance of stratification and rotation. 345

**pressure torque** Pressure torque refers to the role of pressure gradients along boundaries in producing a source or sink of relative vorticity. Boundary pressure torques play a particularly important role in the vorticity budget for the depth integrated ocean flows, with the surface pressure contribution referred to as the atmospheric pressure torque, and the bottom pressure term is the bottom pressure torque. 248

**process method** In water mass transformation theory, the process method refers to an analysis of interior fluid processes that lead to the transformation of water across surfaces of constant water mass coordinates. Example processes for buoyancy-based water mass analysis include mixing and internal sources. In general, there are two complementary means to view interior water mass transformation: the *process method* and the *kinematic method*. The process method focuses on physical processes leading to movement of fluid across the water mass coordinate surface. The kinematic and process methods are mathematically identical and so they offer two means to compute the same transformation. The kinematic method tells us *how* transformation happens, whereas we make use of the process to help understand *why* transformation happens. 706

**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

**quasi-geostrophy** Quasi-geostrophy refers to the theory whereby the horizontal geostrophic flow is evolved via the leading order effects from the ageostrophic flow. Quasi-geostrophy is formulated in terms of the quasi-geostrophic potential vorticity, from which the flow, pressure, and buoyancy can be diagnosed. The equations are formulated by taking an asymptotic expansion of the  $\beta$ -plane Boussinesq ocean equations with a small Rossby number and a large Richardson number. 312, 313, 375

**quasi-Stokes transport** The quasi-Stokes transport refers to the eddy-driven transport arising from use of the TRM kinematic framework. 612, 620, 631

**Rayleigh drag** Rayleigh drag is a particular form of friction that makes use of the acceleration,  $-\gamma \mathbf{v}$ , with  $\gamma^{-1}$  having dimensions of time. Rayleigh drag adds an acceleration that drags all flow towards a state of rest. Notably, Rayleigh drag has zero scale selectivity,

which contrasts to Laplacian friction suggested by kinetic theory, whereby friction acts preferentially at the small scales. 143, 182, 201, 205, 245

**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

**reciprocity condition** The reciprocity condition refers to properties of Green's functions according to their behavior when the source and field points are swapped. Green's functions for self adjoint operators, such as the Laplacian, have full symmetry under such swaps. For parabolic operators, such as the diffusion equation, reciprocity involves the adjoint Green's function and more care is needed. 145, 573, 576

**rectification** Rectification occurs if oscillatory or turbulent motions produce a mean flow or other slowly varying effect through nonlinear terms in the governing equations. 485, 591

**reduced gravity** Reduced gravity is Earth's gravitational acceleration scaled by the fractional density contrast, capturing how weakly buoyancy restores motion in a nearly incompressible ocean. It is the effective gravitational acceleration that directly affects motions driven by density differences. For two constant density layers,  $\rho_1$  (upper) and  $\rho_2$  (lower), the reduced gravity is  $g' = g(\rho_2 - \rho_1)/\rho_2 \approx g(\rho_2 - \rho_1)/\rho_{\text{ref}}$ , where we commonly use the Boussinesq reference density for the denominator. The reduced gravity measures the strength that buoyancy acts on an interface. 16, 17

**reduced gravity model** The reduced gravity model describes a dynamically active layer of uniform density,  $\rho_1$ , above a stagnant layer of density,  $\rho_2$ , and below a fluid of zero density,  $\rho_0 = 0$ . It is often referred to as the 1.5 layer model. This theoretical model has been used, to some success, as an idealization of the upper ocean circulation whereby an active layer (e.g., the region above the ocean pycnocline), sits above an inactive layer (the abyss) of much smaller motion (here assumed to be zero motion). 15

**relative potential vorticity** The relative potential vorticity is the quasi-geostrophic potential vorticity minus the planetary vorticity,  $f$ . 330

**relative vorticity** The relative vorticity is the vorticity in a fluid as computed relative to the moving planet:  $\omega = \nabla \times \mathbf{v}$ . 97

**residual mean velocity** The residual mean velocity is the sum of the Eulerian mean plus the eddy-induced velocity,  $\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*$ . The residual mean provides the net advective effect on tracers from eddy advection plus mean advection. 553, 628, 662

**Reynolds decomposition** A Reynolds decomposition is a method to decompose the flow field,  $\Psi$ , into a mean,  $\bar{\Psi}$ , plus a departure from the mean (an eddy),  $\Psi'$ . This decomposition satisfies the following three properties: (i)  $\bar{\Psi}' = 0$  (mean of eddy vanishes); (ii)  $\bar{\bar{\Psi}} = \bar{\Psi}$  (mean of mean equals mean); (iii)  $A\bar{\Psi}' = A\bar{\Psi}$  (mean of constant equals the constant).. 594

**Reynolds stress tensor** The turbulent contribution to the kinetic stress tensor is known as the Reynolds stress tensor. 493

**Reynolds transport theorem** The Reynolds transport theorem is the Leibniz-Reynolds transport theorem applied to a Lagrangian region. 107, 535, 586

**Richardson number** The Richardson number is the non-dimensional ratio of the vertical stratification to the vertical shear of the horizontal flow. The Richardson number appears throughout our study of stratified geophysical fluid flows. 343

**rigid lid approximation** The rigid lid approximation is a boundary condition that assumes vertical fluid velocity vanishes at the top boundary of the fluid column. It is commonly used for studies of the large-scale ocean circulation. 27, 129

**Robin boundary condition** The Robin or mixed boundary condition combines aspects of the Neumann and Dirichlet boundary conditions by specifying the sum of the normal derivative along the boundary (like the Neumann condition), as well as the value along the boundary (like the Dirichlet condition). 574

**Rossby effect** The Rossby effect is a northward drift of vortices found in a rotating reference frame. 140

**Rossby number** The Rossby number is the non-dimensional ratio of the horizontal material acceleration (acceleration of a fluid particle) to the Coriolis acceleration. Geostrophic flows in the ocean and atmosphere are characterized by Rossby numbers much smaller than unity. The Rossby number offers a measure of the departure of the flow from solid body rotation, as determined by the radial component of the vorticity. 40, 312

**Rossby potential vorticity** The Rossby potential vorticity (also referred to as the shallow water potential vorticity), is given by  $(f + \zeta)/h$ , where  $f$  is the Coriolis parameter,  $\zeta$  is the relative vorticity, and  $h$  is the layer thickness. It is the form of potential vorticity appearing in the shallow water models, and is the discrete limit of the potential vorticity found in continuous isopycnal coordinates. 163, 169, 175

**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. 141, 178, 200, 323

**rotation tensor** The rotation tensor is the anti-symmetric portion of the velocity gradient tensor, with elements given by twice the vorticity,  $R^{mn} = -\epsilon^{mnp} \omega_p/2$  with  $\boldsymbol{\omega} = \nabla \times \mathbf{v}$  the vorticity. The rotation tensor measures the ability of fluid flow to rotate without dilating fluid elements. 91, 110, 225, 255

**salinity** Salinity is 1000 times the mass concentration of salt within a sample of seawater. More precisely, the salinity,  $S$ , is the *Absolute Salinity*. Absolute Salinity is distinct from the *practical salinity* determined by conductivity measurements. IOC et al. (2010) provides a full accounting of the theory and practice of ocean salinity, with ? providing a summary of their interpretation within ocean models. 659, 660, 669, 691

**scalar mechanics** Scalar fields provide a number at each point in space and time. Scalar mechanics is the study of scalar fields, such as tracers, density, and other thermodynamical tracers. Mathematically, the advection-diffusion equation is the most prominent equation arising in scalar mechanics. 508

**scale selective** Scale selective refers to the property of a differential operator to act preferentially on a particular scale relative to another. This term is used most commonly in reference to Laplacian-like operators used to model molecular or turbulent diffusion, whereby small scales in the concentration field are dissipated more rapidly than large scales. 527

**screened Poisson equation** The screened Poisson equation is  $(-\nabla^2 + \mu^2)\psi = \Lambda$ , with  $\mu^2 > 0$ . This equation occurs in the Yukawa theory of mesons and the Debye-Hückel screening of ionic solutions. Although not of direct use for our studies of geophysical fluid mechanics, we see in VOLUME 4 that there is a direct connection to the Helmholtz equation (where  $\mu^2 < 0$ ), which is important for geophysical wave theory. 146

**secondary circulation** Secondary circulation is a relatively weak, typically cross-stream (transverse) circulation that accompanies a dominant primary flow, often arising from imbalances (e.g., ageostrophic flow), friction, or nonlinearities. It consists of motions in the plane perpendicular to the main flow direction. While usually much smaller in magnitude, secondary circulations play an outsized role in transporting mass, momentum, tracers, and vorticity across flow structures. 140, 142, 633

**self-adjoint** A self-adjoint operator is a linear operator that equals to its adjoint, where adjoint is defined relative to a chosen inner product. For real matrices, the adjoint is the transpose, so that a symmetric matrix is a self-adjoint matrix. For complex operators, self-adjoint refers to Hermitian operators (such as considered in quantum mechanics). For partial differential equations, self-adjointness is determined by properties of the differential operator, the boundary conditions, and the inner product. The eigenfunctions of self-adjoint operators form a complete set over the space of functions that satisfy the same boundary conditions as the eigenfunctions. We make particular use of this property when performing an eigenfunction expansion of Green's functions. 145, 537

**semi-geostrophic** The semi-geostrophic equations extend geostrophic balance to include finite ageostrophic effects, especially those associated with fronts, jets, and strong horizontal strain. It sits between quasi-geostrophic theory and the primitive equations. 324

**shallow water approximation** The shallow water approximation assumes fluid to be vertically partitioned into homogeneous layers (with constant density), and with the pressure to be hydrostatic. As a result, fluid motion is comprised of vertically coherent columns moving within each layer, with the columns able to expand or contract according to horizontal flow convergence within each layer. 2, 5

**shallow water model** The shallow water model is a depth-integrated fluid model that describes the horizontal motion of a layer whose thickness is small compared with its horizontal scales, retaining free-surface gravity and rotation while filtering out vertical structure. 2, 3, 39, 88, 452, 471, 476

**shallow water potential vorticity** The shallow water potential vorticity, also referred to as the Rossby potential vorticity, is given by  $(f + \zeta)/h$ , where  $f$  is the Coriolis parameter,  $\zeta$  is

the relative vorticity, and  $h$  is the layer thickness. It is the form of potential vorticity appearing in the shallow water models, and is the discrete limit of the potential vorticity found in continuous isopycnal coordinates. 163, 169, 175

**shallow water Sverdrup balance** The shallow water Sverdrup balance is given mathematically by  $\beta v = Q^{\text{pg}} \mathbf{u} \cdot \nabla h + \hat{z} \cdot (\nabla \times \mathbf{F}^{\text{wind}})$ . This balance states how horizontal advection of layer thickness (first right hand side term) plus the wind stress curl (second term) balance meridional motion for flow on a rotating sphere (beta effect on left hand side). 184

**sharp representation** The sharp representation refers to the  $(2,0)$  representation of second order tensors. For example, the sharp representation of the diffusion tensor is  $\mathbf{K} = K^{mn} \mathbf{e}_n \otimes \mathbf{e}_m$ . The sharp representation follows the musical nomenclature for second order tensors, where we have sharp, flat, and natural representations. 531

**sifting property** The sifting property is a basic property of the Dirac delta, whereby the integral of the Dirac delta times another function yields the function evaluated at the point where the Dirac delta is singular. In one dimension, this property reads  $\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$ . 579

**simple ideal gas** A simple ideal gas is an ideal gas in which the internal energy is a linear function of temperature, so that the heat capacities are constants. 257

**skew diffusion** Skew diffusion is the process of transporting tracers through the convergence of skew tracer fluxes. Skew tracer fluxes result from contracting the tracer concentration gradient with the anti-symmetric skew diffusion tensor, and they differ by a total curl from the advective tracer fluxes. Hence, skew fluxes and advective fluxes have the same convergence, so that the skew diffusion process and the advection process affect the same time tendency on a tracer. 554, 627, 665

**skew tracer flux** The skew tracer flux,  $-\nabla C \times \rho \Psi^*$ , has a convergence equal to that of the advective tracer flux,  $\rho C \mathbf{v}$ , with  $\mathbf{v} = \nabla \times \Psi^*$ . The skew tracer flux is neither upgradient nor downgradient. Rather, it is oriented parallel to isosurfaces of tracer concentration. In some contexts, such as when studying the effects from mesoscale eddies, it can be more convenient to focus on the skew flux than the advective flux. 553, 554

**skew-symmetric tensor** A skew-symmetric (also anti-symmetric) tensor equals to minus its transpose,  $\mathbf{A} = -\mathbf{A}^T$ . 627

**skewsion** Skewsion is any process that leads to tracer transport via skew fluxes, with skew diffusion a particular example. It is a stylized name that emphasizes its connection to advection. 555

**slope Burger number** The slope Burger number is a non-dimensional number given by  $Bu = [(N/f) \tan \varphi]^2$ , where  $\varphi$  is the slope of the ocean bottom topography. It is commonly encountered when studying flow next to the bottom. 344

**solenoid** A solenoid is a tube region in the fluid that is perpendicular to both  $\nabla \rho$  and  $\nabla p$ . There are no solenoids for barotropic flows, whereby  $p = p(\rho)$ . For baroclinic flow, solenoids are associated with a torque that affects vorticity. 222

**Soret effect** The Soret effect is the diffusion of matter due to a temperature gradient. 513, 661

**specific thickness** The specific thickness is the partial derivative,  $\partial s/\partial z$ , of the generalized vertical coordinate,  $s(x, y, z, t)$ , with respect to the vertical. It provides the Jacobian of transformation between  $z$ -coordinates and generalized vertical coordinates. 610

**splat** Splat is given by the doubly-contracted strain rate tensor,  $\mathbf{S}: \mathbf{S} = S^{mn} S_{mn} \geq 0$ . This term appears as part of the source term for pressure in non-divergent flows. It is so-named since it is large when a fluid element is squashed in a manner increasing fluid strains, akin to how strains appear when a fluid impacts or “splats” against a solid obstacle. Imagine a water balloon thrown against a wall. 131

**steric** Steric effects generally refer to properties of a substance associated with the space occupied by atoms. In the sea level context, steric effects refer to changes in sea level associated with density changes, with changes in density associated with changes in the volume occupied by seawater molecules. Changes in global mean sea level arising from changes in the global mean density are called global steric sealevel changes. 679

**steric sea level** The steric sea level is the thickness of a fluid layer bounded below by an isobar and above by the sea surface. It is commonly used for ocean circulation studies to map geostrophic flows. The dynamic topography is also referred to as the ocean dynamic topography. 679

**stirring** Stirring is the tracer transport process affected by advection. Stirring generally acts to reversibly rearrange fluid elements and to increase tracer gradients, particularly when the flow is turbulent. By increasing gradients, and thus increasing the surface area of tracer contours, stirring enhances the ability of molecular diffusion to affect irreversible mixing that acts to dissipate gradients. 545

**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. 485, 595, 600

**Stokes mean** The Lagrangian mean (mean computed on fluid elements) minus the Eulerian mean (mean computed at a fixed point) defines the Stokes mean. Note that the literature sometimes refers to the Stokes mean as the “Stokes correction”. We avoid that terminology in order to avoid the spurious notion that one type of mean operator is more correct than the other. Instead, a mean operator is subjectively chosen based on its suitability to a particular question. 595

**Stokes' theorem** The vector calculus form of Stokes' theorem says that for a simply connected surface,  $\oint_{\partial\Omega} \mathbf{A} \cdot d\mathbf{x} = \int_{\Omega} (\nabla \times \mathbf{A}) \cdot \hat{\mathbf{n}} d\mathcal{S}$ , where  $\mathbf{A}$  is a vector,  $\Omega$  is a simply connected surface with boundary  $\partial\Omega$ ,  $d\mathbf{x}$  is a line increment around the boundary, and  $\hat{\mathbf{n}}$  is an outward normal to the surface and which defines the orientation. 88, 94, 150, 429

**Stommel model** The Stommel model The Stommel model is a minimal theory explaining why wind-driven ocean gyres have a narrow, strong western boundary current (like the Gulf Stream), rather than a symmetric circulation, with the asymmetry holding even when the winds are symmetric. Planetary vorticity variation (the beta effect) plus linear bottom friction forces the return flow of a wind-driven gyre to be more intense (and thus narrow) along the western side of the basin. 187

**storage term** In the analysis of budgets, we sometimes refer to time tendency terms as the storage term, which contrasts to fluxes crossing boundaries of the budget region. 711

**strain rate tensor** The strain rate tensor,  $\mathbf{S}$ , is the symmetric portion of the velocity gradient tensor. The strain rate tensor measures the ability of fluid flow to deform fluid elements through stretching, tilting, straining, and dilation. 91, 97, 109–111, 134, 225, 255, 552

**streamfunction** In two-dimensional non-divergent flow, we can define a streamfunction, whose curl yields the velocity field at each time instance. Consequently, isolines of the streamfunction are streamlines. The streamfunction is arbitrary up to a constant, thus representing a simple form of gauge symmetry. In three-dimensional non-divergent flow we can introduce a vector streamfunction, whose curl yields the velocity field. The vector streamfunction is arbitrary up to the gradient of a scalar field, which represents a form of gauge symmetry. 283

**streamlines** A streamline is a curve in space that is tangent, at each time instance, to the fluid particle velocity. Streamlines and pathlines are identical only for steady flows, whereas they are generally distinct for unsteady flows. 83, 99

**streamtube** A streamtube is a bundle of streamlines crossing through an arbitrary closed curve. At each time instance, the flow velocity is tangent to the streamtube sides. Furthermore, when the flow is steady then streamlines are identical to material particle pathlines. Hence, a streamtube is a material tube for steady flow, in which case no fluid particles cross the streamtube boundary. 100

**stress tensor** The stress tensor is a second order tensor that organizes the stresses acting within a fluid. The stress tensor components have dimension of force per unit area, and each component is oriented by both the direction of the stress and the outward normal of the surface on which the stress acts. 526

**subgrid scale** Subgrid scale (SGS) refers to processes happening at scales smaller than the grid (length) scale of either a numerical model or field measurement. The parameterization of subgrid scale process remains an active area of ocean and atmospheric physics, given the importance of such processes for the large scale circulation. 31, 623, 662

**surface transformation** That portion of water mass transformation theory concerned with how surface boundary fluxes contribute to water mass transformation. Surface transformation theory forms the focus of many studies of water mass transformation because it only requires surface boundary information, which is generally more accessible than information from interior ocean mixing processes or bottom geothermal processes. Furthermore, much of the transformation of water occurs in surface regions since this region is home to large contributions from surface boundary fluxes and associated ocean mixing. 719

**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. The traditional form of the Sverdrup balance arises when we assume the flow takes place over a flat bottom and the free surface undulations are negligible, so that  $\beta v = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}^{\text{wind}})$ . This balance helps to explain the steady equatorward ocean circulation appearing in the eastern portion of middle latitude gyres. xi, 184, 251, 351, 363, 365, 366

**tangent plane** The tangent plane makes use of Cartesian coordinates defined local to a point on the rotating planet. The  $f$ -plane and  $\beta$ -plane are the two common forms of the tangent

plane approximation encountered in this book. It is important to note that the tangent plane approximation is not based on assuming a locally flat sphere. Rather, the tangent plane approximation as based on assuming a locally flat geopotential. 83, 115, 147

**tangent plane Cartesian coordinates** Tangent plane Cartesian coordinates are Cartesian coordinates defined local to a point on the rotating planet, with the approximation based on assuming a locally flat geopotential. The  $f$ -plane and  $\beta$ -plane are the two forms of the tangent plane approximation encountered in this book. It is important to note that the tangent plane approximation is not based on assuming a locally flat sphere. Rather, the tangent plane approximation as based on assuming a locally flat geopotential. Consequently, the approximation makes use of the effective gravitational acceleration (minus the gradient of the geopotential) that includes both central gravity acceleration plus the planetary centrifugal acceleration. 115, 147

**Taylor column** A Taylor column refers to a vertical column of fluid that is stiffened due to the effects of the rotating planet. Taylor columns are aligned with the axis of rotation, which for a spherical planet means they are aligned with the polar axis. 15

**Taylor-Bretherton identity** In a horizontally non-divergent barotropic fluid, the Taylor-Bretherton identity provides a connection between the vorticity flux and the momentum flux. For a quasi-geostrophic fluid, the Taylor–Bretherton identity states that eddy buoyancy fluxes are exactly equivalent to eddy PV fluxes projected along mean isopycnals. In this manner, quasi-geostrophic eddies transport buoyancy only by transporting potential vorticity along mean isopycnals. 124

**Taylor-Proudman effect** The Taylor-Proudman effect refers to the vertical stiffening of fluid columns that appears in an unstratified, homogeneous, inviscid fluid in geostrophic balance. 15, 178, 185, 351, 378

**tensor product** A tensor product between two tensors produces a third tensor whose order is the sum of the terms. For example, the tensor product of two  $(1, 0)$  vectors produces a  $(2, 0)$  tensor, such as the second order kinetic stress tensor,  $\mathbb{T}^{\text{kinetic}} = -\rho \mathbf{v} \otimes \mathbf{v}$ . In components, we have  $(\mathbf{v} \otimes \mathbf{v})^{ab} = v^a v^b$ . The tensor product is an expression for tensors of the outer product from linear algebra. 46

**thermal expansion coefficient** The thermal expansion coefficient is a response function that measures minus the relative change in density as the temperature is altered while holding the pressure and matter concentration fixed. It is typically positive, so that density decreases as temperature increases. Freshwater near its freezing point is an important counterexample, where the negative thermal expansion coefficient allows for solid ice to float on liquid water. 661

**thermal wind balance** Thermal wind balance refers to diagnostic balance between the vertical shear in the horizontal geostrophic flows and horizontal gradient in the density field. 42, 351, 401

**thermobaricity** Thermobaricity is a physical process that results in density changes due to the pressure dependence of the ratio of the thermal expansion coefficient to the haline contraction coefficient. 664–667, 684

**thermodynamic equilibrium** In our study of equilibrium thermodynamics, we are concerned with macroscopic fluid systems whose evolution tends toward states whose properties are determined by intrinsic factors rather than depending on memory of previous external influences. These particular macrostates are known as thermodynamic equilibria. At a basic level, a system in thermodynamic equilibrium could remain in that state for all time, with details of the equilibrium dependent on the constraints imposed on the system. When constraints are removed, then a system generally transitions to another thermodynamic equilibria. Note that “for all time” is a loaded term. More precisely, we mean “for a time extremely long compared to any time scale relevant to the physical system under consideration”. [511](#)

**thermosteric** The temperature component of steric sea level change is referred to as the thermosteric sea level. For global mean sea level, thermosteric effects dominate over halosteric effects, with global halosteric effects negligible. Locally, however, thermosteric and halosteric effects can be sizable. [679](#)

**thickness weighted averaging** Thickness weighted averaging refers to the use of a layer thickness when performing a low pass filter average for dynamical fields. It is commonly used for shallow water models and isentropic vertical coordinate models. [58](#), [405](#), [471](#), [483](#), [610](#)

**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](#)

**topographic beta effect** Topographic slopes can act just like planetary beta to impart a source to the vorticity field. In analogy to planetary beta effect, we use the term topographic beta effect, with topographic beta the reason there are low-frequency and rotation-dominated waves that propagate along sloping topography (continental slopes, ridges, seamount flanks). Namely, the depth variation of the bottom acts like a  $\beta$ -effect for potential vorticity. [177](#)

**topographic form stress** Topographic form stress is the contribution from pressure acting to produce a horizontal acceleration and that acts at the interface between a fluid and solid-earth topography. Topographic form stress plays an important role in the force balance of flows next to sloping solid earth boundaries, such as in the Southern Ocean and western boundaries. [78](#), [360](#)

**topographic Rossby waves** Topographic Rossby waves are low-frequency, rotation-dominated waves that propagate along sloping topography (continental slopes, ridges, seamount flanks) because depth variation of the bottom acts like a  $\beta$ -effect for potential vorticity (topographic beta effect). [178](#)

**topographic topographic Sverdrup balance** Topographic Sverdrup balance is the Sverdrup balance with the added contribution from bottom pressure torques. [252](#), [368](#)

**tracer equation** The tracer equation is a partial differential equation describing the evolution of tracer fields, such as material tracers and thermodynamic tracers (potential enthalpy). Tracers are intensive scalar fluid properties that evolve through advection (in an Eulerian description), diffusion, and source/sinks. [456](#)

**Traditional Approximation** The traditional approximation comprises three approximations that come as a package in order to maintain physical consistency. (A) It sets to zero the Coriolis terms in the horizontal momentum equations that involve the vertical velocity, thus retaining only the local vertical component of the earth's angular rotation vector. (B) It drops the metric terms,  $uw/r$  and  $vw/r$ , associated with the vertical velocity as they appear in the horizontal momentum equations. (C) The shallow fluid approximation and both parts of the traditional approximation must be taken together in order to maintain a consistent energy and angular momentum conservation principle for the resulting equations. 8, 237, 238, 462

**transformation** In the context of water mass analysis, transformation refers to the process of moving water across surfaces defined by water mass coordinates. For example, with water mass configuration space defined by surfaces of constant Conservative Temperature,  $\Theta$ , then when water moves across  $\Theta$  surfaces we say that the water has been transformed. Processes leading to transformation arise from mixing, solar radiation, and chemical reactions. Correspondingly, we measure zero motion along a water mass coordinate axis when the property defining that axis remains materially unchanged. For example, adiabatic and isohaline processes such as linear waves can render nontrivial motion in geographical/depth space whereas they lead to no motion in  $(S, \Theta)$  space. 692–694

**transformed residual mean** The transformed residual mean is the Eulerian mean field that when evaluated at the mean vertical position of an isopycnal is equal to the isopycnal thickness weighted average of the field. 611

**turbulence closure** When we measure fluid motions in the laboratory or field, we generally do not measure the motions at scales on the order of  $L_{\text{macro}} \approx 10^{-4}$  m. That is, our measurement devices generally have a spatial resolution coarser than  $L_{\text{macro}}$ , so that  $L_{\text{measure}} \gg L_{\text{macro}}$ . Likewise, numerical simulations are generally designed using discrete grids with length scales  $L_{\text{numerical}} \gg L_{\text{macro}}$ . The equations describing motions at the measurement/simulation length scales involve effects from fluctuations occurring at the smaller (unmeasured) scales. The reason for this coupling is that the fluid equations are nonlinear, and with the nonlinearities leading to an interaction across spatial scales. These fluctuations, generally associated with turbulent or chaotic motions, have statistical correlations that can play a role, sometimes a dominant role, in the evolution of flow features at the measured/simulated scales. The parameterization of these correlations in terms of measured/simulated motions constitutes the turbulence closure problem. 596

**turbulent diffusion** Diffusion is the physical process by which a field, such as a tracer, spreads in space over time due to random motion. Turbulent diffusion arises from random motion within a turbulent flow. The net flux moves from regions of higher concentration to regions of lower. The efficiency for diffusive transport in a turbulent fluid is many orders higher than that from molecular diffusion. Chapter 13 in [Vallis \(2017\)](#) provides a pedagogical discussion of turbulent diffusion in geophysical flows. 521

**turbulent flow** Turbulent flow is characterized by a quasi-random fluid flow that supports an enhanced amount of mixing and transport of fluid properties. Geophysical fluid flows are predominantly turbulent across multiple scales. 521

**turbulent velocity scale** The turbulent velocity scale is a measure of the velocity of random turbulent motion that leads to turbulent diffusion. The turbulent velocity scale is much

smaller than molecular speeds. Determination of turbulent length and velocity scales is subject to large uncertainties and variations given the multiple regimes of turbulence exhibited by geophysical flows. As a result, tracer transport by turbulent flows has remained a topic of much research since the early 20th century. 525

**two dimensional turbulence** Two-dimensional turbulence is the chaotic motion of two-dimensional non-divergent fluid flow, and it is characterized by a cascade of kinetic energy to the large scales (inverse cascade) whereas potential enstrophy cascades to the small scales (direct cascade). These properties of two-dimensional turbulence are shared by quasi-geostrophic turbulence. 128

**upwind advective flux** An upwind advective flux is the discretized transport of a fluid property (e.g., tracer or momentum) across a grid cell face that uses the upstream value of the quantity, determined by the sign of the velocity normal to that face. 32

**vector invariant** The vector-invariant form of the velocity equation is based on replacing the material time derivative with the vorticity and kinetic energy, through use of the identity  $(\mathbf{v} \cdot \nabla)\mathbf{v} = \boldsymbol{\omega} \times \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2$ , where  $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ . 27, 28, 461

**velocity gradient tensor** The velocity gradient tensor,  $\mathbf{G}$ , is the Eulerian means to quantify flow deformation. Its elements are the spatial derivatives of the velocity field. It is commonly decomposed into the symmetric strain rate tensor (also known as the deformation rate tensor) plus the anti-symmetric rotation tensor. 109, 551

**vertical gauge** The vertical gauge refers to the choice to set to zero the third component of the vector streamfunction, whose curl yields a non-divergent transport velocity. The vertical gauge is commonly used for parameterized eddy-induced velocities. 557

**vortex filament** A vortex tube is the accumulation of vortex lines passing through a closed loop. A related, but distinct, concept arises when a vortex tube is surrounded by irrotational flow, with such tubes referred to as vortex filaments (see Section 1.4 of [Saffman \(1992\)](#)). 99, 100

**vortex line** A vortex line is a curve in the fluid that is instantaneously tangent to the vorticity at each point along the curve. As noted in Section 1.4 of [Saffman \(1992\)](#), a vortex filament is a vortex tube surrounded by irrotational fluid, which contrasts to the more general concept of a vortex line. 99

**vortex stretching** Vortex stretching is the process whereby a material fluid region is stretched or compressed, in which case vorticity of the region is either increased in magnitude (stretching) or decreased (compression). Vortex stretching is an important part of planetary flows given that motion through the planetary vorticity field can induce stretching or compression to thus alter the fluid's relative vorticity, as per the Sverdrup balance. The changes in relative vorticity arising from vortex stretching are linked to the conservation of angular momentum. 170, 228

**vortex tilting** Vortex tilting is the process whereby the projection of the vorticity vector is altered when a material fluid region moves through a region with shears. 170

**vortex tube** A vortex tube is the accumulation of vortex lines passing through a closed loop. 99, 100

**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, 88, 89, 428

**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, 782

**Walin analysis** Walin analysis is another name for water mass transformation theory, which focuses on mass distribution functions and follows the methods introduced by [Walin \(1977\)](#) and [Walin \(1982\)](#). 702

**water mass** In ocean physics, a water mass refers to a region of seawater characterized by a suite of physical properties. Water masses often originate through extremely large buoyancy fluxes at the high latitudes that form waters such as the Antarctic Bottom Water and North Atlantic Deep Water. As these waters enter the ocean interior they are transported over basin scales while they are eroded or transformed by irreversible mixing or sources. Water masses and their properties offer a conceptual means to partition or bin the fluid into distinct classes whose origin, movement, and transformation can be measured, modeled, and studied. Scalar properties generally used to classify water masses are simpler to measure than vector properties such as velocity. Hence, a water mass perspective offers the means to infer ocean circulation within the space of ocean properties without directly measuring vector fields. 689

**water mass configuration space** Water mass configuration space (denoted by  $\mathbf{q}$ -space) is the space we work within to study water mass transformations. This space has some or all of its coordinates set by properties other than geographic/depth coordinates. Operationally, we fill  $\mathbf{q}$ -space by forming histograms that result in a  $\mathbf{q}$ -space distribution of the fluid properties. For example, a one-dimensional  $\mathbf{q}$ -space results from binning the ocean according to potential density, whereas retaining latitudinal information along with potential density renders a two-dimensional  $\mathbf{q}$ -space. Typically  $\mathbf{q}$ -space has three or fewer dimensions, given the three dimensionality of  $\mathbf{x}$ -space. There is no implied constraint that any of the  $\mathbf{q}$ -space coordinates are monotonic with respect to  $\mathbf{x}$ -space. Indeed, there is no presumption that points in  $\mathbf{q}$ -space maintain a 1-to-1 relation to points in  $\mathbf{x}$ -space. For example, many points in  $\mathbf{x}$ -space may fall into a single point (or bin) within  $\mathbf{q}$ -space. 691

**water mass transformation** Water mass transformation theory examines the budgets for fluid mass and tracer mass within layers or classes defined by properties such as Archimedean buoyancy, Conservative Temperature, salinity, or biogeochemical tracers. The theory is concerned with how processes affect the evolution of fluid within property space and in the characterization of circulation inferred from this evolution. This kinematic lens is distinct from the Eulerian and Lagrangian kinematics, with particular use for examining questions where the irreversible transformation of properties plays a central role. The theory originates from the work of [Walin \(1977\)](#) and [Walin \(1982\)](#), and it has found many advances in both formalism and application since then (see [Groeskamp et al. \(2019\)](#) for a review). 689, 691, 702

**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](#)

**zonal mean** The zonal mean refers to a spatial averaging operation computed by a line integral of a field over the full longitudinal extent of the domain, and then a division by the zonal length of the domain. The zonal mean is particularly relevant for regions that are zonally periodic, such as the atmosphere and the ocean in the latitudes around the Drake Passage. Notably, the resulting zonally averaged field is independent of longitude. [594](#)



---

## Appendix B

### LIST OF ACRONYMS

**AI** artificial intelligence xvii

**GFD** geophysical fluid dynamics ix

**GFM** geophysical fluid mechanics vii, 592

**GLM** generalized Lagrangian mean 591, 595, 596

**GVC** generalized vertical coordinate xx, 404, 407–409, 412–416, 418, 419, 423, 424, 426, 430, 433, 438

**PG** planetary geostrophy 178, 312, 313

**QG** quasi-geostrophy 312, 313, 375

**TRM** transformed residual mean 611

**TWA** thickness weighted average 405, 471, 472, 483–485, 488, 491–494, 499, 500, 502, 610



## 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 $A^{mn} = -A^{nm}$
$A^\nu$	Avogadro's number: $A^\nu = 6.0222 \times 10^{23}$ 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: $Q_m > 0 \text{ enters ocean}$
$Q_m$	mass flux (mass per surface area per time) across ocean surface: $Q_m dS = Q_m dA$
$Q_c$	turbulent tracer flux (tracer per horiz area per time) across ocean surface: $Q_c > 0 \text{ enters ocean}$
$Q_c$	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

- Abernathay, R., D. Ferreira, and A. Klocker, Diagnostics of isopycnal mixing in a circumpolar channel, *Ocean Modelling*, 72(0), 1 – 16, doi:10.1016/j.ocemod.2013.07.004, 2013. 625, 638, 645
- Acheson, D., *Elementary Fluid Dynamics*, Oxford Applied Mathematics and Computing Science Series, Oxford, Oxford, 1990. 102
- Adcock, S. T., and D. P. Marshall, Interactions between geostrophic eddies and the mean circulation over large-scale bottom topography, *Journal of Physical Oceanography*, 30, 3223–3238, 2000. 202
- Adcroft, A., and J.-M. Campin, Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models, *Ocean Modelling*, 7, 269–284, doi:10.1016/j.ocemod.2003.09.003, 2004. 411
- Adcroft, A., R. Hallberg, and M. Harrison, A finite volume discretization of the pressure gradient force using analytic integration, *Ocean Modelling*, 22, 106–113, doi:10.1016/j.ocemod.2008.02.001, 2008. 462, 463, 466, 474
- Adcroft, A., W. Anderson, C. Blanton, M. Bushuk, C. O. Dufour, J. P. Dunne, S. M. Griffies, R. W. Hallberg, M. J. Harrison, I. Held, M. Jansen, J. John, J. P. Krasting, A. Langenhorst, S. Legg, Z. Liang, C. McHugh, B. G. Reichl, A. Radhakrishnan, T. Rosati, B. Samuels, A. Shao, R. J. Stouffer, M. Winton, A. T. Wittenberg, B. Xiang, N. Zadeh, and R. Zhang, The GFDL global ocean and sea ice model OM4.0: Model description and simulation features, *Journal of Advances in Modeling Earth Systems*, 11, 3167–3211, doi:10.1029/2019MS001726, 2019. 37
- Aiki, H., T. Jacobson, and T. Yamagata, Parameterizing ocean eddy transports from surface to bottom, *Journal of Geophysical Research*, 31, L19 302, doi:10.1029/2004GL020703, 2004. 634, 639
- Anderson, J., Ludwig Prandtl's boundary layer, *Physics Today*, 58, 42–48, doi:10.1063/1.2169443, 2005. 768
- 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, An exact theory of nonlinear waves on a Lagrangian-mean flow, *Journal of Fluid Mechanics*, 89, 609–646, doi:10.1017/S0022112078002773, 1978a. 591, 603

- Andrews, D. G., and M. E. McIntyre, On wave action and its relatives, *Journal of Fluid Mechanics*, 89, 647–664, doi:10.1017/S0022112078002785, 1978b. 591, 603
- Arakawa, A., and V. R. Lamb, The UCLA general circulation model, in *Methods in Computational Physics: General Circulation Models of the Atmosphere*, vol. 17, edited by J. Chang, pp. 174–265, Academic Press, 1977. 112
- Badin, G., and F. Crisciani, *Variational formulation of fluid and geophysical fluid dynamics*, Springer International Publishing, Switzerland, 218 pages, 2018. xii
- Balwada, D., Q. Xiao, K. S. Smith, R. Abernathey, and A. R. Gray, Vertical fluxes conditioned on vorticity and strain reveal submesoscale ventilation, *Journal of Physical Oceanography*, 51(9), 2883–2901, doi:10.1175/JPO-D-21-0016.1, 2021. 552
- Barthel, A., A. Hogg, S. Waterman, and S. Keating, Jet-topography interactions affect energy pathways to the deep Southern Ocean, *Journal of Physical Oceanography*, 47(1799–1816), doi:10.1175/JPO-D-16-0220.1, 2017. 58
- Batchelor, G. K., On steady laminar flow with closed streamlines at large Reynolds number, *Journal of Fluid Mechanics*, 1, 177–190, doi:10.1017/S0022112056000123, 1956. 572
- Becker, J. M., Effect of a western continental slope on the wind-driven circulation, *Journal of Physical Oceanography*, 29, 512–518, doi:10.1175/1520-0485(1999)029<0512:EOAWCS>2.0.CO;2, 1999. 194
- Becker, J. M., and R. Salmon, Eddy formation on a continental slope, *Journal of Marine Research*, 55, 181–200, 1997. 191, 194
- Beckers, J.-M., H. Burchard, J.-M. Campin, E. Deleersnijder, and P. P. Mathieu, Another reason why simple discretizations of rotated diffusion operators cause problems in ocean models: Comments on isoneutral diffusion in a  $z$ -coordinate ocean model, *Journal of Physical Oceanography*, 28, 1552–1559, doi:10.1175/1520-0485(1998)028<1552:ARWSDO>2.0.CO;2, 1998. 643
- Beckers, J.-M., H. Burchard, E. Deleersnijder, and P. P. Mathieu, Numerical discretization of rotated diffusion operators in ocean models, *Monthly Weather Review*, 128, 2711–2733, doi:10.1175/1520-0493(2000)128<2711:NDORDO>2.0.CO;2, 2000. 644
- Belmadani, A., N. A. Maximenko, J. P. McCreary, R. Furue, O. V. Melnichenko, N. Schneider, and E. D. Lorenzo, Linear wind-forced beta plumes with application to the Hawaiian Lee Countercurrent, *Journal of Physical Oceanography*, 43, 2071–2094, doi:10.1175/JPO-D-12-0194.1, 2013. 117
- Benthuyzen, J., and L. N. Thomas, Friction and diapycnal mixing at a slope: Boundary control of potential vorticity, *Journal of Physical Oceanography*, 42, 1509–1523, doi:10.1175/JPO-D-11-0130.1, 2012. 309
- Black, T. L., The new NMC mesoscale eta model: description and forecast examples, *Weather and Forecasting*, 9, 265–278, doi:10.1175/1520-0434(1994)009<0265:TNNMEM>2.0.CO;2, 1994. 411
- Bleck, R., Finite difference equations in generalized vertical coordinates. Part I: Total energy conservation, *Contributions to Atmospheric Physics*, 51, 360–372, 1978. 404

Boltzmann, L., Further studies on the thermal equilibrium of gas molecules, in *Kinetic Theory, Vol. 2: Irreversible Processes*, edited by S. G. Brush, pp. 88–175, Pergamon Press, Oxford, english translation of Boltzmann (1872), “Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen”, 1966. 541, 762

Böning, C. W., W. R. Holland, F. O. Bryan, G. Danabasoglu, and J. C. McWilliams, An overlooked problem in model simulations of the thermohaline circulation and heat transport in the Atlantic Ocean, *Journal of Physical Oceanography*, 8, 515–523, doi:10.1175/1520-0442(1995)008<0515:AOPIMS>2.0.CO;2, 1995. 640

Bradshaw, P., Possible origin of prandt's mixing-length theory, *Nature*, 249, 135–136, doi:10.1038/249135b0, 1974. 524

Bretherton, C., and C. Schär, Flux of potential vorticity substance: a simple derivation and uniqueness property, *Journal of the Atmospheric Sciences*, 50, 1834–1836, doi:10.1175/1520-0469(1993)050<1834:FOPVSA>2.0.CO;2, 1993. 294, 297

Bretherton, F. P., Critical layer instability in baroclinic flows, *Quarterly Journal of the Royal Meteorological Society*, 92, 325–334, 1966. 393

Bryan, K., A numerical investigation of a nonlinear model of a wind-driven ocean, *Journal of Atmospheric Sciences*, 20, 594–606, doi:10.1175/1520-0469(1963)020<0594:ANIOAN>2.0.CO;2, 1963. 194

Bryan, K., A numerical method for the study of the circulation of the world ocean, *Journal of Computational Physics*, 4, 347–376, doi:10.1016/0021-9991(69)90004-7, 1969. 135, 371

Bühler, O., *Waves and mean flows*, 2nd ed., Cambridge University Press, Cambridge, UK, doi:10.1017/CBO9781107478701, 2014a. 2, 591, 597, 603

Bühler, O., A gentle stroll through EP flux theory, *European Journal of Mechanics B*, 47, 12–15, doi:10.1016/j.euromechflu.2014.01.010, 2014b. 484

Buzzicotti, M., B. A. Storer, H. Khatri, S. M. Griffies, and H. Aluie, Spatio-temporal coarse-graining decomposition of the global ocean geostrophic kinetic energy, *Journal of Advances in Modeling Earth Systems*, 15, doi:10.1029/2023MS003693, 2023. 594

Callies, J., Restratification of abyssal mixing layers by submesoscale baroclinic eddies, *Journal of Physical Oceanography*, 48, 1995–2010, doi:10.1175/JPO-D-18-0082.1, 2018. 404, 431

Callies, J., and R. Ferrari, Dynamics of an abyssal circulation driven by bottom-intensified mixing on slopes, *Journal of Physical Oceanography*, 48, 1257–1282, doi:10.1175/JPO-D-17-0125.1, 2018. 290, 431

Cane, M., V. Kamenkovich, and A. Krupitsky, On the utility and disutility of JEBar, *Journal of Physical Oceanography*, 28, 519–526, doi:10.1175/1520-0485(1998)028<0519:OTUADO>2.0.CO;2, 1998. 373

Carnevale, G. F., R. C. Kloosterziel, and G. J. F. van Heijst, Propagation of barotropic vortices over topography in a rotating tank, *Journal of Fluid Mechanics*, 233, 119–139, doi:10.1017/S0022112091000411, 1991. 142

- Carslaw, H. S., and J. C. Jaeger, *Conduction of Heat in Solids*, Oxford University Press, 1959. 508
- Cauchy, A.-L., *Résumé des leçons données à l'école royale polytechnique sur le calcul infinitésimal*, Imprimerie royale, 1823. 734
- Charney, J. G., R. Fjörtoft, and J. von Neumann, Numerical integration of the barotropic vorticity equation, *Tellus*, 2, 237–254, doi:10.3402/tellusa.v2i4.8607, 1950. 115, 128
- Chatwin, P. C., The vorticity equation as an angular momentum equation, *Mathematical Proceedings of the Cambridge Philosophical Society*, 74, 365–367, doi:10.1017/S0305004100048131, 1973. 111
- Cole, S. T., C. Wortham, E. Kunze, and W. B. Owens, Eddy stirring and horizontal diffusivity from argo float observations: Geographic and depth variability, *Geophysical Research Letters*, 42(10), 3989–3997, doi:10.1002/2015GL063827, 2015GL063827, 2015. 625
- Courant, R., and D. Hilbert, *Methods of Mathematical Physics Volume I*, Wiley-Interscience, New York, 1953. 537, 539
- Courant, R., and D. Hilbert, *Methods of Mathematical Physics Volume II: Partial Differential Equations*, Wiley-Interscience, 1962. 537, 539
- Cox, M., and K. Bryan, A numerical model of the ventilated thermocline, *Journal of Physical Oceanography*, 14, 674–687, 1984. 640
- Cox, M. D., Isopycnal diffusion in a  $z$ -coordinate ocean model, *Ocean Modelling*, 74, 1–5, 1987. 642, 643, 644, 655
- Crank, J., *The Mathematics of Diffusion*, Oxford University Press, 1956. 508
- Csanady, G. T., *Turbulent Diffusion in the Environment*, D. Reidel Publishing Company, 1973. 523, 526
- Cushman-Roisin, B., and J.-M. Beckers, *Introduction to Geophysical Fluid Dynamics*, Academic Press, Amsterdam, 828, 2011. 16, 24
- Danabasoglu, G., R. Ferrari, and J. McWilliams, Sensitivity of an ocean general circulation model to a parameterization of near-surface eddy fluxes, *Journal of Climate*, 21, 1192–1208, doi:10.1175/2007JCLI1508.1, 2008. 625, 645
- Davidson, P., *Turbulence: an introduction for scientists and engineers*, Oxford University Press, Oxford, UK, 2015. 111
- Davies-Jones, R., Comments on "A Generalization of Bernoulli's Theorem", *Journal of the Atmospheric Sciences*, 60, 2039–2041, doi:10.1175/1520-0469(2003)060<2039:COAGOB>2.0.CO;2, 2003. 294
- de Lavergne, C., G. Madec, Julien Le Sommer, A. J. G. Nurser, and A. C. Naveira Garabato, On the consumption of Antarctic Bottom Water in the abyssal ocean, *Journal of Physical Oceanography*, 46, 635–661, doi:10.1175/JPO-D-14-0201.1, 2016. 733
- DeSzeoke, R. A., and A. F. Bennett, Microstructure fluxes across density surfaces, *Journal of Physical Oceanography*, 23, 2254–2264, 1993. 617

Drake, H. F., S. Bailey, R. Dussin, S. M. Griffies, J. Krasting, G. MacGilchrist, G. Stanley, J.-E. Tesdal, and J. D. Zika, Water mass transformation budgets in finite-volume generalized vertical coordinate ocean models, *Journal of Advances in Modeling Earth Systems (JAMES)*, 17, doi:10.1029/2024MS004383, 2025. 707

Drijfhout, S. S., D. P. Marshall, and H. A. Dijkstra, Conceptual models of the wind-driven and thermohaline circulation, 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. 257–282, Academic Press, 2013. 373

Dukowicz, J. K., and R. D. Smith, Stochastic theory of compressible turbulent fluid transport, *Physics of Fluids*, 9, 3523–3529, doi:10.1063/1.869460, 1997. 645

Eckart, C., An analysis of the stirring and mixing processes in incompressible fluids, *Journal of Marine Research*, 7, 265–275, 1948. 545, 546, 607, 623, 625

Einstein, A., Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen [on the motion of small particles suspended in liquids at rest as required by the molecular-kinetic theory of heat], *Annalen der Physik*, 17, 549–560, doi:10.1002/andp.19053220806, 1905a. 511, 514, 520, 521, 744, 747

Einstein, A., Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt [on a heuristic viewpoint concerning the production and transformation of light], *Annalen der Physik*, 17, 132–148, doi:10.1002/andp.19053220607, 1905b. 511

Einstein, A., Zur elektrodynamik bewegter körper [on the electrodynamics of moving bodies], *Annalen der Physik*, 17, 891–921, doi:10.1002/andp.19053221004, 1905c. 511

Eliassen, A., and E. Palm, On the transfer of energy in stationary mountain waves, *Geofysiske Publikasjoner*, 22(1–23), 1960. 484

Ertel, H., Ein neuer hydrodynamicscher Wirbelsatz, *Meteorol. Z. (Braunschweig)*, 59, 271–281, 1942. 263, 267

Ferrari, R., J. C. McWilliams, V. M. Canuto, and M. Dubovikov, Parameterization of eddy fluxes near oceanic boundaries, *Journal of Climate*, 21, 2770–2789, doi:10.1175/2007JCLI1510.1, 2008. 625, 639, 645

Ferrari, R., S. M. Griffies, A. J. G. Nurser, and G. K. Vallis, A boundary-value problem for the parameterized mesoscale eddy transport, *Ocean Modelling*, 32, 143–156, doi:10.1016/j.ocemod.2010.01.004, 2010. 639, 640, 645

Ferreira, D., and J. Marshall, Formulation and implementation of a residual-mean ocean circulation model, *Ocean Modelling*, 13, 86–107, 2006. 635

Feynman, R. P., R. B. Leighton, and M. L. Sands, *The Feynman lectures on physics*, Addison-Wesley Publishing Company, Reading, Mass, 1963. 758

Fofonoff, N. P., Steady flow in a frictionless homogeneous ocean, *Journal of Marine Research*, 13, 254–262, 1954. 193

- Fox-Kemper, B., R. Lumpkin, and F. O. Bryan, Lateral transport 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. 185–209, Academic Press, 2013. [647](#)
- Fraser, N. J., A. D. Fox, S. Cunningham, W. Rath, F. U. Schwarzkopf, and A. Biastoch, Vertical velocity dynamics in the North Atlantic and implications for AMOC, *Journal of Physical Oceanography*, 54, 2011–2024, 2024. [211](#)
- Fraser, N. J., A. D. Fox, and C. W. Hughes, Pycnocnestic coordinates and their application to meridional overturning, *Journal of Physical Oceanography*, submitted, 2026. [408](#)
- Gardiner, C. W., *Handbook of stochastic methods for physics, chemistry and the natural sciences*, 2nd ed., 442 pp., Springer-Verlag, Berlin, Heidelberg, New York, 1985. [521](#)
- Garrett, C., Turbulent dispersion in the ocean, *Progress in Oceanography*, 70, 113–125, 2006. [547](#), [606](#), [607](#)
- Garrett, C., P. MacCready, and P. Rhines, Boundary mixing and arrested ekman layers: Rotating stratified flow near a sloping boundary, *Annual Review of Fluid Mechanics*, 25, 291–323, doi:10.1146/annurev.fl.25.010193.001451, 1993. [404](#), [431](#)
- Gavriel, N., and Y. Kaspi, The number and location of Jupiter's circumpolar cyclones explained by vorticity dynamics, *Nature Geoscience*, 14, 559–563, doi:10.1038/s41561-021-00781-6, 2021. [142](#)
- Gebbie, G., Cancelation of deglacial thermosteric sea level rise by a barosteric effect, *Journal of Physical Oceanography*, 50, 3623–3639, doi:10.1175/JPO-D-20-0173.1, 2020. [681](#), [682](#), [684](#)
- Gent, P., The energetically consistent shallow-water equations, *Journal of the Atmospheric Sciences*, 50, 1323–1325, doi:10.1175/1520-0469(1993)050<1323:TECSWE>2.0.CO;2, 1993. [34](#), [36](#)
- Gent, P. R., and J. C. McWilliams, Isopycnal mixing in ocean circulation models, *Journal of Physical Oceanography*, 20, 150–155, doi:10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2, 1990. [31](#), [627](#), [643](#), [644](#), [758](#)
- Gent, P. R., J. Willebrand, T. J. McDougall, and J. C. McWilliams, Parameterizing eddy-induced tracer transports in ocean circulation models, *Journal of Physical Oceanography*, 25, 463–474, doi:10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2, 1995. [31](#), [558](#), [617](#), [627](#), [631](#), [632](#), [633](#), [634](#), [635](#), [636](#), [638](#), [639](#), [640](#), [654](#), [655](#), [656](#), [758](#)
- Gilbert, A. D., and J. Vanneste, Geometric approaches to Lagrangian averaging, *Annual Review of Fluid Mechanics*, 57, 117–140, doi:10.1146/annurev-fluid-030524-095913, 2025. [591](#), [592](#), [603](#)
- Gill, A., *Atmosphere-Ocean Dynamics*, *International Geophysics Series*, vol. 30, Academic Press, London, 662 + xv pp, 1982. [625](#)
- Gnanadesikan, A., A global model of silicon cycling: Sensitivity to eddy parameterization and dissolution, *Global Biogeochemical Cycles*, 13, 199–220, 1999. [644](#)

Godfrey, J. S., A Sverdrup model of the depth-integrated flow for the World Ocean allowing for island circulations, *Geophysical and Astrophysical Fluid Dynamics*, 45, 89–112, doi: 10.1080/03091928908208894, 1989. 197, 199, 759

Graham, F., and T. McDougall, Quantifying the nonconservative production of Conservative Temperature, potential temperature, and entropy, *Journal of Physical Oceanography*, 43, 838–862, doi:10.1175/JPO-D-11-0188.1, 2013. 626

Gray, A. R., and S. R. Riser, A global analysis of Sverdrup balance using absolute geostrophic velocities from Argo, *Journal of Physical Oceanography*, 44, 1213–1229, doi: 10.1175/JPO-D-12-0206.1, 2014. 366

Greatbatch, R. J., Exploring the relationship between eddy-induced transport velocity, vertical momentum transfer, and the isopycnal flux of potential vorticity, *Journal of Physical Oceanography*, 28, 422–432, doi:10.1175/1520-0485(1998)028<0422:ETRBEI>2.0.CO;2, 1998. 501, 502

Greatbatch, R. J., and K. G. Lamb, On parameterizing vertical mixing of momentum in non-eddy resolving ocean models, *Journal of Physical Oceanography*, 20, 1634–1637, doi: 10.1175/1520-0485(1990)020<1634:OPVMOM>2.0.CO;2, 1990. 501, 635, 636

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. 681, 682, 745

Griffies, S. M., The Gent-McWilliams skew-flux, *Journal of Physical Oceanography*, 28, 831–841, doi:10.1175/1520-0485(1998)028<0831:TGMSF>2.0.CO;2, 1998. 555, 633

Griffies, S. M., *Fundamentals of Ocean Climate Models*, Princeton University Press, Princeton, USA, 518+xxxiv pages, 2004. 34, 135, 462, 537, 539, 555, 592, 617, 639, 640, 647

Griffies, S. M., and R. J. Greatbatch, Physical processes that impact the evolution of global mean sea level in ocean climate models, *Ocean Modelling*, 51, 37–72, doi:10.1016/j.ocemod.2012.04.003, 2012. 7, 592, 628, 631, 668, 685, 686, 764

Griffies, S. M., and R. Hallberg, Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models, *Monthly Weather Review*, 128, 2935–2946, doi:10.1175/1520-0493(2000)128<2935:BFWASL>2.0.CO;2, 2000. 34, 539

Griffies, S. M., A. Gnanadesikan, R. C. Pacanowski, V. Larichev, J. K. Dukowicz, and R. D. Smith, Isoneutral diffusion in a  $z$ -coordinate ocean model, *Journal of Physical Oceanography*, 28, 805–830, doi:10.1175/1520-0485(1998)028<0805:IDIAZC>2.0.CO;2, 1998. 537, 539, 643, 644

Griffies, S. M., C. W. Böning, F. O. Bryan, E. P. Chassignet, R. Gerdes, H. Hasumi, A. Hirst, A.-M. Treguier, and D. Webb, Developments in ocean climate modelling, *Ocean Modelling*, 2, 123–192, doi:10.1016/S1463-5003(00)00014-7, 2000. 463, 638

Griffies, S. M., J. Yin, P. J. Durack, P. Goddard, S. Bates, E. Behrens, M. Bentsen, D. Bi, A. Biastoch, C. W. Böning, A. Bozec, C. Cassou, E. Chassignet, G. Danabasoglu, S. Danilov, C. Domingues, H. Drange, R. Farneti, E. Fernandez, R. J. Greatbatch, D. M. Holland,

- M. Ilicak, J. Lu, S. J. Marsland, A. Mishra, W. G. Large, K. Lorbacher, A. G. Nurser, D. Salas y Mélia, J. B. Palter, B. L. Samuels, J. Schröter, F. U. Schwarzkopf, D. Sidorenko, A.-M. Treguier, Y. Tseng, H. Tsujino, P. Uotila, S. Valcke, A. Voldoire, Q. Wang, M. Winton, and Z. Zhang, An assessment of global and regional sea level for years 1993–2007 in a suite of interannual CORE-II simulations, *Ocean Modelling*, 78, 35–89, doi:10.1016/j.ocemod.2014.03.004, 2014. 19, 683, 764
- Griffies, S. M., A. Adcroft, and R. W. Hallberg, A primer on the vertical lagrangian-remap method in ocean models based on finite volume generalized vertical coordinates, *Journal of Advances in Modeling Earth Systems*, 12, doi:10.1029/2019MS001954, 2020. 16, 27, 239, 404, 452, 462, 469, 501, 645, 650
- Griffiths, D. J., *Introduction to Electrodynamics*, Prentice-Hall, Englewood Cliffs, New Jersey, USA, 479 pp, 1981. 127, 556
- Groeskamp, S., J. Zika, T. McDougall, B. Sloyan, and F. Laliberté, The representation of ocean circulation and variability in thermodynamic coordinates, *Journal of Physical Oceanography*, 44, 1735–1750, doi:10.1175/JPO-D-13-0213.1, 2014. 692
- Groeskamp, S., R. P. Abernathey, and A. Klocker, Water mass transformation by cabbeling and thermobaricity, *Geophysical Research Letters*, doi:10.1002/2016GL070860, 2016. 668
- Groeskamp, S., S. M. Griffies, D. Iudicone, R. Marsh, A. G. Nurser, and J. D. Zika, The water mass transformation framework for ocean physics and biogeochemistry, *Annual Review of Marine Science*, 11, 1–35, doi:10.1146/annurev-marine-010318-095421, 2019. 668, 689, 693, 726, 727, 782
- Gula, J., M. Molemaker, and J. McWilliams, Gulf Stream dynamics along the southeastern U.S. seaboard, *Journal of Physical Oceanography*, 45, 690–715, doi:10.1175/JPO-D-14-0154.1, 2015. 250, 262
- Haine, T. W. N., and A. Fuller, Boundary  $\beta$ -plumes and their vorticity budgets, *Quarterly Journal of the Royal Meteorological Society*, 142, 2758–2767, doi:DOI:10.1002/qj.2866, 2016. 146
- Haine, T. W. N., S. M. Griffies, G. Gebbie, and W. Jiang, A review of Green's function methods for tracer timescales and pathways in ocean models, *Journal of Advances in Modeling Earth Systems (JAMES)*, 17, doi:10.1029/2024MS004637, 2025. 543, 585
- Hallberg, R., A thermobaric instability in Lagrangian vertical coordinate ocean models, *Ocean Modelling*, 8, 227–300, 2005. 474
- Hallberg, R., and P. Rhines, Buoyancy-driven circulation in a ocean basin with isopycnals intersectin the sloping boundary, *Journal of Physical Oceanography*, 26, 913–940, 1996. 247, 252, 309
- Haltiner, G. T., and R. T. Williams, *Numerical Prediction and Dynamic Meteorology*, John Wiley and Sons, New York, USA, 1980. 128
- Haney, R. L., On the pressure gradient force over steep topography in sigma-coordinate ocean models, *Journal of Physical Oceanography*, 21, 610–619, 1991. 463

- Haynes, P. H., and M. E. McIntyre, On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces, *Journal of Atmospheric Sciences*, 44, 828–841, 1987. 290, 291, 292, 294
- Haynes, P. H., and M. E. McIntyre, On the conservation and impermeability theorems for potential vorticity, *Journal of Atmospheric Sciences*, 47, 2021–2031, 1990. 290, 294
- Held, I., The general circulation of the atmosphere, in *WHOI 2000 Summer Study Program in Geophysical Fluid Dynamics*, Woods Hole Oceanographic Institute, 2000. 153, 155
- 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. 393
- Holland, G. J., Tropical cyclone motion: environmental interaction plus a beta effect, *Journal of the Atmospheric Sciences*, 40, 328–342, doi:10.1175/1520-0469(1983)040<0328:TCMEIP>2.0.CO;2, 1983. 142
- Holloway, G., Eddy transport of thickness and momentum in layer and level models, *Journal of Physical Oceanography*, 27, 1153–1157, 1997. 638
- Holloway, G., and P. Rhines, Angular momenta of modeled ocean gyres, *Journal of Geophysical Research*, 27, 843–846, doi:10.1029/90JC02256, 1991. 158
- Holmes, R. M., C. De Lavergne, and T. J. McDougall, Tracer transports within abyssal mixing layers, *Journal of Physical Oceanography*, 49, 2669–2695, doi:10.1175/JPO-D-19-0006.1, 2019a. 404, 431, 434
- Holmes, R. M., J. Zika, and M. H. England, Diathermal heat transport in a global ocean model, *Journal of Physical Oceanography*, 49, 141–161, doi:10.1175/JPO-D-18-0098.1, 2019b. 702, 731
- Holton, J. R., and G. J. Hakim, *An Introduction to Dynamic Meteorology*, 5th ed., Academic Press, Waltham, Mass, USA, 532 pp, 2013. 139, 180, 233, 257, 335, 337, 339, 384, 386
- Hoskins, B., I. Draghici, and H. Davies, A new look at the  $\omega$  equation, *Quarterly Journal of the Royal Meteorological Society*, 104, 31–38, 1978. 387
- Hoskins, B., I. James, and G. White, The shape, propagation, and mean flow interaction of large-scale weather systems, *Journal of the Atmospheric Sciences*, 40, 1595–1612, doi:10.1175/1520-0469(1983)040<1595:TSPAMF>2.0.CO;2, 1983. 121
- Hoskins, B. J., The geostrophic momentum approximation and the semi-geostrophic equations, *Journal of Atmospheric Sciences*, 32, 233–242, 1975. 337, 386
- Hoskins, B. J., Towards a PV- $\theta$  view of the general circulation, *Tellus*, 43AB, 27–35, 1991. 263, 269, 312
- Huang, K., *Statistical Mechanics*, John Wiley and Sons, New York, 493 pp, 1987. 513
- Huang, R. X., and R. W. Schmitt, Goldsbrough-Stommel circulation of the World Oceans, *Journal of Physical Oceanography*, 23, 1277–1284, 1993. 363
- Hughes, C. W., A theoretical reason to expect inviscid western boundary currents in realistic oceans, *Ocean Modelling*, 2, 73–83, doi:10.1016/S1463-5003(00)00011-1, 2000. 191, 194, 221

- Hughes, C. W., and B. de Cueves, Why western boundary currents in realistic oceans are inviscid: A link between form stress and bottom pressure torques, *Journal of Physical Oceanography*, 31, 2871–2885, 2001. 78, 191, 194, 247, 252
- Hughes, C. W., I. Fukumori, S. M. Griffies, J. M. Huthnance, S. Minobe, P. Spence, K. R. Thompson, and A. Wise, Sea level and the role of coastal trapped waves in mediating the influence of the open ocean on the coast, *Surveys in Geophysics*, 40, 1467–1492, doi: 10.1007/s10712-019-09535-x, 2019. 201
- Huw Davies, J., Global map of solid Earth surface heat flow, *Geochemistry, Geophysics, Geosystems*, 14, 4608–4622, doi:10.1002/ggge.20271, 2013. 680
- IOC, SCOR, and IAPSO, *The international thermodynamic equation of seawater-2010: calculation and use of thermodynamic properties*, Intergovernmental Oceanographic Commission, Manuals and Guides No. 56, UNESCO, doi:10.25607/OPB-1338, 196pp, 2010. 626, 660, 661, 669, 773
- Jackson, J. D., *Classical Electrodynamics*, John Wiley and Sons, New York, USA, 848 pp, 1975. 127, 556, 749
- Jackson, L., C. W. Hughes, and R. G. Williams, Topographic control of basin and channel flows: the role of bottom pressure torques and friction, *Journal of Physical Oceanography*, 36, 1786–1805, doi:10.1175/JPO2936.1, 2006. 191, 251
- Jansen, M., A. Adcroft, S. M. Griffies, and I. Grooms, The averaged hydrostatic Boussinesq equations in generalized vertical coordinates, *Journal of Advances in Modeling Earth Systems (JAMES)*, 16, doi:10.1029/2024MS004506, 2024. 37, 483, 592, 617, 620, 635
- Jaynes, E. T., Information theory and statistical mechanics, *Physical Review*, 106(4), 620–630, doi:10.1103/PhysRev.106.620, 1957. 762
- Jeevanjee, N., and D. Romps, Effective buoyancy, inertial pressure, and the mechanical generation of boundary layer mass flux by cold pools, *Journal of Atmospheric Sciences*, 72, 3199–3213, doi:10.1175/JAS-D-14-0349.1, 2015. 135
- Kantha, L. H., and C. A. Clayson, *Small Scale Processes in Geophysical Fluid Flows*, Academic Press, New York, USA, 883 pp, 2000. 625
- Kasahara, A., Various vertical coordinate systems used for numerical weather prediction, *Monthly Weather Review*, 102, 509–522, doi:10.1175/1520-0493(1974)102<0509:VVCSUF>2.0.CO;2, 1974. 404
- Killworth, P. D., On the parameterization of eddy transfer Part I: Theory, *Journal of Marine Research*, 55, 1171–1197, 1997. 638
- Kiss, A. E., Potential vorticity crises, adverse pressure gradients, and western boundary separation, *Journal of Marine Research*, 60, 779–803, 2002. 193
- Kiss, A. E., Potential vorticity dynamics in a domain with closed geostrophic contours. I: steady linear flows, *Journal of Marine Research*, 62, 461–489, 2004. 193
- Klinger, B., and T. Haine, *Ocean Circulation in Three Dimensions*, Cambridge University Press, 466, 2019. 199

Klocker, A., and R. Abernathey, Global patterns of mesoscale eddy properties and diffusivities, *Journal of Physical Oceanography*, 44, 1030–1046, 2014. [625](#)

Klocker, A., and T. J. McDougall, Influence of the nonlinear equation of state on global estimates of dianeutral advection and diffusion, *Journal of Physical Oceanography*, 40, 1690–1709, doi:10.1175/2010JPO4303.1, 2010. [668](#)

Kooloth, P., L. M. Smith, and S. N. Stechmann, Conservation laws for potential vorticity in a salty ocean or cloudy atmosphere, *Geophysical Research Letters*, 49, doi:10.1029/2022GL100009, 2022. [216](#), [284](#), [771](#)

Kraichnan, R. H., and D. Montgomery, Two-dimensional turbulence, *Reports on Progress in Physics*, 43, 547–619, 1980. [128](#)

Kundu, P., I. Cohen, and D. Dowling, *Fluid Mechanics*, Academic Press, 921 + xxiv pp, 2016. [523](#)

Kushner, P. J., and I. M. Held, Potential vorticity thickness fluxes and wave-mean flow interaction, *Journal of Atmospheric Sciences*, 56, 948–958, 1999. [617](#)

Langevin, P., Sur la théorie du mouvement brownien [on the theory of brownian motion], *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 146, 530–533, 1908. [518](#), [521](#)

Lapeyre, G., P. Klein, and B. L. Hua, Does the tracer gradient vector align with the strain eigenvectors in 2d turbulence?, *Physics of Fluids*, 11(12), 3729–3737, doi:10.1063/1.870234, 1999. [552](#)

Large, W., J. McWilliams, and S. Doney, Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization, *Reviews of Geophysics*, 32, 363–403, doi:10.1029/94RG01872, 1994. [625](#), [677](#)

Larson, V. E., The relationship between the transilient matrix and the Green's function for the advection-diffusion equation, *Journal of Atmospheric Sciences*, 56(14), 2447–2453, doi:10.1175/1520-0469(1999)056<2447:trbtm>2.0.co;2, 1999. [574](#)

Le Sommer, J., F. d'Ovidio, and G. Madec, Parameterization of subgrid stirring in eddy resolving ocean models. part 1: Theory and diagnostics, *Ocean Modelling*, 39(1-2), 154–169, doi:10.1016/j.ocemod.2011.03.007, 2011. [647](#)

LeCorre, M., J. Gula, and A.-M. Treguier, Barotropic vorticity balance of the North Atlantic subpolar gyre in an eddy-resolving model, *Ocean Science*, 16, 451–468, doi:10.5194/os-16-451-2020, 2020. [250](#), [251](#)

Lee, M.-M., D. Marshall, and R. Williams, On the eddy transfer of tracers: advective or diffusive?, *Journal of Marine Research*, 55, 483–505, 1997. [485](#)

Lemarié, F., L. Debreu, A. F. Shchepetkin, and J. C. McWilliams, On the stability and accuracy of the harmonic and biharmonic isoneutral mixing operators in ocean models, *Ocean Modelling*, 52-53, 9–35, doi:10.1016/j.ocemod.2012.04.007, 2012. [643](#), [644](#)

- Lemons, D. S., and A. Gythiel, Paul Langevin's 1908 paper "on the theory of Brownian motion", *American Journal of Physics*, 65(11), 1079–1081, doi:10.1119/1.18725, includes English translation of Langevin (1908): *Sur la théorie du mouvement brownien*, C. R. Acad. Sci. (Paris) 146, 530–533, 1997. 518
- Li, X., K. Klingbeil, and H. Burchard, Introducing the coincidence index: A novel diagnostic and production-based framework for water body interaction analysis and its application to study the interaction of river plumes, *Journal of Advances in Modeling Earth Systems (JAMES)*, submitted, 2026. 590
- Lin, S. J., A finite volume integration method for computing pressure gradient force in general vertical coordinates, *Quarterly Journal of the Royal Meteorological Society*, 123, 1749–1762, doi:10.1002/qj.49712354214, 1997. 462, 463, 466
- Loose, N., S. Bachman, I. Grooms, and M. Jansen, Diagnosing scale-dependent energy cycles in a high-resolution isopycnal ocean model, *Journal of Physical Oceanography*, doi:10.1175/JPO-D-22-0083.1, 2022. 70, 472, 483
- Loose, N., G. M. Marques, A. Adcroft, S. Bachman, S. M. Griffies, I. Grooms, R. W. Hallberg, and M. Jansen, Comparing two parameterizations for the restratification effect of mesoscale eddies in an isopycnal ocean model, *Journal of Advances in Modeling Earth Systems (JAMES)*, 15, doi:10.1029/2022MS003518, 2023. 472, 483, 635, 636
- Lorbacher, K., S. J. Marsland, J. A. Church, S. M. Griffies, and D. Stammer, Rapid barotropic sea-level rise from ice-sheet melting scenarios, *Journal of Geophysical Research*, 117, C06003, doi:10.1029/2011JC007733, 2012. 684
- Lowe, J. A., and J. M. Gregory, Understanding projections of sea level rise in a Hadley centre coupled climate model, *Journal of Geophysical Research: Oceans*, 111(C11), doi:10.1029/2005JC003421, 2006. 681
- MacCready, P., and P. Rhines, Slippery bottom boundary layers on a slope, *Journal of Physical Oceanography*, 23, 5–22, doi:10.1175/1520-0485(1993)023(0005:SBBLOA)2.0.CO;2, 1993. 344
- 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. 625, 626
- Maddison, J., and D. Marshall, The Eliassen-Palm flux tensor, *Journal of Fluid Mechanics*, 729, 69–102, doi:10.1017/jfm.2013.259, 2013. 49, 58, 397, 484, 498
- Marion, J. B., and S. T. Thornton, *Classical Dynamics of Particles and Systems*, Harcourt Brace Jovanovich, San Diego, USA, 602 pp, 1988. x, 167
- Markowski, P., and Y. Richardson, *Mesoscale Meteorology in Midlatitudes*, Wiley-Blackwell Publishers, Oxford, UK, 2010. 257
- Marques, G. M., N. Loose, E. Yankovsky, J. M. Steinberg, C.-Y. Chang, N. Bhamidipati, A. Adcroft, B. Fox-Kemper, S. M. Griffies, R. W. Hallberg, M. F. Jansen, H. Khatri, and L. Zanna, NeverWorld2: An idealized model hierarchy to investigate ocean mesoscale

- eddies across resolutions, *Geoscientific Model Development*, 15, 6567–6579, doi:10.5194/gmd-15-6567-2022, 2022. 483
- Marshall, D., Vertical fluxes of potential vorticity and the structure of the thermocline, *Journal of Physical Oceanography*, 30, 3102–3112, 2000. 282, 297
- Marshall, D. P., J. R. Maddison, and P. S. Berlov, A framework for parameterizing eddy potential vorticity fluxes, *Journal of Physical Oceanography*, 42, 539–557, doi:10.1175/JPO-D-11-048.1, 2012. 202
- Marshall, J., and R. A. Plumb, *Atmosphere, Ocean, and Climate Dynamics*, 1st ed., Academic Press, Amsterdam, 319pp, 2008. 82, 83
- Marshall, J., D. Jamous, and J. Nilsson, Entry, flux, and exit of potential vorticity in ocean circulation, *Journal of Physical Oceanography*, 31, 777–789, doi:10.1175/1520-0485(2001)031<0777:EFAEOP>2.0.CO;2, 2001. 282, 294, 296, 297, 308
- Marshall, J. C., E. Shuckburgh, H. Jones, and C. Hill, Estimates and implications of surface eddy diffusivity in the Southern Ocean derived from tracer transport, *Journal of Physical Oceanography*, 36, 1806–1821, 2006. 734
- McComb, W., *The Physics of Fluid Turbulence*, 1st ed., Oxford University Press, Oxford, 602pp, 1990. 522, 523
- McDougall, T. J., Neutral surfaces, *Journal of Physical Oceanography*, 17, 1950–1967, doi:10.1175/1520-0485(1987)017<1950:NS>2.0.CO;2, 1987a. 625, 629
- McDougall, T. J., Thermobaricity, cabbeling, and water-mass conversion, *Journal of Geophysical Research*, 92, 5448–5464, doi:10.1029/JC092iC05p05448, 1987b. 625, 664, 665, 666, 667, 684, 748
- McDougall, T. J., Potential enthalpy: a conservative oceanic variable for evaluating heat content and heat fluxes, *Journal of Physical Oceanography*, 33, 945–963, doi:10.1175/1520-0485(2003)033<0945:PEACOV>2.0.CO;2, 2003. 660
- McDougall, T. J., Thermodynamic concepts used in physical oceanography, *EGU sphere [preprint]*, doi:10.5194/egusphere-2025-4556, 2025. 668
- McDougall, T. J., and P. C. McIntosh, The temporal-residual-mean velocity. Part II: isopycnal interpretation and the tracer and momentum equations, *Journal of Physical Oceanography*, 31, 1222–1246, doi:10.1175/1520-0485(2001)031<1222:TTRMVP>2.0.CO;2, 2001. 565, 592, 609, 611, 617, 620, 632
- McDougall, T. J., and Y. You, Implications of the nonlinear equation of state for upwelling in the ocean interior, *Journal of Geophysical Research*, 95, 13,263—13,276, 1990. 668
- McDougall, T. J., S. Groeskamp, and S. M. Griffies, On geometric aspects of interior ocean mixing, *Journal of Physical Oceanography*, 44, 2164–2175, doi:10.1175/JPO-D-13-0270.1, 2014. 625, 627, 629, 631, 641, 665
- McDougall, T. J., P. M. Barker, R. M. Holmes, R. Pawlowicz, S. M. Griffies, and P. J. Durack, The interpretation of temperature and salinity variables in numerical ocean model output, and the calculation of heat fluxes and heat content, *Geoscientific Model Development*, 14, 6445–6466, doi:10.5194/gmd-14-6445-2021, 2021.

- McIntosh, P. C., and T. J. McDougall, Isopycnal averaging and the residual mean circulation., *Journal of Physical Oceanography*, 26, 1655–1660, 1996. 617, 632
- McWilliams, J., *Fundamentals of Geophysical Fluid Dynamics*, Cambridge University Press, Cambridge, Cambridge, UK, 2006. 115, 126, 157
- McWilliams, J. C., and G. Flierl, On the evolution of isolated, nonlinear vortices, *Journal of Physical Oceanography*, 9, 1155–1182, doi:10.1175/1520-0485(1979)009<1155:OTEON>2.0.CO;2, 1979. 142
- McWilliams, J. C., M. J. Molemaker, and P. Damien, Baroclinic sea-level, *Journal of Advances in Modeling Earth Systems (JAMES)*, *in review*, 2024. 211
- Mellor, G. L., and T. Yamada, Development of a turbulent closure model for geophysical fluid problems, *Reviews of Geophysics*, 20, 851–875, 1982. 625
- Mellor, G. L., L.-Y. Oey, and T. Ezer, Sigma coordinate pressure gradient errors and the seamount problem, *Journal of Atmospheric and Oceanic Technology*, 15, 1122–1131, 1998. 463
- Mermin, N. D., What's wrong with these equations?, *Physics Today*, 42, 9–11, doi:10.1063/1.2811173, 1989. viii
- Mertz, G., and D. G. Wright, Interpretations of the JE BAR term, *Journal of Physical Oceanography*, 22, 301–305, doi:10.1175/1520-0485(1992)022<0301:IOTJT>2.0.CO;2, 1992. 373
- Middleton, J. F., and J. W. Loder, Skew fluxes in polarized wave fields, *Journal of Physical Oceanography*, 19, 68–76, doi:10.1175/1520-0485(1989)019<0068:SFIPWF>2.0.CO;2, 1989. 555, 591, 606, 607, 640
- Minobe, S., M. Terada, B. Qiu, and N. Schneider, Western boundary sea level: A theory, rule of thumb, and application to climate models, *Journal of Physical Oceanography*, 47, 957–977, doi:10.1175/JPO-D-16-0144.1, 2017. 201
- Misner, C., K. Thorne, and J. Wheeler, *Gravitation*, W.H. Freeman and Co., 1279 pages, 1973. 766
- Moffatt, H., Transport effects associated with turbulence with particular attention to the influence of helicity, *Reports on Progress in Physics*, 46, 621–664, doi:10.1088/0034-4885/46/5/002, 1983. 555, 607
- Moffatt, H., Helicity and singular structures in fluid dynamics, *Proceedings of the National Academy of Sciences*, 111, 3663–3670, doi:10.1073/pnas.1400277111, 2014.
- Molemaker, M., J. McWilliams, and W. Dewar, Submesoscale instability and generation of mesoscale anticyclones near a separation of the California Undercurrent, *Journal of Physical Oceanography*, 45, 613–629, doi:10.1175/JPO-D-13-0225.1, 2015. 262
- Montgomery, R. B., Circulation in upper layers of southern North Atlantic, *Papers in Physical Oceanography*, 6, 55, 1938. 640
- Morel, Y., J. Gula, and A. Ponte, Potential vorticity diagnostics based on balances between volume integral and boundary conditions, *Ocean Modelling*, 138, 23–35, doi:10.1016/j.ocemod.2019.04.004, 2019. 297, 298, 302

- Müller, P., Ertel's potential vorticity theorem in physical oceanography, *Reviews of Geophysics*, 33, 67–97, doi:10.1029/94RG03215, 1995. 270, 290
- Müller, P., and C. Garrett, From stirring to mixing in a stratified ocean, *Oceanography*, 15, 12–19, 2002. 547, 625
- Munk, W., and E. Palmén, Note on the dynamics of the Antarctic Circumpolar Current, *Tellus*, 3, 53–55, 1951. 74, 77
- Munk, W. H., On the wind-driven ocean circulation, *Journal of Meteorology*, 7, 79–93, 1950. 193
- Nakamura, N., A new look at eddy diffusivity as a mixing diagnostic, *Journal of the Atmospheric Sciences*, 58(24), 3685–3701, doi:10.1175/1520-0469(2001)058<3685:ANLAED>2.0.CO;2, 2001. 625
- Naveira Garabato, A. C., E. E. Frajka-Williams, C. P. Spingys, S. A. Legg, K. L. Polzin, A. Forryan, E. P. Abrahamsen, C. E. Buckingham, S. M. Griffies, S. D. McPhail, K. W. Nicholls, L. F. Thomas, and M. P. Meredith, Rapid mixing and exchange of deep-ocean waters in an abyssal boundary current, *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1904087116, 2019. 309
- Nurser, A. G., and S. M. Griffies, Relating diffusive surface salinity fluxes to boundary freshwater and salt fluxes, *Journal of Physical Oceanography*, 49, 2365–2376, doi:10.1175/JPO-D-19-0037. 1, 2019. 582, 673, 709
- Nurser, A. G., and M.-M. Lee, Isopycnal averaging at constant height. Part I: The formulation and a case study, *Journal of Physical Oceanography*, 34, 2721–2739, 2004a. 617
- Nurser, A. G., and M.-M. Lee, Isopycnal averaging at constant height. Part II: Relating to the residual streamfunction in Eulerian space, *Journal of Physical Oceanography*, 34, 2740–2755, 2004b. 617
- Nurser, A. G., R. Marsh, and R. Williams, Diagnosing water mass formation from air-sea fluxes and surface mixing, *Journal of Physical Oceanography*, 29, 1468–1487, doi:10.1175/1520-0485(1999)029<1468:DWMFFA>2.0.CO;2, 1999. 709, 718
- Okubo, A., Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences, *Deep-Sea Research and Oceanographic Abstracts*, 17, 445–454, doi:10.1016/0011-7471(70)90059-8, 1970. 137, 552
- Olbers, D. J., J. Willebrand, and C. Eden, *Ocean Dynamics*, 1st ed., Springer, Berlin, Germany, 704 pages, 2012. xii, 79, 626
- Otto, A., F. Otto, O. Boucher, J. Church, G. Hegerl, P. Forster, N. Gillett, J. Gregory, G. Johnson, R. Knutti, N. Lewis, U. Lohmann, J. Marotzke, G. Myhre, D. Shindell, B. Stevens, and M. Allen, Energy budget constraints on climate response, *Nature Geosciences*, 6, 415–416, doi:10.1038/ngeo1836, 2013. 681
- Patmore, R. D., P. R. Holland, D. R. Munday, A. C. Naveira Garabato, D. P. Stevens, and M. P. Meredith, Topographic control of Southern Ocean gyres and the Antarctic Circumpolar Current: a barotropic perspective, *Journal of Physical Oceanography*, 49, 3221–3244, doi:10.1175/JPO-D-19-0083.1, 2019. 191, 251

- Pedlosky, J., L. Pratt, M. Spall, and K. Helfrich, Circulation around islands and ridges, *Journal of Marine Research*, 55, 1199–1251, 1997. 199
- Perrin, J., Mouvement brownien et réalité moléculaire [brownian motion and molecular reality], *Annales de Chimie et de Physique (8ème série)*, 18, 5–114, no DOI assigned; see also English translation reviewed in *Nature* (1910)., 1909. 521
- Peterson, H. G., and J. Callies, Rapid spin up and spin down of flow along slopes, *Journal of Physical Oceanography*, doi:10.1175/JPO-D-21-0173.1, 2022. 344, 405, 412, 469
- Peterson, K. A., and R. J. Greatbatch, Vorticity fluxes in shallow water ocean models, *Atmosphere-Ocean*, 39, 1–14, doi:10.1080/07055900.2001.9649662, 2001. 502
- Phillips, N. A., A coordinate system having some special advantages for numerical forecasting, *Journal of Meteorology*, 14, 184–185, doi:10.1175/1520-0469(1957)014<0184:ACSHSS>2.0.CO;2, 1957. 411
- Phillips, O. M., On flows induced by diffusion in a stably stratified fluid, *Deep-Sea Research: Oceanography Abstracts*, 17, 435–440, doi:10.1016/0011-7471(70)90058-6, 1970. 404, 431
- Plumb, R. A., Eddy fluxes of conserved quantities by small-amplitude waves, *Journal of Atmospheric Sciences*, 36, 1699–1704, 1979. 607
- Polton, J., and D. Marshall, Overturning cells in the Southern Ocean and subtropical gyres, *Ocean Science*, 3, 17–30, 2007. 282, 297
- Polzin, K. L., and T. J. McDougall, Mixing at the ocean's bottom boundary, in *Ocean Mixing : Drivers, Mechanisms and Impacts*, edited by M. Meredith and A. N. Garabato, pp. 145–180, Elsevier, doi:10.1016/B978-0-12-821512-8.00014-1, 2021. 305
- Polzin, K. L., J. M. Toole, J. R. Ledwell, and R. W. Schmitt, Spatial variability of turbulent mixing in the abyssal ocean, *Science*, 276, 93–96, 1997. 625
- Prandtl, L., Bericht über untersuchungen zur ausgebildeten turbulenz, *Zeitschrift für Angewandte Mathematik und Mechanik*, 5, 136–139, doi:10.1002/zamm.19250050212, 1925a.
- Prandtl, L., Bericht über untersuchungen zur ausgebildeten turbulenz [report on investigations of fully developed turbulence], *Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM)*, 5(2), 136–139, doi:10.1002/zamm.19250050212, 1925b. 524
- Prandtl, L., Report on investigation of developed turbulence, *Tech. Rep. Technical Memorandum 1231*, National Advisory Committee for Aeronautics (NACA), english translation of Prandtl (1925), ZAMM 5(2):136–139, 1949. 524
- 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. 45, 256
- Pudig, M. P., W. Zhang, K. S. Smith, and L. Zhang, Parameterizing isopycnal mixing via kinetic energy backscattering in an eddy-permitting ocean model, *Journal of Advances in Modeling Earth Systems, in review*, 2026. 472, 593
- Redi, M. H., Oceanic isopycnal mixing by coordinate rotation, *Journal of Physical Oceanography*, 12, 1154–1158, doi:10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2, 1982. 640, 642, 644

Reif, F., *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York, 1965. 513

Rhines, P. B., Lectures on geophysical fluid dynamics, in *Fluid Dynamics in Astrophysics and Geophysics*, edited by N. R. Lebovitz, pp. 3–58, American Mathematical Society, 1980. 146, 185

Rhines, P. B., Basic dynamics of the large-scale geostrophic circulation, in *WHOI 1982 Summer Study Program*, Woods Hole Oceanographic Institute, 1982. 486, 614

Rhines, P. B., and W. R. Holland, A theoretical discussion of eddy-driven mean flows, *Dynamics Of Atmospheres And Oceans*, 3, 289–325, doi:10.1016/0377-0265(79)90015-0, 1979. 588

Rhines, P. B., and W. R. Young, Homogenization of potential vorticity in planetary gyres, *Journal of Fluid Mechanics*, 122, 347–367, 1982. 569, 572

Roberts, M. J., and D. Marshall, Do we require adiabatic dissipation schemes in eddy-resolving ocean models?, *Journal of Physical Oceanography*, 28, 2050–2063, doi:10.1175/1520-0485(1998)028<2050:DWRADS>2.0.CO;2, 1998. 656

Rooth, C., Hydrology and the ocean circulation, *Progress in Oceanography*, 11, 131–149, doi:10.1016/0079-6611(82)90006-4, 1982. 640

Roquet, F., G. Madec, L. Brodeau, and J. Nycander, Defining a simplified yet realistic equation of state for seawater, *Journal of Physical Oceanography*, 45, 2464–2579, doi:10.1175/JPO-D-15-0080.1, 2015. 683

Rossby, C.-G., Planetary flow patterns in the atmosphere, *Quarterly Journal of the Royal Meteorological Society (suppl)*, 66, 68–87, 1940. 88, 175

Rossby, C.-G., On displacements and intensity changes of atmospheric vorticity, *Journal of Marine Research*, 7, 175–187, 1948. 140

Ruan, X., and R. Ferrari, Diagnosing diapycnal mixing from passive tracers, *Journal of Physical Oceanography*, 51, 757–767, doi:10.1175/JPO-D-20-0194.1, 2021. 588

Sadourny, R., and C. Basdevant, Parameterization of subgrid scale barotropic and baroclinic eddies in quasi-geostrophic models: anticipated potential vorticity method, *Journal of Atmospheric Sciences*, 42, 1353–1363, 1985. 159, 744

Saffman, P. G., *Vortex Dynamics*, Cambridge University Press, Cambridge, England, 1992. 99, 100, 111, 781

Salmon, R., *Lectures on Geophysical Fluid Dynamics*, Oxford University Press, Oxford, England, 378 + xiii pp., 1998. xii, 3, 169

Samelson, R., *The Theory of Large-Scale Ocean Circulation*, Cambridge University Press, Cambridge, UK, 193 pp., 2011. 362

Schär, C., A generalization of Bernoulli's Theorem, *Journal of the Atmospheric Sciences*, 50, 1437–1443, doi:10.1029/2003JC001823, 1993. 282, 296, 297

Schmitt, R. W., Double diffusion in oceanography, *Annual Review of Fluid Mechanics*, 26, 255–285, 1994. 625, 665

- Schneider, T., I. M. Held, and S. Garner, Boundary effects in potential vorticity dynamics, *Journal of Atmospheric Sciences*, 60, 1024–1040, doi:10.1175/1520-0469(2003)60<1024:BEIPVD>2.0.CO;2, 2003. [397](#)
- Shannon, C. E., A mathematical theory of communication, *Bell System Technical Journal*, 27(3), 379–423, doi:10.1002/j.1538-7305.1948.tb01338.x, reprinted with corrections in Bell Syst. Tech. J., 27, 623–656 (1948)., 1948. [762](#)
- Shao, A., A. Adcroft, R. W. Hallberg, and S. M. Griffies, A general-coordinate, nonlocal neutral diffusion operator, *Journal of Advances in Modeling Earth Systems*, 12, doi:10.1029/2019MS001992, 2020. [644](#), [650](#)
- Shchepetkin, A., and J. McWilliams, A method for computing horizontal pressure-gradient force in an ocean model with a non-aligned vertical coordinate, *Journal of Geophysical Research*, 108, 35.1–35.34, 2002. [463](#)
- Smagorinsky, J., General circulation experiments with the primitive equations: I. The basic experiment, *Monthly Weather Review*, 91, 99–164, doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2, 1963. [761](#)
- Smith, K. S., and J. Marshall, Evidence for enhanced eddy mixing at middepth in the southern ocean, *Journal of Physical Oceanography*, 39, 50–69, 2009. [638](#), [645](#)
- Smith, K. S., and G. K. Vallis, The scales and equilibration of midocean eddies: freely evolving flow, *Journal of Physical Oceanography*, 31, 554–570, doi:10.1175/1520-0485(2001)031<0554:TSAEOM>2.0.CO;2, 2001. [128](#)
- Smith, K. S., and G. K. Vallis, The scales and equilibration of midocean eddies: forced-dissipative flow, *Journal of Physical Oceanography*, 32, 1699–1721, 2002. [638](#)
- Smith, R. B., A hurricane beta drift law, *Journal of Atmospheric Sciences*, 50, doi:10.1175/1520-0469(1993)050<3213:AHBDL>2.0.CO;2, 1993. [142](#)
- Smith, R. D., The primitive equations in the stochastic theory of adiabatic stratified turbulence, *Journal of Physical Oceanography*, 29, 1865–1880, doi:10.1175/1520-0485(1999)029<1865:TPEITS>2.0.CO;2, 1999. [620](#)
- Smith, R. D., and P. R. Gent, Anisotropic Gent-McWilliams parameterization for ocean models, *Journal of Physical Oceanography*, 34, 2541–2564, 2004. [647](#), [648](#), [650](#), [651](#), [652](#)
- Smyth, W. D., and J. R. Carpenter, *Instability in Geophysical Flows*, Cambridge University Press, Cambridge, UK, 327 pp, 2019. [541](#), [542](#)
- Solomon, H., On the representation of isentropic mixing in ocean models, *Journal of Physical Oceanography*, 1, 233–234, doi:10.1175/1520-0485(1971)001<0233:OTROIM>2.0.CO;2, 1971. [640](#)
- Spence, P., O. A. Saekno, W. Sijp, and M. England, The role of bottom pressure torques on the interior pathways of North Atlantic Deep Water, *Journal of Physical Oceanography*, 42, 110–125, doi:10.1175/2011JPO4584.1, 2012. [250](#)
- Stakgold, I., *Boundary value problems of mathematical physics, volume I*, SIAM, Philadelphia, 340 pp, 2000.

- Starr, V. P., A quasi-Lagrangian system of hydrodynamical equations, *Journal of Meteorology*, 2, 227–237, doi:10.1175/1520-0469(1945)002<0227:AQLSOH>2.0.CO;2, 1945. 404
- Stern, M., *Ocean circulation physics, International Geophysics Series*, vol. 19, Academic Press, New York, New York, 246 pp, 1975. 204
- Stewart, A., and P. Dellar, The role of the complete Coriolis force in cross-equatorial flow of abyssal ocean currents, *Ocean Modelling*, 38(3-4), 187 – 202, doi:10.1016/j.ocemod.2011.03.001, 2011. 237
- Stewart, A., J. McWilliams, and A. Solococh, On the role of bottom pressure torques in wind-driven gyres, *Journal of Physica Oceanography*, 51, 1441–1464, doi:10.1175/JPO-D-20-0147.1, 2021. 192, 194
- Stommel, H., The westward intensification of wind-driven ocean currents, *Transactions of the American Geophysical Union*, 29, 202–206, 1948. 145, 183, 193
- Straub, D. N., On thermobaric production of potential vorticity in the ocean, *Tellus*, 51A, 314–325, 1999. 273
- Sun, S., R. Bleck, C. Rooth, J. Dukowicz, E. Chassignet, and P. D. Killworth, Inclusion of thermobaricity in isopycnic-coordinate ocean models, *Journal of Physical Oceanography*, 29, 2719–2729, 1999. 474
- Sverdrup, H., Wind-driven currents in a baroclinic ocean: with application to the equatorial currents of the eastern Pacific, *Proc. Natl. Aca. Sci. USA*, 33, 318–326, 1947. 689
- Symon, K., *Mechanics*, Addison-Wesley Publishing Co., Reading, MA, USA, 639 pp, 1971. x
- Talley, L. D., G. L. Pickard, W. J. Emery, and J. H. Swift, *Descriptive Physical Oceanography*, 6th ed., Elsevier, 555pp, 2011. 17
- Taylor, G. I., Diffusion by continuous movements, *Proceedings of the London Mathematical Society*, 20(1), 196–212, doi:10.1112/plms/s2-20.1.196, read June 10, 1920, 1922. 511, 521, 523, 625, 747
- Tennekes, H., and J. Lumley, *A First Course in Turbulence*, 300 pp., MIT Press, Cambridge, USA, 1972. 93
- 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. 280, 309
- 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. 280, 309
- Thorne, K., and R. Blandford, *Modern Classical Physics*, Princeton University Press, Princeton, USA, 1511 + xl pp, 2017. 223
- Thorpe, S., *The Turbulent Ocean*, 439 pp., Cambridge University Press, Cambridge, UK, 2005. 625
- Tomczak, M., and J. S. Godfrey, *Regional Oceanography: An Introduction*, Pergamon Press, Oxford, England, 422 + vii pp, 1994. 19, 199

- 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
- Treguier, A. M., I. M. Held, and V. D. Lariehev, 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. 625, 638, 645
- Trenberth, K. E., An imperative for climate change planning: Tracking earth's global energy, *Current Opinion in Environmental Sustainability*, 1(1), 19–27, doi:10.1016/j.cosust.2009.06.001, 2009. 685
- Tromp, J., *Theoretical and Computational Seismology*, Princeton University Press, Princeton, USA, 2025. xv
- Vallis, G., and B.-L. Hua, Eddy viscosity of the anticipated potential vorticity method, *Journal of the Atmospheric Sciences*, 45, 617–627, doi:10.1175/1520-0469(1988)045<0617:EVOTAP>2.0.CO;2, 1988. 160
- Vallis, G. K., *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*, 1st ed., Cambridge University Press, Cambridge, 745 + xxv pp, 2006. 38, 83
- Vallis, G. K., *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*, 2nd ed., Cambridge University Press, Cambridge, 946 + xxv pp, 2017. 3, 16, 19, 24, 79, 128, 140, 160, 189, 193, 230, 273, 302, 329, 393, 522, 523, 539, 555, 569, 623, 627, 780
- Van Roekel, L., A. Adcroft, G. Danabasoglu, S. M. Griffies, B. Kauffman, W. Large, M. Levy, B. Reichl, T. Ringler, and M. Schmidt, The KPP boundary layer scheme for the ocean: revisiting its formulation and benchmarking one-dimensional simulations relative to LES, *Journal of Advances in Modeling Earth Systems*, 10, 2647–2685, doi:10.1029/2018ms001336, 2018. 677
- Venaille, A., G. K. Vallis, and S. M. Griffies, The catalytic role of beta effect in barotropization processes, *Journal of Fluid Dynamics*, doi:10.1017/jfm.2012.344, 2012. 128
- Veronis, G., The role of models in tracer studies, in *Numerical Models of Ocean Circulation*, pp. 133–146, National Academy of Sciences, 1975. 640
- Waldman, R., and H. Giordani, Ocean barotropic vorticity balances: theory and application to numerical models, *Journal of Advances in Modeling Earth Systems*, 15(4), e2022MS003,276, doi:10.1029/2022MS003276, 2023. 373
- Waln, G., A theoretical framework for the description of estuaries, *Tellus*, 29(2), 128–136, doi:10.1111/j.2153-3490.1977.tb00716.x, 1977. 689, 702, 711, 782
- Waln, G., On the relation between sea-surface heat flow and thermal circulation in the ocean, *Tellus*, 34, 187–195, doi:10.1111/j.2153-3490.1982.tb01806.x, 1982. 689, 702, 711, 782
- Wallace, J., and P. Hobbs, *Atmospheric Science: An Introductory Survey*, Academic Press, 2006. 83
- Ward, M., and A. Hogg, Establishment of momentum balance by form stress in a wind-driven channel, *Ocean Modelling*, 40(133–146), doi:10.1016/j.ocemod.2011.08.004, 2011. 58

- Waterhouse, A. F., J. A. MacKinnon, J. D. Nash, M. H. Alford, E. Kunze, H. L. Simmons, K. L. Polzin, L. C. St. Laurent, O. M. Sun, R. Pinkel, L. D. Talley, C. B. Whalen, T. N. Huussen, G. S. Carter, I. Fer, S. Waterman, A. C. Naveira Garabato, T. B. Sanford, and C. M. Lee, Global Patterns of Diapycnal Mixing from Measurements of the Turbulent Dissipation Rate., *Journal of Physical Oceanography*, 44(7), 1854–1872, doi:10.1175/JPO-D-13-0104.1, 2014. 625
- Waterman, S., and B. J. Hoskins, Eddy shape, orientation, propagation, and mean flow feedback in western boundary current jets, *Journal of Physical Oceanography*, 43, 1666–1690, 2013. 121
- Waterman, S., and J. Lilly, Geometric decomposition of eddy feedbacks in barotropic systems, *Journal of Physical Oceanography*, 45, 1009–1024, 2015. 121
- Webb, D. J., and B. A. de Cueves, On the fast response of the Southern Ocean to changes in the zonal wind, *Ocean Science*, 3, 417–427, doi:10.5194/os-3-417-2007, 2007. 79
- Weiss, J., The dynamics of enstrophy transfer in two-dimensional hydrodynamics, *Physica D: Nonlinear Phenomena*, 48, 273–294, doi:10.1016/0167-2789(91)90088-Q, 1991. 137, 552
- Welander, P., Wind-driven ocean circulation in one and two layer oceans of variable depth, *Tellus*, 20, 1–16, doi:10.1111/j.2153-3490.1968.tb00347.x, 1968. 146, 192, 201
- Wenegrat, J. O., and L. N. Thomas, Centrifugal and symmetric instability during ekman adjustment of the bottom boundary layer, *Journal of Physical Oceanography*, 50, 1793–1812, doi:10.1175/JPO-D-20-0027.1, 2020. 431
- Wenegrat, J. O., L. N. Thomas, J. Gula, and J. C. McWilliams, Effects of the submesoscale on the potential vorticity budget of the ocean mode waters, *Journal of Physical Oceanography*, 48, 2141–2165, doi:10.1175/JPO-D-17-0219.1, 2018. 309
- Whalen, C. B., L. D. Talley, and J. A. MacKinnon, Spatial and temporal variability of global ocean mixing inferred from argo profiles, *Geophysical Research Letters*, 39(18), n/a–n/a, doi:10.1029/2012GL053196, 2012. 625
- Wise, A., C. W. Hughes, and J. Polton, Bathymetric influence on the coastal sea level response to ocean gyres at western boundaries, *Journal of Physical Oceanography*, 481, 2949–2964, doi:10.1175/JPO-D-18-0007.1, 2018. 199, 201
- Wise, A., C. W. Hughes, J. A. Polton, and J. M. Huthnance, Leaky slope waves and sea level: Unusual consequences of the beta effect along western boundaries with bottom topography and dissipation, *Journal of Physical Oceanography*, 50, 217–237, doi:10.1175/JPO-D-19-0084.1, 2020a. 201
- Wise, A., J. A. Polton, C. W. Hughes, and J. M. Huthnance, Idealised modelling of offshore-forced sea level hot spots and boundary waves along the north american east coast, *Ocean Modelling*, 155, doi:10.1016/j.ocemod.2020.101706, 2020b. 201
- Wolfe, C., Approximations to the ocean's residual circulation in arbitrary tracer coordinates, *Ocean Modelling*, 75, 20–35, doi:10.1016/j.ocemod.2013.12.004, 2014. 617
- Wunsch, C., On oceanic boundary mixing, *Deep-Sea Research*, 17, 291–307, doi:10.1016/0011-7471(70)90022-7, 1970. 404

- 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. 393
- Yeager, S., Topographic coupling of the Atlantic overturning and gyre circulations, *Journal of Physical Oceanography*, 45, 1258–1284, doi:10.1175/JPO-D-14-0100.1, 2015. 369
- Young, W. R., Stirring and mixing, in *WHOI 1999 Summer Study Program in Geophysical Fluid Dynamics*, Woods Hole Oceanographic Institute, 1999. 521, 523, 547, 551, 744
- Young, W. R., An exact thickness-weighted average formulation of the Boussinesq equations, *Journal of Physical Oceanography*, 42, 692–707, doi:10.1175/JPO-D-11-0102.1, 2012. 483, 484, 490, 491, 492, 493, 502, 592, 617, 620, 635
- Zeitlin, V., *Geophysical Fluid Dynamics: Understanding (almost) everything with rotating shallow water models*, Oxford University Press, 2018. xix
- Zhang, X., M. Nikurashin, B. Peña-Molino, S. R. Rintoul, and E. Doddridge, Maintenance of the zonal momentum balance in the Antarctic Circumpolar Current by barotropic dynamics, *Journal of Physical Oceanography*, *in review*, 2024. 79
- Zhao, R., and G. K. Vallis, Parameterizing mesoscale eddies with residual and Eulerian schemes, and a comparison with eddy-permitting models, *Ocean Modelling*, 23, 1–12, doi: 10.1016/j.ocemod.2008.02.005, 2008. 635
- Zinnser, W., *Writing to Learn*, Harper Perennial, 1993. xiv

# INDEX

- AABW, 689  
absolute circulation, 231  
absolute vorticity, 122, 165  
    impermeability, 294  
absolute vorticity invariance, 137  
action, xi  
active tracer, 567  
adiabatic flow, 568  
advection, 555  
    maths, 551  
advection equation, 14, 548  
    geometric, 549  
advection-diffusion equation, 572  
advective boundary transport, 674  
advective flux, 670  
advective time scale, 317  
advective tracer flux, 553, 631  
ageostrophic components, 378  
ageostrophic velocity, 326  
ageostrophic vertical velocity, 386  
angular momentum, 106, 108, 167, 173, 221  
    and strain, 106, 109  
    and vorticity, 106, 109  
barotropic model, 158  
    free vortex, 96  
    shalllow water, 80  
anisotropic Gent-McWilliams, 651  
anisotropic neutral diffusion, 647  
anticipated potential vorticity, 159  
aspect ratio, 317, 319  
association versus causation, xi  
asymptotic methods, 325  
atmospheric pressure torque, 189, 242, 248  
atomic bomb, 685  
available potential energy, 633  
QG fluid, 399  
averaging, 488  
    Reynolds, 489  
thickness equation, 491  
thickness weighted, 490  
thickness weighted tracer, 492  
balanced model, 375  
balances, xi  
baroclinic, 221  
baroclinic velocity, 57  
baroclinicity, 122, 211, 221, 244, 267  
    Boussinesq, 239  
    generalized vertical coordinates, 467  
    ideal gas, 214, 257  
    seawater, 271  
    solenoid, 221  
barotropic, 221  
barotropic model, 115, 503  
barotropic velocity, 57, 213  
barycentric velocity, 211, 548, 669  
Batchelor scale, 623  
Bernoulli  
    function, 295  
    potential, 477  
    potential for Boussinesq hydrostatic, 282  
Bernoulli theorem  
    shallow water, 63  
beta drift, 140  
beta effect, 122, 138, 169, 183, 188, 233, 248,  
    357  
    topographic, 177, 329  
    two-dimensional example, 234  
beta gyre, 141  
beta plume, 145, 188

- beta-plane approximation, 121  
Biot-Savart law, 127  
Bjerknes circulation theorem, 233  
body forces, **xix**  
    homogeneous layer, 182  
bolus velocity, 485, 488, 492, 494, 614  
bottom  
    following coordinates, 412  
    geostrophic velocity, 248  
    kinematic boundary condition, 11  
    pressure torque, 189, 224, 242, 244, 248, 358  
    vertical velocity, 359, 360  
    water mass transformation, 706  
bottom drag  
    shallow water, 34  
boundary  
    conditions, 573  
    propagator, 583  
    water mass transformation, 712  
boundary condition  
    natural, 537  
    Neumann, 537  
Boussinesq ocean  
    non-dimensional, 340, 347  
Brownian motion, 514, 625  
Buckingham-II theorem, 314  
budget analysis, **ix**  
buoyancy, 79, 659, 673  
    boundary condition, 673  
    coordinates, 412  
    frequency, 21  
    globally referenced, 18  
    surface ocean budget, 675  
buoyancy work  
    shallow water, 61  
Burger function, 377  
Burger number, 320, 343, 377, 400  
    slope, 344  
cabbeling, 664, 665, 667  
cabbeling parameter, 666, 684  
causal relations, **ix**  
center of mass velocity, 670  
centrifugal acceleration, 131  
channel, 71  
circulation, 94  
    absolute, 231  
    around a streamline, 182  
    free vortex, 97  
    friction effects, 181  
    Kelvin's theorem and work, 212  
    Rayleigh drag, 182  
    rigid-body rotation, 98  
    rotating fluids, 230  
    shallow water, 181  
    tornado, 101  
    wind stress, 182  
circulation induction, 233  
circulation theorem, 181  
coarse graining, 594  
cold core eddy, 44  
column vorticity, 194  
compatibility  
    total mass + tracer mass, 551  
composition property  
    Green's function, 579  
Conservative Temperature, 659  
conservative tracers, 508, 529, 543, 545  
contact forces, **xix**  
contact pressure force, 47, 49  
converging flow  
    PV constraints, 180  
Coulomb gauge, 556  
covariant divergence, 531  
cross-diffusion, 626  
cyclonic eddy, 44  
deformation radius, 344, 400  
    internal, 377  
    shallow water, 320, 332  
delta sheets, 393  
density  
    evolution, 626, 660  
dia-surface flow, 567  
dia-surface transport, 433, 438  
    non-divergent, 442  
    shallow water, 29  
dia-surface velocity, 441  
    compared to vertical, 445  
diabatic process, 568  
diagnostic equation, **xi**  
dianeutral direction, 628  
diapycnal velocity component, 475  
differentiable manifold, 692  
diffusion, 513, 662, 669

- dissipation functional, 537  
Fick's law, 523  
fine scale, 626  
horizontal, 534  
isotropic, 534  
molecular, 513, 626  
momentum, 526  
neutral, 534, 539, 664  
operator, 527, 533, 537  
skew, 554  
temperature, 525  
tensor, 533  
tracer moments, 535, 536  
tracer powers, 540  
tracer variance, 535  
turbulent, 521  
diffusion operator, 641  
diffusion tensor, 627, 670  
diffusion tensor transformation, 648  
diffusively driven flow, 354  
diffusivity, 670  
air, 524  
dynamic, 523  
kinematic, 523  
molecular, 523  
temperature, 525  
turbulent, 523  
dimensional analysis, xii  
dimensionless numbers, 314  
Dirac delta, 392, 393, 538, 706, 720  
Dirichlet boundary condition, 574, 575  
double integral identity, 734  
downscale cascade, 623  
dual form stress, 49, 54  
dual pressure form stress, 497  
dynamical pressure, 28  
eddy tracer fluxes, 603  
eddy-induced  
mass flux, 552  
mass streamfunction, 552  
velocity, 552, 627, 654  
effective beta, 177, 330  
effective free surface height, 7  
Ekman  
generalized Ekman velocity, 307  
velocity, 359  
Eliassen-Palm
- flux, 55, 57, 484, 498, 499  
flux tensor, 48, 498, 499  
emergent scales, xii  
energetics  
shallow water, 58  
shallow water kinetic, 60–62  
shallow water mechanical, 63  
shallow water potential, 59  
ensemble average, 489  
ensemble averaging  
ensemble, 489  
ensemble mean, 594  
epineutral diffusion, 640  
equation of state, 664  
equivalent barotropic depth, 127  
equivalent barotropic flow, 127  
Ertel potential vorticity, 263, 266, 289  
Eulerian  
mean, 595  
Eulerian mean, 593  
Exner function, 214  
exterior moment of  $\lambda$ , 701  
external  
moment, 701  
tracer mass, 729  
velocity, 57  
external scales, xii, 314  
fetch, 76  
Fick's law, 523  
fine scale mixing, 625, 626  
fluid  
dynamics, xviii  
kinematics, xviii  
mass distribution, 697  
flux-form conservation law, 274  
form stress, 47, 49, 52, 53, 62, 464, 635  
dual, 49, 54  
Fourier's law, 525  
Fréchet derivative, 538  
free  
vortex, 95, 111  
free surface patterns, 199  
freshwater  
boundary condition, 670, 671  
velocity, 670  
friction  
shallow water, 34

- torque, 363
- Froude number, 316, 317, 320
- functional derivative, 538
- functional variations, 537
- Galilean
- invariance, 159
- gauge
- freedom, 276, 280, 308, 335, 384
  - function, 176, 276
  - invariance, 555
  - symmetry, 552
- gauge function, 230
- Gauss's divergence theorem, 462
- general vertical coordinates
- basis one-forms, 414
  - basis vectors, 412
  - circulation, 429
  - common confusion, 409
  - contravariant velocity, 419
  - covariant velocity, 419
  - diffusion operator, 426
  - divergence, 426
  - divergence theorem, 426
  - examples, 435
  - Jacobian, 421
  - layer integrated diffusion, 427
  - Levi-Civita tensor, 423
  - material time derivative, 425
  - metric tensor, 420
  - partial derivatives, 423
  - position vector, 417
  - related to Cartesian, 409
  - specific thickness, 421
  - triple product identity, 415
  - vector, 418
  - vector cross product, 430
  - velocity, 419
  - volume element, 423
  - vorticity, 428
- generalized Lagrangian mean, 591, 598
- kinematics, 595
  - tracers, 602
- generalized vertical coordinates, 404
- Gent-McWilliams
- anisotropic, 651
  - available potential energy, 633
  - boundary value problems, 639
- effect, 632
- form stress, 635
- parameterization, 630
- secondary circulation, 632
- thickness diffusion, 636, 638
- geostrophic
- advection, 395
  - balance, xi
  - contours, 178, 569, 588
  - eddies, 43
  - momentum, 205
  - number, 342
  - streamfunction, 326, 329
  - Sverdrup balance, 365
  - transport, 42
- geostrophic streamfunction, 477
- geostrophy
- isopycnal models, 477
  - shallow water, 40
- global conservation law, 274
- global instability, xxi
- global mean sea level, 677
- Godfrey's island rule, 197
- gradient
- generalized vertical, 425
- gradient operator
- Cartesian, 423
  - general vertical coordinate, 423, 424
- gradient wind balance
- barotropic flow, 130
- gravity waves, 92
- Green's function
- advection-diffusion equation, 575
  - composition, 579
  - method, 145
  - non-closed reciprocity, 578
  - passive tracers, 572
  - reciprocity, 576
- gyre circulation, 187
- gyres and channels, 78
- halosteric sea level, 681
- Hamilton's principle, xi
- Hankel function, 146
- harmonic function, 93
- Haynes-McIntyre PV flux, 291
- Heaviside step function, 393
- helicity, 260

- Helmholtz  
first theorem, 101  
second theorem, 102  
third theorem, 102
- Hughes gyre model, 193
- hydrodynamics, vii
- hydrostatic  
approximate balance, 14  
number, 342  
pressure, 5
- ice skater, 167
- ideal gas  
baroclinicity, 214
- impermeability  
compare to material invariance, 294
- impermeability theorem, 176, 289, 291, 304, 398
- absolute vorticity components, 294  
confusions, 294  
kinematics, 292  
planetary geostrophy, 356  
seawater, 294
- impulse response function, 584
- incropping, 27
- induction vector, 397
- inertial  
oscillations, 344
- information entropy, 541
- instability  
global, xxi  
wave, xxi
- integration in  $\lambda$ -space, 696
- interfacial form stress, 49, 52, 494
- interior moment of  $\lambda$ , 701
- interior water mass transformation, 706
- internal  
moment, 701  
tracer mass, 729  
velocity, 57
- internal scales, 314
- inverse barometer sea level, 7
- inverse energy cascade, 128
- invertibility, 375
- irreversible process, vii
- irrotational, 91  
flow, 92, 104
- isopycnal  
coordinates, 412  
ensemble, 607  
layer transport, 607  
mean, 609  
primitive equations, 473
- Jacobian  
on a curve, 588  
operator, 395, 588
- Jacobian operator, 123
- JEBAR, 371
- jets, 140
- Kelvin's circulation theorem, 211, 222, 264, 268
- Kelvin-Helmholtz instability, 343
- kinematic  
water mass transformation, 703
- kinematic boundary condition, 11, 670
- kinetic stress, 45, 493
- kinetic stress tensor, 121
- Kronecker delta, 530
- Lagrangian  
mean, 596
- Lagrangian mean, 593
- Langevin equation, 518
- Laplace's equation, 93
- Laplace-Beltrami operator, 532
- layer mass continuity, 452  
compressible, 452  
pressure coordinates, 454
- Leibniz's rule, 242
- Leibniz-Reynolds transport theorem, 304
- level of no motion, 15
- Levi-Civita tensor, 130
- lid pressure, 27, 115, 128
- linear momentum, 107
- longwave radiation, 674
- MacDonald's function, 146
- macro-turbulence, 43
- Magnus acceleration, 28
- Margules' relation, 25, 41, 44
- Marshall potential vorticity flux, 295
- mass  
 $\lambda$ -layer, 696
- mass budget

- $\lambda$ -layer, 711  
mass continuity  
generalized vertical coordinates, 450  
layer integrated, 452  
mass distribution function, 697  
mass equation  
general vertical coordinates, 460  
mass-Sverdrup, 693  
material  
invariance, 169  
line elements, 224  
material time derivative  
general vertical coordinates, 444  
isopycnal models, 475  
material tracers, 543  
mean free path, 524  
mean tracer equation, 617  
meridional  
overturning circulation, 457  
overturning transport, 655  
transport, 357  
mesoscale eddies, 623, 627  
mesoscale eddy mixing, 625  
metric tensor, 530  
mixed boundary condition, 574  
mixing, 627  
surface boundary, 671  
tensor, 627  
mixing length, 524  
modified mean, 609  
modified mean density, 608  
molecular  
diffusion, 625, 626  
moment of inertia, 110, 112, 167  
moments of  $\lambda$ , 701  
momentum equation, 45  
flux form, 461  
general vertical coordinates, 460  
horizontal, 460  
vertical, 460  
momentum-based, x  
Montgomery potential, 474  
shallow water, 24  
Munk gyre model, 193  
NADW, 689  
natural boundary condition, 537  
natural coordinates  
vorticity, 103  
Neumann boundary condition, 537, 574, 575  
neutral  
anisotropic neutral diffusion, 647  
density, 693  
diffusion, 640  
directions, 640  
slope, 628  
tangent plane, 640  
tangent plane coordinates, 648  
neutral diffusion, 539, 664  
neutrality condition, 644, 649  
Newton's  
third law, 53  
Newtonian friction, 526  
non-advection  
flux, 670  
transport, 674  
non-dimensionalization, xii, 320  
non-divergent  
barotropic model, 115, 117, 503  
non-hydrostatic  
pressure, 8  
notation  
slanted, 529  
upright, 529  
obduction, 449  
ocean  
buoyancy, 659  
free surface, 435  
gyres, 183  
gyres and topographic form stress, 189  
mesoscale eddy, 43  
mixed layer base, 436  
mixing processes, 625  
omega equation, 386  
one-forms, 414  
outcropping, 27  
outer product, 45  
overturning streamfunction, 457  
parameterizations, 623  
particle displacement, 604  
passive  
tracer, 543, 572  
tracer source, 575  
perfect fluid, vii, 548

- periodic channel, 71  
permutation symbol, 130  
phase  
    averaging, 594  
Phillip's layering instability, 541  
planetary  
    induction, 233  
    vorticity, xix  
planetary geostrophy, 251, 322  
    continuously stratified, 351  
    energetics, 354  
    equations, 352  
    potential vorticity, 178, 184, 185, 322, 355  
point vortex, 125  
Poisson equation, 125, 129  
potential energy  
    shallow water layer, 497  
potential enstrophy, 268  
potential flow, 92  
potential vorticity, 428  
    substance, 275  
    baroclinic fluid, 267  
    barotropic fluid, 264  
    Bernoulli potential, 283  
    Boussinesq, 468  
    concentration, 297  
    cylinder, 167  
    derivation, 279  
    dynamical tracer, 269  
    Ertel, 123, 163, 263, 266, 289, 387  
    flux, 176, 275, 280, 283, 289, 291, 295  
    flux-form budget, 274  
    flux-form equation, 175  
    gauge choice, 283  
    generalized vertical coordinates, 466  
    Haynes-McIntyre PV flux, 291  
    hydrostatic Boussinesq, 279  
    impermeability, 289, 356  
    impermeability kinematics, 292  
    induction vector, 397  
    integrated substance, 297  
    inversion, 312  
    isopycnal coordinates, 479  
    iterated, 268  
    Kelvin circulation theorem, 171  
    kinematic derivation, 281  
layer budget, 302  
Marshall PV flux, 295  
material invariance, 289  
name, 171, 269  
non-conservative processes, 175, 273  
non-divergent barotropic flow, 123, 270  
ocean layer, 302  
pancake, 284  
perfect fluid, 264  
planetary geostrophy, 178, 184, 185, 322, 332, 355  
potential density, 271  
preferred flux forms, 308  
quasi-geostrophy, 328, 329, 332, 387  
relative, 330  
Rossby, 123, 163, 169, 171  
seawater, 294  
seawater equation of state, 271  
shallow water, 169  
stuff, 176  
substance, 289, 291  
symmetric instability, 308  
thickness weighted average, 501  
two-dimensional non-divergent, 122  
potential vorticity flux  
    air-sea boundary, 306  
    land-sea boundary, 305  
    steady state constraints, 296  
potential vorticity inversion, 332  
Prandtl number, 526  
Prandtl ratio, 344, 396  
pressure  
    contact force, 47, 49, 464  
    contact stress, 462  
    coordinates, 410  
    gradient force, 462  
    Lagrange multiplier, x  
    lid, 27, 115, 128  
    Poisson equation, 129  
    shallow water dynamic, 24  
    time derivative, 661  
    two-dimensional flow, 131  
pressure form stress  
    interfacial stress, 465  
pressure form stress, 47, 49, 62  
    dual, 497  
    shallow water, 52

- pressure gradient  
  external, 79  
  internal, 79
- pressure gradient work  
  shallow water, 60
- pressure scale  
  geostrophy, 343
- pressure source  
  Coriolis, 134  
  friction, 134  
  irrotational flow, 133  
  rigid-body flow, 131  
  rotation, 130  
  self-advection, 130  
  strain, 130
- pressure torque, 56, 247, 358  
  atmosphere, 363  
  bottom, 363
- process water mass transformation, 703
- prognostic equation, x
- pseudo vector, 90
- pycnocline, 43
- quasi-geostrophy, 323, 375  
  potential vorticity, 328, 329
- quasi-Stokes transport, 612, 616, 617
- radius of curvature, 103
- Rayleigh drag, 182, 201, 245
- re-entrant channel, 71
- real fluid, vii
- reciprocity, 576
- rectification, 485, 591
- reduced gravity, 15, 17, 18, 21, 41  
  model, 15, 38, 43
- reductio ad absurdum, 569
- reference  
  shallow water density, 21
- rescaled geopotential coordinate, 455
- residual mean  
  transport, 628  
  velocity, 628
- Reynolds  
  average, 489  
  stress, 493
- Reynolds transport theorem, 535
- Richardson number, 343
- rigid lid, 115, 128
- approximation, 26, 371
- rigid lid ocean models, 135
- rigid-body  
  rotation, 97, 111, 112
- Robin boundary condition, 574
- Rossby  
  deformation radius, 344  
  effect, 140  
  potential vorticity, 175
- Rossby number, 317, 319, 320, 342
- rotating tank, 80, 133
- rotation matrix, 648
- rotation tensor, 91, 109, 130, 225
- rotational  
  fluxes, 605  
  tracer flux, 553, 631
- salinity, 659
- salt  
  boundary condition, 670, 671  
  concentration, 668  
  velocity, 670
- scalar  
  mechanics, 508
- scale analysis, xii, 314
- scale selectivity, 527
- scales  
  emergent, xii  
  external, xii, 314  
  internal, 314
- sea breeze, 257
- sea level, 677  
  Boussinesq ocean, 685  
  budget, 678  
  steric, 679  
  thermosteric, 679, 680
- seawater  
  equation of state, 271
- secondary circulation, 141
- self-adjoint operator, 537, 577
- self-advection, 130
- semi-geostrophy, 324
- shallow water  
   $N$ -layer equations, 50  
  columns, 15  
  cross-layer flow, 29  
  hydrostatic approximation, 14  
  model formulation, 4

- momentum, 8  
momentum equation, 45  
non-dimensional, 320  
planetary geostrophy, 322  
potential vorticity, 169  
potential vorticity invariance, 171  
pressure, 5  
quasi-geostrophy, 323  
reference density, 17  
stacked model, 19  
subgrid scale, 29  
Sverdrup balance, 184  
thickness, 9  
two-layer, 19  
vertical velocity, 13  
vorticity, 164, 229  
shortwave radiation, 674  
skew  
    diffusion, 554, 627, 662  
    diffusion tensor, 554  
    skew diffusion and skewson, 555  
    symmetric tracer fluxes, 604  
    tracer flux, 553, 631  
slope Burger number, 344  
solenoid, 221  
sound speed  
    ocean, 272  
Southern Ocean, 79  
    channel, 71  
specific  
    thickness, 436, 610, 613  
specific thickness, 421  
splat, 130  
steady state, 78  
stirring, 627  
Stokes  
    drift, 485, 486, 595, 600, 601  
    mean, 595, 600  
    theorem, 94  
Stommel gyre model, 193  
Stommel model, 187  
storage, 711  
strain rate tensor, 91, 109, 130  
    vorticity source, 225  
streamfunction, 117  
    meridional- $\sigma$ , 457  
quasi-geostrophy, 329  
stress  
    shallow water, 34  
stretched vertical coordinate, 396  
stretching, 226  
subduction, 449  
subgrid scale transport, 31  
subgrid scales, 623  
surface  
    gravity waves, 8  
surface quasi-geostrophy, 392  
surface water mass transformation, 706, 717, 719  
    buoyancy, 720  
    circulation, 719  
Sverdrup, 693  
Sverdrup balance, xi, 184, 251, 363, 365  
    geostrophic, 365  
    topographic, 251, 367  
symmetric  
    tracer fluxes, 604  
symmetry  
    condition, 644  
Taylor  
    columns, 15  
Taylor-Bretherton identity, 124, 484, 501  
Taylor-Proudman effect, 185  
tensor  
    (1, 1) representation, 530  
    (2, 0) representation, 530  
    0, 2) representation, 530  
    flat representation, 530  
    musical nomenclature, 530  
    natural representation, 530  
    sharp representation, 530  
tensor product, 45  
terrain following coordinates, 411  
test  
    paddle wheel, 90  
thermal wind balance, 44, 380  
    shallow water, 25, 41  
thermobaricity, 664, 665, 667  
thermobaricity parameter, 666, 684  
thermosteric sea level, 684  
thickness, 613  
    equation, 455  
    isopycnal thickness equation, 475  
    specific, 436

- specific thickness, 475  
weighted average, 490, 617  
weighted velocity, 477  
weighting, 610  
thickness weighted averaging, 483  
tilting, 226  
time  
    mean, 593  
    tendency, x  
topographic beta, 177, 329, 330  
topographic form stress, 189  
    gyres, 189  
topographic nonlinear balance, 251  
topographic Sverdrup balance, 251  
tornado, 101  
torque, 221, 357  
tracer  
    active, 567  
    material, 543  
    mechanics, xx  
    parameterization, 623  
    passive, 543  
    transport tensor, 619  
    variance, 535  
tracer equation  
    general vertical coordinates, 460  
    layer integrated, 456  
    mean, 617  
    shallow water, 14  
tracer fluxes  
    rotational, 605  
    skew symmetric, 604  
    symmetric, 604  
tracer mass  
    external, 729  
    internal, 729  
tracer mass analysis, 689, 723  
    external, 729  
    internal, 729  
    special cases, 727  
tracer moments, 535, 589  
tracers  
    conservative, 508, 529, 543, 545  
trajectory  
    generalized vertical coordinates, 447  
transformation matrix  
    general vertical coordinates, 409  
transformed residual mean, 610  
velocity, 612  
translation motion, 97  
turbulent cascade, 230  
    vortex stretching, 230  
    vorticity, 93  
turbulent transport, 627, 674  
two-dimensional  
    flow, 115, 503  
    non-divergent flow, 234  
    turbulence, 128  
upwind tracer flux, 32  
vanishing layers, 479  
vector  
    invariant velocity equation, 28, 216  
velocity  
    circulation, 94  
    gradient tensor, 109  
    self-advection, 131, 216  
velocity equation  
    vector-invariant, 253, 461, 477  
velocity vector  
    generalized vertical coordinates, 447  
vertical  
    gauge, 556  
    stiffness, 185  
vertical velocity, 445, 492  
    decomposed, 445  
    shallow water, 13  
    two-dimensional non-divergent, 120  
viscosity, 526  
    molecular, 526  
volume  
    between isosurfaces, 265  
volume-Sverdrup, 693  
vortex  
    filament, 99  
    line, 99, 226  
    point, 125  
    stretching, 170, 357  
    tilting, 170  
    tube, 99, 226  
vortex line, 224  
vortex tubes  
    shallow water, 170  
vortical flow, 312

- vorticity, 90, 109, 121  
absolute, 97, 122, 165  
baroclinicity, 211  
barotropic flow, 213  
beta effect, 357  
bottom boundary, 243  
Boussinesq, 237  
curvature, 104, 137  
depth averaged velocity, 369  
depth integrated, 242  
depth integrated velocity, 366  
dynamics, 164, 216  
flux vector, 242  
free vortex, 95, 104  
frozen-in, 226  
Gaussian jet, 105  
general vertical coordinate, 461, 466  
hydrostatic, 237  
isopycnal coordinates, 480, 504  
isopycnal models, 477  
line element rotation, 91  
meridional transport, 362  
natural coordinates, 103  
non-divergent, 92  
of a column, 194  
orbital, 104, 137  
planetary, 97  
quasi-geostrophy, 328  
relative, 97, 461  
rigid-body, 104  
rigid-body rotation, 98  
rotating fluids, 230  
rotating reference frame, 91  
shallow water, 164, 229  
shear, 104, 137  
stretching, 226, 357  
surface boundary, 245  
thickness weighted average, 501  
tilting, 226  
torques, 221  
turbulent cascade, 93  
two-dimensional non-divergent, 122  
vertical component, 241
- vorticity budget  
depth integrated, 357  
flux-form, 218  
normal component, 220  
vorticity equation, 217  
hydrostatic, 238  
planetary geostrophy, 241
- warm core eddy, 44  
water mass, 689  
water mass analysis  
mathematics, 696  
tracers, 723  
water mass configuration space, 691  
water mass formation, 692, 711  
buoyancy, 693  
water mass transformation, 689, 691, 692, 702  
boundary, 712  
buoyancy, 693, 720  
dia-surface flux, 702  
interior, 702  
kinematic, 703  
kinematic method, 706  
process, 703  
process method, 706  
processes, 694, 715  
surface, 717, 719  
wave instability, xxi  
waves  
polarized, 605  
western boundary layer, 139  
western intensification, 183, 187  
wind stress  
homogeneous layer, 182  
shallow water, 34  
work  
on circulation, 212
- zonal  
mean, 594  
re-entrant, 71  
vorticity constraints on zonal flow, 137