

A HANDBOOK OF GEOPHYSICAL FLUID MECHANICS

WITH SELECTED TOPICS IN PHYSICAL OCEANOGRAPHY

Stephen M. Griffies

Princeton University Program in Atmospheric and Oceanic Sciences

smg@princeton.edu

Draft from December 6, 2019

COPYRIGHT ©2019 BY STEPHEN M. GRIFFIES

ALL RIGHTS RESERVED

GRIFFIES, STEPHEN M., 1962-

A HANDBOOK OF GEOPHYSICAL FLUID MECHANICS WITH SELECTED TOPICS IN PHYSICAL
OCEANOGRAPHY / STEPHEN M. GRIFFIES.

THIS BOOK WAS TYPESET USING LATEX.

Contents

PREFACE	vii
1 Princeton University AOS571	1
2 A continuum basis for fluid mechanics	11
Part I. Geophysical fluid mathematics	21
3 Cartesian tensor algebra	25
4 Cartesian tensor calculus	39
5 Partial differential equations	61
6 Geometry of curves and surfaces	81
7 General tensors in brief	89
8 General tensor algebra	97
9 General tensor calculus	107
10 Orthogonal coordinates	115
11 Generalized vertical coordinates	129
12 Tracer coordinates	149
Part II. Geophysical particle mechanics	155
13 Particle kinematics	157
14 Particle dynamics	177
15 Symmetries and conservation laws	183
Part III. Fluid kinematics	201
16 Kinematics of fluid motion	203
17 Mass conservation	229
18 Material tracers and their budgets	249
19 Incompressible flow	265

20 Material fluid objects	277
21 General vertical coordinate kinematics	291
Part IV. Basics of geophysical fluid dynamics	313
22 Momentum dynamics	315
23 Thermodynamics	327
24 Energy dynamics	351
25 Buoyancy	361
26 Stress	377
27 Filtered equations	401
28 Oceanic Boussinesq fluid	415
29 Geostrophy and thermal wind	429
30 Balanced inviscid horizontal flows	451
31 Ekman mechanics	465
Part V. Selected topics in ocean dynamics	477
32 General vertical coordinate dynamics	479
33 Time stepping the ocean equations	487
34 Space-time dependent gravity	507
35 Surface tension	519
36 Surface gravity waves and Stokes drift	527
Part VI. Tracer mechanics	543
37 Advection and diffusion	545
38 Tracer kinematics	569
39 Subgrid scale tracer transport	599
40 Ocean buoyancy	621
41 Water mass phase space	633

Part VII. Shallow water mechanics	649
42 Shallow water models	651
43 Shallow water dynamics	669
44 Gravity waves and geostrophic adjustment	687
45 Isopycnal models	701
Part VIII. Vorticity and potential vorticity	709
46 Vorticity and circulation	711
47 Shallow water vorticity and potential vorticity	721
48 Vorticity mechanics	739
49 Angular momentum, vorticity, and strain	759
50 Potential vorticity mechanics	767
51 Impermeability theorem and PV budgets	781
52 Hydrostatic and Boussinesq ocean potential vorticity	797
Part IX. Balanced models	811
53 Two-dimensional barotropic flows	813
54 Shallow water planetary and quasi-geostrophy	819
55 Continuously stratified planetary geostrophy	835
56 Continuously stratified quasi-geostrophy	853
57 Local stability of fronts	871

Contents

PREFACE

Synopsis of this book

This book treats geophysical fluid mechanics (GFM) as a sub-discipline of physics with a focus on physical processes and emergent phenomena exhibited by fluid motion where rotation and/or gravitation are important. The primary application considered here is to the mechanistic understanding of natural fluid motion on rotating and gravitating planets. We encounter mathematics, kinematics, dynamics, thermodynamics, and tracer mechanics as part of formulating a variety of physical GFM models. A key goal of this book is to develop a suite of physical concepts and mathematical tools for the practitioner of geophysical fluid mechanics. We do so by studying perfect fluids, which only consider reversible mechanical processes, as well as real fluids, which include multiple matter constituents and thermodynamically irreversible processes such as mixing. We also consider specialized topics from physical oceanography that exemplify how geophysical fluid mechanics can be used to describe a suite of ocean processes.

This book appreciates the mathematical structure of the governing fluid mechanical equations and the associated conservation laws. Doing so supports physical understanding and development of methods for analysis and simulation. The presentation aspires to be deductive with a degree of rigor and generality realized through the use of physically informed mathematics. We provide detailed derivations of all formula both to support a deeper understanding of geophysical fluid mechanics and to provide guidance to the practitioner interested in mastering the details. The material is targeted to students with a sound undergraduate training in physics, engineering, or applied mathematics, with an assumed exposure to Newtonian mechanics, equilibrium thermodynamics, fluid mechanics (or another continuum field theory such as classical electromagnetism), vector calculus, tensor analysis, and partial differential equations. Most of these pre-requisites are reviewed within this book. Hence, for readers finding the entry fee rather expensive, patience and persistence with this book will be rewarded.

Two pillars of geophysical fluid mechanics

We conceive of two general pillars to theoretical geophysical fluid mechanics. The first comprises the physical and mathematical formulation of conceptual models used to garner insight into rotating and stratified fluids. Here one is concerned with the physical basis for rationalizing fluid flows, aiming to develop transparent and practical viewpoints and analysis methods through the use of physical principles and mathematical tools. We refer to this aspect as the *conceptual model*

formulation pillar and it forms the main focus to this book. As part of our development of this pillar, we emphasize the physical foundations and methods of use to support an understanding of the mechanics of phenomena emerging from the conceptual models. The additional study of *emergent geophysical fluid phenomena*, such as waves, instabilities, turbulence and large-scale circulation, forms the second pillar of GFM. We offer here a modest discussion of waves and instabilities, yet minimally treat other emergent phenomena such as turbulence and general circulation. These subjects are each worthy of separate treatments that go beyond the scope of this book.

Organization and style

This book is organized into parts according to their particular focus, with each chapter starting with a brief overview and guide pointing to dependencies of the material in that chapter. Some chapters focus on topics required for a basic understanding of the subject and offer exercises to test that understanding. This material is suitable for a graduate-level course. Other chapters offer monograph-style special topics that further the foundations and exemplify applications with a particular bias towards physical oceanography.

One aim of writing this book is to offer a handbook or reference guide for those interested in an accessible and physically meaningful presentation of GFM foundations and methods. Here the reader will find many of the details necessary to become a practitioner of the discipline. Genuinely addressing this aim meant writing many pages. However, there is no pretense that any reader will penetrate all topics nor read the book cover-to-cover. This recognition is particularly keen in a world where research agendas commonly spread rather than focus attention. With that point in mind, an attempt has been made to facilitate picking up the book at a variety of starting points. One strategy is to offer relatively succinct chapters with mostly bite-sized topics. This approach affords the reader many intellectual breathers of use for self-study and preparing class lectures. Furthermore, each chapter and/or part is written in a reasonably self-contained manner even at the cost of some redundancy. Yet when redundancy becomes onerous, generous use of cross-referencing identifies allied material treated elsewhere in the book.

No book is an island. This book certainly stands on the shoulders of other books, review articles, research articles, and online tutorials. Furthermore, many readers find value in studying a subject from a variety of perspectives, thus justifying the proliferation of books that may cover much of the same material. For example, in other areas of physics there are countless texts on quantum mechanics, classical mechanics, general relativity, etc., each offering useful insights that provide the student with a suite of resources to dive into the subject. The study of geophysical fluid mechanics is less mature in this regard. Nonetheless, there is a growing suite of pedagogical and authoritative resources available that greatly enhances the opportunities to learn the subject. In an attempt to support a broad study of the field, throughout this book we provide pointers to written and/or video presentations to support further studies. Many more resources are available through a quick internet search.

Emphasizing the synergy between physics and maths

It is not uncommon for treatments of rotating and stratified fluids, and applications to circulating atmospheres and oceans, to eschew certain of the foundational physical concepts and mathematical methods. Conversely, those treatments that venture into the foundations often assume a rather specialized knowledge outside the purview of many practitioners and can be somewhat distant from phenomenology. This perception motivated the style for this book, which presents geophysical fluid mechanics from a mathematical physics perspective at a level accessible to the entering graduate student. Doing so is partly an intellectual exercise in weaving together a variety

of physical and mathematical topics. It is also a practical exercise that provides the reader with a theoretical platform for research and teaching, thus justifying the term “handbook” in the title. There is a synergy between the physical content of geophysical fluid mechanics and its mathematical expression, whereby physics informs the maths and maths reveals the physics. This synergy pervades theoretical physics and this book is a humble attempt to exemplify that synergy.

What inspired the writing of this book?

This book grew from class notes developed for a first semester graduate course in geophysical fluid mechanics at Princeton University. The course offers an overview of fluid mechanics and the formulation of a variety of geophysical fluid models. Although there are tremendous existing resources in the literature, such as [Vallis \(2017\)](#) and its more succinct offspring [Vallis \(2019\)](#), each teacher has a distinct style and focus that generally leads to the development of unique class notes. Even so, not all class notes evolve into a book. The following provides a summary of what motivated this book.

[Gill \(1982\)](#) and [Pedlosky \(1987\)](#) pioneered the systematic presentation of dynamical ideas falling under the moniker of *geophysical fluid dynamics* (GFD). Their treatments of GFD offer a powerful and elegant application of Newton’s laws of motion to rotating and stratified fluids, with an emphasis on large-scale motion and the use of applied mathematical methods. The study of rotating and stratified fluids has matured since the 1980s through deepening its foundations, refining its mathematical and numerical methods, extending its applications, and increasing its observational measurements. What has emerged is a recognition that a fruitful study of rotating and stratified fluids makes use of ideas that go beyond the traditional notions of “dynamics”. A practitioner thus develops insights by weaving together concepts and tools from a variety mathematical, kinematical, thermodynamical, tracer mechanical, and dynamical perspectives. Acknowledging this broadening of the practice motivates the term “mechanics” for this book’s title, rather than the more focused “dynamics”. It is a trivial change in verbiage but a nontrivial change in perspective.

Further acknowledging the expansion of the purview for this book from GFD to GFM suggests the following thesis that underlies the present treatment. Namely, geophysical fluid mechanics is a discipline within physics that provides a mechanistic framework for studying the variety of planetary fluid phenomena. Correspondingly, this book aims, at least in part, to link geophysical fluid mechanics to the broader subject matter of fluid and continuum mechanics. This aim is distinct from books primarily focused on describing observed phenomena. Though here guided by observed phenomena, an equal weight is given to appreciating geophysical fluid mechanics as a subject within physics, and as such we develop the theory and methods to a degree unnecessary when primarily focusing on phenomenological goals. This perspective makes this book somewhat distinct from others that might sit nearby on a bookcase.

Geophysical Fluid Mechanics and Climate Science 2.0

Fluid mechanics has a history of applications that span science and engineering, from blood flow to the stability of galaxies. A key 21st century application of geophysical fluid mechanics concerns the questions of climate science associated with the greenhouse gas experiment pursued by industrialized civilization’s carbon centered energy use. Leading order questions about climate warming have been sufficiently addressed to recognize that the planet has reached a crisis point threatening the viability of the biosphere. Even so, mechanistic answers to a number of questions remain at the cutting edge of climate science 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? Are there feasible and sustainable geo-engineering options that equitably reduce the negative impacts of climate warming without introducing new problems? These questions, and many others, constitute the scientific challenges of *Climate Science 2.0*.

Answers to Climate Science 2.0 questions generally require basic research using concepts and tools from geophysical fluid mechanics. Furthermore, geophysical fluid mechanics fundamentals become even more essential as numerical climate models admit enhanced details of the complex multi-scaled fluid flow, and as observational field measurements become more refined and process-oriented. One area where a strong foundation in GFM is critical concerns the development of robust numerical models along with mechanistic interpretations of their output, which in turn further support confidence in model projections for future climate. Another area concerns the design of field measurement campaigns, where advanced technologies facilitate investigating geophysical fluid mechanical questions within the natural environment as a complement to the laboratory. These examples highlight the relevance of geophysical fluid mechanics as part of the scientific investigation of the earth's climate, thus serving an essential role in society's quest for sustainability and equity.

Caveats and limitations

Although growing to a nontrivial size, this book remains a work in progress that is not ready for publication. For example, the following topics are planned for inclusion: Hamilton's Principle, Rossby waves, barotropic and baroclinic instability theory. More effort is also needed to unify notation, build an index, extend the exercises, enhance figures, improve discussions, and correct errors. Feedback is solicited particularly where the reader identifies poor writing or incorrect presentations.

1

Princeton University AOS571

1.1	Welcome	1
1.2	Class structure and expectations	2
1.2.1	Class notes	2
1.2.2	Grade = class participation (10%) + homeworks (65%) + final exam (25%)	2
1.2.3	During the class	3
1.3	Written and spoken communication	3
1.3.1	Clear thinking leads to clear communication	4
1.3.2	Empathy is key	4
1.3.3	Clarity helps, but some material is just tough	4
1.4	Pointers on problem solving	4
1.4.1	Dimensional analysis	4
1.4.2	Tensorial consistency	5
1.4.3	Words and pictures	5
1.4.4	Mathematical sophistication	5
1.4.5	More than one path to a solution	6
1.4.6	Balance between thorough and brief	6
1.4.7	Questions for clarification	6
1.4.8	Stay positive	6
1.5	Specifics for marking assignments	6
1.6	Course syllabus	7

1.1 Welcome

Welcome to Princeton University's AOS571 for the autumn 2019 semester. In this class we explore elements of geophysical fluid mechanics, which is a physics discipline concerned with the motion of rotating and stratified fluids such as those in the earth's atmosphere and ocean. This subject is rich in physics, maths, and phenomenology. Our explorations will reveal basic features about the observed patterns of planetary fluid motion, and provide a mathematical physics foundation for later courses and research. Here are details for your teacher:

Dr. Stephen M. Griffies
SMG@princeton.edu
<https://stephengriffies.github.io/>

Office hours by appointment

Class Monday and Wednesday 3:00pm-4:30pm + make-up lectures on selected Fridays

Room 154 Guyot

1.2 Class structure and expectations

My goal for this course is help you learn how to formulate and to solve problems in geophysical fluid mechanics, and to convince yourself and others that you understand the material. Garnering the needed skills requires effort and practice. You will have plenty of opportunities to develop the necessary brain muscle assuming you maintain the discipline to keep up with the material. You should also ask questions, preferably in class, when you are unsure and to sincerely attempt to solve the exercises, either alone or collaboratively with other students in the course. Here are some particulars in regards to class notes, grading, and class etiquette.

1.2.1 Class notes

The class lectures closely follow selected material from [the Griffies \(2019\) course notes](#) available online. These notes started in 2014 as annotations to the textbooks from [Vallis \(2017\)](#) and [Vallis \(2019\)](#). Both of these books remain highly recommended references for the course. However, the class notes have grown over recent years into an autonomous entity from which the course material directly derives.

Material for each class comes from the class notes as well as online videos. You are expected to read through the class notes and to view the videos **prior to the class**. During class we will discuss salient points from the assigned notes. To allow sufficient time for interactive discussion and questions, not all of the assigned reading material will be directly covered in class. In this way, the class is partially "flipped". To make this process work requires students to come prepared for class by doing the reading (and video viewing) prior to each class.

1.2.2 Grade = class participation (10%) + homeworks (65%) + final exam (25%)

There are three areas where grades are earned: class participation, homework exercises, and take-home final exam.

Class participation = 10%

There are three areas where students are expected to contribute during class.

- DISCUSSIONS AND QUESTIONS
- LECTURE REPORTS: To help motivate keeping up with the material, at the start of each class a randomly chosen student will briefly report on material from the prior class and provide an advertisement for the material to be covered in the current class. You can use notes and the chalkboard to help with this report. These lecture reports are roughly 5-10 minutes in duration. They offer a means to learn how to talk in front of people and for me to gauge how well you are digesting the course material.
- HOMEWORK REPORTS: We will discuss solutions to homework exercises at the start of the day they are due, particularly those exercises that offered the most difficulty. To start the discussion, a randomly chosen student will present a synopsis of their solution on the chalkboard.

Class participation credit will be graded pass/fail with any sincere attempt to contribute deemed a pass with full 10% credit. This credit will be given even if you say something incorrect or present an incorrect homework solution. Note that excessive absentees will compromise opportunities to garner a pass.

Worked homework exercises = 65%

Worked homework solutions are due at the start of class one week after the exercises are assigned (unless otherwise noted). Students can make use of any resources for solving homework exercise, including other people. Clarification of questions can also be obtained via email to me. If you find a solution from a source other than your own head, then be sure that you fully understand both the essence and the detail of the solution. Although you are encouraged to discuss the problems with other students, you are cheating yourself if you merely parrot another person's answer without fully grokking it yourself.

There is no pretense that the exercises offered in class are clearly formulated. Indeed, a certain degree of ambiguity reflects the *status quo* in research, where formulating a novel and insightful question is generally the most difficult part of the research process. Additionally, the solutions may not be 100% correct or ideal from a pedagogical perspective. Rather, they represent a work in progress. If mistakes or ambiguities are found, then please share your questions and concerns.

Take-home final exam = 25%

There is a take-home final exam during exam period in January. You are asked to do the work as a solo student, with no help or consultation from another human. However, you can make use of books, notes, online resources, etc. The questions are generally taken from published papers with references provided to the student.

1.2.3 During the class

The class time is comprised of iPad and chalkboard lectures based on the class notes along with discussion. The pace will be gauged on questions during the lectures and my sense for how well the class is grokking the material. You are encouraged to follow lectures by having a copy of the class notes on-hand, either electronically or printed.

To support your learning and teaching experience, and those of your classmates, please ensure that you turn all electronic devices into "airplane mode" so that you are not tempted to divert attention to non-class issues.

1.3 Written and spoken communication

Reading class notes, research articles, and books exposes one to sound and clear writing, as well as to unsound and obscure writing. Attending lectures and seminars in person or online also exposes one to a range of speakers. Some speakers offer the benefits of a sleeping aide without the cost and side-effects of a pill. Others rush through more material than even they can digest, whereas others are happy to pedagogically engage in a lecture even without any prior knowledge of the material. To succeed in research and teaching, you will need to master elements of both written and spoken communication. It is therefore critical to nurture these skills throughout your career whether you seek employment in academics or elsewhere.

1.3.1 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 compose and organize thoughts and, if time allows, to edit and edit yet again. Others are better at speaking, where spontaneous and interactive reflections and experience can bolster the clarity of a presentation. Both modes of communication are important in science and engineering.

As inspiration, for both clear and obscure, pick up one of your textbooks or class lecture notes and analyze the presentation for clarity. Where are you confused? Where is the material crystal clear? Then pick up a journal article and perform the same analysis. What do you like? What do you dislike? Then go to YouTube and find a science or engineering lecture, old or new. What makes the speaker engaging and clear, or boring and obscure?

1.3.2 Empathy is key

A basic tenet of effective communication is empathy. Place your mind inside that of an interested and smart reader or listener. Identify with their quest to understand what you wish to communicate. What assumptions are you making? Are the assumptions justified based on the audience? How compelling is your scientific story? As a start along the path towards clear written communication, I have accumulated some pointers in an online document: [Elements of Style for Journal Papers](#). You will not need this document for homeworks or exams in AOS571. However, it will come in handy when you begin the process of writing scientific documents and preparing scientific presentations.

1.3.3 Clarity helps, but some material is just tough

Although poor communication certainly does hinder our ability to digest new ideas and material, it is conversely important to appreciate that some material is, quite honestly, 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 one or two generations before some material can be sufficiently digested to allow for the core nugget to be uncovered. So do ask for clear communication, but do not expect clarity to remove the struggles we all experience when learning.

1.4 Pointers on problem solving

Most people are not born with *a priori* physics problem solving skills. Rather, it takes extensive practice to develop the necessary brain muscle. Here are some general pointers to keep in mind when diving into a physics problem, whether it is one given for a class or one forming part of a broader research question.

1.4.1 Dimensional analysis

The symbols we use in mathematical physics correspond to geometrical objects (e.g., points, vectors, tensors) describing a physical concept (e.g., location, momentum, stress). Hence, the symbols generally carry physical dimensions. The three physical dimensions we are concerned with in this course are length (L), time (T), and mass (M). Physical dimensions of the equations must be self-consistent. For example, if one writes an equation

$$A = B, \tag{1.1}$$

where A and B have different physical dimensions, then the equation makes no sense physically. Something is wrong. Although not always sufficient to uncover errors, dimensional analysis is an important means to “debug” the maths.

1.4.2 Tensorial consistency

In the same way that mathematical equations in physics need to maintain dimensional consistency, they must also respect basic tensor rules. For example, the equation

$$A = B, \tag{1.2}$$

makes mathematical sense if A and B are both scalars. Likewise,

$$\mathbf{A} = \mathbf{B} \tag{1.3}$$

makes sense if both \mathbf{A} and \mathbf{B} are vectors. However, if both \mathbf{A} and \mathbf{B} are vectors, then the equation

$$\mathbf{A} = \nabla \cdot \mathbf{B} \tag{1.4}$$

does not make sense because the left hand side is a vector and the right hand side is a scalar. Maintaining basic tensorial rules can be considered the next level of sophistication above dimensional analysis.

1.4.3 Words and pictures

It is important to explain the problem and your solution using words and pictures. Hence, it is good practice to liberally include words/sentences in between the key equations, with the purpose to explain what the maths means using English. Here are some practical payoffs for this style of presentation.

- In the process of trying to explain the maths using words and pictures, you generally must dive deeper into the logic of the problem. In doing so, you often identify weak points and errors in the solution. This process is a very important learning stage in preparing to stand in front of people 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.

1.4.4 Mathematical sophistication

Students enter a graduate level course in geophysical fluid mechanics with differing levels of maths training. For the more math-centric problems, it may prove useful to exercise your maths brain muscle. Do so if you feel it useful for the solution presentation. However, solutions will generally not be marked down if you fail to present the full depth of the maths available for a problem. Having said that, as this course progresses you are expected to evolve in your math skills, just as your understanding of the physics matures.

1.4.5 More than one path to a solution

In physics, there is often more than one path to a solution. Pursuing distinct paths offers added physical and mathematical insight, exposes assumptions, and allows one to double-check a solution. Some of the most profound findings 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 Newton's formulation, we may well have had a very different intellectual evolution of 19th and 20th century physics.

1.4.6 Balance between thorough and brief

There is often a conflict between showing full mastery of a problem and keeping the solution write-up brief. In general, there is no need to re-derive equations already presented in the class lecture notes or in [Vallis \(2017\)](#). Proper referencing of the equation is all that you need; i.e., tell me something like “starting from equation (X.YY) from the notes.” Additionally, when presenting a derivation, you may choose to show just the key steps rather than all intermediate steps. Determining what is a “key” step is largely up to you, but it should be something you learn to do in time. Nonetheless, as per the previous pointer, you are encouraged to show more than one approach to a solution.

1.4.7 Questions for clarification

Questions for points of clarification will be entertained if you feel the problem is ill-posed or if you are totally lost. Email is the most efficient means to communicate to me. Responses will generally be sent through BlackBoard so that all students can see the response, thus keeping everyone with the same information. Correspondingly, questions within 24 hours of the deadline are generally not entertained so to ensure that all students have time to see the response.

1.4.8 Stay positive

Everyone makes mistakes, some more than others. The toughest part about making mistakes is the self-imposed shame or embarrassment. Please try to keep a positive mind about your mistakes. As you will learn, making mistakes offers significant opportunities for learning! I am a poster-child for this process!

So do not fret if you find many marks on your homeworks and exams. But do be sure to use mistakes as learning opportunities. That is how life in academics (life in general!) works. Furthermore, be completely honest with yourself to candidly identify weaknesses. I will do my best to work personally with you if something remains uncertain or you feel there is a weakness in your skills that needs some extra help. Please seek help should you wish it. And finally, please do question my marks should you feel they are unfair or incorrect.

1.5 Specifics for marking assignments

Assignments are generally marked using the following rules, with grading less forgiving as the class progresses.

- SIGN ERRORS: Sign errors are a nuisance. We all must spend time to uncover them. One means of detecting errors is to try explaining the maths to yourself or someone else. Does the result make sense? If not, then perhaps there is a sign error. I am generally not too upset

with sign errors if they have minimal physical relevance. But when they indicate a physical misconception, then I will mark it more harshly.

- **MATH ERRORS:** Math errors, such as those associated with basic calculus mistakes, are generally marked down by two points.
- **DIMENSIONAL AND TENSORIAL ERRORS:** I am relatively unforgiving of dimensional mistakes and tensorial inconsistencies, generally removing three points when they occur.
- **PHYSICALLY MISSING THE POINT:** Evidence of physically missing the point will generally invoke the most negative marks, depending on the depth of the misconception. The best way to convince me you grasp the basic physics is to use words and pictures. If the maths is missing or totally wrong, but you present some sensible words and pictures, then that will help earn nonzero credit.
- **PRESENTATION OF THE SOLUTION:** Please write clearly and legibly. There is generally no need to submit solutions in L^AT_EX. But if your handwriting is horrible, then consider learning L^AT_EX to typeset the solutions. If the equations and words are sloppy, you will find teachers less forgiving of errors. You must convince the teacher that you understand the solution and present the maths in a legible manner. Good communication skills are key to being a good scientist or engineer.
- **DEADLINES:** Please do your best to be on time with handing in homeworks and exams, with disasters, personal tragedy, and accidents the only excuses for late assignments. Fairness is the fundamental reason to insist on this rule, particularly since we will generally work through the solutions on the day the homeworks are handed in.
 - You may hand in assignments via email, but within the same deadline as for class hand-in.
 - You may hand-in homework late but only with prior arrangement. Otherwise, the homework set will receive zero credit.

1.6 Course syllabus

The following is a syllabus for 24 lectures of roughly 80-minutes each. The course material consists of the class notes plus a selection of online videos that are required viewing. The notes must be read and the videos watched *prior to the class*. Salient point related to the material will be discussed during class, along further discussion and questions. The syllabus is subject to slight modification as the semester progresses.

- **LECTURE 1:** Intro to the course and mathematics refresher
 - [9-minute video from Dr. N. Hall](#)
 - Course introduction
 - Survey of maths from Part I
 - Chapter 3: cartesian tensors
 - Chapter 4: vector calculus
- **LECTURE 2:** Geophysical particle mechanics

- Chapter 13: kinematics of a particle moving around a rotating sphere, including position, velocity, acceleration, Cartesian and spherical coordinates, rotating reference frame, Coriolis acceleration, centrifugal acceleration
 - Chapter 14: Newton's equation of motion for particle moving around rotating sphere, rotating reference frames, gravitational geopotential
- LECTURE 3: Symmetries, conservation laws, and constrained motion
 - Chapter 15: mechanical energy, potential momentum, inertial oscillations, axial angular momentum
- LECTURE 4: Continuum hypothesis and fluid kinematics
 - [4-minute video on Eulerian and Lagrangian descriptions](#) from Prof. A. Hogg
 - [27-minute video on Eulerian and Lagrangian descriptions](#) from Prof. J. Lumley.
 - Chapter 2: the continuum hypothesis used for describing fluids as a continuous media
 - Chapter 16: Fluid kinematics, Eulerian and Lagrangian descriptions, Galilean invariance, material time derivative, flow lines
- LECTURE 5: Mass and tracer conservation
 - [5-minute video on mass conservation](#) from Prof. A. Hogg.
 - Chapter 17: continuity equation, mass budget for fluid elements and finite regions, kinematic boundary conditions
 - Chapter 18: barycentric velocity, tracer equation, budgets for infinitesimal fluid elements, budgets for finite fluid regions, Leibniz-Reynolds transport theorem, boundary conditions
- LECTURE 6: Kinematics of incompressible flow
 - [4-minute video on streamlines](#) from Prof. A. Hogg.
 - Chapter 19: scalar streamfunction, vector streamfunction, area and volume conservation, meridional-depth overturning circulation
- LECTURE 7: momentum dynamics
 - [6-minute video on momentum](#) from Prof. A. Hogg.
 - Chapter 22: momentum dynamics, accelerations, contact forces, body forces, special forms of the momentum equation, axial angular momentum
- LECTURE 8: thermodynamics
 - Section 4.8: exact and inexact differentials
 - Chapter 23: First law of thermodynamics, thermodynamic potentials, ideal gas atmosphere
- LECTURE 9: thermodynamics, energy dynamics, and buoyancy
 - Chapter 23: thermodynamics of a moving fluid, potential temperature
 - Chapter 24: mechanical energy, internal energy, total energy

- Chapter 25: Archimedes' principle, buoyancy, stratification, gravitational stability, mass density for perfect and realistic fluids
- LECTURE 10: stress in fluids
 - [2.5-minute video on stress and strain](#) from Prof. A. Hogg.
 - [8-minute video on stress](#) from Prof. A. Hogg.
 - Chapter 26: stresses and the stress tensor, linear momentum budget, relating stress to strain, form stress, boundary conditions
- LECTURE 11: filtered equations
 - [8-minute video on hydrostatic pressure](#) from Prof. A. Hogg.
 - Chapter 27: primitive equations, hydrostatic approximation, tangent plane approximation, basics of time evolving a fluid state
 - Chapter 28: oceanic Boussinesq approximation, equations for an ocean model
- LECTURE 12: Geostrophy and thermal wind
 - [26-minute video from Prof. D. Fultz](#) for an overview of rotating fluids.
 - [4-minute video from the UCLA SpinLab](#) for examples of Taylor columns.
 - Chapter 29: Rossby number, geostrophy, planetary geostrophy, Taylor-Proudman, thermal wind, isopycnal form stress
- LECTURE 13: Balanced inviscid horizontal flows
 - Chapter 30: natural coordinates; centripetal, centrifugal, and Coriolis accelerations; exact geostrophic flow; inertial motion of fluid particles; cyclostrophic balance; gradient wind balance
- LECTURE 14: Ekman layer mechanics
 - Start around the 23-minute mark [of this video from Prof. Fultz](#) for his discussion of Ekman layers.
 - Chapter 31: natural coordinates, spiral motion across isobars, non-dimensionalization and the Ekman number, net mass transport
- LECTURE 15: Advection and diffusion
 - Chapter 37: advection maths and physics; diffusion maths and physics
- LECTURE 16: Formulation of shallow water models
 - [30-minute video](#) on shallow water model from Dr. N. Hall.
 - Chapter 42: thickness equation, momentum equation, reduced gravity model, stacked shallow water layers, shallow water layer in a rotating tank
- LECTURE 17: Shallow water dynamics
 - Chapter 43: geostrophy, form stress, mechanical energy including available potential energy, angular momentum in a tank

- LECTURE 18: gravity waves and geostrophic adjustment
 - [14-minute video on gravity waves](#) from Dr. N. Hall.
 - Chapter 44: gravity waves in single layer, geostrophic adjustment
- LECTURE 19: vorticity, circulation, and potential vorticity
 - [23-minute video on vorticity](#) from Prof. A. Shapiro
 - Chapter 46: vorticity and circulation
 - Chapter 47: shallow water vorticity and potential vorticity
- LECTURE 20: vorticity mechanics
 - [21-minute video on vorticity](#) from Prof. A. Shapiro.
 - [5-minute video on vortex rings and Helmholtz's theorems](#) from the Physics Girl.
 - Chapter 48: vortex lines and tubes, Kelvin's circulation theorem, mechanics of baroclinicity, β -effect
- LECTURE 21: potential vorticity mechanics
 - Chapter 50: PV material invariance for perfect fluid; PV evolution with friction and heating; impermeability theorem; isopycnal layer integrated PV
- LECTURE 22: balanced models I (single layer)
 - Chapter 53: barotropic vorticity equation
 - Chapter 54: Buckingham's Π theorem, asymptotic expansion in terms of small Rossby number, shallow water planetary geostrophy and quasi-geostrophy
- LECTURE 23: balanced models II (continuous stratification)
 - Chapter 55: asymptotic derivation of continuously stratified planetary geostrophy and properties of these equations.
 - Chapter 56: asymptotic derivation of continuously stratified quas-geostrophy and properties of these equations.
- LECTURE 24: Course wrap up

2

A continuum basis for fluid mechanics

Viewed macroscopically, a fluid deforms continuously when applying a force so that a fluid has no preferred shape. Correspondingly, a fluid cannot sustain a shearing stress in the absence of motion. Ordinary gases and liquids are canonical examples of fluids, with gases filling any container with its molecules widely separated whereas molecules in liquids are much closer together so that liquids are far less compressible than gases.

For geophysical fluid mechanics, we are concerned with the atmosphere (mostly a gas) and the ocean (mostly a liquid). We are furthermore interested in macroscopic properties of fluid motion, with no interest in describing molecular degrees of freedom. Nor do we consider rarefied gas dynamics, which is a subject appropriate for the upper bounds of the atmosphere. For these reasons we pursue a phenomenological approach that makes use of conservation laws describing the motion of a continuous fluid media. This treatment is based on the *continuum hypothesis*, which assumes that mathematical limits for fluid volumes tending to zero are reached on length and time scales very large compared to molecular scales. The temporal version of the continuum hypothesis corresponds to *quasi-static processes*, which forms the basis for quasi-equilibrium thermodynamics. Quasi-static processes refer to macroscopic motion that evolves with time scales far larger than time scales of molecular motions, so that when treated thermodynamically a fluid evolves smoothly from one local thermodynamic equilibrium state to another.

Operationally, the continuum hypothesis allows us to make use of differential calculus for describing the mechanics of fluid motion. That is, the continuum hypothesis makes fluid mechanics a continuous space-time field theory.

READER'S GUIDE TO THIS CHAPTER

This chapter presents salient points concerning a continuum description of fluid mechanics. Our goal is to physically and concisely unpack the dictum “macroscopically small yet microscopically large”, which summarizes the regime considered by the continuum and quasi-static hypotheses. For this purpose we borrow from the kinetic theory of gases as treated in statistical physics books such as [Reif \(1965\)](#) and [Huang \(1987\)](#). However, no prior exposure to these treatments is necessary, nor do we dive into the details. Every chapter makes use of the continuum hypothesis and quasi-static hypothesis, thus warranting their presentation at the start of this book.

2.1	A variety of length scales	12
2.1.1	Molecular and macroscopic length scales	12
2.1.2	Continuous fields rather than discrete molecules	13
2.1.3	Reynolds number and the continuum length scale	14
2.1.4	Resolution of measurements and simulations	14
2.1.5	Comments	15
2.2	Results from kinetic theory	15
2.2.1	A mole and Avogadro's number	15
2.2.2	Ideal gas law	16
2.2.3	Molecular mean free path	17
2.2.4	Root mean square molecular speed	17
2.2.5	Basis for the quasi-static approximation	17
2.2.6	Macroscopically small and microscopically large	18
2.2.7	Whence a rigorous treatment?	18
2.2.8	Further study	19

2.1 A variety of length scales

Matter is comprised of molecules. However, fluid mechanics is not concerned with the motion of individual molecular degrees of freedom. Rather, fluid mechanics is concerned with phenomenological conservation laws satisfied by a continuous fluid material. This approach represents an idealization that is supported by centuries of successful descriptions of macroscopic fluid motion in the environment and laboratory.

We here outline the essential features of the continuum hypothesis. This hypothesis supports our macroscopic description of a fluid in terms of continuous fields rather than discrete molecules. More details are offered in Section 2.2, although a full discussion is outside the subject of fluid mechanics, instead resting deep within the field of statistical physics.

2.1.1 Molecular and macroscopic length scales

A fluid mechanical description focuses on fluid regions that are macroscopically small (e.g., $L_{\text{macro}} \sim 10^{-3}$ m) yet microscopically large (e.g., $L_{\text{macro}} \gg L_{\text{mfp}} \sim 10^{-7}$ m, where L_{mfp} is the molecular mean free path). A region of air with volume L_{macro}^3 contains roughly 10^{16} air molecules, whereas that same volume in water contains roughly 10^{19} water molecules. These numbers (justified in Section 2.2) illustrate the notions of macroscopically small yet microscopically large. It is only when reaching length scales on the order of the molecular mean free path that we need to be concerned with the discrete nature of matter. Figure 2.1 offers a schematic to illustrate these quite distinct length scales.

The huge number of molecules within a macroscopically tiny region justifies our assumption that physical properties are homogeneous over regions of size L_{macro} . In essence, this *continuum hypothesis* works with small but finite sized fluid elements whose mean dynamical properties (e.g., velocity, vorticity) and thermodynamical properties (e.g., mass density, matter concentration, temperature, specific entropy) are defined locally at any point within the continuous fluid media.

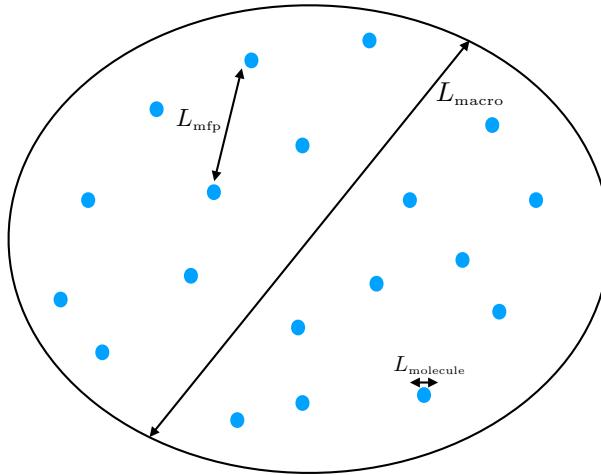


Figure 2.1: Schematic to illustrate the three length scales considered when making the continuum hypothesis. The blue circles represent molecules with diameter L_{molecule} . On average, molecules are separated by the mean free path, $L_{\text{mfp}} \approx 1000 L_{\text{molecule}}$. The smallest macroscopic length scale of interest for fluid mechanics is $L_{\text{macro}} \sim 10^{-3} \text{ m}$, which is roughly $L_{\text{macro}} = 10^4 L_{\text{mfp}}$ for an ideal gas at standard conditions. A region of air with volume L_{macro}^3 contains roughly 10^{16} air molecules, whereas that same volume in water contains roughly 10^{19} water molecules. For either case, the Law of Large Numbers greatly helps in taking the continuum limit. Note that this schematic is not drawn to scale!

2.1.2 Continuous fields rather than discrete molecules

When measured on length scales of the mean free path, material properties exhibit very large fluctuations on time scales of order $L_{\text{mfp}}/v_{\text{rms}}$, where v_{rms} is the root-mean-square speed of a fluid molecule (see Section 2.2.4). However, on macroscopic scales encompassing many molecular degrees of freedom, fluid matter appears continuous in both space and time.

The ratio of the mean free path to the macroscopic length scale is known as the Knudsen number

$$\text{Kn} = \frac{L_{\text{mfp}}}{L_{\text{macro}}}. \quad (2.1)$$

Large mean free paths occur for certain rarefied gases such as in the outer regions of the earth's atmosphere. Under these conditions, there are very few molecular collisions due to the tiny number density of molecules, thus supporting relatively large mean free paths. For our purposes, we are concerned only with fluid conditions where the mean free path is microscopic so that the Knudsen number is tiny

$$\text{Kn} \ll 1. \quad (2.2)$$

For tiny Knudsen numbers, we are led to make use of the continuum hypothesis.

When making the continuum hypothesis, we employ fluid properties that take values at each point within a space and time continuum, (\mathbf{x}, t) . For example, we make use of the mass density, $\rho(\mathbf{x}, t)$, fluid velocity, $\mathbf{v}(\mathbf{x}, t)$, pressure $p(\mathbf{x}, t)$, temperature $T(\mathbf{x}, t)$, tracer concentration, $C(\mathbf{x}, t)$, and other thermodynamic fields. These continuous fluid properties, or fields, formally represent the mean of molecular properties estimated over a linear dimension L_{macro} that is large microscopically but small macroscopically

$$L_{\text{molecule}} \ll L \ll L_{\text{macro}}. \quad (2.3)$$

As noted in Section 2.2.6, $L_{\text{macro}} \sim 10^{-3} \text{ m}$, which is much larger than the molecular mean free path.

2.1.3 Reynolds number and the continuum length scale

The continuum field equations of fluid mechanics are formally established for fluid motions with length scales on the order of L_{macro} and larger. We stated earlier that L_{macro} is on the order of a millimetre, with that length loosely based on noting that there are a huge number of molecules in a volume with this size. Furthermore, most macroscopic measurements have a resolution no finer than a millimetre. We here describe another means to determine this length scale.

Namely, we set L_{macro} to the length scale at which the Reynolds number is order unity

$$Re_{\text{macro}} = \frac{U L_{\text{macro}}}{\nu} \sim 1. \quad (2.4)$$

In this equation, $\nu > 0$ is the kinematic viscosity (dimensions squared length per time), and U is the scale for a macroscopic fluid velocity fluctuation. The Reynolds number measures the ratio of inertial accelerations (accelerations felt by fluid elements) to frictional accelerations from viscous forces (forces due to the rubbing of fluid elements against one another). When this ratio is on the order of unity, viscous forces play a leading role in the acceleration of the fluid. In particular, it is at this scale that viscous accelerations are able to dissipate kinetic energy and to thus put a halt to the cascade of kinetic energy moving to the smaller scales through nonlinear turbulent processes. Without viscous effects, the turbulent cascade would lead to an ultraviolet energy catastrophe in which mechanical energy unboundedly accumulates at the small scales. We are thus motivated to let the length scale where viscosity is important determine L_{macro} .

For air, the kinematic viscosity is (page 75 of [Gill \(1982\)](#))

$$\nu_{\text{air}} = \frac{1.7 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}}{1.3 \text{ kg m}^{-3}} = 1.3 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}, \quad (2.5)$$

and a typical fluid velocity fluctuation has a scale 10^{-1} m s^{-1} , so that

$$L_{\text{macro}} \approx 10^{-4} \text{ m} = 0.1 \text{ mm}. \quad (2.6)$$

Water has a kinematic viscosity (page 75 of [Gill \(1982\)](#))

$$\nu_{\text{air}} = \frac{10^{-3} \text{ kg m}^{-1} \text{ s}^{-1}}{1000 \text{ kg m}^{-3}} \approx 10^{-6} \text{ m}^2 \text{ s}^{-1}, \quad (2.7)$$

and a fluid velocity fluctuation about 10 times smaller than air. Hence, the macroscopic length scale for water is on the order of that for air.

2.1.4 Resolution of measurements and simulations

When we measure fluid motions in the laboratory or field, we generally do not measure the motions at scales on the order of L_{macro} . That is, our measurement devices generally have a spatial resolution much coarser than L_{macro} , so that $L_{\text{measure}} \gg L_{\text{macro}}$. Likewise, numerical simulations are generally made with discrete grid cells with length scales $L_{\text{numerical}} \gg L_{\text{macro}}$. The equations describing motions at the measurement/simulation length scales involve correlations of nonlinear fluctuations occurring at the smaller (unmeasured) scales. The correlations play a role, sometimes a dominant role, in the evolution of the measured scales. The parameterization of these correlations in terms of measured/simulated motions constitutes the turbulence closure problem. We do not study turbulence closure theories in this book, though we do identify its implications at certain points.

It is important to acknowledge the limitations of measurements in accurately characterizing fine scale motions. For this purpose define a gradient length scale

$$L_{\text{gradient}} = \frac{\mathbf{v}}{|\nabla \mathbf{v}|}, \quad (2.8)$$

where \mathbf{v} is the velocity of a fluid element and $|\nabla \mathbf{v}|$ is the magnitude of velocity gradients. Decomposing fluctuations into Fourier modes allows us to see that an accurate measurement of velocity fluctuations with length scales L_{gradient} requires a measurement length scale that satisfies

$$2\pi L_{\text{measure}} \leq L_{\text{gradient}}. \quad (2.9)$$

This constraint means that to measure velocity fluctuations on a scale L_{gradient} requires a finer measurement sampling with $L_{\text{measure}} = L_{\text{gradient}}/(2\phi)$.

2.1.5 Comments

The above discussion of length scales transfers seamlessly over to time scales through dividing the length scale by the velocity scale. Correspondingly, fluctuations with time scales shorter than $2\pi T_{\text{measure}}$ cannot be accurately measured.

2.2 Results from kinetic theory

If the reader is content to accept the continuum hypothesis on face value, then the material in this section can be readily skipped. For others, this section outlines results from the kinetic theory of ideal gases in support of the continuum hypothesis. Deductive treatments that transition from molecular mechanics to macroscopic fluid mechanics is a topic of the kinetic theory of gases and liquids, which is well outside our scope. In Section 2.2.8, we provide literature pointers for those wishing more rigor concerning the continuum limit of matter.

2.2.1 A mole and Avogadro's number

There are a tremendous number of molecules in the tiniest drop of water or puff of air. Just how many? To answer this question, we introduce the notion of a mole of matter. A mole is defined as the mass of a material substance that contains Avogadro's number of that substance, where

$$A_v = 6.022 \times 10^{23} \text{ mole}^{-1}. \quad (2.10)$$

Avogadro's number, A_v , is the proportionality constant converting from one molar mass of a substance to the mass of a substance. Avogadro's number is conventionally specified so that one mole of the carbon isotope ^{12}C contains exactly 12 grams. Hence, 12 grams of ^{12}C contains 6.022×10^{23} atoms of ^{12}C . Avogadro's number provides a connection between scales active in the microscopic world of molecules to the macroscopic world of everyday experience.

Dry air (air with no water vapor) is comprised of oxygen molecules O_2 , at roughly 22% by molecular mass, and nitrogen molecules N_2 , at roughly 78% molecular mass.¹ The molar mass of dry air is thus

$$M_{\text{air}} = 0.22 * 32 \text{ g mole}^{-1} + 0.78 * 28 \text{ g mole}^{-1} \approx 28.8 \text{ g mole}^{-1}. \quad (2.11)$$

¹We here ignore the presence of other trace gases, such as CO_2 , although these gases are critical for understanding atmospheric radiation and hence the earth's energy budget.

Pure (fresh) water is comprised of two hydrogen atoms and one oxygen atom. The molar mass of pure water is thus given by

$$M_{\text{water}} = 2 * 1 \text{ g mole}^{-1} + 16 \text{ g mole}^{-1} = 18 \text{ g mole}^{-1}. \quad (2.12)$$

2.2.2 Ideal gas law

The ideal gas law is given by

$$pV = nRT, \quad (2.13)$$

where p is the pressure, V is the volume, n is the number of moles, R is the universal gas constant, and T is the absolute temperature (temperature relative to absolute zero). Measuring the temperature in Kelvin leads to the universal gas constant

$$R = 8.314 \text{ J mole}^{-1} \text{ K}^{-1} = 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1}, \quad (2.14)$$

where the second equality replaced the energy unit, Joule, by its MKS equivalent,

$$\text{J} = \text{kg m}^2 \text{ s}^{-2}. \quad (2.15)$$

Use of the ideal gas law (2.13) says that one mole of ideal gas at standard temperature ($0^\circ\text{C} = 273.15 \text{ K}$) and standard atmospheric pressure ($101.325 \times 10^3 \text{ Pa}$) occupies the following volume

$$V = \frac{nRT}{p} \quad (2.16a)$$

$$= \frac{(1 \text{ mole})(8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1})(273.15 \text{ K})}{101.325 \times 10^3 \text{ kg m}^{-1} \text{ s}^{-2}} \quad (2.16b)$$

$$\approx 2.25 \times 10^{-2} \text{ m}^3, \quad (2.16c)$$

where we introduced the MKS units for pressure (force per unit area)

$$\text{Pa} = \text{N m}^{-2} = \text{kg m}^{-1} \text{ s}^{-2}. \quad (2.17)$$

Hence, the number density (number of molecules) for a mole of ideal gas is given by

$$n_{\text{gas}} = \frac{\text{number per mole}}{\text{volume per mole}} \quad (2.18a)$$

$$= \frac{A_v}{V} \quad (2.18b)$$

$$= \frac{6.022 \times 10^{23}}{2.25 \times 10^{-2} \text{ m}^3} \quad (2.18c)$$

$$= 2.68 \times 10^{25} \text{ m}^{-3}. \quad (2.18d)$$

Specializing to air, we compute the mass density of air at standard temperature and pressure as

$$\rho_{\text{air}} = \frac{M_{\text{air}}}{V} = \frac{28.8 \times 10^{-3} \text{ kg}}{2.25 \times 10^{-2} \text{ m}^3} = 1.28 \text{ kg m}^{-3}, \quad (2.19)$$

where we set $M_{\text{air}} = 28.8 \times 10^{-3} \text{ kg}$ according to equation (2.11). This ideal gas density is close to the 1.225 kg m^{-3} density measured for air at standard conditions, thus giving us confidence for using the ideal gas law for dry air. Differences arise from trace constituents in air as well as inter-molecular forces (an ideal gas has no inter-molecular forces).

2.2.3 Molecular mean free path

We are in search of length scales relevant for molecular motion. One length scale is that of the molecule itself. Another is set by the distance between molecular collisions. The molecular mean free path is the mean distance that a molecule travels before colliding with another molecule. Arguments from kinetic theory of gases, applied to an ideal gas, lead to the expression

$$L_{\text{mfp}} = \frac{1}{\pi \sqrt{2} n_v d^2} \quad (2.20)$$

where d is the diameter of the molecule. The mean diameter of air molecules is roughly

$$d_{\text{molecule air}} \approx 2 \times 10^{-10} \text{ m.} \quad (2.21)$$

The mean free path for air molecules is thus given by

$$L_{\text{mfp}} = \frac{1}{\pi \sqrt{2} n_v d^2} \quad (2.22a)$$

$$= \frac{1}{\pi \sqrt{2} (2.68 \times 10^{25} \text{ m}^{-3}) (2 \times 10^{-10} \text{ m})^2} \quad (2.22b)$$

$$= 2 \times 10^{-7} \text{ m.} \quad (2.22c)$$

The mean free path for an air molecule is roughly 1000 times larger than the molecular diameter (e.g., Figure 2.1).

2.2.4 Root mean square molecular speed

What is the mean speed for molecules moving through a gas? Again, kinetic theory for ideal gases offers an explicit expression, here written in terms of the pressure and density of the gas

$$v_{\text{rms}} = \sqrt{\frac{3p}{\rho}} = \sqrt{\frac{3RT}{M}}. \quad (2.23)$$

Note the direct relation between pressure, temperature, and speed. That is, molecules move faster at higher temperature, and thus impart larger pressure on their surroundings. At standard pressure and temperature, the root-mean-square speed for an air molecule is given by

$$v_{\text{rms}} = \sqrt{\frac{3p_{\text{stand}}}{\rho_{\text{air}}}} \quad (2.24a)$$

$$= \sqrt{\frac{3(101.325 \times 10^3 \text{ kg m}^{-1} \text{ s}^{-2})}{1.28 \text{ kg m}^{-3}}} \quad (2.24b)$$

$$= 487 \text{ m s}^{-1}. \quad (2.24c)$$

To get a sense for the relative scale of this speed, note that the speed of sound in air at standard temperature and pressure is 331 m s^{-1} . So these molecules are moving faster than sound!

2.2.5 Basis for the quasi-static approximation

Assuming one collision occurs within a mean free path, and the molecules are moving at the root-mean-square speed, we can estimate the time between collision according to

$$t_{\text{collision}} = \frac{L_{\text{mfp}}}{v_{\text{rms}}} \quad (2.25)$$

The corresponding time for air is given by

$$t_{\text{air}} = \frac{2 \times 10^{-7} \text{ m}}{487 \text{ m s}^{-1}} = 4.1 \times 10^{-10} \text{ s.} \quad (2.26)$$

Inverting this number, we see that there are roughly $t_{\text{air}}^{-1} = 2.5 \times 10^9 \text{ s}^{-1}$ collisions per second.

The huge number of collisions per second means that for all macroscopic processes, including geophysical fluid flow, the dynamical time scales for the macroscopic motion are much larger than the time scales for molecular equilibration. We are thus led to define a *quasi-static macroscopic process* as one that occurs through a series of thermodynamic equilibrium states. Consequently, we can use equilibrium thermodynamic relations while allowing for time evolution of the macroscopic system. We have more to say on this topic in Chapter 23 when studying thermodynamics, with particular emphasis on the implications of the quasi-static approximation in Section 23.2.2.

2.2.6 Macroscopically small and microscopically large

For environmental measurements of the atmosphere and ocean, or for conventional measurements in fluid laboratories, we can detect differences in fluid properties (e.g., mass density, velocity, tracer concentration, thermodynamic state properties) for length scales on the order of

$$L_{\text{macro}} = 10^{-3} \text{ m.} \quad (2.27)$$

For macroscopic purposes, fluid properties are homogeneous over regions smaller than this length. Although macroscopically rather tiny, a fluid region of volume L_{macro}^3 is huge microscopically. We can see so by computing the number of molecules in this region.

At standard conditions, a volume of air of size L_{macro} contains

$$N_{\text{air molecules}} = V n_{\text{gas}} = (10^{-3} \text{ m})^3 (2.68 \times 10^{25} \text{ m}^{-3}) \approx 3 \times 10^{16} \text{ air molecules.} \quad (2.28)$$

To compute the number of water molecules in this same volume, we first use the water mass density of $\rho \approx 10^3 \text{ kg m}^{-3}$ to determine the water mass in this region

$$M_{\text{water}} = \rho_{\text{water}} V = (1000 \text{ kg m}^{-3}) (10^{-9} \text{ m}^3) = 10^{-6} \text{ kg.} \quad (2.29)$$

Water has a molar mass of $0.018 \text{ kg mole}^{-1}$, so a volume of $(10^{-3} \text{ m})^3$ contains²

$$N_{\text{water molecules}} = \left(\frac{10^{-6} \text{ kg}}{0.018 \text{ kg mole}^{-1}} \right) \times 6.022 \times 10^{23} \text{ molecules mole}^{-1} = 3 \times 10^{19} \text{ water molecules.} \quad (2.30)$$

Water thus has roughly 10^3 more molecules in this volume than air at standard pressure, which reflects the roughly 10^3 times larger mass density for water. Regardless, both water and air contain a huge number of molecules in this macroscopically tiny region.

2.2.7 Whence a rigorous treatment?

A rigorous derivation of continuum field theory, starting from molecular dynamics, is nontrivial even for an ideal gas, and largely non-existent for liquids. Indeed, some say a Nobel Prize awaits the person providing a fully deductive theory. For our purpose, we remain satisfied to postulate that a continuum description is valid for fluid mechanics of the atmosphere and ocean. A means for evaluating this postulate is to perform experimental measures and compare to the continuum theory. Centuries of experiments with fluid motions in the environment and laboratory lend credence to the continuum description. We consider these tests to be sufficient motivation to pursue the continuum approach for fluid mechanics and geophysical fluid dynamics.

²The calculation on page 9 of [Griffies \(2004\)](#) has a factor of 10^6 error.

2.2.8 Further study

Pedagogical treatments of the ideal gas law and kinetic theory can be found in most books on introductory physics or chemistry. [Vallis \(2017\)](#) provides extensions of the ideal gas law for an atmosphere with moisture.

For discussions of the continuum hypothesis reflecting that given here, see the terse discussion on page 1 of [Olbers et al. \(2012\)](#), or the more thorough treatments given in Section 1.2 of [Batchelor \(1967\)](#) and Section 1.4 of [Kundu et al. \(2012\)](#). Chapter 1 of [Salmon \(1998\)](#) offers an even more thorough treatment, touching on elements from kinetic theory and details for how to coarse grain average over molecular degrees of freedom (see his pages 3 and 4 and Sections 9, 10, and 11). A rigorous account of kinetic theory is offered in many treatments of statistical mechanics. That given by [Reif \(1965\)](#) and [Huang \(1987\)](#) are accessible to those with a physics undergraduate training. When reading the statistical mechanics literature, look for discussions of the “hydrodynamical limit,” which concerns the transition from discrete particle mechanics to continuum mechanics.

Part I

Geophysical fluid mathematics

Fluid mechanics is a classical field theory based on Newton's laws of mechanics and classical thermodynamics, both applied to a continuous fluid media. Geophysical fluid mechanics (GFM) is concerned with buoyancy stratified fluids of multiple constituents moving on a rotating sphere. Rotation, stratification, multiple constituents, and spherical geometry each influence the maths encountered in GFM. Our goal for this part of the book is to review mathematical topics with a focus on how they are useful for GFM. The general reader is encouraged to study Tier-I material and skim Tier-II material, whereas those aiming to delve deeper into the theory are encouraged to study the many Tier-II chapters. In either case, much of the Tier-II material can be readily returned to if needed to fill in the gaps in later chapters.

PHYSICS PROVIDES RELATIONS BETWEEN GEOMETRIC OBJECTS

Mathematical objects of use for the study of fluid mechanics include scalar fields (e.g., temperature, mass density, specific entropy), vector fields (e.g., velocity, vorticity), and second order tensor fields (e.g., diffusion tensor, stress tensor). These and other fields have an existence independent of the arbitrary coordinate choices used for their description. Thinking abstractly, they are geometric objects such as points, vectors, surfaces, volumes, etc. In the study of geophysical fluid mechanics, we thus use physical principles to develop equations relating geometric objects. The tools of tensor analysis are then used to compute numbers as required to compare with experiments and field measurements.

The above perspective of “physics as geometry” is foundational to theoretical physics (e.g., [Thorne and Blandford \(2017\)](#)) and it has conceptual and practical use for our study. It furthermore provides the framework for this part of the book, in which we develop mathematical tools that are later used to formulate a variety of theoretical geophysical fluid models. A key focus of this book concerns the development of mathematical tools to help unpack the physics encapsulated by the equations. This focus extends to those cases where analytical solutions are unavailable, which is the norm for nonlinear field theories such as fluid mechanics. In turn, such qualitative and conceptual tools are of great use for understanding and for prediction

TENSOR ANALYSIS AND GEOPHYSICAL FLUID MECHANICS

There are many occasions where a geophysical fluid system is more physically transparent when using a particular coordinate description or reference frame. However, there is no *a priori* choice that fits all systems. Thus, being adept at transforming from one description to another eases our study. Tensor analysis is the proven means for systematically performing such transformations, hence motivating its use for this book.

The following is an incomplete list of geophysical fluid systems where various coordinate descriptions or reference frames are encountered, and thus where tensor analysis can be put to use. Granted, each system listed here can be studied without the formalism of tensor analysis. However, by doing so one often encounters clumsy and burdensome manipulations that can obfuscate the underlying physical concepts. Indeed, imagine the tedium required to write field equations in multiple dimensions prior to vector analysis! That situation is akin to the tedium and awkward nature required to work across multiple coordinate systems and reference frames absent the formalism of tensor analysis. Hence, an adept use of vector analysis, and its generalization to tensor analysis, reveals how maths can inform the physics and how physics can be transparently embodied by the maths.

- There is a duality in fluid kinematics between Eulerian and Lagrangian descriptions of fluid motion. To develop an understanding of this duality we make use of tensor analysis to facilitate the transformation between the two descriptions.

- Geophysical fluids move on a spherical planet, making spherical coordinates the preferred choice for studying and modeling planetary flows. We make use of tensor methods to transform between planetary Cartesian coordinates (origin at the center of the planet) and spherical coordinates.
- Rotating laboratory fluids move in a circular tank, with cylindrical polar coordinates of use to respect symmetry of the domain. We make use of tensor methods to transform between Cartesian and cylindrical polar coordinates when considering rotating tank systems.
- Geophysical fluids move around a rotating earth close to solid-body motion. Terrestrial observers also move in near solid-body motion. We are thus motivated to study geophysical fluids from a rotating reference frame. We use rudimentary tensor methods to transform between a fixed inertial frame and the non-inertial rotating reference frame, with this transformation revealing non-inertial accelerations that impact on the observed fluid flow.
- Geophysical fluids move in a gravitational field that acts to stratify the fluid according to its local buoyancy. For many purposes it can be useful to describe the vertical position of a fluid particle according to its buoyancy rather than its height. This “isopycnal” vertical coordinate choice leads to a non-orthogonal coordinate description of the fluid motion. There are other vertical coordinates that can be of use for other situations. Transforming between a Cartesian and a generalized vertical coordinate description requires the mathematical precision of general tensors.

SUMMARY OF THE MATHS CHAPTERS

Some of the chapters in this part of the book are essential for nearly all subsequent chapters, whereas many others target the aficionados and serve somewhat limited, albeit interesting and compelling, purposes. All readers are encouraged to take a close look at each chapter if only to know where to find topics that might be of use later in the book or later in one’s career.

- **CARTESIAN TENSOR ALGEBRA:** Chapter 3 provides a synopsis of Cartesian tensor analysis. This topic provides a systematization of ideas from Cartesian geometry and linear algebra. Material in this chapter is essential for nearly every topic in this book.
- **CARTESIAN TENSOR CALCULUS:** Chapter 4 extends the algebraic ideas from Chapter 3 to differential and integral calculus. This chapter provides a resume of multivariate calculus of use for fluid mechanics. Material in this chapter is essential for nearly every topic in this book.
- **PARTIAL DIFFERENTIAL EQUATIONS:** Chapter 5 provides a summary of linear partial differential equations (PDEs) commonly encountered in mathematical physics. Even though the equations of fluid mechanics are nonlinear PDEs, their linear counterparts offer much insight into fluid behavior.
- **GEOMETRY OF CURVES AND SURFACES:** Chapter 6 introduces rudimentary differential geometry used to describe properties of curves (such as fluid particle trajectories) and surfaces (such as isopycnals).
- **GENERAL TENSOR INTRODUCTION:** Chapter 7 provides an introduction to general tensor analysis and its applications to geophysical fluids. The discussion is accessible to anyone who has read Chapter 3, and is recommended for all readers of this book, even for those who do not wish to study the details of general tensors in Chapters 8 and 9.

- GENERAL TENSOR ALGEBRA: Chapter 8 extends the Cartesian tensor algebra from Chapter 3 to allow for the use of arbitrary, or general, coordinates. This chapter is essential to understand the mathematics underlying non-Cartesian coordinates, such as spherical and isopycnal coordinates.
- GENERAL TENSOR CALCULUS: Chapter 9 extends the Cartesian tensor calculus from Chapter 4 to the case of general coordinates. Again, this chapter is essential to understand the mathematics underlying non-Cartesian coordinates, such as spherical and isopycnal coordinates.
- ORTHOGONAL COORDINATES: Chapter 10 offers a reference for various locally orthogonal coordinate systems (Cartesian, spherical, cylindrical) used in this book and how various mathematical objects appear when written in these coordinates.
- GENERALIZED VERTICAL COORDINATES: Chapter 11 offers a reference for the mathematics of generalized vertical coordinates. These non-orthogonal coordinates, such as isopycnal coordinates, are commonly used for conceptual and numerical models of stratified flows.

WHY SO MUCH MATHS?

Some may consider mathematics to be a burden to avoid rather than a joyful intellectual process to embrace. Consequently, there is a nonzero chance that certain readers will choose to completely skip this material, even the Tier-I chapters. Doing so comes at an unfortunate price since there is a nonzero level of mathematical acumen required to describe and to understand the physics of geophysical fluids. At some point, both the pedestrian and practitioner of geophysical fluid mechanics must allow mathematics to become a trusted and welcome friend. The treatment offered here is an opportunity to initiate that friendship and, if so choosing, to foster a certain degree of sophistication.

3

Cartesian tensor algebra

In this chapter we introduce the formalism of Cartesian tensor analysis, focusing here on the basic algebraic relations. The use of Cartesian tensors means we are only concerned with Cartesian coordinates and their orthogonal transformations via rotations. We follow standard treatments of Cartesian tensors such as that in Chapter 2 of *Aris (1962)*. The discussion should be accessible to those having studied undergraduate calculus and linear algebra. For geophysical fluid mechanics, mastery of Cartesian tensors is nearly sufficient for mastery of general tensors.

READER'S GUIDE TO THIS CHAPTER

This chapter is basic to all of the maths in this book. It should be familiar to those having taken an undergraduate linear algebra class, though here couched in the language of Cartesian tensors.

3.1	Introduction to tensors and tensor fields	26
3.2	Points and vectors	27
3.3	Distance and the scalar product	28
3.3.1	Distance between points	28
3.3.2	Magnitude of a vector and the scalar product	29
3.4	Vector product	29
3.4.1	Basis vector orientation and the Levi-Civita tensor	29
3.4.2	Orthogonality relations between cross products	30
3.4.3	Vector product of arbitrary vectors	31
3.4.4	Geometric interpretation of the vector product	31
3.4.5	Generalization to arbitrary vectors	32
3.5	Measuring volume	32
3.5.1	Volume defined by three vectors	32
3.5.2	Cartesian volume element for integration	33
3.5.3	n -space volumes and the Levi-Civita tensor	33
3.6	Example vector identities using the Levi-Civita tensor	34
3.6.1	Double vector product	34
3.6.2	Scalar product of two vector products	34
3.7	Transforming the coordinate representations	35
3.7.1	Inverse transformation	35
3.7.2	Orthogonal transformation	35
3.7.3	Geometric interpretation of orthogonal transformations	36
3.7.4	Transforming the coordinate representation of a vector	36
3.7.5	Form invariance of the scalar product	37
3.7.6	Transforming the coordinate representation of a second order tensor	37
3.7.7	Importance of distinguishing between tensors and matrices	37
3.8	Exercises	38

3.1 Introduction to tensors and tensor fields

Fluid mechanics involves fields of scalars, vectors, and tensors. We generically refer to all of these geometric objects as *tensors*, with a scalar a zero order tensor and a vector a first order tensor. A scalar field at a point provides a single number and a scalar field provides a number at each point in space-time. Example scalar fields include temperature, mass density, entropy, salinity, humidity, and mechanical energy. A vector connects two points and is specified by a direction and a magnitude, with a vector field providing a vector at each point in space-time. Example vector fields include the fluid velocity and forces acting on fluid elements. A second order tensor can be represented by a matrix, with a tensor field providing a matrix at each point in space-time. The stress tensor and the diffusion tensor are examples encountered in this book.

Geophysical fluids are embedded in the Newtonian world of universal time and flat Euclidean space. This space-time introduces the familiar Euclidean norm when measuring the spatial distance between points, whether the points are on a plane, a sphere, or an arbitrary surface within the fluid such as a surface of constant specific entropy. We can thus make use of Cartesian coordinates as the starting point for a mathematical formulation of geophysical fluid mechanics. Transformations to alternative coordinates are made when they lend insight to the symmetry of the flow or the geometry of the surface on which the flow occurs. It is for this reason that we devote this chapter to developing the formalism of Cartesian tensor analysis. Furthermore, it is for this reason that Cartesian tensors gives us nearly all of the formalism necessary to study general tensors in geophysical fluid mechanics.

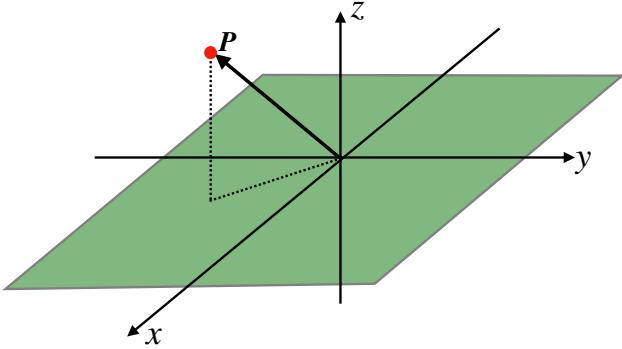


Figure 3.1: An arbitrary point in space, \mathcal{P} , has an objective existence independent of our subjective choice of coordinate system used to describe its position. We here represent its position with respect to the origin of a right-handed Cartesian coordinate system. The Cartesian representation of the position vector is $\mathcal{P} \mapsto \vec{P} = \mathbf{P} = \hat{\mathbf{x}} P_1 + \hat{\mathbf{y}} P_2 + \hat{\mathbf{z}} P_3$, with the Cartesian basis vectors the normalized unit vectors $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ and (P_1, P_2, P_3) the Cartesian coordinates. There are an infinite number of such Cartesian coordinate systems that are rotated and/or translated with respect to the one shown here.

3.2 Points and vectors

Consider a point, \mathcal{P} , in three dimensional Euclidean space \mathbb{R}^3 . We can represent its spatial position by providing its Cartesian coordinates relative to an arbitrary origin. As such, the position is a vector whose tail is at the origin and head at the point as show in Figure 3.1. We write this coordinate representation as

$$\mathcal{P} \mapsto \vec{P} = \hat{\mathbf{x}} P_1 + \hat{\mathbf{y}} P_2 + \hat{\mathbf{z}} P_3. \quad (3.1)$$

Vectors are denoted by an arrow. The right hand side of equation (3.1) provides the representation of the position vector in terms of its Cartesian coordinates, (P_1, P_2, P_3) , that measure distance along their corresponding Cartesian unit vectors, $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. The Cartesian unit vectors form a basis for three dimensional Euclidean space.¹ Hence, the position vector for any point in space can be represented in terms of these three basis vectors.

We find it useful to make use of alternative notations in which the position vector is written

$$\vec{P} = \mathbf{P} \quad (3.2a)$$

$$= \hat{\mathbf{x}} P_1 + \hat{\mathbf{y}} P_2 + \hat{\mathbf{z}} P_3 \quad (3.2b)$$

$$= \sum_{a=1}^3 \vec{e}_a P_a \quad (3.2c)$$

$$= \vec{e}_a P_a. \quad (3.2d)$$

The first equality introduced the boldface notation, which we commonly use for the representation of vectors in Cartesian coordinates. Although less convenient with the general tensors of Chapter 7, we still maintain the boldface given its common usage in the literature. Equation (3.2c) introduced a generic notation for the basis vectors

$$\vec{e}_1 = \hat{\mathbf{x}} \quad \vec{e}_2 = \hat{\mathbf{y}} \quad \vec{e}_3 = \hat{\mathbf{z}}. \quad (3.3)$$

Equation (3.2d) introduced the Einstein summation convention in which repeated indices are summed over their range, thus allowing us to drop the summation symbol.

¹The unit vectors are sometimes denoted $(\hat{i}, \hat{j}, \hat{k})$ in the literature.

We emphasize that the tensor labels denote components of coordinates, P_a , and members from the set of basis vectors, \vec{e}_a . These labels are not to be confused with partial derivative operations.² We sometimes write the vector components and basis vectors in the form of a list

$$P_a = (P_1, P_2, P_3) \quad \vec{e}_a = (\vec{e}_1, \vec{e}_2, \vec{e}_3) = (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}). \quad (3.4)$$

Use of a hat or carot symbol over a vector signifies that the vector is normalized to unity. For Cartesian coordinates we generally work with the unit basis vectors (3.3). Furthermore, a normalized vector can change only through rotation since by definition it remains of unit norm (see Section 4.1.4). Note that for the general tensors of Chapter 7, the most convenient basis vectors are not necessarily normalized.

3.3 Distance and the scalar product

In defining a vector to have unit magnitude, we are assuming we know how to measure the magnitude of a vector. We here make this notion precise.

3.3.1 Distance between points

Consider two points in Euclidean space, \mathcal{P} and $\mathcal{P} + d\mathcal{P}$, separated by a small distance and specified by their respective position vectors

$$\mathcal{P} \mapsto \vec{P} = \vec{e}_a P_a \quad (3.5a)$$

$$\mathcal{P} + d\mathcal{P} \mapsto \vec{P} + d\vec{x} = \vec{e}_a (P_a + dx_a). \quad (3.5b)$$

Euclidean space is afforded a metric whereby the squared distance between two points is measured via the Pythagorean Theorem

$$[\text{distance}(\mathcal{P}, \mathcal{P} + d\mathcal{P})]^2 = (\vec{P} + d\vec{x} - \vec{P}) \cdot (\vec{P} + d\vec{x} - \vec{P}) \quad (3.6a)$$

$$= (\mathbf{P} + d\mathbf{x} - \mathbf{P}) \cdot (\mathbf{P} + d\mathbf{x} - \mathbf{P}) \quad (3.6b)$$

$$= dx_a dx_b (\vec{e}_a \cdot \vec{e}_b) \quad (3.6c)$$

$$= dx_a dx_b \delta_{ab} \quad (3.6d)$$

$$= dx_a dx_a \quad (3.6e)$$

$$= (dx_1)^2 + (dx_2)^2 + (dx_3)^2. \quad (3.6f)$$

To reach this result we introduced the components to the Kronecker delta tensor, which can be represented by the 3×3 identity matrix

$$\vec{e}_a \cdot \vec{e}_b = \delta_{ab} = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.7)$$

The Kronecker tensor provides the Cartesian coordinate representation of the *metric* for Euclidean space. The metric provides the means to measure the distance between points on a manifold, and how to measure the length of a vector. It thus allows us to *normalize* a vector to have unit magnitude, motivating some to use the term *norm* rather than metric. In Section 8.1 we introduce alternative representations for the metric based on the use of non-Cartesian coordinates and non-Euclidean manifolds.

²To help avoid confusion we generally eschew the notation where partial derivatives are denoted by a subscript.

3.3.2 Magnitude of a vector and the scalar product

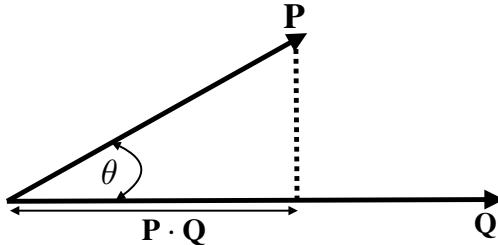


Figure 3.2: Illustrating the geometry associated with forming the scalar product between two vectors, $\mathbf{P} \cdot \mathbf{Q} = |\mathbf{P}| |\mathbf{Q}| \cos \theta$.

By defining the distance between two points, we in turn have a prescription for defining the squared magnitude of a vector

$$|\mathbf{P}|^2 = \mathbf{P} \cdot \mathbf{P} = P_a P_b (\vec{e}_a \cdot \vec{e}_b) = P_a P_a = (P_1)^2 + (P_2)^2 + (P_3)^2. \quad (3.8)$$

Correspondingly, we have the scalar (or dot) product between two arbitrary vectors

$$\mathbf{P} \cdot \mathbf{Q} = P_a Q_b (\vec{e}_a \cdot \vec{e}_b) = P_a Q_a. \quad (3.9)$$

Given our expression for the scalar product and the magnitude of vectors, we can introduce a geometrical interpretation by defining the angle between the vectors according to

$$\cos \theta \equiv \frac{\mathbf{P} \cdot \mathbf{Q}}{|\mathbf{P}| |\mathbf{Q}|} = \frac{P_a Q_a}{\sqrt{P_a P_a} \sqrt{Q_b Q_b}}. \quad (3.10)$$

We illustrate this equation in Figure 3.2. It is useful to convince oneself that this definition is consistent with $-1 \leq \cos \theta \leq 1$.

3.4 Vector product

The scalar product provides a means to measure the magnitude of a vector and the distance between two points. We here introduce the vector (or cross) product, which provides a means to measure area associated with two vectors and to specify the orientation of that area.

3.4.1 Basis vector orientation and the Levi-Civita tensor

Consider a flat plane defined by any two of the Cartesian basis vectors, \vec{e}_a and \vec{e}_b . We seek a means to specify what side of the plane is up and what side is down. Doing so allows us to orient objects within space. Notably, there is no objective means for this specification, since “up” and “down” are relative to a chosen orientation. Therefore, we must choose a convention. For that purpose, we follow the *right hand rule*, in which the out-stretched thumb, index, and middle fingers of the right hand orient the three Cartesian basis vectors.

We algebraically specify the right hand rule for the basis vectors through the relation³

$$\vec{e}_a \wedge \vec{e}_b = \epsilon_{abc} \vec{e}_c. \quad (3.11)$$

³Many authors choose the symbol \times for the vector product rather than the wedge symbol, \wedge . The wedge is used here as it lends itself to less confusion with the coordinate x .

The left hand side introduces the vector (or cross) product of two basis vectors. The right hand side algebraically defines the vector product as the contraction of the Levi-Civita tensor with another basis vector. The Cartesian components of the Levi-Civita tensor are given by the totally anti-symmetric permutation symbol

$$\epsilon_{123} = 1 \quad (3.12a)$$

$$\epsilon_{abc} = \begin{cases} 1, & \text{even permutation of } abc \text{ (123, 312, 231)} \\ -1, & \text{odd permutation of } abc \text{ (321, 132, 213)} \\ 0, & \text{all other } abc. \end{cases} \quad (3.12b)$$

Exchanging indices (an odd permutation) flips the sign of the permutation symbol

$$\epsilon_{abc} = -\epsilon_{bac} = -\epsilon_{acb}, \quad (3.13)$$

whereas cycling indices (an even permutation) preserves the sign

$$\epsilon_{abc} = \epsilon_{cab} = \epsilon_{bca}. \quad (3.14)$$

3.4.2 Orthogonality relations between cross products

As defined, the permutation symbol ensures that $\vec{e}_a \wedge \vec{e}_b$ is orthogonal to both \vec{e}_a and \vec{e}_b

$$\vec{e}_a \cdot (\vec{e}_a \wedge \vec{e}_b) = \vec{e}_a \cdot \epsilon_{abc} \vec{e}_c \quad \text{definition} \quad (3.15a)$$

$$= \epsilon_{abc} \vec{e}_a \cdot \vec{e}_c \quad \text{rearrangement} \quad (3.15b)$$

$$= \epsilon_{cba} \vec{e}_c \cdot \vec{e}_a \quad \text{relabel } a \text{ to } c \text{ and } c \text{ to } a \quad (3.15c)$$

$$= -\epsilon_{abc} \vec{e}_c \cdot \vec{e}_a \quad \text{cba is an odd permutation of abc} \quad (3.15d)$$

$$\Rightarrow \vec{e}_a \cdot (\vec{e}_a \wedge \vec{e}_b) = 0. \quad (3.15e)$$

To fully digest step (3.15c) it can be useful to reintroduce the summation symbol so that

$$\epsilon_{abc} \vec{e}_a \cdot \vec{e}_c = \sum_{a=1}^3 \sum_{c=1}^3 \epsilon_{abc} \vec{e}_a \cdot \vec{e}_c \quad \text{summation symbols reintroduced} \quad (3.16a)$$

$$= \sum_{c=1}^3 \sum_{a=1}^3 \epsilon_{cba} \vec{e}_c \cdot \vec{e}_a \quad \text{swap } a \text{ and } c \quad (3.16b)$$

$$= \epsilon_{cba} \vec{e}_c \cdot \vec{e}_a \quad \text{reintroduce summation convention} \quad (3.16c)$$

Additionally, to digest step (3.15d) we step through the permutations

$$\epsilon_{cba} = -\epsilon_{bca} \quad \text{swap } c \text{ with } b \text{ to pick up a minus sign} \quad (3.17a)$$

$$= \epsilon_{bac} \quad \text{swap } c \text{ with } a \text{ to pick up a minus sign} \quad (3.17b)$$

$$= -\epsilon_{abc} \quad \text{swap } b \text{ with } a \text{ to pick up a minus sign.} \quad (3.17c)$$

The same procedure shows that $\vec{e}_b \cdot (\vec{e}_a \wedge \vec{e}_b) = 0$. Hence, the vector product is orthogonal to the plane specified by any two of the basis vectors. That is, the vector product points orthogonal to that plane and in a direction determined by the right hand rule. We note that this proof reveals a general property. Namely, there is a zero contraction of a symmetric tensor (e.g., the scalar product $\vec{e}_c \cdot \vec{e}_a$) with an anti-symmetric tensor (see Exercise 3.2).

3.4.3 Vector product of arbitrary vectors

The expression (3.11) for the vector product of two basis vectors renders the vector product of arbitrary vectors.

$$\mathbf{P} \wedge \mathbf{Q} = P_a \vec{e}_a \wedge Q_b \vec{e}_b \quad (3.18a)$$

$$= P_a Q_b \vec{e}_a \wedge \vec{e}_b \quad (3.18b)$$

$$= P_a Q_b \epsilon_{abc} \vec{e}_c \quad (3.18c)$$

$$= (P_2 Q_3 - P_3 Q_2) \vec{e}_1 + (P_3 Q_1 - P_1 Q_3) \vec{e}_2 + (P_1 Q_2 - P_2 Q_1) \vec{e}_3. \quad (3.18d)$$

We can write the vector product as a determinant

$$\mathbf{P} \wedge \mathbf{Q} = \det \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \end{bmatrix}. \quad (3.19)$$

As with the basis vectors, the vector product is orthogonal to both of the individual vectors

$$\mathbf{P} \cdot (\mathbf{P} \wedge \mathbf{Q}) = (P_d \vec{e}_d) \cdot (P_a Q_b \epsilon_{abc} \vec{e}_c) \quad (3.20a)$$

$$= P_c P_a Q_b \epsilon_{abc} \quad (3.20b)$$

$$= 0, \quad (3.20c)$$

where the final equality follows since the product $P_c P_a$ is symmetric on the labels ac , whereas ϵ_{abc} is anti-symmetric.

3.4.4 Geometric interpretation of the vector product

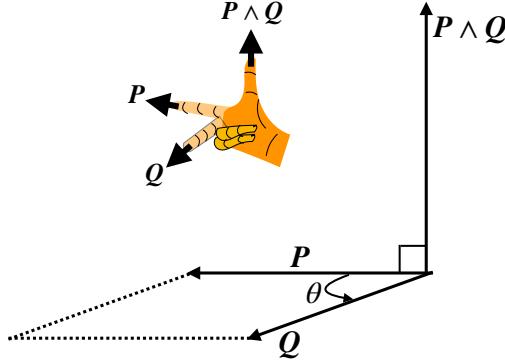


Figure 3.3: The magnitude for the vector product between two vectors is given by the product of their magnitudes and the sine of the angle between them, $|\mathbf{P} \wedge \mathbf{Q}| = |\mathbf{P}| |\mathbf{Q}| \sin \theta$. This magnitude equals to the area of the parallelogram subtended by the two vectors. The vector product is directed perpendicular to the plane determined by the two vectors and oriented according to the right hand rule. We depict the right hand rule with the insert, whereby the index finger orients the first vector, the middle finger the second vector, and the thumb orients their vector product.

The expression (3.18d) leads to the identity

$$|\mathbf{P} \wedge \mathbf{Q}|^2 = |\mathbf{P}|^2 |\mathbf{Q}|^2 - (\mathbf{P} \cdot \mathbf{Q})^2 \quad (3.21a)$$

$$= |\mathbf{P}|^2 |\mathbf{Q}|^2 (1 - \cos^2 \theta), \quad (3.21b)$$

$$= |\mathbf{P}|^2 |\mathbf{Q}|^2 \sin^2 \theta, \quad (3.21c)$$

where we used the scalar product expression (3.10) to introduce the angle subtended by the two vectors. Trigonometry indicates that the area of the parallelogram defined by \mathbf{P} and \mathbf{Q} is given by $|\mathbf{P}| |\mathbf{Q}| \sin \theta$. Hence, the vector product has a magnitude given by this area

$$\text{area}(\mathbf{P}, \mathbf{Q}) = |\mathbf{P}| |\mathbf{Q}| \sin \theta = |\mathbf{P} \wedge \mathbf{Q}|. \quad (3.22)$$

Since $\mathbf{P} \wedge \mathbf{Q}$ is orthogonal to the plane defined by \mathbf{P} and \mathbf{Q} , we can write the vector product in the purely geometric manner

$$\mathbf{P} \wedge \mathbf{Q} = \hat{\mathbf{n}} \text{area}(\mathbf{P}, \mathbf{Q}) = \hat{\mathbf{n}} |\mathbf{P}| |\mathbf{Q}| \sin \theta, \quad (3.23)$$

where $\hat{\mathbf{n}}$ is a unit vector pointing normal to the area and in a direction given by the right hand rule. This formula is illustrated in Figure 3.3.

3.4.5 Generalization to arbitrary vectors

Thus far the discussion has considered vectors to represent the position of a point in space. As such, the vectors have the physical dimensions of length and $\text{area}(\mathbf{P}, \mathbf{Q})$ has dimensions of area. However, the vector analysis is general, so that the above notions extend to vectors of arbitrary physical dimensions, such as velocity. In these more general cases the physical dimensions must be adjusted accordingly.

3.5 Measuring volume

The vector product offers a means to measure area defined by two vectors. We now extend that result to measure the volume determined by three non-parallel vectors. This result has particular relevance to the volume element used for integration over space.

3.5.1 Volume defined by three vectors

Consider the scalar product of an arbitrary vector with the vector product, $(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}$. This scalar product projects that portion of the vector \mathbf{R} onto the direction parallel to the normal to the plane defined by $\mathbf{P} \wedge \mathbf{Q}$. Given that $|(\mathbf{P} \wedge \mathbf{Q})|$ is the area of the parallelogram defined by \mathbf{P} and \mathbf{Q} , we conclude that $(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}$ is the volume of the parallelepiped defined by the three vectors. However, note that this volume is not positive definite since the sign depends on the relative orientation of $\mathbf{P} \wedge \mathbf{Q}$ and \mathbf{R} . So more precisely, we need to apply an absolute value around the triple product to get the volume.

We can prove cyclic symmetry of $(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}$ through the following manipulations

$$(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R} = (P_a \vec{e}_a \wedge Q_b \vec{e}_b) \cdot R_d \vec{e}_d \quad (3.24a)$$

$$= P_a Q_b (\vec{e}_a \wedge \vec{e}_b) \cdot R_d \vec{e}_d \quad (3.24b)$$

$$= P_a Q_b (\epsilon_{abc} \vec{e}_c) \cdot \vec{e}_d R_d \quad (3.24c)$$

$$= P_a Q_b \epsilon_{abc} (\vec{e}_c \cdot \vec{e}_d) R_d \quad (3.24d)$$

$$= P_a Q_b \epsilon_{abc} \delta_{cd} R_d \quad (3.24e)$$

$$= P_a Q_b \epsilon_{abc} R_c \quad (3.24f)$$

$$= R_c P_a Q_b \epsilon_{abc} \quad (3.24g)$$

$$= R_a P_b Q_c \epsilon_{bca} \quad (3.24h)$$

$$= R_a P_b Q_c \epsilon_{abc} \quad (3.24i)$$

$$= (\mathbf{R} \wedge \mathbf{P}) \cdot \mathbf{Q}. \quad (3.24j)$$

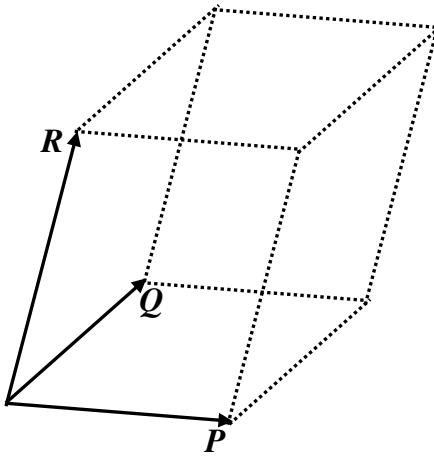


Figure 3.4: Three linearly independent position vectors determine a volume given by $|(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}| = |(\mathbf{R} \wedge \mathbf{P}) \cdot \mathbf{Q}| = |(\mathbf{Q} \wedge \mathbf{R}) \cdot \mathbf{P}|$

We thus have the geometric result illustrated in Figure 3.4

$$\text{volume}(\mathbf{P}, \mathbf{Q}, \mathbf{R}) = |(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}| = |(\mathbf{R} \wedge \mathbf{P}) \cdot \mathbf{Q}| = |(\mathbf{Q} \wedge \mathbf{R}) \cdot \mathbf{P}|. \quad (3.25)$$

3.5.2 Cartesian volume element for integration

We are in need the volume of an infinitesimal region when performing an integration over space. When making use of Cartesian coordinates we need the volume of a rectangular prism defined by infinitesimal distances along each of the axes. We thus set

$$\mathbf{P} = \hat{x} \, dx \quad \mathbf{Q} = \hat{y} \, dy \quad \mathbf{R} = \hat{z} \, dz, \quad (3.26)$$

in which case the volume element is

$$dV = (\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R} = dx \, dy \, dz \, (\hat{x} \wedge \hat{y}) \cdot \hat{z} = dx \, dy \, dz. \quad (3.27)$$

This expression for the volume element could have been written down without the formalism of a vector triple product. However, in Chapter 7 we find the general relation $(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}$ provides a useful starting point for deriving the volume element with arbitrary coordinates.

3.5.3 n -space volumes and the Levi-Civita tensor

We combine the geometric specification of the vector product as a means to measure area, (3.23), with the algebraic specification (3.18d) by writing

$$\text{2-volume} = \epsilon(\mathbf{P}, \mathbf{Q}) = \epsilon_{ab} P_a Q_b = \det \begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix}. \quad (3.28)$$

In this equation, ϵ_{ab} is the totally anti-symmetric 2×2 tensor. It has Cartesian components that can be organized as a matrix according to

$$\begin{bmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{21} & \epsilon_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (3.29)$$

In words, the first equality in equation (3.28) states that the ϵ -tensor in two dimensions takes two vectors as its argument and produces a 2-volume (i.e., an area). The three dimensional generalization yields

$$\text{3-volume} = \epsilon(\mathbf{P}, \mathbf{Q}, \mathbf{R}) = \epsilon_{abc} P_a Q_b R_c = \det \begin{bmatrix} P_1 & Q_1 & R_1 \\ P_2 & Q_2 & R_2 \\ P_3 & Q_3 & R_3 \end{bmatrix}. \quad (3.30)$$

Suppressing the first vector argument in the 3-volume produces a vectorial surface area defined by the other two vectors

$$\text{surface area} = \epsilon(\ , \mathbf{Q}, \mathbf{R}). \quad (3.31)$$

By construction, the vectorial surface area is orthogonal to both \mathbf{Q} and \mathbf{R} .

3.6 Example vector identities using the Levi-Civita tensor

The Levi-Civita tensor is a versatile tool for deriving vector identities. We illustrated some of these features in the previous discussion and here illustrate two more.

3.6.1 Double vector product

Consider the double vector product

$$\mathbf{P} \wedge (\mathbf{Q} \wedge \mathbf{R}) = P_a Q_b R_c \vec{e}_a \wedge (\vec{e}_b \wedge \vec{e}_c) \quad (3.32a)$$

$$= P_a Q_b R_c \vec{e}_a \wedge (\epsilon_{bcd} \vec{e}_d) \quad (3.32b)$$

$$= P_a Q_b R_c \epsilon_{bcd} \epsilon_{ade} \vec{e}_e \quad (3.32c)$$

$$= -P_a Q_b R_c \epsilon_{bcd} \epsilon_{aed} \vec{e}_e. \quad (3.32d)$$

Explicit substitution verifies that the product $\epsilon_{bcd} \epsilon_{aed}$ equals to

$$\epsilon_{bcd} \epsilon_{aed} = \delta_{ba} \delta_{ce} - \delta_{be} \delta_{ca}. \quad (3.33)$$

This identity then leads to

$$\epsilon_{bcd} \epsilon_{aed} \vec{e}_e = \delta_{ba} \vec{e}_c - \delta_{ca} \vec{e}_b \quad (3.34)$$

so that

$$\mathbf{P} \wedge (\mathbf{Q} \wedge \mathbf{R}) = -P_a Q_b R_c (\delta_{ba} \vec{e}_c - \delta_{ca} \vec{e}_b) \quad (3.35a)$$

$$= -(\mathbf{P} \cdot \mathbf{Q}) \mathbf{R} + (\mathbf{P} \cdot \mathbf{R}) \mathbf{Q}. \quad (3.35b)$$

3.6.2 Scalar product of two vector products

We make further use of the Levi-Civita identity (3.33) to write

$$(\mathbf{P} \wedge \mathbf{Q}) \cdot (\mathbf{R} \wedge \mathbf{S}) = (\epsilon_{abc} p_a q_b) (\epsilon_{dec} R_d s_e) \quad (3.36a)$$

$$= p_a q_b R_d s_e \epsilon_{abc} \epsilon_{dec} \quad (3.36b)$$

$$= p_a q_b R_d s_e (\delta_{ad} \delta_{be} - \delta_{ae} \delta_{bd}) \quad (3.36c)$$

$$= (\mathbf{P} \cdot \mathbf{R}) (\mathbf{Q} \cdot \mathbf{S}) - (\mathbf{P} \cdot \mathbf{S}) (\mathbf{Q} \cdot \mathbf{R}). \quad (3.36d)$$

3.7 Transforming the coordinate representations

The Cartesian basis vectors are mutually orthogonal and fixed in space. However, the orientation of the basis vectors is arbitrary. We thus consider an alternative specification to the basis vectors by performing a linear transformation

$$\vec{e}_{\bar{a}} = \mathcal{R}_{\bar{a}b} \vec{e}_b. \quad (3.37)$$

This expression introduced components to the transformation matrix moving between the unbarred and the barred Cartesian coordinates

$$\mathcal{R}_{\bar{a}b} = \begin{bmatrix} \mathcal{R}_{\bar{1}1} & \mathcal{R}_{\bar{1}2} & \mathcal{R}_{\bar{1}3} \\ \mathcal{R}_{\bar{2}1} & \mathcal{R}_{\bar{2}2} & \mathcal{R}_{\bar{2}3} \\ \mathcal{R}_{\bar{3}1} & \mathcal{R}_{\bar{3}2} & \mathcal{R}_{\bar{3}3} \end{bmatrix}. \quad (3.38)$$

In Cartesian tensor analysis the transformation is assumed to be independent of space.⁴ Although the transformation matrix carries two indices, it is not a tensor. Instead, it is a matrix operator used to transform from one set of basis vectors to another. We now deduce some constraints on this transformation matrix.

3.7.1 Inverse transformation

Assuming the transformation is invertible leads to the inverse transformation

$$\vec{e}_a = (\mathcal{R}^{-1})_{a\bar{b}} \vec{e}_{\bar{b}}. \quad (3.39)$$

As a self-consistency check we combine this relation with equation (3.37) thus rendering

$$\vec{e}_a = (\mathcal{R}^{-1})_{a\bar{b}} \vec{e}_{\bar{b}} = (\mathcal{R})_{a\bar{b}}^{-1} (\mathcal{R}_{\bar{b}c} \vec{e}_c). \quad (3.40)$$

This relation holds since

$$(\mathcal{R}^{-1})_{a\bar{b}} \mathcal{R}_{\bar{b}c} = \delta_{ac}, \quad (3.41)$$

or as a matrix identity

$$\mathcal{R}^{-1} \mathcal{R} = I. \quad (3.42)$$

3.7.2 Orthogonal transformation

We now assume that the two sets of Cartesian basis vectors are orthonormal. That assumption leads to the following constraint on the transformation matrix

$$\delta_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} \quad (3.43a)$$

$$= \mathcal{R}_{\bar{a}a} \vec{e}_a \cdot \mathcal{R}_{\bar{b}b} \vec{e}_b \quad (3.43b)$$

$$= \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}b} \vec{e}_a \cdot \vec{e}_b \quad (3.43c)$$

$$= \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}b} \delta_{ab} \quad (3.43d)$$

$$= \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}a} \quad (3.43e)$$

$$= \mathcal{R}_{\bar{a}a} (\mathcal{R}^T)_{a\bar{b}}, \quad (3.43f)$$

where \mathcal{R}^T is the matrix transpose with components

$$(\mathcal{R}^T)_{a\bar{b}} = \mathcal{R}_{\bar{b}a}. \quad (3.44)$$

⁴The transformation matrix is a function of space and time for the general tensors considered in Chapter 8.

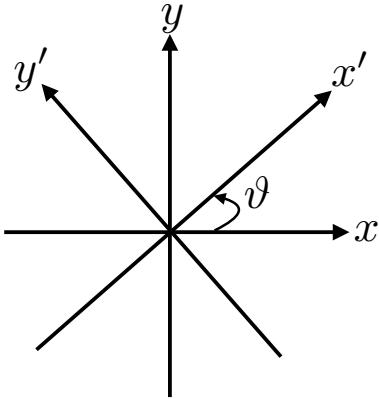


Figure 3.5: Counter-clockwise rotation of horizontal Cartesian axes through an angle ϑ .

Written as a matrix equation we see that

$$\mathcal{R} \mathcal{R}^T = I. \quad (3.45)$$

This relation defines an *orthogonal transformation*, whereby the inverse matrix equals to the matrix transpose

$$\mathcal{R}^{-1} = \mathcal{R}^T. \quad (3.46)$$

3.7.3 Geometric interpretation of orthogonal transformations

Orthogonal transformations convert one set of Cartesian coordinates to another. Geometrically, an orthogonal transformation corresponds to a rotation, with Figure 3.5 illustrating this axis rotation in two dimensions. For this two dimensional example, the rotation matrix can be written in terms of the cosine of the angles between the unit vectors; i.e., the *direction cosines*

$$\mathcal{R}_{\bar{a}a} = \begin{bmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{bmatrix} = \begin{bmatrix} \cos \vartheta & \cos(\pi/2 - \vartheta) \\ \cos(\pi/2 + \vartheta) & \cos \vartheta \end{bmatrix} = \begin{bmatrix} \vec{e}_1 \cdot \vec{e}_1 & \vec{e}_1 \cdot \vec{e}_2 \\ \vec{e}_2 \cdot \vec{e}_1 & \vec{e}_2 \cdot \vec{e}_2 \end{bmatrix}. \quad (3.47)$$

The final form of the rotation matrix reveals that it is built from the projection of the rotated basis vectors onto the original basis vectors. This result holds for rotations in three dimensions as well, thus leading to

$$\mathcal{R}_{\bar{a}a} = \begin{bmatrix} \vec{e}_1 \cdot \vec{e}_1 & \vec{e}_1 \cdot \vec{e}_2 & \vec{e}_1 \cdot \vec{e}_3 \\ \vec{e}_2 \cdot \vec{e}_1 & \vec{e}_2 \cdot \vec{e}_2 & \vec{e}_2 \cdot \vec{e}_3 \\ \vec{e}_3 \cdot \vec{e}_1 & \vec{e}_3 \cdot \vec{e}_2 & \vec{e}_3 \cdot \vec{e}_3 \end{bmatrix}. \quad (3.48)$$

In summary, Cartesian tensor analysis considers arbitrary Cartesian coordinates as related through a rotation matrix built from the *direction cosines*.

3.7.4 Transforming the coordinate representation of a vector

We introduced the transformation (3.39) according to how it acts on the basis vectors. Now consider how it acts on the coordinate representation of an arbitrary vector by moving brackets

$$\mathbf{P} = P_a \vec{e}_a \quad (3.49a)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} \vec{e}_{\bar{a}} \quad (3.49b)$$

$$\equiv P_{\bar{a}} \vec{e}_{\bar{a}}, \quad (3.49c)$$

where we defined the transformation of the vector components

$$P_{\bar{a}} = P_a (\mathcal{R}^T)_{a\bar{a}} = \mathcal{R}_{\bar{a}a} P_a. \quad (3.50)$$

3.7.5 Form invariance of the scalar product

The above properties of an orthogonal transformation ensure that the scalar product

$$\mathbf{P} \cdot \mathbf{Q} = P_a Q_a \quad (3.51)$$

is form invariant

$$\mathbf{P} \cdot \mathbf{Q} = P_a \vec{e}_a \cdot Q_b \vec{e}_b \quad (3.52a)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} Q_b (\mathcal{R}^T)_{\bar{b}\bar{b}} (\vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}}) \quad (3.52b)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} Q_b (\mathcal{R}^T)_{\bar{b}\bar{b}} \delta_{\bar{a}\bar{b}} \quad (3.52c)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} Q_b (\mathcal{R}^T)_{b\bar{a}} \quad (3.52d)$$

$$= P_{\bar{a}} Q_{\bar{a}}. \quad (3.52e)$$

We return to form invariance in Section 7.2, where it is referred to as *general covariance* in the context of general tensor analysis.

3.7.6 Transforming the coordinate representation of a second order tensor

The diffusion tensor is introduced in Chapter 37 and the stress tensor in Chapter 22. These tensors are second order, with second order tensors having a coordinate representation given by

$$\mathcal{T} = T_{ab} \vec{e}_a \vec{e}_b, \quad (3.53)$$

with T_{ab} the Cartesian representation of the second order tensor \mathcal{T} . Notably, there is no scalar product between the basis vectors. We determine how the components T_{ab} transform by following the above procedure for the basis vectors, only now with two basis vectors to carry around

$$\mathcal{T} = T_{ab} \vec{e}_a \vec{e}_b \quad (3.54a)$$

$$= T_{ab} (\mathcal{R}^T)_{a\bar{a}} \vec{e}_{\bar{a}} (\mathcal{R}^T)_{\bar{b}\bar{b}} \vec{e}_{\bar{b}} \quad (3.54b)$$

$$= T_{ab} (\mathcal{R}^T)_{a\bar{a}} (\mathcal{R}^T)_{\bar{b}\bar{b}} \vec{e}_{\bar{a}} \vec{e}_{\bar{b}} \quad (3.54c)$$

$$\equiv T_{\bar{a}\bar{b}} \vec{e}_{\bar{a}} \vec{e}_{\bar{b}}. \quad (3.54d)$$

The final equality introduced the transformed tensor components

$$T_{\bar{a}\bar{b}} = T_{ab} (\mathcal{R}^T)_{a\bar{a}} (\mathcal{R}^T)_{\bar{b}\bar{b}} = T_{ab} \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}b}. \quad (3.55)$$

The transformation of the components to higher order tensors follows analogously.

3.7.7 Importance of distinguishing between tensors and matrices

A matrix is an ordered array of objects. Hence, matrices are useful for organizing the coordinate components to a tensor. For example, the coordinate components to a first-order tensor (a vector) can be organized into a row or column matrix. Likewise, the coordinate components to a second-order tensor can be organized into a matrix. Consequently, the algebra of Cartesian tensors

shares much with the matrices familiar from linear algebra. However, tensors are not equivalent to matrices. The key distinction concerns how the components to tensors transform under changes to coordinates.

Namely, tensor components transform in a precise manner when modifying coordinates. In contrast, elements of an arbitrary matrix may or may not transform, with details depending on what the matrix elements represent. One means to help maintain focus on the distinction is to recall that a tensor is a geometric object (Section 3.1) that can be represented using arbitrary coordinates. Since the tensor has an existence independent of coordinates, its coordinate components are constrained to transform in a precise manner under changes to the coordinates.

3.8 Exercises

EXERCISE 3.1: PRODUCT OF SYMMETRIC MATRICES

Let $A = A^T$ and $B = B^T$ be two symmetric matrices. Under what condition is $AB = (AB)^T$?

EXERCISE 3.2: PRODUCT OF SYMMETRIC AND ANTI-SYMMETRIC TENSORS

Let $A = -A^T$ and $S = S^T$ be an anti-symmetric and a symmetric matrix, respectively. Show that the trace of their product vanishes: $\text{Tr}(AS) = 0$. Alternatively, in terms of tensors, show that the full contraction of an anti-symmetric tensor with a symmetric tensor vanishes: $A_{mn}S_{mn} = 0$.

4

Cartesian tensor calculus

READER'S GUIDE TO THIS CHAPTER

This chapter presents a synopsis of differential and integral calculus of use for our study of fluid mechanics. We build from the Cartesian tensor algebra of Chapter 3 to develop elements of Cartesian tensor calculus. Everything should be familiar to readers having taken undergraduate calculus. The material is used throughout this book so the reader is encouraged to master the basics here. For further study consult any book on calculus with analytic geometry. Particularly effective treatments, with applications to physics, are given in the following.

FEYNMAN LECTURES: Chapters 2 and 3 in Volume II of the [Feynman Lectures](#) offers insightful discussions of vector differential calculus. Although written for students of electrodynamics, many of the examples are drawn from fluid mechanics.

DIV, GRAD, CURL AND ALL THAT ([Schey, 2004](#)): This text pedagogically presents the methods and theorems of vector calculus in a manner that greatly assists the development of intuition.

THEORY AND PROBLEMS OF VECTOR ANALYSIS ([Spiegel, 1974](#)): This “Schaum’s Outline Series” book has nearly 500 worked exercises. Hence, it is a particularly useful resource for those wishing a refresher on problem solving in vector calculus. Some of the exercises here in Section 4.9 are drawn from this book.

4.1	The gradient of a scalar field	40
4.1.1	Direction of steepest ascent	41
4.1.2	Tangent to an isosurface	41
4.1.3	Normal to an isosurface	42
4.1.4	Unit vectors change only by rotation	42
4.1.5	Showing that $\delta\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 0$	43
4.2	The divergence of a vector field	43
4.2.1	Divergence of a scalar field times a vector field	44
4.2.2	Laplacian of a scalar field	44
4.3	The curl of a vector field	44
4.3.1	Computing the curl	44
4.3.2	Curl-free vector fields	45
4.3.3	Curl-free and divergence-free fields	46
4.3.4	Identities involving the curl	46
4.4	Path integral of a scalar function	48
4.5	Path integral of a vector function	49
4.5.1	Circulation	50
4.5.2	Circulation example	50
4.5.3	Fundamental theorem of calculus	50
4.6	Stokes' theorem	50
4.6.1	Statement of Stokes' theorem	51
4.6.2	Stokes' theorem for a rectangular region	51
4.7	Gauss's divergence theorem	52
4.7.1	An example rectangular volume	52
4.7.2	Divergence theorem for scalar fields	53
4.7.3	First and second form of Green's identities	53
4.7.4	Integral of a curl over a closed surface	54
4.8	Exact and inexact differentials	55
4.8.1	Exact differentials	55
4.8.2	Inexact differentials	55
4.8.3	Integrating factors	56
4.8.4	An example using the velocity field	56
4.8.5	Heuristic physics of exact and inexact differential operations	56
4.9	Exercises	57

4.1 The gradient of a scalar field

Consider a scalar field $\psi(\mathbf{x})$ defined on Euclidean space with position $\mathbf{x} = x_a \vec{e}_a$. For example, this field may be the temperature at a point, the mass density, or the specific entropy. We may estimate its value at an adjacent point $\mathbf{x} + d\mathbf{x}$ through a Taylor series

$$\psi(\mathbf{x} + d\mathbf{x}) = \psi(\mathbf{x}) + \frac{\partial\psi}{\partial x_1} dx_1 + \frac{\partial\psi}{\partial x_2} dx_2 + \frac{\partial\psi}{\partial x_3} dx_3 + \mathcal{O}(d\mathbf{x} \cdot d\mathbf{x}) \quad (4.1a)$$

$$\approx [1 + dx_a \partial_a] \psi(\mathbf{x}), \quad (4.1b)$$

where we dropped higher order terms to reach the final approximate expression, and introduced the shorthand notation for the partial derivative operator

$$\partial_a = \frac{\partial}{\partial x_a}. \quad (4.2)$$

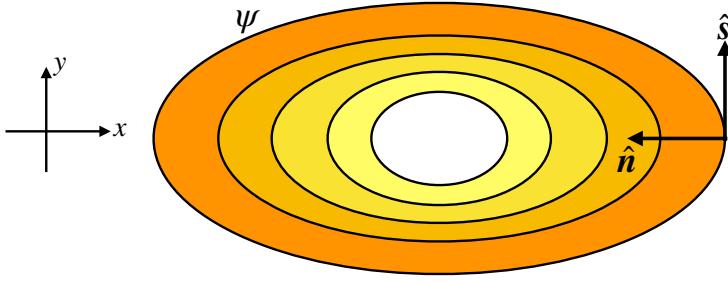


Figure 4.1: Contours of a scalar field $\psi(x, y)$, with values increasing toward the center. At any point in space, $\nabla\psi$ points in the direction of steepest increase (ascent) and orients the normal vector $\hat{\mathbf{n}} = |\nabla\psi|^{-1} \nabla\psi$. The unit tangent vector, $\hat{\mathbf{s}}$, points in a direction tangent to a ψ isosurface so that it follows the surface of constant ψ and it is orthogonal to the direction of steepest ascent: $\hat{\mathbf{n}} \cdot \hat{\mathbf{s}} = 0$.

We can introduce the gradient operator according to

$$\nabla = \vec{e}_a \partial_a = \hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} \partial_z \quad (4.3)$$

in which case

$$\psi(\mathbf{x} + d\mathbf{x}) \approx (1 + d\mathbf{x} \cdot \nabla) \psi(\mathbf{x}). \quad (4.4)$$

Note that in some treatments, ∇ is referred to as *Hamilton's operator*.

4.1.1 Direction of steepest ascent

Using the approximate relation (4.4), and the geometric expression (3.10) for the scalar product, renders

$$\psi(\mathbf{x} + d\mathbf{x}) - \psi(\mathbf{x}) \approx |d\mathbf{x}| |\nabla\psi| \cos \theta, \quad (4.5)$$

where θ is the angle between the differential increment $d\mathbf{x}$ and the gradient $\nabla\psi$. Orienting the increment $d\mathbf{x}$ so that $\theta = 0$ ensures that $\psi(\mathbf{x} + d\mathbf{x}) - \psi(\mathbf{x})$ is maximal. Consequently, $\nabla\psi$ points in the direction of *steepest ascent* across constant ψ isosurfaces (Figure 4.1). The opposite direction is that of *steepest descent*, where $\theta = \pi$.

4.1.2 Tangent to an isosurface

Consider a family of isosurfaces defined by points satisfying

$$\psi(\mathbf{x}) = \text{constant} \quad (4.6)$$

Figure 4.1 shows a two dimensional example where the isosurfaces are lines where ψ is a constant. As another example, consider $\psi(\mathbf{x}) = \psi(r)$, where $r^2 = \mathbf{x} \cdot \mathbf{x}$ is the squared radius of a sphere. Isosurfaces for this spherically symmetric function are spherical shells of radius r .

In general, moving along an isosurface keeps the scalar field unchanged. Let $\hat{\mathbf{s}}$ be a unit vector that points in the direction tangent to the isosurface at any point \mathbf{x} . By construction

$$\psi(\mathbf{x} + \hat{\mathbf{s}} ds) - \psi(\mathbf{x}) = 0, \quad (4.7)$$

where ds is an infinitesimal increment. In words, this identity says that if we move an infinitesimal distance in the direction tangent to the isosurface, then the function ψ does not change its value. Now expanding this identity in a Taylor series leads to

$$\hat{\mathbf{s}} \cdot \nabla\psi = \frac{\partial\psi}{\partial s} = 0. \quad (4.8)$$

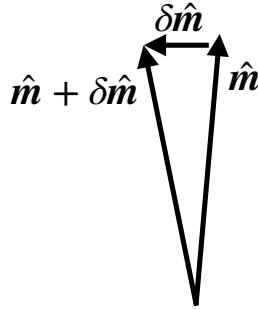


Figure 4.2: The infinitesimal change to a unit vector is itself perpendicular to the unit vector: $\delta\hat{\mathbf{m}} \cdot \hat{\mathbf{m}} = 0$. The reason is that the unit vector is constrained to remain unit length, so that the only way that it can change is to change its direction.

That is, isosurfaces of a function ψ are defined by directions along which the partial derivative of the function vanishes. For the spherically symmetric function, $\psi(\mathbf{x}) = \psi(r)$, the tangent vector points in either of the two angular directions along the spherical surface.

4.1.3 Normal to an isosurface

We may normalize the direction of maximal ascent, in which case we define the normal direction

$$\hat{\mathbf{n}} = |\nabla\psi|^{-1} \nabla\psi. \quad (4.9)$$

By construction, the gradient computed in the $\hat{\mathbf{n}}$ direction yields the maximum change for the function

$$\hat{\mathbf{n}} \cdot \nabla\psi = \frac{\partial\psi}{\partial n} = |\nabla\psi|. \quad (4.10)$$

For the spherically symmetric example,

$$\hat{\mathbf{n}} = \frac{\mathbf{x}}{|\mathbf{x}|} = \hat{\mathbf{r}}, \quad (4.11)$$

where $\hat{\mathbf{r}}$ is the unit vector pointing radially outward from the origin. In this case the normal derivative is equal to the radial derivative

$$\hat{\mathbf{n}} \cdot \nabla\psi = \frac{\partial\psi}{\partial r} \quad \text{spherically symmetric } \psi. \quad (4.12)$$

4.1.4 Unit vectors change only by rotation

Consider an arbitrary unit vector, $\hat{\mathbf{m}}$. The defining feature of a unit vector is that it has unit magnitude

$$\hat{\mathbf{m}} \cdot \hat{\mathbf{m}} = 1. \quad (4.13)$$

Unit vectors can only be modified through changes in their direction since their magnitude is fixed at unity. Hence, they are only modified by rotations. An important consequence of this constraint is that changes in unit vectors are perpendicular to the unit vector itself (see Figure 4.2). We see this property through considering an arbitrary change, symbolized by δ , in which

$$0 = \delta(1) = \delta(\hat{\mathbf{m}} \cdot \hat{\mathbf{m}}) = 2\hat{\mathbf{m}} \cdot \delta\hat{\mathbf{m}}. \quad (4.14)$$

In Section 13.5, we formally show that the constraint

$$\delta\hat{\mathbf{m}} \cdot \hat{\mathbf{m}} = 0 \quad (4.15)$$

means that unit vector changes can only arise from rotations, thus supporting the above assertion.

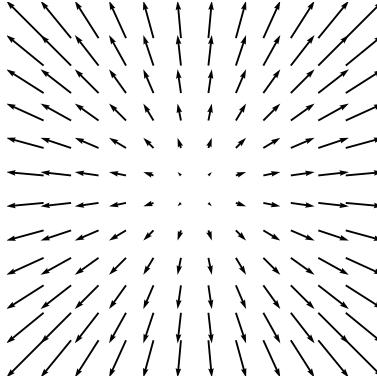


Figure 4.3: A vector field with a non-zero horizontal divergence. With $\mathbf{F} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$ the field diverges from the origin with a spatially constant divergence $\nabla \cdot \mathbf{F} = 2$. Note that this vector field has zero curl, $\nabla \wedge \mathbf{F} = 0$.

4.1.5 Showing that $\delta\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 0$

As an illustration of the constraint (4.15), let us verify that it holds for the special case of a unit normal vector (4.9) defined according to surfaces of constant scalar field

$$\hat{\mathbf{n}} = |\nabla\psi|^{-1} \nabla\psi. \quad (4.16)$$

The proof follows first by writing

$$\delta\hat{\mathbf{n}} = |\nabla\psi|^{-1} [\delta(\nabla\psi) - \hat{\mathbf{n}} \delta|\nabla\psi|], \quad (4.17)$$

so that

$$|\nabla\psi| \hat{\mathbf{n}} \cdot \delta\hat{\mathbf{n}} = \hat{\mathbf{n}} \cdot \delta(\nabla\psi) - \delta|\nabla\psi| \quad (4.18a)$$

$$= \frac{\nabla\psi \cdot \delta(\nabla\psi)}{|\nabla\psi|} - \delta|\nabla\psi| \quad (4.18b)$$

$$= \frac{|\nabla\psi| \delta|\nabla\psi|}{|\nabla\psi|} - \delta|\nabla\psi| \quad (4.18c)$$

$$= 0. \quad (4.18d)$$

The last step made use of the identity

$$\delta(|\nabla\psi|) = \delta(\sqrt{\nabla\psi \cdot \nabla\psi}) = \frac{1}{2\sqrt{\nabla\psi \cdot \nabla\psi}} \delta(\nabla\psi \cdot \nabla\psi) = \frac{\nabla\psi \cdot \delta(\nabla\psi)}{|\nabla\psi|}. \quad (4.19)$$

4.2 The divergence of a vector field

The divergence of a vector field, \mathbf{F} , is the scalar product of the divergence operator with the vector

$$\text{div}(\mathbf{F}) = \nabla \cdot \mathbf{F} = \partial_a F_a \begin{cases} > 0 \Rightarrow \text{diverging vector field}, \\ < 0 \Rightarrow \text{converging vector field}. \end{cases} \quad (4.20)$$

If the vector field in the surrounding neighborhood of a point is directed away from that point, then the vector field is diverging as if there is a source at the point (Figure 4.3). In this case the divergence of the vector field is positive. The converse occurs for a vector field converging to a point as if there is a sink. If the vector field under consideration is the velocity field of a fluid, then these considerations are directly related to the conservation of matter (see Chapter 17).

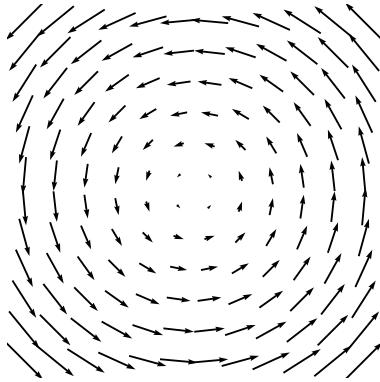


Figure 4.4: A horizontal vector field with a constant curl and zero divergence: $\mathbf{F} = -y \hat{x} + x \hat{y}$, $\implies \nabla \wedge \mathbf{F} = 2 \hat{z}$ and $\nabla \cdot \mathbf{F} = 0$.

4.2.1 Divergence of a scalar field times a vector field

We have many opportunities to make use of properties of the divergence operator following from application of the chain rule. Making use of the chain rule indicates that the divergence of a scalar field times a vector field is given by

$$\nabla \cdot (\phi \mathbf{F}) = \partial_a(\phi F_a) \quad (4.21a)$$

$$= \partial_a(\phi) F_a + \phi \partial_a F_a \quad (4.21b)$$

$$= \mathbf{F} \cdot \nabla \phi + \phi \nabla \cdot \mathbf{F}. \quad (4.21c)$$

4.2.2 Laplacian of a scalar field

The Laplacian of a scalar field is the divergence of the gradient

$$\nabla^2 \psi = \nabla \cdot \nabla \psi. \quad (4.22)$$

Scalar fields that have a vanishing Laplacian are said to be *harmonic*

$$\nabla^2 \psi = 0 \quad \text{harmonic function.} \quad (4.23)$$

Familiar examples of harmonic functions are the sines and cosines used for Fourier analysis in flat space, and the spherical harmonics used for Fourier analysis on the sphere. The name *harmonic* originates from the relation of harmonic functions to characteristic vibrational modes of a taut string such as those found on musical instruments (when played with skill). Furthermore, harmonic functions play a central role in the mathematical discipline of complex analysis.

4.3 The curl of a vector field

The curl characterizes how a vector field rotates at a point. In fluid mechanics we make much of use of the vorticity field, which is defined as the curl of the velocity field (Chapter 48).

4.3.1 Computing the curl

We measure the curl of a vector by computing the cross product of the divergence operator and the vector field. Hence, just like the cross product from Section 3.4, the curl is specified by both a

magnitude and a direction

$$\text{curl}(\mathbf{F}) = \nabla \wedge \mathbf{F} \quad (4.24a)$$

$$= \vec{e}_a \partial_a \wedge \vec{e}_b F_b \quad (4.24b)$$

$$= \vec{e}_a \wedge \vec{e}_b \partial_a F_b + \vec{e}_a F_b \wedge \partial_b \vec{e}_a \quad (4.24c)$$

$$= \epsilon_{abc} \vec{e}_c \partial_a F_b \quad (4.24d)$$

$$= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \hat{x} + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) \hat{y} + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \hat{z}. \quad (4.24e)$$

To reach this result we set $\partial_b \vec{e}_a = 0$ since the Cartesian basis vectors are fixed in space.¹ We also made use of the relation (3.11) for the cross product of basis vectors. We can express the curl from equation (4.24e) as a determinant

$$\nabla \wedge \mathbf{F} = \det \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ F_1 & F_2 & F_3 \end{bmatrix}. \quad (4.25)$$

The horizontal vector field $\mathbf{F} = x \hat{x} + y \hat{y}$ shown in Figure 4.3 has zero curl yet non-zero divergence. Figure 4.4 shows another vector field, $\mathbf{F} = -y \hat{x} + x \hat{y}$, with zero divergence yet nonzero curl $\nabla \wedge \mathbf{F} = 2 \hat{z}$. As seen in Section 46.5, this vector field corresponds to the velocity due to solid-body motion on a rotating planet, with its curl determining the planetary vorticity.

4.3.2 Curl-free vector fields

There are some cases of physically relevant vector fields that have a vanishing curl

$$\nabla \wedge \mathbf{F} = 0. \quad (4.26)$$

In fluid mechanics a curl-free velocity field has zero vorticity, which is a property maintained by linear gravity waves in the absence of rotation (Section 44.1). We illustrate a curl-free vector field in Figure 4.5, where the scalar potential is given by $\psi = \sin x \sin y$.

The curl of a gradient vanishes

$$\nabla \wedge \nabla \psi = \vec{e}_a \partial_a \wedge \vec{e}_b \partial_b \psi \quad (4.27a)$$

$$= (\vec{e}_a \wedge \vec{e}_b) \partial_a \partial_b \psi \quad (4.27b)$$

$$= 0, \quad (4.27c)$$

where the final equality follows since $\vec{e}_a \wedge \vec{e}_b$ is anti-symmetric on the labels ab whereas $\partial_a \partial_b$ is symmetric. This property allows us to introduce a scalar field ψ for curl-free vector fields so that

$$\mathbf{F} = -\nabla \psi \quad \text{scalar potential.} \quad (4.28)$$

The scalar ψ is known as the *scalar potential*. In the specific case of \mathbf{F} representing the gravitational force, then ψ is called the gravitational potential (see Section 14.1 and Chapter 34).

¹Basis vectors corresponding to non-Cartesian coordinates are spatially dependent (see Chapters 7 and 9), thus making this step invalid for general tensors. We will find a “fix” for this step in Section 9.9 by defining the *covariant curl operator*.

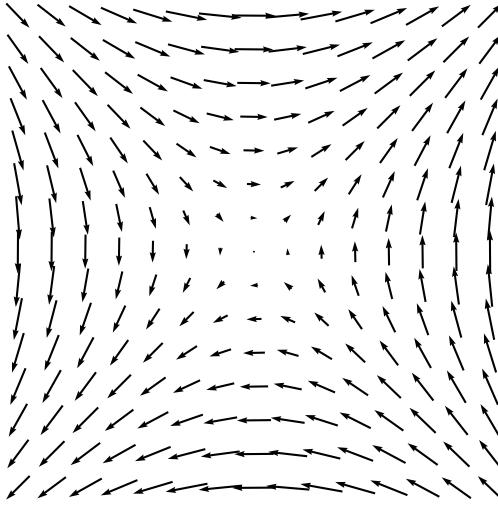


Figure 4.5: A horizontal vector field with a zero curl, where $\mathbf{F} = -\nabla\psi$ with $\psi = \sin(x/10) \sin(y/10)$.

4.3.3 Curl-free and divergence-free fields

Consider a vector field that has zero curl *and* zero divergence. The curl-free property means that

$$\nabla \wedge \mathbf{F} = 0 \Rightarrow \mathbf{F} = -\nabla\psi. \quad (4.29)$$

The divergence-free property means that ψ is a harmonic function (Section 4.2.2)

$$\nabla \cdot \nabla\psi = \nabla^2\psi = 0. \quad (4.30)$$

The velocity field arising from a linear non-rotating gravity wave (Section 44.1) in a Boussinesq fluid (Section 28.1) maintains zero vorticity and zero divergence. Furthermore, curl-free and divergence-free velocity fields are commonly encountered in aerodynamics.

4.3.4 Identities involving the curl

We close this section by deriving a suite of identities involving the curl operator. These identities are especially useful when developing dynamical equations for vorticity. Furthermore, by making use of the rules for general tensor analysis developed in Chapters 7, 8, and 9, these formula take on the same form regardless the coordinate choice.

Divergence of the curl vanishes

The divergence of the curl vanishes, as seen through the following

$$\nabla \cdot (\nabla \wedge \mathbf{F}) = \partial_a(\epsilon_{abc} \partial_b F_c) \quad (4.31a)$$

$$= \epsilon_{abc} \partial_a \partial_b F_c \quad (4.31b)$$

$$= 0. \quad (4.31c)$$

The final equality holds since $\partial_a \partial_b$ is symmetric on ab whereas ϵ_{abc} is anti-symmetric.

Divergence of a cross product

We now derive an expression for the divergence of a cross product

$$\nabla \cdot (\mathbf{F} \wedge \mathbf{E}) = \mathbf{E} \cdot (\nabla \wedge \mathbf{F}) - \mathbf{F} \cdot (\nabla \wedge \mathbf{E}) \quad (4.32)$$

through the following manipulations

$$\nabla \cdot (\mathbf{F} \wedge \mathbf{E}) = \vec{e}_a \partial_a (F_b \vec{e}_b \wedge E_c \vec{e}_c) \quad (4.33a)$$

$$= \vec{e}_a \cdot (\vec{e}_b \wedge \vec{e}_c) \partial_a (F_b E_c) \quad (4.33b)$$

$$= \epsilon_{abc} \partial_a (F_b E_c) \quad (4.33c)$$

$$= F_b \epsilon_{abc} \partial_a E_c + E_c \epsilon_{abc} \partial_a F_b \quad (4.33d)$$

$$= -\mathbf{F} \cdot (\nabla \wedge \mathbf{E}) + \mathbf{E} \cdot (\nabla \wedge \mathbf{F}). \quad (4.33e)$$

Curl of a scalar times a vector

We can compute the curl of a scalar field $\psi \mathbf{F}$ through the following steps

$$\nabla \wedge (\psi \mathbf{F}) = \vec{e}_a \partial_a \wedge \psi \vec{e}_b F_b \quad (4.34a)$$

$$= (\vec{e}_a \wedge \vec{e}_b) \partial_a (\psi F_b) \quad (4.34b)$$

$$= \epsilon_{abc} \vec{e}_c (\psi \partial_a F_b + F_b \partial_a \psi) \quad (4.34c)$$

$$= \psi \nabla \wedge \mathbf{F} + \nabla \psi \wedge \mathbf{F}. \quad (4.34d)$$

Curl of a cross product

The curl of a cross product of two vectors is given by

$$\nabla \wedge (\mathbf{F} \wedge \mathbf{E}) = \vec{e}_a \partial_a \wedge (\vec{e}_b F_b \wedge \vec{e}_c E_c) \quad (4.35a)$$

$$= \vec{e}_a \wedge (\vec{e}_b \wedge \vec{e}_c) \partial_a (F_b E_c) \quad (4.35b)$$

$$= \vec{e}_a \wedge (\epsilon_{bcd} \vec{e}_d) \partial_a (F_b E_c) \quad (4.35c)$$

$$= \epsilon_{ade} \epsilon_{bcd} \vec{e}_e \partial_a (F_b E_c) \quad (4.35d)$$

$$= -\epsilon_{aed} \epsilon_{bcd} \partial_a (F_b E_c) \quad (4.35e)$$

$$= -(\delta_{ab} \delta_{ec} - \delta_{ac} \delta_{eb}) \vec{e}_e \partial_a (F_b E_c) \quad (4.35f)$$

$$= (-\delta_{ba} \vec{e}_c + \delta_{ca} \vec{e}_b) \partial_a (F_b E_c) \quad (4.35g)$$

$$= \mathbf{F} (\nabla \cdot \mathbf{E}) + (\mathbf{E} \cdot \nabla) \mathbf{F} - \mathbf{F} (\nabla \cdot \mathbf{F}) - (\mathbf{F} \cdot \nabla) \mathbf{E}. \quad (4.35h)$$

Curl of a curl

A special case of the identity (4.35h) allows us to write the curl of a curl as

$$\nabla \wedge (\nabla \wedge \mathbf{F}) = \vec{e}_a \wedge (\vec{e}_b \wedge \vec{e}_c) \partial_a \partial_b F_c \quad (4.36a)$$

$$= (-\delta_{ba} \vec{e}_c + \delta_{ca} \vec{e}_b) \partial_a \partial_b F_c \quad (4.36b)$$

$$= \nabla (\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}. \quad (4.36c)$$

Relating advection, curl, and kinetic energy

We now apply some of the previous manipulations to derive a relation required to derive the vorticity equation (Section 48.4.1). Here, we aim to show that

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = \boldsymbol{\omega} \wedge \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2, \quad (4.37)$$

where

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (4.38)$$

is the vorticity, $\mathbf{v} \cdot \mathbf{v}/2$ is the kinetic energy per mass, and \mathbf{v} is the fluid velocity field. We here show all the steps along with their justification

$$\begin{aligned} \boldsymbol{\omega} \wedge \mathbf{v} &= (\nabla \wedge \mathbf{v}) \wedge \mathbf{v} && \text{insert } \boldsymbol{\omega} = \nabla \wedge \mathbf{v} && (4.39a) \\ &= (\vec{e}_a \partial_a \wedge \vec{e}_b v_b) \wedge \vec{e}_c v_c && \text{Cartesian representation of } \mathbf{v} \text{ and } \nabla && (4.39b) \\ &= (\vec{e}_a \wedge \vec{e}_b) \wedge \vec{e}_c (\partial_a v_b) v_c && \text{rearrange} && (4.39c) \\ &= \epsilon_{abd} (\vec{e}_d \wedge \vec{e}_c) (\partial_a v_b) v_c && \text{first cross product expanded} && (4.39d) \\ &= \epsilon_{abd} \epsilon_{dce} \vec{e}_e (\partial_a v_b) v_c && \text{second cross product expanded} && (4.39e) \\ &= \epsilon_{abd} \epsilon_{ced} \vec{e}_e (\partial_a v_b) v_c && \text{arrange indices to prepare for next step} && (4.39f) \\ &= (\delta_{ac} \delta_{be} - \delta_{ae} \delta_{bc}) \vec{e}_e (\partial_a v_b) v_c && \text{use identity (3.33)} && (4.39g) \\ &= \vec{e}_a v_c \partial_c v_a - \vec{e}_a v_c \partial_a v_c && \text{contract the Kronecker deltas} && (4.39h) \\ &= \vec{e}_a [(\mathbf{v} \cdot \nabla) v_a - \partial_a \mathbf{v}^2/2] && \text{re-express as Cartesian tensor} && (4.39i) \\ &= (\mathbf{v} \cdot \nabla) \mathbf{v} - \nabla[\mathbf{v} \cdot \mathbf{v}/2] && \text{rearrange.} && (4.39j) \end{aligned}$$

Note that Section 4.4.4 of [Griffies \(2004\)](#) exhibits these steps making use of a general coordinate framework.

4.4 Path integral of a scalar function

Consider the integral of a scalar function, ψ , over an arbitrary one-dimensional path in space, C

$$\mathcal{I} = \int_C \psi(\varphi) d\varphi. \quad (4.40)$$

A path is a one-dimensional curve, so that a point along the path can be specified by a single parameter, denoted here by φ . Now lay down a Cartesian coordinate system with an arbitrary origin. The corresponding Cartesian coordinate representation of a point along the path is written

$$C \mapsto \mathbf{x}(\varphi) = \hat{\mathbf{x}} x(\varphi) + \hat{\mathbf{y}} y(\varphi) + \hat{\mathbf{z}} z(\varphi) \quad (4.41)$$

so that the path integral can be written

$$\mathcal{I} = \int_C \psi(\varphi) d\varphi = \int_C \psi[\mathbf{x}(\varphi)] d\varphi. \quad (4.42)$$

Arc length to parameterize the path

A particularly common special case for path parameterization is where we choose φ to be the arc length along the path. For Euclidean space using Cartesian coordinates, the differential increment of arc length is given by

$$ds = \sqrt{d\mathbf{x} \cdot d\mathbf{x}} = ds \sqrt{\frac{d\mathbf{x}}{ds} \cdot \frac{d\mathbf{x}}{ds}} = ds |\mathbf{x}'(s)|, \quad (4.43)$$

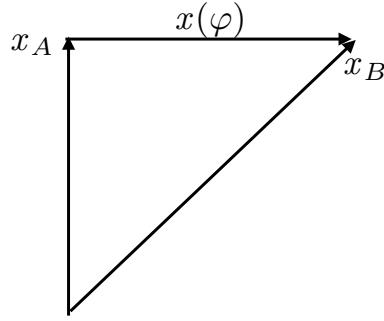


Figure 4.6: A linear path, $\mathbf{x}(\varphi)$ extending from \mathbf{x}_A to \mathbf{x}_B that is parameterized by a non-dimensional parameter $\varphi \in [0, 1]$ via $\mathbf{x}(\varphi) = \mathbf{x}_A + (\mathbf{x}_B - \mathbf{x}_A)\varphi$. Alternatively it can be parameterized by the arc-length along the path via $\mathbf{x}(s) = \mathbf{x}_A + \hat{\mathbf{s}} s$ with $s \in [0, L]$.

so that the path integral takes the form

$$\mathcal{I} = \int_C \psi[\mathbf{x}(s)] |\mathbf{x}'(s)| ds. \quad (4.44)$$

Linear path example

As a specific example, consider a line between two points, \mathbf{x}_A and \mathbf{x}_B , as in Figure 4.6. We can parameterize the line using a dimensionless parameter φ according to

$$\mathbf{x}(\varphi) = \mathbf{x}_A + (\mathbf{x}_B - \mathbf{x}_A) \varphi \quad \varphi \in [0, 1]. \quad (4.45)$$

Alternatively, we can parameterize using the arc length

$$\mathbf{x}(s) = \mathbf{x}_A + \hat{\mathbf{s}} s \quad s \in [0, L], \quad (4.46)$$

where $L = \int_A^B ds = |\mathbf{x}_B - \mathbf{x}_A|$ is the total arc length of the line, and where $\hat{\mathbf{s}}$ is the unit tangent vector pointing along the path from \mathbf{x}_A to \mathbf{x}_B

$$\hat{\mathbf{s}} = \frac{\mathbf{x}'(s)}{|\mathbf{x}'(s)|} = \frac{\mathbf{x}_B - \mathbf{x}_A}{|\mathbf{x}_B - \mathbf{x}_A|}. \quad (4.47)$$

As defined we have $|\mathbf{x}'(s)| = |\hat{\mathbf{s}}| = 1$, so that the path integral is given by $\mathcal{I} = \int_0^L \psi[\mathbf{x}(s)] ds$.

4.5 Path integral of a vector function

Generalizing to a vector field, $\mathbf{F}(\mathbf{x})$, we could conceivably integrate each component of the vector along the curve independently using the scalar result we just found. In practice, however, this quantity rarely appears in physics. Instead, we more commonly wish to integrate that component of $\mathbf{F}(\mathbf{x})$ that projects onto the curve

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C \mathbf{F} \cdot \frac{d\mathbf{x}}{ds} ds \quad (4.48)$$

where $d\mathbf{x}/ds$ is tangent to the curve.

4.5.1 Circulation

For the case of a closed curve, we refer to the above path integral as the *circulation* and use the convention of putting an arrowed circle on the integral sign

$$\text{circulation of vector field} = \oint_C \mathbf{F} \cdot d\mathbf{x}. \quad (4.49)$$

The arrow indicates that we conventionally traverse the closed path in a counter-clockwise (right hand) manner.

4.5.2 Circulation example

Consider the vector field $\mathbf{F} = -y \hat{\mathbf{x}} + x \hat{\mathbf{y}}$ shown in Figure 4.4. What is the circulation for this field computed around a circle of radius r whose center is the origin? To compute this circulation we make use of plane polar coordinates, in which $x = r \cos \varphi$ and $y = r \sin \varphi$, with $\varphi \in [0, 2\pi]$ the polar angle measured from the positive x -axis. The position of a point on the circle is thus written $\mathbf{x}(\varphi) = r(\hat{\mathbf{x}} \cos \varphi + \hat{\mathbf{y}} \sin \varphi)$, and the tangent to the circle is $d\mathbf{x}(\varphi)/d\varphi = r(-\hat{\mathbf{x}} \sin \varphi + \hat{\mathbf{y}} \cos \varphi)$. The integrand to the circulation (4.48) thus takes the form

$$\mathbf{F} \cdot \frac{d\mathbf{x}(\varphi)}{d\varphi} = r(y \sin \varphi + x \cos \varphi) = r^2. \quad (4.50)$$

Hence, the circulation around the constant radius circle is given by twice the area of the circle

$$\oint_C \mathbf{F} \cdot d\mathbf{x} = 2\pi r^2. \quad (4.51)$$

This result has application for geophysical fluids when computing the vorticity induced by the rotating planet (see Section 48.6.2).

4.5.3 Fundamental theorem of calculus

The special case of $\mathbf{F} = -\nabla\psi$ for a scalar field ψ recovers the fundamental theorem of calculus

$$\psi(\mathbf{x}_A) - \psi(\mathbf{x}_B) = \int_{\mathbf{x}_A}^{\mathbf{x}_B} d\psi = \int_{\mathbf{x}_A}^{\mathbf{x}_B} \nabla\psi \cdot d\mathbf{x}. \quad (4.52)$$

It follows that for any closed curve with $\mathbf{x}_A = \mathbf{x}_B$, the circulation of $\nabla\psi$ vanishes

$$\oint_C d\psi = \oint_C \nabla\psi \cdot d\mathbf{x} = 0. \quad (4.53)$$

4.6 Stokes' theorem

Stokes' theorem relates a vector field's behavior on the boundary of a surface to its behavior within the area of the surface. It is used extensively in our study of circulation and vorticity (Chapter 48).

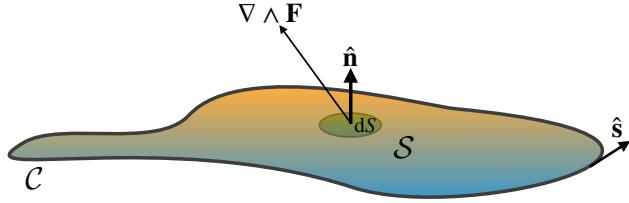


Figure 4.7: Illustrating the geometry of Stokes' Theorem. The unit normal $\hat{\mathbf{n}}$ points outward from the surface, \mathcal{S} , with $(\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}}$ the projection of the curl of a vector field onto the surface normal. The outer boundary of the area, $\partial\mathcal{S}$, is traversed counterclockwise following a tangent vector $\hat{\mathbf{s}}$ when computing the circulation.

4.6.1 Statement of Stokes' theorem

Given an oriented surface \mathcal{S} with a boundary $\partial\mathcal{S}$, Stokes' Theorem says that the circulation around the boundary equals to the area integrated curl projected onto the surface outward normal

$$\oint_{\partial\mathcal{S}} \mathbf{F} \cdot d\mathbf{x} = \int_{\mathcal{S}} (\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (4.54)$$

In this equation, $d\mathbf{x}$ is the vector line element along the path, $\partial\mathcal{S}$ is the closed path defining the boundary to an oriented two-dimensional surface \mathcal{S} , $\hat{\mathbf{n}}$ is the outward normal vector on the surface, and $d\mathcal{S}$ is the infinitesimal surface area element. The orientation of the outward normal is determined by the right hand rule according to the circulation direction.

4.6.2 Stokes' theorem for a rectangular region

To build experience with Stokes' theorem, consider the case of a rectangle in the x-y plane with dimensions $L_x \times L_y$. In this case $\hat{\mathbf{n}} = \hat{\mathbf{z}}$, so that

$$(\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{z}} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}, \quad (4.55)$$

in which case the right hand side of Stokes' theorem reduces to

$$\int_{\mathcal{S}} (\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_0^{L_x} \int_0^{L_y} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy. \quad (4.56)$$

Integration around the rectangle then leads to a direct verification of Stokes' Theorem

$$\int_0^{L_x} \int_0^{L_y} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy = \int_0^{L_y} F_2(x, y) \Big|_{x=0}^{x=L_x} dy - \int_0^{L_x} F_1(x, y) \Big|_{y=0}^{y=L_y} dx \quad (4.57a)$$

$$= \int_0^{L_x} F_1(x, 0) dx + \int_0^{L_y} F_2(L_x, y) dy + \int_{L_x}^0 F_1(x, L_y) dx + \int_{L_y}^0 F_2(0, y) dy \quad (4.57b)$$

$$= \oint_{\partial\mathcal{S}} \mathbf{F} \cdot d\mathbf{x}. \quad (4.57c)$$

We can generalize this result to verify Stokes' Theorem for an arbitrary surface. We do so by breaking the surface into a lattice of tiny rectangles. Integrating around the tiny rectangles and summing their contributions leads to a cancellation of the line integrals over all interior boundaries. The cancellation occurs since an internal edge of a rectangle is integrated once in each direction thus cancelling its contribution. The only nonzero contribution comes from integration over the external boundary.

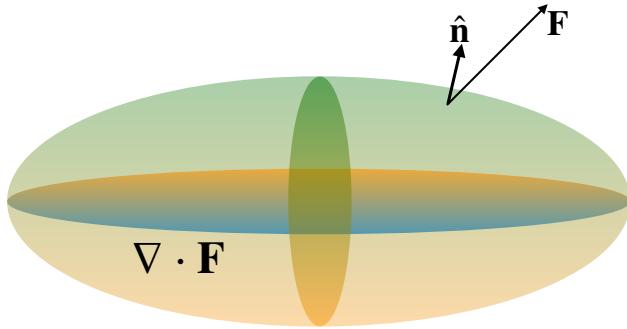


Figure 4.8: Illustrating the geometry of Gauss's divergence theorem for an ellipsoidal volume. The outward unit normal $\hat{\mathbf{n}}$ is projected onto the vector field \mathbf{F} via the scalar product, $\mathbf{F} \cdot \hat{\mathbf{n}}$. Integrating this scalar product over the closed surface \mathcal{S} yields the same result as computing the volume integral of the divergence, $\nabla \cdot \mathbf{F}$, over the region, \mathcal{R} , bounded by the closed surface.

4.7 Gauss's divergence theorem

Gauss's divergence theorem, also known as Gauss's law in electrodynamics, relates the divergence of a vector field integrated over a volume to the area integrated normal projection of the vector field through the closed surface bounding the volume. For a vector field \mathbf{F} , the divergence theorem states that

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{F} dV = \oint_{\partial\mathcal{R}} \mathbf{F} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (4.58)$$

where $\hat{\mathbf{n}}$ is the outward normal to the boundary surface and $d\mathcal{S}$ is the surface area element. We follow the convention that $\oint_{\partial\mathcal{R}}$ refers to a surface integral over a closed surface that bounds a volume. This notation is used to contrast with the surface integral $\int_{\mathcal{S}}$ that generally does not enclose a volume. Figure 4.8 illustrates the geometry of Gauss's divergence theorem. In fluid mechanics jargon, we say that the divergence of a vector field, $\nabla \cdot \mathbf{F}$, integrated over a volume equals to the flux of that vector field, $\mathbf{F} \cdot \hat{\mathbf{n}}$, integrated over the area bounding the volume.

4.7.1 An example rectangular volume

To build intuition for Gauss's divergence theorem, consider a rectangular volume with dimensions $L_x \times L_y \times L_z$. The volume integral on the left hand side of equation (4.58) gives

$$\int_{\mathcal{R}} \left[\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \right] dx dy dz. \quad (4.59)$$

Focusing on just the leftmost term, integration in x gives

$$\int_{\mathcal{R}} \frac{\partial F_1}{\partial x} dx dy dz = \int_{y=0}^{y=L_y} \int_{z=0}^{z=L_z} [F_1(L_x, y, z) - F_1(0, y, z)] dy dz \quad (4.60a)$$

$$= \int_{S_1+S_2} \mathbf{F} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (4.60b)$$

where S_1 is the rectangle's face with outward normal $\hat{\mathbf{n}} = \hat{\mathbf{x}}$ and S_2 is the rectangle's face with normal $\hat{\mathbf{n}} = -\hat{\mathbf{x}}$. Repeating this procedure on the other terms in equation (4.59) gives the area integrated flux (the *transport* using fluid mechanics jargon) through the full boundary. To verify the theorem for a general volume V , we take the approach used to prove Stokes' theorem. First, divide the volume into many rectangular sub-volumes. Then apply the above result to each sub-volume

and sum up the result. The fluxes through internal rectangular faces cancel to zero. Therefore, the sum of all the fluxes equals just the flux through the external boundary, yielding the divergence theorem.

4.7.2 Divergence theorem for scalar fields

We consider two corollaries of the divergence theorem, the first of which arises from the special case of a vector field $\mathbf{F} = \phi \mathbf{c}$ with \mathbf{c} an arbitrary *constant* vector. Substitution into the divergence theorem (4.58) yields

$$\oint_{\partial\mathcal{R}} \phi \mathbf{c} \cdot \hat{\mathbf{n}} dS = \int_{\mathcal{R}} \nabla \cdot (\phi \mathbf{c}) dV. \quad (4.61)$$

Pulling the constant vector out of the integrals and rearrangement leads to

$$\mathbf{c} \cdot \left[\oint_{\partial\mathcal{R}} \phi \hat{\mathbf{n}} dS - \int_{\mathcal{R}} \nabla \phi dV \right] = 0. \quad (4.62)$$

Since \mathbf{c} is an arbitrary vector, this equality is true in general only when

$$\oint_{\partial\mathcal{R}} \phi \hat{\mathbf{n}} dS = \int_{\mathcal{R}} \nabla \phi dV. \quad (4.63)$$

In words, this result says that the integral of a scalar field over a closed surface, when weighted by the outward normal to the surface, equals to the volume integral of the gradient. We make use of this result in Section 22.1.3 when formulating the contribution of stresses to the motion of a fluid element, in particular when considering how pressure affects motion.

4.7.3 First and second form of Green's identities

The second corollary to the divergence theorem arises from considering another special vector field

$$\mathbf{F} = \psi \nabla \phi, \quad (4.64)$$

with ψ and ϕ scalar fields. Substitution into the divergence theorem (4.58) leads to

$$\oint_{\partial\mathcal{R}} \psi \frac{\partial \phi}{\partial n} dS = \int_{\mathcal{R}} [\nabla \psi \cdot \nabla \phi + \psi \nabla^2 \phi] dV \quad \text{Green's first integral identity.} \quad (4.65)$$

We can make this result more symmetric by swapping ψ and ϕ and then subtracting, thus yielding

$$\oint_{\partial\mathcal{R}} \left[\psi \frac{\partial \phi}{\partial n} - \phi \frac{\partial \psi}{\partial n} \right] dS = \int_{\mathcal{R}} [\psi \nabla^2 \phi - \phi \nabla^2 \psi] dV \quad \text{Green's second integral identity.} \quad (4.66)$$

Finally, setting $\phi = 1$ renders

$$\oint_{\partial\mathcal{R}} \frac{\partial \psi}{\partial n} dS = \int_{\mathcal{R}} \nabla^2 \psi dV \iff \oint_{\partial\mathcal{R}} \nabla \psi \cdot \hat{\mathbf{n}} dS = \int_{\mathcal{R}} \nabla \cdot \nabla \psi dV. \quad (4.67)$$

As seen in Sections 5.7 and 41.6, these identities are fundamental to the Green's function method for solving linear partial differential equations.

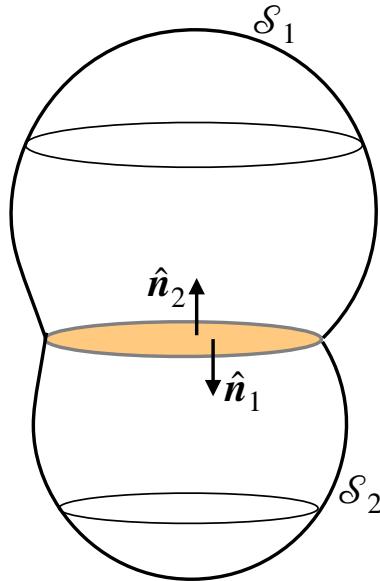


Figure 4.9: The integral of the curl of a function vanishes when integrated over an oriented closed surface. This result follows from both Gauss's divergence theorem as well as Stokes' curl theorem. Here we depict a closed volume that has been split into two surfaces \mathcal{S}_1 and \mathcal{S}_2 . Separately applying Stokes' theorem to the surfaces leads to the calculation of the circulation around the common boundary. Yet the orientation of the outward normal along the closed surface is opposite for the two regions, thus leading to a cancellation of the circulations.

4.7.4 Integral of a curl over a closed surface

Application of Gauss's divergence theorem leads us to conclude that the following integral vanishes

$$\oint_{\partial\mathcal{R}} (\nabla \wedge \Psi) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \int_{\mathcal{R}} \nabla \cdot (\nabla \wedge \Psi) \, dV = 0, \quad (4.68)$$

where the final equality follows since the divergence of a curl vanishes. Hence, the integral of the curl of a function over an oriented closed surface vanishes. We can understand this result geometrically by splitting the closed volume into two regions and then applying Stokes' theorem separately to the two regions (see Figure 4.9)

$$\oint_{\partial\mathcal{R}} (\nabla \wedge \Psi) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \int_{\mathcal{S}_1} (\nabla \wedge \Psi) \cdot \hat{\mathbf{n}} \, d\mathcal{S}_1 + \int_{\mathcal{S}_2} (\nabla \wedge \Psi) \cdot \hat{\mathbf{n}} \, d\mathcal{S}_2 \quad (4.69a)$$

$$= \oint_{\partial\mathcal{S}_1} \Psi \cdot d\mathbf{x} - \oint_{\partial\mathcal{S}_2} \Psi \cdot d\mathbf{x} \quad (4.69b)$$

$$= 0. \quad (4.69c)$$

The minus sign appearing in front of $\oint_{\partial\mathcal{S}_2}$ occurs since the orientation of the circulation integral is opposite that for $\oint_{\partial\mathcal{S}_1}$. We are thus left with a cancellation of the circulations. When applied to the vorticity of a fluid (Chapter 46), we see that

$$\oint_{\partial\mathcal{R}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \int_{\mathcal{R}} \nabla \cdot \boldsymbol{\omega} \, dV = 0, \quad (4.70)$$

where $\boldsymbol{\omega} = \nabla \wedge \mathbf{v}$ is the vorticity and \mathbf{v} is the fluid velocity.

4.8 Exact and inexact differentials

Thus far in this chapter all differentials have been exact. However, the thermodynamics discussed in Chapter 23 makes use of both exact and inexact differentials. We here introduce the mathematics of such differentials.

4.8.1 Exact differentials

As in Section 4.1, consider an arbitrary function of space, $\Phi(\mathbf{x})$. A differential increment for that function, computed between two close points \mathbf{x} and $\mathbf{x} + d\mathbf{x}$, is given by

$$d\Phi(\mathbf{x}) = \Phi(\mathbf{x} + d\mathbf{x}) - \Phi(\mathbf{x}) \quad (4.71a)$$

$$= d\mathbf{x} \cdot \nabla \Phi, \quad (4.71b)$$

where we dropped higher order terms due to the infinitesimal nature of the increments. It follows that we can determine the finite increment between two points through integration

$$\Phi(\mathbf{x}_B) - \Phi(\mathbf{x}_A) = \int_{\mathbf{x}_A}^{\mathbf{x}_B} d\Phi(\mathbf{x}) = \int_{\mathbf{x}_A}^{\mathbf{x}_B} d\mathbf{x} \cdot \nabla \Phi. \quad (4.72)$$

These results are familiar from elementary calculus, with the increment $d\Phi$ given by equation (4.71b) termed an *exact* differential. Importantly, the finite increment, $\Phi(\mathbf{x}_B) - \Phi(\mathbf{x}_A)$, depends only on the endpoint values of Φ . It does not depend on the path taken to go from \mathbf{x}_A to \mathbf{x}_B . Correspondingly, the integral of an exact differential around a closed loop vanishes

$$\oint d\Phi = 0. \quad (4.73)$$

4.8.2 Inexact differentials

Consider now a general differential expression written as

$$\mathbf{A} \cdot d\mathbf{x} = A dx + B dy + C dz, \quad (4.74)$$

where $\mathbf{A} = A \hat{\mathbf{x}} + B \hat{\mathbf{y}} + C \hat{\mathbf{z}}$ is an arbitrary vector function. If $\nabla \wedge \mathbf{A} = 0$, then \mathbf{A} can be written as the gradient of a scalar

$$\nabla \wedge \mathbf{A} = 0 \implies \mathbf{A} = \nabla \Psi, \quad (4.75)$$

in which case we have an exact differential expression

$$\mathbf{A} \cdot d\mathbf{x} = \nabla \Psi \cdot d\mathbf{x} = d\Psi. \quad (4.76)$$

That is, the differential $\mathbf{A} \cdot d\mathbf{x}$ is exact if

$$\nabla \wedge \mathbf{A} = 0 \implies \mathbf{A} \cdot d\mathbf{x} \text{ exact differential.} \quad (4.77)$$

In the more general case where $\nabla \wedge \mathbf{A} \neq 0$, then $\mathbf{A} \cdot d\mathbf{x}$ is an inexact differential. We make use of the following notation for inexact differentials

$$d\Psi = \mathbf{A} \cdot d\mathbf{x}. \quad (4.78)$$

Notably, the path integral of an inexact differential depends on the path taken between the endpoints. Correspondingly, the integral of an inexact differential around a closed loop does not generally vanish

$$\oint d\Psi \neq 0. \quad (4.79)$$

4.8.3 Integrating factors

Consider again the inexact differential $d\Psi = \mathbf{A} \cdot d\mathbf{x}$. Let us presume there exists a function τ so that the product $\tau^{-1} d\Psi$ is an exact differential. For τ to exist it must be such that

$$\nabla \wedge (\mathbf{A} \tau^{-1}) = 0. \quad (4.80)$$

Consequently, we can write

$$\mathbf{A} = \tau \nabla \Phi, \quad (4.81)$$

so that

$$d\Psi = \mathbf{A} \cdot d\mathbf{x} \quad (4.82a)$$

$$= \tau \nabla \Phi \cdot d\mathbf{x} \quad (4.82b)$$

$$= \tau d\Phi. \quad (4.82c)$$

The function τ is known as an *integrating factor*. As seen in Section 23.2.3, pressure is the integrating factor for mechanical work, temperature is the integrating factor for heat, and the chemical potential is the integrating factor for chemical work.

4.8.4 An example using the velocity field

Consider the product $\mathbf{v} \cdot d\mathbf{x}$, where \mathbf{v} is the velocity field for a fluid and $d\mathbf{x}$ is a differential increment in space directed along a path. Furthermore, introduce the curl of the velocity, which defines the vorticity (Section 46.1) $\boldsymbol{\omega} = \nabla \wedge \mathbf{v}$. For cases where the vorticity vanishes, $\boldsymbol{\omega} = 0$, then $d\Psi = \mathbf{v} \cdot d\mathbf{x}$ is an exact differential. Consequently, Stokes' theorem means that the circulation² around an arbitrary closed loop vanishes (Section 46.3)

$$C \equiv \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0. \quad (4.83)$$

Another way to see this result is to note that a vanishing curl means that the velocity field can be expressed as the gradient of a scalar, $\mathbf{v} = \nabla \psi$, so that $d\Psi = \nabla \psi \cdot d\mathbf{x}$, which is manifestly exact.

4.8.5 Heuristic physics of exact and inexact differential operations

Consider a hiker climbing a mountain. The mechanical work, which is force applied over a distance, is a function of the path taken. Some paths are smooth and well marked, whereas others are rough and poorly marked. Likewise, the frictional heating (of the hiker's feet, for example) depend on details of the path (and the shoes!). So although the start and finish points are fixed, the work exerted and heat generated in going between these points is a function of the path.

In contrast, the change in gravitational potential energy between the start and finish points is a function only of the relative elevation; it does not depend on the path between the points. So the gravitational potential energy increment between the two points is an exact differential, with the potential energy for each point a function of the elevation at the point. Analogously, the First Law of thermodynamics says that the sum of path-dependent processes (work and heat) used in going from one thermodynamic state to another equals to the difference in the internal energy between the two states. That is, the sum of the inexact differentials for heat and work equal to the exact differential for internal energy.

²The circulation around a closed loop is distinct from the transport of fluid moving in a direction normal to a line, with the transport introduced in Section 19.4.3.

4.9 Exercises

Throughout these exercises we consider a point whose Cartesian position vector, relative to an origin, is written

$$\mathbf{x} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + z \hat{\mathbf{z}} \quad (4.84)$$

and whose squared distance from the origin is

$$r^2 = \mathbf{x} \cdot \mathbf{x} = x^2 + y^2 + z^2. \quad (4.85)$$

Likewise, we consider closed surfaces \mathcal{S} and their enclosed volume V .

EXERCISE 4.1: PRACTICE WITH THE GRADIENT OPERATOR

Prove the following identities:

- (a) $\nabla(|\mathbf{x}|) = \mathbf{x} |\mathbf{x}|^{-1} \equiv \hat{\mathbf{r}}$
- (b) $\nabla \ln |\mathbf{x}| = \mathbf{x} |\mathbf{x}|^{-2} = \hat{\mathbf{r}} |\mathbf{x}|^{-1}$
- (c) $\nabla |\mathbf{x}|^{-1} = -\mathbf{x} |\mathbf{x}|^{-3} = -\hat{\mathbf{r}} |\mathbf{x}|^{-2}$.

EXERCISE 4.2: PRACTICE WITH THE LAPLACIAN OPERATOR

Show that the Laplacian of the function

$$\psi = \frac{z x^2}{r^2} \quad (4.86)$$

is given by

$$\nabla^2 \psi = \frac{2 z (r^2 - 5 x^2)}{r^4}. \quad (4.87)$$

Perform the proof using both Cartesian coordinates as well as spherical coordinates (see Figure 10.1), making use of the following expressions for Laplacian operator acting on a scalar field

$$\nabla^2 \psi(x, y, z) = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \quad (4.88a)$$

$$\nabla^2 \psi(\lambda, \phi, r) = \frac{1}{r^2 \cos \phi} \left[\frac{1}{\cos \phi} \frac{\partial^2 \psi}{\partial \lambda^2} + \frac{\partial}{\partial \phi} \left(\cos \phi \frac{\partial \psi}{\partial \phi} \right) + \cos \phi \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) \right]. \quad (4.88b)$$

EXERCISE 4.3: MORE PRACTICE WITH OPERATORS

Prove the following identities with $r \neq 0$:

- (a) $\nabla^2 r^{-1} = 0$
- (b) $\nabla \cdot (\mathbf{x}/r^3) = 0$
- (c) $\nabla \cdot (\mathbf{A} \wedge \mathbf{x}) = \mathbf{x} \cdot (\nabla \wedge \mathbf{A})$ for an arbitrary vector field $A(\mathbf{x})$.
- (d) $\nabla \wedge [\mathbf{x} f(r)] = 0$ for an arbitrary function $f(r) = f(|\mathbf{x}|)$.

EXERCISE 4.4: SOLID-BODY ROTATION

Define a velocity field according to

$$\mathbf{v} = \boldsymbol{\Omega} \wedge \mathbf{x} \quad (4.89)$$

with $\boldsymbol{\Omega}$ a spatial constant. This velocity field describes solid-body rotation as discussed in Section 46.5. Show that $2\boldsymbol{\Omega} = \nabla \wedge \mathbf{v}$. See also Exercise 46.1.

EXERCISE 4.5: DIVERGENCE-FREE AND IRROTATIONAL VECTOR

Let Φ be a harmonic function so that $\nabla^2\Phi = 0$. Show that $\mathbf{v} = -\nabla\phi$ satisfies

$$(a) \nabla \cdot \mathbf{v} = 0$$

$$(b) \nabla \wedge \mathbf{v} = 0.$$

In this way we prove that all harmonic scalar fields correspond to a divergence-free and curl-free vector field.

EXERCISE 4.6: CONSERVATIVE VECTOR FIELD AND SCALAR POTENTIAL

Show that the curl, $\nabla \wedge \mathbf{F}$, of the following vector field vanishes

$$\mathbf{F} = 2xz\hat{\mathbf{x}} + 2y^2\hat{\mathbf{y}} + (x^2 + 2y^2z - 1)\hat{\mathbf{z}}. \quad (4.90)$$

Hence, deduce that \mathbf{F} is a conservative vector field so that it can be written as the gradient of a scalar potential ψ according to $\mathbf{F} = -\nabla\psi$, where (to within an arbitrary constant)

$$\psi = -[x^2z + (yz)^2 - z]. \quad (4.91)$$

EXERCISE 4.7: BELTRAMI FLOW

Beltrami flow is defined by velocity and vorticity fields satisfying

$$\nabla \cdot \mathbf{v} = 0 \quad (4.92a)$$

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} = \lambda \mathbf{v} \quad (4.92b)$$

where λ is a constant. Show that the following velocity field is a Beltrami flow

$$\mathbf{v} = (A \sin z + C \cos y)\hat{\mathbf{x}} + (B \sin x + A \cos z)\hat{\mathbf{y}} + (C \sin y + B \cos x)\hat{\mathbf{z}}, \quad (4.93)$$

where A, B, C are constants. Hint: the solution follows directly from computing

$$u = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \quad v = \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \quad w = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \quad (4.94)$$

EXERCISE 4.8: PRACTICE WITH PATH INTEGRALS

Consider the vector field

$$\mathbf{F} = xy^2\hat{\mathbf{x}} + 2\hat{\mathbf{y}} + x\hat{\mathbf{z}}. \quad (4.95)$$

Let L be a path parameterized by

$$x = ct \quad y = c/t \quad z = d \quad t \in [1, 2], \quad (4.96)$$

where c and d are constants. Show that the following identities hold

$$\int_L \mathbf{F} dt = c^3 \ln 2 \hat{\mathbf{x}} + 2\hat{\mathbf{y}} + \frac{3c}{2}\hat{\mathbf{z}} \quad (4.97a)$$

$$\int_L \mathbf{F} dy = -\frac{3c^4}{8}\hat{\mathbf{x}} - c\hat{\mathbf{y}} - c^2 \ln 2\hat{\mathbf{z}} \quad (4.97b)$$

$$\int_L \mathbf{F} \cdot d\mathbf{x} = c^4 \ln 2 - c, \quad (4.97c)$$

where $d\mathbf{x} = \hat{\mathbf{x}}dx + \hat{\mathbf{y}}dy + \hat{\mathbf{z}}dz$. Although all three integrals are computed along the same path, they are not necessarily of the same type. In particular, the first two integrals are vector fields, whereas the third integral is a scalar.

EXERCISE 4.9: STOKES' THEOREM ON A PLANE

Show that

$$\mathcal{I} = \oint_{\partial\mathcal{S}} [y(4x^2 + y^2) dx + x(2x^2 + 3y^2) dy] = \frac{\pi}{2} b a^3 \quad (4.98)$$

when integrating around the ellipse \mathcal{S} defined by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (4.99)$$

where a, b are constants. Hint: make use of Stokes' Theorem on a plane, otherwise known as Green's Theorem. Also make use of the substitution $x = a \cos \phi$ and the identity

$$\int_{\pi}^0 \sin^2(2\phi) d\phi = -\frac{\pi}{2}. \quad (4.100)$$

EXERCISE 4.10: PRACTICE WITH GAUSS'S DIVERGENCE THEOREM

We here demonstrate the validity of Gauss's divergence theorem for a particular vector field

$$\mathbf{F} = \frac{\alpha \mathbf{x}}{(r^2 + a^2)^{3/2}}, \quad (4.101)$$

where α and a are constants and $r^2 = \mathbf{x} \cdot \mathbf{x}$ is the squared radial distance to a point. Using fluid mechanics jargon, we think of \mathbf{F} as a matter flux with physical dimensions of $M L^{-2} T^{-1}$ (mass length $^{-2}$ time $^{-1}$). Now compute the transport of \mathbf{F} through a spherical surface, \mathcal{S} , of radius $|\mathbf{x}| = a\sqrt{3}$

$$\Phi = \oint_{|\mathbf{x}|=a\sqrt{3}} \mathbf{F} \cdot \hat{\mathbf{n}} d\mathcal{S} = \frac{3\pi\alpha\sqrt{3}}{2}. \quad (4.102)$$

With \mathbf{F} a matter flux then Φ has physical dimensions of $M T^{-1}$, so that it is the mass transport through the spherical surface. Next, show that this transport is equal to the integral of the divergence over the volume of the sphere

$$\Phi = \int_{|\mathbf{x}|=a\sqrt{3}} \nabla \cdot \mathbf{F} dV. \quad (4.103)$$

We thus verify, for this particular vector field, the divergence theorem

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{F} dV = \oint_{\partial\mathcal{R}} \mathbf{F} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (4.104)$$

where $\hat{\mathbf{n}}$ is the outward normal on the bounding surface \mathcal{S} .

EXERCISE 4.11: MORE PRACTICE WITH GAUSS'S DIVERGENCE THEOREM

Prove the following identities, which are readily shown using Gauss's divergence theorem.

- (a) $\oint_{\partial\mathcal{R}} \mathbf{x} \cdot \hat{\mathbf{n}} d\mathcal{S} = 3 \int_V dV$, where V is the volume of the closed region bounded by \mathcal{S} .
- (b) $\oint_{\partial\mathcal{R}} (\hat{\mathbf{n}} \wedge \mathbf{F}) d\mathcal{S} = \int_{\mathcal{R}} \nabla \wedge \mathbf{F} dV$, for an arbitrary vector field \mathbf{F} and with $\hat{\mathbf{n}}$ the outward normal on the bounding surface \mathcal{S} . Hint: in a manner similar to the result shown in Section 4.7.2, make use of Gauss's theorem with $\mathbf{A} = \mathbf{F} \wedge \mathbf{C}$ where \mathbf{C} is a constant vector.
- (c) Let \mathcal{S} be a closed surface and \mathbf{x} denote the position vector of a point measured from the origin. Prove the following

$$\oint_{\partial\mathcal{R}} \frac{\hat{\mathbf{n}} \cdot \mathbf{x}}{r^3} = \begin{cases} 0 & \text{if the origin lies outside of } \mathcal{S} \\ 4\pi & \text{if the origin lies inside of } \mathcal{S}. \end{cases} \quad (4.105)$$

5

Partial differential equations

Fluid mechanics is a classical field theory whose mathematical description involves nonlinear partial differential equations (PDEs). Nonetheless, an understanding of linear PDEs provides important insight into the behaviour of geophysical fluids. For example, as seen in Chapter 37, the concentration of trace matter satisfies the advection-diffusion equation. When the tracer is passive it has no impact on the flow so that the advection-diffusion equation is a linear PDE for the tracer concentration. Additionally, a gently perturbed fluid responds through wave fluctuations that are described by a linear PDE (e.g., Chapters 36 and 44). Relatedly, in the study of fluid instabilities we consider whether an infinitesimal perturbation of the linear system grows or decays. The above examples, and many others, motivate a rudimentary study of linear PDEs. This chapter provides a small start to that study.

READER'S GUIDE TO THIS CHAPTER

One can penetrate much of this book without reading this chapter. However, the reader will be far less appreciative of the mathematical structure and physical behavior of the equations describing geophysical fluid mechanics. Indeed, it is remarkable how useful it is, both mathematically and physically, to develop a rudimentary understanding of linear PDEs. Furthermore, those aiming to develop solution methods, either analytical or numerical, should have a working knowledge of this chapter.

There are many resources devoted to the study and application of PDEs throughout physics, engineering, and applied mathematics. Chapter 8 of [Hildebrand \(1976\)](#) offers a pedagogical starting point; chapter 7 of [Morse and Feshbach \(1953\)](#) is an authoritative physics discussion of Green's function methods; [Stakgold \(2000a,b\)](#) thoroughly develops the theory and methods available for boundary value problems encountered in physics, including Green's function methods; and [Duchateau and Zachmann \(1986\)](#) concisely summarize PDEs and provide nearly 300 worked exercises. Much of the presentation in this chapter follows [Duchateau and Zachmann \(1986\)](#).

5.1	The advection equation	62
5.1.1	Constant advection velocity	63
5.1.2	Specifying the arbitrary functions resulting from PDEs	63
5.1.3	Further study	64
5.2	Characteristic curves for first order PDEs	64
5.2.1	General formulation	64
5.2.2	Examples	65
5.3	Classifying second order PDEs	66
5.4	Features of elliptic equations	67
5.4.1	Some general features of Laplace's equation	67
5.4.2	Mean-value property of harmonic functions	67
5.4.3	Laplace's boundary value problem	68
5.4.4	Poisson's equation	69
5.4.5	Extended max-min principle for Poisson's equation	69
5.4.6	Poisson's boundary value problem	70
5.5	Features of the heat/diffusion equation	70
5.5.1	Initial and initial-boundary value problems	71
5.5.2	Smoothing property	71
5.5.3	Duhamel's superposition integral for the heat equation	72
5.5.4	Further study	73
5.6	Features of hyperbolic equations	73
5.6.1	Initial value problem for the infinite-domain wave equation	74
5.6.2	Domain of influence for wave signals	75
5.6.3	Helmholz equation	75
5.6.4	Duhamel's superposition integral for the wave equation	76
5.6.5	Further study	78
5.7	Green's function method	78
5.7.1	Dirac delta (generalized) function	78
5.7.2	Poisson's equation	79
5.7.3	An example Green's function for the heat equation	80
5.7.4	Comments and further study	80

5.1 The advection equation

Consider a tracer concentration, C , which for present purposes is a scalar field that is a function of space and time. As derived in Section 37.4, the tracer concentration in the absence of diffusion satisfies the advection equation

$$\frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C = 0. \quad (5.1)$$

The highest derivatives in both space and time are first order, indicating that the advection equation is a first order partial differential equation. It is a nonlinear PDE for *active* tracers such as temperature, where active tracers affect the velocity through changes to density and hence to pressure. In contrast, the advection equation is linear for *passive* tracers (e.g., dust), which are tracers that do not significantly alter velocity (Section 18.2). We limit the present discussion to passive tracers.

5.1.1 Constant advection velocity

To expose the gist of the advection equation, consider one space dimension and let the advection velocity be constant in space and time,

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} = 0 \quad (5.2)$$

where U is a constant velocity in the \hat{x} direction. An inspired guess reveals that

$$C(x, t) = \Gamma(x - Ut) \quad (5.3)$$

is a general solution to equation (5.2), where Γ is an arbitrary differentiable function that is determined by the initial conditions of the tracer field. Verification of this result is readily found by noting

$$\frac{\partial C}{\partial x} = \frac{d\Gamma(x - Ut)}{d(x - Ut)} \frac{\partial(x - Ut)}{\partial x} = \Gamma' \quad (5.4a)$$

$$\frac{\partial C}{\partial t} = \frac{d\Gamma(x - Ut)}{d(x - Ut)} \frac{\partial(x - Ut)}{\partial t} = -\Gamma' U. \quad (5.4b)$$

The functional dependence $x - Ut$ reveals that as time progresses with $U > 0$, an observer that moves in the positive \hat{x} direction with a speed U maintains a constant argument to the tracer concentration. This behavior means that the tracer concentration is transported by advection with a speed U without changing its structure. We illustrate this behavior in Figure 5.1.

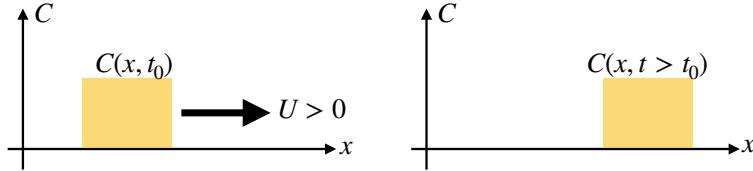


Figure 5.1: Illustrating the advection of a scalar field resulting from a constant advection velocity $v = U\hat{x}$ with $U > 0$. The initially square pulse of tracer is translated, unchanged, by the constant advection velocity.

5.1.2 Specifying the arbitrary functions resulting from PDEs

As revealed from the above example, the solution to a PDE is typically given in terms of an arbitrary function with a specified dependence on the dependent variables. The function itself is unspecified without additional information from initial and/or boundary conditions. For example, prescribe the initial tracer concentration in the form of a sine-wave

$$C(x, t = 0) = C_0 \sin x, \quad (5.5)$$

allow the domain to be infinite in extent (no boundaries), and assume a constant advection velocity. In this case the solution is a wave moving in the positive \hat{x} direction with speed U

$$C(x, t) = C_0 \sin(x - Ut). \quad (5.6)$$

The arbitrary functional degree of freedom is reminiscent of ordinary differential equations, whose solutions are specified up to unknown constants with values set by initial and/or boundary conditions.

5.1.3 Further study

Advection plays a fundamental role in the transport of matter, energy, and momentum within fluids. As seen in our discussion of fluid kinematics in Chapter 16, advection appears in the fluid mechanical equations when viewing the fluid from the fixed laboratory or *Eulerian* reference frame. We thus encounter advection throughout this book, with further development of the mathematics and physics provided in Sections 37.4 and 37.5.

5.2 Characteristic curves for first order PDEs

The advection equation is the canonical first order PDE commonly found in fluid mechanics. A more general form for a first order PDE in one space dimension is given by

$$P(x, t, \psi) \frac{\partial \psi}{\partial x} + Q(x, t, \psi) \frac{\partial \psi}{\partial t} = R(x, t, \psi), \quad (5.7)$$

where P , Q , and R are arbitrary smooth functions. This PDE is linear if P , Q , and R are independent of ψ , and quasi-linear if P and Q are independent of ψ and R is at most a linear function of ψ . In this section we develop a formalism that allows us to determine the functional dependence of the solutions to the PDEs. This *method of characteristics* is quite useful for exposing general properties of the solutions, even for those cases where the solution is not analytically available.

5.2.1 General formulation

In the first order PDE given by equation (5.7), assume there is a functional relation

$$\Upsilon(x, t, \psi) = \text{constant} \quad (5.8)$$

that determines ψ consistent with the PDE (5.7). We refer to Υ as an *integral surface*, with this integral surface specifying a solution to the PDE. For Υ to indeed specify an integral surface it must satisfy

$$\frac{d\Upsilon}{dt} = 0 = \frac{\partial \Upsilon}{\partial \psi} \frac{\partial \psi}{\partial t} + \frac{\partial \Upsilon}{\partial t} \quad (5.9a)$$

$$\frac{d\Upsilon}{dx} = 0 = \frac{\partial \Upsilon}{\partial \psi} \frac{\partial \psi}{\partial x} + \frac{\partial \Upsilon}{\partial x}. \quad (5.9b)$$

So long as $\partial \Upsilon / \partial \psi \neq 0$ then the first order PDE (5.7) takes on the equivalent form

$$P \frac{\partial \Upsilon}{\partial x} + Q \frac{\partial \Upsilon}{\partial t} + R \frac{\partial \Upsilon}{\partial \psi} = 0. \quad (5.10)$$

Considering the ordered triple, (P, Q, R) , as components to a vector in (x, t, ψ) space, then equation (5.10) reveals that (P, Q, R) is perpendicular to the direction in (x, t, ψ) space that is normal to the integral surface $\Upsilon(x, t, \psi) = \text{constant}$. That is, (P, Q, R) lives in the plane tangent to the integral surface. The solution space fills out a curve on the tangent plane known as the *characteristic curve*. This interpretation takes on a somewhat less abstract form if we consider the function ψ to measure the vertical position z of a surface $\psi(x, t) = z$, so that the integral surface is given by

$$\Upsilon(x, t, z) = \text{constant}. \quad (5.11)$$

Let us parameterize the characteristic curve by its arc-length s and let \mathbf{r} be the position on a characteristic curve so that

$$\frac{d\mathbf{r}}{ds} = \hat{\mathbf{x}} \frac{dx}{ds} + \hat{\mathbf{t}} \frac{dt}{ds} + \hat{\mathbf{z}} \frac{d\psi}{ds}, \quad (5.12)$$

where $\hat{\mathbf{t}}$ points in the direction of increasing time. In order for $d\mathbf{r}/ds$ to point in the direction of the tangent to a characteristic curve requires

$$P = \mu \frac{dx}{ds} \quad Q = \mu \frac{dt}{ds} \quad R = \mu \frac{d\psi}{ds} \quad (5.13)$$

for μ an arbitrary function. These relations in turn imply the following ordinary differential equations for the characteristics

$$\frac{dx}{P} = \frac{dt}{Q} = \frac{d\psi}{R}. \quad (5.14)$$

If any one of the functions P , Q , or R vanish, then we merely remove that piece of the above relations.

5.2.2 Examples

Let us ground the discussion by considering the linear homogeneous advection equation

$$U \frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial t} = 0, \quad (5.15)$$

in which we identify $P = U$ and $Q = 1$. The single ODE defining the characteristic curve is given by

$$\frac{dx}{U} = \frac{dt}{1}, \quad (5.16)$$

so that characteristics are given by the family of space-time lines

$$x - Ut = \alpha \quad (5.17)$$

with α an arbitrary constant. These lines determine the paths in space-time along which advective signals are transmitted.

Now add a constant source to the linear advection equation

$$U \frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial t} = R. \quad (5.18)$$

The two ODEs defining the characteristic curve are

$$\frac{dx}{U} = \frac{dt}{1} = \frac{d\psi}{R}. \quad (5.19)$$

In addition to the relation $x - Ut = \alpha_1$ determined from the homogeneous case, we also have $\psi - Rt = \alpha_2$ for α_2 an arbitrary constant. Hence, the characteristic equations render the general solution of the form

$$\Gamma[x - Ut, \psi - Rt] = 0, \quad (5.20)$$

for Γ an arbitrary function. One example solution is given by

$$\psi = f(x - Ut) + Rt \quad (5.21)$$

for an arbitrary smooth function f . This solution has the form of a traveling signal, $f(x - Ut)$, plus a growing source, Rt .

For the final example, consider the linear advection equation with non-constant coefficients and non-constant source

$$x \frac{\partial \psi}{\partial x} + t \frac{\partial \psi}{\partial t} = \psi, \quad (5.22)$$

in which the ODEs determining the characteristics are given by

$$\frac{dx}{x} = \frac{dt}{t} = \frac{d\psi}{\psi}. \quad (5.23)$$

We are thus led to the relations

$$\frac{t}{x} = \alpha_1 \quad \frac{\psi}{x} = \alpha_2. \quad (5.24)$$

Hence, the general solution of the PDE (5.22) is given by

$$\Gamma(t/x, \psi/x) = 0 \Rightarrow \psi = x F(t/x) \quad (5.25)$$

for an arbitrary smooth function F .

5.3 Classifying second order PDEs

There are many second order PDEs appearing in fluid mechanics, a general form of which in one space dimension is given by

$$A \frac{\partial^2 \psi}{\partial x^2} + B \frac{\partial^2 \psi}{\partial x \partial t} + C \frac{\partial^2 \psi}{\partial t^2} = \sigma. \quad (5.26)$$

For linear PDEs, A, B, C are arbitrary functions of space and time that are independent of ψ and σ is a function of space and time and at most a linear function of ψ and its derivatives. Furthermore, for linear PDEs the most general solution consists of the sum of any *particular solution* and a solution to the homogeneous problem (where $\sigma = 0$).

The terms involving second derivatives in equation (5.26) are of principle importance for determining the character of the solutions, with importance placed on the sign of the discriminant $B^2 - 4AC$. By analogy with conic sections we classify 2nd order PDEs as follows

$$\text{PDE form} = \begin{cases} \text{hyperbolic} & B^2 - 4AC > 0 \\ \text{elliptic} & B^2 - 4AC < 0 \\ \text{parabolic} & B^2 - 4AC = 0. \end{cases} \quad (5.27)$$

We can further motivate this terminology by considering the case of a homogeneous constant coefficient PDE and an assumed solution of the form

$$\psi(x, t) = f(mx + t). \quad (5.28)$$

Plugging into the 2nd order PDE (5.26) with $\sigma = 0$ leads to

$$A m^2 + B m + C = 0. \quad (5.29)$$

The two solutions m_1 and m_2 are both real for the hyperbolic case, conjugate complex for the elliptic case, and a perfect square for the parabolic case.

5.4 Features of elliptic equations

The elliptic case from Section 5.3 has $B^2 - 4AC < 0$, which yields complex conjugate characteristics. The canonical example elliptic equation is Laplace's equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0, \quad (5.30)$$

where we converted from t as an independent variable to the space coordinate y . Formally, this transition is realized by setting $t = iy$, where $i = \sqrt{-1}$. Laplace's equation is satisfied by time-independent (i.e., *steady*) solutions to the heat/diffusion equation (Section 5.5). Another common elliptic equation is Poisson's equation, which results from adding a source to Laplace's equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = \sigma. \quad (5.31)$$

As there is no time present in the Laplace and Poisson equations, their solutions are global. That is, “signals” propagate instantaneously so that their structure is fully determined by specified boundary conditions or boundary *data*. Strictly speaking this behavior is not physical since all signals have a finite propagation speed. However, it can be a useful mathematical construct. For example, acoustic signals in fluids propagate much faster than other waves and particle speeds, and they carry a very small energy.¹ For many purposes it is thus suitable to assume acoustic speeds are infinite, and in so doing to *filter* them out of the dynamical equations. In the process, the hyperbolic equation describing acoustic signals is converted into an elliptic equation.

5.4.1 Some general features of Laplace's equation

Solutions to Laplace's equation, $\nabla^2 \psi = 0$, are known as *harmonic functions*. Example harmonic functions are

$$\psi = x^3 - 3xy^2 \quad \psi = \ln(x^2 + y^2) \quad \psi = e^{\gamma x} \cos(\gamma y) \quad \psi = ax + by, \quad (5.32)$$

for arbitrary constants a, b, γ . Furthermore, with

$$\nabla^2(\psi \phi) = \psi \nabla^2 \phi + 2 \nabla \phi \cdot \nabla \psi + \phi \nabla^2 \psi, \quad (5.33)$$

we see that the product of two harmonic functions ($\nabla^2 \psi = \nabla^2 \phi = 0$) is itself harmonic if and only if their gradients are orthogonal, $\nabla \psi \cdot \nabla \phi = 0$. In the remainder of this section we present, without proof, some general properties of harmonic functions and develop self-consistency conditions for the boundary data appearing in the Laplacian boundary value problem (5.36a)-(5.36b).

5.4.2 Mean-value property of harmonic functions

Harmonic functions possess a remarkable *mean-value* property. This property says that the value of a harmonic function, ψ , at a point \mathbf{x}_o within an open region of \mathcal{R} equals to the average of ψ taken over the surface of a sphere within \mathcal{R} centered at \mathbf{x}_o . In equations this property states that

$$\psi(\mathbf{x}_o) = \frac{\oint_{S_R} \psi(\mathbf{x}) dS}{\oint_{S_R} dS}, \quad (5.34)$$

¹A scuba diver feeling the beat of a ship underwater, or an audience member at a rock concert may question this statement. However, acoustic energy is in fact tiny relative to planetary waves and gravity waves, and utterly negligible for studies of large scale geophysical fluid motions.

where \mathcal{S}_R is a sphere with radius R centered at \mathbf{x}_o with “area” given by

$$\oint_{\mathcal{S}_R} d\mathcal{S} = \begin{cases} 2\pi R & n = 2 \text{ space dimensions} \\ 4\pi R^2 & n = 3 \text{ space dimensions.} \end{cases} \quad (5.35)$$

We illustrate this property in Figure 5.2.

The mean-value property of harmonic functions holds anywhere within the domain where $\nabla^2\psi = 0$, so long as the sphere is fully contained within that domain. It implies that there can be no extrema of ψ within the domain, since if there was an extrema then it could not satisfy the mean-value property. Hence, all extrema of harmonic functions must exist on the domain boundary. These properties lend mathematical support for considering harmonic functions to be solutions to continuous physical systems that are in equilibrium or a steady state. As a physical example, consider a temperature field, $T(\mathbf{x})$, in a region with zero heat sources and zero fluid flow. As shown in Section 37.2.4, the steady state temperature satisfies $\nabla^2T = 0$, and as such it is harmonic and hence has no extrema within the domain.

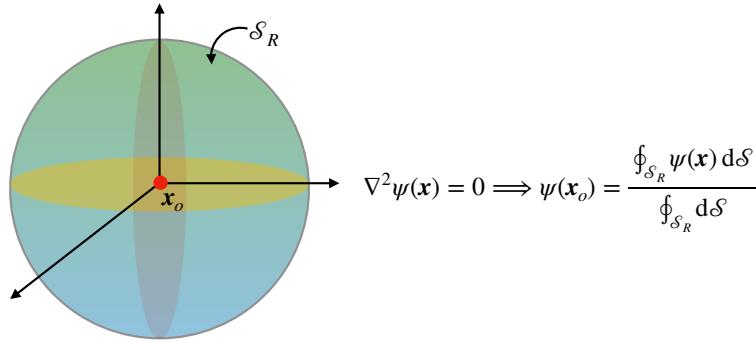


Figure 5.2: The value of a harmonic function at a point \mathbf{x}_o equals to the area average of the function over a sphere centered at \mathbf{x}_o . We here illustrate this property for 3-dimensions, but it holds for arbitrary space dimensions.

5.4.3 Laplace’s boundary value problem

Laplace’s equation requires boundary conditions to fully specify the unknown harmonic function. We here consider the Laplacian *boundary value problem* in the form

$$\nabla^2\psi = 0 \quad \mathbf{x} \in \mathcal{R} \quad (5.36a)$$

$$\alpha\psi + \beta\hat{\mathbf{n}} \cdot \nabla\psi = f \quad \mathbf{x} \in \partial\mathcal{R}, \quad (5.36b)$$

where \mathcal{R} is a smooth and simply connected volume, $\partial\mathcal{R}$ is the boundary surface enclosing \mathcal{R} and with outward normal $\hat{\mathbf{n}}$, and α , β , and f are given boundary data functions.

We can establish constraints on the boundary conditions that lead to a self-consistent Laplacian boundary value problem (5.36a)-(5.36b). We do so through the use of Gauss’s divergence theorem (Section 4.7) in which integration over the full domain leads to

$$0 = \int_{\mathcal{R}} \nabla^2\psi dV = \oint_{\partial\mathcal{R}} \hat{\mathbf{n}} \cdot \nabla\psi d\mathcal{S}. \quad (5.37)$$

In physical applications the boundary condition (5.36b) usually appear with either $\alpha = 0$ or $\beta = 0$, and these two cases are associated with distinct self-consistency constraints.

Dirichlet boundary condition

The case with $\beta = 0$ is referred to as a *Dirichlet* boundary condition whereby

$$\psi = f \quad \mathbf{x} \in \partial\mathcal{R}, \quad (5.38)$$

where we set $\alpha = 1$ without loss of generality. In this case all boundary data result in a self-consistent Laplacian boundary value problem so there is no constraint on f . Thinking again about temperature, this boundary condition specifies the temperature on the domain boundary to equal $T = f$, with all boundary functions f consistent with a harmonic temperature distribution within the domain interior.

Neumann boundary condition

The case with $\alpha = 0$ results in a *Neumann* boundary condition. Without loss of generality we set $\beta = 1$ and reach the following self-consistency condition

$$\oint_{\partial\mathcal{R}} f \, dS = 0. \quad (5.39)$$

That is, a self-consistent boundary condition for Laplace's equation with a Neumann boundary condition must satisfy this surface integral constraint. For the temperature example, this boundary condition says that to realize a steady state harmonic temperature distribution within a region, we can at most apply a zero area averaged boundary heating. If the boundary constraint (5.39) is not satisfied, then the interior temperature field cannot be harmonic so that it will not be in a steady state.

5.4.4 Poisson's equation

The generic boundary value problem for Poisson's equation takes on the form

$$\nabla^2 \psi = \sigma \quad \mathbf{x} \in \mathcal{R} \quad (5.40a)$$

$$\alpha \psi + \beta \hat{\mathbf{n}} \cdot \nabla \psi = f \quad \mathbf{x} \in \partial\mathcal{R}, \quad (5.40b)$$

where $\sigma(\mathbf{x})$ is a specified source function. We here present, without proof, some general properties of solutions to Poisson's equation and develop self-consistency conditions for the boundary data appearing in equation (5.40b).

5.4.5 Extended max-min principle for Poisson's equation

A subharmonic function is one where

$$\nabla^2 \psi = \sigma \geq 0 \quad \mathbf{x} \in \mathcal{R}. \quad (5.41)$$

Here, the source function makes the curvature of a subharmonic function positive. Correspondingly, every point within \mathcal{R} satisfies the minimum principle

$$\psi(\mathbf{x}_o) \leq \frac{\oint_{\mathcal{S}_R} \psi(\mathbf{x}) \, dS}{\oint_{\mathcal{S}_R} dS}, \quad (5.42)$$

for spheres, \mathcal{S}_R , that are fully within \mathcal{R} . The signs switch for superharmonic functions whereby $\nabla^2 \psi \leq 0$ for $\mathbf{x} \in \mathcal{R}$

Returning to the temperature example, consider a temperature field in a region with a positive heat source, $\sigma > 0$. The steady state temperature in the presence of zero fluid flow satisfies Poisson's equation $\nabla^2 T = \sigma \geq 0$ for regions with the heat source. The minimum principle (5.42) means that the temperature at any point within the heating region is less than the temperature averaged over a sphere centered on the point, so long as the sphere remains within the region of heating. It is only in the absence of a heat source or sink, where $\nabla^2 T = 0$, that we recover the mean-value property of harmonic functions given by equation (5.34).

5.4.6 Poisson's boundary value problem

We follow the method in Section 5.4.3 to develop constraints on the boundary conditions applied as part of the Poisson boundary value problem (5.40a)-(5.40b). Use of Gauss's divergence theorem leads to the constraint

$$\oint_{\partial\mathcal{R}} \hat{\mathbf{n}} \cdot \nabla \psi \, d\mathcal{S} = \int_{\mathcal{R}} \sigma \, dV. \quad (5.43)$$

We separately consider the Dirichlet and Neumann cases.

Dirichlet boundary condition

The Dirichlet condition with $\beta = 0$ leads to

$$\psi = f \quad \mathbf{x} \in \partial\mathcal{R}. \quad (5.44)$$

Just as for the Laplacian boundary value problem, all boundary data result in a self-consistent Poisson boundary value problem so there is no constraint on f . Thinking again about temperature, this boundary condition specifies the temperature on the domain boundary to equal $T = f$, with all boundary functions f consistent with the interior heating σ and a temperature field satisfying $\nabla^2 T = \sigma$ within the interior.

Neumann boundary condition

The *Neumann* boundary condition leads to the following self-consistency condition

$$\oint_{\partial\mathcal{R}} f \, d\mathcal{S} = \int_{\mathcal{R}} \sigma \, dV. \quad (5.45)$$

That is, a self-consistent boundary condition for Poisson's equation with a Neumann boundary condition must satisfy this surface integral constraint. For the temperature example, this boundary condition says that the area integrated boundary data must be consistent with the volume integrated source function in order for the temperature to satisfy Poisson's equation. Otherwise, the temperature field will evolve and not realize a steady state.

5.5 Features of the heat/diffusion equation

The parabolic case from Section 5.3, $B^2 - 4AC = 0$, contains a single real characteristic. The canonical example is the *heat equation*, which is also known as the *diffusion equation*

$$\frac{\partial \psi}{\partial t} = \kappa \frac{\partial^2 \psi}{\partial x^2}, \quad (5.46)$$

where $\kappa > 0$ is the kinematic diffusivity.

5.5.1 Initial and initial-boundary value problems

The *Cauchy Problem* is the name given to the initial value problem for the heat equation in the full space of reals, \mathbb{R}^n

$$\frac{\partial \psi}{\partial t} = \kappa \nabla^2 \psi \quad \mathbf{x} \in \mathbb{R}^n, t > 0 \quad (5.47a)$$

$$\psi(\mathbf{x}, t = 0) = f(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^n \quad (5.47b)$$

$$|\psi(\mathbf{x}, t)| < \infty \quad \mathbf{x} \in \mathbb{R}^n, t > 0. \quad (5.47c)$$

The general initial-boundary value problem over a finite domain \mathcal{R} takes the form

$$\frac{\partial \psi}{\partial t} = \kappa \nabla^2 \psi \quad \mathbf{x} \in \mathcal{R}, t > 0 \quad (5.48a)$$

$$\psi(\mathbf{x}, t = 0) = f(\mathbf{x}) \quad \mathbf{x} \in \mathcal{R} \quad (5.48b)$$

$$\alpha(\mathbf{x}) \psi(\mathbf{x}, t) + \beta(\mathbf{x}) \hat{\mathbf{n}} \cdot \nabla \psi(\mathbf{x}, t) = g(\mathbf{x}, t) \quad \mathbf{x} \in \mathcal{S}, t > 0, \alpha \beta \geq 0. \quad (5.48c)$$

We have more occasion to consider the initial-boundary value problem in geophysical fluid mechanics. Following from the discussion of Laplace's and Poisson's boundary value problems, choices for the boundary functions α and β impact on the character of the boundary conditions. The Neumann condition is most commonly applied to set the flux of a tracer or temperature at the boundaries. The alternative use of the Dirichlet condition is less common in geophysical fluid applications.

5.5.2 Smoothing property

The extended max-min principle from Section 5.4.5 holds also for the heat equation, which is consistent with solutions to the heat equation generally decaying their initial condition towards zero by reducing the amplitude of all extrema. Hence, no extrema are introduced in the interior of the domain by the heat equation; extrema only arise via boundary and/or initial conditions. Furthermore, the steady state limit of the heat equation is a harmonic function solution to Laplace's equation whereby the mean-value property holds (Section 5.4.2). These characteristics of the heat equation are generally shared by all linear parabolic PDEs.

Smoothing property for a finite domain initial-boundary value problem

We illustrate the smoothing property for the specific case of the one-dimensional initial-boundary value problem with Dirichlet boundary conditions

$$\frac{\partial \psi}{\partial t} = \kappa \frac{\partial^2 \psi}{\partial x^2} \quad 0 < x < \pi, t > 0 \quad (5.49a)$$

$$\psi(x, t) = f(x) \quad 0 < x < \pi, t = 0 \quad (5.49b)$$

$$\psi(0, t) = \psi(\pi, t) = 0 \quad t > 0, \quad (5.49c)$$

where x and t are non-dimensional space and time variables. A variety of methods can be used to construct the following Fourier series solution

$$\psi(x, t) = \sum_{n=1}^{\infty} f_n e^{-n^2 t} \sin(n x) \quad 0 < x < \pi, t > 0 \quad (5.50a)$$

$$f_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(n x) dx \quad n = 1, 2, \dots \quad (5.50b)$$

As per the smoothing property, note how the amplitude of each Fourier mode decays exponentially in time.

Smoothing property for an initial value problem on the real line

Now consider the one-dimensional heat equation on the real line, with the only boundary conditions being regularity at infinity

$$\frac{\partial \psi}{\partial t} = \kappa \frac{\partial^2 \psi}{\partial x^2} \quad -\infty < x < \infty, t > 0 \quad (5.51a)$$

$$\psi(x, t) = f(x) \quad -\infty < x < \infty, t = 0. \quad (5.51b)$$

One can show by direct differentiation that the following Gaussian is a solution

$$\psi(x, t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} f(y) \exp\left[-\frac{(x-\xi)^2}{4\kappa t}\right] d\xi. \quad (5.52)$$

Again, this function smooths/damps the initial condition function $f(x)$ as time moves forward.

5.5.3 Duhamel's superposition integral for the heat equation

Consider a scalar field that starts from zero initial conditions and evolves in the presence of a source

$$\frac{\partial \Psi}{\partial t} = \kappa \nabla^2 \Psi + f(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^n, t > 0 \quad (5.53a)$$

$$\Psi(\mathbf{x}, t) = 0 \quad \mathbf{x} \in \mathbb{R}^n, t = 0. \quad (5.53b)$$

Now consider the converse, in which another scalar field evolves without a source and yet is initialized according to the source

$$\frac{\partial \psi}{\partial t} = \kappa \nabla^2 \psi \quad \mathbf{x} \in \mathbb{R}^n, t > 0 \quad (5.54a)$$

$$\psi(\mathbf{x}, t) = f(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^n, t = 0. \quad (5.54b)$$

The two scalar fields are related by *Duhamel's superposition integral*

$$\Psi(\mathbf{x}, t) = \int_0^t \psi(\mathbf{x}, t-\tau) d\tau. \quad (5.55)$$

We verify the connection by direct differentiation

$$\frac{\partial \Psi(\mathbf{x}, t)}{\partial t} = \psi(\mathbf{x}, 0) + \int_0^t \frac{\partial \psi(\mathbf{x}, t-\tau)}{\partial t} d\tau = f(\mathbf{x}) + \kappa \nabla^2 \Psi(\mathbf{x}, t). \quad (5.56)$$

Duhamel's superposition integral allows us to move the source from the partial differential operator into the initial conditions, and vice-versa. It says that the forced solution $\Psi(\mathbf{x}, t)$ is built by time integrating the “retarded” values of the unforced solution ψ from the initial time $t = 0$ to the current time t . Note that a more general presentation allows for the source function to be a function of time, $f(\mathbf{x}, t)$, in which case we must develop a family of solutions, $\psi_f(\mathbf{x}, t; \tau)$, generated by reinitializing $\psi_f(\mathbf{x}, t = \tau; \tau) = f(\mathbf{x}, \tau)$ and then superposing the members of this family to generate $\Psi(\mathbf{x}, t)$.

As an example, consider the initial value problem for the heat equation on a line as given by equations (5.51a)-(5.51b)

$$\frac{\partial \psi}{\partial t} = \kappa \frac{\partial^2 \psi}{\partial x^2} \quad -\infty < x < \infty, t > 0 \quad (5.57a)$$

$$\psi(x, t) = f(x) \quad -\infty < x < \infty, t = 0, \quad (5.57b)$$

whose solution is given by the Gaussian function in equation (5.52). Duhamel's superposition integral (5.55) says that

$$\Psi(x, t) = \int_0^t \psi(x, t - \tau) d\tau = \int_0^t \frac{1}{\sqrt{4\pi\kappa(t-\tau)}} \int_{-\infty}^{\infty} f(\xi) \exp\left[-\frac{(x-\xi)^2}{4\kappa(t-\tau)}\right] d\tau d\xi \quad (5.58)$$

satisfies the forced (inhomogeneous) initial value problem with zero initial condition

$$\frac{\partial\Psi}{\partial t} = \kappa \frac{\partial^2\Psi}{\partial x^2} + f(x) \quad -\infty < x < \infty, t > 0 \quad (5.59a)$$

$$\Psi(x, t) = 0 \quad -\infty < x < \infty, t = 0. \quad (5.59b)$$

We make use of this result in our discussion of Green's functions in Section 5.7.

5.5.4 Further study

We examine physical and mathematical properties of the heat/diffusion equation in Sections 37.2 and 37.3. Section 9.11 of [Hildebrand \(1976\)](#) offers a lucid discussion of Duhamel's superposition integral.

5.6 Features of hyperbolic equations

The hyperbolic case from Section 5.3 has $B^2 - 4AC > 0$ and thus contains two real characteristics. The canonical example of a hyperbolic PDE is the linear homogeneous wave equation

$$\frac{\partial^2\psi}{\partial t^2} - U^2 \frac{\partial^2\psi}{\partial x^2} = 0. \quad (5.60)$$

Solutions have the form of a moving wave in both directions (the two wave characteristics)

$$\psi(x, t) = \mathcal{F}(x - Ut) + \mathcal{G}(x + Ut), \quad (5.61)$$

where \mathcal{F} and \mathcal{G} are differentiable functions whose form is determined by the initial conditions. Note that we can factor the differential operator into the form

$$(\partial_t - U \partial_x)(\partial_t + U \partial_x)\psi = 0. \quad (5.62)$$

Consequently, if either one of the linear first-order PDEs are satisfied

$$(\partial_t - U \partial_x)\psi = 0 \quad (5.63a)$$

$$(\partial_t + U \partial_x)\psi = 0 \quad (5.63b)$$

then ψ will satisfy the full wave equation. These first-order PDEs are the one-dimensional advection equations considered in Section 5.2 with opposite advection direction, and each of which has a single characteristic. In this manner, we can think of advection by constant velocity as the square root of the wave equation. Similarly, some disciplines refer to the linear advection equation (5.2), with constant advection speed, as the *one-way wave equation*.

5.6.1 Initial value problem for the infinite-domain wave equation

Since there are two time derivatives, specification of a solution requires initial conditions for the field and its first time derivative. To illustrate the structure of a solution to the wave equation, we develop a solution to the *Cauchy problem*, which is the initial value problem for the one-dimensional wave equation on the real line (no boundary conditions)

$$\frac{\partial^2 \psi}{\partial t^2} = c^2 \frac{\partial^2 \psi}{\partial x^2} \quad -\infty < x < \infty, t > 0 \quad \text{wave equation on a line} \quad (5.64a)$$

$$\psi = F(x) \quad -\infty < x < \infty, t = 0 \quad \text{initial condition} \quad (5.64b)$$

$$\frac{\partial \psi}{\partial t} = G(x) \quad -\infty < x < \infty, t = 0 \quad \text{initial tendency,} \quad (5.64c)$$

where the intial condition data, F, G are arbitrary functions of space and c is a constant wave speed. Following from the discussion of characteristics in Section 5.2, we are motivated to transform the wave equation into wave characteristic coordinates

$$\xi = x + ct \quad \text{and} \quad \eta = x - ct \implies \frac{\xi + \eta}{2} = x \quad \text{and} \quad \frac{\xi - \eta}{2c} = t. \quad (5.65)$$

Wave signals propagate in directions defined by constant ξ and η , so that these coordinates isolate the signal transmission. Furthermore, as we will see, this coordinate transformation facilitates a direct integration of the wave equation.

Transformation to characteristic coordinates

To help organize the transformation to characteristic coordinates, we write equation (5.65) as a matrix-vector equation

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} c & 1 \\ -c & 1 \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} \iff \begin{bmatrix} t \\ x \end{bmatrix} = \frac{1}{2} \begin{bmatrix} c^{-1} & -c^{-1} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}. \quad (5.66)$$

Furthermore, define

$$x^n = (x^0, x^1) = (x, t) \quad \text{and} \quad x^{\bar{n}} = (x^{\bar{0}}, x^{\bar{1}}) = (\xi, \eta) \quad (5.67)$$

and use index notation from Chapter 3 so that the transformation (5.66) takes the tidy form

$$x^{\bar{n}} = \Lambda^{\bar{n}}_m x^m \iff x^n = \Lambda^n_{\bar{m}} x^{\bar{m}}, \quad (5.68)$$

where the transformation matrices are given by

$$\Lambda^{\bar{n}}_m = \begin{bmatrix} c & 1 \\ -c & 1 \end{bmatrix} \quad \text{and} \quad \Lambda^n_{\bar{m}} = \frac{1}{2} \begin{bmatrix} c^{-1} & -c^{-1} \\ 1 & 1 \end{bmatrix}. \quad (5.69)$$

Note the use of an upstairs position for the row index on the transformation matrix, which conforms to the use with general tensors from Chapter 7. For present purposes there is no significance to this position placement; it is only used here for convenience. The coordinate transformation (5.68) and the transformation matrices (5.69) then lead to the partial derivative relationship

$$\partial_{\bar{n}} = \Lambda^m_{\bar{n}} \partial_m \iff \partial_n = \Lambda^{\bar{m}}_n x^{\bar{m}}, \quad (5.70)$$

so that

$$\frac{\partial^2}{\partial x^2} = \left[\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \right]^2 = \frac{\partial^2}{\partial \xi^2} + 2 \frac{\partial}{\partial \xi} \frac{\partial}{\partial \eta} + \frac{\partial^2}{\partial \eta^2} \quad (5.71a)$$

$$c^{-2} \frac{\partial^2}{\partial t^2} = \left[\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta} \right]^2 = \frac{\partial^2}{\partial \xi^2} - 2 \frac{\partial}{\partial \xi} \frac{\partial}{\partial \eta} + \frac{\partial^2}{\partial \eta^2}. \quad (5.71b)$$

General solution for the initial value problem

The operator transformations (5.71a) and (5.71b) bring the initial value problem (5.64a)-(5.64c) into

$$\frac{\partial^2 \psi}{\partial \xi \partial \eta} = 0 \quad -\infty < \eta < \xi < \infty \quad (5.72a)$$

$$\psi(\xi, \eta) = F(\xi) \quad -\infty < \xi < \infty, \xi = \eta \quad (5.72b)$$

$$\frac{\partial \psi}{\partial \xi} - \frac{\partial \psi}{\partial \eta} = c^{-1} G(\xi) \quad -\infty < \xi < \infty, \xi = \eta. \quad (5.72c)$$

Integrating equation (5.72a) in two steps leads to $\partial_\xi \psi = \theta(\xi)$ so that

$$\psi(\xi, \eta) = \Phi(\eta) + \int^\xi \theta(s) ds \equiv \Phi(\eta) + \Theta(\xi), \quad (5.73)$$

for two functions $\Phi(\eta)$ and $\Theta(\xi)$. The initial conditions (5.72b) and (5.72c) determine relations between $\Phi(\eta)$ and $\Theta(\xi)$ and the initial data

$$\Theta(\xi) = \frac{1}{2} \left[F(\xi) + \frac{1}{c} \int^\xi G(s) ds \right] \quad (5.74a)$$

$$\Phi(\eta) = \frac{1}{2} \left[F(\eta) - \frac{1}{c} \int^\eta G(s) ds \right], \quad (5.74b)$$

in which case the general solution to the initial value problem (5.64a)-(5.64c) takes the form

$$\psi(x, t) = \frac{1}{2} [F(x + ct) + F(x - ct)] + \frac{1}{2c} \int_{x-ct}^{x+ct} G(s) ds, \quad (5.75)$$

where we reintroduced the variables (x, t) . This solution is known as the *D'Alembert formula* for the wave equation. Note how the initial profile, $F(x)$ is propagated along the two characteristics, $\xi = x + ct$ and $\eta = x - ct$, without any change. In contrast, the initial tendency, $\partial_t \psi(x, t = 0) = G(x)$, is smoothed through the time integration. This behavior contrasts to the heat equation in Section 5.5, with its single time derivative resulting in a smoothing of the full solution.

5.6.2 Domain of influence for wave signals

The wave solution (5.75) at a point in space time, (x_o, t_o) , depends on data to its past within a causality triangle, or *domain of influence*, as shown in Figure 5.3. The domain of influence is bounded by the two wave characteristics, $\xi_o = x_o + ct_o$ and $\eta_o = x_o - ct_o$. These characteristics are the pathways for propagating information about the initial wave profile, $\psi(x, t = 0) = F(x)$. Points in between the characteristics are causally connected via the initial time tendency, $\partial_t \psi(x, t = 0) = G(x)$, which is integrated over the region $x_o - ct_o \leq x \leq x_o + ct_o$. Points outside the domain of influence are causally disconnected from the point (x_o, t_o) .

5.6.3 Helmholtz equation

Consider the wave equation with a constant wave speed

$$\frac{\partial^2 \psi}{\partial t^2} - c^2 \nabla^2 \psi = 0. \quad (5.76)$$

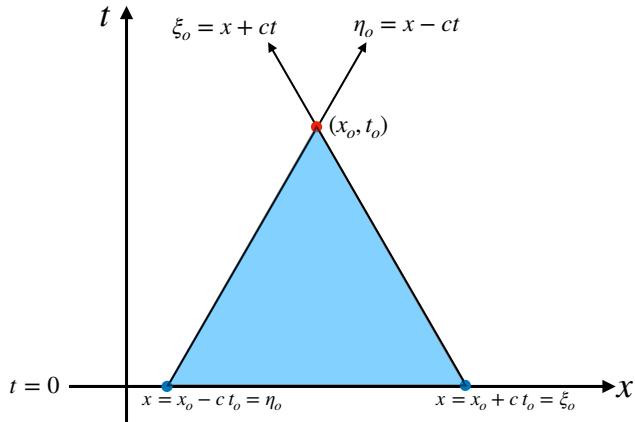


Figure 5.3: According to the wave solution (5.75) to the initial value problem on a line, an arbitrary space-time point, (x_o, t_o) , is causally connected via wave signals to all space-time points within the blue region. This domain of influence is bounded by the two wave characteristics, $\xi_o = x_o + c t_o$ and $\eta_o = x_o - c t_o$, with these characteristics the pathway for propagating information about the initial wave profile, $\psi(x, t = 0) = F(x)$. Points in between the characteristics are causally connected via the initial time tendency, $\partial_t \psi(x, t = 0) = G(x)$, which is integrated over the region $x_o - c t_o \leq x \leq x_o + c t_o$. Points outside the domain of influence are causally disconnected from the point (x_o, t_o) .

Assuming a wave ansatz of the form

$$\psi(\mathbf{x}, t) = e^{i\omega t} \Psi(\mathbf{x}) \quad (5.77)$$

results in the *Helmholz equation* for the amplitude function

$$(\nabla^2 + k^2) \Psi = 0, \quad (5.78)$$

where $k = \omega/c$ is the wavenumber. The Helmholtz equation thus plays a central role in the wave theory.

5.6.4 Duhamel's superposition integral for the wave equation

We here present Duhamel's superposition integral for the wave equation, following from the similar discussion for the heat equation in Section 5.5.3. For that purpose, consider the forced wave equation with time independent forcing

$$\frac{\partial^2 \Psi}{\partial t^2} = c^2 \nabla^2 \Psi + G(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^n, t > 0 \quad (5.79a)$$

$$\Psi(\mathbf{x}, t) = 0 \quad \mathbf{x} \in \mathbb{R}^n, t = 0, \quad (5.79b)$$

and the corresponding unforced wave equation with inhomogeneous initial time tendency

$$\frac{\partial^2 \psi}{\partial t^2} = c^2 \nabla^2 \psi \quad \mathbf{x} \in \mathbb{R}^n, t > 0 \quad (5.80a)$$

$$\psi(\mathbf{x}, t) = 0 \quad \mathbf{x} \in \mathbb{R}^n, t = 0. \quad (5.80b)$$

$$\frac{\partial \psi(\mathbf{x}, t)}{\partial t} = G(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^n, t = 0. \quad (5.80c)$$

The two scalar fields are related by Duhamel's superposition integral

$$\Psi(\mathbf{x}, t) = \int_0^t \psi(\mathbf{x}, t - \tau) d\tau. \quad (5.81)$$

We can verify this formula through direct differentiation

$$\frac{\partial \Psi(\mathbf{x}, t)}{\partial t} = \int_0^t \frac{\partial \psi(\mathbf{x}, t - \tau)}{\partial t} d\tau \quad (5.82a)$$

$$\frac{\partial^2 \Psi(\mathbf{x}, t)}{\partial t^2} = \frac{\partial \psi(\mathbf{x}, 0)}{\partial t} + \int_0^t \frac{\partial^2 \psi(\mathbf{x}, t - \tau)}{\partial t^2} d\tau = G(\mathbf{x}) + c^2 \nabla^2 \Psi(\mathbf{x}, t). \quad (5.82b)$$

As an example, consider the initial value problem for the forced wave equation on a line with time independent forcing

$$\frac{\partial^2 \Psi}{\partial t^2} = c^2 \frac{\partial^2 \Psi}{\partial x^2} + G(x) \quad -\infty < x < \infty, t > 0 \quad (5.83a)$$

$$\Psi(x, t) = 0 \quad -\infty < x < \infty, t = 0 \quad (5.83b)$$

$$\frac{\partial \Psi(x, t)}{\partial t} = 0 \quad -\infty < x < \infty, t = 0. \quad (5.83c)$$

Duhamel's superposition integral says that Ψ is related to the solution of the unforced wave equation with initial time tendency given by the forcing

$$\frac{\partial^2 \psi}{\partial t^2} = c^2 \frac{\partial^2 \psi}{\partial x^2} \quad -\infty < x < \infty, t > 0 \quad (5.84a)$$

$$\psi(x, t) = 0 \quad -\infty < x < \infty, t = \tau \quad (5.84b)$$

$$\frac{\partial \psi(x, t)}{\partial t} = G(x) \quad -\infty < x < \infty, t = \tau > 0. \quad (5.84c)$$

We know from Section 5.6.1 that the solution ψ is given by the D'Alembert formula in equation (5.75), only here with the initial condition function set to zero

$$\psi(x, t) = \frac{1}{2c} \int_{x-ct}^{x+ct} G(s) ds. \quad (5.85)$$

Hence, D'Alembert's formula says that the solution to the forced wave equation (5.83a)-(5.83c) is given by the superposition integral

$$\Psi(x, t) = \frac{1}{2c} \int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} G(s) ds. \quad (5.86)$$

Introducing the *antiderivative* function via

$$\mathcal{G}(s) = \int^s G(s') ds' \iff \frac{\partial \mathcal{G}(s)}{\partial s} = G(s) \quad (5.87)$$

allows us to interpret the solution (5.86) as the superposition of two oppositely traveling waves

$$\Psi(x, t) = \frac{1}{2c} \int_0^t [\mathcal{G}[x + c(t - \tau)] - \mathcal{G}[x - c(t - \tau)]] d\tau. \quad (5.88)$$

5.6.5 Further study

Stakgold (2000a,b) provides a thorough discussion of the wave equation and the related Helmholtz equation.

5.7 Green's function method

There are many methods to solve linear PDEs either analytically or numerically. We here introduce the method of Green's functions. This method offers a very powerful conceptual and analytical framework for solving linear PDEs, and it is encountered in many places throughout physics. In particular, we make use of Green's function methods in Section 41.6 when considering the passive tracer advection-diffusion equation.

The Green's function method exploits the superposition property of linear PDEs by finding a particular solution to a PDE (the Green's function) and convolving this solution with boundary and source terms to determine the solution of the original PDE. The Green's function provides a formal means to invert the linear PDE in a manner reminiscent of how one inverts to find the solution to a matrix-vector problem. Our treatment is terse, offering a mere taste of a vast literature in applied mathematics and mathematical physics.

5.7.1 Dirac delta (generalized) function

Consider a still pool of water subject to a point forcing, or consider an electrostatic potential resulting from a point charge. The Green's function is the resulting solution to the PDEs describing these systems with the Dirac delta (generalized) function the mathematical expression for the point source. The delta function is written $\delta(\mathbf{x} - \mathbf{x}')$, and it is the mathematical limit of a point source with infinite strength that is nonzero only when the field point, \mathbf{x} , equals to the source point, $\mathbf{x} = \mathbf{x}'$. The delta function satisfies two key integral properties

$$\int_{\mathcal{R}} \delta(\mathbf{x} - \mathbf{x}') dV' = 1 \quad \text{normalization} \quad (5.89a)$$

$$\int_{\mathcal{R}} \delta(\mathbf{x} - \mathbf{x}') \psi(\mathbf{x}') dV' = \psi(\mathbf{x}) \quad \text{sifting property,} \quad (5.89b)$$

so long as the source location, \mathbf{x}' , is within the integration region \mathcal{R} . Notably, these two relations imply that the delta function has dimensions of inverse volume (if considering 3-dimensional space). If instead we are working on a two-dimensional area then the delta function has dimensions of inverse area, or when on a line it has dimensions of inverse length.

We may find occasion to introduce a temporal delta function, $\delta(t - t')$, which turns on only at a specified time $t' = t$. It satisfies similar properties to the spatial delta function

$$\int_{-\infty}^{\infty} \delta(t - t') dt' = 1 \quad \text{normalization} \quad (5.90a)$$

$$\int_{-\infty}^{\infty} \delta(t - t') \psi(t') dt' = \psi(t) \quad \text{sifting property.} \quad (5.90b)$$

The temporal delta function has dimensions of inverse time, as required by these two properties.

5.7.2 Poisson's equation

We illustrate the Green's function method by considering the three-dimensional Poisson equation with Neumann boundary conditions

$$\nabla^2\psi = \sigma \quad \mathbf{x} \in \mathcal{R} \quad (5.91a)$$

$$\hat{\mathbf{n}} \cdot \nabla\psi = f \quad \mathbf{x} \in \partial\mathcal{R}. \quad (5.91b)$$

The goal is to determine an integral expression for ψ in terms of the known source function, σ , and the boundary data, f . To do so we introduce the Green's function as the solution to the Poisson equation with a delta function source and a homogeneous boundary condition

$$\nabla^2G(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}') \quad \mathbf{x} \in \mathcal{R} \quad (5.92a)$$

$$\hat{\mathbf{n}} \cdot \nabla G = 0 \quad \mathbf{x} \in \partial\mathcal{R}. \quad (5.92b)$$

The Green's function, $G(\mathbf{x}, \mathbf{x}')$, is harmonic everywhere except at the source point, $\mathbf{x} = \mathbf{x}'$, where it equals to the inverse Laplacian acting on the delta function. Furthermore, the Green's function satisfies the homogeneous version of the boundary condition (5.91b); i.e., it does not feel the boundary function f . In some treatments, the Green's function is known as the *singularity solution* to the linear partial differential operator.

Now assume we have a method to determine the Green's function (there are many methods). How does knowing G help to find ψ ? To answer this question requires the second form of Green's integral identity (4.66) using ψ and G

$$\oint_{\partial\mathcal{R}} [\psi \hat{\mathbf{n}} \cdot \nabla G - G \hat{\mathbf{n}} \cdot \nabla\psi] d\mathcal{S}' = \int_{\mathcal{R}} [\psi \nabla^2 G - G \nabla^2\psi] dV'. \quad (5.93)$$

Integration is taken over all positions of the source point \mathbf{x}' , as denoted for the volume integral by dV' and corresponding surface integral $d\mathcal{S}'$. Making use of the Green's function equation (5.92a) and the sifting property (5.89b) means that the first term on the right hand side of equation (5.93) renders

$$\int_{\mathcal{R}} \psi(\mathbf{x}') \nabla^2 G(\mathbf{x}, \mathbf{x}') dV' = \int_{\mathcal{R}} \psi(\mathbf{x}') \delta(\mathbf{x} - \mathbf{x}') dV' = \psi(\mathbf{x}). \quad (5.94)$$

Use of the Poisson equation (5.91a) for the second term on the right hand side of equation (5.93) leads to

$$\int_{\mathcal{R}} G(\mathbf{x}, \mathbf{x}') \nabla^2\psi(\mathbf{x}') dV' = \int_{\mathcal{R}} G(\mathbf{x}, \mathbf{x}') \sigma(\mathbf{x}') dV'. \quad (5.95)$$

The boundary conditions (5.91b) and (5.92b) allow us to write the boundary integral in the form

$$\oint_{\partial\mathcal{R}} [\psi \hat{\mathbf{n}} \cdot \nabla G - G \hat{\mathbf{n}} \cdot \nabla\psi] d\mathcal{S}' = - \oint_{\partial\mathcal{R}} G(\mathbf{x}, \mathbf{x}') f(\mathbf{x}') d\mathcal{S}'. \quad (5.96)$$

Bringing results together leads to

$$\psi(\mathbf{x}) = \underbrace{\int_{\mathcal{R}} G(\mathbf{x}, \mathbf{x}') \sigma(\mathbf{x}') dV'}_{\text{volume convolution}} - \underbrace{\oint_{\partial\mathcal{R}} G(\mathbf{x}, \mathbf{x}') f(\mathbf{x}') d\mathcal{S}'}_{\text{boundary convolution}}. \quad (5.97)$$

This solution for ψ consists of a volume convolution of the Green's function with the source σ plus a boundary convolution of the Green's function with the boundary function f .

5.7.3 An example Green's function for the heat equation

Recall our discussion of Duhamel's superposition integral in Section 5.5.3 where we considered the one-dimensional unforced heat equation with nonhomogeneous initial conditions

$$\frac{\partial \psi}{\partial t} - \kappa \frac{\partial^2 \psi}{\partial x^2} = 0 \quad -\infty < x < \infty, t > 0 \quad (5.98a)$$

$$\psi(x, t) = f(x) \quad -\infty < x < \infty, t = 0, \quad (5.98b)$$

and noted that the solution is

$$\psi(x, t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} f(y) \exp\left[-\frac{(x-y)^2}{4\kappa t}\right] dy. \quad (5.99)$$

Duhamel's superposition integral was then used to show that

$$\Psi(x, t) = \int_0^t \psi(x, t-\tau) d\tau \quad (5.100a)$$

$$= \int_0^t \frac{1}{\sqrt{4\pi\kappa(t-\tau)}} \int_{-\infty}^{\infty} f(\xi) \exp\left[-\frac{(x-\xi)^2}{4\kappa(t-\tau)}\right] d\xi d\tau \quad (5.100b)$$

$$\equiv \int_0^t \int_{-\infty}^{\infty} G(x, t; \xi, \tau) f(\xi) d\xi d\tau \quad (5.100c)$$

satisfies the forced (inhomogeneous) initial value problem with zero initial condition

$$\frac{\partial \Psi}{\partial t} = \kappa \frac{\partial^2 \Psi}{\partial x^2} + f(x) \quad -\infty < x < \infty, t > 0 \quad (5.101a)$$

$$\Psi(x, t) = 0 \quad -\infty < x < \infty, t = 0. \quad (5.101b)$$

In equation (5.100c) we identified the Green's function for the forced heat equation with zero initial conditions

$$G(x, t; \xi, \tau) = \frac{1}{\sqrt{4\pi\kappa(t-\tau)}} \exp\left[-\frac{(x-\xi)^2}{4\kappa(t-\tau)}\right], \quad (5.102)$$

with the formula (5.100c) expressing Ψ as a space-time convolution of the source function with the Green's function.

5.7.4 Comments and further study

The Green's function allows us to invert the linear differential operator taking into account the initial and boundary conditions. Hence, if we have the Green's function then we have a direct integral solution to the initial-boundary value problem. There are many methods available to determine the Green's function (e.g., LaPlace transforms, Fourier transforms, separation of variables, method of images, depending on the linear operator, the space-time domain, and the initial and/or boundary conditions. In Section 41.6 we consider Green's functions for the passive tracer advection-diffusion equation. For more details, the math savvy reader may wish to penetrate the presentations offered in chapter 7 of [Morse and Feshbach \(1953\)](#); the two volume development of boundary value problems by [Stakgold \(2000a\)](#) and [Stakgold \(2000b\)](#), and/or work through the concise summary of Green's functions given in Chapter 8 of [Duchateau and Zachmann \(1986\)](#).

6

Geometry of curves and surfaces

We encounter curves and surfaces throughout the study of geophysical fluid mechanics, with fluid particle pathlines and isopycnal/isentropic surfaces providing two examples. Indeed, curves and surfaces are encountered throughout physics. Hence, there is a well developed mathematics used to describe the differential geometric properties of these objects, and we here introduce the basics.

Although the curves and surfaces of geophysical fluid mechanics are moving as part of the fluid flow, we are concerned here with their instantaneous spatial properties. Hence, time does not appear in this chapter. Furthermore, although curves and surfaces can overturn and wrap around themselves, we restrict attention to curves and surfaces whose normal direction has a nonzero projection onto the vertical; i.e., they have no overhangs and no wrapping (Figure 6.1). This constraint is shared with the surfaces of constant generalized vertical coordinates (e.g., isopycnal surfaces) considered in Chapter 11 and in many other places in this book. It allows us to make use of coordinates known as the *Monge gauge* in condensed matter physics

READER'S GUIDE TO THIS CHAPTER

This chapter requires an understanding of the Cartesian calculus of Chapter 4. The differential geometry presented here is of some use throughout this book, and yet the casual reader need not penetrate the material. Conversely, the interested reader can find far more development of the mathematics, along with physical applications, by studying the physics of fluctuating membranes. Section 10.4 of the condensed matter physics textbook from *Chaikin and Lubensky (1995)* provides a useful starting point.

6.1	Conventions and notation	82
6.2	Curves in 2d Euclidean space (planar curves)	83
6.2.1	Differential increments along the curve	83
6.2.2	Length along the curve	83
6.2.3	Curvature of a curve	85
6.3	Surfaces in 3d Euclidean space	86
6.3.1	Area on the surface	87
6.3.2	Curvature of a surface	87

6.1 Conventions and notation

In this chapter we write the Cartesian position of a point on a surface as

$$\mathbf{S} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + \eta(x, y) \hat{\mathbf{z}} \quad \text{position on surface,} \quad (6.1)$$

with the vertical position written as

$$z = \eta(x, y) \quad \text{vertical position on surface.} \quad (6.2)$$

If we are instead referring to a point on a planar curve in the x - z -plane, then we drop the y -dependence to have

$$\mathbf{S} = x \hat{\mathbf{x}} + \eta(x) \hat{\mathbf{z}} \quad \text{position on planar curve.} \quad (6.3)$$

Time dependence is dropped throughout this chapter since we focus on the spatial geometry of curves and surfaces at a particular time instance.

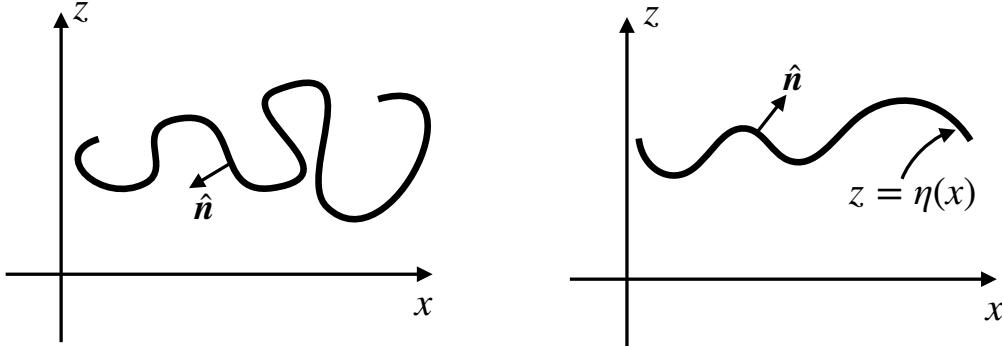


Figure 6.1: Two sample curves on the x - z plane. The left panel shows a curve whose outward normal, $\hat{\mathbf{n}}$, encounters points where $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}} = 0$ and thus where $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}}$ changes sign. This curve, and its generalization to a surface, are not treated in this chapter. The right panel shows a more gently undulating curve where $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}} \neq 0$ everywhere, and thus where $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}}$ is single signed. For these curves we can express the vertical position as a one-to-one function of the horizontal position, $z = \eta(x)$. Again, this curve has its natural generalization to a gently undulating surface whereby $z = \eta(x, y)$ provides a unique mapping between horizontal position and vertical. The assumption regarding no overturning curves and surfaces is consistent with our study of surfaces defined by a constant generalized vertical coordinate (e.g., isopycnals or isentropes) in Chapter 11.

We assume the outward normal direction on the curve or the surface has a nonzero projection into the vertical as shown in Figure 6.1. Indeed, we are only able to write the vertical position as $z = \eta(x, y)$ so long as there are no overturns in the surface, in which case the outward is

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla\eta}{[1 + (\nabla\eta)^2]^{1/2}}. \quad (6.4)$$

Figure 6.2 provides an example surface along with the notation.

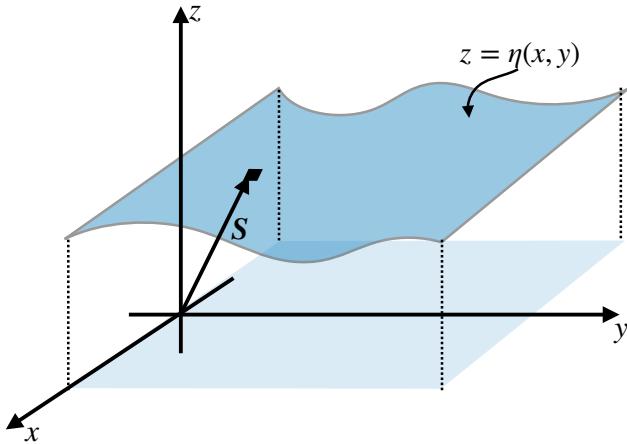


Figure 6.2: A sample surface considered in this chapter. The position of a point on the surface is given by the Cartesian position vector $\mathbf{S} = x \hat{x} + y \hat{y} + \eta(x, y) \hat{z}$. The relation $z = \eta(x, y)$ provides a one-to-one mapping between the horizontal position and the vertical position of a point on the surface. Correspondingly, the surface is uniquely specified by finding the envelope of points where $z - \eta(x, y) = 0$. The shaded region on the horizontal represents the projection of the curved surface onto the flat horizontal plane below.

6.2 Curves in 2d Euclidian space (planar curves)

We here describe the geometry of a curve on the x-z-plane (a *planar curve*) as depicted in Figure 6.3. These curves are 1d objects living in a 2d Euclidean space.

6.2.1 Differential increments along the curve

An arbitrary curve can be parameterized by a single coordinate, referred to here as φ . The differential increment between two infinitesimally close points on the curve is given by

$$\mathbf{S}(\varphi + d\varphi) - \mathbf{S}(\varphi) = d\mathbf{S} = \frac{d\mathbf{S}}{d\varphi} d\varphi \equiv \mathbf{t} d\varphi, \quad (6.5)$$

where

$$\mathbf{t} = \frac{d\mathbf{S}}{d\varphi} \quad (6.6)$$

is tangent to the curve. If $\varphi = s$ is the arc length along the curve, then $\mathbf{t} = \hat{\mathbf{t}}$ is a unit vector

$$\hat{\mathbf{t}} \cdot \hat{\mathbf{t}} = \frac{d\mathbf{S}}{ds} \cdot \frac{d\mathbf{S}}{ds} = 1. \quad (6.7)$$

Recall we made use of the arc-length along a curve in Section 4.4 when describing path integration.

6.2.2 Length along the curve

As in equation (6.3) we can represent the position of a point along the curve using Cartesian coordinates

$$\mathbf{S} = x \hat{x} + \eta(x) \hat{z}. \quad (6.8)$$

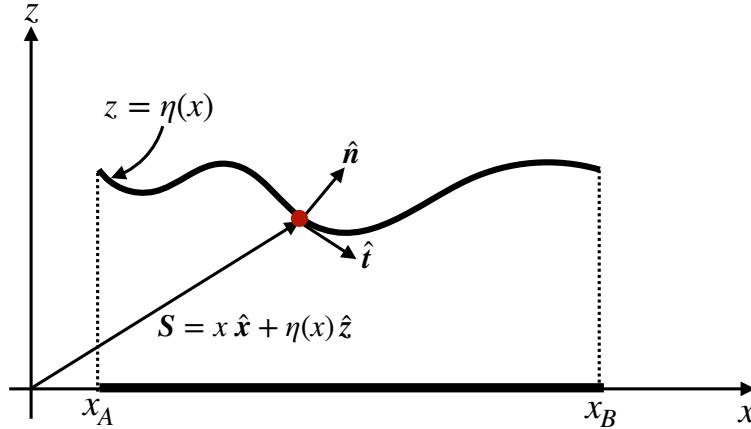


Figure 6.3: A curve in the x - z -plane (planar curve). The Cartesian position of a point on the curve is given by $\mathbf{S} = x \hat{x} + \eta(x) \hat{z}$, where $z = \eta(x)$ is the vertical position of the point. The projection of the curve onto the horizontal x -axis occupies a range $x_A \leq x \leq x_B$. One way to define the curve is by finding the envelope of points where $z - \eta(x) = 0$, in which case we can readily find the normal direction pointing upward as $\hat{\mathbf{n}} = \nabla(z - \eta)/|\nabla(z - \eta)| = [\hat{z} - (\partial\eta/\partial x) \hat{x}] [1 + (\partial\eta/\partial x)^2]^{-1/2}$, and the normalized tangent direction $\hat{\mathbf{t}} = [\hat{x} + (\mathrm{d}\eta/\mathrm{d}x) \hat{z}] [1 + (\mathrm{d}\eta/\mathrm{d}x)^2]^{-1/2}$.

Hence, letting $\varphi = x$ parameterize the curve leads to the representation of the tangent direction

$$\mathbf{t} = \frac{d\mathbf{S}}{dx} = \hat{x} + \frac{d\eta}{dx} \hat{z}, \quad (6.9)$$

which has the magnitude

$$\mathbf{t} \cdot \mathbf{t} = 1 + (\mathrm{d}\eta/\mathrm{d}x)^2, \quad (6.10)$$

so that the normalized tangent vector is

$$\hat{\mathbf{t}} = \frac{\hat{x} + (\mathrm{d}\eta/\mathrm{d}x) \hat{z}}{[1 + (\mathrm{d}\eta/\mathrm{d}x)^2]^{1/2}}. \quad (6.11)$$

Likewise, the curve's normal vector is given by

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{z} - (\partial\eta/\partial x) \hat{x}}{[1 + (\partial\eta/\partial x)^2]^{1/2}}, \quad (6.12)$$

with orthogonality simple to show

$$\hat{\mathbf{t}} \cdot \hat{\mathbf{n}} = 0. \quad (6.13)$$

The squared length of an infinitesimal segment along the curve is given by

$$(ds)^2 = d\mathbf{S} \cdot d\mathbf{S} = \left[\frac{d\mathbf{S}}{dx} \cdot \frac{d\mathbf{S}}{dx} \right] dx dx, \quad (6.14)$$

so that the finite length of the curve is determined by the integral

$$L = \int_0^L ds = \int_{x_A}^{x_B} |d\mathbf{S}/dx| dx = \int_{x_A}^{x_B} [1 + (\mathrm{d}\eta/\mathrm{d}x)^2]^{1/2} dx, \quad (6.15)$$

where $x_A \leq x \leq x_B$ is the range over which x runs for the projection of the curve onto the x -axis (see Figure 6.3).

6.2.3 Curvature of a curve

Curvature measures the amount that the normal direction changes along the curve. For a planar curve, the curvature at a point equals to the inverse radius of a circle that shares the same tangent plane to the curve at the point (see Figure 6.4). We refer to the radius as the *radius of curvature* and the corresponding circle as the *curvature circle*. To formulate an analytic expression for the radius of curvature at a point on a curve, orient the Cartesian coordinate axes so that the point is at the origin and the tangent plane sits along the x-axis as in Figure 6.4. Consequently, the outward normal, \hat{n} , is parallel to the \hat{z} direction.

A Taylor series expansion about the origin tells us that the vertical position of a point along the curve and near to the origin can be written

$$\eta(x) = \eta(0) + x \left[\frac{d\eta}{dx} \right]_{x=0} + \frac{x^2}{2} \left[\frac{d^2\eta}{dx^2} \right]_{x=0} + \mathcal{O}(x^3) \quad (6.16a)$$

$$= \frac{x^2}{2} \left[\frac{d^2\eta}{dx^2} \right]_{x=0} + \mathcal{O}(x^3). \quad (6.16b)$$

This result follows since we placed the origin so that $\eta(0) = 0$, and aligned the x-axis so that it is a tangent plane at the origin, in which case $d\eta/dx = 0$ at $x = 0$. Hence, η has a quadratic behavior near the origin.

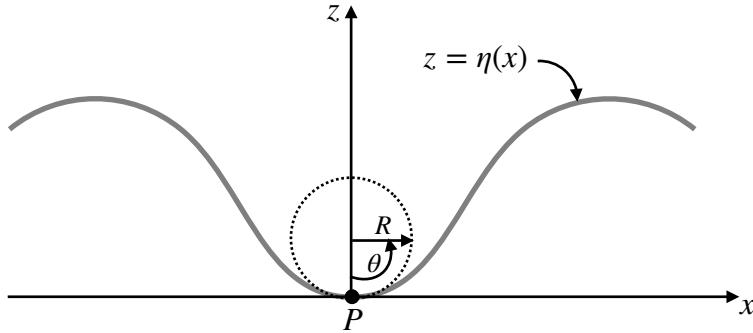


Figure 6.4: The radius of curvature at a point on a curve, P , equals to the radius of the curvature circle that shares the same tangent plane as the curve at the point P . When constructing the curvature circle we make use of the angle, θ , to measure the height of a point along the circle, $h(x) = R(1 - \cos \theta) \approx R\theta^2/2 \approx x^2/(2R)$. Setting $R^{-1} = d^2\eta/dx^2$ provides a second order accurate fit of the curvature circle to the curve at the point P .

Now place a circle with center along the z-axis so that it is tangent to the curve at the origin, as depicted in Figure 6.4. What is the radius, R , of the circle that best fits the curve at the origin? To answer this question note that the height of a point on the circle is given by $h(x) = R(1 - \cos \theta)$, where $\theta = 0$ for a point at the origin and $\theta = \pi$ at the diametrically opposite point. For small θ this height takes the form

$$h(x) \approx R[1 - 1 + \theta^2/2] = x^2/(2R), \quad (6.17)$$

where $\theta = x/R$ near the origin. For the height of a point on the curve (equation (6.16b)) to match the height along the circle, to second order accuracy, requires us to set the circle's radius to

$$\frac{1}{R} = \frac{d^2\eta}{dx^2}. \quad (6.18)$$

Equation (6.18) thus provides an expression for the radius of curvature, R , whose inverse is the curvature

$$\text{curvature} = \frac{1}{R}. \quad (6.19)$$

This result supports our expectation that the second derivative measures the curvature. As R gets larger the curvature decreases as the circle approaches a flat plane. In the opposite limit the curvature grows as R decreases. Note that we could have chosen to orient the circle on the opposite side of the tangent (on the convex side), in which case the radius of curvature is negative. That is, $R > 0$ when the normal points towards the concave side (side where the curve rises towards $\hat{\mathbf{n}}$), whereas $R < 0$ when the normal points towards the convex side (side where the curve falls away from $\hat{\mathbf{n}}$).

In closing this section we note that

$$-\nabla \cdot \hat{\mathbf{n}} = \frac{d^2\eta/dx^2}{[1 + (d\eta/dx)^2]^{3/2}}. \quad (6.20)$$

When evaluated at the point of interest along the curve, we set $d\eta/dx = 0$ so that

$$-\nabla \cdot \hat{\mathbf{n}} = \frac{d^2\eta}{dx^2} = \frac{1}{R}. \quad (6.21)$$

This result supports our earlier statement that curvature measures the change in the normal direction along the curve. In fact, the identity

$$-\nabla \cdot \hat{\mathbf{n}} = \frac{1}{R} \quad (6.22)$$

holds for an arbitrary point along the curve since it is a coordinate invariant statement.

6.3 Surfaces in 3d Euclidean space

We now extend the previous discussion to a 2d surface embedded in 3d Euclidean space such as in Figure 6.2. In general, a 2d surface in 3d space can be parameterized by two variables, φ^1 and φ^2 , so that infinitesimal increments along the surface satisfy

$$d\mathbf{S} = \frac{\partial \mathbf{S}}{\partial \varphi^1} d\varphi^1 + \frac{\partial \mathbf{S}}{\partial \varphi^2} d\varphi^2 = \mathbf{t}_1 d\varphi^1 + \mathbf{t}_2 d\varphi^2. \quad (6.23)$$

The vectors \mathbf{t}_1 and \mathbf{t}_2 are tangent to the surface at the point (φ^1, φ^2) , and yet they are not generally orthogonal to one another.

Making use of the Cartesian expression (6.1)

$$\mathbf{S} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + \eta(x, y) \hat{\mathbf{z}} \quad (6.24)$$

brings the two tangent directions into the form

$$\mathbf{t}_1 = \frac{\partial \mathbf{S}}{\partial x} = \hat{\mathbf{x}} + \frac{\partial \eta}{\partial x} \hat{\mathbf{z}} \quad (6.25a)$$

$$\mathbf{t}_2 = \frac{\partial \mathbf{S}}{\partial y} = \hat{\mathbf{y}} + \frac{\partial \eta}{\partial y} \hat{\mathbf{z}}. \quad (6.25b)$$

Likewise, the surface normal vector is given by

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla\eta}{[1 + (\nabla\eta)^2]^{1/2}}, \quad (6.26)$$

and it is straightforward to show orthogonality with the two tangent vectors

$$\hat{\mathbf{t}}_1 \cdot \hat{\mathbf{n}} = \hat{\mathbf{t}}_2 \cdot \hat{\mathbf{n}} = 0. \quad (6.27)$$

6.3.1 Area on the surface

The area of an infinitesimal surface element with sides $d\varphi_1$ and $d\varphi_2$ is given by

$$dS = \left| \frac{\partial \mathbf{S}}{\partial \varphi^1} \wedge \frac{\partial \mathbf{S}}{\partial \varphi^2} \right| d\varphi_1 d\varphi_2. \quad (6.28)$$

Making use of Cartesian coordinates brings the area to

$$dS = [1 + (\nabla \eta)^2]^{1/2} dx dy, \quad (6.29)$$

so that the area of a finite region is given by the integral

$$\mathcal{S} = \int dS = \int [1 + (\nabla \eta)^2]^{1/2} dx dy, \quad (6.30)$$

where the second integral extends over the region defined by the projection of the surface onto the horizontal (see Figure 6.2).

6.3.2 Curvature of a surface

We now seek an expression for the curvature of a point on the surface. Since the surface has two dimensions, we expect the curvature to be measured by two numbers rather than the single curvature of a curve discussed in Section 6.2.3. The method for developing the curvature is analogous to that used for a curve, yet with a bit more mathematics needed to allow for the extra dimension. Figure 6.5 depicts the situation.

Let $\mathbf{x} = (x_1, x_2) = (x, y)$ be Cartesian coordinates on a tangent plane local to an arbitrary point on the surface, with the origin of the coordinate system taken at the point. Near to the point, we can estimate the vertical distance of a point on the surface from the tangent plane according to the quadratic form

$$\eta \approx \frac{1}{2} x_i \mathbb{K}_{ij} x_j, \quad (6.31)$$

where \mathbb{K} is the second order tensor of second partial derivatives evaluated at the point

$$\mathbb{K} = \begin{bmatrix} \frac{\partial^2 \eta}{\partial x_1^2} & \frac{\partial^2 \eta}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \eta}{\partial x_1 \partial x_2} & \frac{\partial^2 \eta}{\partial x_2^2} \end{bmatrix}. \quad (6.32)$$

As a symmetric tensor, \mathbb{K} is diagonalizable and it has two eigenvalues, R_1^{-1} and R_2^{-1} , along with its associated eigenvectors, \mathbf{e}_1 and \mathbf{e}_2 . The quadratic form (6.31) can thus be written as

$$\eta \approx \frac{1}{2} R_1^{-1} (\mathbf{x} \cdot \mathbf{e}_1)^2 + \frac{1}{2} R_2^{-1} (\mathbf{x} \cdot \mathbf{e}_2)^2. \quad (6.33)$$

R_1 and R_2 are the principle radii of curvature for the surface at the point P . They correspond, respectively, to the radii of the curvature circles in the $\mathbf{n} - \mathbf{e}_1$ and $\mathbf{n} - \mathbf{e}_2$ planes. If the radius of curvature R_i is positive, then the surface curves towards $\hat{\mathbf{n}}$ along the $\mathbf{n} - \mathbf{e}_i$ plane, and conversely if R_i is negative. The surface takes the shape of a saddle when the radii of curvature have opposite signs.

There are two scalar invariants of the tensor \mathbb{K} that commonly arise in applications.

- $\text{Tr}(\mathbb{K}) = R_1^{-1} + R_2^{-1}$, which is twice the mean curvature for the surface. With the normal vector given by equation (6.4), one can show that

$$-\nabla \cdot \hat{\mathbf{n}} = \frac{\nabla^2 \eta}{[1 + (\nabla \eta)^2]^{3/2}}. \quad (6.34)$$

A bit of algebra leads us to conclude that

$$-\nabla \cdot \hat{\mathbf{n}} = \frac{1}{R_1} + \frac{1}{R_2}, \quad (6.35)$$

for any point along the surface, thus generalizing the result (6.21) found for a 1d curve.

- $\det(\mathbb{K}) = 1/(R_1 R_2)$ is known as the *Gaussian curvature*, which is the product of the two curvatures.

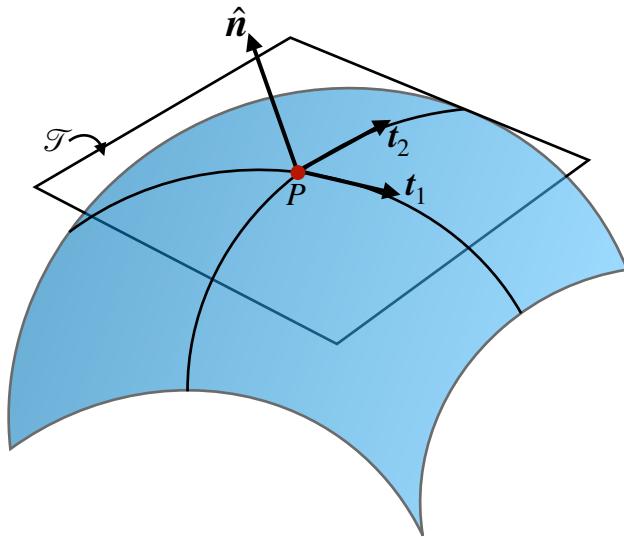


Figure 6.5: Depicting the elements needed to construct the curvature of a surface at an arbitrary point, P . The local normal direction is given by $\hat{\mathbf{n}}$, along with the two tangent vectors \mathbf{t}_1 and \mathbf{t}_2 . The tangent vectors span the space of the tangent plane, \mathcal{T} , shown as a flat surface that is tangent at the chosen point on the surface. In this case the surface falls away from the normal direction, as per a convex surface, so that the two radii of curvature are negative.

7

General tensors in brief

In the study of physics, we aim to uncover objective statements about how physical systems work. That aim is supported by mathematical tools that reflect the underlying objective nature of physical relationships, while also allowing for the quantitative realization of those relationships in particular situations. In this chapter we provide a conceptual platform for general tensor analysis, with details presented in Chapters 8 and 9. This brief chapter offers an overview sufficient to appreciate the power of general tensors yet without going into details.

READER'S GUIDE TO THIS CHAPTER

This chapter builds from the Cartesian tensor analysis of Chapter 3 and introduces mathematical concepts that are central to this book. Even so, the uninterested reader can penetrate the bulk of this book without reading this chapter, and as such it is part of the Tier-II material.

7.1	General tensors and geophysical fluid mechanics	90
7.2	Covariance of physical relations	90
7.2.1	Tensor operations retaining covariance	91
7.2.2	Comments	91
7.3	Points and coordinates	92
7.3.1	Time as a parameter and time as a coordinate	92
7.3.2	The importance of index placement	92
7.4	Example coordinate descriptions	93
7.4.1	Eulerian coordinates	93
7.4.2	Isopycnal coordinates	93
7.4.3	Lagrangian or material coordinates	94
7.4.4	Tracer coordinates	94
7.5	The velocity vector and basis vectors	94
7.5.1	Coordinate representation	95
7.5.2	Basis vectors	95
7.6	Notational conventions	95
7.6.1	Placement of tensor labels	95
7.6.2	Einstein summation convention	95
7.6.3	The boldface notation	96

7.1 General tensors and geophysical fluid mechanics

The Cartesian tensors described in Chapters 3 and 4 are sufficient for many areas of geophysical fluid mechanics. However, there are a number of applications where general tensors prove of great use to ensure that the physics shines through the maths. We thus go beyond Cartesian tensor analysis to enable a more versatile, and precise, mathematical framework for the study of geophysical fluid mechanics.

Geophysical fluid mechanics applications require only a modest level of new formalism in the transition from Cartesian tensors to general tensors. The key reason is that geophysical fluid systems are embedded within Euclidean space, \mathbb{R}^3 . Euclidean space is the familiar space of classical Newtonian mechanics. Notably, \mathbb{R}^3 is flat in that it has zero intrinsic curvature. So although we are concerned with fluid motion on curved manifolds (e.g., spherical planets, isopycnal surfaces), and in describing fluid motion using non-orthogonal coordinate (e.g., isopycnal coordinates), the fluid remains embedded within a background Euclidean space. Through this embedding, the geometry local to the surface inherits features from the background flat Euclidean space such as how to measure distance between points. A further simplification arises since we make use of universal Newtonian time. We thus only require general tensors for the spatial coordinates. Time remains untouched. Therefore, our mathematical needs are far simpler than the general relativist.

7.2 Covariance of physical relations

Physical relations are independent of subjective choices for their mathematical representations. This principle motivates us to seek mathematical expressions between objects whose meaning transcends a particular coordinate representation. This is the essence of *covariance*. In practice, covariance means that a mathematical expression of a physical relation is form invariant when expressed as relations between geometric objects such as points, vectors, and tensors.

Although physics does not care about coordinates, physicists often do. Namely, it is convenient, and sometimes necessary, to work with specific coordinates suited to the symmetry of the physical system. After deriving a physical law in one set of coordinates, it is often of interest to establish the form of the law in another set of coordinates. How does the physical law, typically represented as a differential equation, transform into other coordinates? So long as the equations are written in a proper tensorial form, in which they exemplify covariance, then the equations are form invariant. Consequently, “physics as geometry” has a major practical implication. Namely, we can establish the validity of a physical relation in any convenient set of coordinates, and then extend that relation to all coordinates.

7.2.1 Tensor operations retaining covariance

Extending a mathematical equation to all coordinates requires the equations to respect covariance. Operationally, general covariance means that all tensor indices are properly matched and each derivative is covariant (as specified in Section 9.2). In chapter 9, we provide the details needed to understand general covariance. In this chapter we outline the procedure. The elegance and power rendered by general covariance is the key reason that tensor analysis is ubiquitous in theoretical physics.

To ensure an equation respects covariance requires us to understand certain properties of tensors and operations with tensors that produces components of new tensors. We here summarize the specific properties characterizing covariance (taken after page 153 of *Schutz, 1985*):

1. Manipulations of tensor components are called *permissible tensor operations* if they produce components of new tensors. The following are permissible operations:
 - (a) Multiplication of a tensor by a scalar produces a new tensor of the same type.
 - (b) Addition of components of two tensors of the same type gives components of a new tensor of the same type. In particular, only tensors of the same type can be equal.
 - (c) Multiplication of components of two tensors of arbitrary type gives components of a new tensor whose type is given by the sum of the types for the individual tensors. This operation is called the *outer product* and is denoted by the operator \otimes .
 - (d) Covariant differentiation (Section 9.2) increases by one the order of a tensor, with the covariant derivative operator denoted by ∇ .
 - (e) Contraction on a pair of indices of the components of a tensor reduces by one the order of a tensor.
2. If two tensors of the same type have equal components in a given coordinate system, then they have equal components in all coordinate systems. Hence, the tensors are identical.
3. If a mathematical equation consists of tensors combined only by the permissible tensor operations, and if the equation is true in one coordinate system, then it is true in any coordinate system. If the equations involve covariant derivatives, then the equations remain form invariant under changes in coordinates. For the partial differential equations of geophysical fluid mechanics, covariant differentiation is the key to general covariance.

7.2.2 Comments

The remainder of this chapter, as well as Chapters 8 and 9, provide the needed details for supporting the above notions of covariance. Even without penetrating these details, the reader should be able to appreciate why covariance is so central to physics.

7.3 Points and coordinates

Consider a spatial point \mathcal{P} at a particular time τ . As time progresses, the point traces out a curve in space-time. We call that curve a *trajectory*. The trajectory could be of a point particle following the fluid flow, thus defining the Lagrangian reference frame (Section 16.1). Or it could trace the path of something else such as a fish, balloon, boat, or airplane. As the trajectory is a one-dimensional curve, it is specified mathematically by a single parameter. We choose the time measured by an observer on the trajectory for this parameter, in which case the trajectory is written $\mathcal{P}(\tau)$.

A point in the fluid and its trajectory in space-time are geometric objects that exist independently of any coordinate representation. Even so, we find the need to represent points, trajectories, vectors, and other geometric objects using coordinates. For example, coordinates are needed to make quantitative statements about fluid flow in relation to other observers. What is its speed and direction relative to a chosen reference frame? What is the distance from an origin or from another particle? Tensor analysis provides a formalism that enables us to answer such quantitative questions while maintaining a clear view on the underlying physics and geometry.

7.3.1 Time as a parameter and time as a coordinate

In special and general relativity, there is a mixing of space and time that warrants the use of four-dimensional space-time tensor analysis. In contrast, for classical mechanics forming the foundation of geophysical fluid mechanics, time remains numerically the same throughout space. We thus make use of the same universal (or Newtonian) time since the fluid velocity and wave speeds are far smaller than the speed of light. Correspondingly, we make use of general tensor analysis only for the description of points in space.

The time parameter, τ , specifies a point along a trajectory. The coordinate time, t , measures time for all positions throughout space. This distinction between the time parameter and time coordinate is pedantic given that $\tau = t$ in a Newtonian universe. Nonetheless, it is convenient to make the distinction when measuring how fluid properties change since these changes are subject to motion of the observer. For example, changes following a trajectory, found by computing the trajectory time derivative $\partial/\partial\tau$, are generally distinct from changes found by computing the time derivative $\partial/\partial t$, in which the spatial coordinates are held fixed. When the trajectory is defined by a fluid particle, we refer to $\partial/\partial\tau$ as the material or Lagrangian time derivative. If the spatial coordinates are fixed in space, then $\partial/\partial t$ is an Eulerian time derivative. When alternative spatial coordinates are used, some of which can move (see Section 7.3.2), then $\partial/\partial t$ is perhaps a mixture of Lagrangian and Eulerian or perhaps neither.

7.3.2 The importance of index placement

Much of the formalism of general tensor analysis builds from Cartesian representations of vector and matrix analysis, with generalizations that provide objective statements independent of coordinates. One key point of distinction from Cartesian tensors is that the position (up or down) of a tensor label has significance in general tensor analysis. We follow the standard convention by labeling an arbitrary coordinate as ξ^a . The upper “contra-variant” position of the label is not an exponent. Rather, it is a label running from $a = 1, 2, 3$ for the three dimensional space of Newtonian mechanics. Notably, we refer to ξ^a as the “spatial” coordinates even if they are not traditional coordinates for a point in space. We clarify this comment in the examples of Section 7.4.

7.4 Example coordinate descriptions

We here offer a few examples of coordinates used for describing geophysical fluid systems.

7.4.1 Eulerian coordinates

The Cartesian coordinates for a point are written

$$\xi^a = (x, y, z) \quad \text{Cartesian}, \quad (7.1)$$

whereas for spherical coordinates we write (see Figure 10.1)

$$\xi^a = (r, \lambda, \phi) \quad \text{spherical} \quad (7.2)$$

and polar cylindrical coordinates

$$\xi^a = (r, \lambda, z) \quad \text{cylindrical}. \quad (7.3)$$

These coordinates identify fixed positions in space. We can use these *Eulerian* coordinates to mark the trajectory $\mathcal{P}(\tau)$ as it crosses the spatial point ξ^a at time t . We provide a more complete discussion of Eulerian coordinates in Section 16.2.

As shown in Section 7.5, Cartesian coordinates are notable for having basis vectors maintaining a fixed direction throughout space. This feature lends much simplicity to Cartesian coordinates and its corresponding Cartesian tensor analysis (Chapters 3 and 4). In contrast, the spherical basis vectors are spatially dependent. Likewise, the radial and angular basis vectors for polar cylindrical coordinates are spatially dependent, whereas the vertical direction is fixed. Additionally, the spherical and cylindrical coordinates do not all have the same physical dimensions. Each of these features of spherical and cylindrical coordinates places them outside the purview of Cartesian tensor analysis.

7.4.2 Isopycnal coordinates

In geophysical fluids that are stably stratified in the vertical, it is common to measure the vertical position of a fluid element by specifying its entropy, buoyancy, or potential density depending on the application. We generically write these *isentropic*, *buoyancy*, or *isopycnal* coordinates as

$$\xi^a = (x, y, b) \quad \text{isopycnal coordinates}, \quad (7.4)$$

where $b = b(x, y, z, t)$ is a generic symbol for entropy, buoyancy, or potential density. Entropy, buoyancy, and potential density are materially invariant for perfect fluid flow (flow absent irreversible processes such as mixing). Hence, all fluid particle motion occurs on surfaces of constant b . Under such perfect fluid conditions, isopycnal coordinates are of great use for describing fluid mechanics of stably stratified geophysical flows.

The isopycnal coordinate is generally not orthogonal to the horizontal coordinates x, y . Hence, even if the horizontal coordinates are Cartesian, the use of b to measure the vertical precludes the use of Cartesian tensor analysis. Furthermore, we note the distinct physical dimensions of the three spatial coordinates (x, y, b) , again necessitating the use of general tensor analysis.

7.4.3 Lagrangian or material coordinates

We often conceive of a fluid as a continuum of constant mass fluid elements distinguished by continuum marker coordinates or labels. The initial position for a fluid element offers a common choice for these *material coordinates*. The fluid dynamical equations of motion (i.e., Newton's Law of motion) can be formulated using material coordinates so long as the material coordinate maintains a one-to-one relation to points in space. This kinematical framework is termed *Lagrangian* or *material*. The resulting dynamical equations share much in common with Newtonian particle mechanics, though with the added feature of contact forces acting between the fluid elements. We provide a discussion of Lagrangian coordinates in Section 16.2.

If we represent material coordinates by the Cartesian positions of fluid elements at an arbitrary initial time, then we can make use of Cartesian tensor analysis. However, it is sometimes useful to make use of alternative markers. One example is the isopycnal coordinate mentioned above, whose value remains invariant under perfect fluid motion. In this manner, we refer to the isopycnal coordinates as "quasi-Lagrangian" since its vertical coordinate follows the vertical position of an perfect fluid parcel whereas its horizontal coordinates are Eulerian.

7.4.4 Tracer coordinates

Consider a trio of linearly independent tracer concentrations $C^a = C^a(x, y, z, t)$ (we introduce tracers in Section 18.1). Linear independence means that for any point in space there is a unique intersection of three constant tracer surfaces, so that we can uniquely determine a point in space by specifying the value for the three tracer concentrations. We can thus use tracer concentrations as the spatial coordinates

$$\xi^a = (C^1, C^2, C^3). \quad (7.5)$$

In some cases there are only two linearly independent tracers, in which case the two may be used in combination with a third spatial coordinate such as depth or pressure. Furthermore, the case of one tracer coordinate reduces to the isopycnal coordinate system described above.

7.5 The velocity vector and basis vectors

Consider two points along a trajectory separated by the infinitesimal time increment $d\tau$. The velocity vector for this trajectory is defined by

$$\vec{v}(\tau) = \lim_{\Delta\tau \rightarrow 0} \frac{\mathcal{P}(\tau + \Delta\tau/2) - \mathcal{P}(\tau - \Delta\tau/2)}{\Delta\tau} \quad (7.6a)$$

$$= \frac{d\mathcal{P}(\tau)}{d\tau}. \quad (7.6b)$$

The velocity is a vector pointing in the direction determined by the difference between two points on a trajectory, in the limit as the time separation between the points vanishes. Consequently, the velocity points in a direction tangent to the trajectory. Notably, the above definition for the velocity makes no use of coordinates. Rather, the velocity vector is determined by the geometry of the trajectory and the specification of the trajectory's time parameter. Hence, velocity is fundamentally an arrow with a length and direction; i.e., it is a geometric object.

The definition of velocity as a vector tangent to the trajectory is a general property of all vectors living on a manifold. Namely, a vector at a point on a manifold lives within the tangent plane to the manifold at that point. This observation signals to us the need to be very careful when comparing vectors on a curved manifold such as the sphere or an isopycnal surface.

7.5.1 Coordinate representation

We now establish an arbitrary set of coordinates, ξ^a , to represent points in space. These coordinates are used to measure the spatial position of the trajectory according to

$$\mathcal{P}(\tau) = \mathcal{P}[\xi^a(\tau)], \quad (7.7)$$

where $\xi^a(\tau)$ is the coordinate position on the trajectory at trajectory time τ . This coordinate representation for the trajectory induces a coordinate representation for the velocity through use of the chain rule

$$\vec{v}(\tau) = \frac{d\mathcal{P}(\tau)}{d\tau} \quad (7.8a)$$

$$= \frac{\partial \mathcal{P}}{\partial \xi^a} \frac{d\xi^a}{d\tau} \quad (7.8b)$$

$$\equiv \vec{e}_a v^a. \quad (7.8c)$$

The expansion coefficients

$$v^a = \frac{d\xi^a}{d\tau} \quad (7.9)$$

provide a representation the velocity vector $\vec{v}(\tau)$ within the coordinate system ξ^a .

7.5.2 Basis vectors

For each number v^a there is a corresponding basis vector \vec{e}_a defined by

$$\vec{e}_a = \frac{\partial \mathcal{P}}{\partial \xi^a}. \quad (7.10)$$

The basis vectors are generally a function of position and time, although they are constant for Cartesian coordinates.

7.6 Notational conventions

We here introduce notational conventions that help to simplify many of the manipulations (“index gymnastics”) encountered with general tensors.

7.6.1 Placement of tensor labels

As indicated in Section 7.3.2, the placement of tensor labels has specific meaning with general tensor analysis. It is therefore critical to maintain proper usage to ensure “conservation of labels” across an equals sign. As a first example of this usage, notice how the basis vectors in equation (7.10) inherit a lowered tensor label. This placement follows from the partial derivative operator that carries an upper coordinate label in the denominator of the operator.

7.6.2 Einstein summation convention

For general tensors, the Einstein summation convention assumes that labels are summed over their range when a lower (covariant) label matches an upper (contra-variant) label. In this way we have

$$\vec{v}(\tau) = \sum_{a=1}^3 \vec{e}_a v^a = \vec{e}_a v^a. \quad (7.11)$$

This rule generalizes that used for Cartesian tensors in Chapter 3. For general tensors, contraction is between a lower and an upper label. We return to such contractions in Section 8.2, where we show that the contraction of a vector and a one-form renders a scalar.

7.6.3 The boldface notation

In Cartesian tensors we have no need to distinguish \vec{v} from the boldface \mathbf{v} . One means to extend this notation to general tensors is to organize the velocity vector representation v^a into an ordered list (v^1, v^2, v^3) and to use the boldface notation

$$\mathbf{v} = (v^1, v^2, v^3). \quad (7.12)$$

Likewise, we can organize the basis vectors according to

$$\vec{\mathbf{e}} = (\vec{e}_1, \vec{e}_2, \vec{e}_3) \quad (7.13)$$

and the coordinates as

$$\boldsymbol{\xi} = (\xi^1, \xi^2, \xi^3). \quad (7.14)$$

With this notation the velocity vector representation from equation (7.8c) takes on the form

$$\vec{v} = \vec{e}_a v^a = \vec{\mathbf{e}} \cdot \mathbf{v}. \quad (7.15)$$

Likewise, a trajectory can be represented in terms of a chosen set of coordinates according to

$$\mathcal{P} = \vec{e}_a \xi^a = \vec{\mathbf{e}} \cdot \boldsymbol{\xi}. \quad (7.16)$$

Notice that the arrow symbol over the basis vector remains even when using the boldface. This usage is required since the arrow carries information about the vector nature of the object, whereas the boldface is merely a shorthand for an ordered list.

Although the above notation makes good sense, we more generally allow the boldface to be synonymous with the vector arrow notation

$$\mathbf{v} = \vec{v} = \vec{e}_a v^a. \quad (7.17)$$

Though corrupting the convention in equation (7.15), this notation is readily used in the literature and as such will be employed in this book where the meaning is clear and after our tensor brain-muscle has been exercised.

8

General tensor algebra

In this chapter we develop the algebra of general tensors. General tensor algebra is very similar to Cartesian tensors (Chapter 3), requiring only a modest amount of further effort and precision. Material in this chapter represents a streamlined version of Chapter 20 from [Griffies \(2004\)](#). Other resources include the lucid treatment of tensors for fluid mechanics given by [Aris \(1962\)](#) and that of [Thorne and Blandford \(2017\)](#).

READER'S GUIDE TO THIS CHAPTER

This chapter dives into the details of general tensor analysis. It is necessary for understanding calculus on curved manifolds detailed in Chapter 9, with applications to the mathematics of generalized vertical coordinates discussed in Chapter 11. We assume here a mastery of Cartesian tensor algebra in Chapter 3. Many readers may choose to skip this chapter and come back later to fill in details if needed.

8.1	The metric tensor and coordinate transformations	98
8.1.1	Cartesian coordinates in Euclidean space	98
8.1.2	The metric as a symmetric second order tensor	99
8.1.3	Coordinate representation of the metric tensor	99
8.1.4	Transforming the representation of the metric tensor	99
8.1.5	Finite distance between points	100
8.2	One-forms	100
8.2.1	Coordinate representation of a one-form	100
8.2.2	Basis one-forms and the orthogonality relation	101
8.2.3	Metric as a mapping between vectors and one-forms	101
8.2.4	Transformation of the coordinate representation	101
8.3	Scalar product	102
8.4	The volume element and Jacobian of transformation	102
8.4.1	Jacobian of transformation	103
8.4.2	Relating the Jacobian to the determinant of the metric	103
8.5	The permutation symbol and the determinant	104
8.5.1	Connecting the permutation symbol to the determinant	104
8.5.2	Further identities satisfied by the determinant	104
8.6	The Levi-Civita tensor and the volume element	105
8.6.1	General coordinate representation of the Levi-Civita tensor	105
8.6.2	The Levi-Civita tensor and the volume element	105
8.7	Cross product and biorthogonality relation	106

8.1 The metric tensor and coordinate transformations

In the study of fluid mechanics we find the need to measure the distance between two points in space at a particular time instance. Since we assume all points live on a smooth manifold (e.g., a sphere, an isopycnal), it is sufficient to consider the distance between two infinitesimally close points and use integration to measure finite distances. The measurement of distance requires a metric tensor, which is the topic of this section.

8.1.1 Cartesian coordinates in Euclidean space

Consider a Cartesian coordinate representation for the position of two points, with point \mathcal{P} having coordinates $\xi^a = x^a$ and the other point \mathcal{Q} an infinitesimal distance away at $x^a + dx^a$. Furthermore, let $d\vec{x} = dx^a \vec{e}_a$ be the infinitesimal vector pointing from \mathcal{P} to \mathcal{Q} . Since the space is Euclidean, the squared distance between the two points is based on the Euclidean norm; i.e., the familiar scalar or dot product (Section 3.3)

$$ds^2 = d\vec{x} \cdot d\vec{x} = \vec{e}_a \cdot \vec{e}_b dx^a dx^b = \delta_{ab} dx^a dx^b. \quad (8.1)$$

In this expression,

$$(ds)^2 \equiv ds^2 \quad (8.2)$$

is the squared infinitesimal arc-length separating the two points. The Cartesian representation of the Kronecker symbol, δ_{ab} , is symmetric

$$\delta_{ab} = \delta_{ba}, \quad (8.3)$$

and vanishes when $a \neq b$ and is unity when $a = b$.

8.1.2 The metric as a symmetric second order tensor

As defined by equation (8.1), δ_{ab} forms the Cartesian representation of the *metric tensor* for Euclidean space. The metric is a second order tensor, meaning that its coordinate representation carries two tensor labels. Contracting the metric tensor with two vectors leads to a number, namely the squared distance between the two points. Hence, the metric establishes the means to measure the distance between two points that live on a manifold.

We write this distance-measuring property of the metric tensor in a geometric manner through

$$\text{distance}(\vec{P}, \vec{Q}) = \sqrt{\mathcal{G}(\vec{P}, \vec{Q})}. \quad (8.4)$$

Here, \mathcal{G} is the metric tensor with coordinate representation g_{ab} and \vec{P}, \vec{Q} are infinitesimally close vectors with coordinate representations

$$\vec{P} = \xi^a \vec{e}_a \quad \vec{Q} = \vec{P} + d\xi^a \vec{e}_a. \quad (8.5)$$

Equation (8.4) indicates that the metric tensor takes two vectors as argument and produces a scalar. Furthermore, since

$$\text{distance}(\vec{P}, \vec{Q}) = \text{distance}(\vec{Q}, \vec{P}) \geq 0, \quad (8.6)$$

the metric tensor is a symmetric and positive tensor that produces zero only when $\vec{P} = \vec{Q}$.

8.1.3 Coordinate representation of the metric tensor

Given the geometric expression (8.4) for the metric, we determine its representation in an arbitrary coordinate system by considering the squared distance between the coordinate basis vectors

$$\text{distance}(\vec{e}_a, \vec{e}_b) = \sqrt{\mathcal{G}(\vec{e}_a, \vec{e}_b)}. \quad (8.7)$$

This relation determines the coordinate components of the metric tensor

$$\mathcal{G}(\vec{e}_a, \vec{e}_b) \equiv g_{ab}. \quad (8.8)$$

Furthermore, in Euclidean space this relation is written

$$g_{ab} = \vec{e}_a \cdot \vec{e}_b. \quad (8.9)$$

In this manner we see that the basis vectors determine the metric tensor components. Note that if the basis vectors are orthogonal, then the metric tensor components vanish unless $a = b$.

8.1.4 Transforming the representation of the metric tensor

We find many opportunities to represent the metric tensor in various coordinate systems. Here, we consider the transformation from Cartesian coordinates $\xi^a = x^a$ to arbitrary coordinates $\xi^{\bar{a}}$. Use of the chain rule leads to the equivalent expression for the squared infinitesimal length,

$$ds^2 = \delta_{ab} d\xi^a d\xi^b \quad (8.10a)$$

$$= \delta_{ab} \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} \frac{\partial \xi^b}{\partial \xi^{\bar{b}}} d\xi^{\bar{a}} d\xi^{\bar{b}} \quad (8.10b)$$

$$= \delta_{ab} \Lambda_{\bar{a}}^a \Lambda_{\bar{b}}^b d\xi^{\bar{a}} d\xi^{\bar{b}} \quad (8.10c)$$

$$= g_{\bar{a}\bar{b}} d\xi^{\bar{a}} d\xi^{\bar{b}}, \quad (8.10d)$$

where

$$g_{\bar{a}\bar{b}} = \delta_{ab} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \quad (8.11)$$

defines the components to the metric tensor as represented by the new set of coordinates $\xi^{\bar{a}}$. We also introduced elements to the transformation operator

$$\Lambda^a_{\bar{a}} = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}}. \quad (8.12)$$

When organized as a matrix, we let the row be denoted by a and columns by \bar{a} .¹ The transformation operator is nonsingular for one-to-one invertible coordinate transformations, in which case its determinant, called the *Jacobian of the transformation*, is nonvanishing and single signed.

8.1.5 Finite distance between points

Once the metric is determined, the distance along a curve between two finitely separated points is given by the integration

$$\begin{aligned} L &= \int \sqrt{ds^2} \\ &= \int_{\varphi_1}^{\varphi_2} \left| g_{ab} \frac{d\xi^a}{d\varphi} \frac{d\xi^b}{d\varphi} \right|^{1/2} d\varphi, \end{aligned} \quad (8.13)$$

where φ is a parameter specifying the curve (e.g., the arc length as in Section 4.4), and $\varphi_{1,2}$ are the curve's endpoints.

8.2 One-forms

The metric tensor \mathcal{G} is a function of two vectors. When the metric “eats” the two vectors, the result is the scalar distance between the vectors (equation (8.4))

$$\text{distance}(\vec{A}, \vec{B}) = \sqrt{\mathcal{G}(\vec{A}, \vec{B})}. \quad (8.14)$$

What if the metric only eats one vector? The resulting geometric object is known as a one-form

$$\tilde{A} \equiv \mathcal{G}(\vec{A},), \quad (8.15)$$

with the tilde used to distinguish a one-form from a vector.

8.2.1 Coordinate representation of a one-form

We can determine the coordinate representation of a one-form by eating a basis vector

$$\tilde{A}(\vec{e}_b) = \mathcal{G}(\vec{A}, \vec{e}_b) \quad (8.16a)$$

$$= \mathcal{G}(A^a \vec{e}_a, \vec{e}_b) \quad (8.16b)$$

$$= \mathcal{G}(\vec{e}_a, \vec{e}_b) A^a \quad (8.16c)$$

$$= g_{ab} A^a. \quad (8.16d)$$

¹We generally follow the convention of displacing the lower index to the right to help keep track of which index refers to the column.

To reach this result we pulled the coordinate representation A^a outside of the metric tensor since the tensor eats vectors rather than numbers. This equation defines the coordinate representation of the one-form \tilde{A} in terms of its *dual* vector \vec{A} and the metric tensor

$$A_b = g_{ab} A^a. \quad (8.17)$$

8.2.2 Basis one-forms and the orthogonality relation

Just as for vectors, we find use for a basis of one-forms to specify their coordinate representation. The basis one-form, \tilde{e}^a , are defined through the orthogonality relation

$$\mathcal{G}(\tilde{e}^a, \vec{e}_b) = \tilde{e}^a \cdot \vec{e}_b = \delta_b^a, \quad (8.18)$$

where

$$\delta_b^a = g^{ac} g_{cb} \quad (8.19)$$

are components to the Kronecker delta tensor, taking the value of unity when $a = b$ and zero otherwise

$$\delta_b^a = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (8.20)$$

It is only for Cartesian coordinates that we have

$$\delta_c^a = g^{ab} \delta_{bc} \quad \text{Cartesian coordinates,} \quad (8.21)$$

which follows since $g^{ab} = \delta^{ab}$ in Cartesian coordinates.

8.2.3 Metric as a mapping between vectors and one-forms

We can contract the expression (8.17) with components of the inverse metric tensor, g^{ab} , to render

$$g^{ab} A_b = g^{ab} g_{bc} A^c \quad (8.22a)$$

$$= \delta_c^a A^c \quad (8.22b)$$

$$= A^a. \quad (8.22c)$$

This identity, as well as equation (8.17), show that the metric provides a map between coordinate representations of one-forms and vectors.

In general, to every vector \vec{A} there is a corresponding one-form \tilde{A} . We say that the one-forms and vectors are *dual*, with mapping between one-forms and vectors rendered by the metric tensor. In Cartesian tensor analysis, duality between one-forms and vectors becomes the duality between row vectors and column vectors. Furthermore, as for Cartesian tensors, we construct an inner product by contracting one-forms and vectors to produce a scalar. Finally, the duality relation given by equation (8.17) offers us the means to raise and lower tensor indices in a manner akin to the transpose operation in linear algebra that produces a row vector from a column vector.

8.2.4 Transformation of the coordinate representation

The transformation matrix (8.12) provides the means to convert any arbitrary coordinate representation of a tensor from one coordinate system to another. For example, consider the coordinate representation of a vector, which is realized by letting the vector eat one of the basis one-forms

$$\vec{F}(\tilde{e}^a) = F^a. \quad (8.23)$$

Now consider another coordinate system with basis one-forms $\tilde{e}^{\bar{a}}$, so that the vector has a representation

$$\vec{F}(\tilde{e}^{\bar{a}}) = F^{\bar{a}}. \quad (8.24)$$

Transforming the basis one-form using the transformation matrix leads to

$$F^{\bar{a}} = \vec{F}(\tilde{e}^{\bar{a}}) = \vec{F}(\Lambda_a^{\bar{a}} \tilde{e}^a) = \Lambda_a^{\bar{a}} \vec{F}(\tilde{e}^a) = \Lambda_a^{\bar{a}} F^a. \quad (8.25)$$

Transformation of an arbitrary one-form representation takes place with the inverse transformation matrix

$$F_{\bar{a}} = \tilde{F}(\vec{e}_{\bar{a}}) = \tilde{F}(\Lambda_{\bar{a}}^a \vec{e}_a) = \Lambda_{\bar{a}}^a \tilde{F}(\vec{e}_a) = \Lambda_{\bar{a}}^a F_a. \quad (8.26)$$

8.3 Scalar product

In Section 3.3.2 we defined the scalar product between two Cartesian vectors. The natural generalization is given by

$$\vec{P} \cdot \vec{Q} = P^a Q^b \vec{e}_a \cdot \vec{e}_b = P^a Q^b g_{ab} = P^a Q_a = P_b Q^b, \quad (8.27)$$

where the second equality made use of the metric tensor coordinate representation given by equation (8.9). We can conceive of the scalar product in a somewhat more general manner by recalling that a one-form operates on a vector, $\tilde{P}(\vec{Q})$. Conversely, a vector operates on a one-form, $\vec{Q}(\tilde{P})$. Exposing components leads to

$$\tilde{P}(\vec{Q}) = \tilde{P}(Q^a \vec{e}_a) = Q^a \tilde{P}(\vec{e}_a) = Q^a P_a, \quad (8.28)$$

which equals to

$$\vec{Q}(\tilde{P}) = \vec{Q}(P_a \tilde{e}^a) = P_a \vec{Q}(\tilde{e}^a) = P_a Q^a. \quad (8.29)$$

The scalar product is invariant to coordinate changes, as seen through

$$\vec{Q}(\tilde{P}) = \vec{Q}(P_a \tilde{e}^a) = P_a Q^a = \vec{Q}(P_{\bar{a}} \tilde{e}^{\bar{a}}) = P_{\bar{a}} Q^{\bar{a}}. \quad (8.30)$$

The invariance is also revealed by working just with the coordinate representations and introducing the transformation matrix elements

$$P_a Q^a = (\Lambda_a^{\bar{a}} P_{\bar{a}}) (\Lambda_{\bar{b}}^a Q^{\bar{b}}) = \Lambda_a^{\bar{a}} \Lambda_{\bar{b}}^a P_{\bar{a}} Q^{\bar{b}} = \delta_{\bar{b}}^{\bar{a}} P_{\bar{a}} Q^{\bar{b}} = P_{\bar{a}} Q^{\bar{a}}. \quad (8.31)$$

8.4 The volume element and Jacobian of transformation

Recall from Section 3.5.2 that we derived an expression for the volume of an infinitesimal region of Euclidean space \mathbb{R}^3 using Cartesian coordinates

$$dV = dx dy dz (\hat{x} \wedge \hat{y}) \cdot \hat{z} = dx dy dz. \quad (8.32)$$

This volume element is used for integrating over a region of \mathbb{R}^3 when using Cartesian coordinates. Furthermore, its material fluid expression measures the volume of a fluid element. We now generalize this result to arbitrary coordinates.

8.4.1 Jacobian of transformation

From multi-variate calculus, the relation between $d\xi^1 d\xi^2 d\xi^3$ and $d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}$ for two sets of coordinates is given by

$$d\xi^1 d\xi^2 d\xi^3 = \frac{\partial(\xi^1, \xi^2, \xi^3)}{\partial(\xi^{\bar{1}}, \xi^{\bar{2}}, \xi^{\bar{3}})} d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}} \quad (8.33a)$$

$$= \det(\Lambda^a_{\bar{a}}) d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}, \quad (8.33b)$$

where $\det(\Lambda^a_{\bar{a}})$ is the determinant of the transformation matrix, also known as the *Jacobian of transformation*. The transformation is well defined so long as the Jacobian does not vanish. We maintain labels on the transformation matrix inside the determinant symbol to help indicate the sense for the transformation. This notation also helps maintain proper conservation of tensor indices.

8.4.2 Relating the Jacobian to the determinant of the metric

Recall the expression (8.10d) for the transformation of the metric

$$g_{\bar{a}\bar{b}} = \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} g_{ab}. \quad (8.34)$$

We can write this expression as a matrix equation

$$\bar{\mathcal{G}} = \Lambda^T \mathcal{G} \Lambda \quad (8.35)$$

where Λ^T is the transposed matrix. This equation is valid upon taking determinants of both sides so that

$$\det(\bar{\mathcal{G}}) = \det(\Lambda^T \mathcal{G} \Lambda) \quad (8.36a)$$

$$= \det(\Lambda^T) \det(\mathcal{G}) \det(\Lambda) \quad (8.36b)$$

$$= [\det(\Lambda)]^2 \det(\mathcal{G}). \quad (8.36c)$$

To reach this result we used the property of determinants that $\det(AB) = \det(A)\det(B)$ for any two matrices, and $\det(\Lambda^T) = \det(\Lambda)$. Consequently,

$$\det(\Lambda^a_{\bar{a}}) = \frac{\sqrt{\det(g_{\bar{a}\bar{b}})}}{\sqrt{\det(g_{ab})}} = \frac{\sqrt{\det(\bar{\mathcal{G}})}}{\sqrt{\det(\mathcal{G})}}. \quad (8.37)$$

We are thus led to the equivalent expressions for the volume element

$$dV \equiv \sqrt{\det(\mathcal{G})} d\xi^1 d\xi^2 d\xi^3 = \sqrt{\det(\bar{\mathcal{G}})} d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}. \quad (8.38)$$

This relation provides us with our desired general coordinate expression for the volume element. For the special case when the unbarred coordinates are Cartesian, $g_{ab} = \delta_{ab}$ so that $\det(\mathcal{G}) = 1$ and

$$\det(\Lambda^a_{\bar{a}}) = \sqrt{\det(g_{\bar{a}\bar{b}})} \quad \text{unbarred coordinates are Cartesian.} \quad (8.39)$$

This is a rather useful expression for our purposes, since we can always use Cartesian as the unbarred coordinates given that geophysical fluids move in a background Euclidean space.

8.5 The permutation symbol and the determinant

As discussed in Section 3.4.1, the Cartesian components of the Levi-Civita tensor are given by the permutation symbol, ϵ_{abc} . To help determine the general coordinate representation of the Levi-Civita tensor, we here develop some identities satisfied by the determinant of the transformation matrix.

8.5.1 Connecting the permutation symbol to the determinant

Consider a two-dimensional space with a transformation matrix $\Lambda^a_{\bar{a}}$ between two sets of coordinates with $a = 1, 2$. The determinant of the transformation is given by

$$\det(\Lambda^a_{\bar{a}}) = \Lambda^1_{\bar{1}} \Lambda^2_{\bar{2}} - \Lambda^1_{\bar{2}} \Lambda^2_{\bar{1}}. \quad (8.40)$$

Introducing the permutation symbol ϵ_{ab} allows us to write this expression in a more tidy manner

$$\det(\Lambda^a_{\bar{a}}) = \epsilon_{ab} \Lambda^a_{\bar{1}} \Lambda^b_{\bar{2}} \quad (8.41)$$

with

$$\epsilon_{12} = 1 \quad \epsilon_{21} = -1. \quad (8.42)$$

The permutation symbol is defined to have numerically the same values whether the labels are raised or lowered: $\epsilon^{ab} = \epsilon_{ab}$.

We can generalize the above to any number of dimensions, each of which adds one more label to the permutation symbol and one more number added to the permutation string. We already encountered the three dimensional version in Section 3.4.1 when discussing the vector cross product, in which case the permutation symbol is

$$\epsilon_{123} = 1 \quad (8.43a)$$

$$\epsilon_{abc} = \begin{cases} 0 & \text{if any two labels are the same,} \\ 1 & \text{if } a, b, c \text{ is an even permutation of } 1, 2, 3, \\ -1 & \text{if } a, b, c \text{ is an odd permutation of } 1, 2, 3. \end{cases} \quad (8.43b)$$

Likewise, the determinant of the transformation matrix takes the form

$$\det(\Lambda^a_{\bar{a}}) = \frac{\partial(\xi^1, \xi^2, \xi^3)}{\partial(\bar{\xi}^1, \bar{\xi}^2, \bar{\xi}^3)} = \frac{\partial \boldsymbol{\xi}}{\partial \bar{\boldsymbol{\xi}}} = \epsilon_{abc} \Lambda^a_{\bar{1}} \Lambda^b_{\bar{2}} \Lambda^c_{\bar{3}}. \quad (8.44)$$

8.5.2 Further identities satisfied by the determinant

The following identity in two dimensions can be readily verified through enumeration

$$\epsilon_{ab} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} = \epsilon_{\bar{a}\bar{b}} \det(\Lambda^a_{\bar{a}}). \quad (8.45)$$

It follows directly from the definition of the determinant and can be explicitly verified so long as we assume the permutation symbol $\epsilon_{\bar{a}\bar{b}}$ is numerically identical to ϵ_{ab} . Now contract both sides of this relation with $\epsilon^{\bar{a}\bar{b}}$ to isolate the determinant

$$\frac{1}{2} \epsilon^{\bar{a}\bar{b}} \epsilon_{ab} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} = \det(\Lambda^a_{\bar{a}}), \quad (8.46)$$

where we used

$$\epsilon^{\bar{a}\bar{b}} \epsilon_{\bar{a}\bar{b}} = \epsilon^{\bar{1}\bar{2}} \epsilon_{\bar{1}\bar{2}} + \epsilon^{\bar{2}\bar{1}} \epsilon_{\bar{2}\bar{1}} = 2. \quad (8.47)$$

The three dimensional version takes the form

$$\epsilon_{abc} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} = \epsilon_{\bar{a}\bar{b}\bar{c}} \det(\Lambda^a_{\bar{a}}), \quad (8.48)$$

along with

$$\frac{1}{3!} \epsilon^{\bar{a}\bar{b}\bar{c}} \epsilon_{abc} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} = \det(\Lambda^a_{\bar{a}}). \quad (8.49)$$

8.6 The Levi-Civita tensor and the volume element

The metric tensor introduced in Section 8.1 provides a means to measure distance between two points. The Levi-Civita tensor allows us to compute volumes (or areas for two dimensional manifolds). We make particular use of this tensor to compute the volume element used for integration. This section generalizes the Cartesian coordinate discussion provided in Section 3.5.3.

8.6.1 General coordinate representation of the Levi-Civita tensor

The relations (8.45) and (8.48) indicate that the permutation symbol *does not* transform as the components to a second order covariant tensor, unless the determinant of the transformation is unity. Unit determinants occur for special transformations, such as rotations (i.e., Cartesian to Cartesian coordinate transformation as in Chapter 3) and the identity transformation. However, they are not unity in general, which motivates us to introduce the general coordinate form of the *Levi-Civita tensor*

$$\varepsilon_{abc} = \sqrt{\det(\mathcal{G})} \epsilon_{abc}. \quad (8.50)$$

We highlight the distinct symbols in this definition, with ε the Levi-Civita tensor and ϵ the permutation symbol. By construction, the Levi-Civita tensor components transform as

$$\Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} \varepsilon_{abc} = \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} \sqrt{\det(\mathcal{G})} \epsilon_{abc} \quad (8.51a)$$

$$= \sqrt{\det(\mathcal{G})} \epsilon_{\bar{a}\bar{b}\bar{c}} \det(\Lambda^a_{\bar{a}}) \quad (8.51b)$$

$$= \sqrt{\det(\mathcal{G})} \epsilon_{\bar{a}\bar{b}\bar{c}} \quad (8.51c)$$

$$= \varepsilon_{\bar{a}\bar{b}\bar{c}}, \quad (8.51d)$$

where equations (8.37) and (8.45) were used. Therefore, ε_{abc} transforms as components to a third order covariant tensor. Likewise,

$$\varepsilon^{abc} = \frac{\epsilon^{abc}}{\sqrt{\det(\mathcal{G})}} \quad (8.52)$$

transforms as the components to a third order contravariant tensor. These transformation rules allow us to identify ε as a tensor rather than just a combination of numbers.

8.6.2 The Levi-Civita tensor and the volume element

As a third order tensor, the Levi-Civita tensor takes three vectors as its argument. In particular, for three infinitesimal vectors we have

$$\varepsilon(\vec{e}_1 d\xi^1, \vec{e}_2 d\xi^2, \vec{e}_3 d\xi^3) = d\xi^1 d\xi^2 d\xi^3 \varepsilon(\vec{e}_1, \vec{e}_2, \vec{e}_3) \quad (8.53a)$$

$$= d\xi^1 d\xi^2 d\xi^3 \varepsilon_{123} \quad (8.53b)$$

$$= d\xi^1 d\xi^2 d\xi^3 \sqrt{\det(\mathcal{G})} \epsilon_{123} \quad (8.53c)$$

$$= dV, \quad (8.53d)$$

where we used equation (8.38) for the final step. This result means that geometrically, the Levi-Civita tensor measures the volume defined by three vectors

$$\varepsilon(\vec{A}, \vec{B}, \vec{C}) = \text{volume}(\vec{A}, \vec{B}, \vec{C}). \quad (8.54)$$

This interpretation accords with the Cartesian coordinate discussion of the Levi-Civita tensor in Section 3.5.3.

8.7 Cross product and biorthogonality relation

The cross product of two Cartesian basis vectors yields the third, so that

$$\hat{\mathbf{x}} \wedge \hat{\mathbf{y}} = \hat{\mathbf{z}} \quad \text{cyclic.} \quad (8.55)$$

The coordinate invariant generalization of this relation is given by the biorthogonality relation

$$\vec{e}_a \wedge \vec{e}_b \equiv \varepsilon_{abc} \tilde{e}^c. \quad (8.56)$$

That is, the cross-product of two vectors leads to a one-form. We are thus led to the general coordinate expression for the cross-product of two arbitrary vectors

$$\vec{P} \wedge \vec{Q} = P^a Q^b \vec{e}_a \wedge \vec{e}_b \quad (8.57a)$$

$$= P^a Q^b \varepsilon_{abc} \tilde{e}^c. \quad (8.57b)$$

9

General tensor calculus

In this chapter, we generalize the Cartesian vector calculus of Chapter 4 to develop elements of vector calculus on a curved manifold using arbitrary coordinates. The material in this chapter requires the most patience from the novice, as there are some new elements of technology that must be mastered. Nonetheless, mastery has great payoffs, for example when describing fluid flow using isopycnal coordinates or tracer coordinates.

READER'S GUIDE TO THIS CHAPTER

This chapter assumes understanding of the general tensor algebra in Chapter 8. This chapter is the most specialized of the math chapters in this book. It can be returned to later when and if the need arises. This material is a summary of that given in Chapter 21 of *Griffies* (2004). The discussion will be rudimentary for those having taken a course in general relativity or tensor analysis yet will be challenging for the newcomer.

9.1	Loose threads	108
9.2	The covariant derivative operator	108
9.3	Covariant derivative of a vector	108
9.3.1	Derivative of a vector	109
9.3.2	An alternative derivation	109
9.3.3	Christoffel symbols are not components of a tensor	110
9.4	Covariant derivative of a one-form	110
9.5	Covariant derivative of the metric	110
9.6	Christoffel symbols in terms of the metric	110
9.7	Covariant divergence of a vector	111
9.7.1	Contraction of the Christoffel symbols	111
9.7.2	Exponential of the determinant	111
9.8	Covariant Laplacian of a scalar	112
9.9	Covariant curl of a vector	112
9.10	The Lie derivative	112
9.11	Gauss' divergence theorem	112
9.12	Stokes' theorem	113

9.1 Loose threads

The Lie derivative material in Section 9.10 is unwritten. Build on material from Section 14.2.1 of *Thorne and Blandford (2017)*.

9.2 The covariant derivative operator

Application of the chain rule leads to the transformation of the partial derivative operator

$$\partial_{\bar{a}} = \frac{\partial}{\partial \xi^{\bar{a}}} = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} \frac{\partial}{\partial \xi^a} = \Lambda^a_{\bar{a}} \partial_a. \quad (9.1)$$

Contracting the partial derivative with the basis of one-forms renders the form invariant expression of the gradient

$$\text{grad}(\psi) = \nabla \psi = \tilde{e}^a \partial_a \psi = \tilde{e}^{\bar{a}} \partial_{\bar{a}} \psi. \quad (9.2)$$

We thus define the covariant derivative operator

$$\nabla = \tilde{e}^a \partial_a. \quad (9.3)$$

Equation (9.2) provides the expression for the covariant derivative when acting on a scalar field, which is more commonly known as the gradient of the scalar.

9.3 Covariant derivative of a vector

The covariant derivative operator, ∇ , can act on a vector as well as a scalar, in which case we consider $\nabla \vec{F}$. To perform calculations requires us to unpack the manifestly covariant expression $\nabla \vec{F}$ by introducing a coordinate representation

$$\nabla \vec{F} = (\tilde{e}^b \partial_b) (F^a \vec{e}_a). \quad (9.4)$$

9.3.1 Derivative of a vector

The chain rule leads to the expression for the partial derivative operator acting on a vector field

$$\partial_b \vec{F} = \partial_b(\vec{e}_a F^a) \quad \text{coordinate representation of the vector } \vec{F} \quad (9.5a)$$

$$= (\partial_b F^a) \vec{e}_a + F^a \partial_b \vec{e}_a \quad \text{chain rule} \quad (9.5b)$$

$$= (\partial_b F^a) \vec{e}_a + F^a \Gamma_{ba}^c \vec{e}_c \quad \text{define Christoffel symbols} \quad (9.5c)$$

$$= (\partial_b F^a + F^c \Gamma_{bc}^a) \vec{e}_a \quad \text{reorganize} \quad (9.5d)$$

$$= (\nabla_b F^a) \vec{e}_a \quad \text{define covariant derivative acting on vector component.} \quad (9.5e)$$

In the third equality we introduced the *Christoffel symbols*

$$\partial_b \vec{e}_a = \Gamma_{ba}^c \vec{e}_c. \quad (9.6)$$

The Christoffel symbols carry information about the partial derivatives of the basis vectors. They vanish in Cartesian coordinates yet are generally nonzero. In the final equality we introduced components to the covariant derivative acting on the vector components

$$\nabla_b F^a = \partial_b F^a + \Gamma_{bc}^a F^c. \quad (9.7)$$

Contracting $\partial_b \vec{F}$ with the basis one-form \tilde{e}^b leads to

$$\nabla \vec{F} = (\tilde{e}^b \partial_b) \vec{F} = (\tilde{e}^b \nabla_b F^a) \vec{e}_a. \quad (9.8)$$

9.3.2 An alternative derivation

Recall from elementary calculus that the derivative of a function is computed by comparing the function at two points in space, dividing by the distance between those points, and taking the limit as the points get infinitesimally close. Now apply this operation to a vector field \vec{F} represented by arbitrary coordinates ξ^a , in which case

$$\partial_b \vec{F} = \lim_{\Delta \rightarrow 0} \frac{\vec{F}(\vec{P} + \Delta \vec{e}_b) - \vec{F}(\vec{P})}{\Delta}, \quad (9.9)$$

where $\vec{P} = \vec{e}_a \xi^a$ is the position vector for an arbitrary point and \vec{e}_b specifies the direction for computing the partial derivative. The basis vectors \vec{e}_a are spatially independent for Cartesian coordinates, so that the derivative of a vector is computed merely by taking the derivative of each Cartesian component

$$\partial_b \vec{F} = (\partial_b F^a) \vec{e}_a \quad \text{Cartesian coordinates.} \quad (9.10)$$

However, for general coordinates both the vector components and the basis vectors are spatially dependent, in which case

$$\vec{F}(\vec{P} + \Delta \vec{e}_b) - \vec{F}(\vec{P}) = [F^a + \Delta \partial_b F^b] [\vec{e}_a + \Delta \partial_b \vec{e}_a] - F^a \vec{e}_a \quad (9.11a)$$

$$= \Delta \partial_b (F^a \vec{e}_a) + \mathcal{O}(\Delta^2). \quad (9.11b)$$

This is the same result as found in the first step of the chain rule used in equation (9.5a). Following through that derivation then leads to the same coordinate expression for the covariant derivative acting on a vector field.

9.3.3 Christoffel symbols are not components of a tensor

The Christoffel symbols vanish in Euclidean space when using Cartesian coordinates whereas they are nonzero with other coordinates. As discussed in Section 7.2, a tensor that vanishes in one coordinate system remains zero for all coordinate systems. We thus conclude that the Christoffel symbols are *not* components to a tensor. Rather, they carry information regarding the partial derivatives of the coordinate basis vectors.

9.4 Covariant derivative of a one-form

The gradient acting on the product of a one-form and a vector is given by

$$\nabla(\tilde{E} \cdot \vec{F}) = \tilde{e}^b \partial_b (E_a F^a). \quad (9.12)$$

Expanding the partial derivative yields

$$\partial_b(E_a F^a) = F^a \partial_b E_a + E_a \partial_b F^a \quad (9.13a)$$

$$= F^a \partial_b E_a + E_a (\nabla_b F^a - \Gamma_{bc}^a F^c) \quad (9.13b)$$

$$= F^a (\partial_b E_a - \Gamma_{ba}^c E_c) + E_a \nabla_b F^a \quad (9.13c)$$

$$\equiv F^a \nabla_b E_a + E_a \nabla_b F^a. \quad (9.13d)$$

The last equality defines the covariant derivative when acting on the components to a one form

$$\nabla_b E_a = \partial_b E_a - \Gamma_{ba}^c E_c, \quad (9.14)$$

which leads to

$$\nabla \tilde{E} = (\tilde{e}^b \partial_b) \tilde{E} = (\tilde{e}^b \nabla_b E_a) \tilde{e}^a. \quad (9.15)$$

9.5 Covariant derivative of the metric

When written in Cartesian coordinates, the covariant derivative of the metric for Euclidean space vanishes,

$$\nabla g_{ab} = \nabla \delta_{ab} = 0, \quad (9.16)$$

because the Cartesian representation of the metric is the unit tensor δ_{ab} so that all Christoffel symbols vanish. Previous results establish the tensorial nature of the covariant derivative. Hence, $\nabla g_{ab} = 0$ is a valid result for *all* coordinates. This result is often called the *metricity* condition. It represents a self-consistency condition required for the manifolds considered in geophysical fluid mechanics.

9.6 Christoffel symbols in terms of the metric

We can develop an expression for the covariant derivative when acting on the components to a second order tensor. When applied to the metric tensor, its vanishing covariant derivative (equation (9.16)) then leads to the identity

$$0 = \nabla_c g_{ab} = \partial_c g_{ab} - \Gamma_{ca}^d g_{db} - \Gamma_{cb}^d g_{ad}. \quad (9.17)$$

We can solve this equation for the Christoffel symbols

$$\Gamma_{ab}^c = \frac{1}{2} g^{cd} (\partial_b g_{da} + \partial_a g_{db} - \partial_d g_{ab}). \quad (9.18)$$

This expression exhibits the symmetry property of the lower two indices on the Christoffel symbols

$$\Gamma_{ab}^c = \Gamma_{ba}^c. \quad (9.19)$$

9.7 Covariant divergence of a vector

The covariant divergence of the components to a vector results in a scalar

$$\nabla_a F^a = \partial_a F^a + \Gamma_{ab}^a F^b. \quad (9.20)$$

We now bring this expression into a form more convenient for practical calculations.

9.7.1 Contraction of the Christoffel symbols

Expression (9.18) for the Christoffel symbols yields for the contraction

$$\Gamma_{ab}^a = \frac{1}{2} g^{ad} (\partial_b g_{da} + \partial_a g_{db} - \partial_d g_{ab}) = \frac{1}{2} g^{ad} \partial_b g_{ad} \quad (9.21)$$

where symmetry of the metric tensor and its inverse was used.

9.7.2 Exponential of the determinant

For any symmetric positive definite matrix such as the metric tensor we can write

$$\det(A) = e^{\ln \det(A)} \quad \text{simple identity} \quad (9.22a)$$

$$= e^{\ln(\prod_i \Lambda_i)} \quad \text{determinant related to product of eigenvalues} \quad (9.22b)$$

$$= e^{\sum_i \ln \Lambda_i} \quad \text{simple identity} \quad (9.22c)$$

$$= e^{\text{Tr}(\ln A)} \quad \text{sum of eigenvalues related to trace of matrix.} \quad (9.22d)$$

Each of these identities is trivial to verify using a set of coordinates in which the matrix is diagonal. For any symmetric and positive definite matrix, such a set of coordinates always exists. This result gives

$$\partial_c \ln \det(A) = \partial_c [\text{Tr}(\ln A)] \quad (9.23a)$$

$$= \text{Tr}(\partial_c \ln A) \quad (9.23b)$$

$$= \text{Tr}(A^{-1} \partial_c A). \quad (9.23c)$$

With A now set equal to the metric tensor \mathcal{G} with components g_{ab} , this result yields

$$\partial_c \ln \det(\mathcal{G}) = g^{ab} \partial_c g_{ab} \quad (9.24)$$

which in turn yields for the contracted Christoffel symbol

$$\Gamma_{ac}^a = \partial_c \ln \sqrt{\det(\mathcal{G})}. \quad (9.25)$$

This result brings the covariant divergence of a vector to the form

$$\nabla_a F^a = \partial_a F^a + F^a \partial_a \ln \sqrt{\det(\mathcal{G})} \quad (9.26a)$$

$$= \frac{1}{\sqrt{\mathcal{G}}} \partial_a [\sqrt{\mathcal{G}} F^a]. \quad (9.26b)$$

This is a very convenient result since it requires us to use only partial derivatives in the chosen coordinate system chosen. All coordinate dependent properties are captured by $\sqrt{\det(\mathcal{G})}$.

9.8 Covariant Laplacian of a scalar

Making use of equation (9.26b) with

$$F^a = g^{ab} \partial_b \psi \quad (9.27)$$

leads to the covariant Laplacian of a scalar field

$$\nabla_a (g^{ab} \partial_b \psi) = \frac{1}{\sqrt{\mathcal{G}}} \partial_a [\sqrt{\mathcal{G}} g^{ab} \partial_b \psi]. \quad (9.28)$$

This expression is fundamental to the evolution of scalar fields under the impacts from diffusion (Chapter 37).

9.9 Covariant curl of a vector

The Levi-Civita tensor $\varepsilon_{abc} = \sqrt{\det(\mathcal{G})} \epsilon_{abc}$ from Section 8.6 is useful for generalizing the curl operation from Cartesian coordinates in Euclidean space to arbitrary coordinates on a curved manifold. Consequently, we define the curl according to the coordinate invariant expression

$$\text{curl}(\vec{F}) = \varepsilon^{abc} (\nabla_b F_c) \vec{e}_a = \varepsilon_{abc} (\nabla^b F^c) \vec{e}^a. \quad (9.29)$$

This expression simplifies by making use of equation (9.14) for the covariant derivative $\nabla_b F_c = \partial_b F_c - \Gamma_{cb}^a F_a$. Conveniently, the contraction $\varepsilon^{abc} \Gamma_{cb}^a$ vanishes identically since $\varepsilon^{abc} = -\varepsilon^{acb}$, whereas $\Gamma_{cb}^a = \Gamma_{bc}^a$. Hence, one is left with the general expression for the covariant curl that involves just the partial derivatives

$$\text{curl}(\vec{F}) = \varepsilon^{abc} (\partial_b F_c) \vec{e}_a = \varepsilon^{abc} [\partial_b (g_{cd} F^d)] \vec{e}_a. \quad (9.30)$$

The second equality made use of the equality $F_c = g_{cd} F^d$.

9.10 The Lie derivative

As in Section 14.2.1 of [Thorne and Blandford \(2017\)](#).

9.11 Gauss' divergence theorem

The integral theorems from Cartesian vector analysis transform in a straightforward manner to arbitrary coordinates in arbitrary smooth spaces. An easy way to prove the theorems is to invoke the ideas of general covariance from Section 7.2, in which the integral theorems are written in a tensorially correct manner and then partial derivatives are changed to covariant derivatives. The

divergence theorem offers a particularly simple example. For this purpose, we make use of the volume element (8.38)

$$dV = \sqrt{\det(\mathcal{G})} d\xi^1 d\xi^2 d\xi^3, \quad (9.31)$$

multiplied by the covariant divergence (9.26b). Hence, the volume integral of the divergence is given by

$$\int_{\mathcal{R}} (\nabla_a F^a) dV = \int_{\mathcal{R}} \partial_a [\sqrt{\det(\mathcal{G})} F^a] d\xi^1 d\xi^2 d\xi^3 = \oint_{\partial\mathcal{R}} F^a \hat{n}_a d\mathcal{S}, \quad (9.32)$$

where \hat{n} is the outward normal one-form for the boundary, $\partial\mathcal{R}$, and \hat{n}_a are its covariant components.

9.12 Stokes' theorem

The Cartesian form of Stokes' Theorem from Section 4.6 is generalized in a manner similar to the divergence theorem

$$\oint_{\partial\mathcal{S}} \vec{F} \cdot d\vec{x} = \int_{\mathcal{S}} \text{curl}(\vec{F}) \cdot \hat{n} d\mathcal{S}, \quad (9.33)$$

where $d\vec{x}$ is the vector line element along the path and $\partial\mathcal{S}$ is the closed path defining the boundary to a simply connected two-dimensional surface \mathcal{S} . For the circulation on the left hand side we have

$$\vec{F} \cdot d\vec{x} = F^a \vec{e}_a \cdot \vec{e}_b dx^b = F_b dx^b = F_{\bar{b}} d\xi^{\bar{b}}. \quad (9.34)$$

For the curl on the right hand side we have

$$\text{curl}(\vec{F}) \cdot \hat{n} = \varepsilon^{abc} (\partial_b F_c) \vec{e}_a \cdot \hat{n} = \varepsilon^{abc} (\partial_b F_c) \hat{n}_a = \varepsilon^{\bar{a}\bar{b}\bar{c}} (\partial_{\bar{b}} F_{\bar{c}}) \hat{n}_{\bar{a}}, \quad (9.35)$$

thus leading to the expression of Stokes' theorem in arbitrary coordinates

$$\oint_{\partial\mathcal{S}} F_{\bar{b}} d\xi^{\bar{b}} = \int_{\mathcal{S}} \varepsilon^{\bar{a}\bar{b}\bar{c}} (\partial_{\bar{b}} F_{\bar{c}}) \hat{n}_{\bar{a}} d\mathcal{S}. \quad (9.36)$$

10

Orthogonal coordinates

READER'S GUIDE TO THIS CHAPTER

This chapter compiles mathematical results for Cartesian, spherical, and cylindrical coordinates, thus providing explicit examples of the general tensor analysis machinery from Chapter 9. Many results discussed here are used throughout this book, with this chapter readily used as a reference to return to when certain coordinate dependent details are needed.

10.1	Cartesian coordinates	116
10.1.1	The basics	116
10.1.2	Summary of Cartesian coordinate expressions	117
10.2	Spherical coordinates	117
10.2.1	Transforming between Cartesian and spherical coordinates	118
10.2.2	Basis vectors	119
10.2.3	Basis one-forms	120
10.2.4	Position and velocity	120
10.2.5	Metric tensor	120
10.2.6	Components of a vector field	121
10.2.7	Differential operators	122
10.2.8	Summary of spherical coordinate expressions	122
10.3	Cylindrical-polar coordinates	122
10.3.1	Transforming between Cartesian and cylindrical-polar coordinates	124
10.3.2	Basis vectors	124
10.3.3	Basis one-forms	124
10.3.4	Position and velocity	125
10.3.5	Metric tensor	125
10.3.6	Components of a vector field	126
10.3.7	Differential operators	126
10.3.8	Summary of cylindrical coordinate expressions	127
10.4	General orthogonal coordinates	127

10.1 Cartesian coordinates

Whenever developing a general tensor relation it is useful to check its validity by considering Cartesian coordinates. We here summarize some results from our discussion of Cartesian tensors in Chapters 3 and 4.

10.1.1 The basics

We start by expressing the trajectory in the following equivalent forms

$$\mathcal{P}(\tau) = \vec{e}_1 x(\tau) + \vec{e}_2 y(\tau) + \vec{e}_3 z(\tau) \quad (10.1a)$$

$$= \hat{\mathbf{x}} x(\tau) + \hat{\mathbf{y}} y(\tau) + \hat{\mathbf{z}} z(\tau) \quad (10.1b)$$

$$= \vec{x}(\tau) \quad (10.1c)$$

$$= \mathbf{x}(\tau), \quad (10.1d)$$

with the basis vectors written

$$\vec{e}_1 = \hat{\mathbf{x}} \quad \vec{e}_2 = \hat{\mathbf{y}} \quad \vec{e}_3 = \hat{\mathbf{z}}. \quad (10.2)$$

The boldface notation is used for the position vector in the final equality of equation (10.1d), with the boldface commonly used throughout this book. Notably, the orthogonal unit vectors for Cartesian coordinates are normalized so that

$$\vec{e}_1 \cdot \vec{e}_1 = \vec{e}_2 \cdot \vec{e}_2 = \vec{e}_3 \cdot \vec{e}_3 = 1. \quad (10.3)$$

Furthermore, the basis vectors are identical to the basis one-forms

$$\vec{e}_1 = \tilde{e}^1 = \hat{x}, \quad \vec{e}_2 = \tilde{e}^2 = \hat{y}, \quad \vec{e}_3 = \tilde{e}^3 = \hat{z}. \quad (10.4)$$

Since the Cartesian basis vectors are independent of both space and time, we compute the coordinate representation of the velocity vector through taking the time derivative as

$$\vec{v}(\tau) = \mathbf{v}(\tau) \quad (10.5a)$$

$$= \frac{d\mathcal{P}}{d\tau} \quad (10.5b)$$

$$= \frac{dx}{d\tau} \quad (10.5c)$$

$$= \vec{e}_1 \frac{dx(\tau)}{d\tau} + \vec{e}_2 \frac{dy(\tau)}{d\tau} + \vec{e}_3 \frac{dz(\tau)}{d\tau} \quad (10.5d)$$

$$= \hat{x} v^1(\tau) + \hat{y} v^2(\tau) + \hat{z} v^3(\tau). \quad (10.5e)$$

10.1.2 Summary of Cartesian coordinate expressions

In Cartesian coordinates, mathematical operators and integral theorems take their familiar form. We here list the key ones in forms that are encountered throughout this book.

$$\mathbf{x} = (x^1, x^2, x^3) = (x, y, z) \quad \text{Cartesian coordinates} \quad (10.6)$$

$$\mathbf{F} = \hat{x} F^1 + \hat{y} F^2 + \hat{z} F^3 = \hat{x} F_1 + \hat{y} F_2 + \hat{z} F_3 \quad \text{covariant = contravariant} \quad (10.7)$$

$$\frac{\partial}{\partial x^a} = \partial_{x^a} \quad \text{or} \quad (\partial_x, \partial_y, \partial_z) \quad \text{partial derivative operator} \quad (10.8)$$

$$\nabla = \hat{x} \partial_x + \hat{y} \partial_y + \hat{z} \partial_z \quad \text{gradient operator} \quad (10.9)$$

$$\nabla_z = \hat{x} \partial_x + \hat{y} \partial_y \quad \text{horizontal gradient operator} \quad (10.10)$$

$$\nabla \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y + \partial_z F_z \quad \text{divergence of a vector} \quad (10.11)$$

$$\nabla_z \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y \quad \text{horizontal divergence of a vector} \quad (10.12)$$

$$(\nabla \wedge \mathbf{F})_a = \epsilon_{abc} \partial_b F_c \quad \text{components to curl of a vector} \quad (10.13)$$

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{F} dV = \oint_{\partial\mathcal{R}} \mathbf{F} \cdot \hat{\mathbf{n}} dS \quad \text{divergence (or Gauss's) theorem} \quad (10.14)$$

$$\oint_{\partial\mathcal{S}} \mathbf{F} \cdot d\mathbf{x} = \int_{\mathcal{S}} (\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad \text{Stokes' theorem.} \quad (10.15)$$

10.2 Spherical coordinates

We now consider spherical coordinates defined by Figure 10.1 and related to Cartesian coordinates through

$$x = r \cos \phi \cos \lambda \quad (10.16a)$$

$$y = r \cos \phi \sin \lambda \quad (10.16b)$$

$$z = r \sin \phi. \quad (10.16c)$$

The radial coordinate

$$r = \sqrt{x^2 + y^2 + z^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}} \quad (10.17)$$

measures the distance from the center of the sphere to position of the particle. The spherical angle coordinates

$$0 \leq \lambda \leq 2\pi \quad \text{longitude} \quad (10.18)$$

$$-\pi/2 \leq \phi \leq \pi/2 \quad \text{latitude} \quad (10.19)$$

specify the longitude, measuring the radians of the position east of the prime meridian, and latitude, measuring the radians north or south from the equator. To streamline notation in the following, we introduce the unbarred and barred labels for the Cartesian and spherical coordinates, respectively

$$(x, y, z) = (\xi^1, \xi^2, \xi^3) \equiv \xi^a \quad (10.20)$$

$$(\lambda, \phi, r) = (\xi^{\bar{1}}, \xi^{\bar{2}}, \xi^{\bar{3}}) \equiv \xi^{\bar{a}}. \quad (10.21)$$

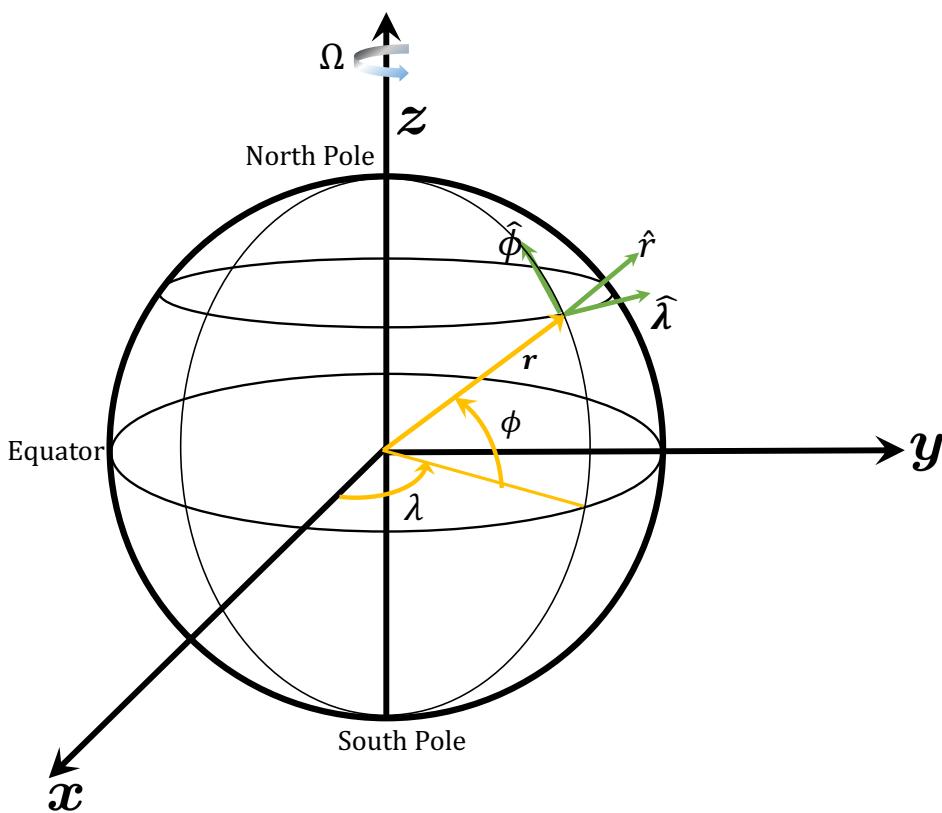


Figure 10.1: This schematic illustrates the geometry and notation for motion around a rotating sphere of radius R . For geophysical applications, the sphere rotates counter-clockwise when looking down from the north polar axis with angular speed Ω . The planetary Cartesian triad of orthonormal basis vectors, $(\hat{x}, \hat{y}, \hat{z})$ points along the orthogonal axes and rotates with the sphere. The planetary spherical triad (also rotating with the sphere) of orthonormal basis vectors, $(\hat{\lambda}, \hat{\phi}, \hat{r})$, makes use of the longitudinal unit vector $\hat{\lambda}$, which points in the longitudinal direction (positive eastward), the latitudinal unit vector $\hat{\phi}$, which points in the latitudinal direction (positive northward) and the radial unit vector \hat{r} , which point in the radial direction (positive away from the center).

10.2.1 Transforming between Cartesian and spherical coordinates

Following the general discussion in Section 8.1.4, we consider the infinitesimal distance along one of the Cartesian coordinate axes, $d\xi^a$. The chain rule allows us to relate this distance to those along

the axes of the spherical coordinate system

$$d\xi^a = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} d\xi^{\bar{a}} = \Lambda_{\bar{a}}^a d\xi^{\bar{a}}. \quad (10.22)$$

The partial derivatives $\partial \xi^a / \partial \xi^{\bar{a}}$ form components to the transformation matrix that transforms between coordinate representations. For the coordinate relation (10.16a)-(10.16c), this transformation matrix is given by

$$\Lambda_{\bar{a}}^a = \begin{bmatrix} \partial \xi^1 / \partial \xi^{\bar{1}} & \partial \xi^1 / \partial \xi^{\bar{2}} & \partial \xi^1 / \partial \xi^{\bar{3}} \\ \partial \xi^2 / \partial \xi^{\bar{1}} & \partial \xi^2 / \partial \xi^{\bar{2}} & \partial \xi^2 / \partial \xi^{\bar{3}} \\ \partial \xi^3 / \partial \xi^{\bar{1}} & \partial \xi^3 / \partial \xi^{\bar{2}} & \partial \xi^3 / \partial \xi^{\bar{3}} \end{bmatrix} = \begin{bmatrix} -r \cos \phi \sin \lambda & -r \sin \phi \cos \lambda & \cos \phi \cos \lambda \\ r \cos \phi \cos \lambda & -r \sin \phi \sin \lambda & \cos \phi \sin \lambda \\ 0 & r \cos \phi & \sin \phi \end{bmatrix}. \quad (10.23)$$

The determinant of the transformation (Jacobian) is given by

$$\det(\Lambda_{\bar{a}}^a) = r^2 \cos \phi. \quad (10.24)$$

The Jacobian vanishes at the north and south poles ($\phi = \pm\pi/2$), where the transformation is singular. Methods familiar from linear algebra render the inverse transformation matrix

$$\Lambda_{\bar{a}}^a = \frac{1}{r^2 \cos \phi} \begin{bmatrix} -r \sin \lambda & r \cos \lambda & 0 \\ -r \cos \phi \sin \phi \cos \lambda & -r \cos \phi \sin \phi \sin \lambda & r \cos^2 \phi \\ r^2 \cos^2 \phi \cos \lambda & r^2 \cos^2 \phi \sin \lambda & r^2 \cos \phi \sin \phi \end{bmatrix}. \quad (10.25)$$

10.2.2 Basis vectors

The spherical coordinate basis vectors, $\vec{e}_{\bar{a}}$, are related to the Cartesian coordinate basis vectors, \vec{e}_a , through the transformation

$$\vec{e}_{\bar{a}} = \Lambda_{\bar{a}}^a \vec{e}_a. \quad (10.26)$$

The transformation matrix (10.23) leads to

$$\vec{e}_{\lambda} = r \cos \phi (-\hat{x} \sin \lambda + \hat{y} \cos \lambda) \quad (10.27a)$$

$$\vec{e}_{\phi} = r (-\hat{x} \sin \phi \cos \lambda - \hat{y} \sin \phi \sin \lambda + \hat{z} \cos \phi) \quad (10.27b)$$

$$\vec{e}_r = \hat{x} \cos \phi \cos \lambda + \hat{y} \cos \phi \sin \lambda + \hat{z} \sin \phi. \quad (10.27c)$$

We can introduce the orthonormal unit vectors through

$$\vec{e}_{\lambda} = r \cos \phi \hat{\lambda} \quad \vec{e}_{\phi} = r \hat{\phi} \quad \vec{e}_r = \hat{r}, \quad (10.28)$$

so that

$$\hat{\lambda} = -\hat{x} \sin \lambda + \hat{y} \cos \lambda \quad (10.29a)$$

$$\hat{\phi} = -\hat{x} \cos \lambda \sin \phi - \hat{y} \sin \lambda \sin \phi + \hat{z} \cos \phi \quad (10.29b)$$

$$\hat{r} = \hat{x} \cos \lambda \cos \phi + \hat{y} \sin \lambda \cos \phi + \hat{z} \sin \phi \quad (10.29c)$$

along with the inverse relations

$$\hat{x} = -\hat{\lambda} \sin \lambda - \hat{\phi} \cos \lambda \sin \phi + \hat{r} \cos \lambda \cos \phi \quad (10.30a)$$

$$\hat{y} = \hat{\lambda} \cos \lambda - \hat{\phi} \sin \lambda \sin \phi + \hat{r} \sin \lambda \cos \phi \quad (10.30b)$$

$$\hat{z} = \hat{\phi} \cos \phi + \hat{r} \sin \phi. \quad (10.30c)$$

10.2.3 Basis one-forms

Since spherical coordinates are orthogonal, we can readily derive the one-form basis through inverting the vector basis

$$\tilde{e}^\lambda = (r \cos \phi)^{-1} \hat{\lambda} \quad \tilde{e}^\phi = r^{-1} \hat{\phi} \quad \tilde{e}^r = \hat{r}, \quad (10.31)$$

which satisfy the orthogonality relation with the basis vectors (Section 8.2.2)

$$\tilde{e}^{\bar{b}} \cdot \vec{e}_{\bar{a}} = \delta_{\bar{a}}^{\bar{b}}. \quad (10.32)$$

10.2.4 Position and velocity

In spherical coordinates, the position of a point is fully specified by the radial position

$$\mathcal{P}(\tau) = r \vec{e}_r = r \hat{r}. \quad (10.33)$$

The velocity requires all three spherical coordinates since the radial basis vector is a function of the angular positions, which are in turn functions of time. Use of the chain rule renders

$$\vec{v}(\tau) = \mathbf{v} \quad (10.34a)$$

$$= \frac{d\mathcal{P}}{d\tau} \quad (10.34b)$$

$$= \vec{e}_r \frac{dr}{d\tau} + r \frac{d\vec{e}_r}{d\tau} \quad (10.34c)$$

$$= \vec{e}_r \frac{dr}{d\tau} + r \frac{\partial \vec{e}_r}{\partial \lambda} \frac{d\lambda}{d\tau} + r \frac{\partial \vec{e}_r}{\partial \phi} \frac{d\phi}{d\tau} \quad (10.34d)$$

$$\equiv \vec{e}_r \frac{dr}{d\tau} + \vec{e}_\lambda \frac{d\lambda}{d\tau} + \vec{e}_\phi \frac{d\phi}{d\tau} \quad (10.34e)$$

$$= \vec{e}_r v^r + \vec{e}_\lambda v^\lambda + \vec{e}_\phi v^\phi. \quad (10.34f)$$

To reach this result we made use of the identities satisfied by the spherical basis vectors

$$\vec{e}_\lambda = r \frac{\partial \vec{e}_r}{\partial \lambda} \quad \vec{e}_\phi = r \frac{\partial \vec{e}_r}{\partial \phi}. \quad (10.35)$$

10.2.5 Metric tensor

The metric tensor for spherical coordinates takes on the diagonal form

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} (r \cos \phi)^2 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (10.36)$$

as does the spherical representation of the inverse metric tensor

$$g^{\bar{a}\bar{b}} = \tilde{e}^{\bar{a}} \cdot \tilde{e}^{\bar{b}} = \begin{bmatrix} (r \cos \phi)^{-2} & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (10.37)$$

Volume element and Levi-Civita tensor

The square root of the determinant of the metric tensor written in spherical coordinates (from equation (10.36)) is given by

$$\sqrt{\det(\bar{G})} = r^2 \cos \phi \quad (10.38)$$

so that the volume element is

$$dV = r^2 \cos \phi dr d\lambda d\phi. \quad (10.39)$$

The covariant Levi-Civita tensor has the spherical representation

$$\varepsilon_{\bar{a}\bar{b}\bar{c}} = (r^2 \cos \phi) \epsilon_{\bar{a}\bar{b}\bar{c}}. \quad (10.40)$$

Cross product of basis vectors

As a check on the formalism for cross products, let us verify the relation (8.56) for the cross product of two basis vectors using spherical coordinates

$$\vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}} \implies \vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = (r^2 \cos \phi) \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}}. \quad (10.41)$$

Making use of the spherical coordinate basis vectors and one-forms renders

$$\vec{e}_r \wedge \vec{e}_{\lambda} = (r \cos \phi) (\hat{r} \wedge \hat{\lambda}) = (r \cos \phi) \hat{\phi} = (r^2 \cos \phi) \tilde{e}^{\phi} = \varepsilon_{r\lambda\phi} \tilde{e}^{\phi} \quad (10.42a)$$

$$\vec{e}_{\lambda} \wedge \vec{e}_{\phi} = (r^2 \cos \phi) (\hat{\lambda} \wedge \hat{\phi}) = (r^2 \cos \phi) \hat{r} = (r^2 \cos \phi) \tilde{e}^r = \varepsilon_{\lambda\phi r} \tilde{e}^r \quad (10.42b)$$

$$\vec{e}_{\phi} \wedge \vec{e}_r = r (\hat{\phi} \wedge \hat{r}) = r \hat{\lambda} = (r^2 \cos \phi) \tilde{e}^{\lambda} = \varepsilon_{\phi r \lambda} \tilde{e}^{\lambda}. \quad (10.42c)$$

To reach these results we made use of the cross products for the spherical coordinate unit vectors

$$\hat{r} \wedge \hat{\lambda} = \hat{\phi} \quad \hat{\lambda} \wedge \hat{\phi} = \hat{r} \quad \hat{\phi} \wedge \hat{r} = \hat{\lambda}. \quad (10.43)$$

10.2.6 Components of a vector field

A vector field \vec{F} has Cartesian components, F^a , related to spherical components, $F^{\bar{a}}$, via the transformation matrix, $F^{\bar{a}} = \Lambda_a^{\bar{a}} F^a$. This transformation leads to

$$F^{\bar{1}} = (r \cos \phi)^{-1} [-F^x \sin \lambda + F^y \cos \lambda] \quad (10.44a)$$

$$F^{\bar{2}} = r^{-1} [-F^x \sin \phi \cos \lambda - F^y \sin \phi \sin \lambda + F^z \cos \phi] \quad (10.44b)$$

$$F^{\bar{3}} = F^x \cos \phi \cos \lambda + F^y \cos \phi \sin \lambda + F^z \sin \phi. \quad (10.44c)$$

Making use of the spherical unit vector (10.29a)-(10.29c) leads to the identities

$$(r \cos \phi) F^{\bar{1}} = \hat{\lambda} \cdot \mathbf{F} \quad (10.45a)$$

$$r F^{\bar{2}} = \hat{\phi} \cdot \mathbf{F} \quad (10.45b)$$

$$F^{\bar{3}} = \hat{r} \cdot \mathbf{F}. \quad (10.45c)$$

10.2.7 Differential operators

In spherical coordinates the gradient operator $\nabla = \tilde{e}^a \partial_a$ takes on the form

$$\nabla = \hat{\lambda} (r \cos \phi)^{-1} \partial_\lambda + \hat{\phi} r^{-1} \partial_\phi + \hat{r} \partial_r \quad (10.46)$$

and the covariant divergence of a vector field is given by

$$\nabla_{\bar{a}} F^{\bar{a}} = (r^2 \cos \phi)^{-1} \partial_{\bar{a}} [r^2 \cos \phi F^{\bar{a}}] \quad (10.47a)$$

$$= (r^2 \cos \phi)^{-1} \left(\partial_\lambda [r^2 \cos \phi F^{\bar{1}}] + \partial_\phi [r^2 \cos \phi F^{\bar{2}}] + \partial_r [r^2 \cos \phi F^{\bar{3}}] \right) \quad (10.47b)$$

$$= \frac{1}{r \cos \phi} \frac{\partial (\hat{\lambda} \cdot \mathbf{F})}{\partial \lambda} + \frac{1}{r \cos \phi} \frac{\partial (\hat{\phi} \cdot \mathbf{F} \cos \phi)}{\partial \phi} + \frac{1}{r^2} \frac{\partial (\hat{r} \cdot \mathbf{F} r^2)}{\partial r} \quad (10.47c)$$

The covariant curl (Section 9.9) takes the form

$$(\text{curl } \vec{F})^{\bar{1}} = (r^2 \cos \phi)^{-1} [\partial_\phi F^{\bar{3}} - \partial_r (r^2 F^{\bar{2}})] \quad (10.48a)$$

$$(\text{curl } \vec{F})^{\bar{2}} = (r^2 \cos \phi)^{-1} [\partial_r (r^2 \cos^2 \phi F^{\bar{1}}) - \partial_\lambda F^{\bar{3}}] \quad (10.48b)$$

$$(\text{curl } \vec{F})^{\bar{3}} = (r^2 \cos \phi)^{-1} [\partial_\lambda (r^2 F^{\bar{2}}) - \partial_\phi (r^2 \cos^2 \phi F^{\bar{1}})], \quad (10.48c)$$

which can be written in the more conventional form (e.g., equation (2.33) of [Vallis \(2017\)](#))

$$r \cos \phi (\text{curl } \vec{F})^{\bar{1}} = \frac{1}{r} \left[\frac{\partial (\hat{r} \cdot \mathbf{F})}{\partial \phi} - \frac{\partial (r \hat{\phi} \cdot \mathbf{F})}{\partial r} \right] \quad (10.49a)$$

$$r (\text{curl } \vec{F})^{\bar{2}} = \frac{1}{r} \left[\frac{\partial (r \hat{\lambda} \cdot \mathbf{F})}{\partial r} - \frac{1}{\cos \phi} \frac{\partial (\hat{r} \cdot \mathbf{F})}{\partial \lambda} \right] \quad (10.49b)$$

$$(\text{curl } \vec{F})^{\bar{3}} = \frac{1}{r \cos \phi} \left[\frac{\partial (\hat{\phi} \cdot \mathbf{F})}{\partial \lambda} - \frac{\partial (\cos \phi \hat{\lambda} \cdot \mathbf{F})}{\partial \phi} \right]. \quad (10.49c)$$

10.2.8 Summary of spherical coordinate expressions

We here summarize the spherical coordinate version of some common mathematical operators.

$$(\lambda, \phi, r) = (x^{\bar{1}}, x^{\bar{2}}, x^{\bar{3}}) \quad \text{spherical coordinates} \quad (10.50)$$

$$(r \cos \phi) F^{\bar{1}} = \hat{\lambda} \cdot \mathbf{F} \quad r F^{\bar{2}} = \hat{\phi} \cdot \mathbf{F} \quad F^{\bar{3}} = \hat{r} \cdot \mathbf{F} \quad \text{vector components} \quad (10.51)$$

$$\nabla = \hat{\lambda} (r \cos \phi)^{-1} \partial_\lambda + \hat{\phi} r^{-1} \partial_\phi + \hat{r} \partial_r \quad \text{gradient} \quad (10.52)$$

$$\nabla_{\bar{a}} F^{\bar{a}} = \frac{1}{r \cos \phi} \frac{\partial (\hat{\lambda} \cdot \mathbf{F})}{\partial \lambda} + \frac{1}{r \cos \phi} \frac{\partial (\hat{\phi} \cdot \mathbf{F} \cos \phi)}{\partial \phi} + \frac{1}{r^2} \frac{\partial (\hat{r} \cdot \mathbf{F} r^2)}{\partial r} \quad \text{divergence} \quad (10.53)$$

$$(\nabla \wedge \mathbf{F})_{\bar{a}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} F^{\bar{c}} \quad \text{see equations (10.49a) -- (10.49c)} \quad \text{curl of a vector} \quad (10.54)$$

10.3 Cylindrical-polar coordinates

Many physical systems exhibit circular symmetry in two-dimensions or cylindrical symmetry in three-dimensions. The primary example encountered in this book is the laboratory motion of liquid in a rotating circular tank. In the following, we emulate the discussion presented for the spherical coordinates in Section 10.2, here focusing on cylindrical-polar coordinates as shown in Figure 10.2. Our task here is somewhat simpler than for the spherical coordinates since the vertical or axial position, z , remains unchanged from its Cartesian value.

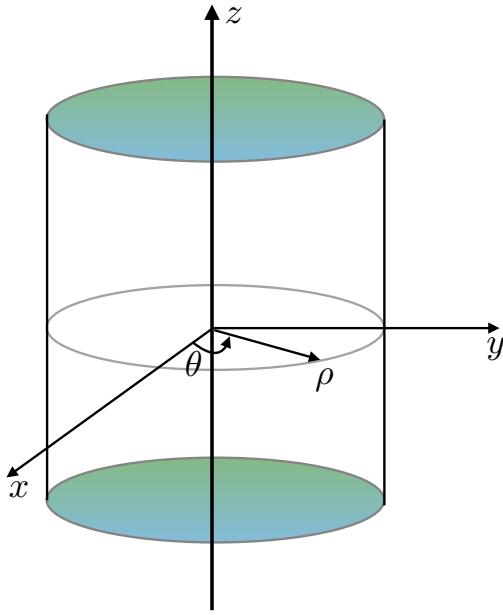


Figure 10.2: This schematic illustrates the geometry and notation for cylindrical-polar coordinates used to describe motion in a rotating laboratory tank. The Cartesian triad of orthonormal basis vectors, $(\hat{x}, \hat{y}, \hat{z})$ points along the orthogonal axes. The cylindrical-polar triad of orthonormal basis vectors, $(\hat{\rho}, \hat{\theta}, \hat{z})$, makes use of the radial unit vector $\hat{\rho}$, which points outward from the vertical axis, the angular unit vector $\hat{\theta}$, which points in the counter-clockwise direction around the circle, and the vertical unit vector \hat{z} .

The coordinate transformation between Cartesian coordinates and cylindrical-polar coordinates is given by

$$x = \rho \cos \theta \equiv \xi^1 \cos \xi^2 \quad (10.55a)$$

$$y = \rho \sin \theta \equiv \xi^1 \sin \xi^2 \quad (10.55b)$$

$$z = \xi^3. \quad (10.55c)$$

The radial coordinate for cylindrical-polar coordinates

$$\rho = \sqrt{x^2 + y^2} \quad (10.56)$$

measures the distance from the vertical z -axis, and the angular coordinate $0 \leq \theta \leq 2\pi$ measures the angle counter-clockwise from the positive x -axis. We introduce the unbarred and barred labels for the Cartesian and cylindrical polar coordinates

$$(x, y, z) = (\xi^1, \xi^2, \xi^3) \equiv \xi^a \quad (10.57)$$

$$(\rho, \theta, z) = (\xi^1, \xi^2, \xi^3) \equiv \xi^{\bar{a}}. \quad (10.58)$$

Although the vertical position z remains the same in both coordinates, and is orthogonal to the other coordinates, it is useful to introduce a distinct symbol ξ^3 and $\xi^{\bar{3}}$ to specify what other coordinates are held fixed when performing derivative operations.

10.3.1 Transforming between Cartesian and cylindrical-polar coordinates

The coordinate relation (10.55a)-(10.55c) leads to the transformation matrix

$$\Lambda_{\bar{a}}^a = \begin{bmatrix} \partial\xi^1/\partial\xi^{\bar{1}} & \partial\xi^1/\partial\xi^{\bar{2}} & \partial\xi^1/\partial\xi^{\bar{3}} \\ \partial\xi^2/\partial\xi^{\bar{1}} & \partial\xi^2/\partial\xi^{\bar{2}} & \partial\xi^2/\partial\xi^{\bar{3}} \\ \partial\xi^3/\partial\xi^{\bar{1}} & \partial\xi^3/\partial\xi^{\bar{2}} & \partial\xi^3/\partial\xi^{\bar{3}} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (10.59)$$

and the inverse transformation is given by

$$\Lambda^{\bar{a}}_a = \frac{1}{\rho} \begin{bmatrix} \rho\cos\theta & \rho\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & \rho \end{bmatrix}. \quad (10.60)$$

The determinant of the transformation (Jacobian) is given by

$$\det(\Lambda_{\bar{a}}^a) = \rho, \quad (10.61)$$

which vanishes along the vertical axis, where the transformation is singular.

10.3.2 Basis vectors

The cylindrical-polar coordinate basis vectors, $\vec{e}_{\bar{a}}$, are related to the Cartesian coordinate basis vectors, \vec{e}_a , through the transformation $\vec{e}_{\bar{a}} = \Lambda_{\bar{a}}^a \vec{e}_a$. The transformation matrix (10.59) leads to

$$\vec{e}_{\rho} = \hat{x}\cos\theta + \hat{y}\sin\theta \quad (10.62a)$$

$$\vec{e}_{\theta} = \rho(-\hat{x}\sin\theta + \hat{y}\cos\theta) \quad (10.62b)$$

$$\vec{e}_{\bar{z}} = \hat{z}. \quad (10.62c)$$

We sometimes make use of the following orthonormal unit vectors $(\hat{\rho}, \hat{\theta}, \hat{z})$

$$\vec{e}_{\rho} = \hat{\rho} \quad \vec{e}_{\theta} = \rho\hat{\theta} \quad \vec{e}_{\bar{z}} = \hat{z} \quad (10.63)$$

along with the inverse relations

$$\hat{x} = \hat{\rho}\cos\theta - \hat{\theta}\sin\theta \quad (10.64a)$$

$$\hat{y} = \hat{\rho}\sin\theta + \hat{\theta}\cos\theta \quad (10.64b)$$

$$\hat{z} = \hat{z}. \quad (10.64c)$$

10.3.3 Basis one-forms

Since cylindrical-polar coordinates are orthogonal, we can readily derive the one-form basis through inverting the vector basis

$$\tilde{e}^{\rho} = \hat{\rho} \quad \tilde{e}^{\theta} = \rho^{-1}\hat{\theta} \quad \tilde{e}^{\bar{z}} = \hat{z}, \quad (10.65)$$

which satisfy the orthogonality relation (Section 8.2.2)

$$\tilde{e}^{\bar{b}} \cdot \vec{e}_{\bar{a}} = \delta_{\bar{a}}^{\bar{b}}. \quad (10.66)$$

10.3.4 Position and velocity

In cylindrical-polar coordinates, the position of a point is specified by the radial position plus the vertical position

$$\mathcal{P}(\tau) = \rho \vec{e}_\rho + z \vec{e}_{\bar{z}}. \quad (10.67)$$

The velocity requires all three coordinates since the radial basis vector is a function of the angular positions, which are in turn functions of time. Use of the chain rule renders

$$\vec{v}(\tau) = \frac{d\mathcal{P}}{d\tau} \quad (10.68a)$$

$$= \vec{e}_\rho \frac{d\rho}{d\tau} + \rho \frac{d\vec{e}_\rho}{d\tau} + \vec{e}_{\bar{z}} \frac{dz}{d\tau} \quad (10.68b)$$

$$= \vec{e}_\rho \frac{d\rho}{d\tau} + \rho \frac{\partial \vec{e}_\rho}{\partial \theta} \frac{d\theta}{d\tau} + \vec{e}_{\bar{z}} \frac{dz}{d\tau} \quad (10.68c)$$

$$= \vec{e}_\rho \frac{d\rho}{d\tau} + \vec{e}_\theta \frac{d\theta}{d\tau} + \vec{e}_{\bar{z}} \frac{dz}{d\tau} \quad (10.68d)$$

$$= \vec{e}_\rho v^\rho + \vec{e}_\theta v^\theta + \vec{e}_{\bar{z}} v^{\bar{z}}. \quad (10.68e)$$

To reach this result we made use of the identity

$$\vec{e}_\theta = r \frac{\partial \vec{e}_\rho}{\partial \theta}. \quad (10.69)$$

10.3.5 Metric tensor

The metric tensor for cylindrical-polar coordinates takes on the diagonal form

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (10.70)$$

as does the inverse metric tensor

$$g^{\bar{a}\bar{b}} = \vec{e}^{\bar{a}} \cdot \vec{e}^{\bar{b}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^{-2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (10.71)$$

Volume element and Levi-Civita tensor

The square root of the determinant of the metric tensor written in cylindrical-polar coordinates (from equation (10.70)) is given by

$$\sqrt{\det(\mathcal{G})} = \rho \quad (10.72)$$

so that the volume element is

$$dV = \rho d\rho d\theta dz. \quad (10.73)$$

The covariant Levi-Civita tensor has the cylindrical-polar representation

$$\epsilon_{\bar{a}\bar{b}\bar{c}} = \rho \epsilon_{\bar{a}\bar{b}\bar{c}}. \quad (10.74)$$

Cross product of basis vectors

As a check on the formalism for cross products, let us verify the relation (8.56) for the cross product of two basis vectors using spherical coordinates

$$\vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}} \implies \vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \rho \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}}. \quad (10.75)$$

Making use of the cylindrical-polar coordinate basis vectors and one-forms renders

$$\vec{e}_\rho \wedge \vec{e}_\theta = \rho (\hat{\rho} \wedge \hat{\theta}) = \rho \tilde{e}^{\bar{z}} = \varepsilon_{\rho\theta\bar{z}} \tilde{e}^{\bar{z}} \quad (10.76a)$$

$$\vec{e}_\theta \wedge \vec{e}_{\bar{z}} = \rho (\hat{\theta} \wedge \hat{z}) = \rho \hat{\rho} = \rho \tilde{e}^\rho = \varepsilon_{\theta z \rho} \tilde{e}^\rho \quad (10.76b)$$

$$\vec{e}_{\bar{z}} \wedge \vec{e}_\rho = \hat{z} \wedge \hat{\rho} = \hat{\theta} = \rho \tilde{e}^\theta = \varepsilon_{\bar{z}\rho\theta} \tilde{e}^\theta. \quad (10.76c)$$

To reach these results we made use of the cross products for the unit vectors

$$\hat{\rho} \wedge \hat{\theta} = \hat{z} \quad \hat{\theta} \wedge \hat{z} = \hat{\rho} \quad \hat{z} \wedge \hat{\rho} = \hat{\theta}. \quad (10.77)$$

10.3.6 Components of a vector field

A vector field \vec{F} has Cartesian components, F^a , related to cylindrical-polar components, $F^{\bar{a}}$, via the transformation matrix, $F^{\bar{a}} = \Lambda_a^{\bar{a}} F^a$. This transformation leads to

$$F^{\bar{1}} = F^x \cos \theta + F^y \sin \theta \quad (10.78a)$$

$$F^{\bar{2}} = \rho^{-1} [-F^x \sin \theta + F^y \cos \theta] \quad (10.78b)$$

$$F^{\bar{3}} = F^z. \quad (10.78c)$$

Introducing the cylindrical-polar unit vectors (10.63) leads to

$$F^{\bar{1}} = \hat{\rho} \cdot \mathbf{F} \quad (10.79a)$$

$$\rho F^{\bar{2}} = \hat{\theta} \cdot \mathbf{F} \quad (10.79b)$$

$$F^{\bar{3}} = \hat{z} \cdot \mathbf{F}. \quad (10.79c)$$

10.3.7 Differential operators

In cylindrical-polar coordinates, the gradient operator $\nabla = \tilde{e}^a \partial_a$ takes on the form

$$\nabla = \hat{\rho} \frac{\partial}{\partial \rho} + \frac{\hat{\theta}}{\rho} \frac{\partial}{\partial \theta} + \hat{z} \frac{\partial}{\partial z} \quad (10.80)$$

and the covariant divergence of a vector field is given by

$$\nabla_{\bar{a}} F^{\bar{a}} = \rho^{-1} \partial_{\bar{a}} (\rho F^{\bar{a}}) \quad (10.81a)$$

$$= \rho^{-1} \left(\partial_\rho [\rho F^{\bar{1}}] + \partial_\theta [\rho F^{\bar{2}}] + \partial_z [\rho F^{\bar{3}}] \right) \quad (10.81b)$$

$$= \frac{1}{\rho} \frac{\partial (\rho \hat{\rho} \cdot \mathbf{F})}{\partial \rho} + \frac{1}{\rho} \frac{\partial (\hat{\theta} \cdot \mathbf{F})}{\partial \theta} + \frac{\partial (\hat{z} \cdot \mathbf{F})}{\partial z}. \quad (10.81c)$$

The covariant curl (Section 9.9) takes the form

$$(\text{curl} \vec{F})^{\bar{1}} = \rho^{-1} [\partial_{\theta} F^{\bar{3}} - \partial_{\bar{z}} (\rho^2 F^{\bar{2}})] \quad (10.82a)$$

$$(\text{curl} \vec{F})^{\bar{2}} = \rho^{-1} [\partial_{\bar{z}} F^{\bar{1}} - \partial_{\rho} F^{\bar{3}}] \quad (10.82b)$$

$$(\text{curl} \vec{F})^{\bar{3}} = \rho^{-1} [\partial_{\rho} (\rho^2 F^{\bar{2}}) - \partial_{\theta} F^{\bar{1}}], \quad (10.82c)$$

which can be written in the more conventional form

$$(\text{curl} \vec{F})^{\bar{1}} = \frac{1}{\rho} \frac{\partial(\hat{z} \cdot \mathbf{F})}{\partial \theta} - \frac{\partial(\hat{\theta} \cdot \mathbf{F})}{\partial z} \quad (10.83a)$$

$$\rho (\text{curl} \vec{F})^{\bar{2}} = \frac{\partial(\hat{\rho} \cdot \mathbf{F})}{\partial z} - \frac{\partial(\hat{z} \cdot \mathbf{F})}{\partial z} \quad (10.83b)$$

$$(\text{curl} \vec{F})^{\bar{3}} = \frac{1}{\rho} \frac{\partial(\rho \hat{\theta} \cdot \mathbf{F})}{\partial \rho} - \frac{1}{\rho} \frac{\partial(\hat{\rho} \cdot \mathbf{F})}{\partial \theta}. \quad (10.83c)$$

10.3.8 Summary of cylindrical coordinate expressions

We here summarize the cylindrical coordinate version of some common mathematical operators.

$$(\rho, \theta, z) = (x^{\bar{1}}, x^{\bar{2}}, x^{\bar{3}}) \quad \text{cylindrical coordinates} \quad (10.84)$$

$$F^{\bar{1}} = \hat{\rho} \cdot \mathbf{F} \quad \rho F^{\bar{2}} = \hat{\theta} \cdot \mathbf{F} \quad F^{\bar{3}} = \hat{z} \cdot \mathbf{F} \quad \text{vector components} \quad (10.85)$$

$$\nabla = \hat{\rho} \frac{\partial}{\partial \rho} + \frac{\hat{\theta}}{\rho} \frac{\partial}{\partial \theta} + \hat{z} \frac{\partial}{\partial z} \quad \text{gradient} \quad (10.86)$$

$$\nabla_{\bar{a}} F^{\bar{a}} = \frac{1}{\rho} \frac{\partial(\rho \hat{\rho} \cdot \mathbf{F})}{\partial \rho} + \frac{1}{\rho} \frac{\partial(\hat{\theta} \cdot \mathbf{F})}{\partial \theta} + \frac{\partial(\hat{z} \cdot \mathbf{F})}{\partial z}. \quad \text{divergence} \quad (10.87)$$

$$(\nabla \wedge \mathbf{F})_{\bar{a}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} F^{\bar{c}} \quad \text{see equations (10.82a) -- (10.82c).} \quad \text{curl of a vector} \quad (10.88)$$

10.4 General orthogonal coordinates

We here generalize the spherical and cylindrical coordinates by considering a nonsingular and orthogonal set of coordinates defined such that the metric tensor takes on the diagonal form

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} h_{\bar{1}} & 0 & 0 \\ 0 & h_{\bar{2}} & 0 \\ 0 & 0 & h_{\bar{3}} \end{bmatrix}, \quad (10.89)$$

where $h_{\bar{a}} > 0$ are “stretching” functions. The corresponding volume element is expressed as

$$dV = h_{\bar{1}} h_{\bar{2}} h_{\bar{3}} d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}. \quad (10.90)$$

These *generalized orthogonal curvilinear coordinates* have a corresponding orthogonal set of basis vectors

$$\vec{e}_{\bar{a}} = h_{\bar{a}} \hat{\mathbf{e}}_{(\bar{a})} \quad \text{no implied sum.} \quad (10.91)$$

The objects $\hat{\mathbf{e}}_{(\bar{a})}$ are the dimensionless unit directions. The corresponding one-form basis is given by

$$\tilde{e}^{\bar{a}} = (h_{\bar{a}})^{-1} \hat{\mathbf{e}}_{(\bar{a})}. \quad (10.92)$$

The index on the unit directions is enclosed in parentheses to advertise that it is not tensorial; i.e., the unit directions do not transform as tensors. Rather, the functions $h_{\bar{a}}$ carry the tensorial properties of the basis vectors $\vec{e}_{\bar{a}}$. Results for the trajectory and velocity are straightforward generalizations of the spherical results in Section 10.2 and cylindrical-polar results from Section 10.3. A thorough examination of the generalized orthogonal coordinates is found in Section 21.11 of [Griffies \(2004\)](#).

11

Generalized vertical coordinates

Generalized vertical coordinates offer a mathematical framework for a variety of topics in stratified fluid mechanics. Hence, they appear in many guises throughout this book. Notably, their non-orthogonality requires extra care beyond that needed with the orthogonal coordinates considered in Chapter 10. GVCs were introduced by [Starr \(1945\)](#) for atmospheric modeling and for ocean modeling by [Bleck \(1978\)](#). There is a growing use of GVC-based numerical ocean and atmospheric models, prompting the need to master their use for analysis, model formulation, and theory.

READER'S GUIDE TO THIS CHAPTER

Particular chapters directly relying on GVC material from this chapter include the fluid kinematics discussed in Chapter 21, the tracer equation diffusion and stirring operators discussed in Chapter 39, and the dynamics discussed in Chapter 32. These related chapters are all part of the book's Tier-II material.

11.1	Introducing GVCs	130
11.1.1	Relating Cartesian and generalized vertical coordinates	131
11.1.2	A common confusion	132
11.2	Spatial basis vectors	132
11.2.1	A few more points on the transformation matrix	133
11.2.2	Expressions for the basis vectors	133
11.3	Basis one-forms	133
11.3.1	A few more points on the inverse transformation matrix	134
11.3.2	GVC basis one-forms	134
11.3.3	Verifying the orthogonality relation	134
11.4	Showing that $\Lambda_{\bar{a}}^{\bar{a}} \Lambda_{\bar{b}}^a = \delta_{\bar{b}}^{\bar{a}}$ and $\Lambda_{\bar{a}}^a \Lambda_{\bar{b}}^{\bar{a}} = \delta_b^a$	134
11.5	Triple product identities	135
11.6	Position vector	136
11.7	Transforming components of a first order tensor	137
11.8	Velocity	138
11.8.1	Contravariant components to the velocity	138
11.8.2	Covariant components to the velocity vector	139
11.8.3	Introducing the material time derivative	139
11.8.4	Equivalence to the Cartesian velocity representation	139
11.8.5	Comments	139
11.9	Metric tensor	140
11.9.1	Jacobian of transformation	140
11.9.2	Covariant and contravariant representations	140
11.10	Volume element and the Levi-Civita tensor	141
11.11	Cross product of basis vectors	141
11.12	Partial derivative operators	141
11.12.1	Analytical derivation	141
11.12.2	Geometrical derivation	142
11.13	Material time derivative	143
11.14	Divergence of a vector and the divergence theorem	144
11.15	The diffusion operator	144
11.15.1	Continuous expression	144
11.15.2	Layer integrated expression	145
11.16	Vorticity	146
11.16.1	The components	146
11.16.2	Transforming from Cartesian coordinates	146
11.17	Velocity circulation	147

11.1 Introducing GVCs

In this chapter we present the mathematics of generalized vertical coordinates (GVC), with Figure 11.1 offering a schematic of how these coordinates monotonically partition the vertical direction. Such coordinates are of particular use for stratified fluid dynamics, where it is often convenient to make use of a vertical coordinate distinct from, but uniquely related to, the geopotential coordinate. For example, in hydrostatic compressible fluids, pressure is a convenient choice since it naturally absorbs the appearance of density in many formula, such as the geostrophic balance given by equation (29.1a) and the mass continuity equation discussed in Section 21.9.3. Hence, a natural expression of the compressible hydrostatic equations of motion make use of pressure rather than geopotential for the vertical coordinate. Two other common vertical coordinates include the isopy-

cnal coordinate and the terrain following coordinate. Notably, the Arbitrary Lagrangian-Eulerian (ALE) method is a relatively recent use for generalized vertical coordinates in numerical models of the ocean and atmosphere, and we offer a sketch of the method in Section 33.4. The ALE method offers the means to evolve the model state even without an explicit specification of the vertical coordinate.

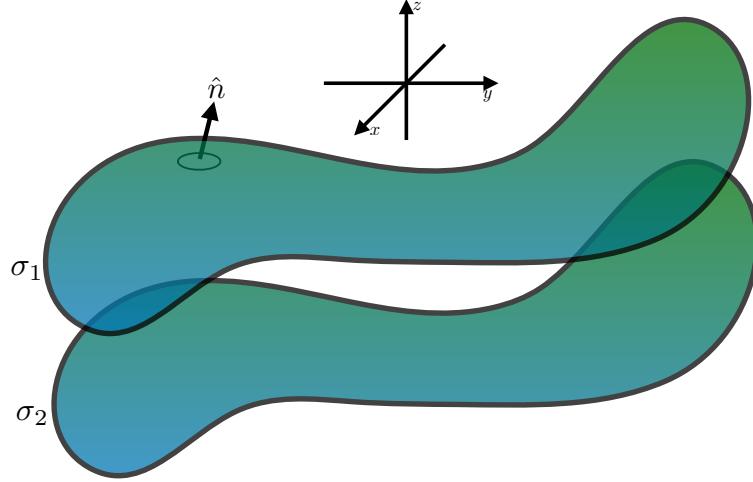


Figure 11.1: This schematic illustrates the geometry of two surfaces of constant generalized vertical coordinate $\sigma(x, y, z, t) = \sigma_1$ and $\sigma(x, y, z, t) = \sigma_2$, here showing patches on two such surfaces. The surfaces are generally undulating in space and time yet are assumed to monotonically layer the fluid. Monotonic layering means that the surface normal, \hat{n} , always has a non-zero projection onto the vertical: $\hat{n} \cdot \hat{z} \neq 0$. That is, the surfaces never become vertical nor do they overturn. It also means that there is a one-to-one invertible relation between σ and geopotential, so that specifying $(x, y, \sigma(x, y, z, t))$ is sufficient to yield a unique z .

11.1.1 Relating Cartesian and generalized vertical coordinates

We make use of the symbol σ for a generalized vertical coordinate. This coordinate is *not* orthogonal to the horizontal spatial coordinates x, y . This is a central property of GVCs that influences nearly all aspects of GVC calculus. To help develop the mathematics for transforming between Cartesian coordinates and GVCs, it is important to distinguish the two coordinate systems. One practical use of this distinction concerns the development of partial derivatives considered in Section 11.12.

To help keep track of the two coordinate systems, we write the time coordinate and spatial Cartesian coordinates according to

$$\xi^\alpha = (\xi^0, \xi^a) = (\xi^0, \xi^1, \xi^2, \xi^3) = (t, x, y, z) \text{ with } \alpha = 0, 1, 2, 3, \text{ and } a = 1, 2, 3. \quad (11.1)$$

As defined, the tensor label a runs over the spatial coordinates 1, 2, 3 whereas α also includes the time coordinate with $\alpha = 0$. For our considerations, we are mostly interested in how the spatial components of tensors transform under coordinate transformations. Hence, we will soon dispense with the time component, thereafter focusing just on the spatial components $a = 1, 2, 3$. However, it is useful to carry through some of the space-time formalism in particular for use in determining how the time partial derivatives are related (see Section 11.12).

The corresponding generalized vertical coordinates are denoted with an overbar

$$\xi^{\bar{\alpha}} = (\bar{\xi}^0, \bar{\xi}^1, \bar{\xi}^2, \bar{\xi}^3) = (\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (11.2)$$

The one-to-one coordinate transformation between Cartesian and GVC coordinates is written

$$\xi^{\bar{0}} = \xi^0 \iff \bar{t} = t \quad (11.3a)$$

$$\xi^{\bar{1}} = \xi^1 \iff \bar{x} = x \quad (11.3b)$$

$$\xi^{\bar{2}} = \xi^2 \iff \bar{y} = y \quad (11.3c)$$

$$\xi^{\bar{3}} = \sigma(t, x, y, z). \quad (11.3d)$$

Since the coordinate transformation is invertible, we can define the inverse to equation (11.3d), thus providing an expression for the vertical position of a given GVC surface

$$\xi^3 = \xi^3(\xi^{\bar{a}}) = z(\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (11.4)$$

This equation says that when we locate a position via the value of a chosen GVC surface (i.e., specify the value for σ), the geopotential position for that surface is a function of time, horizontal position, and the chosen value for the generalized vertical coordinate. For example, the vertical position of a pressure surface of chosen value p is given by the functional relation $\xi^3 = z(\bar{t}, \bar{x}, \bar{y}, p)$. The full inverse coordinate transformation takes the form

$$\xi^0 = \xi^{\bar{0}} \quad (11.5a)$$

$$\xi^1 = \xi^{\bar{1}} \quad (11.5b)$$

$$\xi^2 = \xi^{\bar{2}} \quad (11.5c)$$

$$\xi^3 = z(\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (11.5d)$$

11.1.2 A common confusion

The above discussion exposes a means for confusion. Namely, it is common in the literature to switch between using the symbol z to mean a particular geopotential; i.e., $z = -100\text{m}$, versus the symbol z as the geopotential position of a particular σ surface, $z(\bar{t}, \bar{x}, \bar{y}, \sigma)$. One way to be careful is to always write $\xi^3(\bar{t}, \bar{x}, \bar{y}, \sigma)$ rather than $z(\bar{x}, \bar{y}, \sigma, \bar{t})$. We maintain care in places where it is crucial to keep the meaning clear. However, developing a GVC brain muscle allows us to be relaxed when the meaning is clear.

11.2 Spatial basis vectors

Making use of the tensor formalism from Chapters 8 and 9, consider the transformation of the Cartesian basis vectors into their corresponding GVC representation. This transformation takes on the form

$$\vec{e}_{\bar{a}} = \Lambda_{\bar{a}}^a \vec{e}_a, \quad (11.6)$$

where the transformation matrix takes the form

$$\Lambda_{\bar{a}}^a = \begin{bmatrix} \partial x / \partial \bar{x} & \partial x / \partial \bar{y} & \partial x / \partial \sigma \\ \partial y / \partial \bar{x} & \partial y / \partial \bar{y} & \partial y / \partial \sigma \\ \partial z / \partial \bar{x} & \partial z / \partial \bar{y} & \partial z / \partial \sigma \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial z / \partial \bar{x} & \partial z / \partial \bar{y} & \partial z / \partial \sigma \end{bmatrix}. \quad (11.7)$$

The diagonal unit values for the space-space components arise since a horizontal position in Cartesian and GVCs is the same and the horizontal directions are orthogonal. Likewise, the time coordinate does not change when changing \bar{x} , \bar{y} , or σ . Additionally, $\partial x / \partial \sigma = \partial y / \partial \sigma = 0$ since the

horizontal position remains unchanged when moving across a GVC surface. In contrast, a non-zero value for $\partial z/\partial \bar{t}$ arises since we generally change the vertical position when following the temporal motion of a constant GVC surface. Likewise, $\partial z/\partial \bar{x}$ and $\partial z/\partial \bar{y}$ are nonzero since we generally change vertical position when moving horizontally along a sloped GVC surface. Finally, the element $\partial z/\partial \sigma$ is nonzero due to vertical stratification of the fluid when represented using GVCs.

11.2.1 A few more points on the transformation matrix

To further detail how to produce elements of the transformation matrix (11.7), it is crucial to ensure that the proper variables are held fixed when performing the partial derivatives. For example, consider the top row where we compute derivatives of the time coordinate

$$\Lambda_{\bar{a}}^1 = [[\partial x/\partial \bar{x}]_{\bar{y}, \sigma} \quad [\partial x/\partial \bar{y}]_{\bar{x}, \sigma} \quad [\partial x/\partial \sigma]_{\bar{x}, \bar{y}}] \quad (11.8)$$

Since $x = \bar{x}$, all elements vanish except for the first. Namely, $[\partial x/\partial \bar{y}]_{\bar{x}, \sigma} = 0$ since the x cannot change when \bar{x} is fixed. The same idea leads to the results for y (as well as t).

11.2.2 Expressions for the basis vectors

Use of the transformation matrix (11.7) renders the spatial components of the GVC basis vectors

$$\vec{e}_{\bar{1}} = \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial \bar{x}) \quad (11.9a)$$

$$\vec{e}_{\bar{2}} = \hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial \bar{y}) \quad (11.9b)$$

$$\vec{e}_{\bar{3}} = \hat{\mathbf{z}} (\partial z / \partial \sigma). \quad (11.9c)$$

The basis vectors $\vec{e}_{\bar{1}}$ and $\vec{e}_{\bar{2}}$ have a vertical component due to sloping GVC surfaces. These basis vectors lie within the tangent plane of the GVC surface. The basis vector $\vec{e}_{\bar{3}}$ is purely vertical and has a non-unit magnitude due to the inverse vertical stratification, $\partial z/\partial \sigma$. The left panel of Figure 11.2 illustrates the basis vectors.

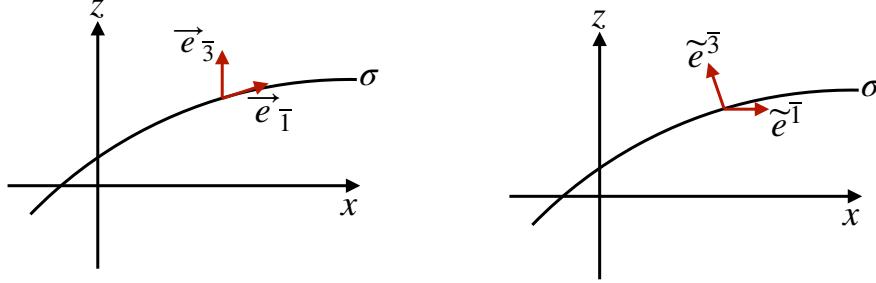


Figure 11.2: Illustrating the basis vectors (left panel) and basis one-forms (right panel) for generalized vertical coordinates. The $\vec{e}_{\bar{3}}$ basis vector is vertical whereas $\vec{e}_{\bar{1}}$ and $\vec{e}_{\bar{2}}$ lie within the tangent plane to the σ surface. As a complement, the basis one-form $\tilde{e}^{\bar{3}}$ is normal to the σ surface whereas the basis one-forms $\tilde{e}^{\bar{1}}$ and $\tilde{e}^{\bar{2}}$ are horizontal.

11.3 Basis one-forms

The basis one-forms are obtained by transforming from Cartesian into GVCs through use of the inverse transformation

$$\tilde{e}^{\bar{a}} = \Lambda_a^{\bar{a}} \tilde{e}^a, \quad (11.10)$$

where the inverse transformation matrix takes the form

$$\Lambda_{\bar{a}}^{\bar{a}} = \begin{bmatrix} \partial \bar{x}/\partial x & \partial \bar{x}/\partial y & \partial \bar{x}/\partial z \\ \partial \bar{y}/\partial x & \partial \bar{y}/\partial y & \partial \bar{y}/\partial z \\ \partial \sigma/\partial x & \partial \sigma/\partial y & \partial \sigma/\partial z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial \sigma/\partial x & \partial \sigma/\partial y & \partial \sigma/\partial z \end{bmatrix}. \quad (11.11)$$

As for the transformation matrix (11.7), 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. Finally, 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.

11.3.1 A few more points on the inverse transformation matrix

When computing elements of the inverse transformation matrix (11.11), it is crucial to ensure that the proper variables are held fixed. For example, consider the top row where we compute

$$\Lambda_a^{\bar{1}} = [[\partial \bar{x}/\partial x]_{y,z} \quad [\partial \bar{x}/\partial y]_{x,z} \quad [\partial \bar{x}/\partial z]_{x,y}]. \quad (11.12)$$

Just as for the transformation matrix (11.8), since $x = \bar{x}$, all but the first element vanish in equation (11.12). 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 (as well as \bar{t}).

11.3.2 GVC basis one-forms

Use of the inverse transformation matrix (11.11) renders the spatial components of the GVC basis one-forms

$$\tilde{e}^{\bar{1}} = \hat{x} \quad (11.13a)$$

$$\tilde{e}^{\bar{2}} = \hat{y} \quad (11.13b)$$

$$\tilde{e}^{\bar{3}} = \tilde{e}^a \partial_a \sigma = \hat{x} (\partial \sigma / \partial x) + \hat{y} (\partial \sigma / \partial y) + \hat{z} (\partial \sigma / \partial z) = \nabla \sigma. \quad (11.13c)$$

The left panel of Figure 11.2 illustrates the basis one-forms.

11.3.3 Verifying the orthogonality relation

The basis one-forms satisfy the orthogonality relation (8.18) with the basis vectors

$$\tilde{e}^{\bar{a}} \cdot \tilde{e}_{\bar{b}} = \delta_{\bar{b}}^{\bar{a}}. \quad (11.14)$$

This identity is trivial to verify for all $\bar{a} = 1, 2, 3$.

11.4 Showing that $\Lambda_{\bar{a}}^{\bar{a}} \Lambda_{\bar{b}}^a = \delta_{\bar{b}}^{\bar{a}}$ and $\Lambda_{\bar{a}}^a \Lambda_{\bar{b}}^{\bar{a}} = \delta_b^a$

We present two arguments to verify that the matrix (11.11) is indeed the inverse of the matrix (11.7). Both rely on writing the coordinate transformation as a composite function

$$\xi^{\bar{a}} = \xi^{\bar{a}}(\xi^a) = \xi^{\bar{a}}[\xi^a(\xi^{\bar{b}})]. \quad (11.15)$$

Taking partial derivatives and using the chain rule renders

$$\delta_{\bar{b}}^{\bar{a}} = \frac{\partial \xi^{\bar{a}}}{\partial \xi^{\bar{b}}} = \frac{\partial \xi^{\bar{a}}}{\partial \xi^a} \frac{\partial \xi^a}{\partial \xi^{\bar{b}}} = \Lambda_a^{\bar{a}} \Lambda_b^a \quad (11.16)$$

as well as

$$\delta_b^a = \frac{\partial \xi^a}{\partial \xi^b} = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} \frac{\partial \xi^{\bar{a}}}{\partial \xi^b} = \Lambda_{\bar{a}}^a \Lambda_b^{\bar{a}} \quad (11.17)$$

11.5 Triple product identities

We find various occasions to make use of a suite of triple product identities that hold for GVCs. For this purpose we write σ as a composite function as in Section 11.4

$$\sigma = \sigma(t, x, y, z) = \sigma[t, x, y, z(\bar{t}, \bar{x}, \bar{y}, \sigma)], \quad (11.18)$$

so that the chain rule leads to the differential increment

$$d\sigma = dt \left[\frac{\partial \sigma}{\partial t} \right]_{x,y,z} + dx \left[\frac{\partial \sigma}{\partial x} \right]_{t,y,z} + dy \left[\frac{\partial \sigma}{\partial y} \right]_{t,x,z} + dz \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y}. \quad (11.19)$$

Likewise, writing $z = z[\bar{t}, \bar{x}, \bar{y}, \sigma]$ leads to the differential increment dz

$$dz = d\bar{t} \left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} + d\bar{x} \left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} + d\bar{y} \left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} + d\sigma \left[\frac{\partial z}{\partial \sigma} \right]_{\bar{t}, \bar{x}, \bar{y}}. \quad (11.20)$$

We note the identities

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \sigma} \right]_{\bar{t}, \bar{x}, \bar{y}} = 1 \quad d\bar{t} = dt \quad d\bar{x} = dx \quad d\bar{y} = dy, \quad (11.21)$$

which follow since $t = \bar{t}$, $x = \bar{x}$, and $y = \bar{y}$. Substituting equation (11.20) into equation (11.19) and making use of the identities (11.21) yields

$$\begin{aligned} 0 &= dt \left[\left[\frac{\partial \sigma}{\partial t} \right]_{x,y,z} + \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} \right] \\ &\quad + dx \left[\left[\frac{\partial \sigma}{\partial x} \right]_{t,y,z} + \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} \right] + dy \left[\left[\frac{\partial \sigma}{\partial y} \right]_{t,x,z} + \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} \right]. \end{aligned} \quad (11.22)$$

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 identities¹

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} = - \left[\frac{\partial \sigma}{\partial t} \right]_{x,y,z} \quad (11.23a)$$

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} = - \left[\frac{\partial \sigma}{\partial x} \right]_{t,y,z} \quad (11.23b)$$

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} = - \left[\frac{\partial \sigma}{\partial y} \right]_{t,x,z} \quad (11.23c)$$

¹These identities are directly analogous to the Maxwell identities from thermodynamics (e.g., [Callen, 1985](#)).

If the vertical stratification, $\partial\sigma/\partial z$, is non-zero, the triple product identities are equivalent to

$$\left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} = - \frac{[\partial\sigma/\partial t]_{x,y,z}}{[\partial\sigma/\partial z]_{t,x,y}} = - \left[\frac{\partial\sigma}{\partial t} \right]_{x,y,z} \left[\frac{\partial z}{\partial\sigma} \right]_{t,x,y} \quad (11.24a)$$

$$\left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} = - \frac{[\partial\sigma/\partial x]_{t,y,z}}{[\partial\sigma/\partial z]_{t,x,y}} = - \left[\frac{\partial\sigma}{\partial x} \right]_{t,y,z} \left[\frac{\partial z}{\partial\sigma} \right]_{t,x,y} \quad (11.24b)$$

$$\left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} = - \frac{[\partial\sigma/\partial y]_{t,x,z}}{[\partial\sigma/\partial z]_{t,x,y}} = - \left[\frac{\partial\sigma}{\partial y} \right]_{t,x,z} \left[\frac{\partial z}{\partial\sigma} \right]_{t,x,y}. \quad (11.24c)$$

Since $t = \bar{t}$, $x = \bar{x}$, and $y = \bar{y}$ we can write these identities in the more succinct form

$$\left[\frac{\partial z}{\partial \bar{t}} \right]_\sigma = - \frac{[\partial\sigma/\partial t]_z}{[\partial\sigma/\partial z]} \quad (11.25a)$$

$$\left[\frac{\partial z}{\partial \bar{x}} \right]_\sigma = - \frac{[\partial\sigma/\partial x]_z}{[\partial\sigma/\partial z]} \quad (11.25b)$$

$$\left[\frac{\partial z}{\partial \bar{y}} \right]_\sigma = - \frac{[\partial\sigma/\partial y]_z}{[\partial\sigma/\partial z]}. \quad (11.25c)$$

These identities are quite useful for manipulating equations involving GVCs. In particular, equations (11.25b) and (11.25c) provide alternate expressions for the slope of constant σ surfaces relative to the horizontal plane (see Section 11.12).

11.6 Position vector

We are familiar with locating a point in space using Cartesian coordinates as in Figure 3.1. What about specifying the position using GVCs? We can do so by making use of the basis vectors (11.9a)-(11.9c) so that the position of an arbitrary point in space is given by

$$\mathcal{P} = \xi^{\bar{a}} \vec{e}_{\bar{a}} \quad (11.26a)$$

$$= \bar{x} [\hat{\mathbf{x}} + (\partial z / \partial \bar{x}) \hat{\mathbf{z}}] + \bar{y} [\hat{\mathbf{y}} + (\partial z / \partial \bar{y}) \hat{\mathbf{z}}] + \sigma (\partial z / \partial \sigma) \hat{\mathbf{z}} \quad (11.26b)$$

$$= \hat{\mathbf{x}} \bar{x} + \hat{\mathbf{y}} \bar{y} + \hat{\mathbf{z}} [\bar{x} (\partial z / \partial \bar{x}) + \bar{y} (\partial z / \partial \bar{y}) + \sigma (\partial z / \partial \sigma)] \quad (11.26c)$$

$$= \hat{\mathbf{x}} \bar{x} + \hat{\mathbf{y}} \bar{y} + \hat{\mathbf{z}} \xi^{\bar{a}} \partial_{\bar{a}} z. \quad (11.26d)$$

We identify the following properties as a means to help understand these expressions, with Figure 11.3 offering a schematic.

- The form (11.26b) has horizontal positions \bar{x} and \bar{y} multiplying the basis vectors $\vec{e}_{\bar{x}}$ and $\vec{e}_{\bar{y}}$, with these vectors oriented parallel to a surface of constant GVC as in Figure 11.3. Likewise, the third term, $\sigma (\partial z / \partial \sigma) \hat{\mathbf{z}}$, positions the point vertically according to the value of the GVC and its inverse stratification.
- Consider the case of $\bar{y} = 0$ so that

$$\mathcal{P} = \bar{x} \hat{\mathbf{x}} + \hat{\mathbf{z}} [\bar{x} (\partial z / \partial \bar{x}) + \sigma (\partial z / \partial \sigma)] \quad (11.27a)$$

$$= \bar{x} \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial \sigma) [\bar{x} (\partial \sigma / \partial z)_x (\partial z / \partial \bar{x})_\sigma + \sigma] \quad (11.27b)$$

$$= \bar{x} \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial \sigma) [-\bar{x} (\partial \sigma / \partial x)_z + \sigma], \quad (11.27c)$$

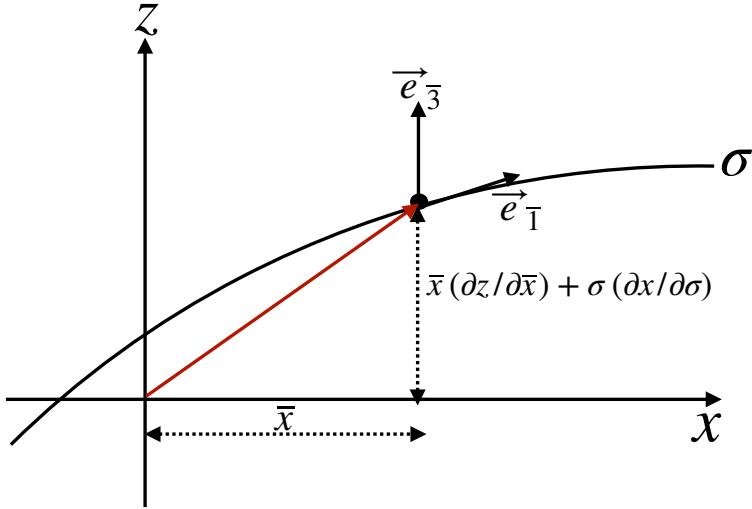


Figure 11.3: The position of a point in space as represented using GVCs following equation (11.27a). For this example, $\bar{y} = 0$ so that the horizontal position is determined by the coordinate $\bar{x} = x$, whereas the vertical position is determined by $\bar{x} (\partial z / \partial \bar{x}) + \sigma (\partial x / \partial \sigma)$.

where we used the triple product identity (11.25b) for the final equality. Consequently, a horizontal position vector is realized using GVC coordinates if $\sigma = \bar{x} (\partial \sigma / \partial x)$. Hence, a horizontal position vector crosses surfaces of constant GVC when the GVC surface has a nonzero horizontal slope.

- The projection of the position vector onto the basis one-forms leads to

$$\mathcal{P} \cdot \tilde{e}^{\bar{b}} = \xi^{\bar{a}} \tilde{e}_{\bar{a}} \cdot \tilde{e}^{\bar{b}} = \xi^{\bar{b}}. \quad (11.28)$$

This result follows from the orthogonality relation (8.18). So the projection of the position vector onto a basis one-form picks out the corresponding coordinate value.

- Equation (11.4) provides the spatial dependence for the vertical position of the surface of constant GVC

$$z = z(\xi^{\bar{a}}). \quad (11.29)$$

At any particular time instance we can perform a Taylor series about a reference depth z_0 , so that

$$z(\xi^{\bar{a}}) \approx z_0 + \xi^{\bar{a}} \partial_{\bar{a}} z. \quad (11.30)$$

We can thus write the position (11.26d) in the form

$$\mathcal{P} = \hat{x} \bar{x} + \hat{y} \bar{y} + \hat{z} [z - z_0]. \quad (11.31)$$

Taking the reference depth to be $z_0 = 0$ recovers the Cartesian expression. Since the position vector is a geometric object, it is reassuring that the GVC representation is the same as the Cartesian representation; it is merely a reorganization of the basis vectors and corresponding coordinate representation.

11.7 Transforming components of a first order tensor

Consider a vector field \vec{F} with Cartesian representation

$$\vec{F} = \mathbf{F} = F^a \vec{e}_a = F^x \hat{x} + F^y \hat{y} + F^z \hat{z}. \quad (11.32)$$

The corresponding GVC components are related through the transformation matrix

$$F^{\bar{a}} = \Lambda_{\bar{a}}^{\bar{a}} F^a. \quad (11.33)$$

Making use of the transformation matrix (11.11) yields the relations between GVC components and Cartesian components

$$F^{\bar{1}} = F^1 \quad F^{\bar{2}} = F^2 \quad F^{\bar{3}} = \nabla \sigma \cdot \mathbf{F}, \quad (11.34)$$

where we wrote

$$\nabla \sigma \cdot \mathbf{F} = (\partial \sigma / \partial x) F^1 + (\partial \sigma / \partial y) F^2 + (\partial \sigma / \partial z) F^3. \quad (11.35)$$

The vector field thus can be represented in GVC coordinates as

$$\vec{F} = F^{\bar{a}} \vec{e}_{\bar{a}} = F^1 \vec{e}_{\bar{1}} + F^2 \vec{e}_{\bar{2}} + (\nabla \sigma \cdot \mathbf{F}) \vec{e}_{\bar{3}}. \quad (11.36)$$

Similarly, the covariant components transform as $F_{\bar{a}} = \Lambda_{\bar{a}}^a F_a$, where use of the inverse transformation matrix (11.11) renders

$$F_{\bar{1}} = F_1 + \frac{\partial z}{\partial \bar{x}} F_3 \quad F_{\bar{2}} = F_2 + \frac{\partial z}{\partial \bar{y}} F_3 \quad F_{\bar{3}} = \frac{\partial z}{\partial \sigma} F_3, \quad (11.37)$$

and the expression for the vector field

$$\vec{F} = F_{\bar{a}} \tilde{e}^{\bar{a}} = [F_1 + (\partial z / \partial \bar{x}) F_3] \tilde{e}^{\bar{1}} + [F_2 + (\partial z / \partial \bar{y}) F_3] \tilde{e}^{\bar{2}} + (\partial z / \partial \sigma) F_3 \tilde{e}^{\bar{3}}. \quad (11.38)$$

Recall also that for Cartesian coordinates the contravariant and covariant components to a vector are identical: $F^a = F_a$.

11.8 Velocity

We now make use of the results from Section 11.7 to represent the velocity vector, considering both covariant and contravariant representations. As for the position vector detailed in Section 11.6, we are assured that both representations lead to the same velocity vector since the velocity is an objective geometric object (i.e., an arrow with a magnitude). In Section 11.8.4 we verify that the transformation formalism indeed respects this equivalence, with the GVC representations equivalent to the Cartesian representation

$$\vec{v} = u \hat{x} + v \hat{y} + w \hat{z}. \quad (11.39)$$

11.8.1 Contravariant components to the velocity

Following Section 11.7, we have the contravariant velocity components

$$v^{\bar{1}} = u \quad v^{\bar{2}} = v \quad v^{\bar{3}} = \mathbf{v} \cdot \nabla \sigma. \quad (11.40)$$

Use of the basis vectors (11.9a)-(11.9c) then leads to

$$\vec{v} = v^{\bar{a}} \vec{e}_{\bar{a}} \quad (11.41a)$$

$$= u \vec{e}_{\bar{x}} + v \vec{e}_{\bar{y}} + (\mathbf{v} \cdot \nabla \sigma) \vec{e}_{\bar{\sigma}} \quad (11.41b)$$

$$= u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] + (\mathbf{v} \cdot \nabla \sigma) (\partial z / \partial \sigma) \hat{z}. \quad (11.41c)$$

11.8.2 Covariant components to the velocity vector

The covariant velocity components are given by

$$v_{\bar{1}} = u + \frac{\partial z}{\partial \bar{x}} w \quad v_{\bar{2}} = v + \frac{\partial z}{\partial \bar{y}} w \quad v_{\bar{3}} = \frac{\partial z}{\partial \sigma} w. \quad (11.42)$$

The one-form basis (11.13a)–(11.13c) thus leads to the velocity vector

$$\vec{v} = v_{\bar{a}} \tilde{e}^{\bar{a}} = [u + (\partial z / \partial \bar{x}) w] \hat{x} + [v + (\partial z / \partial \bar{y}) w] \hat{y} + w (\partial z / \partial \sigma) \nabla \sigma. \quad (11.43)$$

11.8.3 Introducing the material time derivative

The material evolution for the generalized vertical coordinate can be written

$$\frac{D\sigma}{Dt} = \frac{\partial \sigma}{\partial t} + \mathbf{v} \cdot \nabla \sigma = \dot{\sigma}, \quad (11.44)$$

with $\dot{\sigma}$ symbolizing any process contributing to dia-surface transfer. If the GVC is pressure, $\sigma = p$, then $\dot{\sigma} = \dot{p}$ can arise from reversible motion such as linear waves, as well as irreversible processes such as mixing. If the GVC is an isopycnal, then $\dot{\sigma} \neq 0$ typically arises only from irreversible processes such as mixing. Using the expression (11.44) in the velocity vector expression (11.41c) leads to

$$\vec{v} = u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] + (\mathbf{v} \cdot \nabla \sigma) (\partial z / \partial \sigma) \hat{z}. \quad (11.45a)$$

$$= u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] + (\dot{\sigma} - \partial \sigma / \partial t) (\partial z / \partial \sigma) \hat{z} \quad (11.45b)$$

$$= u \hat{x} + v \hat{y} + [\partial z / \partial \bar{t} + \mathbf{u} \cdot \nabla_{\sigma} z + (\partial z / \partial \sigma) \dot{\sigma}] \hat{z}, \quad (11.45c)$$

where the final equality made use of the triple product (11.24a): $(\partial \sigma / \partial t) (\partial z / \partial \sigma) = -\partial z / \partial \bar{t}$. In the steady state and in the absence of material changes to σ , the three dimensional flow lies within a surface of constant σ , whereby $\mathbf{v} \cdot \nabla \sigma = 0$ and

$$\vec{v} = u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] \quad \text{if } \partial_t \sigma = 0 \text{ and } \dot{\sigma} = 0. \quad (11.46)$$

However, in general there are transient fluctuations and material changes to σ so that $\mathbf{v} \cdot \nabla \sigma \neq 0$.

11.8.4 Equivalence to the Cartesian velocity representation

Use of the triple product identities (11.25b)–(11.25c) allows us to manipulate both expressions (11.41c) and (11.43) to recover the Cartesian expression

$$\vec{v} = u \hat{x} + v \hat{y} + w \hat{z}. \quad (11.47)$$

Another way to see this identity is to note that in equation (11.45c), the vertical component is an expression for the material time derivative of the vertical position

$$w = \frac{Dz}{Dt} = \frac{\partial z}{\partial \bar{t}} + \mathbf{u} \cdot \nabla_{\sigma} z + \frac{\partial z}{\partial \sigma} \dot{\sigma}. \quad (11.48)$$

We derive this identity in Section 21.4 where we discuss further kinematic results using GVCs.

11.8.5 Comments

In Chapter 21 we have far more to say on the GVC expressions for the velocity vector.

11.9 Metric tensor

Recall from Section 8.1 that we make use of a metric tensor to measure the distance between two points in space. The GVC representation of the metric tensor is given by

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} 1 + (\partial z / \partial \bar{x})^2 & (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) & (\partial z / \partial \bar{x})(\partial z / \partial \sigma) \\ (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) & 1 + (\partial z / \partial \bar{y})^2 & (\partial z / \partial \bar{y})(\partial z / \partial \sigma) \\ (\partial z / \partial \bar{x})(\partial z / \partial \sigma) & (\partial z / \partial \bar{y})(\partial z / \partial \sigma) & (\partial z / \partial \sigma)^2 \end{bmatrix}, \quad (11.49)$$

and the GVC representation of the inverse metric tensor is given by the somewhat simpler form

$$g^{\bar{a}\bar{b}} = \vec{e}^{\bar{a}} \cdot \vec{e}^{\bar{b}} = \begin{bmatrix} 1 & 0 & \partial \sigma / \partial x \\ 0 & 1 & \partial \sigma / \partial y \\ \partial \sigma / \partial x & \partial \sigma / \partial y & |\nabla \sigma|^2 \end{bmatrix}. \quad (11.50)$$

Proof that $g^{\bar{a}\bar{b}} g_{\bar{b}\bar{c}} = \delta_{\bar{a}}^{\bar{c}}$ requires use of the triple product identities (11.25b) and (11.25c).

11.9.1 Jacobian of transformation

The determinant of the GVC representation of the metric tensor (11.49) is

$$\det(g_{\bar{a}\bar{b}}) = (\partial z / \partial \sigma)^2 \quad (11.51)$$

so that the Jacobian of transformation (Section 8.4) is the specific thickness

$$\frac{\partial(x, y, z)}{\partial(\bar{x}, \bar{y}, \sigma)} = \frac{\partial z}{\partial \sigma}. \quad (11.52)$$

The coordinate transformation from Cartesian to generalized vertical is invertible only so long as the Jacobian remains nonzero and single-signed, meaning the fluid retains a monotonic vertical stratification of GVC surfaces. The invertible relation between z and σ means that each point in the vertical can be uniquely specified by either of the two vertical coordinates, z or σ . For example, the Jacobian for pressure as the generalized vertical coordinate in a hydrostatic fluid is given by²

$$\frac{\partial z}{\partial \sigma} = \frac{\partial z}{\partial p} = -\frac{1}{\rho g}, \quad (11.53)$$

which is indeed single-signed since the mass density is always positive.

11.9.2 Covariant and contravariant representations

The metric tensor allows us to convert between the covariant and contravariant representations of a vector via the identity (Section 8.2.3)

$$F_{\bar{a}} = g_{\bar{a}\bar{b}} F^{\bar{b}}. \quad (11.54)$$

We use triple product identities (11.25b)-(11.25c) to verify that this relation agrees with the transformation matrix approach detailed in Section 11.7. For example,

$$F_{\bar{1}} = g_{\bar{1}\bar{b}} F^{\bar{b}} \quad (11.55a)$$

$$= [1 + (\partial z / \partial \bar{x})^2] F^{\bar{1}} + (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) F^{\bar{2}} + (\partial z / \partial \bar{x})(\partial z / \partial \sigma) F^{\sigma} \quad (11.55b)$$

$$= [1 + (\partial z / \partial \bar{x})^2] F^1 + (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) F^2 + (\partial z / \partial \bar{x})(\partial z / \partial \sigma) \nabla \sigma \cdot \mathbf{F} \quad (11.55c)$$

$$= F^1 + (\partial z / \partial \bar{x}) F^3 \quad (11.55d)$$

$$= F_1 + (\partial z / \partial \bar{x}) F_3, \quad (11.55e)$$

²We derive the hydrostatic balance in Section 27.3.

where the final equality holds since $F^1 = F_1$ and $F^3 = F_3$ for Cartesian tensor components.

11.10 Volume element and the Levi-Civita tensor

The square root of the determinant of the metric tensor (11.49) is

$$\sqrt{\det(g_{\bar{a}\bar{b}})} = \partial z / \partial \sigma \quad (11.56)$$

so that the volume element (Section 8.4) is

$$dV = (\partial z / \partial \sigma) dx dy d\sigma. \quad (11.57)$$

The covariant Levi-Civita tensor (Section 8.6) has the GVC representations

$$\varepsilon_{\bar{a}\bar{b}\bar{c}} = (\partial z / \partial \sigma) \epsilon_{\bar{a}\bar{b}\bar{c}} \quad \varepsilon^{\bar{a}\bar{b}\bar{c}} = (\partial z / \partial \sigma)^{-1} \epsilon^{\bar{a}\bar{b}\bar{c}} \quad (11.58)$$

ϵ is the permutation symbol introduced in Section 3.4.1 with its components independent of coordinate representation.

11.11 Cross product of basis vectors

We now verify the relation (8.56) for the cross product of two basis vectors using GVCs

$$\vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}} \implies \vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = (\partial z / \partial \sigma) \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}}. \quad (11.59)$$

Making use of the basis vectors from Section 11.2 and the basis one-forms from Section 11.3 renders

$$\vec{e}_{\bar{x}} \wedge \vec{e}_{\bar{y}} = \hat{\mathbf{z}} - \hat{\mathbf{x}} (\partial z / \partial \bar{x}) - \hat{\mathbf{y}} (\partial z / \partial \bar{y}) = (\partial z / \partial \sigma) \nabla \sigma = \varepsilon_{\bar{x}\bar{y}\sigma} \tilde{e}^{\sigma} \quad (11.60a)$$

$$\vec{e}_{\bar{y}} \wedge \vec{e}_{\sigma} = \hat{\mathbf{x}} (\partial z / \partial \sigma) = \varepsilon_{\bar{y}\sigma\bar{x}} \tilde{e}^{\bar{x}} \quad (11.60b)$$

$$\vec{e}_{\sigma} \wedge \vec{e}_{\bar{x}} = \hat{\mathbf{y}} (\partial z / \partial \sigma) = \varepsilon_{\sigma\bar{x}\bar{y}} \tilde{e}^{\bar{y}}. \quad (11.60c)$$

11.12 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. Notably, the geometric derivation in Section 11.12.2 requires minimal use of the previous tensor formalism.

11.12.1 Analytical derivation

The partial derivative operators in GVCs are computed via $\partial_{\bar{a}} = \Lambda_{\bar{a}}^a \partial_a$. Including also the time component leads to the relations

$$\partial_{\bar{t}} = \partial_t + (\partial z / \partial \bar{t}) \partial_z \quad \partial_{\bar{x}} = \partial_x + (\partial z / \partial \bar{x}) \partial_z \quad \partial_{\bar{y}} = \partial_y + (\partial z / \partial \bar{y}) \partial_z \quad \partial_{\sigma} = (\partial z / \partial \sigma) \partial_z. \quad (11.61)$$

We can make use of the triple product identities (11.25b) and (11.25c) to express the slope of a constant GVC surface in two equivalent manners

$$\nabla_\sigma z = -\frac{\nabla_z \sigma}{\partial \sigma / \partial z} \quad (11.62)$$

where we introduced the standard shorthand notation

$$\nabla_\sigma = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} \quad \nabla_z = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y}. \quad (11.63)$$

It is important to note that ∇_σ is merely a shorthand for the two partial derivative operators, and that it furthermore only has components in the horizontal directions. It is common to transform between the horizontal gradient operators, in which case we write

$$\nabla_\sigma = \nabla_z + (\nabla_\sigma z) \partial_z \equiv \nabla_z + \mathbf{S} \partial_z, \quad (11.64)$$

where we introduced the notation for the slope of the σ surface relative to the horizontal plane

$$\mathbf{S} = \nabla_\sigma z = -\frac{\nabla_z \sigma}{\partial \sigma / \partial z}. \quad (11.65)$$

11.12.2 Geometrical derivation

We provide a geometric derivation for the lateral derivative operator that complements the previous analytical derivation. 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 11.4. This feature of the horizontal derivative operator is a key aspect of the GVCs' non-orthogonality.

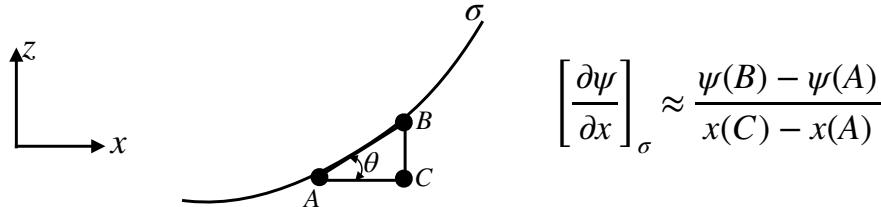


Figure 11.4: A surface of constant generalized vertical coordinate, σ , along with a local tangent plane with a slope $\tan \theta$ with respect to the horizontal plane. This figure illustrates the identities (11.67a)-(11.67d), with these identities relating a lateral derivative taken along the GVC surface to horizontal and vertical derivatives taken along orthogonal Cartesian axes.

Consider the geometry shown in Figure 11.4, which shows a generalized vertical coordinate surface (constant σ surface) along with a sample tangent plane with a slope

$$S^{(x)} = \frac{\text{rise}}{\text{run}} = \tan \theta = \frac{z(B) - z(C)}{x(C) - x(A)} \approx \left[\frac{\partial z}{\partial x} \right]_\sigma = -\frac{(\partial \sigma / \partial x)_z}{(\partial \sigma / \partial z)} \quad (11.66)$$

relative to the horizontal. We readily verify the following identities based on finite difference

operations for an arbitrary function

$$\left[\frac{\partial \psi}{\partial x} \right]_{\sigma} \approx \frac{\psi(B) - \psi(A)}{x(C) - x(A)} \quad (11.67a)$$

$$= \frac{\psi(C) - \psi(A)}{x(C) - x(A)} + \frac{\psi(B) - \psi(C)}{x(C) - x(A)} \quad (11.67b)$$

$$= \frac{\psi(C) - \psi(A)}{x(C) - x(A)} + \left[\frac{z(B) - z(C)}{x(C) - x(A)} \right] \frac{\psi(B) - \psi(C)}{z(B) - z(C)} \quad (11.67c)$$

$$= \left[\frac{\partial \psi}{\partial x} \right]_z + S^{(x)} \left[\frac{\partial \psi}{\partial z} \right]_x. \quad (11.67d)$$

Taking the continuum limit then leads to the relations between horizontal derivatives computed on constant σ surfaces to those computed on constant z surfaces

$$\left[\frac{\partial}{\partial x} \right]_{\sigma} = \left[\frac{\partial}{\partial x} \right]_z + \left[\frac{\partial z}{\partial x} \right]_{\sigma} \frac{\partial}{\partial z} \quad (11.68a)$$

$$\left[\frac{\partial}{\partial y} \right]_{\sigma} = \left[\frac{\partial}{\partial y} \right]_z + \left[\frac{\partial z}{\partial y} \right]_{\sigma} \frac{\partial}{\partial z}, \quad (11.68b)$$

which can be written in the shorthand vector notation

$$\nabla_{\sigma} = \hat{x} \left[\frac{\partial}{\partial x} \right]_{\sigma} + \hat{y} \left[\frac{\partial}{\partial y} \right]_{\sigma} = \nabla_z + (\nabla_{\sigma} z) \partial_z. \quad (11.69)$$

11.13 Material time derivative

Making use of the relations for the partial derivative operators in Section 11.12 allows us to write the material time derivative in the following equivalent forms

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (11.70a)$$

$$= \left[\frac{\partial}{\partial t} \right]_{\sigma} - (\partial z / \partial t) \partial_z + \mathbf{u} \cdot [\nabla_{\sigma} - (\nabla_{\sigma} z) \partial_z] + w \partial / \partial z \quad (11.70b)$$

$$= \left[\frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + [w - \mathbf{u} \cdot \nabla_{\sigma} z - \partial z / \partial t] (\partial \sigma / \partial z) \partial / \partial \sigma \quad (11.70c)$$

$$= \left[\frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + \frac{D\sigma}{Dt} \frac{\partial}{\partial \sigma} \quad (11.70d)$$

$$= \left[\frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} \frac{\partial}{\partial z}. \quad (11.70e)$$

The equality (11.70d) made use of the identity (11.48), which is itself derived in Section 21.4 where we discuss further kinematic results using GVCs. Besides differences in the spatial operators, it is important to note that the time derivative operators are computed on constant geopotential and constant GVC surfaces, respectively. However, the horizontal velocity component is the *same* for both forms of the material time derivative

$$(u, v) = \frac{D(x, y)}{Dt}. \quad (11.71)$$

11.14 Divergence of a vector and the divergence theorem

Making use of the general expression (9.7) for the covariant divergence of a vector renders the GVC expression

$$\nabla_{\bar{a}} F^{\bar{a}} = [\det(g_{\bar{a}\bar{b}})]^{-1/2} \partial_{\bar{a}} \left[[\det(g_{\bar{a}\bar{b}})]^{1/2} F^{\bar{a}} \right] = (\partial z / \partial \sigma)^{-1} \partial_{\bar{a}} [(\partial z / \partial \sigma) F^{\bar{a}}]. \quad (11.72)$$

Recall that the GVC vector components, $F^{\bar{a}}$, are related to the Cartesian components in equation (11.34), and the GVC components of the partial derivative operator, $\partial_{\bar{a}}$, are related to the Cartesian operator in equation (11.63).

When making use of the divergence theorem (Section 9.11), we require the product of the volume element and the covariant divergence. For GVCs this product takes on the form

$$(\nabla_{\bar{a}} F^{\bar{a}}) dV = \partial_{\bar{a}} [(\partial z / \partial \sigma) F^{\bar{a}}] d\bar{x} d\bar{y} d\sigma, \quad (11.73)$$

which reduces to a boundary integral when integrating over a volume.

11.15 The diffusion operator

As an explicit example of the covariant divergence operator (11.72), we here consider the diffusion operator discussed in Chapter 37. The derivation here recovers much of what we just discussed in Section 11.14, yet we make use of a bit less tensor formalism though at the cost of more algebra.

11.15.1 Continuous expression

The diffusion operator is the convergence of the diffusive flux

$$\mathcal{R} = -\nabla \cdot \mathbf{J}, \quad (11.74)$$

where \mathbf{J} is a vector field. Let us convert the pieces of this operator from Cartesian coordinates into GVC coordinates, making use of the transformation of partial derivative operators given in Section 11.12

$$-\mathcal{R} = \nabla \cdot \mathbf{J} \quad (11.75a)$$

$$= \nabla_z \cdot \mathbf{J}^h + \partial_z J^z \quad (11.75b)$$

$$= (\nabla_{\sigma} - \nabla_{\sigma} z \partial_z) \cdot \mathbf{J}^h + (\sigma_z) \partial_{\sigma} J^z \quad (11.75c)$$

$$= \sigma_z \left[z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}^h + (\hat{z} \partial_{\sigma} - \nabla_{\sigma} z \partial_{\sigma}) \cdot \mathbf{J} \right] \quad (11.75d)$$

$$= \sigma_z \left[\nabla_{\sigma} \cdot (z_{\sigma} \mathbf{J}^h) - \mathbf{J}^h \cdot \nabla_{\sigma} (z_{\sigma}) + \partial_{\sigma} J^z - \partial_{\sigma} (\nabla_{\sigma} z \cdot \mathbf{J}) + \mathbf{J} \cdot \partial_{\sigma} (\nabla_{\sigma} z) \right] \quad (11.75e)$$

$$= \sigma_z \left[\nabla_{\sigma} \cdot (z_{\sigma} \mathbf{J}^h) + \partial_{\sigma} J^z - \partial_{\sigma} (\nabla_{\sigma} z \cdot \mathbf{J}^h) \right] \quad (11.75f)$$

$$= \sigma_z \left(\nabla_{\sigma} \cdot (\partial_{\sigma} z \mathbf{J}^h) + \partial_{\sigma} [(\hat{z} - \nabla_{\sigma} z) \cdot \mathbf{J}] \right) \quad (11.75g)$$

$$= \sigma_z \left[\nabla_{\sigma} \cdot (z_{\sigma} \mathbf{J}^h) + \partial_{\sigma} (z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}) \right], \quad (11.75h)$$

where we used

$$z_{\sigma} \nabla_{\sigma} = \hat{z} - \nabla_{\sigma} z \quad (11.76)$$

to reach the final equality, and made use of the shorthand

$$z_\sigma = \frac{\partial z}{\partial \sigma} \quad \sigma_z = \frac{\partial \sigma}{\partial z}. \quad (11.77)$$

Making use of the coordinate transformations in Section 11.7 for vector components reveals that the expression (11.75h) is identical to equation (11.72) derived using formal tensor methods. Likewise, when multiplying by the volume element

$$\delta V = \delta x \delta y \delta z = \delta x \delta y z_\sigma \delta \sigma, \quad (11.78)$$

we are led to

$$-\mathcal{R} \delta V = \left[\nabla_\sigma \cdot (z_\sigma \mathbf{J}^h) + \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] \delta x \delta y \delta \sigma, \quad (11.79)$$

which is identical to the expression (11.73).

11.15.2 Layer integrated expression

The increment $\delta \sigma$ commutes with the horizontal operator ∇_σ , so that

$$-\mathcal{R} \delta V = \left[\nabla_\sigma \cdot (\delta \sigma z_\sigma \mathbf{J}^h) + \delta \sigma \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] \delta x \delta y \quad (11.80a)$$

$$= \frac{1}{\delta z} \left[\nabla_\sigma \cdot (\delta \sigma z_\sigma \mathbf{J}^h) + \delta \sigma \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] \delta x \delta y \delta z \quad (11.80b)$$

$$= \frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \Delta_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] \delta x \delta y h^\sigma, \quad (11.80c)$$

where we introduced the layer thickness

$$h^\sigma = z_\sigma \delta \sigma \quad (11.81)$$

and the non-dimensional differential operator

$$\Delta_\sigma \equiv \delta \sigma \frac{\partial}{\partial \sigma}. \quad (11.82)$$

Cancelling the volume element on both sides leads to the diffusion operator

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \Delta_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right]. \quad (11.83)$$

This form is commonly found in the ocean modeling literature when considering layered models of ocean circulation.

We make the following comments concerning the diffusion operator in equation (11.83).

- Our introduction of the layer thickness $h^\sigma = z_\sigma \delta \sigma$ is treated more formally in Sections 21.9 and 21.10 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, $h^\sigma \mathbf{J}^h$, is oriented within the horizontal plane. However, its contribution to the diffusion operator is computed by taking its convergence using the operator ∇_σ rather than the horizontal operator ∇_z . 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

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) \right] \quad \text{if } \nabla\sigma \cdot \mathbf{J} = 0. \quad (11.84)$$

The neutral diffusion operator of Section 39.3.3 is an example of such an operator, with σ in that case given by the locally referenced potential density.

11.16 Vorticity

As detailed in Chapter 48, vorticity is the curl of the velocity

$$\vec{\omega} = \text{curl}(\vec{v}), \quad (11.85)$$

where the curl has components (Section 9.9)

$$\text{curl}(\vec{v}) = \tilde{e}_a \varepsilon^{abc} \partial_b v_c = \tilde{e}_{\bar{a}} \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}}. \quad (11.86)$$

11.16.1 The components

We identify the contravariant components of the vorticity via

$$\omega^{\bar{a}} = \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}} = (\partial z / \partial \sigma)^{-1} \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}} \quad (11.87)$$

where we made use of equation (11.58) to introduce the permutation symbol. Expanding the components leads to

$$\omega^{\bar{1}} = (\partial \sigma / \partial z) (\partial_{\bar{2}} v_{\bar{3}} - \partial_{\bar{3}} v_{\bar{2}}) \quad (11.88a)$$

$$\omega^{\bar{2}} = (\partial \sigma / \partial z) (\partial_{\bar{3}} v_{\bar{1}} - \partial_{\bar{1}} v_{\bar{3}}) \quad (11.88b)$$

$$\omega^{\bar{3}} = \omega^\sigma = (\partial \sigma / \partial z) (\partial_{\bar{1}} v_{\bar{2}} - \partial_{\bar{2}} v_{\bar{1}}). \quad (11.88c)$$

11.16.2 Transforming from Cartesian coordinates

The above approach works solely with the GVC coordinates. An alternative approach connects the GVC vorticity components and the Cartesian vorticity components. For that purpose we use the transformation matrix via

$$\omega^{\bar{a}} = \Lambda_a^{\bar{a}} \omega^a, \quad (11.89)$$

where ω^a are the Cartesian components

$$\boldsymbol{\omega} = \hat{x} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) + \hat{y} \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right) + \hat{z} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right). \quad (11.90)$$

Making use of the transformation matrix $\Lambda_a^{\bar{a}}$ from equation (11.11) yields (as in Section 11.7)

$$\omega^{\bar{x}} = \omega^x = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \quad \omega^{\bar{y}} = \omega^y = \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \quad \omega^\sigma = \boldsymbol{\omega} \cdot \nabla \sigma. \quad (11.91)$$

Note that for isopycnal coordinates in a Boussinesq fluid, ω^σ equals to the potential vorticity when the vorticity is the absolute vorticity (Section 52.2). That is, the potential vorticity is the isopycnal component of the absolute vorticity.

11.17 Velocity circulation

The velocity circulation (Section 46.3) is given by the closed oriented path integral of the velocity projected into the direction of the path

$$\mathcal{C} \equiv \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} \quad (11.92)$$

where $d\mathbf{r}$ 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 from Section 4.6 leads to the identity

$$\mathcal{C} = \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} = \int_{\mathcal{S}} (\nabla \wedge \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (11.93)$$

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 are all written in a generally covariant manner (Section 7.2) so that they hold for an arbitrary coordinate representation.

As a particular case, consider the circulation around a closed path on a constant σ surface, in which

$$\hat{\mathbf{n}} = \frac{\nabla\sigma}{|\nabla\sigma|} \quad (11.94)$$

is the outward normal and

$$\boldsymbol{\omega} \cdot \hat{\mathbf{n}} = \frac{\omega^\sigma}{|\nabla\sigma|} \quad (11.95)$$

where $\omega^\sigma = \boldsymbol{\omega} \cdot \nabla\sigma$ (equation (11.91)). So long as the vertical stratification remains non-zero ($\partial\sigma/\partial z \neq 0$) we can write the area factor in the form

$$\frac{d\mathcal{S}}{|\nabla\sigma|} = \frac{d\mathcal{S}}{\sqrt{(\partial\sigma/\partial x)^2 + (\partial\sigma/\partial y)^2 + (\partial\sigma/\partial z)^2}} \quad (11.96a)$$

$$= \frac{d\mathcal{S}}{|\partial\sigma/\partial z| \sqrt{[(\partial\sigma/\partial x)/(\partial\sigma/\partial z)]^2 + [(\partial\sigma/\partial y)/(\partial\sigma/\partial z)]^2 + 1}} \quad (11.96b)$$

$$= \frac{d\mathcal{S}}{|\partial\sigma/\partial z| \sqrt{1 + \tan^2 \theta}} \quad (11.96c)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| |\cos \theta| d\mathcal{S} \quad (11.96d)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| dA. \quad (11.96e)$$

The equality (11.96c) introduces the angle, θ , between the boundary surface and the horizontal plane as in Figure 11.4. The squared slope of this surface given by

$$\tan^2 \theta = \frac{\nabla_z \sigma \cdot \nabla_z \sigma}{(\partial\sigma/\partial z)^2} = \nabla_\sigma z \cdot \nabla_\sigma z. \quad (11.97)$$

The equality (11.96d) made use of a trigonometric identity, and the equality (11.96e) introduced the horizontal projection of the area,

$$dA = |\cos \theta| d\mathcal{S}. \quad (11.98)$$

Bringing these results together leads to the expression for circulation around a closed loop on a constant σ surface

$$\mathcal{C}_{\sigma-\text{surface}} = \int_{\mathcal{S}} (\boldsymbol{\omega} \cdot \nabla\sigma) |\partial z/\partial \sigma| dA. \quad (11.99)$$

12

Tracer coordinates

In this chapter we generalize the GVC considerations of Chapter 11 by considering three tracer coordinates. If each tracer is materially constant and linearly independent of the other tracers, then tracer coordinates can be used as Lagrangian coordinates (Chapter 16). However, real fluids contain irreversible mixing so that fluid elements do not maintain a fixed tracer concentration. Hence, in the presence of mixing, a tracer description is not Lagrangian. However, one can still make use of the tracer description for investigating circulation viewed in tracer space. From a mathematical perspective, it is key to note that the tracer coordinates are all mutually non-orthogonal and thus require extra care relative to the orthogonal coordinates considered in Chapter 10 and the generalized vertical coordinates of Chapter 11.

READER'S GUIDE TO THIS CHAPTER

This chapter assumes a working knowledge of material in the generalized vertical coordinate Chapter 11 as well as the treatment of general tensors in Chapters 8 and 9. No other chapter makes use of this chapter.

12.1 Incomplete threads	149
12.2 Relating Cartesian coordinates and tracer coordinates	150
12.3 The transformation matrix and its inverse	150
12.4 Determinant of the transformation	151
12.5 Basis vectors	151
12.6 Basis one-forms	152
12.7 Trajectory or position vector	152
12.8 Velocity	152
12.9 Metric tensor	153
12.10 Vorticity	153
12.11 Comments	153

12.1 Incomplete threads

- Equations of motion
- Mass conservation and the tracer equations

- Diffusion equation
- Potential vorticity
- Angular momentum
- Eddy and mean flow in tracer space

12.2 Relating Cartesian coordinates and tracer coordinates

To develop the formalism we write the three Cartesian coordinates as, $\xi^a = (x, y, z)$, and the three tracer coordinates as, $\xi^{\bar{a}} = (C^{\bar{1}}, C^{\bar{2}}, C^{\bar{3}})$. Each tracer coordinate need not have the same physical dimensions. For example, they could represent temperature, salinity, and age. Including time leads to the coordinate transformation

$$\xi^{\bar{0}} = \xi^0 \iff \bar{t} = t \quad (12.1a)$$

$$\xi^{\bar{1}} = C^{\bar{1}}(\xi^a, t) \quad (12.1b)$$

$$\xi^{\bar{2}} = C^{\bar{2}}(\xi^a, t) \quad (12.1c)$$

$$\xi^{\bar{3}} = C^{\bar{3}}(\xi^a, t). \quad (12.1d)$$

The inverse functional relation yields the position of a point in space in the form

$$t = \xi^0 = \xi^{\bar{0}} \iff t = \bar{t} \quad (12.2a)$$

$$x = \xi^1 = \xi^1(C^{\bar{a}}, \bar{t}) \quad (12.2b)$$

$$y = \xi^2 = \xi^2(C^{\bar{a}}, \bar{t}) \quad (12.2c)$$

$$z = \xi^3 = \xi^3(C^{\bar{a}}, \bar{t}) \quad (12.2d)$$

As for the GVC case of Chapter 11, time remains unchanged. However, we make use of a distinct symbol to help keep track of what coordinates are held fixed when performing partial derivatives. That is, time derivatives are distinct since distinct coordinates are held fixed when computing the derivative.

12.3 The transformation matrix and its inverse

We can write the coordinate transformation from ξ^a to $\xi^{\bar{a}}$ in the generic form

$$\xi^{\bar{a}} = \xi^{\bar{a}}(\xi^a), \quad (12.3)$$

and the inverse transformation as

$$\xi^a = \xi^a(\xi^{\bar{a}}). \quad (12.4)$$

For the spatial coordinates, the coordinate transformation matrix is comprised of the partial derivatives

$$\Lambda_a^{\bar{a}} = \begin{bmatrix} (\partial\xi^1/\partial\xi^{\bar{1}})_{\bar{2},\bar{3}} & (\partial\xi^1/\partial\xi^{\bar{2}})_{\bar{1},\bar{3}} & (\partial\xi^1/\partial\xi^{\bar{3}})_{\bar{1},\bar{2}} \\ (\partial\xi^2/\partial\xi^{\bar{1}})_{\bar{2},\bar{3}} & (\partial\xi^2/\partial\xi^{\bar{2}})_{\bar{1},\bar{3}} & (\partial\xi^2/\partial\xi^{\bar{3}})_{\bar{1},\bar{2}} \\ (\partial\xi^3/\partial\xi^{\bar{1}})_{\bar{2},\bar{3}} & (\partial\xi^3/\partial\xi^{\bar{2}})_{\bar{1},\bar{3}} & (\partial\xi^3/\partial\xi^{\bar{3}})_{\bar{1},\bar{2}} \end{bmatrix} = \begin{bmatrix} \partial x/\partial C^{\bar{1}} & \partial x/\partial C^{\bar{2}} & \partial x/\partial C^{\bar{3}} \\ \partial y/\partial C^{\bar{1}} & \partial y/\partial C^{\bar{2}} & \partial y/\partial C^{\bar{3}} \\ \partial z/\partial C^{\bar{1}} & \partial z/\partial C^{\bar{2}} & \partial z/\partial C^{\bar{3}} \end{bmatrix}. \quad (12.5)$$

The transformation matrix is full since no coordinates are orthogonal. That is, nonzero values generally exist for each matrix component since as we change spatial position $\xi^a = (x, y, z)$ as we move along any direction in tracer space. The inverse transformation matrix takes the form

$$\Lambda^{\bar{a}}_a = \begin{bmatrix} (\partial\xi^{\bar{1}}/\partial\xi^1)_{2,3} & (\partial\xi^{\bar{1}}/\partial\xi^2)_{1,3} & (\partial\xi^{\bar{1}}/\partial\xi^3)_{1,2} \\ (\partial\xi^{\bar{2}}/\partial\xi^1)_{2,3} & (\partial\xi^{\bar{2}}/\partial\xi^2)_{1,3} & (\partial\xi^{\bar{2}}/\partial\xi^3)_{1,2} \\ (\partial\xi^{\bar{3}}/\partial\xi^1)_{2,3} & (\partial\xi^{\bar{3}}/\partial\xi^2)_{1,3} & (\partial\xi^{\bar{3}}/\partial\xi^3)_{1,2} \end{bmatrix} = \begin{bmatrix} \partial C^{\bar{1}}/\partial x & \partial C^{\bar{1}}/\partial y & \partial C^{\bar{1}}/\partial z \\ \partial C^{\bar{2}}/\partial x & \partial C^{\bar{2}}/\partial y & \partial C^{\bar{2}}/\partial z \\ \partial C^{\bar{3}}/\partial x & \partial C^{\bar{3}}/\partial y & \partial C^{\bar{3}}/\partial z \end{bmatrix}. \quad (12.6)$$

To verify that the matrix (12.6) is indeed the inverse of the matrix (12.5), we follow the method used in Section 11.4 by writing the coordinate transformation as a composite function

$$\xi^{\bar{a}} = \xi^{\bar{a}}(\xi^a) = \xi^{\bar{a}}[\xi^a(\xi^{\bar{b}})]. \quad (12.7)$$

Taking partial derivatives and using the chain rule thus renders

$$\delta^{\bar{a}}_{\bar{b}} = \frac{\partial\xi^{\bar{a}}}{\partial\xi^{\bar{b}}} = \frac{\partial\xi^{\bar{a}}}{\partial\xi^a} \frac{\partial\xi^a}{\partial\xi^{\bar{b}}} \quad (12.8)$$

as well as

$$\delta^a_{\bar{b}} = \frac{\partial\xi^a}{\partial\xi^{\bar{b}}} = \frac{\partial\xi^a}{\partial\xi^{\bar{b}}} \frac{\partial\xi^{\bar{b}}}{\partial\xi^b}. \quad (12.9)$$

12.4 Determinant of the transformation

The determinant of the transformation matrix (12.5) is given by

$$\det(\Lambda^{\bar{a}}_a) = \left(\frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} \wedge \frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \right) \cdot \frac{\partial\mathbf{x}}{\partial C^{\bar{3}}} = \left(\frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \wedge \frac{\partial\mathbf{x}}{\partial C^{\bar{3}}} \right) \cdot \frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} = \left(\frac{\partial\mathbf{x}}{\partial C^{\bar{3}}} \wedge \frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} \right) \cdot \frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \quad (12.10)$$

and the determinant of the inverse transformation matrix (12.6) is

$$\det(\Lambda^a_{\bar{a}}) = \left(\frac{\partial\mathbf{C}}{\partial x} \wedge \frac{\partial\mathbf{C}}{\partial y} \right) \cdot \frac{\partial\mathbf{C}}{\partial z} = \left(\frac{\partial\mathbf{C}}{\partial y} \wedge \frac{\partial\mathbf{C}}{\partial z} \right) \cdot \frac{\partial\mathbf{C}}{\partial x} = \left(\frac{\partial\mathbf{C}}{\partial z} \wedge \frac{\partial\mathbf{C}}{\partial x} \right) \cdot \frac{\partial\mathbf{C}}{\partial y}. \quad (12.11)$$

For convenience we introduced a shorthand for the three tracer coordinates

$$\mathbf{C} = (C^{\bar{1}}, C^{\bar{2}}, C^{\bar{3}}) \quad (12.12)$$

and wrote

$$\vec{x} = \mathbf{x} = \hat{\mathbf{x}} x + \hat{\mathbf{y}} y + \hat{\mathbf{z}} z \quad (12.13)$$

for the Cartesian representation of the position vector.

12.5 Basis vectors

We transform the basis vectors from Cartesian into tracer coordinates through the transformation

$$\vec{e}_{\bar{a}} = \Lambda^a_{\bar{a}} \vec{e}_a. \quad (12.14)$$

Use of the transformation matrix (12.5) renders the tracer coordinate basis vectors

$$\vec{e}_{\bar{1}} = \frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} \quad (12.15a)$$

$$\vec{e}_{\bar{2}} = \frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \quad (12.15b)$$

$$\vec{e}_{\bar{3}} = \frac{\partial\mathbf{x}}{\partial C^{\bar{3}}}. \quad (12.15c)$$

12.6 Basis one-forms

The basis one-forms are obtained by transforming from Cartesian coordinates to tracer coordinates through use of the inverse transformation

$$\tilde{e}^{\bar{a}} = \Lambda_{\bar{a}}^{\bar{a}} \tilde{e}^a. \quad (12.16)$$

Use of the inverse transformation matrix (11.11) renders the one-form basis in terms of tracer coordinates

$$\tilde{e}^{\bar{1}} = \hat{x} \frac{\partial C^{\bar{1}}}{\partial x} + \hat{y} \frac{\partial C^{\bar{1}}}{\partial y} + \hat{z} \frac{\partial C^{\bar{1}}}{\partial z} = \nabla C^{\bar{1}} \quad (12.17a)$$

$$\tilde{e}^{\bar{2}} = \hat{x} \frac{\partial C^{\bar{2}}}{\partial x} + \hat{y} \frac{\partial C^{\bar{2}}}{\partial y} + \hat{z} \frac{\partial C^{\bar{2}}}{\partial z} = \nabla C^{\bar{2}} \quad (12.17b)$$

$$\tilde{e}^{\bar{3}} = \hat{x} \frac{\partial C^{\bar{3}}}{\partial x} + \hat{y} \frac{\partial C^{\bar{3}}}{\partial y} + \hat{z} \frac{\partial C^{\bar{3}}}{\partial z} = \nabla C^{\bar{3}}. \quad (12.17c)$$

Use of the chain rule relations (12.8) and (12.9) ensure that the basis one-forms satisfy the orthogonality relation (8.18) with the basis vectors

$$\tilde{e}^{\bar{a}} \cdot \vec{e}_{\bar{b}} = \delta_{\bar{a}}^{\bar{b}}. \quad (12.18)$$

12.7 Trajectory or position vector

The basis vectors (12.15a)-(12.15c) lead to the tracer coordinate representation of the position of a point in the fluid

$$\mathcal{P} = \xi^{\bar{a}} \vec{e}_{\bar{a}} \quad (12.19a)$$

$$= C^{\bar{1}} \frac{\partial \mathbf{x}}{\partial C^{\bar{1}}} + C^{\bar{2}} \frac{\partial \mathbf{x}}{\partial C^{\bar{2}}} + C^{\bar{3}} \frac{\partial \mathbf{x}}{\partial C^{\bar{3}}} \quad (12.19b)$$

$$= (\mathbf{C} \cdot \nabla_C) \mathbf{x}. \quad (12.19c)$$

Following the isopycnal coordinate representation of the trajectory given by equation (11.31), we can provide a Taylor series interpretation of the tracer trajectory (12.19c).

12.8 Velocity

We determine the tracer coordinate representation of the velocity vector, $\vec{v} = d\mathcal{P}/d\tau$, by transforming from the Cartesian components of the velocity to the tracer components according to

$$v^{\bar{a}} = \lambda_{\bar{a}}^{\bar{a}} v^a, \quad (12.20)$$

with the inverse transformation matrix, $\lambda_{\bar{a}}^{\bar{a}}$, given by equation (12.6). The result of the calculation yields

$$\vec{v} = (\mathbf{v} \cdot \nabla) C^{\bar{1}} \frac{\partial \mathbf{x}}{\partial C^{\bar{1}}} + (\mathbf{v} \cdot \nabla) C^{\bar{2}} \frac{\partial \mathbf{x}}{\partial C^{\bar{2}}} + (\mathbf{v} \cdot \nabla) C^{\bar{3}} \frac{\partial \mathbf{x}}{\partial C^{\bar{3}}} \quad (12.21)$$

where we have the Cartesian inner product

$$\mathbf{v} \cdot \nabla C^{\bar{a}} = u (\partial C^{\bar{a}} / \partial x) + v (\partial C^{\bar{a}} / \partial y) + w (\partial C^{\bar{a}} / \partial z). \quad (12.22)$$

12.9 Metric tensor

The tracer coordinate representation of the metric tensor is given by

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} \frac{\partial \mathbf{x}}{\partial C^1} \cdot \frac{\partial \mathbf{x}}{\partial C^1} & \frac{\partial \mathbf{x}}{\partial C^1} \cdot \frac{\partial \mathbf{x}}{\partial C^2} & \frac{\partial \mathbf{x}}{\partial C^1} \cdot \frac{\partial \mathbf{x}}{\partial C^3} \\ \frac{\partial \mathbf{x}}{\partial C^2} \cdot \frac{\partial \mathbf{x}}{\partial C^1} & \frac{\partial \mathbf{x}}{\partial C^2} \cdot \frac{\partial \mathbf{x}}{\partial C^2} & \frac{\partial \mathbf{x}}{\partial C^2} \cdot \frac{\partial \mathbf{x}}{\partial C^3} \\ \frac{\partial \mathbf{x}}{\partial C^3} \cdot \frac{\partial \mathbf{x}}{\partial C^1} & \frac{\partial \mathbf{x}}{\partial C^3} \cdot \frac{\partial \mathbf{x}}{\partial C^2} & \frac{\partial \mathbf{x}}{\partial C^3} \cdot \frac{\partial \mathbf{x}}{\partial C^3} \end{bmatrix}, \quad (12.23)$$

and the tracer coordinate representation of the inverse metric tensor is given by

$$g^{\bar{a}\bar{b}} = \vec{e}^{\bar{a}} \cdot \vec{e}^{\bar{b}} = \begin{bmatrix} \nabla C^1 \cdot \nabla C^1 & \nabla C^1 \cdot \nabla C^2 & \nabla C^1 \cdot \nabla C^3 \\ \nabla C^2 \cdot \nabla C^1 & \nabla C^2 \cdot \nabla C^2 & \nabla C^2 \cdot \nabla C^3 \\ \nabla C^3 \cdot \nabla C^1 & \nabla C^3 \cdot \nabla C^2 & \nabla C^3 \cdot \nabla C^3 \end{bmatrix}. \quad (12.24)$$

Proof that $g^{\bar{a}\bar{b}} g_{\bar{b}\bar{c}} = \delta_{\bar{a}}^{\bar{c}}$ requires use of the chain rule relations (12.8) and (12.9).

12.10 Vorticity

Following from the discussion of vorticity in generalized vertical coordinates (Section 11.16), we determine the tracer coordinate representation through transforming $\omega^{\bar{a}} = \Lambda^{\bar{a}}_a \omega^a$ from the Cartesian expression $\boldsymbol{\omega} = \omega^1 \hat{\mathbf{x}} + \omega^2 \hat{\mathbf{y}} + \omega^3 \hat{\mathbf{z}}$ and using the transformation matrix (12.6). The resulting expression is given by

$$\omega^{\bar{1}} = \boldsymbol{\omega} \cdot \nabla C^{\bar{1}} \quad (12.25a)$$

$$\omega^{\bar{2}} = \boldsymbol{\omega} \cdot \nabla C^{\bar{2}} \quad (12.25b)$$

$$\omega^{\bar{3}} = \boldsymbol{\omega} \cdot \nabla C^{\bar{3}}, \quad (12.25c)$$

revealing that the tracer coordinate representation of the vorticity is obtained by projecting the Cartesian representation onto the directions perpendicular to the tracer surfaces.

12.11 Comments

In his discussion of potential vorticity, [Salmon \(1998\)](#) considered the case of three materially invariant scalar fields. In turn, he noted that there are three corresponding potential vorticity fields. For geophysical fluids, we generally refer to a single potential vorticity corresponding to buoyancy or specific entropy as the scalar field (see Chapter 50).

Part II

Geophysical particle mechanics

Rotation and gravitation of the planet are the two defining features of atmospheric and oceanic fluid mechanics. Since geophysical fluids are in near solid-body motion (i.e., they exhibit motions that are small relative to that of the rotating planet), and since humans experience the fluids from the rotating terrestrial frame, we find it convenient, if not essential, to make use of a rotating (non-inertial) terrestrial reference frame.

In this part of the book, we employ Newtonian mechanics to describe the motion of a point particle moving around a rotating and gravitating sphere as viewed in the terrestrial reference frame. To be relevant for geophysical fluids, we generally assume the particle is in near solid-body motion with the planet, as per a geosynchronous satellite. This analysis exposes concepts that later appear in geophysical fluid mechanics, such as trajectories, linear momentum, angular momentum, body forces, non-inertial accelerations, symmetries, planetary Cartesian coordinates and planetary spherical coordinates. We also make use of spherical coordinates to develop expressions for position, velocity, and acceleration, with the mathematical treatment for particles sufficient for geophysical fluids.

Why bother with classical particle mechanics in a book on geophysical fluid mechanics? Geophysical fluid mechanics involves concepts from both fluid mechanics and classical point particle mechanics. Hence, in our development of geophysical fluid mechanics, we find it pedagogically useful to first examine rotational and gravitational particle mechanics; i.e., to study *geophysical particle mechanics*. Doing so also allows those less versed in classical mechanics to learn, or perhaps relearn, basic notions from particle mechanics also of use for understanding their more complex realizations in fluids. In this manner, geophysical particle mechanics provides a useful foundation for geophysical fluid mechanics.

13

Particle kinematics

We here consider the kinematics of a point particle of fixed mass and zero electric charge moving around a rotating and gravitating sphere. The motion of the gravitating sphere is prescribed with a fixed kinetic energy and fixed angular momentum around its axis of rotation. Hence, the only interesting mechanics is that for the moving particle. We examine that motion in a spherical geometry as observed from a non-inertial reference frame. Two forms of non-inertial accelerations (centrifugal and Coriolis) appear in this reference frame. We make use of both the planetary Cartesian and planetary spherical coordinates to represent the position, velocity, and acceleration vectors.

READER'S GUIDE TO THIS CHAPTER

This chapter makes use of basic features of both Cartesian and general tensor algebra as presented in Chapters 3, 7, 8, and Section 10.2. However, we offer the salient features of tensor technology in this chapter where needed, thus allowing for a reasonably self-contained presentation. Further chapters make use of the kinematics detailed here, such as Chapter 14 focused on particle dynamics on the sphere, and Chapter 22 considering fluid dynamics on the sphere.

This chapter is notable for working through, in full glory, some relatively tedious algebra as part of developing the spherical coordinate acceleration. Readers can skim these manipulations without losing much conceptual content.

13.1	The rotating earth	158
13.2	Reference frames and non-inertial accelerations	159
13.3	A few points from tensor algebra	160
13.3.1	Why we need general tensors	160
13.3.2	The coordinate representation of a vector	161
13.4	Galilean invariance for particle motion	161
13.5	Rotationally generated changes to a vector	162
13.5.1	Change in direction: brief derivation	163
13.5.2	Change in direction: detailed derivation for planar motion	163
13.5.3	Comments	164
13.6	The velocity vector	164
13.6.1	Coordinate velocity	164
13.6.2	Changes to the basis vectors	165
13.7	Inertial acceleration and its decomposition	165
13.8	Coordinate representation of the position vector	166
13.9	Coordinate representation of the velocity vector	167
13.9.1	Planetary Cartesian coordinate representation of velocity	167
13.9.2	Spherical coordinate representation	167
13.9.3	Axial angular momentum	168
13.10	Planetary Cartesian representation of acceleration	168
13.10.1	Planetary Cartesian representation	169
13.10.2	Summary of acceleration in terms of planetary Cartesian coordinates	170
13.10.3	Further study	170
13.11	Spherical representation of the acceleration vector	171
13.11.1	Decomposing the acceleration	172
13.11.2	Spherical coordinate acceleration	172
13.11.3	Metric acceleration	172
13.11.4	Centrifugal acceleration	172
13.11.5	Coriolis acceleration	173
13.11.6	Coriolis acceleration for large-scale motions	173
13.12	Exercises	174

13.1 The rotating earth

The earth's angular velocity is comprised of two main contributions: the spin of the earth about its axis and the orbit of the earth about the sun (see Figure 13.1). Other astronomical motions can be neglected for geophysical fluid mechanics. Therefore, in the course of a single period of 24 hours, or $24 \times 3600 = 86400$ seconds, the earth experiences an angular rotation of $(2\pi + 2\pi/365.24)$ radians. As such, the angular velocity of the earth is given by

$$\Omega = \frac{2\pi + 2\pi/365.24}{86400\text{s}} = \left[\frac{\pi}{43082} \right] \text{s}^{-1} = 7.2921 \times 10^{-5} \text{ s}^{-1}. \quad (13.1)$$

The earth's angular velocity, both its direction and its magnitude, is assumed constant in time for purposes of geophysical fluid mechanics (see Section 13.1)

$$\frac{d\vec{\Omega}}{dt} = 0. \quad (13.2)$$

The angular velocity (13.1) seems quite small. However, a terrestrial reference frame on the

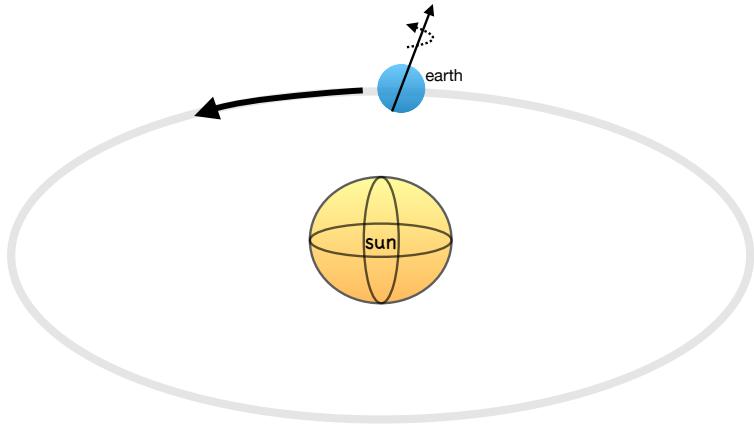


Figure 13.1: The angular velocity of the earth arises from the spin about the polar axis plus the orbit of the planet around the sun. This angular velocity determines the strength of the Coriolis acceleration and the centrifugal acceleration.

surface of the earth, undergoing *solid body motion*, moves with linear velocity

$$U_{\text{solid-body}} = \Omega R_e \approx 465 \text{ m s}^{-1} = 1672 \text{ km hr}^{-1}, \quad (13.3)$$

where we set the earth's radius to

$$R_e = 6.367 \times 10^6 \text{ m}. \quad (13.4)$$

Motion of the atmosphere and ocean fluids are generally quite close to solid-body, so that it makes sense to describe their motion from the non-inertial rotating reference frame that moves at this rather fast speed relative to the “fixed” stars.

13.2 Reference frames and non-inertial accelerations

To describe the motion of geophysical fluids, we make use of both inertial and non-inertial reference frames. An inertial reference frame is one in which an object that experiences no external forces either remains at rest or moves with a constant linear velocity. Two inertial reference frames can differ at most by a constant velocity. Consequently, accelerations measured in one inertial frame are the same as in another inertial frame. This property of inertial reference frames is known as *Galilean invariance* (see Section 13.4). Inertial reference frames are well suited for describing motion and the causes (i.e., forces) for the motion. In particular, when described from an inertial reference frame, Newton’s second law state that objects change their linear momentum only through the imposition of external forces.

Observers on the earth are in a non-inertial frame since the earth frame rotates, and rotational motion is accelerating motion. Furthermore, as noted earlier, motion of geophysical fluids deviates relatively little from solid-body motion. That is, motion of fluid elements deviates relatively little from motion of the earth itself. For these reasons, the preferred frame for studying geophysical motion is the rotating planetary frame. Figure 13.2 illustrates the case for the position vector of a particle relative to the origin of a rotating sphere.

A set of basis vectors is needed to use coordinates for representing vectors. The basis vectors hold two key pieces of information, with the first being details of the coordinates. The second concerns the reference frame, whereby non-inertial accelerations arise from time dependence to the

basis vectors. When represented in terms of non-inertial reference frame coordinates, the inertial acceleration (a vector) is decomposed into the sum of relative acceleration (relative to the non-inertial frame), centrifugal acceleration, and Coriolis acceleration.

When multiplied by mass, non-inertial accelerations can be interpreted as non-inertial forces. However, these forces are not imparted by an external body or force field. Rather, they arise from accelerated motion of the non-inertial reference frame. In this sense, non-inertial accelerations are often termed “fictitious”. Nonetheless, a terrestrial observer describes motion as undergoing non-inertial accelerations, with non-inertial accelerations playing a central role in rationalizing observed planetary fluid motions.

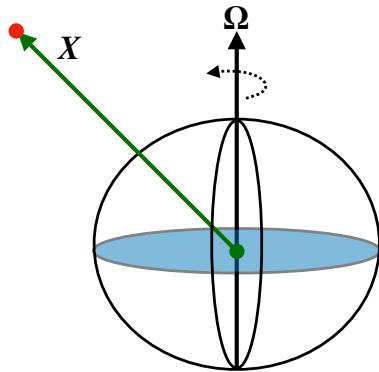


Figure 13.2: The position vector, $\mathbf{X}(t)$, for a particle moving around a rotating sphere with coordinate origins at the center of the sphere. The rotation axis is through the north pole, with angular velocity vector $\boldsymbol{\Omega}$. The sphere rotates in a positive right hand sense through the north polar axis (counter clockwise from above). The rotating frame has a “solid-body” velocity $\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X}$ (equation (13.28)). For a particle on the earth’s surface at the equator, the solid body speed is $\Omega R = 7.2921 \times 10^{-5} \text{ s}^{-1} \times 6.371 \times 10^6 \text{ m} \approx 465 \text{ m s}^{-1} = 1672 \text{ km hr}^{-1}$. This speed is much greater than that of a fluid element relative to the moving earth so that geophysical fluids are moving in near solid-body motion with the planet.

13.3 A few points from tensor algebra

In Part I of this book, we detailed the use of tensor analysis for geophysical fluid mechanics. We here summarize the salient points for the reader who skimmed that earlier discussion.

13.3.1 Why we need general tensors

Cartesian tensors are sufficient for many purposes of fluid mechanics, such as when using Cartesian coordinates for a tangent plane approximation to study geophysical fluid motion (e.g., Section 27.2). However, we make routine use of spherical coordinates when describing geophysical motion, and cylindrical-polar coordinates for studies of rotating tank experiments (see Exercise 27.5 and Section 42.3). Finally, we use generalized vertical coordinates in the description of stratified flows (Chapters 11, 21, and 32). The basis vectors for curvilinear coordinates and generalized vertical coordinates change direction when moving through space. In contrast, Cartesian basis vectors always point in the same direction. This distinction between the basis vectors is the key reason curvilinear coordinates and generalized vertical coordinates require a more general formalism than afforded by Cartesian tensors.

13.3.2 The coordinate representation of a vector

The coordinate representation of a vector follows from decomposing the vector into components aligned according to a set of basis vectors. In particular, the coordinate representation of the position vector is given by

$$\vec{X} = \sum_{a=1}^3 \xi^a \vec{e}_a = \xi^a \vec{e}_a, \quad (13.5)$$

where the Einstein summation convention is defined by the final equality. In this equation,

$$\vec{e}_a = (\vec{e}_1, \vec{e}_2, \vec{e}_3) \quad (13.6)$$

is a set of linearly independent basis vectors, and ξ^a are the corresponding coordinate representations of the position vector \vec{X} . The basis vectors may be normalized to unit magnitude, as in the case of Cartesian coordinates, or may be unnormalized as for spherical coordinates (see Section 10.2.2). Note that we commonly make use of the boldface notation for a vector rather than the arrowed symbol (Section 7.6)

$$\boldsymbol{X} = \vec{X} = \xi^a \vec{e}_a. \quad (13.7)$$

The basis vectors in equation (13.5) have a lower index while the coordinate representation of a vector has an upper index. Why? For arbitrary coordinates (e.g., spherical), we make a distinction between a coordinate representation with an index upstairs (contravariant) versus the downstairs (covariant) representation. Moving between the covariant and contravariant representations requires a metric tensor. For much of our work we can keep this mathematical framework at a modest distance, with exposure only in selected places. The key crutch we are relying on is that the planet is assumed to be embedded in a background Euclidean space \mathbb{R}^3 . That is, we are not considering the curved space-time of general relativity. This assumption simplifies much of our work.

When working with general coordinates, it is necessary to distinguish between a basis vector \vec{e}_a and its dual partner known as a *one-form*, \tilde{e}^a . Duality here is defined using the familiar (Euclidean) inner product

$$\vec{e}_a \cdot \tilde{e}^b = \delta_a^b, \quad (13.8)$$

with δ_a^b the Kronecker delta tensor

$$\delta_a^b = \begin{cases} 1 & \text{if } b = a \\ 0 & \text{if } b \neq a. \end{cases} \quad (13.9)$$

In linear algebra, a row vector is dual to its column vector, with that analog appropriate for the present context. Cartesian basis vectors equal to the basis one-forms, in which case there is no distinction between contravariant and covariant. However, the distinction is important for the general coordinates used in geophysical fluids such as encountered in Chapters 10 and 11.

13.4 Galilean invariance for particle motion

Consider a reference frame moving at a constant velocity relative to another reference frame. In classical non-relativistic physics, there is no experiment that can distinguish the two reference frames. Consequently, if one reference frame is inertial, then so is the other. Correspondingly, the mathematical expression of physical laws in the two inertial frames is the same. This property of inertial reference frames is known as *Galilean invariance*. Two inertial reference frames can at most

be moving relative to one another by a constant velocity. Otherwise, at least one of them must be accelerating, which in turn would mean that it is not an inertial frame.

For the point particle, Galilean invariance means that the acceleration of the particle in one reference frame equals to that in the other frame. This equivalence holds since the two frames are moving with a constant velocity relative to one another. Although rather trivial, we illustrate Galilean invariance through a bit of formalism. Doing so offers us practice for the less trivial case of a rotating reference frame.

Mathematically, the Galilean transformation is given by

$$\bar{t} = t \quad (13.10)$$

$$\bar{\mathbf{X}} = \mathbf{X} + \mathbf{U}t, \quad (13.11)$$

where the barred position vector is measured in the moving reference frame. Time remains unchanged, whereas the position of the particle in the new frame equals to that in the original reference frame plus a contribution from the constant velocity \mathbf{U} . Some refer to the barred reference frame as a “boosted” frame. The particle velocity in the moving (boosted) reference frame is given by

$$\bar{\mathbf{V}} = \frac{d\bar{\mathbf{X}}}{dt} = \frac{d\mathbf{X}}{dt} + \frac{d(\mathbf{U}t)}{dt} = \mathbf{V} + \mathbf{U}, \quad (13.12)$$

where we set

$$\frac{d\mathbf{U}}{dt} = 0 \quad (13.13)$$

since \mathbf{U} has a fixed magnitude and direction (as per our assumption that it is a constant vector). As expected, the velocity is shifted by the constant reference frame velocity \mathbf{U} . The acceleration in the two reference frames is related by

$$\bar{\mathbf{A}} = \frac{d^2\bar{\mathbf{X}}}{dt^2} = \frac{d\mathbf{V}}{dt} = \mathbf{A}. \quad (13.14)$$

The accelerations are indeed identical.

When introducing the equation of motion in Chapter 14, we will see that a Galilean invariant force is one that does not change when changing the velocity. A sufficient condition for such a force is that it be independent of the velocity itself. We return to this point in Section 15.5.1 when discussing kinetic energy. Furthermore, as seen in Section 16.5, Galilean invariance for fluid motion provides a richer statement than it does for a point particle.

13.5 Rotationally generated changes to a vector

How does a vector change under a solid-body rotation such as that shown in Figure 13.3? Answering this question is fundamental to the kinematics of rotational motion. To develop an answer, observe a pure rotation does not change the magnitude of a vector, so that

$$|\mathbf{X}(t)| = |\mathbf{X}(t + \delta t)|. \quad (13.15)$$

The condition (13.15) can be written as

$$\frac{d(\mathbf{X} \cdot \mathbf{X})}{dt} = 0, \quad (13.16)$$

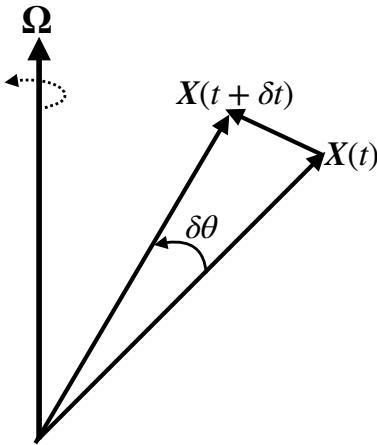


Figure 13.3: The change in a vector under a pure rotation leaves the vector magnitude unchanged, $|\mathbf{X}(t)| = |\mathbf{X}(t + \delta t)|$. Only the vector direction changes, here shown to be $\delta\theta = \Omega \delta t$. Infinitesimal changes generated by the angular velocity $\mathbf{\Omega}$ lead to the vector differences $\mathbf{X}(t + \delta t) - \mathbf{X}(t) = \delta t \mathbf{\Omega} \wedge \mathbf{X}(t)$.

which leads to the constraint

$$\mathbf{X} \cdot \frac{d\mathbf{X}}{dt} = 0. \quad (13.17)$$

That is, the velocity generated by a pure rotation is itself perpendicular to the position. We encountered this result in Section 4.1.4 when showing that unit vectors can only change through rotations.

13.5.1 Change in direction: brief derivation

Referring to Figure 13.3, we see that the infinitesimal difference $\mathbf{X}(t + \delta t) - \mathbf{X}(t)$ equals to the vector cross product of the angular velocity with the position vector

$$\mathbf{X}(t + \delta t) - \mathbf{X}(t) = \delta t \mathbf{\Omega} \wedge \mathbf{X}(t). \quad (13.18)$$

Dividing by δt leads to

$$\frac{d\mathbf{X}}{dt} = \mathbf{\Omega} \wedge \mathbf{X} \quad \text{pure rotation.} \quad (13.19)$$

Note that this evolution satisfies the constraint (13.17) since $\mathbf{X} \cdot (\mathbf{\Omega} \wedge \mathbf{X}) = 0$, meaning that the magnitude of the vector remains fixed.

13.5.2 Change in direction: detailed derivation for planar motion

To determine the change in direction generated by a pure rotation, we first consider the simplified case of planar rotation. Let $\mathbf{\Omega}$ be entirely in the vertical, and let \mathbf{X} be confined to the horizontal plane. In a time δt , the vector $\mathbf{X}(t)$ is rotated by an angle

$$\delta\theta = |\mathbf{\Omega}| \delta t \quad (13.20)$$

to $\mathbf{X}(t + \delta t)$. In the limit of small $\delta\theta$, the difference vector, $\mathbf{X}(t + \delta t) - \mathbf{X}(t)$, is perpendicular to $\mathbf{X}(t)$ and is of magnitude equal to the arc length

$$\delta s = |\mathbf{X}(t)| \delta\theta = |\mathbf{X}(t)| |\mathbf{\Omega}| \delta t. \quad (13.21)$$

We observe that the vector $\boldsymbol{\Omega} \wedge \mathbf{X}(t)$ points in the same direction as $\mathbf{X}(t + \delta t) - \mathbf{X}(t)$ and is of length $|\mathbf{X}(t)| |\boldsymbol{\Omega}|$. We conclude that

$$\mathbf{X}(t + \delta t) - \mathbf{X}(t) = \boldsymbol{\Omega} \wedge \mathbf{X}(t) \delta t. \quad (13.22)$$

Dividing through by δt and taking the limit $\delta t \rightarrow 0$ gives

$$\frac{d\mathbf{X}}{dt} = \boldsymbol{\Omega} \wedge \mathbf{X} \quad \text{pure rotation.} \quad (13.23)$$

The proof for the case in which $\boldsymbol{\Omega}$ has a component along \mathbf{X} is a straightforward generalization. The trajectory is still confined to a plane, but only the component of $\boldsymbol{\Omega}$ normal to the trajectory generates rotation.

13.5.3 Comments

We make use of the result (13.19) in many places when working with rotating physics. One place of particular note concerns the changes to unit vectors under rotation. We emphasize that it is only the Cartesian unit vectors that exhibit changes due to pure rotations, so that, for example,

$$\frac{d\hat{\mathbf{x}}}{dt} = \boldsymbol{\Omega} \wedge \hat{\mathbf{x}}. \quad (13.24)$$

Other unit vectors, such as those for spherical coordinates and polar coordinates, change due to the solid-body rotation just as for the Cartesian unit vectors. In addition, they can change when their orientation is modified relative to the Cartesian coordinate axes at a rate distinct from the solid-body. We further discuss this point in Section 13.6.2 and detail the spherical coordinate case in Sections 13.9.2 and 13.11.

13.6 The velocity vector

The velocity is the time derivative of the position vector

$$\mathbf{V} = \frac{d\mathbf{X}}{dt}. \quad (13.25)$$

In general, both the coordinate representation and the basis vectors are time dependent, so that the velocity has two contributions, one from the time dependence of the coordinates and another from the basis vectors

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = \frac{d(\xi^a \vec{e}_a)}{dt} = \frac{d\xi^a}{dt} \vec{e}_a + \xi^a \frac{d\vec{e}_a}{dt}. \quad (13.26)$$

13.6.1 Coordinate velocity

The first term on the right hand side of equation (13.26) is the velocity as measured within the rotating reference frame using the chosen coordinates

$$\mathbf{V}_{\text{coord}} \equiv \frac{d\xi^a}{dt} \vec{e}_a. \quad (13.27)$$

That is, this is the contribution as measured in the reference frame that moves with the basis vectors. In the context of geophysical motions, this is the velocity measured in the rotating terrestrial reference frame.

13.6.2 Changes to the basis vectors

The second term on the right hand side of equation (13.26) arises from changes to the basis vectors. If the basis vectors are normalized, they can change only through rotation. For a solid-body rotation of the reference frame, the solid-body velocity is given by (see Section 13.5)

$$\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X}. \quad (13.28)$$

Another manner to change a vector through rotation occurs when the basis vector maintains a constant magnitude but changes its direction at a different rate than the solid-body. Finally, if the basis vectors are not normalized, then they can change their magnitude during motion (stretching or compression). We encounter these three sorts of changes when considering coordinate representations later in this chapter.

13.7 Inertial acceleration and its decomposition

The inertial acceleration is given by the time derivative of the inertial velocity, which is the second derivative of the position vector

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2\mathbf{X}}{dt^2}. \quad (13.29)$$

This equation is independent of any coordinate representation. Correspondingly, the physical and geometrical content are manifest. When introducing a coordinate representation, the expression becomes subject to details of the chosen coordinates that can obscure the underlying geometric basis. Consequently, it is important to keep the geometric form in mind when offering an interpretation for coordinate dependent terms.

Introducing a coordinate representation $\mathbf{X} = \xi^a \vec{e}_a$ into the acceleration (13.29), and making use of the chain rule, leads to

$$\mathbf{A} = \frac{d}{dt} \frac{d\mathbf{X}}{dt} \quad (13.30a)$$

$$= \frac{d}{dt} \frac{d(\xi^a \vec{e}_a)}{dt} \quad (13.30b)$$

$$= \frac{d}{dt} \left[\frac{d\xi^a}{dt} \vec{e}_a + \xi^a \frac{d\vec{e}_a}{dt} \right] \quad (13.30c)$$

$$= \frac{d^2\xi^a}{dt^2} \vec{e}_a + 2 \frac{d\xi^a}{dt} \frac{d\vec{e}_a}{dt} + \xi^a \frac{d^2\vec{e}_a}{dt^2}. \quad (13.30d)$$

The first term on the right hand side is the acceleration of the coordinate representation as measured in the rotating reference frame

$$\mathbf{A}_{\text{coord}} \equiv \frac{d^2\xi^a}{dt^2} \vec{e}_a. \quad (13.31)$$

It is the acceleration measured by an observer in the rotating frame using coordinates ξ^a . The remaining two terms arise from changes to the basis vectors, and they give rise to the Coriolis and centrifugal accelerations associated with the rotating reference frame. In non-Cartesian coordinates, they also give rise to a “metric acceleration” arising from the change in directions of the unit vectors associated with motion of the particle relative to the rotating reference frame.

Some presentations of the kinematic result (13.30d) suggest that the factor of two on the middle term (the Coriolis term) is mysterious. In fact, there is nothing mysterious. Rather, the factor of two results from the need to take two time derivatives of the basis vectors as part of a representation of acceleration. It appears throughout rotational physics as part of the Coriolis acceleration.

13.8 Coordinate representation of the position vector

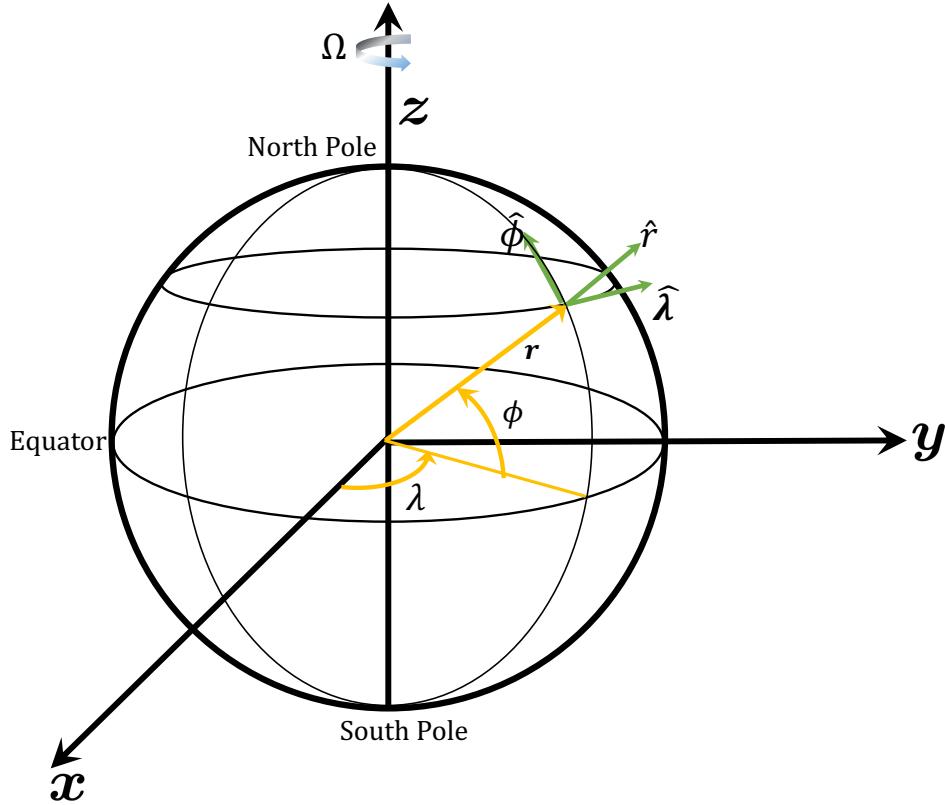


Figure 13.4: Geometry and notation for motion around a rotating sphere. For geophysical applications, the sphere rotates counter-clockwise when looking down from the north polar axis with angular speed Ω . The planetary Cartesian triad of orthonormal basis vectors, $(\hat{x}, \hat{y}, \hat{z})$ points along the orthogonal axes with origin at the sphere's center. The spherical triad of orthonormal basis vectors, $(\hat{\lambda}, \hat{\phi}, \hat{r})$, makes use of the longitudinal unit vector $\hat{\lambda}$, which points in the longitudinal direction (positive eastward), the latitudinal unit vector $\hat{\phi}$, which points in the latitudinal direction (positive northward) and the radial unit vector \hat{r} , which points in the radial direction (positive away from the center).

We make use of some results from Section 10.2 relating Cartesian and spherical coordinates and as furthermore defined by Figure 13.4. Starting with the position vector, we introduce the *planetary* Cartesian basis vectors, $(\hat{x}, \hat{y}, \hat{z})$, and corresponding spherical basis vectors, $(\hat{\lambda}, \hat{\phi}, \hat{r})$. We thus have the suite of equivalent expressions for the position of a particle moving around the sphere

$$\mathbf{X} = x \hat{x} + y \hat{y} + z \hat{z} \quad (13.32a)$$

$$= (r \cos \phi \cos \lambda) \hat{x} + (r \cos \phi \sin \lambda) \hat{y} + (r \sin \phi) \hat{z} \quad (13.32b)$$

$$= r \hat{r} \quad (13.32c)$$

$$= |\mathbf{X}| \hat{r}. \quad (13.32d)$$

Note how the expression for the position vector is quite simple when written in spherical coordinates, as it is merely the distance from the origin with a direction that points radially from the origin to the particle.

13.9 Coordinate representation of the velocity vector

As seen in Section 13.6, the inertial velocity vector has a coordinate representation written as

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = \frac{d(\xi^a \vec{e}_a)}{dt} = \frac{d\xi^a}{dt} \vec{e}_a + \xi^a \frac{d\vec{e}_a}{dt}. \quad (13.33)$$

Contributions arise from both the time changes in the coordinates, ξ^a , and time changes to the basis vectors, \vec{e}_a . We now consider the Cartesian and spherical forms for these changes.

13.9.1 Planetary Cartesian coordinate representation of velocity

The basis vectors for the Cartesian coordinates, $(\vec{e}_1, \vec{e}_2, \vec{e}_3) = (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$, are normalized, so they do not change their magnitude. Furthermore, they move only through solid-body motion of the rotating reference frame. We refer to these coordinates as *planetary Cartesian coordinates* since they are oriented according to the rotating planet. In Section 27.2 we introduce the tangent plane Cartesian coordinates, which are defined according to a tangent plane on the surface of the sphere.

The angular velocity is oriented around the polar axis

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}, \quad (13.34)$$

so that the solid-body velocity only has components in the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ directions

$$\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X} = \Omega (-\hat{\mathbf{x}} y + \hat{\mathbf{y}} x). \quad (13.35)$$

The inertial velocity thus has the following representation in terms of planetary Cartesian coordinates in the rotating reference frame

$$\mathbf{V} = \left[-y \Omega + \frac{dx}{dt} \right] \hat{\mathbf{x}} + \left[x \Omega + \frac{dy}{dt} \right] \hat{\mathbf{y}} + \frac{dz}{dt} \hat{\mathbf{z}} \quad (13.36a)$$

$$= \mathbf{V}_{\text{Cartesian}} + \boldsymbol{\Omega} \wedge \mathbf{X}, \quad (13.36b)$$

where we defined the Cartesian velocity vector

$$\mathbf{V}_{\text{Cartesian}} \equiv \frac{dx}{dt} \hat{\mathbf{x}} + \frac{dy}{dt} \hat{\mathbf{y}} + \frac{dz}{dt} \hat{\mathbf{z}}, \quad (13.37)$$

which is the velocity as measured in the rotating reference frame when using planetary Cartesian coordinates.

13.9.2 Spherical coordinate representation

The position vector in the spherical coordinate representation is given by

$$\mathbf{X} = r \hat{\mathbf{r}}. \quad (13.38)$$

The basis vector $\hat{\mathbf{r}}$ is normalized, so that its evolution arises just from rotational motion. It can rotate either through solid-body motion of the rotating reference frame, or through changes in the spherical angles, λ, ϕ relative to the rotating reference frame. We see these two forms of change by taking the time derivative of $\hat{\mathbf{r}}$ as given by equation (10.29c)

$$\frac{d\hat{\mathbf{r}}}{dt} = \frac{d}{dt} [\hat{\mathbf{x}} \cos \lambda \cos \phi + \hat{\mathbf{y}} \sin \lambda \cos \phi + \hat{\mathbf{z}} \sin \phi] \quad (13.39a)$$

$$= \cos \phi \left[\frac{d\lambda}{dt} + \Omega \right] \hat{\lambda} + \frac{d\phi}{dt} \hat{\phi}. \quad (13.39b)$$

Consequently, the inertial velocity has the following spherical coordinate representation

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} \quad (13.40a)$$

$$= \frac{d(r \hat{\mathbf{r}})}{dt} \quad (13.40b)$$

$$= \frac{dr}{dt} \hat{\mathbf{r}} + r \frac{d\hat{\mathbf{r}}}{dt} \quad (13.40c)$$

$$= r_{\perp} \left[\frac{d\lambda}{dt} + \Omega \right] \hat{\boldsymbol{\lambda}} + r \frac{d\phi}{dt} \hat{\boldsymbol{\phi}} + \frac{dr}{dt} \hat{\mathbf{r}} \quad (13.40d)$$

$$= (u + r_{\perp} \Omega) \hat{\boldsymbol{\lambda}} + v \hat{\boldsymbol{\phi}} + w \hat{\mathbf{r}} \quad (13.40e)$$

$$= \mathbf{V}_{\text{spherical}} + \mathbf{U}_{\text{solid}}. \quad (13.40f)$$

In this equation we introduced the spherical velocity vector

$$\mathbf{V}_{\text{spherical}} = r_{\perp} \frac{d\lambda}{dt} \hat{\boldsymbol{\lambda}} + r \frac{d\phi}{dt} \hat{\boldsymbol{\phi}} + \frac{dr}{dt} \hat{\mathbf{r}} \quad (13.41a)$$

$$= u \hat{\boldsymbol{\lambda}} + v \hat{\boldsymbol{\phi}} + w \hat{\mathbf{r}}, \quad (13.41b)$$

where

$$u = r_{\perp} \frac{d\lambda}{dt} \quad v = r \frac{d\phi}{dt} \quad w = \frac{dr}{dt}, \quad (13.42)$$

are components to the spherical velocity vector, and with

$$r_{\perp} = r \cos \phi \quad (13.43)$$

the distance to the polar axis. The spherical velocity, $\mathbf{V}_{\text{spherical}}$, is the velocity measured in the rotating reference frame when using planetary spherical coordinates. We also noted that the solid-body velocity has the spherical coordinate representation

$$\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X} = r_{\perp} \Omega \hat{\boldsymbol{\lambda}}. \quad (13.44)$$

That is, the solid-body velocity is purely zonal.

13.9.3 Axial angular momentum

As seen in Section 15.7, the zonal component of the inertial velocity equals to the axial angular momentum per unit mass

$$L^z = m \hat{\boldsymbol{\lambda}} \cdot \mathbf{V} = m (u + r_{\perp} \Omega). \quad (13.45)$$

The distance to the rotational axis is given by r_{\perp} , and this is the moment-arm for the axial angular momentum. For cases with rotational symmetry around polar axis, as for motion of a particle around a smooth sphere, the axial angular momentum is a constant of the motion. As discussed in Section 15.7, this conservation law offers a very important constraint on the particle trajectory. It also plays a role in the motion of geophysical fluids (Section 22.4).

13.10 Planetary Cartesian representation of acceleration

The inertial acceleration vector is given by the second time derivative of the position vector

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2\mathbf{X}}{dt^2}. \quad (13.46)$$

We here consider its representation using planetary Cartesian coordinates (x, y, z) and the Cartesian basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3) = (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$.

13.10.1 Planetary Cartesian representation

For our study of geophysical fluid motion, we assume the planetary angular velocity, $\boldsymbol{\Omega}$, is a constant in time

$$\frac{d\boldsymbol{\Omega}}{dt} = 0. \quad (13.47)$$

Making use of the results from Section 13.5 leads to

$$\frac{d\vec{e}_a}{dt} = \boldsymbol{\Omega} \wedge \vec{e}_a \quad (13.48)$$

and

$$\frac{d^2\vec{e}_a}{dt^2} = \frac{d}{dt}(\boldsymbol{\Omega} \wedge \vec{e}_a) = \boldsymbol{\Omega} \wedge \frac{d\vec{e}_a}{dt} = \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \vec{e}_a), \quad (13.49)$$

which yields the inertial acceleration

$$\mathbf{A} = \frac{d}{dt} \frac{d\mathbf{X}}{dt} \quad (13.50a)$$

$$= \frac{d^2\xi^a}{dt^2} \vec{e}_a + 2 \frac{d\xi^a}{dt} (\boldsymbol{\Omega} \wedge \vec{e}_a) + \xi^a \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \vec{e}_a) \quad (13.50b)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} + \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \mathbf{X}) \quad (13.50c)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \Omega^2 (x \hat{\mathbf{x}} + y \hat{\mathbf{y}}) \quad (13.50d)$$

$$= \mathbf{A}_{\text{Cartesian}} - \mathbf{A}_{\text{Coriolis}} - \mathbf{A}_{\text{centrifugal}}. \quad (13.50e)$$

The inertial acceleration is thus decomposed into three terms. The first contribution is the Cartesian expression

$$\mathbf{A}_{\text{Cartesian}} = \frac{d^2x}{dt^2} \hat{\mathbf{x}} + \frac{d^2y}{dt^2} \hat{\mathbf{y}} + \frac{d^2z}{dt^2} \hat{\mathbf{z}}, \quad (13.51)$$

which is the coordinate acceleration measured in the rotating frame using planetary Cartesian coordinates. The second contribution on the right hand side of equation (13.50e) is minus the Coriolis acceleration

$$\mathbf{A}_{\text{Coriolis}} = -2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}}. \quad (13.52)$$

The Coriolis acceleration plays a fundamental role in geophysical fluid mechanics and will be central to our development in this book. The third contribution is the centripetal acceleration, which is also minus the centrifugal acceleration

$$\mathbf{A}_{\text{centripetal}} = -\mathbf{A}_{\text{centrifugal}} = \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \mathbf{X}) = -\Omega^2 \mathbf{X}. \quad (13.53)$$

The centrifugal acceleration points outward from (perpendicular to) the polar axis of rotation whereas the centripetal acceleration points inward; they are action/reaction pairs. They can be written as the gradient of a potential

$$\mathbf{A}_{\text{centrifugal}} = -\nabla \Phi_{\text{centrifugal}} \iff \Phi_{\text{centrifugal}} \equiv -\frac{\Omega^2 r_\perp^2}{2} = -\frac{\Omega^2 (x^2 + y^2)}{2}, \quad (13.54)$$

which is a form to be used in the following. The centripetal acceleration (pointing towards the rotational axis) is that part of the inertial acceleration that keeps the rotating particle from flying away from the rotational axis. Its opposing partner, the centrifugal acceleration, is a non-inertial acceleration that accounts for the slight equatorial bulge of the planet. It is what pulls one outward from the center of a rotating merry-go-round.

13.10.2 Summary of acceleration in terms of planetary Cartesian coordinates

For the purpose of formulating the equation of motion in the rotating terrestrial frame, we write the rotating frame acceleration as

$$\mathbf{A}_{\text{Cartesian}} = \mathbf{A} + \mathbf{A}_{\text{Coriolis}} + \mathbf{A}_{\text{centrifugal}} \quad (13.55a)$$

$$= \mathbf{A} - 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \nabla\Phi_{\text{centrifugal}} \quad (13.55b)$$

and summarize the following accelerations (force per unit mass).

- **INERTIAL:** Newton's Law of motion is formulated within an inertial reference frame. It is the inertial acceleration, \mathbf{A} , that is directly affected by forces such as gravitation.
- **CORIOLIS:** The Coriolis acceleration

$$\mathbf{A}_{\text{Coriolis}} = -2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} = -2\Omega \hat{\mathbf{z}} \wedge \mathbf{V}_{\text{Cartesian}} = -2\Omega \left[-\frac{dy}{dt} \hat{\mathbf{x}} + \frac{dx}{dt} \hat{\mathbf{y}} \right], \quad (13.56)$$

arises from our choice to describe motion within the rotating reference frame. The Coriolis acceleration gives rise to a rich suite of fundamentally new phenomena relative to non-rotating motion. It has components only in the horizontal planetary Cartesian plane spanned by the unit vectors $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$. This geometric results is to be expected since the Coriolis acceleration arises from rotation about the polar $\hat{\mathbf{z}}$ axis.

- **CENTRIFUGAL:** The centrifugal acceleration

$$\mathbf{A}_{\text{centrifugal}} = -\nabla\Phi_{\text{centrifugal}} = \Omega^2 \mathbf{r}_\perp = \Omega^2 (x \hat{\mathbf{x}} + y \hat{\mathbf{y}}) \quad (13.57)$$

is another term arising from the rotating reference frame. As for the Coriolis acceleration, the centrifugal acceleration has components only in the horizontal planetary Cartesian plane spanned by the unit vectors $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$. The centrifugal acceleration is directed outward from (perpendicular to) the polar axis of rotation. We see this orientation in Figure 14.1 to be discussed later. Furthermore, the centrifugal acceleration is nonzero even when the particle is fixed relative to the rotating planet, whereas the Coriolis acceleration is zero when the particle has zero motion relative to the planet.

The centrifugal acceleration can be written as the gradient of a scalar potential, $\mathbf{A}_{\text{centrifugal}} = -\nabla\Phi_{\text{centrifugal}}$ where $\Phi_{\text{centrifugal}} = -\Omega^2(x^2 + y^2)/2$ (equation (13.54)). Hence, the centrifugal acceleration can be combined with the gravitational acceleration in the equation of motion (see Section 14.1). The resulting "effective gravity" leads to a conservative force field that is modified relative to the central gravitational field of the non-rotating spherical planet. We detail these points in Section 14.1.2.

13.10.3 Further study

Section 3.5 of [Apel \(1987\)](#) offers an insightful presentation of the Coriolis acceleration. Visualizations from rotating tank experiments are useful to garner an intuitive understanding of the Coriolis acceleration. The first few minutes of this [this video from Prof. Dave Fultz of the University of Chicago](#) is particularly insightful. We will further build up our understanding of the Coriolis acceleration as the book develops further.

13.11 Spherical representation of the acceleration vector

The spherical representation of the inertial velocity is given by equation (13.40f)

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = (u + r_{\perp} \Omega) \hat{\lambda} + v \hat{\phi} + w \hat{r} = \mathbf{V}_{\text{sphere}} + r_{\perp} \Omega \hat{\lambda}, \quad (13.58)$$

where we introduced the spherical velocity from equation (13.42)

$$\mathbf{V}_{\text{sphere}} \equiv u \hat{\lambda} + v \hat{\phi} + w \hat{r}. \quad (13.59)$$

We will also make use of the notation for the zonal component of the inertial velocity,

$$u_I = u + r_{\perp} \Omega. \quad (13.60)$$

Just as for computing the inertial velocity vector, the inertial acceleration must take into account changes in both the spherical coordinates and spherical basis vectors

$$\mathbf{A} = \frac{d}{dt} (u_I \hat{\lambda} + v \hat{\phi} + w \hat{r}) \quad (13.61a)$$

$$= \frac{du_I}{dt} \hat{\lambda} + \frac{dv}{dt} \hat{\phi} + \frac{dw}{dt} \hat{r} + u_I \frac{d\hat{\lambda}}{dt} + v \frac{d\hat{\phi}}{dt} + w \frac{d\hat{r}}{dt}. \quad (13.61b)$$

The spherical unit vectors change due to both the solid-body rotation of the rotating reference frame, plus motion of the particle relative to the rotating frame. Making use of the expressions given in Section (10.2.2), a bit of algebra yields the time derivatives

$$\frac{d\hat{\lambda}}{dt} = \left[\Omega + \frac{d\lambda}{dt} \right] (\hat{\phi} \sin \phi - \hat{r} \cos \phi) \quad (13.62)$$

$$\frac{d\hat{\phi}}{dt} = -\hat{\lambda} \left[\Omega + \frac{d\lambda}{dt} \right] \sin \phi - \hat{r} \dot{\phi} \quad (13.63)$$

$$\frac{d\hat{r}}{dt} = \hat{\lambda} \left[\frac{d\lambda}{dt} + \Omega \right] \cos \phi + \frac{d\phi}{dt} \hat{\phi}. \quad (13.64)$$

We are thus led to the inertial acceleration components

$$\mathbf{A} \cdot \hat{\lambda} = \left[\frac{du_I}{dt} + \left(\frac{d\lambda}{dt} + \Omega \right) (w \cos \phi - v \sin \phi) \right] \quad (13.65a)$$

$$\mathbf{A} \cdot \hat{\phi} = \left[\frac{dv}{dt} + \left(\frac{d\lambda}{dt} + \Omega \right) u_I \sin \phi + w \frac{d\phi}{dt} \right] \quad (13.65b)$$

$$\mathbf{A} \cdot \hat{r} = \left[\frac{dw}{dt} - \left(\frac{d\lambda}{dt} + \Omega \right) u_I \cos \phi - v \frac{d\phi}{dt} \right]. \quad (13.65c)$$

Use of the identities

$$u = r_{\perp} \frac{d\lambda}{dt} \quad u_I = u + r_{\perp} \Omega \quad \frac{du_I}{dt} = \frac{du}{dt} + \Omega (w \cos \phi - v \sin \phi) \quad (13.66)$$

and some reorganization leads to

$$\begin{aligned} \mathbf{A} = & \hat{\lambda} \left[\frac{du}{dt} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) \right] \\ & + \hat{\phi} \left[\frac{dv}{dt} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi + r_{\perp} \Omega^2 \sin \phi \right] \\ & + \hat{r} \left[\frac{dw}{dt} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi - r_{\perp} \Omega^2 \cos \phi \right]. \end{aligned} \quad (13.67)$$

13.11.1 Decomposing the acceleration

We decompose the inertial acceleration (13.67) into the following terms

$$\mathbf{A} = \mathbf{A}_{\text{sphere}} + \mathbf{A}_{\text{metric}} - \mathbf{A}_{\text{Coriolis}} - \mathbf{A}_{\text{centrifugal}}. \quad (13.68)$$

We chose signs so that in the rotating frame the acceleration is written

$$\underbrace{\mathbf{A}_{\text{sphere}} + \mathbf{A}_{\text{metric}}}_{\text{net spherical acceleration}} = \mathbf{A} + \mathbf{A}_{\text{Coriolis}} + \mathbf{A}_{\text{centrifugal}}. \quad (13.69)$$

We identify the net spherical acceleration as the sum of the coordinate acceleration and metric acceleration. In the absence of rotation, this sum provides an expression for the inertial acceleration as represented by spherical coordinates. The Coriolis and centrifugal terms arise from rotation.

13.11.2 Spherical coordinate acceleration

The spherical coordinate acceleration is given by the time change in the spherical velocity components

$$\mathbf{A}_{\text{sphere}} = \frac{du}{dt} \hat{\lambda} + \frac{dv}{dt} \hat{\phi} + \frac{dw}{dt} \hat{r}. \quad (13.70)$$

This term has no contribution from changes to the spherical unit vectors.

13.11.3 Metric acceleration

The metric acceleration arises from changes to the spherical unit vectors due to our use of spherical coordinates. It is given by

$$\mathbf{A}_{\text{metric}} = \hat{\lambda} \left[\frac{u(w - v \tan \phi)}{r} \right] + \hat{\phi} \left[\frac{v w + u^2 \tan \phi}{r} \right] - \hat{r} \left[\frac{u^2 + v^2}{r} \right] \quad (13.71a)$$

$$= \hat{\lambda} \left[\frac{u(w \cos \phi - v \sin \phi)}{r \cos \phi} \right] + \hat{\phi} \left[\frac{v w \cos \phi + u^2 \sin \phi}{r \cos \phi} \right] - \hat{r} \left[\frac{u^2 + v^2}{r} \right] \quad (13.71b)$$

$$= \frac{1}{r} [u \tan \phi (\hat{r} \wedge \mathbf{V}_{\text{sphere}}) + w \mathbf{U}_{\text{sphere}} - \hat{r} \mathbf{U}_{\text{sphere}} \cdot \mathbf{U}_{\text{sphere}}], \quad (13.71c)$$

where we wrote the horizontal (angular) and vertical (radial) components of the spherical velocity according to

$$\mathbf{V}_{\text{sphere}} = \mathbf{U}_{\text{sphere}} + \hat{r} w = \hat{\lambda} u + \hat{\phi} v + \hat{r} w. \quad (13.72)$$

For purposes of developing a kinetic energy budget (see Section 15.5), note that

$$\mathbf{V}_{\text{sphere}} \cdot \mathbf{A}_{\text{metric}} = 0, \quad (13.73)$$

so that

$$\mathbf{V} \cdot \mathbf{A}_{\text{metric}} = \Omega u r_{\perp} (w - v \tan \phi) = \Omega u r (w \cos \phi - v \sin \phi). \quad (13.74)$$

13.11.4 Centrifugal acceleration

The spherical coordinate representation of the centrifugal acceleration is given by

$$\mathbf{A}_{\text{centrifugal}} = -\nabla \Phi_{\text{centrifugal}} \quad (13.75a)$$

$$= \Omega^2 (x \hat{x} + y \hat{y}) \quad (13.75b)$$

$$= r_{\perp} \Omega^2 (\hat{\phi} \sin \phi - \hat{r} \cos \phi). \quad (13.75c)$$

The centrifugal acceleration points outward from the axis of rotation (see Figure 14.1 to be discussed later), so that it has no component in the longitudinal direction. For purposes of developing a kinetic energy budget (Section 15.5), note that

$$\mathbf{V} \cdot \mathbf{A}_{\text{centrifugal}} = \mathbf{V}_{\text{sphere}} \cdot \mathbf{A}_{\text{centrifugal}} = -\Omega^2 r \cos \phi (w \cos \phi - v \sin \phi). \quad (13.76)$$

13.11.5 Coriolis acceleration

The spherical coordinate representation of the Coriolis acceleration makes use of

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}} = \Omega (\hat{\phi} \cos \phi + \hat{\mathbf{r}} \sin \phi) \quad (13.77)$$

to reach the form

$$\mathbf{A}_{\text{Coriolis}} = -2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{sphere}} \quad (13.78a)$$

$$= -2 \Omega (\hat{\phi} \cos \phi + \hat{\mathbf{r}} \sin \phi) \wedge \mathbf{V}_{\text{sphere}} \quad (13.78b)$$

$$= -2 \Omega (\hat{\phi} \cos \phi + \hat{\mathbf{r}} \sin \phi) \wedge (u \hat{\lambda} + v \hat{\phi} + w \hat{\mathbf{r}}) \quad (13.78c)$$

$$= -2 \Omega [\hat{\lambda} (w \cos \phi - v \sin \phi) + \hat{\phi} u \sin \phi - \hat{\mathbf{r}} u \cos \phi]. \quad (13.78d)$$

Note that

$$\mathbf{V}_{\text{sphere}} \cdot \mathbf{A}_{\text{Coriolis}} = 0, \quad (13.79)$$

so that

$$\mathbf{V} \cdot \mathbf{A}_{\text{Coriolis}} = -2 \Omega^2 r \cos \phi (w \cos \phi - v \sin \phi). \quad (13.80)$$

We find it convenient to introduce a shorthand notation

$$\mathbf{f} = 2 \Omega \sin \phi \hat{\mathbf{r}} \quad (13.81)$$

$$\mathbf{f}^* = 2 \Omega \cos \phi \hat{\phi}, \quad (13.82)$$

so that the Coriolis acceleration takes the form

$$\mathbf{A}_{\text{Coriolis}} = -(\mathbf{f} + \mathbf{f}^*) \wedge \mathbf{V}_{\text{sphere}}. \quad (13.83)$$

There are two contributions to the Coriolis acceleration: one from the radial and one from the meridional component of the earth's rotation vector.

13.11.6 Coriolis acceleration for large-scale motions

Let us again write the Coriolis acceleration in equation (13.78d), only now underlining two terms

$$\mathbf{A}_{\text{Coriolis}} = -2 \Omega [\hat{\lambda} (\underline{w \cos \phi} - \underline{v \sin \phi}) + \hat{\phi} \underline{u \sin \phi} - \hat{\mathbf{r}} \underline{u \cos \phi}]. \quad (13.84)$$

For many applications in geophysical fluid dynamics, the term $\hat{\mathbf{r}} (2 \Omega u \cos \phi)$ is much smaller than the competing gravitational acceleration that also contributes to the radial acceleration, thus prompting $\hat{\mathbf{r}} (2 \Omega u \cos \phi)$ to be dropped from the $\hat{\mathbf{r}}$ equation of motion.¹ Furthermore, the vertical velocity term is generally much smaller than the horizontal velocity term appearing in the $\hat{\lambda}$ component. Dropping these two terms results in the form for the Coriolis acceleration used for

¹The term $\hat{\mathbf{r}} (2 \Omega u \cos \phi)$ is called the Eötvös correction in the study of marine gravity.

large-scale dynamics, such as when considering the hydrostatic primitive equations for geophysical fluids (Section 27.1)

$$\mathbf{A}_{\text{Coriolis}}^{\text{large-scale}} \equiv -2\Omega \sin \phi (-\hat{\lambda} v + \hat{\phi} u) \quad (13.85\text{a})$$

$$\equiv -f \hat{\mathbf{r}} \wedge \mathbf{V}_{\text{sphere}}. \quad (13.85\text{b})$$

For the last equality we introduced the Coriolis parameter

$$f \equiv 2\Omega \sin \phi. \quad (13.86)$$

As illustrated in Figure 13.5, we see that it is the local vertical component of the earth's angular rotation that plays the most important role in large-scale fluid mechanics

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}} = \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi) \approx \Omega \sin \phi \hat{\mathbf{r}} = \mathbf{f}/2. \quad (13.87)$$

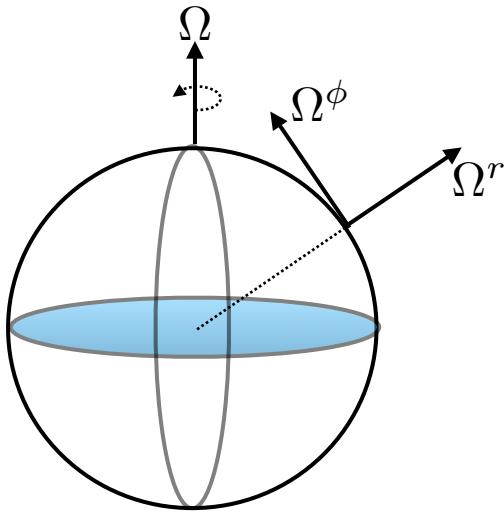


Figure 13.5: This figure illustrates the two components of the earth's rotational velocity, $\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}} = \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi)$. The local vertical component, $\boldsymbol{\Omega} \approx \Omega \sin \phi \hat{\mathbf{r}}$, is the most important component for large scale geophysical fluid dynamics.

13.12 Exercises

EXERCISE 13.1: WORKING THROUGH THE SPHERICAL ACCELERATION

Convince yourself that the spherical form of the acceleration given by equation (13.67) is indeed correct.

EXERCISE 13.2: VELOCITY AND ACCELERATION IN CYLINDRICAL-POLAR COORDINATES

In Section 10.3 we worked through the transformation from Cartesian coordinates to cylindrical-polar coordinates for describing motion from a rotating reference frame. Here we develop an expression for the position, velocity, and acceleration vectors in a frame rotating about the vertical axis using cylindrical-polar coordinates. The cylindrical-polar coordinates are useful when describing physical systems such as rotating fluid columns (e.g., fluids in a rotating circular tank as in Section 43.5) or when studying the cyclostrophically balanced flow.

- (A) Determine the representation of the inertial velocity vector, $\mathbf{V} = d\mathbf{X}/dt$, in terms of cylindrical-polar coordinates.
- (B) Determine the representation of the inertial acceleration vector, $\mathbf{A} = d\mathbf{V}/dt$, in terms of cylindrical-polar coordinates.
- (C) Writing the inertial acceleration in the form

$$\mathbf{A} = \mathbf{A}_{\text{cylindrical-polar}} - \mathbf{A}_{\text{centrifugal}} - \mathbf{A}_{\text{Coriolis}}, \quad (13.88)$$

give the mathematical expressions for the relative acceleration written in cylindrical-polar coordinates, $\mathbf{A}_{\text{cylindrical-polar}}$, the centrifugal acceleration, $\mathbf{A}_{\text{centrifugal}}$, and the Coriolis acceleration, $\mathbf{A}_{\text{Coriolis}}$.

EXERCISE 13.3: VELOCITY PROJECTED ONTO ACCELERATION

The kinetic energy per mass of a particle is given by

$$\mathcal{K} = \mathbf{V} \cdot \mathbf{V}/2, \quad (13.89)$$

where \mathbf{V} is the inertial velocity of a particle. In an inertial reference frame it is trivial to show that

$$\frac{1}{2} \frac{d\mathcal{K}}{dt} = \mathbf{V} \cdot \mathbf{A} \quad (13.90)$$

through use of the chain rule, where $\mathbf{A} = d\mathbf{V}/dt$ is the inertial acceleration. Verify that this identity also holds in the rotating reference frame. For simplicity make use of planetary Cartesian coordinates.

14

Particle dynamics

We here develop the dynamical equations for a point particle of fixed mass and zero electric charge moving around a rotating and gravitating sphere. Geophysical fluids are close to solid-body motion. So for the particle motion to correspond to motion of a geophysical fluid, we are most interested in motion relative to that of the moving sphere. From an inertial reference frame, the only force acting on the particle is from the gravitational field of the sphere (ignoring friction and other forces). A particle at rest in the sphere's rotating reference frame has both kinetic energy and angular momentum due to the solid-body motion. We study particle dynamics as viewed in this non-inertial rotating frame. Doing so provides a useful introduction to rotating dynamics that will serve us when moving to geophysical fluids.

READER'S GUIDE TO THIS CHAPTER

We make use of the mathematics and kinematics introduced for the particle motion in Chapter 13, thus making this chapter relatively brief. Later chapters on rotating fluid dynamics make use of material introduced here.

14.1	Gravitational force and potential energy	178
14.1.1	Newton's Gravitational Law	178
14.1.2	Effective gravitational force from the geopotential	179
14.1.3	Further study	179
14.2	Newton's law of motion	180
14.2.1	Cartesian coordinate representation	180
14.2.2	Spherical coordinate representation	180
14.2.3	Geopotential coordinates for slightly oblate spheroids	181
14.2.4	Further study	182
14.3	Exercises	182

14.1 Gravitational force and potential energy

The mechanical energy for a particle moving around the sphere consists of the gravitational potential energy plus the kinetic energy. Since the point particle contains no internal structure and it has no surface area, the total energy for the particle equals to the mechanical energy; i.e., it has no internal energy. We here discuss the gravitational potential energy and the associated gravitational force. This force is the only inertial force that we consider for the point particle.

14.1.1 Newton's Gravitational Law

For a point particle of mass m moving around the sphere, potential energy (SI units $\text{kg m}^2 \text{s}^{-2}$) is associated with motion through the gravitational field. We write this potential energy as

$$P = m \Phi_e, \quad (14.1)$$

where Φ_e is the gravitational potential (SI units $\text{m}^2 \text{s}^{-2}$) determined from Newton's Law of Gravity. For a spherical mass, the gravitational potential is given by

$$\Phi_e = -\frac{G M}{r} \quad (14.2)$$

where M is the mass of the sphere and

$$G = 6.674 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2} = 6.674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2} \quad (14.3)$$

is Newton's gravitational constant. The gradient of the gravitational potential gives the gravitational acceleration

$$\mathbf{g}_e = -\nabla \Phi_e = -\frac{G M}{r^2} \hat{\mathbf{r}} \quad (14.4)$$

so that the gravitational force acting on the particle is given by

$$\mathbf{F}_{\text{gravity}} = m \mathbf{g}_e = -m \nabla \Phi_e. \quad (14.5)$$

For atmospheric and oceanic fluid dynamics, it is often sufficient to assume the gravitational acceleration is constant and equal to its value at the surface of the sphere.¹ In this case

$$\mathbf{g}_e = -g_e \hat{\mathbf{r}}, \quad (14.6)$$

¹The assumption of constant gravitational field is not appropriate when considering details of oceanic or atmospheric tidal motions or when aiming for precise measures of sea level. We consider more general gravitational fields in Chapter 34.

where

$$g_e = \frac{G M_e}{R_e^2} \approx 9.8 \text{ m s}^{-2}. \quad (14.7)$$

To reach this value, we assumed a sphere of mass equal to the earth mass

$$M_e = 5.977 \times 10^{24} \text{ kg}, \quad (14.8)$$

and radius

$$R_e = 6.371 \times 10^6 \text{ m} \quad (14.9)$$

determined so that the sphere has the same volume as the earth. The corresponding gravitational potential is given by

$$\Phi_e = g_e r, \quad (14.10)$$

and the gravitational potential energy is

$$m \Phi_e = m g_e r. \quad (14.11)$$

14.1.2 Effective gravitational force from the geopotential

We can combine the potential for the centrifugal acceleration as given by equation (13.57) with the gravitational potential (14.2), thus resulting in the *geopotential*

$$\Phi = r [g_e - \mathbf{U}_{\text{solid}}^2/(2r)]. \quad (14.12)$$

The contribution from the centrifugal term can be estimated by making use of terrestrial values, in which $R = R_e = 6.371 \times 10^6 \text{ m}$ (equation (14.9)), and from Section 13.1

$$\Omega_e = 7.292 \times 10^{-5} \text{ s}^{-1}. \quad (14.13)$$

The centrifugal term is its largest at the equator, $\phi = 0$, where

$$\frac{\mathbf{U}_{\text{solid}}^2}{2 R_e} \approx 0.017 \text{ m s}^{-2}, \quad (14.14)$$

so that the ratio of the gravitational to centrifugal accelerations is (at most)

$$\frac{g_e}{\mathbf{U}_{\text{solid}}^2/(2R_e)} = \frac{M_e G / R_e^2}{\Omega_e^2 R_e / 2} \approx 576. \quad (14.15)$$

The geopotential is thus dominated by the earth's gravitational potential. Even so, the centrifugal acceleration leads to a slight equatorial bulge on the earth. To account for this slight non-sphericity, geophysical fluid models generally interpret the radial direction $\hat{\mathbf{r}}$ as pointing parallel to $\nabla\Phi$ rather than parallel to $\nabla\Phi_e$. We have more to say on this topic in Section 14.2.3.

14.1.3 Further study

Newton's Gravitational Law is standard material from freshman physics. Some commonly used physical properties of the earth are summarized in Appendix Two of [Gill \(1982\)](#).

14.2 Newton's law of motion

Newton's law of motion (the 2nd law) says that in an inertial reference frame, time changes to the linear momentum arise only from externally applied forces. For a constant mass particle, momentum changes arise from velocity changes; i.e., accelerations. With gravity the only inertial force acting on the particle, Newton's 2nd law says

$$m \mathbf{A} = -m \nabla \Phi_e. \quad (14.16)$$

We now move to the rotating reference frame of terrestrial observers, thus encountering Coriolis and centrifugal accelerations.

14.2.1 Cartesian coordinate representation

The inertial acceleration using Cartesian coordinates is given by equation (13.55b)

$$\mathbf{A} = \mathbf{A}_{\text{Cartesian}} - \mathbf{A}_{\text{Coriolis}} - \mathbf{A}_{\text{centrifugal}} \quad (14.17a)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} + \nabla \Phi_{\text{centrifugal}}, \quad (14.17b)$$

so that the rotating frame Cartesian equation of motion is given by

$$\mathbf{A}_{\text{Cartesian}} = -\nabla \Phi_e - 2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \nabla \Phi_{\text{centrifugal}} \quad (14.18a)$$

$$= -2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \nabla \Phi, \quad (14.18b)$$

where the geopotential is the sum of the gravitational and centrifugal potentials (equation (14.12))

$$\Phi = \Phi_e + \Phi_{\text{centrifugal}}. \quad (14.19)$$

We can write the equation of motion in component form by exposing the indices (Cartesian tensors) and using a dot for time derivative

$$\ddot{X}_a + 2 \epsilon_{abc} \boldsymbol{\Omega}_b \dot{X}_c = -\partial_a \Phi. \quad (14.20)$$

We can write this equation in the standard vector form

$$\frac{d^2 \mathbf{X}}{dt^2} + 2 \boldsymbol{\Omega} \wedge \dot{\mathbf{X}} = -\nabla \Phi, \quad (14.21)$$

where the basis vectors need not be time differentiated again since their change has already been taken care of when exposing the Coriolis and centrifugal accelerations. This equation of motion is the standard form that will recur for a fluid, with the addition of contact forces from pressure and friction (Chapter 22).

14.2.2 Spherical coordinate representation

We now make use of the acceleration written in spherical coordinates as given in Section 13.11

$$\mathbf{A}_{\text{sphere}} + \mathbf{A}_{\text{metric}} = \mathbf{A}_{\text{Coriolis}} + \mathbf{A} + \mathbf{A}_{\text{centrifugal}} \quad (14.22a)$$

$$= -2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{sphere}} - \nabla \Phi. \quad (14.22b)$$

The effective gravitational force is not a central force, and it is not central due to the contribution from the centrifugal acceleration. We see this more explicitly by using the equations in Section 13.11 to write the spherical equations

$$\dot{u} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) = 0 \quad (14.23)$$

$$\dot{v} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi = -r_{\perp} \Omega^2 \sin \phi \quad (14.24)$$

$$\dot{w} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi = r_{\perp} \Omega^2 \cos \phi - g_e. \quad (14.25)$$

The presence of Ω^2 terms in both the meridional and radial equations signals the non-central nature of the effective gravitational acceleration.

14.2.3 Geopotential coordinates for slightly oblate spheroids

As we saw in Section 14.1.1, the radius of a sphere that best fits the volume of the earth is given by

$$R_e = 6.367 \times 10^6 \text{ m}. \quad (14.26)$$

However, the non-central nature of the effective gravitational force leads to an oblate spheroidal shape for planets such as the earth. The result is a distinction between the equatorial and polar radii

$$R_{\text{equator}} = 6.378 \times 10^6 \text{ m} \quad \text{and} \quad R_{\text{pole}} = 6.357 \times 10^6 \text{ m}, \quad (14.27)$$

with a corresponding ratio

$$1 - \frac{R_{\text{pole}}}{R_{\text{equator}}} \approx 3 \times 10^{-3}. \quad (14.28)$$

An oblate spheroid shape does a better job fitting the actual earth shape than a sphere, thus motivating the use of oblate spheroid coordinates for planetary scale mechanics. In this case, the radial coordinate is constant on the oblate spheroid shaped geopotential.

Even though oblate spheroidal coordinates are better than spherical for describing geopotentials, it is possible, to a high degree of accuracy, to describe the earth's geometry as spherical. Doing so simplifies the mathematics since oblate spheroidal coordinates are less convenient than spherical. We are thus led to assume that the radial coordinate measures distances perpendicular to the geopotential, yet to use geometric/metric functions based on spherical coordinates. The error in this approach is small for the earth, and well worth the price since we no longer have a non-radial component to the effective gravitational force. We illustrate the situation in Figure 14.1. Absorbing the centrifugal term into an effective gravitational potential then leads to the effective gravitational acceleration vector

$$-\nabla\Phi = -g\hat{\mathbf{r}}, \quad (14.29)$$

with g the effective gravitational acceleration. Using this convention, the particle equations of motion take the following form

$$\dot{u} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) = 0 \quad (14.30)$$

$$\dot{v} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi = 0 \quad (14.31)$$

$$\dot{w} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi = -g. \quad (14.32)$$

Notably, the effective gravitational acceleration only impacts the radial equation of motion.

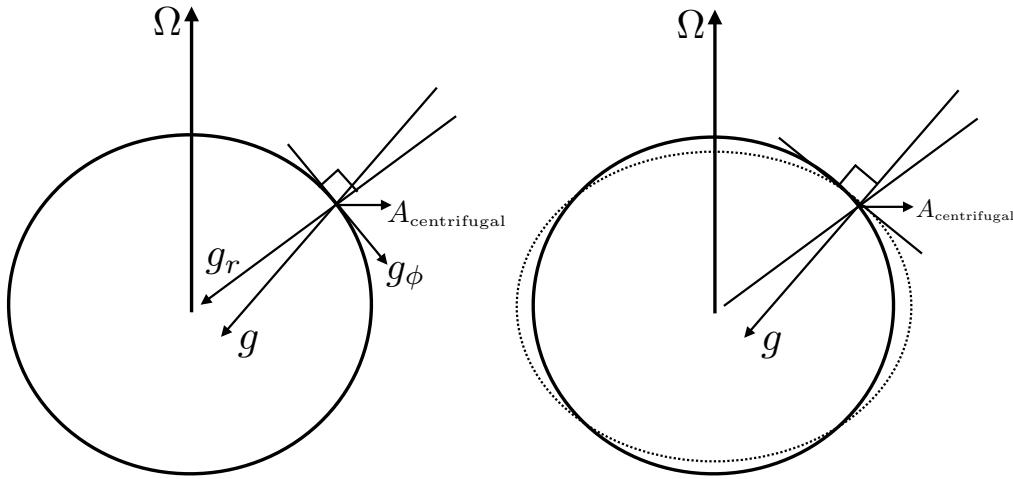


Figure 14.1: This figure illustrates the geopotential vertical coordinate system used to study geophysical fluids. The left panel shows the non-central nature of the effective gravitational force \mathbf{g} , which is given by the sum of the gravitational acceleration \mathbf{g}_r plus centrifugal acceleration $\mathbf{A}_{\text{centrifugal}}$. The gravitational acceleration points radially to the center of the sphere, whereas the centrifugal acceleration points outward from the axis of rotation. The right panel shows a reinterpreted vertical coordinate $r = R + z$ as measuring distance perpendicular to the geopotential surface (dotted surface). Using this geopotential vertical coordinate, the effective gravitational force is aligned with the vertical coordinate and so has no “horizontal” component ($g_\phi = 0$). This figure is not drawn to scale, with the oblate nature highly exaggerated compared to the real earth system (see equation (14.28)).

14.2.4 Further study

Section 4.12 of [Gill \(1982\)](#) and section 2.2.1 of [Vallis \(2017\)](#) present the relevant terrestrial scaling to justify spherical coordinates with a radial effective gravitational potential. [Morse and Feshbach \(1953\)](#) and [Veronis \(1973\)](#) present details of spheroidal coordinates.

14.3 Exercises

EXERCISE 14.1: GEOMETRY OF CONSTANT GEOPOTENTIAL SURFACES

Here we examine some properties of the geopotential given by equation (14.12), where the squared solid-body speed is $\mathbf{U}_{\text{solid}}^2 = (\Omega r \cos \phi)^2$.

- (A) Sketch surfaces of constant geopotential according to the expression (14.12).
- (B) By equating the geopotential going around the pole to that going around the equator, show that the polar radius is less than the equatorial radius when $\Omega > 0$.
- (C) Taking the terrestrial values of g_e , R_{equator} , and Ω , what is the polar radius R_{pole} ? Compare to the measured value of the polar radius given by equation (14.27).

EXERCISE 14.2: SCALING TO JUSTIFY USE OF GEOPOTENTIAL COORDINATES

Summarize the argument that justifies the use of geopotential coordinates while retaining the spherical geometry. Make use of your favorite textbook discussion such that given in Chapter 2 of [Vallis \(2017\)](#).

15

Symmetries and conservation laws

This chapter explores the constraints on particle motion due to conservation laws resulting from space-time symmetries. These ideas are fundamental to how we garner a qualitative understanding of motion, whereby it is generally more useful to know the dynamically conserved properties shared by all trajectories rather than details of any particular trajectory. Conservation laws provide useful predictive capabilities and help in designing analytical and numerical solution methods. These ideas extend to our study of geophysical fluid mechanics.

READER'S GUIDE TO THIS CHAPTER

This chapter is an extension of the particle dynamics discussed in Chapter 14. We make use of the ideas introduced in this chapter when considering the motion of a geophysical fluid in Chapter 22.

15.1	Trajectories and dynamical constraints	184
15.1.1	Connecting symmetries to conservation laws	185
15.1.2	Further study	185
15.2	Potential momentum	185
15.3	Inertial oscillations: motion conserving potential momentum	186
15.3.1	Oscillator equation	187
15.3.2	Particle trajectory and velocity	187
15.3.3	Period of inertial oscillations	187
15.3.4	Comments and further study	188
15.4	Dynamical constraints arising from geometric symmetries	188
15.4.1	Linear momentum conservation	188
15.4.2	Potential momentum conservation	189
15.4.3	Angular momentum conservation	189
15.5	Work and kinetic energy	190
15.5.1	Kinetic energy and Galilean invariance	190
15.5.2	Cartesian expression in a rotating reference frame	191
15.5.3	Kinetic energy using spherical coordinates: Part I	191
15.5.4	Kinetic energy using spherical coordinates: Part II	191
15.6	Mechanical energy conservation	192
15.6.1	General expression for kinetic energy evolution	192
15.6.2	Conservation of mechanical energy	192
15.6.3	Comments and further study	193
15.7	Axial angular momentum conservation	193
15.7.1	Angular momentum	193
15.7.2	Conservation of axial angular momentum	194
15.8	Coriolis & axial angular momentum conservation	194
15.8.1	Constraints imposed by axial angular momentum conservation	195
15.8.2	Axial angular momentum conservation yields zonal Coriolis acceleration	196
15.8.3	Zonal acceleration induced by meridional motion	196
15.8.4	Zonal acceleration induced by radial motion	197
15.8.5	Zonal acceleration derived from axial angular momentum conservation	197
15.8.6	Nearly horizontal motions	198
15.9	More facets of motion on the rotating sphere	198
15.9.1	Zonal motion is not inertial motion	199
15.9.2	Coriolis acceleration in the meridional equation	199
15.9.3	Comments	200
15.10	Exercise	200

15.1 Trajectories and dynamical constraints

From Newton's law of motion, the trajectory of a particle is specified so long as we know the forces acting on the particle and the particle's initial position and initial velocity. The trajectory encapsulates all dynamical information about the moving particle. However, it is often difficult to unpack that information to understand the nature of the motion. Knowledge of the trajectory is not always the best route to dynamical insight.

For dynamical insight it is generally more useful to develop an understanding of constraints respected by the motion, with dynamical constraints manifesting as conservation laws. For example, does the motion conserve mechanical energy? What about angular momentum? If dynamical constraints are present, then all trajectories, regardless their complexity, satisfy the constraints.

Knowledge of the constraints can reveal dynamical insights often hidden when just having information about the trajectory. Constraints also provide predictive statements of great value when studying the stability of motion and for developing numerical methods for simulations.

15.1.1 Connecting symmetries to conservation laws

The discovery of conservation laws often comes from inspired manipulations of the equations of motion. However, there is a more robust and fundamental means to deduce conservation laws through their connection to symmetries, with a symmetry manifesting as an operation that leaves the physical system unchanged. For example, does the physical system remain unchanged when shifting the origin of time? If so, then mechanical energy is a constant of the motion. Likewise, if there is rotational symmetry around an axis, then the associated angular momentum is a constant of the motion. The connection between symmetries (kinematics) and conservation laws (dynamics) was made by E. Noether in 1918. Noether's Theorem is fundamental to all areas of physics. We will not delve into the mathematical details of Noether's Theorem. Instead, it is sufficient for our study to make use of it as a conceptual framework for understanding conservation laws. Namely, if there is a symmetry, then there is a conservation law.

It is very useful to identify conserved quantities as a means to understand and to constrain the motion. This perspective holds even when the symmetries giving rise to conserved quantities are broken in realistic cases. For example, friction breaks time translation symmetry and so leads to the dissipation of mechanical energy. Nonetheless, understanding the frictionless motion, and the associated energy conservation law, offers insights for the frictional case as well. Indeed, for many purposes, knowledge of the trajectory is less important than knowledge of conserved, or partially conserved, dynamical quantities. In this chapter, we offer two examples to support this point: the case of mechanical energy conservation and axial angular momentum conservation. These conservation laws also hold in a modified form when moving to the continuum fluid (e.g., Chapter 22). Additional conservation properties also arise that are unique to the continuum, with conservation of potential vorticity the most notable one for geophysical fluids (Chapter 50).

15.1.2 Further study

The notions of conservation laws and symmetries in classical mechanics are lucidly discussed in Chapters 1 and 2 of [Landau and Lifshitz \(1976\)](#). A pedagogical lecture on these topics can be found in this [online lecture from the Space Time series](#). Finally, the student is encouraged to read [this essay about Emmy Noether](#), whose work united symmetry and conservation laws.

15.2 Potential momentum

We here introduce the notion of *potential momentum*, which is a constant of the motion for a particle moving on a time independent geopotential in a direction where the geopotential does not change. That is, the conservation of potential momentum arises from a symmetry of the geopotential.

To start the analysis, we note that if the planetary rotation is a constant in time, then the Cartesian equation of motion (14.21) can be written

$$\frac{d}{dt} \left[\dot{\mathbf{X}} + 2\boldsymbol{\Omega} \wedge \mathbf{X} \right] = -\nabla\Phi. \quad (15.1)$$

This form suggests we introduce the *potential momentum* per mass

$$\mathbf{M} \equiv \frac{d\mathbf{X}}{dt} + 2\boldsymbol{\Omega} \wedge \mathbf{X} = \hat{\mathbf{x}}(u - 2\Omega y) + \hat{\mathbf{y}}(v + 2\Omega x) + \hat{\mathbf{z}}w, \quad (15.2)$$

in which case the momentum equation is written

$$\frac{d\mathbf{M}}{dt} = -\nabla\Phi. \quad (15.3)$$

Now let $\hat{\mathbf{s}}$ be a unit vector tangent to the geopotential surface so that $\hat{\mathbf{s}} \cdot \nabla\Phi = 0$. Assuming the geopotential surface is time independent, the equation of motion (15.3) leads to

$$\frac{d(\hat{\mathbf{s}} \cdot \mathbf{M})}{dt} = 0. \quad (15.4)$$

That is, the projection of the potential momentum onto a static geopotential surface is a constant of motion. This dynamical constraint arises since we cannot distinguish one point on the geopotential from another; i.e., there is a symmetry associated with motion along the static geopotential. Noether's Theorem (Section 15.1.1) then says that this geometric symmetry leads to a constant of the motion, here given by that component of potential momentum within the geopotential surface. We illustrate this situation in Figure 15.1 with a horizontal geopotential surface.

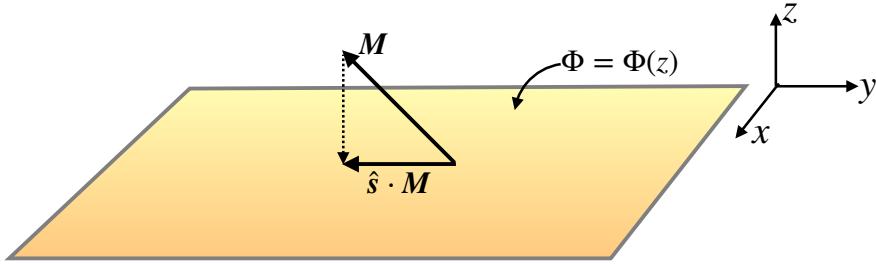


Figure 15.1: The projection of the potential momentum onto the geopotential surface is a constant of the motion, $d(\hat{\mathbf{s}} \cdot \mathbf{M})/dt = 0$. Here the geopotential surface is the x - y -plane so that $\hat{\mathbf{x}} \cdot \mathbf{M} = u - 2\Omega y$ and $\hat{\mathbf{y}} \cdot \mathbf{M} = v + 2\Omega x$ are the two conserved components of potential momentum.

Consider a particle with potential momentum \mathbf{M} and move it from an arbitrary point to a reference position with $\mathbf{X} = 0$. Upon reaching the reference position, the horizontal velocity of the particle must equal to \mathbf{M} in order to maintain the same potential momentum. This example motivates the term “potential”, since \mathbf{M} measures the potential for relative motion contained in the particle as it moves along a geopotential.

15.3 Inertial oscillations: motion conserving potential momentum

In Section 27.2 we introduce the tangent plane approximation for motion on a rotating sphere. In this approximation, motion occurs on a rotating geopotential surface with the surface approximated as horizontal. The f -plane approximation furthermore sets the Coriolis parameter

$$f = 2\Omega \sin \phi_0 \quad (15.5)$$

to a constant, where ϕ_0 is a chosen latitude. Consequently, a particle moving on the f -plane maintains a constant horizontal potential momentum

$$\frac{dM_x}{dt} = \frac{d(u - fy)}{dt} = 0 \quad (15.6a)$$

$$\frac{dM_y}{dt} = \frac{d(v + fx)}{dt} = 0, \quad (15.6b)$$

where we introduced the horizontal velocity components $(u, v) = (\dot{x}, \dot{y})$. These two conservation laws greatly constrain the motion of the particle moving on a constant geopotential surface.

15.3.1 Oscillator equation

Taking the time derivative of the zonal equation (15.6a) and using the meridional equation (15.6b) leads to

$$\ddot{u} - f \dot{v} = \ddot{u} + f^2 u = 0. \quad (15.7)$$

Similar manipulations for the meridional velocity equation render the free oscillator equation for each component of the horizontal velocity

$$\frac{d^2 \mathbf{U}}{dt^2} + f^2 \mathbf{U} = 0. \quad (15.8)$$

Motions that satisfy this equation are termed *inertial oscillations*.

15.3.2 Particle trajectory and velocity

Time integrating the equation of motion (15.8) renders the particle trajectory and its velocity

$$\mathbf{X}(t) = \frac{U}{f} \left[\hat{\mathbf{x}} \sin(ft) + \hat{\mathbf{y}} \cos(ft) \right] \quad (15.9a)$$

$$\mathbf{U}(t) = U \left[\hat{\mathbf{x}} \cos(ft) - \hat{\mathbf{y}} \sin(ft) \right], \quad (15.9b)$$

where U is the particle speed, which is a constant, and we assumed the initial conditions

$$\mathbf{X}(0) = \frac{U}{f} \hat{\mathbf{y}} \quad (15.10a)$$

$$\mathbf{U}(0) = U \hat{\mathbf{x}}. \quad (15.10b)$$

Motion is circular with radius

$$R = |U| f^{-1}. \quad (15.11)$$

Northern hemisphere ($f > 0$) inertial oscillations occur in the clockwise direction whereas southern hemisphere motion is counter-clockwise. Consequently, particle motion undergoing inertial oscillations occurs in an anti-cyclonic sense (opposite to the sense of the rotating reference frame). As discussed in Section 30.3.2, inertial oscillations arise from a balance between the Coriolis acceleration of the rotating frame and the centrifugal acceleration of the particle's circular motion. The only way to realize this balance is for the particle to move anti-cyclonically, with the Coriolis acceleration pointing towards the inside of the inertial circle and the centrifugal acceleration pointing outside. Finally, note that the potential momentum for inertial oscillations vanishes since

$$\mathbf{M}(t) = \mathbf{U}(t) + f \hat{\mathbf{z}} \wedge \mathbf{X}(t) = 0. \quad (15.12)$$

Adding an arbitrary constant to the initial position makes the potential momentum equal to a nonzero constant.

15.3.3 Period of inertial oscillations

Inertial oscillations possess a constant speed and move around the inertial circle with a period

$$T_{\text{inertial}} = \frac{2\pi}{f} = \frac{11.97}{|\sin \phi_0|} \text{ hour}, \quad (15.13)$$

where we set $\Omega = 7.292 \times 10^{-5} \text{ s}^{-1}$ (equation (13.1)). This period is the time it takes to go around the circle. It is smallest at the poles, where the latitude $\phi_0 = \pm\pi/2$ and $T_{\text{smallest}} \approx 12$ hour. At the equator, $\phi_0 = 0$, so that the radius of the inertial circle is infinite and inertial oscillation is available. Furthermore, T_{inertial} is the time for a Foucault pendulum to turn through π radians, so that T_{inertial} is sometimes referred to as one-half a pendulum day.

15.3.4 Comments and further study

Inertial oscillations of fluid parcels are described by the above constant potential momentum equation of motion. Such oscillations are commonly measured by ocean current meters, especially in higher latitude regions where diurnal (day-night) variations in wind forcing have a strong projection onto the inertial period. This resonant forcing puts energy into inertial or near-inertial motions. It is quite amazing that such oscillations are indeed found in the ocean, given that we have ignored pressure and friction which also impact on fluid parcels. The main reason we can observe this motion in the ocean is that upper ocean currents are often generated by winds even in the absence of horizontal pressure gradients, thus allowing us to drop the pressure gradient from the momentum equation.

We encounter inertial motions again in Section 30.3 as part of our characterization of horizontal fluid motion according to the balance between forces. We arrive at inertial motion by balancing the Coriolis acceleration with the centrifugal acceleration due to the curved motion of a fluid particle. Note that the name “inertial” does not here refer to motion in an inertial reference frame (Section 13.2). Instead, it refers to the balance between accelerations arising only when the particle is in motion (i.e., has inertia), with these accelerations being the Coriolis and centrifugal.

A rotating tank offers a useful controlled setting to observe inertial oscillations, such as shown near the 18 minute mark in [this video from Prof. Dave Fultz of the University of Chicago](#).

15.4 Dynamical constraints arising from geometric symmetries

We here position the conservation of potential momentum among two other dynamical conservation laws by summarizing the spatial symmetries that lead to the conservation laws via Noether’s Theorem.

15.4.1 Linear momentum conservation

Linear momentum remains constant for a free particle moving without any forces acting on it. The conservation of linear momentum is most readily viewed within the particle’s inertial reference frame, where a vanishing inertial acceleration leads to a constant inertial velocity

$$\mathbf{A} = 0. \tag{15.14}$$

When viewed from a rotating frame using Cartesian coordinates, a vanishing inertial acceleration means that the Cartesian acceleration balances Coriolis and centrifugal accelerations

$$\ddot{\mathbf{X}} = -2\boldsymbol{\Omega} \wedge \dot{\mathbf{X}} - \nabla\Phi_{\text{centrifugal}}. \tag{15.15}$$

This equation can be written in terms of the potential momentum

$$\frac{d\mathbf{M}}{dt} = -\nabla\Phi_{\text{centrifugal}}. \tag{15.16}$$

Hence, if $\Phi_{\text{centrifugal}}$ is static, then motion maintaining constant inertial momentum also maintains $\hat{\mathbf{s}} \cdot \mathbf{M} = 0$ constant, where $\hat{\mathbf{s}}$ is a vector tangent to constant $\Phi_{\text{centrifugal}}$ surfaces.

15.4.2 Potential momentum conservation

Again, the conservation of potential momentum arises from symmetry of particle motion on a constant geopotential surface. The conservation law is most readily viewed within the rotating frame, whereby (equation (15.4))

$$\frac{d(\hat{s} \cdot \mathbf{M})}{dt} = 0. \quad (15.17)$$

A geopotential is a two-dimensional surface so that this conservation law corresponds to two dynamical constraints such as shown in Figure 15.1.

15.4.3 Angular momentum conservation

As detailed in Section 15.7, the angular momentum computed with respect to the axis of rotation is a constant of the motion (Figure 15.2). This conservation law arises from rotational symmetry of the system about the rotational axis. Axial angular momentum conservation takes the form

$$\frac{dL^z}{dt} = 0, \quad (15.18)$$

where the axial angular momentum is

$$L^z = m r_{\perp}^2 (\dot{\lambda} + \Omega) \quad \text{with} \quad r_{\perp} = \sqrt{x^2 + y^2}, \quad (15.19)$$

r_{\perp} is the distance from the rotation axis; i.e., it is the *moment arm*, λ is the longitude that measures the angle in the counter-clockwise direction from the positive x -axis, and $\dot{\lambda}$ is the time change of the longitude.

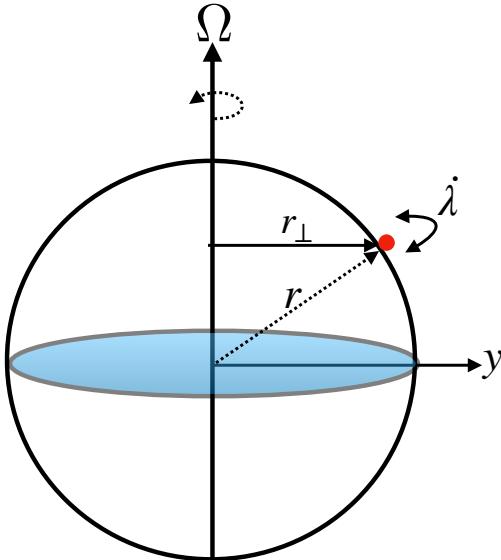


Figure 15.2: Axial angular momentum, $L^z = m r_{\perp}^2 (\dot{\lambda} + \Omega)$, is the moment of the zonal momentum around the sphere, with $r_{\perp} = \sqrt{x^2 + y^2}$ the moment arm length; i.e., the distance of the particle to the axis of rotation, whereas $r^2 = \sqrt{x^2 + y^2 + z^2}$ is the radial distance to the center of the sphere. L^z is a constant of the motion for the particle moving in the absence of friction.

15.5 Work and kinetic energy

Consider the equation of motion for a constant mass particle

$$m \frac{d\mathbf{V}}{dt} = \mathbf{F}. \quad (15.20)$$

Now project this equation onto the velocity vector to render

$$\frac{d\mathcal{K}}{dt} = \mathbf{F} \cdot \mathbf{V}, \quad (15.21)$$

where we identified the particle's kinetic energy

$$\mathcal{K} = \frac{m}{2} \mathbf{V} \cdot \mathbf{V}, \quad (15.22)$$

which is the energy contained in the motion of the particle with respect to an inertial reference frame. Time integration of equation (15.21) leads to

$$\mathcal{K}(t_2) - \mathcal{K}(t_1) = \int_{t_1}^{t_2} \mathbf{F} \cdot \mathbf{V} dt. \quad (15.23)$$

The integrand on the right hand side is the time change of *work* applied to the particle, which is referred to as the *power*. Hence, the time integral of the power is the work done on the particle during its motion, which can be written

$$\mathcal{K}(t_2) - \mathcal{K}(t_1) = \int_{t_1}^{t_2} \mathbf{F} \cdot d\mathbf{X}, \quad (15.24)$$

where $\mathbf{V} dt = d\mathbf{X}$. Hence, the time difference in the particle's kinetic energy is determined by the work applied to the particle over the time interval.

15.5.1 Kinetic energy and Galilean invariance

The kinetic energy is *not* Galilean invariant since movement to another inertial reference frame leads to the change

$$\bar{\mathbf{V}} = \mathbf{V} + \mathbf{U} \implies \bar{\mathcal{K}} = \mathcal{K} + \frac{m}{2} (2 \mathbf{V} \cdot \mathbf{U} + \mathbf{U} \cdot \mathbf{U}), \quad (15.25)$$

where \mathbf{U} is a constant boost velocity so that $d\mathbf{U}/dt = 0$. Even so, the time change of the kinetic energy in the new inertial frame is still given by the power in the new frame

$$\frac{d\bar{\mathcal{K}}}{dt} = \frac{d\mathcal{K}}{dt} + m \mathbf{A} \cdot \mathbf{U} \quad (15.26a)$$

$$= \mathbf{F} \cdot \mathbf{V} + \mathbf{F} \cdot \mathbf{U} \quad (15.26b)$$

$$= \mathbf{F} \cdot \bar{\mathbf{V}}. \quad (15.26c)$$

We generally consider external (inertial) forces that have no velocity dependence, such as the gravitational force. These forces are identical for every inertial observer.

15.5.2 Cartesian expression in a rotating reference frame

We now consider the expression for the kinetic energy when introducing the velocity of the rotating reference frame. Writing the inertial velocity in the planetary Cartesian form

$$\mathbf{V} = \mathbf{V}_{\text{Cartesian}} + \mathbf{U}_{\text{solid}}, \quad (15.27)$$

leads to the kinetic energy

$$\mathcal{K} = \frac{m}{2} [\mathbf{V}_{\text{Cartesian}} \cdot \mathbf{V}_{\text{Cartesian}} + 2 \mathbf{V}_{\text{Cartesian}} \cdot \mathbf{U}_{\text{solid}} + \mathbf{U}_{\text{solid}} \cdot \mathbf{U}_{\text{solid}}] \quad (15.28)$$

The first term arises from motion of the particle relative to the rotating sphere; the second arises from coupling between relative velocity and solid-body velocity; and the third arises from solid-body motion of the sphere.

15.5.3 Kinetic energy using spherical coordinates: Part I

To expose spherical symmetry of the physical system, we express the kinetic energy in terms of the planetary spherical coordinates defined in Figure 13.4. Doing so for the solid body velocity leads to equation (13.44)

$$\mathbf{U}_{\text{solid}} = \Omega r \cos \phi (-\sin \lambda \hat{\mathbf{x}} + \cos \lambda \hat{\mathbf{y}}). \quad (15.29)$$

Likewise, the velocity components measured in the rotating frame are given by

$$\dot{X} = \frac{d(r \cos \phi \cos \lambda)}{dt} = \dot{r} \cos \phi \cos \lambda - r \dot{\phi} \sin \phi \cos \lambda - r \dot{\lambda} \cos \phi \sin \lambda \quad (15.30a)$$

$$\dot{Y} = \frac{d(r \cos \phi \sin \lambda)}{dt} = \dot{r} \cos \phi \sin \lambda - r \dot{\phi} \sin \phi \sin \lambda + r \dot{\lambda} \cos \phi \cos \lambda \quad (15.30b)$$

$$\dot{Z} = \frac{d(r \sin \phi)}{dt} = \dot{r} \sin \phi + r \dot{\phi} \cos \phi. \quad (15.30c)$$

Bringing terms together then leads to the kinetic energy in terms of spherical coordinates

$$\mathcal{K} = (m/2) \left[(\dot{r}^2 + r^2 \dot{\phi}^2 + \dot{\lambda}^2 r^2 \cos^2 \phi) + (2 \Omega r^2 \dot{\lambda} \cos^2 \phi) + (\Omega r \cos \phi)^2 \right]. \quad (15.31)$$

15.5.4 Kinetic energy using spherical coordinates: Part II

An alternative means for deriving the kinetic energy in equation (15.31) makes use of the spherical coordinate form of the inertial velocity given by equation (13.40f), in which case

$$\mathbf{V} = (u + r_{\perp} \Omega) \hat{\lambda} + v \hat{\phi} + w \hat{r}, \quad (15.32)$$

so that

$$\mathcal{K} = \frac{m}{2} [(u + r_{\perp} \Omega)^2 + v^2 + w^2], \quad (15.33)$$

where $r_{\perp} = r \cos \phi$. Additionally, as discussed in Section 15.7, the axial angular momentum is given by

$$L^z = m(u + r_{\perp} \Omega). \quad (15.34)$$

Consequently, the kinetic energy can be written

$$\mathcal{K} = \frac{(L^z)^2}{2m} + \frac{m}{2} (v^2 + w^2). \quad (15.35)$$

15.6 Mechanical energy conservation

Does the particle know anything about the time origin? Since the angular velocity of the planet and the gravitational acceleration are both assumed constant in time, then changing the time will leave the physical system unaltered. That is, the physical system remains unchanged if we shift all clocks by a constant amount. Through Noether's Theorem, this symmetry in time leads to mechanical energy conservation. That is, the particle's mechanical energy is fixed by the initial conditions. We here prove that mechanical energy is constant by manipulating the equations of motion. Many of the manipulations also occur when considering the mechanical conservation laws for a continuum fluid discussed in Chapter 22.

15.6.1 General expression for kinetic energy evolution

As seen in Section 15.5, the time derivative of the kinetic energy is given by

$$\frac{d\mathcal{K}}{dt} = m \mathbf{V} \cdot \frac{d\mathbf{V}}{dt} \quad (15.36a)$$

$$= m \mathbf{V} \cdot \mathbf{A} \quad (15.36b)$$

$$= -m \mathbf{V} \cdot \nabla \Phi_e. \quad (15.36c)$$

For the final equality we introduced the gravitational potential given that the particle only feels an external force from gravity as per equation (14.16).

15.6.2 Conservation of mechanical energy

The gravitational potential is given by (see equation (14.10))

$$\Phi_e = g_e r, \quad (15.37)$$

so that

$$\frac{d\mathcal{K}}{dt} = -m \mathbf{V} \cdot \nabla \Phi_e = -m g_e \dot{r}. \quad (15.38)$$

This result means that kinetic energy is reduced when moving the particle away from the earth center ($\dot{r} > 0$). Moving away from the earth requires work to overcome the gravitational attraction. This work to overcome the gravitational attraction is taken away from the kinetic energy of the particle. Furthermore, the work is added to the gravitational potential energy, whose evolution is given by (see equation (14.11))

$$\frac{d\mathcal{P}_e}{dt} = m g_e \dot{r}, \quad (15.39)$$

where we assumed a constant gravitational acceleration g_e . Consequently, as the particle moves away from the earth center, its reduction in kinetic energy is exactly compensated by an increase in potential energy. Hence, the mechanical energy for the particle remains constant throughout the motion

$$\frac{d(\mathcal{K} + \mathcal{P}_e)}{dt} = 0, \quad (15.40)$$

where the mechanical energy is the sum of the inertial kinetic energy plus the gravitational potential energy

$$m = \mathcal{K} + \mathcal{P}_e \quad (15.41a)$$

$$= \frac{m}{2} \mathbf{V} \cdot \mathbf{V} + m \Phi_e \quad (15.41b)$$

$$= \frac{m}{2} [(u + r_{\perp} \Omega)^2 + v^2 + w^2] + m g_e r \quad (15.41c)$$

$$= \frac{(L^z)^2}{2m} + \frac{m}{2} (v^2 + w^2) + m g_e r. \quad (15.41d)$$

15.6.3 Comments and further study

Rotation plays no role in the evolution of either the kinetic energy or gravitational potential energy. That is, there are no appearance of the Coriolis and centrifugal accelerations in the energy evolution equations. This result is not surprising, since the Coriolis and centrifugal accelerations arise through our subjective choice to view the motion from a rotating reference frame, and that choice plays no role in the evolution of total mechanical energy. Various forms of the discussion in this section can be found in books on classical mechanics, with a lucid and pedagogical treatment given by *Marion and Thornton* (1988).

15.7 Axial angular momentum conservation

Does the particle know anything about the longitudinal angle, λ ? Since we assume that the sphere is smooth (i.e., no mountains), and since the sphere is rotating around the polar axis, there is an arbitrariness in how we choose the zero for the longitudinal angle. That is, the physical system remains unchanged if we shift the longitudinal angle by a constant. Noether's Theorem then says that this rotational symmetry leads to a corresponding angular momentum conservation. Hence, the particle's angular momentum around the rotational axis remains fixed by the initial conditions. We here prove that axial angular momentum is constant by manipulating the equations of motion. Many of the manipulations also occur when considering angular conservation for a continuous fluid (see Section 22.4).

15.7.1 Angular momentum

The angular momentum of the particle, computed with respect to the origin of the sphere, is given by

$$\mathbf{L} = m \mathbf{X} \wedge \mathbf{V}. \quad (15.42)$$

This is the moment of the linear momentum computed relative to the origin. We write the angular momentum computed along the polar axis as

$$L^z = \mathbf{L} \cdot \hat{\mathbf{z}} \quad (15.43a)$$

$$= m (\mathbf{X} \wedge \mathbf{V}) \cdot \hat{\mathbf{z}} \quad (15.43b)$$

$$= m (\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{V} \quad (15.43c)$$

$$= m r \cos \phi (\hat{\boldsymbol{\lambda}} \cdot \mathbf{V}) \quad (15.43d)$$

$$= m r_{\perp} (\hat{\boldsymbol{\lambda}} \cdot \mathbf{V}). \quad (15.43e)$$

Hence, the angular momentum about the polar axis equals to the component of the linear momentum in the longitudinal direction, multiplied by the distance to the polar rotational axis (the moment-arm)

$$r_{\perp} = r \cos \phi. \quad (15.44)$$

In deriving equation (15.43e), we made use of the identity

$$\hat{\mathbf{z}} \wedge \mathbf{X} = r_{\perp} \hat{\boldsymbol{\lambda}}, \quad (15.45)$$

which is useful for proving that axial angular momentum is a constant of the motion.

We now write the axial angular momentum in equation (15.43e) in terms of the rotating frame quantities. To do so, introduce the inertial velocity written using spherical coordinates according to equation (13.40f), which yields

$$L^z = m r_{\perp} (\hat{\boldsymbol{\lambda}} \cdot \mathbf{V}) \quad (15.46a)$$

$$= m r_{\perp}^2 (\dot{\lambda} + \Omega) \quad (15.46b)$$

$$= m r_{\perp} (u + r_{\perp} \Omega). \quad (15.46c)$$

When measured from the rotating terrestrial frame, the axial angular momentum consists of two terms: one from the solid-body motion of the planet, and the other from the zonal velocity of the particle relative to the planet.

15.7.2 Conservation of axial angular momentum

The time derivative of the axial angular momentum is given by

$$\frac{dL^z}{dt} = m \frac{d}{dt} [(\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{V}] \quad (15.47a)$$

$$= m (\hat{\mathbf{z}} \wedge \mathbf{V}) \cdot \mathbf{V} + m (\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{A} \quad (15.47b)$$

$$= m (\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{A} \quad (15.47c)$$

$$= m r_{\perp} \hat{\boldsymbol{\lambda}} \cdot \mathbf{A}. \quad (15.47d)$$

To reach this result we noted that the polar axis direction, $\hat{\mathbf{z}}$, is time independent, and we used the identity (15.45) for the final step. The inertial acceleration arises just from the central-force gravitational field (equation (14.16))

$$\mathbf{A} = -\nabla \Phi_e = -g_e \hat{\mathbf{r}}. \quad (15.48)$$

Since $\hat{\boldsymbol{\lambda}} \cdot \hat{\mathbf{r}} = 0$, we have axial angular momentum conservation

$$\frac{dL^z}{dt} = 0. \quad (15.49)$$

15.8 Coriolis & axial angular momentum conservation

Axial angular momentum conservation greatly constrains the particle motion and in turn it offers a means for interpreting and predicting motion on the sphere. We here explore how these constraints manifest through thought experiments illustrated in Figure 15.3. In so doing we show that the zonal acceleration induced by axial angular momentum conserving motion is identical to the zonal Coriolis acceleration.

Throughout the following we write the axial angular momentum as

$$L^z = m l^z = m r_{\perp}^2 (\dot{\lambda} + \Omega) = m r_{\perp} (u + r_{\perp} \Omega), \quad (15.50)$$

where

$$l^z = L^z/m \quad \text{and} \quad u = r_{\perp} \dot{\lambda} \quad (15.51)$$

are the axial angular momentum per mass and zonal velocity component. Since the mass of the particle is constant, conservation of $L^z = m l^z$ implies conservation of l^z . Note that for most geophysical applications, $l^z > 0$ since solid-body motion dominates over the relative zonal velocity:

$$r_{\perp} \Omega > |u| \quad \text{for most terrestrial motions.} \quad (15.52)$$

15.8.1 Constraints imposed by axial angular momentum conservation

Conservation of axial angular momentum says that it is not possible to change u or $\dot{\lambda}$ without also changing r_{\perp} in such a manner to ensure that l^z remains fixed. To determine the relation between these changes we set $\delta l^z = 0$ thus leading to

$$\delta \dot{\lambda} = -\frac{2 l^z}{r_{\perp}^2} \frac{\delta r_{\perp}}{r_{\perp}} \iff \delta u = -\Omega \delta r_{\perp} \left[1 + \frac{l^z}{\Omega r_{\perp}^2} \right], \quad (15.53)$$

where we set $\delta \Omega = 0$ since the earth's rotation rate is assumed to be fixed. Furthermore, we found it convenient to write the changes in terms of l^z since it is a constant of the motion. As noted above, $l^z > 0$ is generally the case for geophysical fluid motion. Consequently, axial angular momentum conserving motion that brings the particle closer to the rotation axis ($\delta r_{\perp} < 0$) leads to an eastward acceleration ($\delta \dot{\lambda} > 0$ and $\delta u > 0$). The opposite occurs for motion with $\delta r_{\perp} > 0$. These results hold in both the northern and southern hemispheres.

Since $r_{\perp} = r \cos \phi$, the distance to the rotational axis can change either by changing the radial position or the meridional position

$$\delta r_{\perp} = (\cos \phi) \delta r - (r \sin \phi) \delta \phi. \quad (15.54)$$

Assuming these changes occur over a small time increment δt allows us to write

$$\delta r_{\perp}/\delta t = w \cos \phi - v \sin \phi, \quad (15.55)$$

where

$$v = r \delta \phi/\delta t = r \dot{\phi} \quad \text{and} \quad w = \delta r/\delta t = \dot{r} \quad (15.56)$$

introduced the meridional and vertical components to the particle velocity according to equation (13.42). For geophysical fluid motion, changes in vertical distance δr are far smaller than the distance to the earth's center

$$\delta r \ll r. \quad (15.57)$$

Consequently, when $\phi \neq 0$, δr_{\perp} is affected much more by meridional motion at constant radial position (second term on right hand side of equation (15.55)) than by vertical motion at constant latitude (first term). We return to this observation in Section 15.8.6 when discussing nearly horizontal motions, which are the basis for the shallow fluid approximation used to develop the primitive equations in Section 27.1.2.

15.8.2 Axial angular momentum conservation yields zonal Coriolis acceleration

Consider the zonal acceleration induced for a particle at rest whose angular momentum per mass equals $l^z = r_\perp^2 \Omega$. In this case the zonal acceleration in equation (15.53) is

$$\delta\dot{\lambda} = -2\Omega \frac{\delta r_\perp}{r_\perp} \iff \dot{u} = -2\Omega \delta r_\perp = -2\Omega (w \cos \phi - v \sin \phi). \quad (15.58)$$

The expression for \dot{u} is precisely the same as the Coriolis acceleration appearing in the zonal momentum equation (14.30)

$$\dot{u} + \frac{u(w - v \tan \phi)}{r} = \underbrace{-2\Omega (w \cos \phi - v \sin \phi)}_{\text{Coriolis acceleration}}. \quad (15.59)$$

Hence, the Coriolis acceleration appearing in the zonal momentum equation is identical to the zonal acceleration induced by constraining the motion to conserve axial angular momentum. That is, by unpacking the constraint of axial angular momentum conservation to reveal the zonal momentum equation, the Coriolis acceleration is revealed as part of that package. We pursue this connection in Section 15.8.5 by deriving the zonal momentum equation from the axial angular momentum equation.

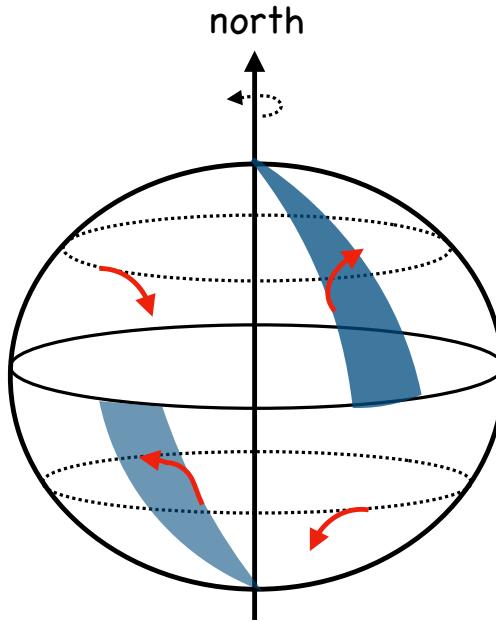


Figure 15.3: In the northern hemisphere, the Coriolis acceleration deflects a particle to its right whereas it deflects a particle to the left in the southern hemisphere. This deflection is also a result of axial angular momentum conservation.

15.8.3 Zonal acceleration induced by meridional motion

Consider a particle moving meridionally ($\delta\phi \neq 0$) while maintaining a constant radial position ($\delta r = 0$). The axial angular momentum constraint (15.58) induces a zonal acceleration

$$\dot{u} = 2\Omega v \sin \phi, \quad (15.60)$$

which, as seen by equation (15.59), is the Coriolis acceleration appearing in the zonal momentum equation arising from the meridional motion. For poleward motion in either hemisphere, $v \sin \phi >$

0. Hence, axial angular momentum conserving motion towards either pole induces an eastward acceleration, whereas a westward acceleration is induced for equatorward motion. For the northern hemisphere, the induced acceleration deflects the particle to the right when looking downstream whereas in the southern hemisphere the induced acceleration deflects the particle to the left. These deflections are illustrated in Figure 15.3.

15.8.4 Zonal acceleration induced by radial motion

Now consider a particle moving radially while holding the latitude fixed ($\delta r \neq 0$ and $\delta\phi = 0$). The axial angular momentum constraint (15.58) induces a zonal acceleration

$$\dot{u} = -2\Omega w \cos\phi, \quad (15.61)$$

which, as seen by equation (15.59), is the Coriolis acceleration appearing in the zonal momentum equation arising from the vertical motion. Hence, for vertically downward motion ($w < 0$), axial angular momentum conservation induces a positive zonal acceleration, $\dot{u} > 0$, which we expect since the particle is moving closer to the rotation axis.

15.8.5 Zonal acceleration derived from axial angular momentum conservation

The above discussion can be formalized by analyzing how the conservation of axial angular momentum leads to an expression for the zonal acceleration. For this purpose, compute the time derivative of the first form of the axial angular momentum in equation (15.50), in which case

$$\frac{1}{m} \frac{dL^z}{dt} = \frac{d[(r \cos\phi)^2 (\dot{\lambda} + \Omega)]}{dt} \quad (15.62a)$$

$$= 2(\dot{r} \cos\phi - r \dot{\phi} \sin\phi)(\dot{\lambda} r \cos\phi + r \Omega \cos\phi) + (r \cos\phi)^2 \ddot{\lambda} \quad (15.62b)$$

$$= 2(w \cos\phi - v \sin\phi)(u + r \Omega \cos\phi) + (r \cos\phi)^2 \ddot{\lambda}, \quad (15.62c)$$

where we introduced the (u, v, w) velocity components according to equation (13.42). With the zonal velocity $u = \dot{\lambda} r \cos\phi$, we have

$$r \cos\phi \ddot{\lambda} = \dot{u} + \frac{u}{r \cos\phi} (v \sin\phi - w \cos\phi), \quad (15.63)$$

so that equation (15.62c) thus takes the form

$$\frac{1}{m} \frac{dL^z}{dt} = 2(w \cos\phi - v \sin\phi)(u + r \Omega \cos\phi) + (r \cos\phi)^2 \ddot{\lambda} \quad (15.64a)$$

$$= (w \cos\phi - v \sin\phi)(u + 2r \Omega \cos\phi) + \dot{u} r \cos\phi. \quad (15.64b)$$

Setting $dL^z/dt = 0$ and rearranging then leads to a prognostic equation for the zonal velocity

$$\frac{du}{dt} = \left[\frac{u}{r \cos\phi} + 2\Omega \right] (v \sin\phi - w \cos\phi). \quad (15.65)$$

The first term in the bracket arises from curvature of the sphere (the “metric acceleration”) whereas the second term is the Coriolis acceleration.

The same result can be obtained by performing the time derivative on the second form of the axial angular momentum in equation (15.50), in which case

$$\frac{1}{m} \frac{dL^z}{dt} = \frac{d[u r \cos\phi + \Omega(r \cos\phi)^2]}{dt} \quad (15.66a)$$

$$= \dot{u} r \cos\phi + u \dot{r} \cos\phi - u r \dot{\phi} \sin\phi + 2\Omega r \cos\phi (\dot{r} \cos\phi - r \dot{\phi} \sin\phi). \quad (15.66b)$$

Again, setting $dL^z/dt = 0$ and rearranging leads to the zonal velocity equation (15.65).

15.8.6 Nearly horizontal motions

We here consider two approximations relevant to large scale geophysical fluid dynamics.

1. The particle kinetic energy is dominated by horizontal motions on the sphere (i.e., motion at constant radial position).
2. Vertical excursions are much smaller than the earth radius.

When applied to a fluid, the first assumption leads to the hydrostatic approximation (Section 27.3), and the second assumption leads to the shallow fluid approximation (Section 27.1.2). Self-consistency of the equations of motion means that these two assumptions must be applied together.

Dropping the vertical velocity component to the kinetic energy leads to

$$\mathcal{K} \approx \frac{m}{2} [(u + r_{\perp} \Omega)^2 + v^2]. \quad (15.67)$$

The second assumption means that the axial angular momentum takes the approximate form

$$L^z \approx m R_{\perp} (u + \Omega R_{\perp}) = m R_{\perp}^2 (\dot{\lambda} + \Omega), \quad (15.68)$$

where

$$r = R + z \approx R \quad \text{and} \quad R_{\perp} = R \cos \phi. \quad (15.69)$$

The approximate angular momentum (15.68) ignores contributions from vertical motion in changing the moment-arm. Indeed, as noted in Section 15.8.1, vertical movements within the atmosphere and ocean (relatively thin fluid layers over the earth's surface) lead to a relatively small modification to the moment-arm, so the assumption that $r_{\perp} \approx R \cos \phi$ is reasonable. With $r \approx R$, the zonal acceleration (15.65) is modified to the form

$$\frac{du}{dt} = v \left[\frac{u \tan \phi}{R} + f \right] \quad \text{where} \quad f = 2 \Omega \sin \phi. \quad (15.70)$$

That is, we dropped the vertical velocity component, w , from the general form of the acceleration (15.65). Correspondingly, the meridional momentum equation takes the form

$$\frac{dv}{dt} = -u \left[\frac{u \tan \phi}{R} + f \right]. \quad (15.71)$$

These approximate forms for the zonal and meridional accelerations appear in the primitive equations of geophysical fluid dynamics presented in Section 27.1.

15.9 More facets of motion on the rotating sphere

In this section we expose a few more points to keep in mind regarding motion on a rotating sphere. These points build from the material in Section 15.8, but here also examine the meridional and vertical equations of motion, (14.31) and (14.32).

15.9.1 Zonal motion is not inertial motion

Consider a particle moving zonally with $u \neq 0$. In the absence of any forces, constant zonal particle motion can presumably continue freely without acceleration, $\dot{u} = 0$ (see equation (15.59)). Indeed, this is the case for motion on a non-rotating plane. However, this presumption is false for motion on a rotating sphere where rotation and curvature both mean that zonal motion is not inertial motion. This point is also seen from our earlier discussions in Chapter 13 of particle kinematics on a rotating sphere.

More precisely, zonal motion leads to nonzero meridional and vertical Coriolis accelerations as well as metric accelerations. We see such accelerations by examining the meridional and vertical velocity equations (14.31) and (14.32)

$$\dot{v} + \frac{vw + u^2 \tan \phi}{r} + 2\Omega u \sin \phi = 0 \quad (15.72)$$

$$\dot{w} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi = -g. \quad (15.73)$$

So the presence of $u \neq 0$, even with $v = w = 0$, introduces accelerations to the meridional and vertical velocity equations. These motions then induce a nonzero zonal acceleration.

As noted above, the two underlying reasons for these accelerations are (i) rotation and (ii) curvature of the sphere. Namely, a constant zonal velocity on the sphere, $\mathbf{v} = u \hat{\lambda}$, is not inertial. In this case, the absolute velocity (velocity seen by an inertial reference frame fixed relative to the stars) is given by

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x} = u \hat{\lambda} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (15.74)$$

whose time derivative is nonzero

$$\frac{d\mathbf{v}_a}{dt} = \dot{u} \hat{\lambda} + u \frac{d\hat{\lambda}}{dt} + \boldsymbol{\Omega} \wedge \mathbf{v}. \quad (15.75)$$

So even if \dot{u} is initially zero, the zonal particle motion is associated with nonzero accelerations due to changes in the zonal unit vector (“metric acceleration”) and due to rotation (Coriolis acceleration).

15.9.2 Coriolis acceleration in the meridional equation

What happens when zonal motion is perturbed in a meridional direction? Following the angular momentum constraint (15.58), a poleward perturbation, $\delta\phi > 0$, to a northern hemisphere particle will have $\dot{u} > 0$, with this zonal acceleration corresponding to a rightward deflection relative to the $\delta\phi > 0$ perturbation. Likewise, for $\delta\phi < 0$ a northern hemisphere particle will have $\dot{u} < 0$, which corresponds to a rightward deflection relative to the $\delta\phi < 0$ perturbation.

We can see the correspondence to the Coriolis acceleration by examining the meridional momentum equation (14.31)

$$\dot{v} + \frac{vw + u^2 \tan \phi}{r} = -\underline{2\Omega u \sin \phi}, \quad (15.76)$$

with the Coriolis acceleration underlined on the right hand side. In the northern hemisphere ($\sin \phi > 0$), the Coriolis acceleration gives rise to a rightward (equatorward) acceleration, $-2\Omega u \sin \phi < 0$, when the particle is moving eastward, $u > 0$ (super-rotating), thus inducing a negative meridional acceleration, $\dot{v} < 0$. Conversely, if the particle is moving to the west so that $u < 0$, then the Coriolis acceleration is again to the right, only this time it deflects the particle in the poleward direction with $\dot{v} > 0$. The analogous considerations hold in the southern hemisphere where the particle is deflected to the left by the Coriolis acceleration. These motions are reflected in Figure 15.3.

15.9.3 Comments

To orient oneself according to the Coriolis acceleration, one needs to remember one fact: in the northern hemisphere motion on the sphere at constant radius ($\delta r = 0$) leads to a rightward deflection whereas in the southern hemisphere the particle experiences a leftward deflection. Figure 15.3 depicts some sample deflections.

15.10 Exercise

EXERCISE 15.1: WESTWARD MOTION FASTER THAN PLANETARY MOTION

Reconsider the discussion in Section 15.8.3, only now allow for the particle to move zonally westward at a speed greater than the planetary rotation

$$\dot{\lambda} + \Omega < 0 \iff u + \Omega r_{\perp} < 0. \quad (15.77)$$

- (A) Discuss what happens to the particle as it is deflected poleward while conserving axial angular momentum.
- (B) Is fluid particle motion with $u + \Omega r_{\perp} < 0$ relevant for the terrestrial atmosphere and ocean? Why? To help answer this question, what is Ωr_{\perp} for $\phi = \pi/4$ and $r = R_e$?

Part III

Fluid kinematics

Kinematics is concerned with the intrinsic properties of motion whereas dynamics is concerned with causes of motion. Consider Newton's Law of motion, $\mathbf{F} = m\mathbf{a}$. Kinematics deduces the forces by knowing the acceleration whereas dynamics deduces the acceleration by knowing the forces. In this part of the book we develop the kinematics of classical fluid mechanics, with all signal speeds non-relativistic (far slower than the speed of light). Furthermore, the fluid kinematic results considered in this part of the book are applicable to both rotating and non-rotating reference frames.

A fundamental element of fluid kinematics concerns the choice of reference frame for describing motion. The Eulerian and Lagrangian reference frames provide dual kinematic descriptions of fluid motion. The Eulerian frame is fixed relative to the laboratory whereas the Lagrangian frame follows a fluid particle. The Eulerian description is concerned with the fluid velocity as observed from a point fixed in space whereas the Lagrangian description is concerned with the velocity of a fluid particle. Having two descriptions of the same motion provides a synergy that is missing with either alone, thus exemplifying an axiom for the practitioner of theoretical physics: *if you can solve a problem more than one way, then do so!* It is thus very useful to have skills at moving between the two descriptions, with tools from mathematical transformation theory of Part I of great use. Developing skills and tools related to Eulerian and Lagrangian kinematics forms the focus for Chapter 16.

The conservation of mass plays a central role in physics. For fluids, mass conservation constrains the fluid flow independently of forces acting on the fluid. Hence, mass conservation is included as part of fluid kinematics rather than dynamics. Mass conservation, and its expression as volume conservation for incompressible flows, are the subjects of Chapters 17, 18 and 19. Chapter 2 provides a resume of the continuum hypothesis that forms the basis for ignoring molecular degrees of freedom in our treatment of fluid motion. Chapter 20 explores the Lagrangian kinematics of material lines, areas, and volumes. Finally, Chapter 21 explores the kinematics of fluid motion when described with the generalized vertical coordinates introduced in Chapter 11.

16

Kinematics of fluid motion

Fluid motion is very complex. Hence, it is essential to avail oneself with more than one means to describe the motion. For this purpose we make use of both the Eulerian and Lagrangian reference frames, where the Eulerian frame is fixed in the laboratory and the Lagrangian frame moves with a fluid particle. These dual descriptions form the foundation for fluid kinematics.

Eulerian (laboratory) and Lagrangian (material) descriptions can be found in any textbook on fluid mechanics, though the Lagrangian approach is typically given less attention and sometimes ignored altogether. However, a Lagrangian description offers useful insights into the theoretical foundations of the subject as it is the natural frame to formulate dynamical laws. Indeed, the Lagrangian description is the natural extension to continuous media of the point particle mechanics discussed in Part II of this book. [Salmon \(1998\)](#) provides an elegant and accessible treatment of Lagrangian fluid mechanics, and [Bennett \(2006\)](#) provides a theoretical treatment along with many applications. Chapter 4 of [Aris \(1962\)](#) offers a lucid treatment in the context of tensor analysis. Much of our treatment here follows Chapters 1 and 2 of the ocean fluid mechanics book of [Olbers et al. \(2012\)](#).

READER'S GUIDE TO THIS CHAPTER

To keep the discussion focused, we assume that spatial positions and fluid particle trajectories are represented using Cartesian coordinates. Even so, we require elements of the tensor transformation theory from Part I to transform between the Eulerian and Lagrangian descriptions. We review the salient formalism in this chapter to help keep the discussion reasonably self-contained. This is a relatively long chapter that introduces many concepts and tools used in nearly every subsequent chapter of this book.

16.1	A variety of fluid systems	204
16.1.1	Defining the fluid systems	204
16.1.2	Comments	206
16.2	Lagrangian and Eulerian reference frames	206
16.3	Material and position coordinates	207
16.3.1	Trajectories of fluid particles	208
16.3.2	Material coordinates	208
16.3.3	Transforming between material and position coordinates	209
16.3.4	Transformation matrix and its Jacobian	209
16.3.5	A discrete algorithm to compute the transformation matrix	210
16.3.6	Jacobian as the ratio of volumes	211
16.3.7	Further study	212
16.4	Lagrangian and Eulerian time derivatives	212
16.4.1	Infinitesimal space-time increment of a function	212
16.4.2	Total time derivative of a function	213
16.4.3	Eulerian: evolution measured in the laboratory frame	213
16.4.4	Lagrangian: evolution measured in the material frame	213
16.4.5	Lagrangian time derivative formulated from the material frame	214
16.4.6	Sample material time derivative operations	215
16.4.7	Summarizing some terminology for equations	216
16.5	Galilean invariance	217
16.5.1	Galilean transformation	217
16.5.2	Transformation of the material time derivative	217
16.6	Invariance of the material time derivative	218
16.6.1	Invariance based on definition of the material time derivative	219
16.6.2	Invariance for a rotating reference frame	219
16.6.3	Comments	221
16.7	Fluid flow lines	221
16.7.1	Material pathlines from fluid particle trajectories	221
16.7.2	Fluid streamlines and streamtubes	222
16.7.3	Distinguishing streamlines from pathlines for unsteady flow	223
16.7.4	Fluid streaklines	223
16.7.5	An analytic example of flow lines	224
16.7.6	Further study	226
16.8	Exercises	227

16.1 A variety of fluid systems

The continuum hypothesis (Chapter 2) allows us to consider fluid flow from a field theoretic perspective, whereby physical properties are described by fields that take on values at each point of a space and time continuum. As part of this continuum description, we make use of conceptual systems for framing the kinematic and dynamic description of the motion. These systems are used throughout this book.

16.1.1 Defining the fluid systems

Fluid particle

A *fluid particle* is a non-extensive massless point that follows the fluid flow, with the flow specified by the fluid's velocity field (left panel in Figure 16.1). A fluid particle is uniquely specified by its

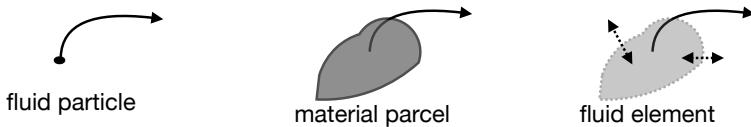


Figure 16.1: Schematic for the fluid systems used in our considerations of fluid kinematics. Left panel: fluid particle that tests the fluid flow field, \mathbf{v} , without altering it. Fluid particles have no mass and no extension. Their trajectories define the Lagrangian reference frame. Middle panel: a material fluid parcel comprised of a fixed material content and thus a fixed mass. Fluid parcels are deformable and tiny regions of the fluid that move with the fluid velocity. Right panel: fluid element, which is a generalization of the fluid parcel and is comprised of a fixed mass but with matter exchanged across its boundary. The fluid element moves with the barycentric velocity (see Section 18.1), which is the center of mass velocity. Finite extensions of the material parcel are referred to as *material fluid regions*, and finite extensions of fluid elements are referred to as *fluid regions*.

material coordinate and time (we discuss material coordinates in Section 16.3.1).

A fluid particle does not refer to a molecule. Rather, a fluid particle is a mathematical construct afforded to us by the continuum hypothesis. The fluid particle has zero mass and zero extent, so that it does not affect the flow. Rather, the sole purpose of a fluid particle is to sample the fluid flow at an arbitrary point in the fluid continuum, and fluid particle trajectories define the Lagrangian reference frame (Section 16.2). A fluid particle is thus directly analogous to *test particles* in Newtonian gravitation that are used to determine gravitational field lines, and test charges in electromagnetism used for the same purpose.

Material fluid parcel

A *material fluid parcel* is an infinitesimal deformable fluid region that follows the fluid flow as specified by the velocity field (middle panel in Figure 16.1). A material parcel maintains a fixed mass, a fixed matter content, and a fixed specific entropy. Hence, it does not exchange matter, heat, or entropy with other fluid parcels. Furthermore, it does not experience irreversible exchange of momentum arising from friction. Its only interaction with adjacent fluid parcels is through reversible mechanical exchanges from pressure.

Just like a fluid particle, a material fluid parcel is uniquely specified by its material coordinate and time. However, a material parcel is not a point. Rather, it has an infinitesimal volume, $\delta V > 0$, that deforms with the flow. Its mass is written $\delta M = \rho \delta V$, with ρ the mass density. Think of a material fluid parcel as a tiny perfectly insulated slippery bag full of gas or liquid. Even so, we never have reason in this book to specify the boundary of a fluid parcel. Rather, we only make use of the conceptual framework provided by fluid parcels as part of a Lagrangian formulation of fluid kinematics and dynamics.

Finite sized material fluid region

A material region is a finite volume generalization of a material fluid parcel. Conversely, a material fluid parcel is the infinitesimal limit of a material fluid region. That is, a material region is comprised of fixed mass and fixed matter content. Hence, as the material region moves through the fluid there is zero exchange of matter across its boundary. In contrast to the material fluid parcel, we are generally concerned with details of the boundary to material fluid regions.

Fluid element

A fluid element is an infinitesimal and deformable fluid region of fixed mass yet non-fixed matter and non-fixed specific entropy (right panel in Figure 16.1). For a homogeneous fluid comprised

of a single matter constituent and no irreversible processes, then a fluid element is identical to a material fluid parcel. However, there is a distinction for non-homogeneous fluids, such as the ocean (e.g., fresh water, salts, biogeochemical tracers) and the atmosphere (e.g., air, water, dust, chemical species). The exchange of matter across the boundary of a fluid element arises from the irreversible diffusive mixing of trace constituents within the fluid (Sections 18.1 and 37.2). As detailed in Section 18.1, diffusive matter exchange leaves the mass of the fluid element unchanged since the fluid element velocity is determined by its center of mass. As for a material fluid parcel, we have no reason in this book to specify the boundary of a fluid element.

Fluid region

A fluid region is the most general subsystem within a fluid, whereby we consider an arbitrary finite region whose boundaries are open to the exchange of matter, mechanical forces, and thermodynamic properties with the surrounding environment. Here, we specify the boundary of the region and provide details for transport of properties across that boundary.

16.1.2 Comments

The fluid particle's sole purpose is to determine trajectories and the associated fluid pathlines. The material fluid parcel has the added feature of nonzero volume and an associated kinematic description leading to the continuity equation (Section 17.1). Mechanics of material fluid parcels are the focus of perfect fluid mechanics, where the fluid matter is comprised of a single homogeneous constituent and there are no irreversible processes.

Following page 3 of [Olbers et al. \(2012\)](#), we introduced the fluid element as the next most general infinitesimal fluid system beyond a material fluid parcel. Fluid elements allow us to consider real fluids with more than one matter constituent, with such multi-component fluids generally allowing for matter and other properties to be irreversibly exchanged between the elements. Much of the kinematics in this chapter holds for both material fluid parcels and fluid elements. We thus refer to “fluid parcels” for brevity. In Chapter 17 and elsewhere, we generally make the distinction when studying the kinematics of multi-constituent fluids.

Notably, we can consider a fluid particle for both homogeneous and multi-component fluids. Again, the fluid particle tracks the trajectories of fluid as defined by the velocity field. For the multi-component fluid the velocity field is the barycentric velocity (see Section 18.1), which is the center of mass velocity. So fluid particles in a multi-component fluid track the barycentric velocity.

Many authors do not distinguish between material fluid parcels and fluid elements. For our purposes, we make the distinction since the perfect fluid mechanics of material parcels is sufficient for only a limited number of fluid systems considered in this book. Fluid elements are required to formulate the kinematics and dynamics of multi-component fluids, in which irreversible exchanges occur between fluid elements thus leading to the diffusion of matter, heat, momentum, and other properties. Such exchanges are fundamental features of geophysical fluid mechanics.

16.2 Lagrangian and Eulerian reference frames

There are two reference frames commonly used as the basis for describing motion of a fluid continuum.

- **LAGRANGIAN OR MATERIAL REFERENCE FRAME:** This reference frame is defined by that of moving material fluid particles. A mechanical description in this reference frame aims to

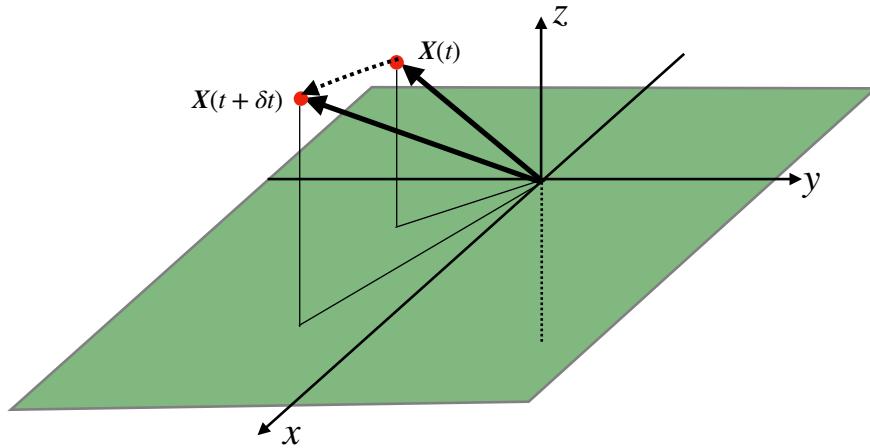


Figure 16.2: Sample trajectory of a fluid particle with endpoints $\mathbf{X}(t)$ and $\mathbf{X}(t + \delta t)$. The trajectory passes through the point $\mathbf{x} = \mathbf{X}(t)$ at time t and $\mathbf{x} + \delta\mathbf{x} = \mathbf{X}(t + \delta t)$ at time $t + \delta t$. Eulerian kinematics describes the fluid flow from the perspective of an observer fixed with respect to the laboratory frame. Lagrangian kinematics describes the fluid flow from the perspective of an observer moving in the frame of fluid particles.

determine the trajectory for each fluid particle. The material approach is commonly termed *Lagrangian*.

- **EULERIAN OR LABORATORY REFERENCE FRAME:** The second reference frame is based on observing the fluid from a fixed spatial position, \mathbf{x} , within a “laboratory”. This *Eulerian* approach measures fluid properties as the fluid streams by a fixed observer. It is not concerned with determining fluid particle trajectories. Instead, the focus of Eulerian fluid mechanics is on fluid properties determined as a function of position \mathbf{x} and time t .

Notably, the “laboratory” used to observe the fluid may itself be moving, such as when on a rotating planet. The present discussion is not concerned with transforming from a non-inertial laboratory frame to an inertial laboratory frame (see Chapter 13).

The Eulerian and Lagrangian approaches complement one another. For example, the Lagrangian approach lends itself to fruitful physical insights since we can borrow freely from the point particle mechanics of Part II. In contrast, the Eulerian approach is often more straightforward when developing numerical methods for simulations and it is typically simpler when making laboratory or field measurements. In general, we make use of both Eulerian and Lagrangian kinematics. A goal of this chapter is to provide the foundation for these two perspectives and to develop tools for transforming from one to the other.

16.3 Material and position coordinates

A material description is suggested by the Lagrangian reference frame, whereby fluid particles are labeled with a material coordinate. This description complements the Eulerian, whereby each point in space is labeled by its position. The trajectory of a fluid particle provides the mathematical transformation between the material coordinates and position coordinates. We explore this transformation in this section.

16.3.1 Trajectories of fluid particles

In describing the motion of a classical point particle (Chapter 13), we specify its spatial position according to a time dependent position vector \mathbf{X} that is a function of time, t . At a given time t the position vector is located at a space point denoted by the Cartesian position \mathbf{x} , in which case we write

$$\mathbf{x} = \mathbf{X}(t) \quad \text{point particle.} \quad (16.1)$$

A sample trajectory is shown in Figure 16.2. We emphasize the notation convention used here, which may seem pedantic but in later discussions proves essential. Namely, the time dependent position of a particle is denoted with the capital $\mathbf{X}(t)$, whose instantaneous space position is denoted by the lowercase \mathbf{x} . This convention aims to distinguish functions, such as $\mathbf{X}(t)$, from the evaluated value of these functions, \mathbf{x} .

When there are N discrete particles, we distinguish the various particle trajectories by introducing a discrete label

$$\mathbf{x} = \mathbf{X}(n, t) \quad n = 1, N \quad \text{point particles.} \quad (16.2)$$

When the matter is a continuum, such as for a rigid body, an elastic solid, or a fluid, then the discrete label becomes a continuous vector

$$\mathbf{x} = \mathbf{X}(\mathbf{a}, t) \quad \text{continuum of matter.} \quad (16.3)$$

The vector \mathbf{a} is referred to as the *material* coordinate. It labels a point of matter within the continuum fluid.

16.3.2 Material coordinates

In classical fluid mechanics, we ignore special relativistic effects of fluid particles. Consequently, both the material reference frame and the laboratory reference frame measure the same universal time, t . In contrast, the spatial coordinates are distinct for the Eulerian and Lagrangian references frames. Again, the spatial coordinates for the Eulerian frame are given by the position relative to a fixed laboratory frame, whereas the three components of a material Lagrangian coordinate remains unchanged as the fluid particle moves. Additionally, the three coordinates for both the Eulerian and Lagrangian description must be linearly independent to allow for a unique specification of the fluid particle.

One common choice for material coordinate is to define it as the spatial position of a fluid particle at an arbitrary time

$$\mathbf{a} = \mathbf{X}(t = t_0). \quad (16.4)$$

As a slightly more refined example, consider a perfect fluid (single material component with no irreversible processes). For this fluid, the specific entropy of each fluid parcel remains fixed at its initial value. When the fluid is placed in a gravitational field, layers of constant specific entropy are generally found to be monotonically stacked, or *stratified*, in the vertical direction (Chapter 25). As a result, we can uniquely specify a fluid parcel by giving its horizontal coordinate position, (x, y) , as well as the specific entropy. The material coordinates for a parcel can thus be written as

$$\mathbf{a} = (X, Y, \theta)_{t=t_0}, \quad (16.5)$$

where we write θ as a measure of the specific entropy (or potential temperature; see Section 23.7). As indicated by this example, the physical dimensions of material coordinates can generally differ. It is this generality that necessitates the use of general tensor methods when developing the mechanical equations using Lagrangian coordinates. The mathematics and physics of these *generalized vertical coordinates* are detailed in the Tier-II Chapters 11, 21, and 32.

16.3.3 Transforming between material and position coordinates

Motion of a fluid continuum is described by a *point transformation*. A point transformation is a mathematical way of saying that the fluid motion causes a fluid particle labeled by \mathbf{a} to continuously move from an initial position $\mathbf{x}_0 = \mathbf{X}(\mathbf{a}, t_0)$ to another position $\mathbf{x} = \mathbf{X}(\mathbf{a}, t)$ at time $t > 0$. The point transformation is defined by the vector relation

$$\mathbf{x} = \mathbf{X}(\mathbf{a}, t), \quad (16.6)$$

which is written in component form as¹

$$x^m = X^m(a^i, t). \quad (16.7)$$

In fluid mechanics, the relation (16.6) defines a single-valued and invertible transformation from material coordinates (\mathbf{a}, t) to position coordinates (\mathbf{x}, t) . That is, for each material coordinate \mathbf{a} and time t , there is a unique spatial point \mathbf{x} , with this point specified by the trajectory $\mathbf{X}(\mathbf{a}, t)$. Conversely, for each space-time point (\mathbf{x}, t) there exists a unique material coordinate found by inverting equation (16.6)²

$$\mathbf{a} = \mathbf{A}(\mathbf{x}, t). \quad (16.8)$$

In this equation, \mathbf{A} is the inverse function that specifies the material coordinate \mathbf{a} given (\mathbf{x}, t) . The single-valued property means that a fluid particle trajectory does not split, nor do two trajectories occupy the same point at the same time. This assumption is fundamental to the continuum hypothesis and the associated use of fluid particles to map out pathlines of the fluid flow.

16.3.4 Transformation matrix and its Jacobian

In our analysis of fluid motions, we make routine use of the position coordinates of an Eulerian description and material coordinates of a Lagrangian description. We here introduce the tensor analysis framework providing a means to transform from one description to the other. Fundamental to that framework is the transformation matrix.

In Section 10.2.1, we encountered the transformation matrix between Cartesian and spherical coordinates. We will later consider a transformation matrix for moving between different references frames in Sections 16.5 and 16.6. Here, we develop the transformation matrix between coordinates in position-space, termed \mathbf{x} -space, and coordinates in material-space, termed \mathbf{a} -space.

The continuum of trajectories, $\mathbf{X}(\mathbf{a}, t)$, is fundamental to this transformation. Namely, the trajectories as given by equation (16.6) provide a transformation between position coordinates (\mathbf{x}, t) and material coordinates (\mathbf{a}, t) . The transformation is invertible so long as the Jacobian of the transformation matrix remains nonzero. Note that our formulation makes use of Cartesian tensors. However, we can make use of the general tensor formalism detailed in Part I of this book to extend the results to arbitrary coordinate for either \mathbf{x} -space and/or \mathbf{a} -space.

The transformation matrix is given by the matrix of partial derivatives, and we choose to organize this matrix according to the following convention

$$F_i^m = \frac{\partial X^m}{\partial a^i} \equiv \begin{bmatrix} \partial X^1 / \partial a^1 & \partial X^1 / \partial a^2 & \partial X^1 / \partial a^3 \\ \partial X^2 / \partial a^1 & \partial X^2 / \partial a^2 & \partial X^2 / \partial a^3 \\ \partial X^3 / \partial a^1 & \partial X^3 / \partial a^2 & \partial X^3 / \partial a^3 \end{bmatrix}. \quad (16.9)$$

¹We choose tensor labels m, n, p for spatial coordinates and trajectories, and i, j, k for material coordinates.

²The use of \mathbf{A} for the inverse function in equation (16.8) should not be confused with the acceleration, also written as \mathbf{A} elsewhere in this book. We will not have much use for equation (16.8), thus minimizing the opportunity for confusion.

As defined by equation (16.9), the upper label, m , denotes the row and the lower label, i , is the column. As seen in Section 20.2, the transformation matrix F_i^m is also known as the *deformation tensor*, as it provides a means to measure how trajectories are deformed by the flow. The Jacobian of the transformation matrix is the determinant

$$\det(F_i^m) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \det \begin{bmatrix} \partial X^1 / \partial a^1 & \partial X^1 / \partial a^2 & \partial X^1 / \partial a^3 \\ \partial X^2 / \partial a^1 & \partial X^2 / \partial a^2 & \partial X^2 / \partial a^3 \\ \partial X^3 / \partial a^1 & \partial X^3 / \partial a^2 & \partial X^3 / \partial a^3 \end{bmatrix}. \quad (16.10)$$

We make use of the notation $\partial \mathbf{X} / \partial \mathbf{a}$ for the Jacobian as it offers a useful means to distinguish between the transformation and its inverse.

16.3.5 A discrete algorithm to compute the transformation matrix

To help further our understanding of the transformation matrix (16.9), it is useful to develop a computational algorithm for its discrete approximation. We illustrate an algorithm for two-dimensional flow and write the trajectory using Cartesian coordinates

$$\mathbf{X}(t) = X^1(t) \hat{\mathbf{x}} + X^2(t) \hat{\mathbf{y}}, \quad (16.11)$$

and use a Cartesian representation for the material coordinate

$$\mathbf{a} = a^1 \hat{\mathbf{x}} + a^2 \hat{\mathbf{y}}. \quad (16.12)$$

Now lay down a two-dimensional lattice with discrete indices (e, f) for each of the nodal points (grid points) on the lattice, and with corresponding spatial coordinates

$$\mathbf{x}(e, f) = x(e, f) \hat{\mathbf{x}} + y(e, f) \hat{\mathbf{y}}. \quad (16.13)$$

Initialize fluid particles at each of the lattice grid points,

$$\mathbf{X}(e, f; t=0) = \mathbf{x}(e, f) = \mathbf{a}(e, f), \quad (16.14)$$

with the discrete material coordinates defined by the initial positions. Then time step the trajectories using the velocity field to compute the particle pathlines $\mathbf{X}(e, f; t)$ as illustrated in Figure 16.3. At any particular time, the position of a fluid particle is found by interpolating from the lattice grid points. Setting the material coordinates equal to the initial position then leads to the finite difference approximation to the transformation matrix

$$F_i^m = \begin{bmatrix} F_1^1 & F_2^1 \\ F_1^2 & F_2^2 \end{bmatrix} \approx \begin{bmatrix} \frac{X^1(e+1,f;t)-X^1(e-1,f;t)}{X^1(e+1,f;0)-X^1(e-1,f;0)} & \frac{X^1(e,f+1;t)-X^1(e,f-1;t)}{X^2(e,f+1;0)-X^2(e,f-1;0)} \\ \frac{X^2(e+1,f;t)-X^2(e-1,f;t)}{X^1(e+1,f;0)-X^1(e-1,f;0)} & \frac{X^2(e,f+1;t)-X^2(e,f-1;t)}{X^2(e,f+1;0)-X^2(e,f-1;0)} \end{bmatrix}. \quad (16.15)$$

If the grid is regular in both directions, then the initial positions have a separation given by the grid spacing in which case

$$F_i^m \approx \begin{bmatrix} \frac{X^1(e+1,f;t)-X^1(e-1,f;t)}{\Delta} & \frac{X^1(e,f+1;t)-X^1(e,f-1;t)}{\Delta} \\ \frac{X^2(e+1,f;t)-X^2(e-1,f;t)}{\Delta} & \frac{X^2(e,f+1;t)-X^2(e,f-1;t)}{\Delta} \end{bmatrix}. \quad (16.16)$$

This algorithm illustrates how the transformation matrix provides a measure of trajectory spreading as fluid particles move away from their initial positions.

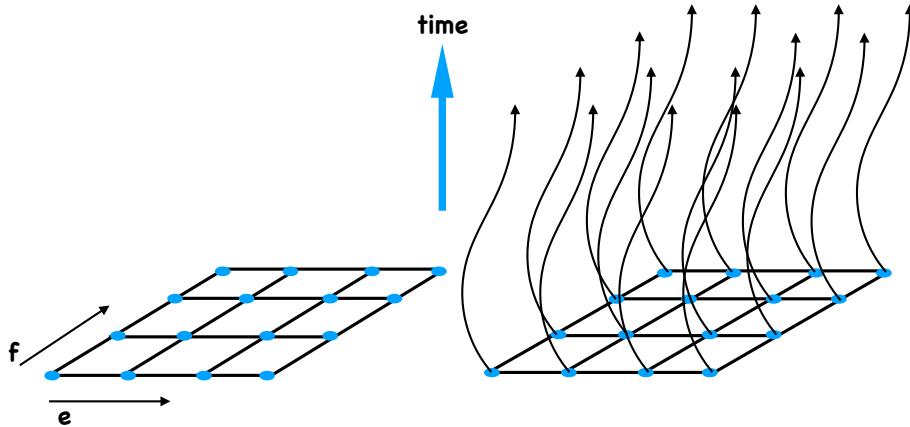


Figure 16.3: Illustrating the computational algorithm of Section 16.3.5 used to approximate the transformation matrix (also known as the deformation tensor) F_i^m . The left panel shows the two-dimensional grid with nodal points defining the initial positions for fluid particles. Each position on the grid is labeled by a unique integer (e, f) . The initial position of each particle is taken as the material coordinate, with the discrete label (e, f) maintained by the particles as they evolve. The right panel shows the pathlines for the fluid particles after time $t > 0$. When working on a discrete grid, the position of the fluid particles is found by interpolating from the node points.

16.3.6 Jacobian as the ratio of volumes

We here establish the Jacobian as the ratio of volume elements written in position space and material space. This property holds at each point within the continuum fluid, and thus holds on fluid parcels.

Volume of an infinitesimal region of space within the fluid continuum

Consider the volume of an infinitesimal region of space. For simplicity, write this volume in terms of Cartesian coordinates

$$dV(\mathbf{x}) = dx dy dz. \quad (16.17)$$

The \mathbf{x} argument is introduced on the left hand side to distinguish this volume, which has dimensions L^3 , from the volume written in material coordinates

$$dV(\mathbf{a}) = da db dc, \quad (16.18)$$

where $\mathbf{a} = (a^1, a^2, a^3) = (a, b, c)$ are coordinates in material space. Note that $dV(\mathbf{a})$ does not generally have dimension L^3 , since the dimension for each component of the material coordinates is not necessarily length (e.g., see the example in equation (16.5)).

The two volumes are related by the Jacobian of transformation between the two coordinate systems

$$dV(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} dV(\mathbf{a}). \quad (16.19)$$

This relation indicates that the Jacobian measures the ratio of the volume written in terms of position coordinates to the volume written in terms of material coordinates

$$\frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{dV(\mathbf{x})}{dV(\mathbf{a})}. \quad (16.20)$$

This is a key result of great value for transforming between Eulerian and Lagrangian coordinates.

Volume of an infinitesimal fluid element

The above results have been formulated for an arbitrary region of the fluid continuum. Hence, the expressions also hold when evaluated on an arbitrary material fluid parcel or fluid element. We use the notation δV for the parcel/element volume, in terms of which the above relations take the form

$$\delta V(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \delta V(\mathbf{a}) \Rightarrow \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{\delta V(\mathbf{x})}{\delta V(\mathbf{a})}. \quad (16.21)$$

Hence, when evaluated on a moving material parcel, the Jacobian measures the ratio of the parcel volume written in terms of position coordinates to the parcel volume written in terms of material coordinates. In the particular case where the material coordinates are the initial fluid particle positions, then the Jacobian measures the ratio of the instantaneous volume of a fluid element to its initial volume

$$\frac{\partial \mathbf{X}}{\partial \mathbf{X}(0)} = \frac{\delta V(\mathbf{x})}{\delta V(0)}. \quad (16.22)$$

16.3.7 Further study

This classic video from the *National Committee for Fluid Mechanics Films*, featuring Prof. John Lumley, offers insightful visualizations to help understand Eulerian and Lagrangian fluid descriptions.

16.4 Lagrangian and Eulerian time derivatives

As noted in Section 16.3.2, we assume non-relativistic motion so that the material reference frame and the laboratory reference frame both measure the same universal time, t . However, when computing time derivatives, the laboratory frame does so by fixing the space point \mathbf{x} , whereas the material frame computes time derivatives by fixing the material coordinate, \mathbf{a} . These two time derivatives generally measure distinct changes in the fluid. Relating their changes constitutes a key result of fluid kinematics.

16.4.1 Infinitesimal space-time increment of a function

Consider a fluid property as represented by a space-time dependent field, Φ . For example, Φ could be the temperature, mass density, or velocity of the fluid. When measured at a fixed point in space this fluid property is written mathematically as

$$\Phi = \Phi(\mathbf{x}, t). \quad (16.23)$$

The difference between $\Phi(\mathbf{x}, t)$ and $\Phi(\mathbf{x} + d\mathbf{x}, t + dt)$ delivers the differential increment, computed to leading order via a Taylor series expansion

$$d\Phi = \Phi(\mathbf{x} + d\mathbf{x}, t + dt) - \Phi(\mathbf{x}, t) \quad (16.24a)$$

$$= dt \frac{\partial \Phi}{\partial t} + d\mathbf{x} \cdot \nabla \Phi. \quad (16.24b)$$

In this equation, dt is the infinitesimal time increment, and $d\mathbf{x}$ is the vector of infinitesimal space increments. For example, making use of Cartesian coordinates leads to the increment

$$d\mathbf{x} = \hat{\mathbf{x}} dx + \hat{\mathbf{y}} dy + \hat{\mathbf{z}} dz. \quad (16.25)$$

We ignore higher order terms in equation (16.24b) since the space and time increments are infinitesimal.

16.4.2 Total time derivative of a function

In fluid mechanics, it is common to sample properties of the fluid from moving reference frames. In this case, the sampling position is a function of time. Consequently, the total time derivative of Φ is determined by dividing both sides of equation (16.24b) by the infinitesimal time increment

$$\frac{d\Phi}{dt} = \frac{\partial\Phi}{\partial t} + \frac{dx}{dt} \cdot \nabla\Phi. \quad (16.26)$$

The first term measures the time derivative of Φ at the point \mathbf{x} . The second term accounts for changes in Φ arising from movement of the reference frame relative to a point \mathbf{x} according to the velocity $d\mathbf{x}/dt$. Expression (16.26) holds in general since the velocity of the moving frame is arbitrary. We next specialize to the two common cases in fluid mechanics.

16.4.3 Eulerian: evolution measured in the laboratory frame

The Eulerian time derivative considers the evolution of a fluid property when sampled at a fixed space point

$$\text{Eulerian time derivative} = \frac{\partial\Phi(\mathbf{x}, t)}{\partial t}. \quad (16.27)$$

This result follows from specializing the total time derivative in equation (16.26) to the case of fixed spatial points, so that $d\mathbf{x}/dt = 0$. In the geophysical fluids literature, the Eulerian time derivative is often termed the *time tendency* and flows with a nonzero time tendency are said to be *developing flows*. When the Eulerian time derivative vanishes everywhere the flow is said to be in a *steady state* or in a *steady flow* condition, with all points in the laboratory frame measuring a zero time change for fluid properties.

16.4.4 Lagrangian: evolution measured in the material frame

The Lagrangian or material time derivative measures the evolution of a fluid property sampled along the trajectory of a moving fluid particle. The Lagrangian time derivative for a field is thus written

$$\text{Lagrangian time derivative} = \frac{D\Phi}{Dt} = \frac{\partial\Phi}{\partial t} + \mathbf{v} \cdot \nabla\Phi. \quad (16.28)$$

The second equality follows by setting $d\mathbf{x}/dt = \mathbf{v}$ in equation (16.26) since we are sampling points along the fluid particle trajectory $\mathbf{x} = \mathbf{X}(\mathbf{a}, t)$. The operator $\partial/\partial t$ is the Eulerian time derivative from equation (16.27), whereas $\mathbf{v} \cdot \nabla$ is referred to as the *advection* operator. Use of the capital D for the material time operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (16.29)$$

signals that the time derivative is computed along a fluid particle trajectory. This notation distinguishes the material time derivative from the more generic total time derivative of equation (16.26).

Equation (16.29) provides an Eulerian expression (right hand side) to the material time derivative D/Dt . There are two Eulerian contributions: the local (fixed space point) time tendency $\partial/\partial t$ and advection, $\mathbf{v} \cdot \nabla$. Advection arises in the Eulerian reference frame due to the fluid passing by the fixed laboratory observer, whereas it is absent from the material reference frame since it moves with the fluid. Figure 16.4 illustrates the differences between the Eulerian and Lagrangian perspectives.

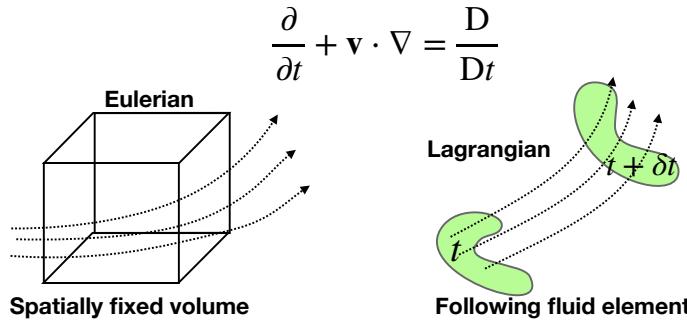


Figure 16.4: Illustrating the distinctions between the Eulerian (laboratory) and Lagrangian (material) reference frames for describing fluid motion. For the Eulerian description we consider a fixed control volume in the laboratory frame and observe fluid moving through the volume. For the Lagrangian description we tag fluid particles and observe their motion through the fluid. The Eulerian representation of the material time derivative has two terms, one due to time changes local to the fixed laboratory point, and one due to the advection of properties that are swept by the local position.

Notably, a steady flow (zero Eulerian time derivative for all flow properties) does not imply a vanishing Lagrangian time derivative for flow properties. Rather, steady flow is a statement about the Eulerian (laboratory) nature of the fluid motion. It does not necessarily mean that properties of fluid particles are constant, since there can be a nonzero Lagrangian evolution (via advection) even when the Eulerian time tendency vanishes.

16.4.5 Lagrangian time derivative formulated from the material frame

Rather than start from the total time derivative (16.26), it is instructive to work in the moving material frame *a priori*. For this purpose, we measure the function Φ on a fluid particle, in which case it is sometimes useful to introduce the shorthand notation

$$\Phi[\mathbf{X}(\mathbf{a}, t), t] \equiv \Phi^L(\mathbf{a}, t) \quad \Leftarrow \text{sampling } \Phi \text{ on a trajectory } \mathbf{x} = \mathbf{X}(\mathbf{a}, t) \text{ at time } t. \quad (16.30)$$

In words, $\Phi^L(\mathbf{a}, t)$ is the function Φ evaluated on a fluid particle trajectory $\mathbf{x} = \mathbf{X}(\mathbf{a}, t)$. That is, Φ^L is the Lagrangian version of the function. For example, the Lagrangian velocity is given by

$$\mathbf{v}^L(\mathbf{a}, t) \equiv \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] = \left[\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \right]_{\mathbf{a}}. \quad (16.31)$$

The \mathbf{a} subscript emphasizes that the partial time derivative is computed while holding \mathbf{a} fixed.

The time derivative along a fluid particle trajectory is the material derivative. We introduce finite differences along the trajectory to estimate the material time derivative

$$\left[\frac{\partial \Phi[\mathbf{X}(\mathbf{a}, t), t]}{\partial t} \right]_{\mathbf{a}} = \lim_{\Delta t \rightarrow 0} \left[\frac{\Phi[\mathbf{X}(\mathbf{a}, t + \Delta t/2), t + \Delta t/2] - \Phi[\mathbf{X}(\mathbf{a}, t - \Delta t/2), t - \Delta t/2]}{\Delta t} \right]. \quad (16.32)$$

Expanding the numerator in a Taylor series, and keeping just the leading terms, yields

$$\left[\frac{\partial \Phi[\mathbf{X}(\mathbf{a}, t), t]}{\partial t} \right]_{\mathbf{a}} = \lim_{\Delta t \rightarrow 0} \left[\frac{\Phi[\mathbf{X}(\mathbf{a}, t + \Delta t/2), t + \Delta t/2] - \Phi[\mathbf{X}(\mathbf{a}, t - \Delta t/2), t - \Delta t/2]}{\Delta t} \right] \quad (16.33a)$$

$$= \left[\left(\frac{\partial}{\partial t} \right)_{\mathbf{X}} + \left(\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \right)_{\mathbf{a}} \cdot \nabla \right] \Phi[\mathbf{X}(\mathbf{a}, t), t] \quad (16.33b)$$

$$= \left[\left(\frac{\partial}{\partial t} \right)_{\mathbf{X}} + \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \cdot \nabla \right] \Phi[\mathbf{X}(\mathbf{a}, t), t]. \quad (16.33c)$$

We included a subscript on the derivative operators on the right hand side to be explicit about what variables are held fixed during differentiation. This extra notation can generally be dropped, since a partial derivative operation is based on holding all variables fixed except for the variable being differentiated. Evaluating the trajectory at the spatial point $\mathbf{X}(\mathbf{a}, t) = \mathbf{x}$ allows us to dispense with the trajectory notation to recover the more succinct expression (16.28). Even so, it is important to keep in mind the underlying trajectory basis for the material time derivative.

16.4.6 Sample material time derivative operations

The material time derivative operator is perhaps the most important operator in fluid mechanics, and its relation to the Eulerian time derivative plus advection is a key result of fluid kinematics. Therefore, it is critical to develop experience with this operator and its generalizations. The examples here offer a starting point.

Linear wave characteristics

Consider the vector function $\Phi(\mathbf{x}, t) = \mathbf{x} - \mathbf{c}t$, where \mathbf{c} is a constant velocity vector. In the study of linear waves, lines of constant Φ represent surfaces on which the wave phase remains constant. In the theory of partial differential equations, these lines form characteristics for the linear advection equation (Section 5.1). Now consider the material time derivative

$$\frac{D(\mathbf{x} - \mathbf{c}t)}{Dt} = \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] (\mathbf{x} - \mathbf{c}t) \quad (16.34a)$$

$$= (\mathbf{v} \cdot \nabla) \mathbf{x} - \frac{\partial(t \mathbf{c})}{\partial t} \quad (16.34b)$$

$$= \mathbf{v}(\mathbf{x}, t) - \mathbf{c}. \quad (16.34c)$$

The first term says that the velocity of a point following a fluid particle is none other than the fluid velocity. This result, which follows by definition, serves as a useful means to verify internal consistency of the formalism. The second term arises from the constant velocity, \mathbf{c} . If the velocity of a fluid parcel, \mathbf{v} , moves with the wave velocity, \mathbf{c} , then surfaces of constant Φ are material. However, such constant phase surfaces are generally not material since fluid particles typically do not move with the wave phase velocity.

Material invariant

Consider a scalar function that remains constant on a material trajectory so that its material time derivative vanishes

$$\frac{D\Phi}{Dt} = 0. \quad (16.35)$$

Material constancy is generally referred to as *material invariance* in this book. We may also say that the property Φ is materially conserved. At a fixed point in space, a materially invariant property has its Eulerian time derivative arising only via advection

$$\frac{\partial\Phi}{\partial t} = -\mathbf{v} \cdot \nabla\Phi. \quad (16.36)$$

Recall the case of a wave characteristic whereby $\partial\Phi/\partial t = \mathbf{c}$. Wave characteristics are material if the fluid particle velocity matches the wave phase velocity, $\mathbf{v} = \mathbf{c}$. However, most waves do not have a phase speed that corresponds to fluid particle motion so that wave characteristics are generally not material lines.

We garner geometric insight into relation (16.36) by introducing the unit normal vector to the surface of constant Φ

$$\hat{\mathbf{n}} = \frac{\nabla\Phi}{|\nabla\Phi|}. \quad (16.37)$$

Material constancy of Φ thus means that the normalized Eulerian time tendency equals to the negative of the projection of the velocity into the direction normal to constant Φ surfaces

$$\frac{\partial\Phi/\partial t}{|\nabla\Phi|} = -\mathbf{v} \cdot \hat{\mathbf{n}}. \quad (16.38)$$

That is, the fluid particle velocity, \mathbf{v} , is matched precisely to the velocity of the moving surface of constant Φ . As detailed in Section 17.4.2, this result means there are no fluid parcels that cross surfaces of constant Φ . That is, constant Φ surfaces are material.

Time derivative measured in a general moving frame

Now consider a reference frame moving at an arbitrary velocity $\mathbf{v}^{(s)}$. Examples include the quasi-Lagrangian reference frames of a float in the ocean or balloon in the atmosphere. Due to their finite size and associated drag effects, these objects only approximate material particle motion, so that $\mathbf{v}^{(s)} \neq \mathbf{v}$. Returning to the general expression (16.26) for the total time derivative, we have the time derivative operator as measured in this non-material moving reference frame

$$\frac{D^{(s)}}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla. \quad (16.39)$$

A function that remains constant within this general moving frame thus satisfies

$$\frac{D^{(s)}\Phi}{Dt} = 0 \Rightarrow \frac{\partial\Phi}{\partial t} = -\mathbf{v}^{(s)} \cdot \nabla\Phi. \quad (16.40)$$

Introducing the normal direction $\hat{\mathbf{n}} = |\nabla\Phi|^{-1} \nabla\Phi$ leads to

$$\frac{\partial\Phi/\partial t}{|\nabla\Phi|} = -\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}, \quad (16.41)$$

which is an analog to the material constancy condition (16.38).

16.4.7 Summarizing some terminology for equations

We here summarize certain terminology used in this book in reference to the variety of equations in geophysical fluid mechanics. Some of this terminology was introduced in this chapter, whereas others will be encountered in subsequent development.

- PROGNOSTIC: This is an equation that determines the time tendency (Eulerian evolution) of a quantity such as the temperature or velocity.
- DIAGNOSTIC: This is an equation that determines the value of a field at a particular instance in time. An example is the non-divergence condition placed on the fluid velocity in an incompressible fluid (see Section 19.1). There are no time derivatives appearing in diagnostic equations.
- STEADY STATE: All Eulerian time derivatives vanish, so that all fluid properties are time independent when measured in the laboratory frame.
- MATERIAL INVARIANT: The Lagrangian time derivative vanishes for a material invariant or for a material constant.

16.5 Galilean invariance

Galilean invariance means that the laws of motion are the same in all inertial reference frames. Furthermore, two inertial reference frames can only be moving with a constant velocity relative to one another. We discussed Galilean invariance for a point particle in Section 13.4. Here we consider its expression for a fluid. As for the particle, Galilean invariance means that the material acceleration of a fluid particle remains the same when viewed in an arbitrary inertial reference frame. Some care is required when translating this invariance into a mathematical statement when decomposing the material acceleration into its Eulerian components. Our considerations here provide a useful warmup to the more general discussion in Section 16.6, where we transform space and time derivative operators between an inertial frame and a rotating frame.

16.5.1 Galilean transformation

A Galilean transformation is given by the linear transformation

$$\bar{t} = t \quad (16.42)$$

$$\bar{\mathbf{x}} = \mathbf{x} + \mathbf{U}t \quad (16.43)$$

$$\bar{\mathbf{v}} = \mathbf{v} + \mathbf{U}. \quad (16.44)$$

By convention, we say that the barred coordinates are those measured in the moving reference frame and the unbarred are measured in the rest frame. However, since both reference frames are inertial, there is no experiment that can determine which frame is at rest or which is moving. Instead, what is relevant for our considerations is that the two inertial frames are moving relative to one another. Furthermore, note that time remains unchanged (non-relativistic physics), whereas the position of a point in the new frame equals to that in the original reference frame plus a contribution from the constant velocity \mathbf{U} . The inverse transformation is trivially given by

$$t = \bar{t} \quad (16.45)$$

$$\mathbf{x} = \bar{\mathbf{x}} - \mathbf{U}\bar{t} \quad (16.46)$$

$$\mathbf{v} = \bar{\mathbf{v}} - \mathbf{U}. \quad (16.47)$$

16.5.2 Transformation of the material time derivative

We find it useful practice to make use of the transformation matrix formalism used in Section 10.2.1 to transform from Cartesian to spherical coordinates. Additionally, it is sufficient to work in the 1+1 dimensional case with time and one space dimension. Writing $(t, x) = (x^0, x^1)$ and $(\bar{t}, \bar{x}) = (\bar{x}^0, \bar{x}^1)$ renders the transformation of derivatives

$$\frac{\partial}{\partial x^{\bar{\alpha}}} = \frac{\partial x^{\alpha}}{\partial \bar{x}^{\bar{\alpha}}} \frac{\partial}{\partial x^{\alpha}}, \quad (16.48)$$

where $\alpha = 0, 1$ is a tensor index that incorporates the 0 value for the time axis. The transformation matrix for the Galilean transformation is

$$\frac{\partial x^{\bar{\alpha}}}{\partial x^{\alpha}} = \begin{bmatrix} \partial \bar{x}^0 / \partial x^0 & \partial \bar{x}^0 / \partial x^1 \\ \partial \bar{x}^1 / \partial x^0 & \partial \bar{x}^1 / \partial x^1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ U & 1 \end{bmatrix}, \quad (16.49)$$

and the inverse is

$$\frac{\partial x^{\alpha}}{\partial \bar{x}^{\bar{\alpha}}} = \begin{bmatrix} 1 & 0 \\ -U & 1 \end{bmatrix}. \quad (16.50)$$

Consequently, the Eulerian time derivative as measured in the moving frame is given by

$$\frac{\partial}{\partial x^0} = \frac{\partial x^0}{\partial \bar{x}^0} \frac{\partial}{\partial \bar{x}^0} + \frac{\partial x^1}{\partial \bar{x}^0} \frac{\partial}{\partial \bar{x}^1} \quad (16.51)$$

$$= \frac{\partial}{\partial \bar{x}^0} - U \frac{\partial}{\partial \bar{x}^1} \quad (16.52)$$

$$= \frac{\partial}{\partial t} - U \frac{\partial}{\partial x}. \quad (16.53)$$

In words, this identity says that the time derivative computed between two inertial reference frames differs due to an advective term arising from the relative motion of the two inertial observers. The space derivatives are related by

$$\frac{\partial}{\partial x^1} = \frac{\partial x^0}{\partial \bar{x}^1} \frac{\partial}{\partial \bar{x}^0} + \frac{\partial x^1}{\partial \bar{x}^1} \frac{\partial}{\partial \bar{x}^1} \quad (16.54)$$

$$= \frac{\partial}{\partial \bar{x}^1}, \quad (16.55)$$

so that the space derivative operator remains form invariant under a Galilean transformation. This result holds also for the other two space dimensions. The material time derivative operator is therefore form invariant under a Galilean transformation

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (16.56a)$$

$$= \frac{\partial}{\partial \bar{t}} + \mathbf{U} \cdot \bar{\nabla} + (\bar{\mathbf{v}} - \mathbf{U}) \cdot \bar{\nabla} \quad (16.56b)$$

$$= \frac{\partial}{\partial \bar{t}} + \bar{\mathbf{v}} \cdot \bar{\nabla} \quad (16.56c)$$

$$= \frac{D}{D\bar{t}}, \quad (16.56d)$$

where we used the shorthand

$$\bar{\mathbf{v}} \cdot \bar{\nabla} = \bar{u} \frac{\partial}{\partial \bar{x}^1} + \bar{v} \frac{\partial}{\partial \bar{x}^2} + \bar{w} \frac{\partial}{\partial \bar{x}^3}. \quad (16.57)$$

Even though the material time derivative operator is form invariant under an arbitrary Galilean transformation, its individual pieces are separately modified.

16.6 Invariance of the material time derivative

In our discussion of Galilean invariance in Section 16.5, we showed that the material time derivative operator remains form invariant under changes to the inertial reference frame. Consequently, the acceleration of a fluid particle is identical when measured in any inertial reference frame. We here consider the more general case of non-inertial reference frames that differ by both rotations and translations. We already know that the form for fluid particle accelerations differs between an inertial frame and a non-inertial frame. Nonetheless, we show here that the material time derivative operator remains form invariant. This is a result of great practical relevance, as it means that the scalar equations of fluid mechanics (e.g., mass continuity and tracer equations) remain form-invariant when changing reference frames.

16.6.1 Invariance based on definition of the material time derivative

The material time derivative measures time changes of a fluid property in the reference frame of a moving fluid particle. The Lagrangian reference frame follows fluid particles, so it is the natural reference frame for measuring material time changes. In contrast, the Eulerian reference frame is fixed in a laboratory. The material time derivative computed from the laboratory reference frame is composed of an Eulerian time tendency plus an advection operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (16.58)$$

Importantly, this expression holds regardless the choice of laboratory reference frames, either inertial or non-inertial. Our choice of laboratory frames only impacts on the form of the Eulerian time derivative and on the advection operator. The sum of the two terms returns the same material time derivative operator, no matter what laboratory frame is chosen.

Again for emphasis, the reason for the form invariance is that the material time derivative operator is, by definition, computed in the fluid particle reference frame. The particle reference frame is unconcerned with the subjective choice made by the observer in the laboratory reference frame. In the following, we exhibit how the mathematics respects this invariance. Namely, we show how the Eulerian expression for the material time derivative remains form invariant when changing reference frames.

16.6.2 Invariance for a rotating reference frame

Consider two laboratory frames. The first is at rest and so serves as an inertial frame, whereas the second is rotating with rotational axis aligned with the vertical direction as in Figure 13.4. Introduce Cartesian coordinates for the inertial frame, with corresponding basis vectors $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. Let these inertial frame unit vectors be related to rotating frame unit vectors according to

$$\hat{\mathbf{x}} = \hat{\mathbf{x}} \cos \theta - \hat{\mathbf{y}} \sin \theta \quad (16.59a)$$

$$\hat{\mathbf{y}} = \hat{\mathbf{x}} \sin \theta + \hat{\mathbf{y}} \cos \theta \quad (16.59b)$$

$$\hat{\mathbf{z}} = \hat{\mathbf{z}}, \quad (16.59c)$$

and let time be the same in the two reference frames. The angle θ measures the counter-clockwise angle between the $\hat{\mathbf{x}}$ axis and the moving $\hat{\mathbf{x}}$, with this angle a linear function of time

$$\theta = \Omega t. \quad (16.60)$$

The above relations between the two sets of basis vectors translates into the same relations between the corresponding coordinate representations for an arbitrary vector. Including time, we have the relation between inertial coordinates (the barred frame) and rotating coordinates (unbarred frame)

$$\bar{t} = t \quad (16.61a)$$

$$\bar{x} = x \cos \theta - y \sin \theta \quad (16.61b)$$

$$\bar{y} = x \sin \theta + y \cos \theta \quad (16.61c)$$

$$\bar{z} = z. \quad (16.61d)$$

The inverse transformation can be easily found

$$t = \bar{t} \quad (16.62a)$$

$$x = \bar{x} \cos \theta + \bar{y} \sin \theta \quad (16.62b)$$

$$y = -\bar{x} \sin \theta + \bar{y} \cos \theta \quad (16.62c)$$

$$z = \bar{z}. \quad (16.62d)$$

We are now prepared to make use of the transformation formalism used for the Galilean transformation in Section 16.5, as well as in Section 10.3 to transform from Cartesian to cylindrical-polar coordinates. Here, we include time as part of the formalism by introducing the Greek label $\alpha = 0, 1, 2, 3$ so that the transformation matrix between the inertial frame and rotating frame is given by

$$\frac{\partial x^{\bar{\alpha}}}{\partial x^{\alpha}} = \begin{bmatrix} \partial x^{\bar{0}}/\partial x^0 & \partial x^{\bar{0}}/\partial x^1 & \partial x^{\bar{0}}/\partial x^2 & \partial x^{\bar{0}}/\partial x^3 \\ \partial x^{\bar{1}}/\partial x^0 & \partial x^{\bar{1}}/\partial x^1 & \partial x^{\bar{1}}/\partial x^2 & \partial x^{\bar{1}}/\partial x^3 \\ \partial x^{\bar{2}}/\partial x^0 & \partial x^{\bar{2}}/\partial x^1 & \partial x^{\bar{2}}/\partial x^2 & \partial x^{\bar{2}}/\partial x^3 \\ \partial x^{\bar{3}}/\partial x^0 & \partial x^{\bar{3}}/\partial x^1 & \partial x^{\bar{3}}/\partial x^2 & \partial x^{\bar{3}}/\partial x^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\Omega \bar{y} & \cos \theta & -\sin \theta & 0 \\ \Omega \bar{x} & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (16.63)$$

Similarly, the inverse transformation is given by

$$\frac{\partial x^{\alpha}}{\partial x^{\bar{\alpha}}} = \begin{bmatrix} \partial x^0/\partial x^{\bar{0}} & \partial x^0/\partial x^{\bar{1}} & \partial x^0/\partial x^{\bar{2}} & \partial x^0/\partial x^{\bar{3}} \\ \partial x^1/\partial x^{\bar{0}} & \partial x^1/\partial x^{\bar{1}} & \partial x^1/\partial x^{\bar{2}} & \partial x^1/\partial x^{\bar{3}} \\ \partial x^2/\partial x^{\bar{0}} & \partial x^2/\partial x^{\bar{1}} & \partial x^2/\partial x^{\bar{2}} & \partial x^2/\partial x^{\bar{3}} \\ \partial x^3/\partial x^{\bar{0}} & \partial x^3/\partial x^{\bar{1}} & \partial x^3/\partial x^{\bar{2}} & \partial x^3/\partial x^{\bar{3}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \Omega y & \cos \theta & \sin \theta & 0 \\ -\Omega x & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (16.64)$$

The derivative operators transform according to

$$\frac{\partial}{\partial x^{\alpha}} = \frac{\partial x^{\bar{\alpha}}}{\partial x^{\alpha}} \frac{\partial}{\partial x^{\bar{\alpha}}}, \quad (16.65)$$

in which case

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \bar{t}} + (\boldsymbol{\Omega} \wedge \bar{\boldsymbol{x}}) \cdot \bar{\nabla} \quad (16.66a)$$

$$\frac{\partial}{\partial x} = \cos \theta \frac{\partial}{\partial \bar{x}} + \sin \theta \frac{\partial}{\partial \bar{y}} \quad (16.66b)$$

$$\frac{\partial}{\partial y} = -\sin \theta \frac{\partial}{\partial \bar{x}} + \cos \theta \frac{\partial}{\partial \bar{y}} \quad (16.66c)$$

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial \bar{z}}. \quad (16.66d)$$

The velocity vector components transform according to

$$v^{\alpha} = \frac{\partial x^{\alpha}}{\partial x^{\bar{\alpha}}} v^{\bar{\alpha}}, \quad (16.67)$$

so that

$$v^0 = v^{\bar{0}} \quad (16.68a)$$

$$u = \Omega y + \bar{u} \cos \theta + \bar{v} \sin \theta \quad (16.68b)$$

$$v = -\Omega x - \bar{u} \sin \theta + \bar{v} \cos \theta \quad (16.68c)$$

$$w = \bar{w}. \quad (16.68d)$$

Bringing these result together leads to the transformation of the horizontal advection operator

$$u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} = (\bar{\mathbf{u}} - \boldsymbol{\Omega} \wedge \bar{\mathbf{x}}) \cdot \bar{\nabla}. \quad (16.69)$$

Combining this result with the transformed Eulerian time derivative leads to the material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (16.70a)$$

$$= \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \quad (16.70b)$$

$$= \frac{\partial}{\partial t} + (\boldsymbol{\Omega} \wedge \bar{\mathbf{x}}) \cdot \bar{\nabla} + (\bar{\mathbf{u}} - \boldsymbol{\Omega} \wedge \bar{\mathbf{x}}) \cdot \bar{\nabla} + \bar{w} \frac{\partial}{\partial z} \quad (16.70c)$$

$$= \frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \bar{\nabla}. \quad (16.70d)$$

As advertised, the operator is form invariant under time dependent transformations to a non-inertial reference frame.

16.6.3 Comments

As argued at the start of this section, the invariance of the material time derivative to changes in the laboratory reference frame is rather obvious: why would a time derivative computed in a material frame be concerned with the nature of the laboratory frame? Even so, it is satisfying to see the tools of coordinate transformations put to use verifying this result. It is this sort of exercise that nurtures trust in tensor analysis, at which point it becomes a reliable tool for exploration where the answer is not *a priori* known.

16.7 Fluid flow lines

There are three types of flow lines commonly used to visualize fluid motion: pathlines, streamlines, and streaklines. These flow lines are identical for time independent (steady) flow, but they differ for unsteady flow. They each offer complementary information about the flow field, and have uses in both theoretical and experimental contexts. We have use mostly for pathlines and streamlines, though also introduce streaklines for completeness.

16.7.1 Material pathlines from fluid particle trajectories

As introduced in Section 16.3.1, a fluid particle traces out a *trajectory* as it moves through space (Figure 16.2). We use the term material *pathline* for a fluid particle trajectory, with a collection of pathlines providing a means to visualize fluid particle motion throughout the flow.

Mathematically, a fluid particle trajectory is a curve $\mathbf{x} = \mathbf{X}(\mathbf{a}, t)$ in space that is traced by fixing the material coordinate, \mathbf{a} , and letting time advance. Trajectories are computed by time integrating the ordinary differential equation

$$\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (16.71a)$$

$$\mathbf{X}(\mathbf{a}, t = 0) = \mathbf{a}, \quad (16.71b)$$

where

$$\mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] = \mathbf{v}^L(\mathbf{a}, t) \quad (16.72)$$

is Lagrangian velocity of the fluid particle (see equation (16.31)), and we have assumed the material coordinates are determined by the initial position. Again, the partial time derivative is computed with the material coordinate held fixed, so that the material coordinate distinguishes between particle trajectories.

In the laboratory, we can insert tiny trace particles (e.g., dust, colored dye) into the fluid to offer a means for flow visualization. A time exposed photograph of the trace particles provides a visualization of fluid pathlines. Trace particles provide an increasingly accurate estimate of fluid particle pathlines if the trace particles do not disperse through diffusion (see Chapter 37). Another example offers further experience with pathliness, where here we consider cars moving at night. A time exposed photograph reveals pathlines for the cars as formed by their lights. Like cars, the material pathlines in a fluid can intersect, cross, and become quite complex, particularly when the flow is turbulent.

16.7.2 Fluid streamlines and streamtubes

Streamlines are curves whose tangent is parallel to the instantaneous fluid velocity field. Streamlines can intersect only at a stagnation point; i.e., a point where the fluid is not moving. Let

$$d\mathbf{x} = \hat{\mathbf{x}} dx + \hat{\mathbf{y}} dy + \hat{\mathbf{z}} dz \quad (16.73)$$

be an infinitesimal increment along a streamline written using Cartesian coordinates. The family of streamlines at a given time t satisfy the tangent constraint

$$\mathbf{v} \wedge d\mathbf{x} = 0, \quad (16.74)$$

which is equivalent to

$$\frac{dx}{u(\mathbf{x}, t)} = \frac{dy}{v(\mathbf{x}, t)} = \frac{dz}{w(\mathbf{x}, t)}. \quad (16.75)$$

Alternatively, we can introduce a pseudo-time parameter, s , that determines a position along a streamline. Streamlines are the curves $\mathbf{x} = \mathbf{X}(s; \mathbf{a}, t)$ computed with (\mathbf{a}, t) held fixed, but with the pseudo-time varied

$$\frac{\partial \mathbf{X}(s; \mathbf{a}, t)}{\partial s} = \mathbf{v}[\mathbf{X}(s; \mathbf{a}, t), t] \quad (16.76a)$$

$$\mathbf{X}(s = 0; \mathbf{a}, t) = \mathbf{a}. \quad (16.76b)$$

Again, both the material coordinate \mathbf{a} and time t are held fixed when determining streamlines, so that (\mathbf{a}, t) act as parameters to distinguish streamlines. Streamlines thus do not know about the time evolution of unsteady flow. Instead, streamlines only sample a snapshot of the velocity field; they are freshly computed at each time instance.

A streamtube is a bundle of streamlines crossing through an arbitrary closed curve (see Figure 16.5). Hence, at each time instance, streamtube sides are parallel to the velocity vector. Furthermore, when the flow is steady then streamlines are identical to material pathlines. A streamtube is therefore a material tube for steady flow, in which case no fluid particles cross the streamtube boundary.

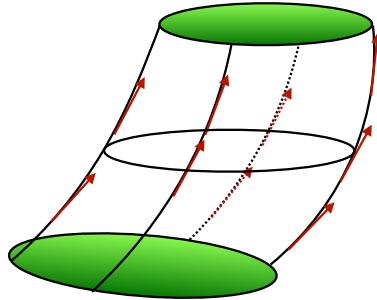


Figure 16.5: This image shows an example streamtube. The side boundaries of a streamtube consist of streamlines. At each point of a streamline, the local tangent vector equals to the velocity field (see equation (16.74)). Streamlines are identical to pathlines only for steady flow; they differ for unsteady flows. Hence, for unsteady flows, particle trajectories generally cross through the streamtube boundary.

16.7.3 Distinguishing streamlines from pathlines for unsteady flow

The tangent to a streamline gives the velocity at a single point in time, whereas the tangent to a material pathline (i.e., a trajectory) gives the velocity at subsequent times. These tangents are identical when the flow is steady. However, if the flow is time dependent (unsteady), then streamlines differ from material pathlines. Furthermore, for unsteady flow, the pseudo-time parameter, s , determining the streamlines in equation (16.76a) is not equal to the time, t , used to compute fluid particle trajectories in equation (16.71a). Consequently, the condition $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$ satisfied at each instance by a streamline does not ensure that fluid particles do not cross streamlines. The reason is that a material line moves with the fluid in such a way that

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \mathbf{v}^{\text{line}} \cdot \hat{\mathbf{n}} \quad \text{material lines,} \quad (16.77)$$

where \mathbf{v}^{line} is the velocity of a point on the material line. The material line thus moves so that no fluid particles cross it. Only when the flow is steady, so that $\mathbf{v}^{\text{line}} \cdot \hat{\mathbf{n}} = 0$, will material lines and streamlines be equal. That is, the streamline constraint $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$ is not a material constraint when $\mathbf{v}^{\text{line}} \cdot \hat{\mathbf{n}} \neq 0$. The key point is that streamlines do not probe the time behaviour of the flow, so they do not know whether the velocity is steady or unsteady.

The distinction between streamlines and pathlines for unsteady flow is sometimes not appreciated in the literature. For example, on top of page 77 of [Williams and Elder \(1989\)](#), they state that since the flow is instantaneously tangent to a streamtube boundary, there can be no flow across the streamtube. [Kundu et al. \(2012\)](#) make a similar statement on their page 84. However, as emphasized above, streamlines equal to material pathlines only for steady flow in which the velocity has a zero time tendency.

16.7.4 Fluid streaklines

A streakline is a curve obtained by connecting the positions for all fluid particles that emanate from a fixed point in space (see Figure 16.6). Streaklines are simple to define conceptually and to realize experimentally. However, they are a bit convoluted to specify mathematically. We thus present two formulations.

At any time t , the streakline passing through a fixed point \mathbf{y} is a curve going from \mathbf{y} to $\mathbf{X}(\mathbf{y}, t)$, the position reached by the particle initialized at $t = 0$ at the point \mathbf{y} . A particle is on the streakline if it passed the fixed point \mathbf{y} at some time between 0 and t . If this time was s , then the material coordinate of the particle would be given by $\mathbf{a}(\mathbf{y}, s)$ (see equation (16.8) relating the

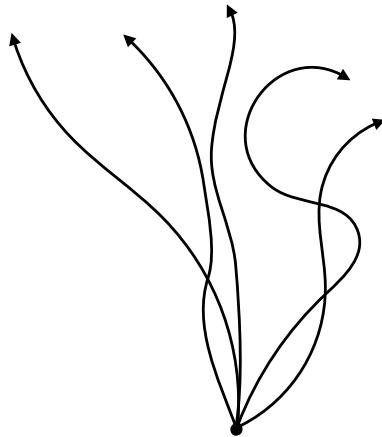


Figure 16.6: A suite of trajectories emanating from a single point. Common realizations include the paths of fluid particles that leave from a chimney, or the smoke from a point source. A streakline is defined as the accumulation of positions at time t of particles that passed through the common point at some earlier time $s < t$.

material coordinate to its corresponding laboratory position). Furthermore, at time t , this particle is at \mathbf{x} , so that the equation of the streakline at time t is

$$\mathbf{x} = \mathbf{X}[\mathbf{a}(\mathbf{y}, s), t] \quad 0 \leq s \leq t. \quad (16.78)$$

We can connect the streakline specification to that given for a pathline and streamline through the following. A streakline at some time instance \tilde{t} is a curve defined by fixing \tilde{t} and varying s over $s \leq \tilde{t}$ in the function $\mathbf{X}(s; \mathbf{a}, \tilde{t})$. We determine the curves $\mathbf{x} = \mathbf{X}(s; \mathbf{a}, \tilde{t})$ by solving the following set of initial value problems for trajectories with initial conditions imposed at $t = s$ rather than $t = 0$

$$\frac{\partial \mathbf{X}(s; \mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(s; \mathbf{a}, t), t] \quad (16.79a)$$

$$\mathbf{X}(t = s; \mathbf{a}, t) = \mathbf{a}. \quad (16.79b)$$

Note that \mathbf{a} remains fixed, as we start all trajectories determining a streakline from the same initial point (e.g., the chimney does not move). A streakline can thus be generated by emitting a dye from a point over a time interval equal to the range of s , with the dye following fluid particle trajectories.

16.7.5 An analytic example of flow lines

Consider the following two-dimensional example as taken from Section 4.13 of [Aris \(1962\)](#). Let the Eulerian velocity field be given by

$$u = \frac{x}{\tau + t} \quad (16.80a)$$

$$v = \frac{y}{\tau} \quad (16.80b)$$

$$w = 0, \quad (16.80c)$$

where $\tau > 0$ is a constant with the dimensions of time.

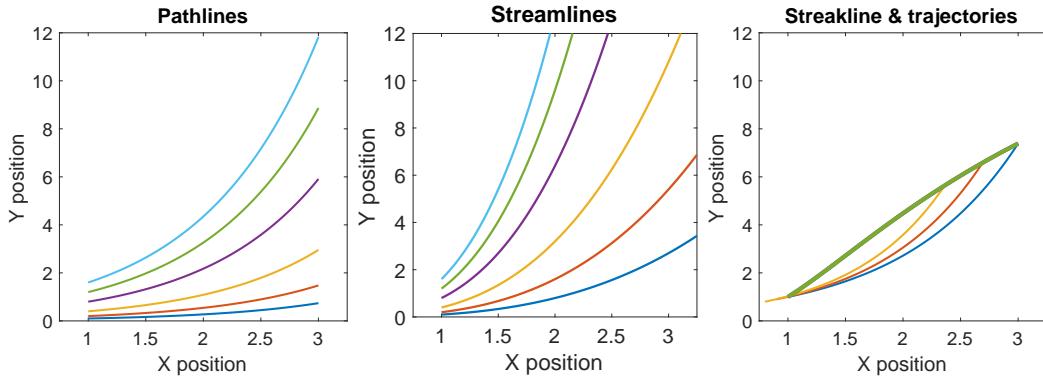


Figure 16.7: Left panel: sample pathlines $X(t) = X_0(1+t/\tau)$ and $Y(t) = Y_0 e^{t/\tau}$ (see equations (16.82a) and (16.82b)) during times $t \in [0, 2]$. The trajectories drawn here all start at $X_0 = 1$ and set the parameter $\tau = 1$. Note that those pathlines with $X_0 = 0$ remain on the y-axis, and those with $Y_0 = 0$ remain on the x-axis. Middle panel: Sample streamlines $X(s; t) = X_0 e^{s/(\tau+t)}$ and $Y(s; t) = Y_0 e^{s/\tau}$ (see equations (16.85a) and (16.85b)). We set $t = 2$ and let the pseudo-time run from $s \in [0, 4]$. All streamlines shown here start at $X_0 = 1$. Note that those that start with $X_0 = 0$ remain on the y-axis, and those that start with $Y_0 = 0$ remain on the x-axis. Right panel: sample analytic streakline (dark bold line) at $t = 2$ according to equations (16.88a) and (16.88b). This streakline is determined by the position of particles at $t = 2$ that pass through $(X, Y) = (1, 1)$ during times $t \in (-\infty, 2)$. We show three sample trajectories that fall onto the streakline. The longest trajectory starts at $(X, Y) = (1, 1)$ at $t = 0$, whereas the two shorter trajectories pass through $(X, Y) = (1, 1)$ at some time $0 < t < 2$. Notice the distinction between all three flow lines, which is to be expected since the flow field is unsteady.

Pathlines

Pathlines are determined by solving the trajectory equations

$$\frac{dX(t)}{dt} = \frac{X(t)}{\tau + t} \quad (16.81a)$$

$$\frac{dY(t)}{dt} = \frac{Y(t)}{\tau} \quad (16.81b)$$

$$\frac{dZ(t)}{dt} = 0, \quad (16.81c)$$

which are found to be

$$X(t) = X_0(1 + t/\tau) \quad (16.82a)$$

$$Y(t) = Y_0 e^{t/\tau} \quad (16.82b)$$

$$Z(t) = Z_0, \quad (16.82c)$$

where $\mathbf{X}(t = 0) = \mathbf{X}_0$. Sample trajectories are shown in Figure 16.7 over time $t \in [0, 2]$. We can eliminate time to yield a curve in the horizontal (x, y) plane

$$y = Y_0 e^{(x-X_0)/X_0}. \quad (16.83)$$

Streamlines

Streamlines are determined by solving the differential equations

$$\frac{dX(s; t)}{ds} = \frac{X(s; t)}{\tau + t} \quad (16.84a)$$

$$\frac{dY(s; t)}{ds} = \frac{Y(s; t)}{\tau} \quad (16.84b)$$

$$\frac{dZ(s; t)}{ds} = 0, \quad (16.84c)$$

where time, t , is a fixed parameter whereas the pseudo-time, s , is varied. Integration renders the streamlines

$$X(s; t) = X_0 e^{s/(\tau+t)} \quad (16.85a)$$

$$Y(s; t) = Y_0 e^{s/\tau} \quad (16.85b)$$

$$Z(s; t) = Z_0. \quad (16.85c)$$

Sample streamlines are shown in Figure 16.7. Note that we can eliminate the pseudo-time s to render a curve in the horizontal (x, y) plane

$$y = Y_0 \left(\frac{x}{X_0} \right)^{(\tau+t)/\tau} \quad (16.86a)$$

$$z = Z_0. \quad (16.86b)$$

Streaklines

For streaklines, invert the trajectory expressions (16.82a)-(16.82b) to find the material coordinates $\mathbf{a}(y, s)$ in the form

$$a_1 = \frac{y_1}{1 + s/\tau} \quad (16.87a)$$

$$a_2 = y_2 e^{-s/\tau} \quad (16.87b)$$

$$a_3 = y_3. \quad (16.87c)$$

We next evaluate the trajectory expressions (16.82a)-(16.82b) with \mathbf{a} as the initial positions to find the streaklines

$$X(s; \mathbf{a}, t) = \frac{y_1 (1 + t/\tau)}{1 + s/\tau} \quad (16.88a)$$

$$Y(s; \mathbf{a}, t) = y_2 e^{(t-s)/\tau} \quad (16.88b)$$

$$Z(s; \mathbf{a}, t) = y_3. \quad (16.88c)$$

Figure 16.7 illustrates the streakline for a particular point $(X, Y) = (1, 1)$.

16.7.6 Further study

A discussion of flow lines can be found in most books on fluid mechanics. The presentation here borrows from Sections 4.11-4.13 of [Aris \(1962\)](#), Section 3.3 of [Kundu et al. \(2012\)](#), and online lecture notes on fluid kinematics from Professor McIntyre of Cambridge University.

16.8 Exercises

EXERCISE 16.1: MATERIAL EVOLUTION OF THE PARTIAL DERIVATIVE OF A FUNCTION

In this exercise we establish some properties of the material time derivative operator when acting on spatial derivatives of a scalar field.

- (a) If a scalar field Φ is materially constant, prove that the material evolution of its spatial derivative is given by

$$\frac{D(\partial_m \Phi)}{Dt} = -\partial_m \mathbf{v} \cdot \nabla \Phi. \quad (16.89)$$

For example, if $D\Phi/Dt = 0$, then the zonal partial derivative $\partial_x \Phi$ has a material time derivative given by

$$\frac{D(\partial \Phi / \partial x)}{Dt} = -\frac{\partial \mathbf{v}}{\partial x} \cdot \nabla \Phi. \quad (16.90)$$

Hint: use Cartesian tensors for convenience.

- (b) What is the material time derivative of $\nabla \Phi$ for the case that Φ is not materially constant?

17

Mass conservation

We assume that mass is neither created nor destroyed anywhere within the fluid domain. This assumption holds independently of the forces acting on the fluid, and so it forms a topic within fluid kinematics. Hence, the only means for changing the mass of the fluid domain is to transport mass across domain boundaries. In this chapter, we derive the mathematical expressions for mass conservation and the associated kinematic constraints placed on the fluid motion. These constraints are examined both in the interior of the fluid as well as at the boundaries. We examine a variety of fluid systems including infinitesimal and finite, moving (Lagrangian) and static (Eulerian). Note that the equation for mass conservation is often referred to as the *continuity equation*, which is a generic name for a local conservation equation.

READER'S GUIDE TO THIS CHAPTER

Spatial positions and trajectories are represented in this chapter using Cartesian coordinates to simplify the maths. Nonetheless, the results hold for general coordinates by making use of general covariance as detailed in Chapters 8 and 9. We presume an understanding of the Eulerian and Lagrangian kinematic descriptions detailed in Chapter 16. Nearly all material in this chapter is used throughout the remainder of the book.

17.1	Material fluid parcels	230
17.1.1	Lagrangian expression for mass conservation	230
17.1.2	Alternative derivation based on the Jacobian	231
17.1.3	Summary of parcel kinematic equations	231
17.2	Eulerian fluid regions	231
17.2.1	Differential expression	232
17.2.2	Finite volume expression	232
17.2.3	Arbitrary Eulerian region	233
17.3	Material fluid regions	234
17.3.1	Evolution of volume	234
17.3.2	Mass conservation	234
17.3.3	Mass conservation using Lagrangian methods	235
17.3.4	Reynolds transport theorem	236
17.3.5	Comments on notation	237
17.4	Kinematic boundary conditions	237
17.4.1	Static material surface	238
17.4.2	Moving material surface	239
17.4.3	Dynamic and permeable surface	242
17.5	Mass budget for a column of ocean fluid	245
17.6	Exercises	246

17.1 Material fluid parcels

We here derive the differential expressions for mass conservation of a material parcel within a Lagragian reference frame. The differential expressions for volume and density arise as a corollary.

17.1.1 Lagrangian expression for mass conservation

The mass of an infinitesimal fluid parcel is written

$$\delta M = \rho \delta V, \quad (17.1)$$

where δV is the parcel volume and

$$\rho = \frac{\delta M}{\delta V} \quad (17.2)$$

is the mass density. By definition, the parcel has a constant mass, so that its material time derivative vanishes

$$\frac{D(\delta M)}{Dt} = 0. \quad (17.3)$$

Equation (17.3) is the most basic form of mass conservation for a material parcel. However, one often has need to express this result in terms of parcel density and parcel volume

$$\frac{D(\delta M)}{Dt} = \frac{D(\rho \delta V)}{Dt} \quad (17.4a)$$

$$= \delta M \left[\frac{1}{\rho} \frac{D\rho}{Dt} + \frac{1}{\delta V} \frac{D(\delta V)}{Dt} \right]. \quad (17.4b)$$

Making use of equation (17.23) derived in Section 17.3.1 for the material evolution of the parcel volume then leads to

$$\frac{1}{\delta M} \frac{D(\delta M)}{Dt} = \frac{1}{\rho} \frac{D\rho}{Dt} + \nabla \cdot \mathbf{v}. \quad (17.5)$$

Setting $D(\delta M)/Dt = 0$ then leads to the mass conservation equation, also known as the continuity equation

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}. \quad (17.6)$$

The parcel volume contracts in regions where the velocity converges (we prove that property in Section 17.3.1). As seen by the mass continuity equation (17.6), such regions are also where the parcel density increases. The opposite occurs for regions where the velocity diverges.

17.1.2 Alternative derivation based on the Jacobian

An alternative approach to deriving the mass conservation equation makes use of the material time evolution of the Jacobian (equation (20.51c)), which then leads to the material evolution for the fluid parcel mass

$$\frac{D}{Dt} [\rho \delta V(\mathbf{x})] = \frac{D}{Dt} \left[\rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \delta V(\mathbf{a}) \right] \quad (17.7a)$$

$$= \left[\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right] \left[\frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right] \delta V(\mathbf{a}) \quad (17.7b)$$

$$= \left[\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right] \delta V(\mathbf{x}). \quad (17.7c)$$

We recover the mass conservation equation (17.6) when noting that the mass of a material parcel is constant.

17.1.3 Summary of parcel kinematic equations

Let us now summarize the variety of differential evolution equations for mass, volume, and density as viewed from a material reference frame

$$\frac{D(\delta M)}{Dt} = 0 \quad \text{parcel mass} \quad (17.8a)$$

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v} \quad \text{parcel volume} \quad (17.8b)$$

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v} \quad \text{parcel density.} \quad (17.8c)$$

Note that the parcel volume equation is derived in Section 17.3.1 below. To help remember the proper signs, recall that as the fluid diverges from a point ($\nabla \cdot \mathbf{v} > 0$), it expands the boundaries of the material parcel and so increases the parcel volume. This process in turn causes the material parcel density to decrease ($-\nabla \cdot \mathbf{v} < 0$).

17.2 Eulerian fluid regions

We now develop expressions for the mass budget within an Eulerian region, both infinitesimal and finite. Recall that Eulerian regions are fixed in space and thus have constant volumes.

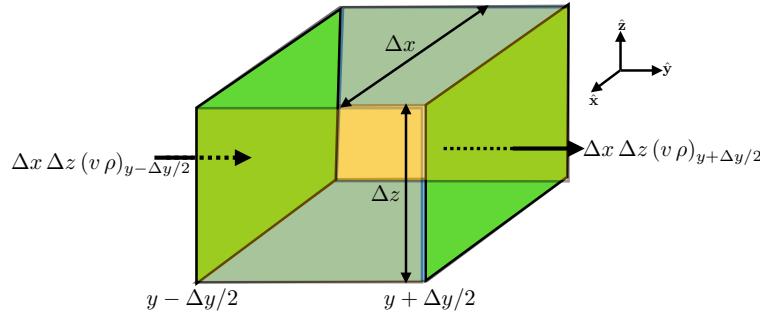


Figure 17.1: A finite sized cube (also known as a cell) of fixed dimensions and position (an Eulerian cube) used to formulate the Eulerian form of mass conservation. We highlight two cell faces with area $\Delta x \Delta z$ and with meridional mass transport crossing the faces given by $\Delta x \Delta z (v \rho)_{y-\Delta y/2}$ and $\Delta x \Delta z (v \rho)_{y+\Delta y/2}$. To establish signs we assume the meridional velocity is positive, $v > 0$, so that mass enters the face at $y - \Delta y/2$ and leaves the face at $y + \Delta y/2$. Differences between these two transports leads to an accumulation of mass within the cell. Note that the resulting mass budget holds regardless the direction of the flow velocity.

17.2.1 Differential expression

The Eulerian form of mass continuity results from introducing the Eulerian expression for the material time derivative operator (equation (16.28)) into the mass continuity equation (17.6). The resulting *flux-form* Eulerian mass continuity equation is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (17.9)$$

This equation is in the form of a flux-form conservation law, in which the local time tendency of a field is determined by the convergence of a flux

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}). \quad (17.10)$$

The flux

$$\text{mass flux} = \rho \mathbf{v} \quad \text{dimensions } M \text{ } L^{-2} \text{ } T^{-1} \quad (17.11)$$

measures the mass per time of matter crossing a unit area oriented with an outward normal in each of the three Cartesian directions. If more mass flux comes into a point than leaves, then the density increases (mass converges), and vice versa for a mass flux that diverges from a point.

17.2.2 Finite volume expression

Now consider a finite sized cube region that is fixed in space; i.e., an Eulerian region such as shown in Figure 17.1. The mass contained within the cube is given by

$$m = \rho \Delta V = \rho \Delta x \Delta y \Delta z, \quad (17.12)$$

where the cube volume, $\Delta V = \Delta x \Delta y \Delta z$, is constant in time. As we will be taking the limit as the size of the cube gets smaller, it is sufficient to approximate the density as that at the cube center, $\rho = \rho(x, y, z, t)$. In the absence of mass sources within the fluid, the mass within the cube changes only through the transport of mass across the six cube faces. Focusing on the mass transported in the meridional direction as illustrated in Figure 17.1, the accumulation of mass within the cube through this transport is determined by the difference in mass transport crossing the two adjacent cell faces

$$\text{mass change from meridional transport} = (\Delta x \Delta z) [(v \rho)_{y-\Delta y/2} - (v \rho)_{y+\Delta y/2}]. \quad (17.13)$$

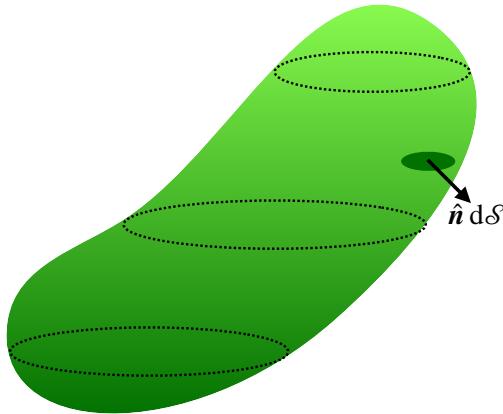


Figure 17.2: An arbitrarily shaped simply closed region, \mathcal{R} , within the fluid. If the region is fixed in space, then it represents a general Eulerian region for considering mass budgets. A surface area element, $d\mathcal{S}$, is oriented according to the outward normal, \hat{n} .

Expanding the difference into a Taylor series leads to

$$\text{mass change from meridional transport} \approx -(\Delta x \Delta y \Delta z) \frac{\partial(v \rho)}{\partial y}. \quad (17.14)$$

The same analysis for the zonal and vertical directions leads to the mass budget for the cube

$$\frac{\partial m}{\partial t} = -\Delta V \left[\frac{\partial(u \rho)}{\partial x} + \frac{\partial(v \rho)}{\partial y} + \frac{\partial(w \rho)}{\partial z} \right]. \quad (17.15)$$

Hence, the cube mass changes according to the convergence of mass across the cube boundaries. Cancelling the constant volume ΔV renders the flux-form continuity equation (17.9), $\partial\rho/\partial t = -\nabla \cdot (\rho \mathbf{v})$.

17.2.3 Arbitrary Eulerian region

The previous discussion for a cube can be generalized by making use of the Divergence theorem (Section 4.7). For that purpose, consider an arbitrary static and simply closed region within the fluid such as in Figure 17.2. Integrating the continuity equation (17.9) over that region leads to

$$\int_{\mathcal{R}} \frac{\partial \rho}{\partial t} dV = - \int_{\mathcal{R}} \nabla \cdot (\rho \mathbf{v}) dV. \quad (17.16)$$

Since the region is static we can move the partial time derivative outside on the left hand side. Furthermore, the Divergence theorem can be applied to the right hand side to convert the volume integral to a surface integral over the boundaries of the static domain. The resulting mass budget is given by

$$\frac{\partial}{\partial t} \int_{\mathcal{R}} \rho dV = - \oint_{\partial\mathcal{R}} \rho \mathbf{v} \cdot \hat{n} d\mathcal{S}, \quad (17.17)$$

where \hat{n} is the outward normal vector along the closed boundary of the region, and $d\mathcal{S}$ is the surface area element along that boundary. This equation says that the mass within a fixed region of the fluid changes in time (left hand side) according to the accumulation of mass crossing the region boundary (right hand side).

17.3 Material fluid regions

We now extend the kinematics of material fluid parcels to finite sized material fluid regions. As for material fluid parcels, the finite sized material fluid region retains the same matter content, and thus maintains a constant mass. We contrast the discussion here with that for the Eulerian regions (fixed in space) considered in Section 17.2. One key operational distinction between the Eulerian and material domains is that partial time derivative operators commute with integration over a fixed Eulerian domain, whereas material time derivative operators commute with integration over a material domain as per Reynolds Transport Theorem proven in Section 17.3.4.

17.3.1 Evolution of volume

Consider a finite material region, $\mathcal{R}(\mathbf{v})$, whose volume is given by the integral

$$V = \int_{\mathcal{R}(\mathbf{v})} dV, \quad (17.18)$$

with dV the volume element. The region changes its shape according to the motion of the fluid particles fixed to the boundary of the material region. We designate this volume as

$$\mathcal{R}(\mathbf{v}) = \text{region following flow}, \quad (17.19)$$

with the \mathbf{v} argument emphasizing that the volume moves with the flow. The material region expands when the flow moves outward, and contracts when the flow moves inward. These statements take on the following mathematical expression

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} dV = \oint_{\partial\mathcal{R}(\mathbf{v})} \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (17.20)$$

where $\hat{\mathbf{n}}$ is the outward normal on the region's closed boundary, $d\mathcal{S}$ is the area element on the boundary, and

$$\mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = \text{volume transport (volume per time) at the boundary } \partial\mathcal{R}. \quad (17.21)$$

Use of Gauss's divergence theorem then leads to the equivalent expression

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} dV = \int_{\mathcal{R}(\mathbf{v})} \nabla \cdot \mathbf{v} dV. \quad (17.22)$$

We now take the limit as the material region becomes a material parcel, in which case we recover the differential expression

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v}, \quad (17.23)$$

where we make use of D/Dt since the infinitesimal volume is moving with the fluid. This equation is also derived in Section 20.4.1 using different methods.

17.3.2 Mass conservation

The mass of fluid contained in a finite material region is given by

$$M = \int_{\mathcal{R}(\mathbf{v})} \rho dV. \quad (17.24)$$

As a material fluid region, it maintains a constant mass as it moves through the fluid

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \rho dV = 0. \quad (17.25)$$

Just as for the volume in Section 17.3.1, taking the limit as the material region becomes infinitesimally small, the region mass conservation statement (17.25) becomes the parcel mass conservation statement (17.3)

$$\frac{D(\delta M)}{Dt} = \frac{D(\rho \delta V)}{Dt} = 0. \quad (17.26)$$

17.3.3 Mass conservation using Lagrangian methods

Rather than take the limit as the finite material region $\mathcal{R}(\mathbf{v})$ becomes infinitesimal, we develop some formalism that specifies how to move the time derivative across the integral over the finite sized material region in equation (17.25). As part of this discussion we introduce two coordinate representations of the material region. The first is the Cartesian \mathbf{x} -space representation (e.g., assume the region is rectangular)

$$\int_{\mathcal{R}} \rho dV = \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} \rho dx dy dz \equiv \int_{\mathcal{R}[\mathbf{v}(\mathbf{x}, t)]} \rho dV(\mathbf{x}), \quad (17.27)$$

where the second equality defines the notation where the region bounds and volume element are moving with the fluid velocity, \mathbf{v} , as specified using the \mathbf{x} -space representation, $\mathbf{v}(\mathbf{x}, t)$. When using the \mathbf{x} -space representation, it is notable that the bounds on the integral are functions of time since the material region is moving with the fluid.

The second representation makes use of the \mathbf{a} -space material coordinates. For this representation we perform a coordinate transformation from \mathbf{x} -space to \mathbf{a} -space, which necessitates the Jacobian of transformation. To capture the gist of this transformation, consider the one-dimensional case in which

$$\int_{x_1(t)}^{x_2(t)} \rho dx = \int_{x_1(t)}^{x_2(t)} \rho \frac{\partial X}{\partial a} da = \int_{a_1}^{a_2} \rho \frac{\partial X}{\partial a} da. \quad (17.28)$$

The first equality introduced the Jacobian, $\partial X / \partial a$, for the one-dimensional coordinate transformation from x -space to a -space. The second equality wrote the integral bounds in terms of the material coordinate, which we can do since there is a one-to-one relation between a -space and x -space. Since we are considering a material region, it follows fluid particles. Hence, the integral bounds have fixed material coordinate values, a_1 and a_2 . These material coordinate values do not change in time since they mark the moving endpoints of the material line. Generalizing to three dimensions then renders

$$\int_{\mathcal{R}[\mathbf{v}(\mathbf{x}, t)]} \rho dV(\mathbf{x}) = \int_{\mathcal{R}[\mathbf{v}(\mathbf{a})]} \rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} dV(\mathbf{a}). \quad (17.29)$$

Again, we distinguish the \mathbf{x} and \mathbf{a} arguments for the material region, with the \mathbf{x} representation also requiring a time argument whereas the material \mathbf{a} representation has no time dependence.

We now make use of the equality (17.29) to take the time derivative of the mass contained in the material region. Since the time derivative follows the material region, we are motivated to make

use of a material space coordinate representation as part of the manipulations

$$\frac{d}{dt} \left[\int_{\mathcal{R}[\mathbf{v}(\mathbf{x}, t)]} \rho dV(\mathbf{x}) \right] = \frac{d}{dt} \left[\int_{\mathcal{R}[\mathbf{v}(\mathbf{a})]} \rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} dV(\mathbf{a}) \right] \quad (17.30a)$$

$$= \int_{\mathcal{R}[\mathbf{v}(\mathbf{a})]} \frac{D}{Dt} \left[\rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right] dV(\mathbf{a}) \quad (17.30b)$$

$$= \int_{\mathcal{R}[\mathbf{v}(\mathbf{x}, t)]} \left[\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right] dV(\mathbf{x}) \quad (17.30c)$$

$$= \int_{\mathcal{R}[\mathbf{v}(\mathbf{x}, t)]} \left[\frac{1}{\rho} \frac{D\rho}{Dt} + \nabla \cdot \mathbf{v} \right] \rho dV(\mathbf{x}). \quad (17.30d)$$

When expressing the integral bounds using \mathbf{a} -space coordinates, the integral bounds are static, thus allowing us to move the time derivative inside of the integral sign to reach the second equality. Upon entering the integral the time derivative is written as a material time derivative, D/Dt , since it is following fluid particles. The third equality made use of equation (17.7c) and converted back to \mathbf{x} -space. As the material region \mathcal{R} has a materially constant mass, we recover the mass continuity equation (17.6) by setting the integrand in equation (17.30d) to zero. This technique of moving between Eulerian (\mathbf{x} -space) and Lagrangian (\mathbf{a} -space) representations is commonly used in fluid mechanics. It is formalized by the Reynolds transport theorem discussed next.

17.3.4 Reynolds transport theorem

Manipulations leading to the mass conservation statement (17.30d) can be generalized by considering the material time derivative of a mass-weighted field ψ (e.g., a tracer concentration as in Section 18.1)

$$\frac{D(\psi \rho \delta V)}{Dt} = \frac{D\psi}{Dt} \rho \delta V + \psi \frac{D(\rho \delta V)}{Dt} \quad (17.31a)$$

$$= \rho \delta V \left[\frac{D\psi}{Dt} + \frac{\psi}{\rho} \frac{D\rho}{Dt} + \psi \nabla \cdot \mathbf{v} \right] \quad (17.31b)$$

$$= \delta V \left[\frac{\partial(\rho \psi)}{\partial t} + \nabla \cdot (\rho \psi \mathbf{v}) \right]. \quad (17.31c)$$

The first equality used the product rule, which holds for material time derivatives. Mass conservation means that the material derivative $D(\rho \delta V)/Dt = 0$. However, we choose to write mass conservation in the form of equation (17.7c), which allows us to introduce the flux-form Eulerian expression after replacing the material time derivative with its Eulerian form from equation (16.28). Another, perhaps more direct means, to derive this result is to write

$$\rho \frac{D\psi}{Dt} = \rho \left[\frac{\partial \psi}{\partial t} + \mathbf{v} \cdot \nabla \psi \right] \quad (17.32a)$$

$$= \rho \left[\frac{\partial \psi}{\partial t} + \mathbf{v} \cdot \nabla \psi \right] + \psi \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{v} \rho) \right] \quad (17.32b)$$

$$= \frac{\partial(\rho \psi)}{\partial t} + \nabla \cdot (\rho \mathbf{v} \psi). \quad (17.32c)$$

Following the discussion in Section 17.3.3, we can extend the material parcel result (17.31c) to a finite size material region. Again, each point in the material region is following a fluid particle.

The result is known as the *Reynolds Transport Theorem*

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \psi \rho dV = \int_{\mathcal{R}(\mathbf{v})} \frac{D\psi}{Dt} \rho dV \quad (17.33a)$$

$$= \int_{\mathcal{R}(\mathbf{v})} \left[\frac{\partial(\rho\psi)}{\partial t} + \nabla \cdot (\rho\psi \mathbf{v}) \right] dV \quad (17.33b)$$

$$= \int_{\mathcal{R}(\mathbf{v})} \frac{\partial(\rho\psi)}{\partial t} dV + \oint_{\partial\mathcal{R}(\mathbf{v})} \rho\psi \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (17.33c)$$

where the final equality follows from Gauss's divergence theorem. Additionally, we returned to the notation $\mathcal{R}(\mathbf{v})$ for material region as introduced in Section 17.3.1. This notation is sufficient to designate that the region is following fluid elements as per the fluid velocity field, \mathbf{v} . Note that for the surface integral term, $\mathbf{v} \cdot \hat{\mathbf{n}}$ generally does not vanish. Rather, it is given by $\mathbf{v} \cdot \hat{\mathbf{n}} = \mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}$, where $\mathbf{v}^{(s)}$ is the velocity of a point on the boundary of the material region. We detail this result in Section 17.4.2 when discussing the kinematics of a moving material surface.

17.3.5 Comments on notation

In this book we write d/dt for the time derivative operator acting on an integral. Furthermore, when the domain is specialized to follow fluid particles, we identify the special nature of such domains by introducing the fluid velocity argument to the domain name, $\mathcal{R}(\mathbf{v})$. This notation designates that all points in the domain, \mathcal{R} , move with the fluid velocity, \mathbf{v} . However, many authors choose an alternative notation by using the material time derivative, D/Dt , when acting on an integral over a material region. We thus have the following equality across the two notational conventions

$$\frac{D}{Dt} \int_{\mathcal{R}} \psi \rho dV = \frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \psi \rho dV. \quad (17.34)$$

The use of one convention versus the other is a matter of taste. We follow Section 2.1 of [Batchelor \(1967\)](#) by restricting the D/Dt operator to act only on space-time fields, such as $\psi(\mathbf{x}, t)$. Hence, the D/Dt operator is not used when acting on integrals over spatial regions. Following this convention leads us to write $\mathcal{R}(\mathbf{v})$ for a region that moves with the fluid flow and to retain d/dt when acting on the integral over that region.

Admittedly, the $\mathcal{R}(\mathbf{v})$ notation is rarely used in the literature, with authors generally dropping the \mathbf{v} and thus letting words designate whether a region follows the flow or otherwise. As we have occasion in this book to consider a variety of fluid regions, we find it essential to introduce the somewhat more pedantic notation $\mathcal{R}(\mathbf{v})$ to denote what sort of region is being considered. This usage helps the reader safely follow the mathematical flow rather than swimming alone in rough seas introduced by notational confusion.

17.4 Kinematic boundary conditions

When a fluid encounters a boundary, either at the edge of the fluid region or within the fluid itself, the fluid must accommodate the boundary. Alternatively, the boundary must accommodate the fluid. Some boundaries are impermeable, so that they do not allow matter to cross. For material boundaries, any fluid originally in contact with the boundary stays in contact; at most this fluid can move tangential to the boundary without leaving it. Other boundaries are permeable, thus allowing matter to cross. Our goal in this section is to develop the various kinematic boundary conditions appropriate for the variety of cases encountered in fluid mechanics. These boundary conditions, and the considerations arising in their derivation recur throughout fluid mechanics.

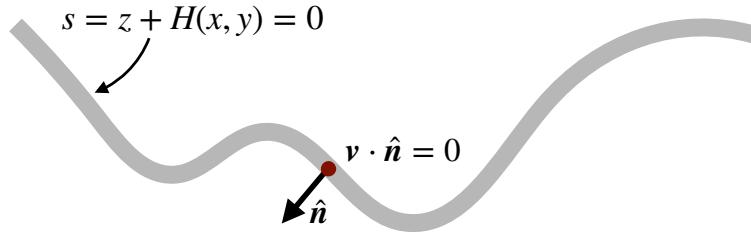


Figure 17.3: Illustrating the no-normal flow boundary condition maintained for a solid boundary, on which $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$ (equation (17.35)). When the solid boundary denotes the ground or ocean bottom, then the position of the interface is written $s(x, y, z) = z + H(x, y) = 0$ (equation (17.36)).

17.4.1 Static material surface

When a fluid encounters a static material surface, such as the solid-earth, the normal component of the fluid velocity must vanish since there is no fluid crossing the boundary (see Figure 17.3)

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{no-flux condition on static material boundary.} \quad (17.35)$$

Recall our discussion of streamlines in Section 16.7.2, where $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$ along a streamline. We thus see that the static material boundary is a flow streamline. That is, fluid that is in contact with the boundary will remain in contact. As the boundary is static, this result holds even in the case of a time dependent flow. Note that specification of the tangential velocity along a material boundary requires dynamical information unavailable from the purely kinematics analysis presented here.

For many cases in practice, the material surface is monotonic in the vertical, meaning there are no overturns. In this case, it is useful to introduce some differential geometry (at the level of introductory calculus) to unpack the boundary condition (17.35). Doing so helps to develop a geometric formalism especially useful for the more complicated moving boundary conditions in Sections 17.4.2 and 17.4.3. For this purpose, introduce a coordinate expression for the boundary according to

$$s(x, y, z) = z + H(x, y) = 0 \quad \text{static material boundary,} \quad (17.36)$$

with $z = -H(x, y)$ the vertical position of the boundary. The outward normal vector at the boundary is thus given by

$$\hat{\mathbf{n}} = -\frac{\nabla s}{|\nabla s|} = -\frac{\nabla(z + H)}{|\nabla(z + H)|} = -\frac{\hat{\mathbf{z}} + \nabla H}{\sqrt{1 + (\nabla H)^2}}. \quad (17.37)$$

Consequently, the no-flux boundary condition (17.35) takes the form

$$w + \mathbf{u} \cdot \nabla H = 0 \quad \text{at } z = -H(x, y), \quad (17.38)$$

where the velocity is decomposed into its horizontal and vertical components,

$$\mathbf{v} = (\mathbf{u}, w). \quad (17.39)$$

Hence, to maintain the no-flux boundary condition requires the vertical velocity component to precisely balance the projection of the horizontal velocity onto the slope of the material surface. If the material surface is flat, so that $\nabla H = 0$, then the kinematic boundary condition reduces to $w = 0$.

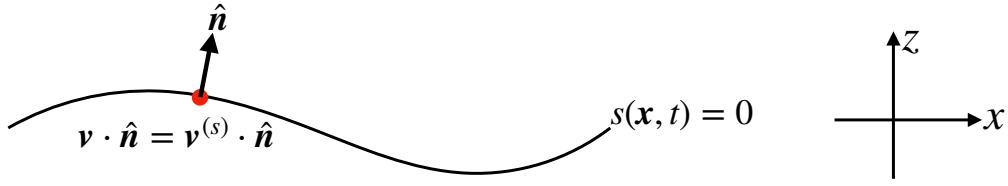


Figure 17.4: Illustrating the boundary condition for a moving material surface, on which $\mathbf{v} \cdot \hat{\mathbf{n}} = \mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}$ (equation (17.40)). For many cases, we can specify the surface by the value of a function that is a constant on the surface: $s(x, y, z, t) = s_{\text{const}}$ (equation (17.41)), in which case the normal direction is given by $\hat{\mathbf{n}} = |\nabla s|^{-1} \nabla s$ (equation (17.42)). For a multi-component fluid, \mathbf{v} is the barycentric velocity of Section 18.1, so that the material surface allows matter to be exchanged across it in the presence of diffusion.

17.4.2 Moving material surface

We now consider the kinematic constraints imposed by a material surface moving with the flow. Such material surfaces follow the flow as defined by the barycentric velocity. Consequently, they do not allow net mass to cross the surface although for multi-component fluids they may allow matter to be exchanged in the presence of diffusion.

General expression of the kinematic boundary condition

To ensure no net flow crosses the surface, the surface must have a velocity that matches that of the fluid. More precisely, the normal component of the surface velocity must match the normal component of the fluid. We are thus led to the kinematic boundary condition for a moving material surface

$$(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{moving material boundary condition.} \quad (17.40)$$

In this equation, $\mathbf{v}^{(s)}$ is the velocity of a point fixed on the moving material surface and \mathbf{v} is the barycentric velocity. Note that this constraint does not mean \mathbf{v} and $\mathbf{v}^{(s)}$ are identical. It only says that their normal components are the same when evaluated on the surface. We illustrate this boundary condition in Figure 17.4. Furthermore, as a Corollary to the boundary condition (17.40), we see that $\mathbf{v} \cdot \hat{\mathbf{n}}$ is not generally zero so that a moving material boundary does *not* coincide with a flow streamline (see discussion in Sections 16.7.2 and 16.7.3).

Specialized expression of the boundary condition

Now specialize the kinematic condition (17.40) to the case of a material surface specified by a function that takes on a constant value on the surface

$$s(\mathbf{x}, t) = s_{\text{const}}. \quad (17.41)$$

Correspondingly, the surface normal vector is given by

$$\hat{\mathbf{n}} = |\nabla s|^{-1} \nabla s. \quad (17.42)$$

From Section 16.4.6, we know that a point fixed on an arbitrary surface has a velocity that satisfies (see equation (16.41))

$$\frac{\partial s}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla s = 0 \quad \text{on an iso-surface } s(\mathbf{x}, t) = s_0. \quad (17.43)$$

Use of the identity

$$\frac{\partial s}{\partial t} = \frac{Ds}{Dt} - \mathbf{v} \cdot \nabla s \quad (17.44)$$

renders

$$\frac{Ds}{Dt} - \mathbf{v} \cdot \nabla s + \mathbf{v}^{(s)} \cdot \nabla s = \frac{Ds}{Dt} + (\mathbf{v}^{(s)} - \mathbf{v}) \cdot \nabla s = 0. \quad (17.45a)$$

Since $(\mathbf{v}^{(s)} - \mathbf{v}) \cdot \nabla s = 0$ from the boundary condition (17.40), we are left with the material constancy condition

$$\frac{Ds}{Dt} = \frac{\partial s}{\partial t} + \mathbf{v} \cdot \nabla s = 0 \quad \text{on material surface } s(\mathbf{x}, t) = s_0. \quad (17.46)$$

Consequently, no net matter crosses a surface of constant s as long as s is materially constant. This is an important kinematic property that reappears in many forms throughout fluid mechanics.

Boundary condition for a material interface

The expression (17.46) of the kinematic boundary condition is quite useful for many applications. For example, consider the interface between two immiscible fluids. Assume this surface is monotonic in the vertical (i.e., no breaking waves), so that we can express its vertical position as

$$s(x, y, z, t) = z - \eta(x, y, t) = s_o. \quad (17.47)$$

Without loss of generality, let the constant $s_o = 0$. The function $\eta(x, y, t)$ is the vertical deviation of the interface relative to the horizontal. The kinematic boundary condition (17.46) thus takes the form

$$\frac{Ds}{Dt} = \frac{D(z - \eta)}{Dt} = 0. \quad (17.48)$$

Hence, the vertical velocity component at the interface equals to the material time derivative of the interface displacement

$$\frac{Dz}{Dt} = \frac{D\eta}{Dt} \Rightarrow w = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta \quad \text{material b.c. at interface } z = \eta(x, y, t). \quad (17.49)$$

This boundary condition can be equivalently written in the form

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \frac{\partial \eta / \partial t}{\sqrt{1 + |\nabla \eta|^2}}, \quad (17.50)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{-\nabla \eta + \hat{\mathbf{z}}}{\sqrt{1 + |\nabla \eta|^2}} \quad (17.51)$$

is the outward normal at the material surface. These are expressions for the boundary condition placed on the ocean free surface when there is no rain or evaporation penetrating the surface. In general, they provide kinematic expressions for the motion of an interface between two immiscible fluid layers.

Another derivation of the material boundary condition

The material invariance condition $Ds/Dt = 0$ is a key kinematic result. We thus offer yet another derivation to help solidify its meaning. As before, define the surface according to

$$s(\mathbf{x}, t) = z - \eta(x, y, t) = 0, \quad (17.52)$$

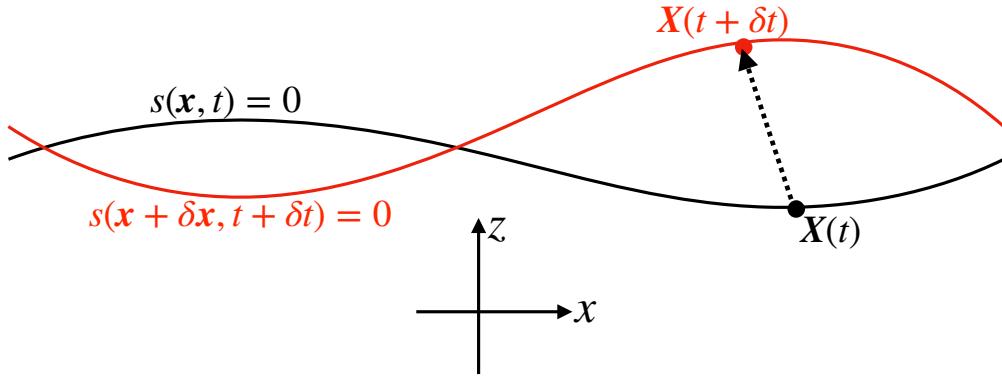


Figure 17.5: A surface that separates two fluid regions as realized at two time instants, along with the position of a sample point on the surface located at $\mathbf{X}(t)$ and $\mathbf{X}(t + \delta t)$. The equation $s(\mathbf{x}, t) = z - \eta(x, y, t) = 0$ specifies the vertical position for points on the surface as a function of horizontal position and time. At both time instances the vertical position is determined by $s(\mathbf{x}, t) = s(\mathbf{x} + \delta\mathbf{x}, t + \delta t) = 0$.

which specifies the vertical position of a point on the surface at time t . Now consider the position of the surface after a small time interval, $t + \delta t$ (see Figure 17.5). The vertical position of the surface at the new time is determined by

$$s(\mathbf{x} + \delta\mathbf{x}, t + \delta t) = 0, \quad (17.53)$$

where $\mathbf{X}(t + \delta t) = \mathbf{x} + \delta\mathbf{x}$ is the displaced position of a point on the surface that started at $\mathbf{X}(t) = \mathbf{x}$. Expanding equation (17.53) in a Taylor series to leading order yields

$$s(\mathbf{x}, t) + \delta\mathbf{x} \cdot \nabla s + \delta t \partial_t s = 0. \quad (17.54)$$

Since $s(\mathbf{x}, t) = 0$ from equation (17.52) we thus have

$$\frac{\partial s}{\partial t} + \frac{\delta\mathbf{x}}{\delta t} \cdot \nabla s = \frac{\partial s}{\partial t} + \frac{\delta\mathbf{x}}{\delta t} \cdot \hat{\mathbf{n}} |\nabla s| = 0, \quad (17.55)$$

where $\hat{\mathbf{n}} = |\nabla s|^{-1} \nabla s$ is the surface normal direction. This result means that when positioned at a fixed point in space, it is the normal component of the displacement that corresponds to a temporal modification of $s(\mathbf{x}, t)$

$$\frac{\partial s}{\partial t} = -\frac{\delta\mathbf{x}}{\delta t} \cdot \hat{\mathbf{n}} |\nabla s|. \quad (17.56)$$

In contrast, any tangential displacement along a constant s surface leaves $s(\mathbf{x}, t)$ unchanged, which in turn leaves the vertical position of a point on the surface unchanged. We are interested in motion of the surface that leads to nonzero vertical displacements. Hence, when following motion of points on the surface, we are only concerned with motion determined by the normal component of the velocity of that point, $\hat{\mathbf{n}} (\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}})$. It is this velocity component that corresponds to movement of the surface normal to itself, thus leading to nonzero vertical motion.

The above considerations mean that we are just concerned with the normal component of the velocity of a point on the surface, and so can write for the displacement

$$\delta\mathbf{x} = \delta t \hat{\mathbf{n}} (\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}). \quad (17.57)$$

Making use of this result in equation (17.55) leads to the differential equation satisfied by a point fixed on the moving surface

$$\frac{\partial s}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla s = 0, \quad (17.58)$$

where we used the identity

$$\frac{\delta \mathbf{x}}{\delta t} \cdot \nabla s = \hat{\mathbf{n}} (\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}) \cdot \nabla s = \mathbf{v}^{(s)} \cdot \nabla s. \quad (17.59)$$

Assuming the surface is material means that

$$\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}}, \quad (17.60)$$

so that vertical motion of the surface is identical to that of the fluid. This identity in equation (17.58) renders the material invariance condition $Ds/Dt = 0$.

17.4.3 Dynamic and permeable surface

We now consider the kinematic boundary condition for a moving permeable surface that separates two fluid media (e.g., ocean and atmosphere) or two regions within a single media (e.g., surfaces of constant specific entropy in the atmosphere or of constant potential density in the ocean). As before, the kinematic boundary condition is a statement about the mass transport through the boundary. Whereas the previous conditions enforced a zero net mass transport through the boundary, here we allow for a generally non-zero net transport (mass per time). We write this transport condition as

$$\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = -\mathcal{Q}_m d\mathcal{S} \quad \text{moving non-material boundary condition.} \quad (17.61)$$

In this equation, $d\mathcal{S}$ is an infinitesimal area element on the surface, and \mathcal{Q}_m is the mass per time per surface area crossing the boundary. The minus sign is a convention that will be motivated in the following. We now massage this kinematic boundary condition into alternative forms of use for a variety of purposes.

Coordinate representation of the permeable surface

The expression (17.43) for $\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}$ holds for a point on an arbitrary surface so that

$$\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = -\frac{\partial s / \partial t}{|\nabla s|}. \quad (17.62)$$

Furthermore, the projection of the barycentric fluid velocity onto the normal direction can be written

$$\frac{Ds}{Dt} = \frac{\partial s}{\partial t} + \mathbf{v} \cdot \nabla s \Rightarrow \mathbf{v} \cdot \hat{\mathbf{n}} = \frac{1}{|\nabla s|} \left(\frac{Ds}{Dt} - \frac{\partial s}{\partial t} \right). \quad (17.63)$$

Bringing these results together leads to

$$\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \frac{\rho d\mathcal{S}}{|\nabla s|} \frac{Ds}{Dt}. \quad (17.64)$$

This equation says that the net mass transport crossing the surface is proportional to the material time derivative of the surface coordinate. The material time derivative vanishes when there is no net transport across the surface (see discussion in Section 17.4.2).

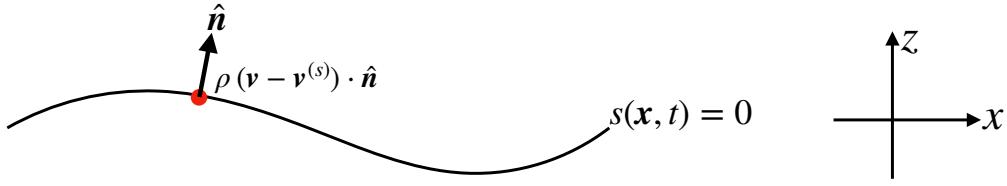


Figure 17.6: Illustrating the boundary condition for a moving permeable surface, such as the interface between two miscible fluid layers. On this surface, the boundary condition states that $\rho(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} d = -\mathcal{Q}_m$ (equation (17.61)). In the special case of an ocean free surface with no overturns, this boundary condition reduces to the surface kinematic boundary condition (17.74).

In terms of the horizontal projection of the surface area

Assume that the surface is not vertical, so that its normal direction has a nonzero component in the vertical (e.g., waves that do not overturn). This assumption means that

$$\frac{\partial s}{\partial z} \neq 0, \quad (17.65)$$

so that we can further massage the boundary condition (17.64) by writing the area factor in the form

$$\frac{ds}{|\nabla s|} = \frac{ds}{\sqrt{(\partial s / \partial x)^2 + (\partial s / \partial y)^2 + (\partial s / \partial z)^2}} \quad (17.66a)$$

$$= \frac{ds}{|\partial s / \partial z| \sqrt{[(\partial s / \partial x) / (\partial s / \partial z)]^2 + [(\partial s / \partial y) / (\partial s / \partial z)]^2 + 1}} \quad (17.66b)$$

$$= \frac{ds}{|\partial s / \partial z| \sqrt{1 + \tan^2 \theta}} \quad (17.66c)$$

$$= \left| \frac{\partial z}{\partial s} \right| |\cos \theta| ds \quad (17.66d)$$

$$= \left| \frac{\partial z}{\partial s} \right| dA. \quad (17.66e)$$

The equality (17.66c) introduced the angle, θ , between the boundary surface and the horizontal plane. The squared slope of this surface given by

$$\tan^2 \theta = \frac{\nabla_z s \cdot \nabla_z s}{(\partial s / \partial z)^2} = \nabla_s z \cdot \nabla_s z \quad (17.67)$$

with

$$\nabla_z = \hat{x} \left[\frac{\partial}{\partial x} \right]_{y,z} + \hat{y} \left[\frac{\partial}{\partial y} \right]_{x,z} \quad (17.68)$$

the horizontal gradient operator on constant z surfaces, and

$$\nabla_s = \hat{x} \left[\frac{\partial}{\partial x} \right]_{y,s} + \hat{y} \left[\frac{\partial}{\partial y} \right]_{x,s} \quad (17.69)$$

the horizontal gradient operator on constant s surfaces.¹ The equality (17.66d) made use of a trigonometric identity, and the equality (17.66e) introduced the horizontal projection of the area,

$$dA = |\cos \theta| dS. \quad (17.70)$$

¹We discuss such generalized vertical coordinates in Chapter 11.

These results bring the kinematic boundary condition (17.64) into the form

$$\rho(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = -Q_m d\mathcal{S} \quad (17.71a)$$

$$= \rho \frac{Ds}{Dt} \left| \frac{\partial z}{\partial s} \right| dA \quad (17.71b)$$

$$\equiv -Q_m dA. \quad (17.71c)$$

As defined, the flux Q_m is the net mass per time per horizontal area crossing the boundary surface

$$Q_m = -\rho(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} \left(\frac{d\mathcal{S}}{dA} \right) = -\rho \frac{Ds}{Dt} \left| \frac{\partial z}{\partial s} \right|. \quad (17.72)$$

The minus sign is a convention that we motivate through the ocean free surface example in the following.

Kinematic boundary condition at the ocean free surface

Consider the ocean free surface located at

$$s(x, y, z, t) = z - \eta(x, y, t) = 0 \quad \text{ocean free surface.} \quad (17.73)$$

For this boundary, $\partial s / \partial z = 1$, so that the boundary condition (17.71b) takes on the form

$$\rho \left[\frac{D(z - \eta)}{Dt} \right] = -Q_m \Rightarrow w + \rho^{-1} Q_m = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta. \quad (17.74)$$

To motivate the sign convention in equation (17.71c), consider the special case of a flat free surface and a resting fluid with $\mathbf{v} = 0$. Adding mass to the ocean raises the free surface, so that $\partial \eta / \partial t > 0$. Hence, our sign convention means that $Q_m > 0$ corresponds to mass added to the ocean.

Kinematic boundary condition on an isopycnal

Now consider the boundary surface to be a surface of constant potential density in the ocean (or analogously a surface of constant specific entropy in the atmosphere). These surfaces are known as isopycnals, and we use the symbol

$$s = \sigma(x, y, z, t) \quad (17.75)$$

for a particular isopycnal σ . The mass transport crossing the isopycnal is written

$$Q_m = \rho \frac{Ds}{Dt} \left| \frac{\partial z}{\partial \sigma} \right| \equiv \rho w^{(\sigma)}, \quad (17.76)$$

where we introduced the *diapycnal velocity*

$$w^{(\sigma)} \equiv \frac{Ds}{Dt} \left| \frac{\partial z}{\partial \sigma} \right|. \quad (17.77)$$

A key focus of physical oceanography concerns the development of theories for what causes a non-zero diapycnal velocity. Examples include breaking waves, which act to mix matter across density surfaces; i.e., to *entrain* water from one density class to another.

17.5 Mass budget for a column of ocean fluid

We close this chapter by deriving the mass budget for a column of ocean fluid such as that shown in Figure 17.7. This derivation requires much of the formalism discussed earlier, thus serving as a useful close to this chapter. In outline form, the derivation proceeds by vertically integrating the mass continuity equation (17.6) over the depth of an ocean column, from $z = -H(x, y)$ at the bottom to $z = \eta(x, y, t)$ at the free surface. Use of the bottom and surface kinematic boundary conditions renders a kinematic expression for the free surface time tendency.

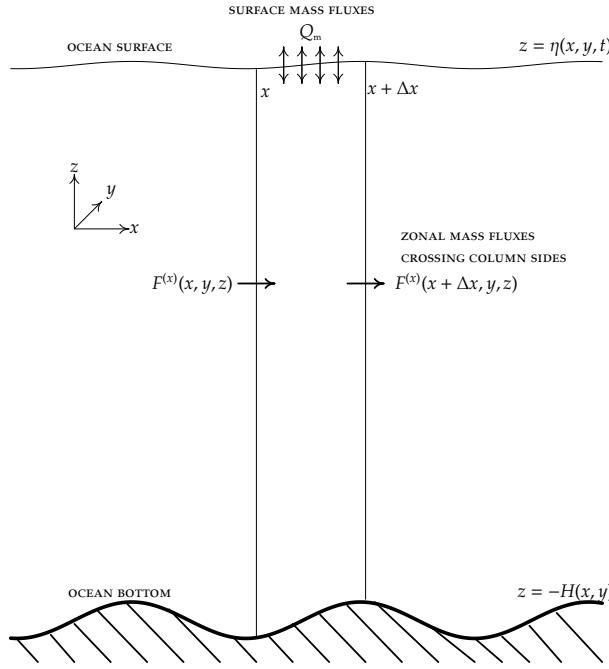


Figure 17.7: A longitudinal-vertical slice of ocean fluid from the surface at $z = \eta(x, y, t)$ to bottom at $z = -H(x, y)$. The horizontal boundaries of the column x and $x + \Delta x$ are static, whereas the free surface is time dependent thus making the horizontal cross-sectional area for the fluid column time independent. The ocean bottom at the solid-earth boundary, $z = -H(x, y)$, is also static with no mass crossing this interface. The ocean surface at $z = \eta(x, y, t)$ is time dependent with mass flux Q_m crossing this interface.

Vertically integrating the continuity equation (17.6) for a compressible fluid renders

$$-\int_{-H}^{\eta} \frac{1}{\rho} \frac{D\rho}{Dt} dz = \int_{-H}^{\eta} \nabla \cdot \mathbf{v} dz \quad (17.78a)$$

$$= w(\eta) - w(-H) + \int_{-H}^{\eta} \nabla_z \cdot \mathbf{u} dz \quad (17.78b)$$

$$= w(\eta) - w(-H) + \nabla_z \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right] - \mathbf{u}(\eta) \cdot \nabla_z \eta - \mathbf{u}(-H) \cdot \nabla_z H \quad (17.78c)$$

$$= [w(\eta) - \mathbf{u}(\eta) \cdot \nabla_z \eta] - [w(-H) + \mathbf{u}(-H) \cdot \nabla_z H] + \nabla_z \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right], \quad (17.78d)$$

where we made use of Leibniz's Rule to move the horizontal divergence outside of the integral. Also note that $\nabla \cdot \mathbf{u} = \nabla_z \cdot \mathbf{u}$ since \mathbf{u} is the horizontal velocity. Likewise for ∇H and $\nabla \eta$ since H and η are both spatially two-dimensional functions.

Use of the surface kinematic boundary condition (17.74) and no-normal flow bottom boundary

condition yield

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho(\eta)} - \nabla \cdot \mathbf{U} - \int_{-H}^{\eta} \frac{1}{\rho} \frac{D\rho}{Dt} dz \quad (17.79)$$

where

$$\mathbf{U} = \int_{-H}^{\eta} \mathbf{u} dz \quad (17.80)$$

is the depth integrated horizontal transport. Hence, as deduced from the mass continuity equation, the ocean free surface time tendency is affected by the passage of mass across the surface boundary (as normalized by the surface density), the convergence of depth integrated flow, and the depth integral of the material changes in density. [Griffies and Greatbatch \(2012\)](#) provide a more complete analysis of the sea surface height budget (17.79) by unpacking the physical processes leading to the material evolution of density.

17.6 Exercises

EXERCISE 17.1: CENTER OF MASS MOTION

Consider a material fluid region, $\mathcal{R}(\mathbf{v})$, with constant mass written as

$$M = \int_{\mathcal{R}(\mathbf{v})} \rho dV. \quad (17.81)$$

- (a) Show mathematically that the centre of mass for the region moves with the region's total linear momentum

$$\frac{d}{dt} \left[\frac{1}{M} \int_{\mathcal{R}(\mathbf{v})} \rho \mathbf{x} dV \right] = \frac{1}{M} \int_{\mathcal{R}(\mathbf{v})} \rho \mathbf{v} dV. \quad (17.82)$$

Precisely describe the reasoning behind each step. Note: a very brief solution is sufficient, so long as the reasoning is sound.

- (b) Show mathematically (or precisely describe why) that the time change in the linear momentum for the region is given by

$$\frac{d}{dt} \left[\int_{\mathcal{R}(\mathbf{v})} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}(\mathbf{v})} \frac{D\mathbf{v}}{Dt} \rho dV. \quad (17.83)$$

Precisely describe the reasoning behind each step. Note: a very brief solution is sufficient, so long as the reasoning is sound.

EXERCISE 17.2: MASS BUDGET FOR A FLUID COLUMN

We here derive the equation for mass conservation over a column of fluid, such as a seawater column extending from the ocean bottom to its surface. This exercise shares much with Section 17.5, but we come at the problem differently and arrive at a slightly different (though equivalent) form for column mass balance. Figure 17.7 provides a schematic of the setup. This is a long problem to state, but the solution is actually quite minimal.

The mass within an arbitrary fluid region is given by

$$m = \int \rho dV. \quad (17.84)$$

Consider the fluid mass within the column shown in Figure 17.7. In this column, the vertical side-walls are fixed in time, the bottom surface, $z = -H(x, y)$, is at the solid-earth boundary, and the top, $z = \eta(x, y, t)$, is the fluctuating ocean free surface. Convince yourself that the mass for this column can be written

$$m = \int dx dy \int_{-H(x,y)}^{\eta(x,y,t)} \rho dz, \quad (17.85)$$

where the horizontal (x, y) integrals extend over the horizontal area of the column. Mass conservation for this column means that the change in mass arises just through boundary fluxes, so that

$$\frac{dm}{dt} = - \int \rho \Delta \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (17.86)$$

where $\hat{\mathbf{n}}$ is the outward normal to the surface of the fluid region, $d\mathcal{S}$ is the area of an infinitesimal element on the surface, and the minus sign means that fluid leaving the region contributes to a reduction in mass within the region. The term

$$\Delta \mathbf{v} = \mathbf{v} - \mathbf{v}^{(s)} \quad (17.87)$$

is the velocity of the fluid relative to the velocity of the boundary.

Mass transported in the zonal direction (\hat{x}) that crosses the column's vertical boundary at x is given by

$$F^{(x)}(x, t) = \int dy \int_{-H(x,y)}^{\eta(x,y,t)} u(x, y, z, t) \rho(x, y, z, t) dz. \quad (17.88)$$

What are the physical dimensions for the mass transport $F^{(x)}$? There is a similar expression for mass crossing each of the other vertical faces of the column in the two horizontal directions (\hat{x}, \hat{y}). Using these expressions for the mass crossing the vertical side boundaries, take the limit as the horizontal cross-sectional area of the column becomes infinitesimally small to show that the evolution equation for the mass per unit area of the column is given by

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} \rho dz \right] = -\nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} \rho dz \right] + Q_m. \quad (17.89)$$

In this expression, Q_m is the mass transport entering the ocean through the surface, per horizontal area, as defined by equation (17.71c), so that

$$\int Q_m dx dy = - \int \rho \Delta \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad \text{at } z = \eta. \quad (17.90)$$

In words, the mass budget in equation (17.89) says that mass changes in a column of fluid if there is a convergence of mass into the column across its vertical boundaries (first term on right hand side), and a mass flux entering the column across the ocean surface (second term on right hand side). What are the physical dimensions of all terms in equation (17.89)?

EXERCISE 17.3: CHANGE IN LINEAR MOMENTUM OF A FLUID REGION

Consider a closed ocean basin with zero boundary fluxes of matter; i.e., zero precipitation/evaporation and zero mass fluxes through the solid-earth bottom. Consequently, this region is bounded by material surfaces and so it maintains constant matter content with fixed mass

$$M = \int_{\mathcal{R}} \rho dV. \quad (17.91)$$

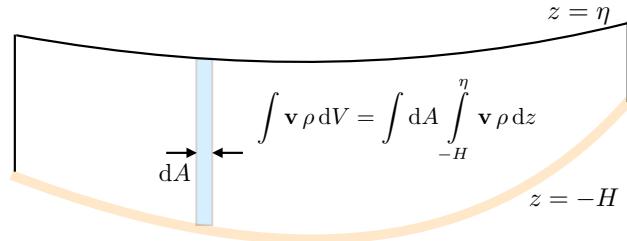


Figure 17.8: Cross-section of the integration region for Exercise 17.3, with the region extending from the ocean bottom at $z = -H(x, y)$ and the free surface at $z = \eta(x, y, t)$. An infinitesimal column is shown with cross-sectional area dA , extending from the bottom to the surface. The cross-sectional area for the column is time independent, so that a time derivative passes across the area integral to act only on the upper limit $z = \eta$ and the integrand in equation (17.93).

Show that the time change in the linear momentum for this ocean basin is given by

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}} \frac{D\mathbf{v}}{Dt} \rho dV. \quad (17.92)$$

This result is identical to that derived in Exercise 17.1. Rather than just repeating the solution method used there, make explicit use of Leibniz's theorem, the kinematic boundary condition detailed in Section 17.4.2, and mass conservation.

Hint: Refer to Figure 17.8 for a schematic of the integration where we have expanded the volume integral into the form

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \frac{d}{dt} \left[\int dA \int_{-H}^{\eta} \rho \mathbf{v} dz \right], \quad (17.93)$$

where the horizontal integral extends over the horizontal area of the basin, $dA = dx dy$ is the time independent horizontal area element, $z = -H(x, y)$ is the solid-earth bottom and $z = \eta(x, y, t)$ is the ocean free surface. Time dependence appears in the upper boundary at $z = \eta$ and within the integrand. Perform the time derivative operation and make use of mass continuity and the kinematic boundary condition. Also make use of the trigonometry presented in Section 17.4.3 (in particular equation (17.70)). Unlike the formulation in Exercise 17.1, there is no use of a material time derivative in this approach. Rather, it is a straightforward use of integration over a domain with fixed horizontal/bottom boundaries and a time dependent free surface boundary.

18

Material tracers and their budgets

As per Chapter 17, we assume that mass is neither created nor destroyed anywhere within the fluid domain, so that mass within the fluid changes only via fluxes crossing the domain boundary. We here extend that discussion to the case of a non-homogeneous fluid comprised of multiple trace matter constituents, thus developing budgets for tracers within the fluid, both for fluid elements and for finite regions.

READER'S GUIDE TO THIS CHAPTER

We assume an understanding of the Eulerian and Lagrangian kinematic descriptions detailed in Chapter 16 and the mass conservation analysis in Chapter 17. Much of the material in this chapter is used throughout the remainder of the book, particularly those subjects involving scalar fields such as heat, potential vorticity, and material tracers.

18.1	Fluid elements and the tracer equation	250
18.1.1	Material regions and material parcels for each constituent	250
18.1.2	Total mass conservation	251
18.1.3	Revisiting the fluid element	251
18.1.4	The tracer equation	252
18.1.5	Compatibility between mass continuity and the tracer equation	253
18.1.6	Comments	253
18.2	Passive tracers	254
18.3	Budgets for arbitrary fluid regions	254
18.3.1	Extensive and intensive properties	254
18.3.2	General form of the finite domain integral	255
18.3.3	Static domain	256
18.3.4	Leibniz-Reynolds Transport Theorem	256
18.3.5	Revisiting Reynolds Transport Theorem	258
18.3.6	Summary and comments on weak and strong formulations	259
18.4	Boundary conditions for the tracer budget	260
18.5	Brute force illustration of Leibniz-Reynolds	262
18.5.1	Leibniz's rule plus kinematic boundary conditions	262
18.5.2	Summarizing the result	263
18.6	Comments on local conservation laws	264

18.1 Fluid elements and the tracer equation

As defined in Section 16.1, a fluid element is an infinitesimal fluid region with constant mass but non-constant material composition. That is, a fluid element is a non-material fluid parcel. Fluid element boundaries are open to the exchange of trace matter (i.e., tracers) with adjacent elements. They are also open to the exchange of thermodynamic properties such as temperature and specific entropy. The kinematics of fluid elements share certain features with material fluid parcels. For example, we can uniquely specify the position of a fluid element by providing a material coordinate and time. Correspondingly, we can make use of Reynold's Transport Theorem for integration over a constant mass fluid region. We make use of fluid elements to develop the mass budgets for non-homogeneous fluids such as the ocean and atmosphere. The constituent mass budgets are commonly referred to as *tracer equations*.

18.1.1 Material regions and material parcels for each constituent

Consider a fluid with $n = 1, N$ matter constituents. For example, seawater has $N = 2$ when concerned just with its freshwater and salt content, whereas $N > 2$ when also concerned with other material constituents such as CO_2 and other biogeochemical species. Now focus on a region of the fluid, \mathcal{R} , with volume V and total mass M . Inside of \mathcal{R} , count the number of molecules of each constituent and determine their corresponding velocities. This information can be used to construct the molecular center of mass velocity for each constituent, $\mathbf{v}^{(n)}$, as well as the mass density,

$$\rho^{(n)} = V^{-1} M^{(n)}, \quad (18.1)$$

with these constituent properties defined over a constituent material region, $\mathcal{R}^{(n)}$. In the continuum limit where the volume and mass in \mathcal{R} get tiny yet the mass density remains finite, then the constituent velocity and mass density are continuous fields whose values are available at each point within the continuum fluid.

The constituent material regions, $\mathcal{R}^{(n)}$, are subsets of the region \mathcal{R} defined above, and they each move according to the corresponding constituent velocity, $\mathbf{v}^{(n)}$. That is, the material regions $\mathcal{R}^{(n)}$ maintain constant constituent mass $M^{(n)}$. In the infinitesimal limit, mass conservation for $\mathcal{R}^{(n)}$ leads the constituent mass continuity equation¹

$$\frac{D^{(n)}\rho^{(n)}}{Dt} = -\rho^{(n)} \nabla \cdot \mathbf{v}^{(n)} \quad \text{for each of the } n = 1, N \text{ constituents,} \quad (18.2)$$

where the constituent material time derivative is given by

$$\frac{D^{(n)}}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^{(n)} \cdot \nabla. \quad (18.3)$$

We thus have N statements of mass conservation corresponding to each constituent material fluid parcel moving according to the velocity $\mathbf{v}^{(n)}$.

18.1.2 Total mass conservation

The component mass continuity equation (18.2) takes on the Eulerian form

$$\frac{\partial \rho^{(n)}}{\partial t} + \nabla \cdot (\rho^{(n)} \mathbf{v}^{(n)}) = 0. \quad (18.4)$$

Summing over all constituents leads to the continuity equation for the total mass

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (18.5)$$

where the total mass density and barycentric velocity are given by

$$\rho = \sum_{n=1}^N \rho^{(n)} \quad \mathbf{v} = \rho^{-1} \sum_{n=1}^N \rho^{(n)} \mathbf{v}^{(n)}. \quad (18.6)$$

Introducing the total material time derivative, $D/Dt = \partial/\partial t + \mathbf{v} \cdot \nabla$, leads to the equivalent material form for the mass conservation equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \quad (18.7)$$

Note that the *barycenter* of a distribution of matter is the center of mass. We choose the term *barycentric velocity* for \mathbf{v} to distinguish \mathbf{v} from the molecular center of mass velocity, $\mathbf{v}^{(n)}$, of each constituent. The barycentric velocity plays a central role the conservation laws for fluid mechanics of multi-component fluids.

18.1.3 Revisiting the fluid element

The mass continuity equation (18.7) motivates us to define a fluid element as an infinitesimal fluid parcel that moves with barycentric velocity \mathbf{v} and maintains a constant total mass

$$\delta M = \sum_{n=1}^N \delta M^{(n)}. \quad (18.8)$$

¹There is no implied summation in equation (18.2).

The fluid element does not maintain a constant mass for each constituent, since the fluid element moves at the barycentric velocity, \mathbf{v} , which generally differs from the constituent velocities $\mathbf{v}^{(n)}$. Consequently, a fluid element boundary is permeable to matter transport that leaves its mass constant but allows for exchanges of matter constituents with adjacent fluid elements. Hence, if some matter leaves the fluid element, then an equal amount must enter the element in order to maintain a constant mass.

18.1.4 The tracer equation

Rather than keep track of each constituent velocity, $\mathbf{v}^{(n)}$, and the corresponding material parcels, it is generally more convenient to focus on the fluid element that moves with the barycentric velocity. For this purpose, we consider again the constituent mass continuity equation (18.4)

$$\left[\frac{\partial}{\partial t} + \mathbf{v}^{(n)} \cdot \nabla \right] \rho^{(n)} = -\rho^{(n)} \nabla \cdot \mathbf{v}^{(n)} \quad (18.9)$$

and insert the barycentric velocity

$$\left[\frac{\partial}{\partial t} + (\mathbf{v} - \mathbf{v} + \mathbf{v}^{(n)}) \cdot \nabla \right] \rho^{(n)} = -\rho^{(n)} \nabla \cdot [\mathbf{v} - \mathbf{v} + \mathbf{v}^{(n)}]. \quad (18.10)$$

Rearrangement leads to

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \rho^{(n)} = -\rho^{(n)} \nabla \cdot \mathbf{v} - \nabla \cdot [\rho^{(n)} (\mathbf{v}^{(n)} - \mathbf{v})], \quad (18.11)$$

which can be written

$$\frac{D\rho^{(n)}}{Dt} = -\rho^{(n)} \nabla \cdot \mathbf{v} - \nabla \cdot \mathbf{J}^{(n)}, \quad (18.12)$$

where we defined the constituent mass flux

$$\mathbf{J}^{(n)} = \rho^{(n)} (\mathbf{v}^{(n)} - \mathbf{v}). \quad (18.13)$$

The dimensions of $\mathbf{J}^{(n)}$ are mass of constituent n per time per area.

The material mass conservation equation (18.12) takes on the Eulerian form

$$\frac{\partial \rho^{(n)}}{\partial t} + \nabla \cdot (\mathbf{v} \rho^{(n)}) = -\nabla \cdot \mathbf{J}^{(n)}. \quad (18.14)$$

Introducing the tracer concentration $C^{(n)}$ according to

$$C^{(n)} = \frac{\rho^{(n)}}{\rho} = \frac{\delta M^{(n)}}{\delta M} = \frac{\text{mass of constituent } n \text{ in fluid element}}{\text{mass of fluid element}}, \quad (18.15)$$

leads to the tracer flux

$$\mathbf{J}^{(n)} = \rho C^{(n)} (\mathbf{v}^{(n)} - \mathbf{v}), \quad (18.16)$$

the Eulerian flux-form tracer budget

$$\frac{\partial(\rho C^{(n)})}{\partial t} + \nabla \cdot [\mathbf{v} \rho C^{(n)} + \mathbf{J}^{(n)}] = 0, \quad (18.17)$$

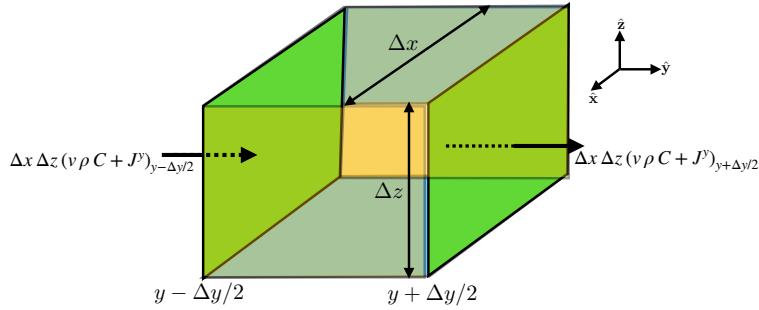


Figure 18.1: A finite sized cube as in Figure 17.1, here used to illustrate the budget of tracer mass over an Eulerian region. In addition to the advective flux of tracer moving with the barycentric velocity, \mathbf{v} , there is a diffusive flux, \mathbf{J} , that contributes to the transfer of tracer across the cell face.

and the corresponding material form of the tracer equation

$$\rho \frac{DC^{(n)}}{Dt} = -\nabla \cdot \mathbf{J}^{(n)}. \quad (18.18)$$

In Figure 18.1 we illustrate the contributions to the tracer evolution as viewed in the Eulerian flux-form tracer equation (18.17).

The above definitions allow us to decompose an advective tracer flux defined according to the tracer velocity into an advective flux based on the barycentric velocity plus a diffusive flux

$$\rho C^{(n)} \mathbf{v}^{(n)} = \rho C^{(n)} (\mathbf{v}^{(n)} - \mathbf{v} + \mathbf{v}) = \mathbf{J}^{(n)} + \rho C^{(n)} \mathbf{v}. \quad (18.19)$$

The diffusive flux vanishes when the tracer velocity equals to the barycentric velocity. The diffusive flux vanishes for a single-component fluid, since in that case there is only one matter component and thus no other matter component for which to diffuse. We have more to say on this topic when discussing passive tracers in Section 18.2.

18.1.5 Compatibility between mass continuity and the tracer equation

By construction, the Eulerian flux-form of the tracer equation (18.17) is compatible with the flux-form continuity equation

$$\frac{\partial(\rho C^{(n)})}{\partial t} + \nabla \cdot [\rho \mathbf{v} C^{(n)} + \mathbf{J}^{(n)}] = 0 \iff \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (18.20)$$

Compatibility is manifest by summing the tracer equation over all constituents and using the identities

$$\sum_{n=1}^N C^{(n)} = 1 \quad \sum_{n=1}^N \mathbf{J}^{(n)} = 0. \quad (18.21)$$

Furthermore, through use of the barycentric velocity (18.6), we are ensured that the continuity equation for the total density of a fluid element is only transported by the barycentric velocity. There is no contribution from $\mathbf{J}^{(n)}$ since $\sum_{n=1}^N \mathbf{J}^{(n)} = 0$.

18.1.6 Comments

The tracer equation expresses the balance of mass for each trace constituent in the fluid. Again, a nonzero tracer flux $\mathbf{J}^{(n)} = \rho C^{(n)} (\mathbf{v}^{(n)} - \mathbf{v})$ arises when the barycentric velocity, \mathbf{v} , differs from

the constituent velocity, $\mathbf{v}^{(n)}$. In that case, matter and thermodynamic properties are exchanged between fluid elements, with the exchange made without altering the mass of a fluid element. In the presence of random motion within a turbulent fluid, or in the presence of random interactions with molecular degrees of freedom, tracer exchange takes the form of a random walk. Such exchange is commonly parameterized by a diffusion process (see Section 37.2). Correspondingly, the mass of trace matter in a fluid element is altered in the presence of tracer concentration gradients.

18.2 Passive tracers

As defined in equation (18.15), the concentration of a material tracer is the mass of the trace constituent per mass of a fluid element. Such material tracers modify the barycentric velocity (18.6) since they carry mass and thus affect the mass density. We here introduce the related construct known as a *passive tracer*. A passive tracer satisfies the advection-diffusion equation, but it has zero impact on the velocity and is thus dynamically passive. Mathematically, we conceive of a passive tracer as a material tracer in the limit where the tracer mass and mass of the fluid element together go to zero. The passive tracer is thus analogous to the massless fluid particle of Section 16.1 whose trajectories define the Lagrangian reference frame. We make use of passive tracers to probe the advective and diffusive features of the flow without in turn modifying the flow. For example, a passive tracer can be used to define pathways (Section 41.5) and time scales for transport between fluid regions (Section 41.6).

In Section 17.4.2 as well as in Chapter 20, we discuss the notion of a material fluid object, which is an object comprised of fluid particles that follow the velocity, \mathbf{v} . In a homogeneous fluid, such material objects are impenetrable to matter, by construction. For a multi-component fluid, trace matter will generally cross the material object through diffusion since $\mathbf{v}^{(n)} \neq \mathbf{v}$. Hence, there is no perfectly impenetrable fluid object in a fluid with any form of diffusion, including molecular diffusion. However, we can conceive of a passive tracer that follows the barycentric velocity and either diffuses or not. Again, these enhanced features are afforded the passive tracer given that it is a massless idealization used to probe the fluid flow properties.

18.3 Budgets for arbitrary fluid regions

Thus far in this chapter we have considered the evolution of mass within a variety of fluid regions, including infinitesimal and finite domains either moving with the fluid or fixed in space. We have also considered similar domains in Chapter 17 where the fluid domains were typically material regions. In this section we synthesize these presentations by considering mass budgets over an arbitrary finite sized domain within non-homogeneous fluids. The resulting mass equations form the basis for matter budget analyses used in geophysical fluid mechanics.

18.3.1 Extensive and intensive properties

Physical properties can be characterized as *extensive* or *intensive*. An extensive property changes when the size of the sample changes. Examples are particle number, mass, length, volume, kinetic energy, entropy, enthalpy. An intensive property generally does not change when removing some of the sample. Examples are number density (number of particles per volume), mass density (mass of substance per unit volume), tracer concentration (mass of tracer per mass of fluid), temperature, velocity (linear momentum per mass), kinetic energy per mass, entropy per mass, and

enthalpy per mass. We have more to say about intensive and extensive properties when considering thermodynamics in Chapter 23.

We are concerned in this section with how extensive properties change as a function of time. Determining the evolution of such properties constitutes a budget analysis. What are the processes responsible for these changes? Where are the changes coming from? Those are the basic questions asked when performing a budget analysis. In addition to physical and biogeochemical processes active within the fluid, details of the region over which one performs a budget have an important impact on the budget. Is the region open to matter and energy transport, or is it closed? Is the region static in time or do boundaries move? If the boundaries move, then do they move with fluid elements or are they moving through some other manner?

In the following let ψ represent an intensive fluid property of a fluid element so that $\psi \rho \delta V$ is the corresponding extensive property for the element

$$\psi = \text{intensive fluid property} \quad (18.22a)$$

$$\psi \rho \delta V = \text{extensive fluid property}. \quad (18.22b)$$

For example, if ψ is the tracer concentration in a fluid element (i.e., mass of tracer per mass of fluid), then the corresponding extensive property, $\psi \rho \delta V$, is the mass of tracer in the fluid element. If ψ is a component of the velocity vector, then the corresponding extensive property, $\psi \rho \delta V$, is the component of linear momentum.

We furthermore assume that ψ satisfies the scalar conservation equation, written here in both its material form and Eulerian flux-form

$$\rho \frac{D\psi}{Dt} = -\nabla \cdot \mathbf{J} \iff \frac{\partial(\rho\psi)}{\partial t} + \nabla \cdot (\rho\psi \mathbf{v} + \mathbf{J}) = 0, \quad (18.23)$$

where \mathbf{J} is a flux such as that associated with the tracer equation derived in Section 18.1.4. Depending on the context, the budget equation (18.23) is sometimes referred to as a conservation law for ψ . Furthermore, in the absence of a flux convergence, $-\nabla \cdot \mathbf{J} = 0$, then the scalar field is constant following a material fluid particle

$$-\nabla \cdot \mathbf{J} = 0 \implies \frac{D\psi}{Dt} = 0. \quad (18.24)$$

In this case we say that ψ is a *material invariant* or a *material constant*.

18.3.2 General form of the finite domain integral

We are concerned here with the evolution of extensive fluid properties integrated over an arbitrary region. Let us make use of the following notation for such integrals

$$\mathcal{I}[\mathcal{R}(t), t] = \int_{\mathcal{R}(t)} \psi \rho dV \equiv \int_{\mathcal{R}(t)} \varphi dV, \quad (18.25)$$

where we introduced the shorthand

$$\varphi = \rho \psi. \quad (18.26)$$

The integrand in equation (18.25) is a function of space and time, $\varphi = \varphi(\mathbf{x}, t)$, and the integration region is generally a function of time, $\mathcal{R}(t)$. In previous sections, \mathcal{R} was a material region of fixed matter content (Section 17.3) or a constant mass fluid region open to the exchange of matter with the surroundings (Section 18.1). In both of these cases the region was denoted by $\mathcal{R}(\mathbf{v})$ since it

moved with the fluid flow. In contrast, here we make no *a priori* assumption about how the region moves.

The total time derivative of \mathcal{I} can be written as

$$\frac{d\mathcal{I}}{dt} = \left[\frac{\partial \mathcal{I}}{\partial t} \right]_{\mathcal{R}} + \frac{d\mathcal{R}}{dt} \left[\frac{\partial \mathcal{I}}{\partial \mathcal{R}} \right]_t. \quad (18.27)$$

The first term on the right hand side is the time derivative of the integral when holding the region fixed in time. The second term accounts for changes due to evolution of the region as weighted by the dependence of the integral on the region itself. How the integral changes in time depends on both the evolution of the fluid property relative to the chosen region and evolution of the fluid region itself. This result is directly analogous to the total time derivative of a field given by equation (16.26).

18.3.3 Static domain

The simplest case is when the domain is static, in which case

$$\frac{d\mathcal{I}}{dt} = \left[\frac{\partial \mathcal{I}}{\partial t} \right]_{\mathcal{R}} = \frac{\partial}{\partial t} \left[\int_{\mathcal{R}} \psi \rho dV \right] = \int_{\mathcal{R}} \left[\frac{\partial(\rho \psi)}{\partial t} \right] dV. \quad (18.28)$$

Movement of the time derivative across the integral sign is available since the domain boundaries are static; i.e., the second term on the right hand side of equation (18.27) vanishes. Furthermore, since the domain is static, the volume element, dV , is a static partition of the total domain volume. Consequently, dV does not appear inside the time derivative. This case corresponds to the Eulerian budgets schematized by Figures 17.1, 17.2, and 18.1.

18.3.4 Leibniz-Reynolds Transport Theorem

Now allow the domain boundaries to be time dependent so that both terms in the total time derivative in equation (18.27) contribute. We here derive the *Leibniz-Reynolds Transport Theorem*, which is a general expression of conservation over an arbitrary region.

Derivation of the transport theorem

Consider a one-dimensional domain with time dependent endpoints. Integrals of this type commonly arise when integrating over the depth of the atmosphere or ocean, in which case the boundary terms are replaced by kinematic boundary conditions (see Section 17.4). The chain rule for differentiating integrals is commonly referred to as *Leibniz's Rule*. It results in the time derivative acting on the upper integral limit, the lower limit, and the integrand

$$\frac{d}{dt} \left[\int_{x_1(t)}^{x_2(t)} \varphi(x, t) dx \right] = \int_{x_1(t)}^{x_2(t)} \frac{\partial \varphi}{\partial t} dx + \frac{d}{dt} \left[\int_{x_1(t)}^{x_2(t)} \right] \varphi(x, t) dx \quad (18.29a)$$

$$= \int_{x_1(t)}^{x_2(t)} \frac{\partial \varphi}{\partial t} dx + \frac{dx_2(t)}{dt} \varphi(x_2, t) - \frac{dx_1(t)}{dt} \varphi(x_1, t) \quad (18.29b)$$

$$= \int_{x_1(t)}^{x_2(t)} \left[\frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial x} \left(\varphi \frac{dx}{dt} \right) \right] dx. \quad (18.29c)$$

As a matter of convenience we brought the boundary terms back inside the integral for the final equality. Generalization of this result to a three-dimensional integral is straightforward

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \varphi dV \right] = \int_{\mathcal{R}} \left[\frac{\partial \varphi}{\partial t} + \nabla \cdot (\varphi \mathbf{v}^{(b)}) \right] dV = \int_{\mathcal{R}} \frac{\partial \varphi}{\partial t} dV + \oint_{\partial \mathcal{R}} \varphi \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (18.30)$$

where the second equality made use of Gauss's divergence theorem to transfer the volume integral into a boundary integral, and where we introduced the shorthand for the velocity of a point on the boundary

$$\mathbf{v}^{(b)} = \frac{d\mathbf{x}}{dt}. \quad (18.31)$$

Setting $\varphi = 1$ yields the time change for the region volume

$$\frac{d}{dt} \left[\int_{\mathcal{R}} dV \right] = \oint_{\partial \mathcal{R}} \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.32)$$

We emphasize that the boundary term in equations (18.30) and (18.32) projects out just the normal component to the boundary velocity; we never make use of information about its tangential component.

Transport theorem for a scalar field

The identity (18.30) is the Leibniz-Reynolds transport theorem. It is of great practical utility for fluid mechanics when interested in the evolution of fluid properties within a moving domain. We here display a corollary that proves useful for budget analyses over moving regions, with generalization to the vector linear momentum provided in Section 26.5.

For this purpose, make use of the Eulerian flux-form of the scalar conservation equation (18.23) so that the transport theorem is written

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \psi dV \right] = - \oint_{\partial \mathcal{R}} [\rho \psi (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.33)$$

Setting $\psi = 1$ gives an expression for the change in mass for the region

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho dV \right] = - \oint_{\partial \mathcal{R}} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.34)$$

The transport theorem (18.33) has a straightforward interpretation. Namely, the left hand side is the time tendency for the total ψ -stuff within the moving region. The right hand side is the surface integral for the transport of ψ -stuff through the boundary of the region. The first transport term arises from the difference between the barycentric fluid velocity and the velocity of the boundary, and the second term arises from the diffusive flux. Both transport terms are projected onto the outward normal at the boundary. Hence, the budget is not affected by transport tangential to the boundary. Finally, for the mass budget (18.34), the diffusive flux vanishes since mass of fluid elements moves according to the barycentric velocity of Section (18.1.2).

In Figure 18.2 we illustrate the transport theorem (18.33) for the special case of a discrete numerical model grid cell. This cell has fixed positions for the vertical sides whereas the top and bottom interfaces are time dependent. This application of the transport theorem provides the framework for finite volume methods in numerical models. We offer further discussion of the kinematics of such *generalized vertical coordinate* models in Chapter 21 and their dynamics in Chapter 32. In particular, in Section 21.3 we connect the advective transport across the moving surface to the dia-surface velocity component, which itself is proportional to material time changes in the generalized vertical coordinate.

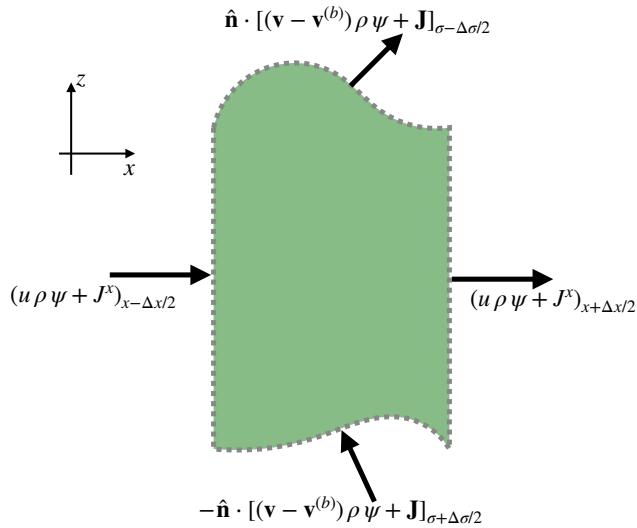


Figure 18.2: This figure depicts the contributions to the Leibnitz-Reynolds Transport Theorem (18.33). The theorem is applied to a domain corresponding to a numerical model grid cell with the top and bottom interfaces defined by generalized vertical coordinates of Chapters 11, 21, and 32. In particular, the vertical cell faces are assumed to have fixed positions, so that $(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}}$ for these cell faces. Hence, the fluxes crossing these faces are due to advection by the barycentric velocity plus the diffusive flux. However, the top and bottom faces of the cell are allowed to move according to the generalized vertical coordinate surfaces. Hence, transport through these faces must take into account the nonzero velocity of the boundaries. Note that numerical models generally assume the top and bottom interfaces have a nonzero projection in the vertical direction so that they never overturn.

18.3.5 Revisiting Reynolds Transport Theorem

We here consider a region that is moving with the fluid flow, in which case we provide a more general derivation of the *Reynolds Transport Theorem* originally derived for material regions in Section 17.3.4. The following results are mere special cases of the general expression (18.33). Even so, it is useful pedagogically to work through the special cases.

Reynolds Transport Theorem

Let us apply the result (18.30) to a region that follows the fluid flow as defined by the barycentric velocity, \mathbf{v} . For this moving region, the time derivative of the region boundaries in equation (18.30) is given by the fluid velocity thus leading to

$$\frac{d}{dt} \left[\int_{\mathcal{R}(\mathbf{v})} \varphi dV \right] = \int_{\mathcal{R}(\mathbf{v})} \left[\frac{\partial \varphi}{\partial t} + \nabla \cdot (\mathbf{v} \varphi) \right] dV = \int_{\mathcal{R}(\mathbf{v})} \left[\frac{D\varphi}{Dt} + \varphi \nabla \cdot \mathbf{v} \right] dV. \quad (18.35)$$

This result is the Reynolds transport theorem. The derivation given here is more general than that in Section 17.3.4 where we assumed the region to be material (i.e., no matter crosses the region boundary). For the present derivation we only assumed that the region boundaries move with the barycentric velocity. We did not assume the region boundaries are material. Consequently, we can make use of Reynolds Transport Theorem for constant mass regions moving with the fluid, where the region either has an impermeable (i.e., material) or permeable (non-material) boundary, with diffusion acting to make boundaries permeable.

Alternative form of Reynolds Transport Theorem

We can put the Reynolds Transport Theorem (18.35) into another useful form by reintroducing $\varphi = \rho\psi$ and making use of mass continuity

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}. \quad (18.36)$$

Doing so yields the rather tidy result

$$\frac{d}{dt} \left[\int_{\mathcal{R}(\mathbf{v})} \psi \rho dV \right] = \int_{\mathcal{R}(\mathbf{v})} \left[\frac{D\varphi}{Dt} + \varphi \nabla \cdot \mathbf{v} \right] dV \quad (18.37a)$$

$$= \int_{\mathcal{R}(\mathbf{v})} \left[\frac{D(\rho\psi)}{Dt} + \rho\psi \nabla \cdot \mathbf{v} \right] dV \quad (18.37b)$$

$$= \int_{\mathcal{R}(\mathbf{v})} \left[\psi \left(\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right) + \rho \frac{D\psi}{Dt} \right] dV \quad (18.37c)$$

$$= \int_{\mathcal{R}(\mathbf{v})} \frac{D\psi}{Dt} \rho dV. \quad (18.37d)$$

Heuristically, this result follows since ρdV is a constant when following the flow, so that passage of the time derivative across the integral only picks up the material derivative of ψ .

We can take the result (18.37d) one more step by inserting the material form of the scalar conservation equation (18.23) so that

$$\frac{d}{dt} \left[\int_{\mathcal{R}(\mathbf{v})} \psi \rho dV \right] = - \oint_{\partial\mathcal{R}(\mathbf{v})} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (18.38)$$

which is a special case of the general transport theorem (18.33) found by setting $(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0$ along the region boundary. This result says that the change in ψ -stuff within a region moving with the barycentric velocity arises only from the area integrated diffusive flux crossing normal to the boundary. This result is a finite volume generalization of the mass conservation statement for a fluid element as discussed in Section 18.1.4. We extend this result to linear momentum in Section 26.5.

18.3.6 Summary and comments on weak and strong formulations

We here summarize the variety of time derivatives acting on integrals of scalar fields

$$\frac{d}{dt} \int_{\mathcal{R}} \rho\psi dV = \begin{cases} \int_{\mathcal{R}} \frac{\partial(\rho\psi)}{\partial t} dV = -\oint_{\partial\mathcal{R}} (\rho\mathbf{v}\psi + \mathbf{J}) \cdot \hat{\mathbf{n}} d\mathcal{S} & \text{static domain} \\ \int_{\mathcal{R}(\mathbf{v})} \rho \frac{D\psi}{Dt} dV = -\oint_{\partial\mathcal{R}(\mathbf{v})} \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S} & \text{fluid flow domain} \\ -\oint_{\partial\mathcal{R}} [\rho\psi(\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} d\mathcal{S} & \text{arbitrary domain,} \end{cases} \quad (18.39)$$

with the scalar fields assumed to satisfy the partial differential equation

$$\rho \frac{D\psi}{Dt} = -\nabla \cdot \mathbf{J} \iff \frac{\partial(\rho\psi)}{\partial t} + \nabla \cdot (\rho\mathbf{v}\psi + \mathbf{J}) = 0. \quad (18.40)$$

In the language of mathematical fluid mechanics, the partial differential equation (18.40) is referred to as the *strong formulation* of the scalar budget, whereas the surface integral expressions in equation (18.39) provide various forms of the *weak formulations*. This terminology arises from the need

to provide more stringent assumptions concerning smoothness of the fields when using the partial differential equations, whereas the integral equations allow for discontinuous distributions of the field properties. The weak formulation is the foundation for developing *finite volume* methods to numerically solve the fluid mechanical equations.

18.4 Boundary conditions for the tracer budget

Here we extend the mass budget considerations from Section 17.5 to develop the budget for tracer mass (or heat content) within a fluid layer such as shown in Figure 18.3. Notably, we are interested in fluid layers that intersect surface (as for the ocean) and/or bottom boundaries (as for the ocean or atmosphere). We commonly think of this layer as defined by isolines of generalized vertical coordinates whose layers are monotonically stacked in the vertical according to the discussion from Sections 11.9.1 and 21.2. However, the treatment given here allows for the layers to be non-monotonic in the vertical (e.g., overturns are allowed), so that these results can be used for the water mass transformation analysis discussed in Chapter 41.

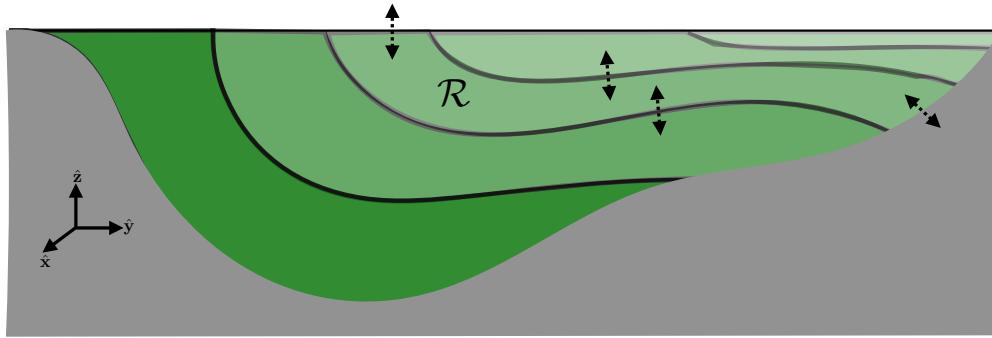


Figure 18.3: A depiction of fluid layers in which we formulate the budget for the total mass of tracer (or total heat content). The tracer mass within the layer, such as that one denoted by \mathcal{R} , is modified by dia-surface transport across interior layer interfaces, as well as transport across the surface and bottom boundaries. Note that an arbitrary layer might never intersect the bottom or surface boundaries. However, the layers depicted here each intersect boundaries, with such layers requiring extra care in formulating their tracer budgets.

The Leibniz-Reynolds transport theorem (18.33) provides our starting point

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = - \int_{\partial \mathcal{R}} [\rho C (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} dS, \quad (18.41)$$

where C is the concentration of a tracer that satisfies the advection-diffusion equation

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v} + \mathbf{J}) = 0. \quad (18.42)$$

The left hand side of equation (18.41) is the time tendency for the mass of tracer within the layer, such as the region \mathcal{R} shown in Figure 18.3. This tendency is affected by transport across the layer boundaries, with three boundaries considered here.

Interior layer boundaries

The boundary transport across interior layer interfaces,

$$\text{interior boundary transport} = [\rho C (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} dS, \quad (18.43)$$

measures the tracer mass transport due to advection across the moving layers (first term) and subgrid scale fluxes at the layer (second term). The advective term is known as the dia-surface transport, with an extensive discussion of its kinematics given in Section 21.3.

Bottom boundary

At the bottom boundary, the no-normal flow condition means that

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0. \quad (18.44)$$

Consider the velocity of a point attached to the layer interface, $\mathbf{v}^{(b)}$, that also tracks the position of the interface as it intersects the bottom boundary. By construction, the movement of this intersection point is tangential to the bottom boundary so that it too is orthogonal to the boundary outward normal direction

$$\mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} = 0. \quad (18.45)$$

Hence, the only contribution to the tracer budget at the bottom boundary comes through the subgrid scale flux \mathbf{J}

$$\text{bottom boundary transport} = \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.46)$$

Geothermal heating is a common example of bottom boundary transport in the ocean. In the atmosphere, planetary boundary layer processes transfer tracer content from the land-atmosphere and sea-atmosphere boundary layer into the troposphere above.

Let us consider the geothermal heating with a bit more detail. In this case heating can be incorporated into the ocean as a boundary flux via

$$\text{geothermal heating} = c_p \mathbf{J}(\Theta) \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (18.47)$$

where c_p is the heat capacity and Θ is the Conservative Temperature (see Chapter 40). If we furthermore assume the Θ flux is downgradient diffusion as per Fourier's Law with diffusivity $\kappa_\Theta > 0$ (Section 37.2.4), then the boundary condition takes the form

$$\text{geothermal heating} = -c_p \rho \kappa_\Theta \nabla \Theta \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.48)$$

Geothermal heating thus supports a non-zero projection of the temperature gradient onto the normal direction. For those cases where the geothermal heating vanishes, or more generally for tracers that have zero boundary flux, then the tracer must satisfy the following no-normal derivative boundary condition

$$\text{no flux boundary} = \nabla C \cdot \hat{\mathbf{n}} = 0. \quad (18.49)$$

In this case, isosurfaces of constant tracer are oriented normal to the boundary as depicted in Figure 18.4. For the dynamical tracers like temperature and salinity, this boundary condition affects flow near the boundary by modifying the density field and thus the pressure.

Upper ocean boundary

We make use of the kinematic boundary condition derived in Section 17.4.3 for the permeable air-sea boundary, where the boundary condition (17.61) is given by

$$\rho \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(b)}) = -\mathcal{Q}_m \quad \text{air-sea boundary} \quad (18.50)$$

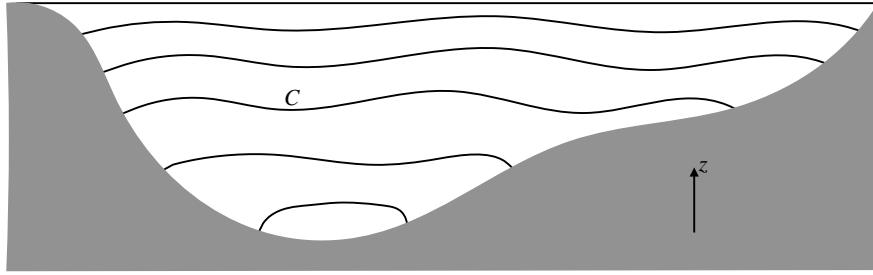


Figure 18.4: In the absence of a boundary tracer flux, isosurfaces of tracer, C , intersect solid boundaries normal to the boundary as per equation (18.49): $\nabla C \cdot \hat{\mathbf{n}} = 0$, where $\hat{\mathbf{n}}$ is the outward boundary normal direction.

with \mathcal{Q}_m the mass per time per surface area crossing the boundary. We are thus led to the air-sea boundary condition for trace matter

$$[\rho C (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} d\mathcal{S} = (C \mathcal{Q}_m + \mathbf{J}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.51)$$

The first term accounts for the transport of trace matter within the mass crossing the boundary. The second term accounts for transport through subgrid scale processes, such as turbulent air-sea boundary fluxes. In Section 40.3 we discuss the variety of salt and heat transports crossing the ocean boundary that contribute to changes in ocean buoyancy, most of which appear as boundary contributions to \mathbf{J} .

18.5 Brute force illustration of Leibniz-Reynolds

The Leibniz-Reynolds transport theorem

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = - \int_{\partial \mathcal{R}} [\rho C (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (18.52)$$

is an incredibly useful and elegant expression of the tracer budget over an arbitrary domain. Correspondingly, we make great use of it throughout this book. To further our understanding, we here consider the tracer budget for an ocean domain such as in Figure 18.3. Rather than make direct use of Leibniz-Reynolds, we use a brute force approach by expanding the volume integral according to

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = \frac{d}{dt} \left[\int_{A(t)} dA \int_{-H}^{\eta} \rho C dz \right]. \quad (18.53)$$

In this equation, $\int_{A(t)} dA$ is an integral over the horizontal area of the domain, with the lateral boundaries of the domain generally a function of time.

18.5.1 Leibniz's rule plus kinematic boundary conditions

Performing the time derivative in equation (18.53) and using Leibniz's rule yields

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = \frac{dA}{dt} \left[\int_{-H}^{\eta} \rho C dz \right]_{\text{bounds}} + \int_{A(t)} [\partial_t \eta (\rho C)_{z=\eta}] dA + \int_{A(t)} dA \int_{-H}^{\eta} \frac{\partial(\rho C)}{\partial t} dz. \quad (18.54)$$

The first term on the right hand side is evaluated along the lateral boundaries of the domain. If the boundaries are fixed in time, as in a box of seawater or a periodic channel, then $dA/dt = 0$.

The more general case has a boundary that is time dependent such as along a beach where fluid moves up and down the sloping shoreline. However, in that case the thickness of fluid vanishes at the lateral boundary, $H + \eta = 0$, thus again revealing that the first term on the right hand side drops from the budget to render

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = \int_{A(t)} [\partial_t \eta (\rho C)_{z=\eta}] dA + \int_{A(t)} dA \int_{-H}^{\eta} \frac{\partial(\rho C)}{\partial t} dz. \quad (18.55)$$

For the second term on the right hand side of equation (18.55) we make use of the tracer equation (18.42) and Leibniz's rule to write

$$\int_{-H}^{\eta} \frac{\partial(\rho C)}{\partial t} dz = - \int_{-H}^{\eta} \nabla_z \cdot (\rho C \mathbf{u} + \mathbf{J}_h) dz - \int_{-H}^{\eta} \frac{\partial(\rho C w + J^z)}{\partial z} dz \quad (18.56a)$$

$$\begin{aligned} &= -\nabla_z \cdot \int_{-H}^{\eta} (\rho C \mathbf{u} + \mathbf{J}_h) dz + \nabla(\eta - z) \cdot (\rho C \mathbf{v} + \mathbf{J})_{z=\eta} \\ &\quad + \nabla(H + z) \cdot (\rho C \mathbf{v} + \mathbf{J})_{z=-H} \end{aligned} \quad (18.56b)$$

where we wrote $\mathbf{J} = \mathbf{J}_h + \hat{\mathbf{z}} J^z$. The surface terms ($z = \eta$) combine with the $\partial_t \eta$ term appearing in equation (18.55) to yield

$$\rho C \left[\frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta - w \right] = C Q_m, \quad (18.57)$$

where we used the surface kinematic boundary condition (17.74) to introduce the surface boundary mass flux Q_m . The bottom kinematic boundary condition eliminates the advective contribution at the bottom, $z = -H$, via the no normal flow condition (17.38)

$$\nabla(H + z) \cdot \mathbf{v} = 0. \quad (18.58)$$

Finally, when integrated over the horizontal extent of the domain, the horizontal convergence term from equation (18.56b) vanishes. The reason it vanishes is because either the thickness of fluid vanishes at the horizontal boundaries (as along a beach); there is a no flux boundary condition if the boundary is a vertical wall; or the domain is periodic.

18.5.2 Summarizing the result

Bringing the results together yields the budget equation

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = \int_{z=\eta} (C Q_m + \nabla(\eta - z) \cdot \mathbf{J}) dA + \int_{z=-H} \nabla(H + z) \cdot \mathbf{J} dA. \quad (18.59)$$

We now use the identity (17.66e) between horizontal area element, $dA = dx dy$, and surface area element

$$\hat{\mathbf{n}} d\mathcal{S} = -\nabla(\eta - z) dA \quad \text{at } z = \eta \quad (18.60a)$$

$$\hat{\mathbf{n}} d\mathcal{S} = -\nabla(H + z) dA \quad \text{at } z = -H, \quad (18.60b)$$

to write

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho C dV \right] = - \int_{z=\eta} (-C Q_m + \hat{\mathbf{n}} \cdot \mathbf{J}) d\mathcal{S} - \int_{z=-H} \hat{\mathbf{n}} \cdot \mathbf{J} d\mathcal{S}, \quad (18.61)$$

where

$$Q_m dA = Q_m d\mathcal{S} = -\rho (\mathbf{v} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (18.62)$$

according to equation (17.71c). The budget for fluid mass is realized by setting C to a constant and thus dropping the subgrid flux

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho dV \right] = \int_{z=\eta} \mathcal{Q}_m d\mathcal{S} = - \int_{z=\eta} \rho (\mathbf{v} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (18.63)$$

The manipulations in this section have succeeded in bringing the tracer and mass budgets into the form of the Leibniz-Reynolds transport theorem (18.52). The process of doing so required far more tedium as compared to the elegance of merely starting from equation (18.52). Even so, our efforts provide a useful means to ground the formalism by unpacking the many steps summarized by Leibniz-Reynolds.

18.6 Comments on local conservation laws

We close this chapter by offering some general comments on conservation laws for field quantities, such as tracers, that satisfy Eulerian flux-form conservations laws in the form

$$\frac{\partial(\rho C)}{\partial t} = -\nabla \cdot (\rho \mathbf{v} C + \mathbf{J}). \quad (18.64)$$

In this chapter we saw how this differential equation leads to finite volume conservation properties for the mass, $\int_{\mathcal{R}} C \rho dV$, of C -matter within the region \mathcal{R} . This result is a consequence of the local conservation of matter within the fluid. That is, the amount of matter changes at a point only through the local convergence of matter fluxes onto that point. There is no matter evolution arising from a non-local transport or a source/sink term that cannot be expressed as the convergence of a flux. The presence of such terms might be relevant for reactions, whereby matter is converted to other forms. But that process is not contained in the conservation law (18.64).

Conservation laws of the form (18.64) recur in our discussion of other properties in this book, such as mechanical energy (Chapter 22), absolute vorticity (Section 51.1.3) and potential vorticity (Chapter 50). These laws are consistent with basic notions of causality and locality that appear throughout physics. A lucid discussion of such conservation laws is offered in Section 27-1 in Volume II of the [Feynman Lectures](#).

19

Incompressible flow

In this chapter, we specialize the general kinematics from Chapters 16 and 17 to the case of an incompressible fluid. The velocity field for an incompressible fluid has zero divergence so that it can be written as the curl of a vector streamfunction. The streamfunction plays a central role for incompressible fluid kinematics.

READER'S GUIDE TO THIS CHAPTER

Spatial positions and trajectories are represented in this chapter using Cartesian coordinates to simplify the maths. Nonetheless, the results hold for general coordinates by making use of general covariance as detailed in Chapters 8 and 9. We presume an understanding of the kinematics of mass conservation from Chapter 17. This is a relatively brief chapter, and yet it introduces many concepts and tools of use for the remainder of the book.

19.1	Introduction to incompressible flow	266
19.2	Kinematic boundary conditions	266
19.3	Kinematic free surface equation	267
19.4	Streamfunction for two-dimensional flow	268
19.4.1	Streamfunction isolines are streamlines	268
19.4.2	Streamfunction is constant on material boundaries	268
19.4.3	Transport between two points	269
19.4.4	Gauge symmetry	269
19.5	Vector streamfunction for three-dimensional flow	270
19.5.1	Connecting to fluid transport through a surface	270
19.5.2	Gauge symmetry	270
19.5.3	Vector streamfunction in the vertical gauge	271
19.6	Evolution of material volume and area	272
19.7	Meridional-depth overturning circulation	272
19.8	Exercises	274

19.1 Introduction to incompressible flow

For many applications in geophysical fluid mechanics, we can make a simplifying assumption regarding the fluid kinematics. For the ocean, the Boussinesq approximation is well maintained (see Section 28.1), whereby the volume of a fluid element is constant. Recalling the expression

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v} \quad (19.1)$$

from Section 20.4.1, we see that a constant volume for a fluid element constrains the velocity field to be non-divergent

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = 0 \Rightarrow \nabla \cdot \mathbf{v} = 0 \quad \text{incompressible.} \quad (19.2)$$

A slightly less onerous constraint arises from the anelastic approximation, whereby

$$\nabla \cdot (\rho \mathbf{v}) = 0. \quad (19.3)$$

The anelastic approximation is sometimes motivated for the atmosphere. However, it is less commonly used for atmospheric dynamics than the Boussinesq approximation is used for the ocean. We thus focus the following on the incompressible case with $\nabla \cdot \mathbf{v} = 0$.

The non-divergence constraint reduces by one the number of functional degrees of freedom possessed by the velocity field. What that means in practice is that we need one fewer velocity component to determine the flow. That is, one velocity component is specified by the other components. This property manifests by our ability to introduce a streamfunction to specify the velocity.

19.2 Kinematic boundary conditions

For incompressible flow, there are slight modifications to the compressible boundary conditions detailed in Section 17.4. Whereas the material conditions remain identical, the non-material conditions are applied with a constant reference density, ρ_0 , rather than the local *in situ* density, ρ .

The reason is that we switch from specifying a mass transport condition as per equation (17.61) to a volume transport condition

$$\rho_0 (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = -Q_m d\mathcal{S}. \quad \text{moving non-material boundary condition.} \quad (19.4)$$

Correspondingly, the kinematic boundary condition (17.74) applied at the ocean free surface takes on the form

$$\rho_0 \frac{D(z - \eta)}{Dt} = -Q_m \Rightarrow w + \frac{Q_m}{\rho_0} = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta. \quad (19.5)$$

19.3 Kinematic free surface equation

We now derive an equation for the volume budget over a column of fluid. This equation provides a kinematic expression for the free surface evolution in an incompressible fluid. For this purpose, we vertically integrate the incompressibility constraint, $\nabla \cdot \mathbf{v} = 0$, over the depth of an ocean column, from $z = -H(x, y)$ at the bottom to $z = \eta(x, y, t)$ at the free surface and use the bottom and surface kinematic boundary conditions. This calculation yields

$$0 = \int_{-H}^{\eta} \nabla \cdot \mathbf{v} dz \quad (19.6a)$$

$$= w(\eta) - w(-H) + \int_{-H}^{\eta} \nabla \cdot \mathbf{u} dz \quad (19.6b)$$

$$= w(\eta) - w(-H) + \nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right] - \mathbf{u}(\eta) \cdot \nabla \eta - \mathbf{u}(-H) \cdot \nabla H \quad (19.6c)$$

$$= [w(\eta) - \mathbf{u}(\eta) \cdot \nabla \eta] - [w(-H) + \mathbf{u}(-H) \cdot \nabla H] + \nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right], \quad (19.6d)$$

where we made use of Leibniz's Rule to move the horizontal divergence outside of the integral. We now make use of the surface kinematic boundary condition (19.5) and the bottom no-flow condition

$$w(\eta) - \mathbf{u} \cdot \nabla \eta = -\frac{Q_m}{\rho_0} + \frac{\partial \eta}{\partial t} \quad z = \eta \quad (19.7a)$$

$$w = -\mathbf{u} \cdot \nabla H \quad z = -H \quad (19.7b)$$

to render the free surface equation for an incompressible fluid

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}, \quad (19.8)$$

where

$$\mathbf{U} = \int_{-H}^{\eta} \mathbf{u} dz \quad (19.9)$$

is the depth integrated horizontal transport. For the special case of a steady state with zero boundary flux, the depth integrated flow is non-divergent

$$\nabla \cdot \mathbf{U} = 0 \quad \text{if } Q_m = 0 \text{ and } \partial \eta / \partial t = 0. \quad (19.10)$$

Comparing the incompressible free surface quation (19.8) to the compressible free surface equation (17.79) indicates that the incompressible case is missing a contribution from the material

changes in density. These changes arise from mixing and boundary fluxes of buoyancy. The absence of an impact from surface buoyancy fluxes means that the free surface in an incompressible fluid is not impacted by global thermal expansion, such as that arising from ocean warming. [Greatbatch \(1994\)](#) and [Griffies and Greatbatch \(2012\)](#) provide a recipe for diagnostically addressing this formulational limitation, thus enabling the study of global mean sea level with Boussinesq (incompressible) ocean models.

19.4 Streamfunction for two-dimensional flow

Vertical stratification and rotation inhibit vertical motion in geophysical flows. Therefore, as an idealization it is useful to assume the flow is horizontal (two-dimensional) and non-divergent. The incompressible constraint for two-dimensional flow can be satisfied by writing the horizontal velocity in the form

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla\psi = -\hat{\mathbf{x}} \frac{\partial\psi}{\partial y} + \hat{\mathbf{y}} \frac{\partial\psi}{\partial x}. \quad (19.11)$$

The constraint $\nabla \cdot \mathbf{u} = 0$ is satisfied since the partial derivative operators commute

$$\frac{\partial^2\psi}{\partial x \partial y} = \frac{\partial^2\psi}{\partial y \partial x}. \quad (19.12)$$

We refer to ψ as the *streamfunction*, with this name motivated by the following considerations.

19.4.1 Streamfunction isolines are streamlines

At any fixed time, the total differential of the streamfunction is

$$d\psi = \frac{\partial\psi}{\partial x} dx + \frac{\partial\psi}{\partial y} dy \quad (19.13a)$$

$$= v dx - u dy, \quad (19.13b)$$

where the second equality follows from equation (19.11). Instantaneous lines along which ψ is a constant satisfy

$$d\psi = 0 \Rightarrow \frac{dx}{u} = \frac{dy}{v}. \quad (19.14)$$

Furthermore, the normal direction to constant ψ lines

$$\hat{\mathbf{n}} = \frac{\nabla\psi}{|\nabla\psi|} = \frac{v \hat{\mathbf{x}} - u \hat{\mathbf{y}}}{|\mathbf{u}|} \quad (19.15)$$

is normal to the velocity

$$\mathbf{u} \cdot \nabla\psi = u v - v u = 0. \quad (19.16)$$

Consequently, at each time instance, lines of constant ψ are streamlines (see Section 16.7.2 for discussion of streamlines). This property in turn motivates the name *streamfunction*.

19.4.2 Streamfunction is constant on material boundaries

As a corollary to the results from Section 19.4.1, we know that the streamfunction is a spatial constant when evaluated along material boundaries where $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$. This property follows from equation (19.16). We can also see it from

$$0 = \mathbf{u} \cdot \hat{\mathbf{n}} = (\hat{\mathbf{z}} \wedge \nabla\psi) \cdot \hat{\mathbf{n}} = (\hat{\mathbf{n}} \wedge \hat{\mathbf{z}}) \cdot \nabla\psi = \hat{\mathbf{t}} \cdot \nabla\psi, \quad (19.17)$$

where $\hat{\mathbf{t}}$ a unit vector pointing tangent to the boundary. The condition $\hat{\mathbf{t}} \cdot \nabla\psi = 0$ means that ψ is a spatial constant along the boundary. Even though spatially constant, ψ along the boundary is generally a function of time.

19.4.3 Transport between two points

Consider an arbitrary curve in the fluid with endpoints \mathbf{x}_1 and \mathbf{x}_2 . At any particular time instance, the difference in streamfunction between these two points is given by

$$\psi(\mathbf{x}_1) - \psi(\mathbf{x}_2) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} d\psi \quad (19.18a)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \left(dx \frac{\partial\psi}{\partial x} + dy \frac{\partial\psi}{\partial y} \right) \quad (19.18b)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \nabla\psi \cdot d\mathbf{x} \quad (19.18c)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \nabla\psi \cdot \hat{\mathbf{t}} ds. \quad (19.18d)$$

For the final equality we wrote

$$d\mathbf{x} = \hat{\mathbf{t}} ds, \quad (19.19)$$

where

$$ds = |d\mathbf{x}| \quad (19.20)$$

is the arc-length element along the curve, and $\hat{\mathbf{t}}$ is a unit vector that points in the direction along the curve from \mathbf{x}_1 to \mathbf{x}_2 . Now introduce the normal vector along the curve according to

$$\hat{\mathbf{t}} = \hat{\mathbf{n}} \wedge \hat{\mathbf{z}}, \quad (19.21)$$

which then allows us to write

$$\psi(\mathbf{x}_1) - \psi(\mathbf{x}_2) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} \nabla\psi \cdot (\hat{\mathbf{n}} \wedge \hat{\mathbf{z}}) ds \quad (19.22a)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} (\hat{\mathbf{z}} \wedge \nabla\psi) \cdot \hat{\mathbf{n}} ds \quad (19.22b)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{u} \cdot \hat{\mathbf{n}} ds. \quad (19.22c)$$

The final equality is an expression for the net area transport of fluid normal to the curve. As the chosen curve connecting the points is arbitrary, we conclude that the difference in streamfunction values between two points measures the transport across any curve connecting the points. Correspondingly, the stronger the gradient in the streamfunction, the larger the transport since

$$|\mathbf{u}| = |\nabla\psi|. \quad (19.23)$$

19.4.4 Gauge symmetry

For a two-dimensional non-divergent flow, the constraint $\nabla \cdot \mathbf{u} = 0$ reduces the functional degrees of freedom from two (the two velocity components (u, v)) to one (the streamfunction). However, the streamfunction is arbitrary up to a constant, k , since

$$\psi' = \psi + k \Rightarrow \mathbf{u}' = \mathbf{u}. \quad (19.24)$$

So the value of the streamfunction at a particular point has no unambiguous physical meaning. Rather, only the difference in streamfunction between two points is physically relevant. The ability to add a constant to the streamfunction is termed a *gauge symmetry*. We return to gauge symmetry for the three dimensional flow in Section 19.5.2.

19.5 Vector streamfunction for three-dimensional flow

A three-dimensional non-divergent velocity, $\nabla \cdot \mathbf{v} = 0$, can be specified by a vector streamfunction

$$\mathbf{v} = \nabla \wedge \Psi. \quad (19.25)$$

The constraint $\nabla \cdot \mathbf{v} = 0$ is trivially satisfied since the divergence of the curl vanishes

$$\nabla \cdot (\nabla \wedge \Psi) = 0. \quad (19.26)$$

19.5.1 Connecting to fluid transport through a surface

In a three-dimensional fluid, the volume transport of fluid across a surface is defined by the area integral

$$\mathcal{T}(\mathcal{S}) = \int_{\mathcal{S}} \mathbf{v} \cdot \hat{\mathbf{n}} \, d\mathcal{S}, \quad (19.27)$$

where $\hat{\mathbf{n}}$ is the outward unit normal vector on the surface. Introducing the vector streamfunction and making use of Stokes' Theorem (Section 4.6) then leads to

$$\mathcal{T}(\mathcal{S}) = \int_{\mathcal{S}} \mathbf{v} \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \int_{\mathcal{S}} (\nabla \wedge \Psi) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = \oint_{\partial\mathcal{S}} \Psi \cdot \hat{\mathbf{t}} \, ds, \quad (19.28)$$

where $\hat{\mathbf{t}} \, ds$ is the arc-distance increment oriented along the boundary of \mathcal{S} . Hence, the transport through the surface depends only on the vector streamfunction on the perimeter of the surface. Furthermore, if the transport through the surface vanishes (e.g., no-flux material surface such as the ocean bottom), then on the surface the vector streamfunction can be written as the gradient of an arbitrary scalar, $\Psi = \nabla\phi$, since

$$\oint_{\partial\mathcal{S}} \Psi \cdot \hat{\mathbf{t}} \, ds = \oint_{\partial\mathcal{S}} \nabla\phi \cdot \hat{\mathbf{t}} \, ds = \oint_{\partial\mathcal{S}} \nabla\phi \cdot d\mathbf{x} = \oint_{\partial\mathcal{S}} d\phi = 0. \quad (19.29)$$

19.5.2 Gauge symmetry

For three-dimensional non-divergent flow, the constraint $\nabla \cdot \mathbf{v} = 0$ reduces the three functional degrees of freedom down to the two available from the vector streamfunction $\mathbf{v} = \nabla \wedge \Psi$. Gauge symmetry manifests through the ability to add the gradient of an arbitrary function to Ψ without altering \mathbf{v} :

$$\Psi' = \Psi + \nabla\lambda \Rightarrow \mathbf{v}' = \mathbf{v}, \quad (19.30)$$

which follows since $\nabla \wedge \nabla\lambda = 0$. Hence, the vector streamfunction has no absolute physical meaning since we can always modify the streamfunction by adding an arbitrary gauge function.

19.5.3 Vector streamfunction in the vertical gauge

We can exploit gauge symmetry to determine the vector streamfunction that is related to the horizontal volume transport of fluid beneath a chosen depth. For this purpose, write the component expression for the velocity

$$u = \frac{\partial \Psi_3}{\partial y} - \frac{\partial \Psi_2}{\partial z} \quad v = \frac{\partial \Psi_1}{\partial z} - \frac{\partial \Psi_3}{\partial x} \quad w = \frac{\partial \Psi_2}{\partial x} - \frac{\partial \Psi_1}{\partial y}, \quad (19.31)$$

where the Cartesian components to the vector streamfunction are

$$\Psi = \hat{x} \Psi_1 + \hat{y} \Psi_2 + \hat{z} \Psi_3. \quad (19.32)$$

We follow studies of ocean mesoscale eddy parameterizations (e.g., see Sections 37.6.1 and 39.1) by choosing the *vertical gauge* whereby

$$\Psi_3 = 0 \iff \text{vertical gauge} \quad (19.33)$$

so that

$$u = -\frac{\partial \Psi_2}{\partial z} \quad v = \frac{\partial \Psi_1}{\partial z} \quad w = \frac{\partial \Psi_2}{\partial x} - \frac{\partial \Psi_1}{\partial y}. \quad (19.34)$$

Vertically integrating the u, v equations from the bottom at $z = -H(x, y)$ up to an arbitrary geopotential leads to

$$\Psi = \int_{-H}^z \mathbf{u} dz' \wedge \hat{z} \equiv \mathbf{U}(z) \wedge \hat{z}, \quad (19.35)$$

where $\mathbf{U}(z) = \int_{-H}^z \mathbf{u} dz'$ is the horizontal transport of fluid beneath a chosen geopotential. By construction $u = -\partial \Psi_2 / \partial z$ and $v = \partial \Psi_1 / \partial z$. We verify that this streamfunction also renders w through noting that

$$\frac{\partial \Psi_2}{\partial x} = - \int_{-H}^z \partial_x u dz' - u(-H) \partial_x H \quad \text{and} \quad \frac{\partial \Psi_1}{\partial y} = \int_{-H}^z \partial_y v dz' + v(-H) \partial_y H, \quad (19.36)$$

so that

$$\frac{\partial \Psi_2}{\partial x} - \frac{\partial \Psi_1}{\partial y} = - \int_{-H}^z \left[\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right] dz' - \mathbf{u}(-H) \cdot \nabla H \quad (19.37a)$$

$$= \int_{-H}^z \frac{\partial w}{\partial z'} dz' + w(-H) \quad (19.37b)$$

$$= w(z), \quad (19.37c)$$

where we used the bottom kinematic boundary condition (17.38) to write $w + \mathbf{u} \cdot \nabla H = 0$ at $z = -H(x, y)$. Hence, knowledge of the vector streamfunction (19.35) offers a means to compute the three velocity components. Of course, the velocity is needed to compute the vector streamfunction! Even though the logic is circular, we are satisfied that the circle closes to thus indicate self-consistency of the formalism. Furthermore, in some contexts it is more suitable to first compute the transport than the velocity, such as in ocean mesoscale eddy parameterizations (e.g., Chapter 9 of [Griffies \(2004\)](#)). We also note that the transport is a smoother field than the velocity given that it is computed as an integral of the velocity.

We close by noting that the transport through the solid-earth bottom at $z = -H(x, y)$ vanishes according to equation (19.28). We can trivially verify this result for the vertical gauge since

$$\Psi(z = -H) = 0, \quad (19.38)$$

so that $\oint_{\partial S} \Psi \cdot \hat{\mathbf{t}} ds = 0$ on the bottom.

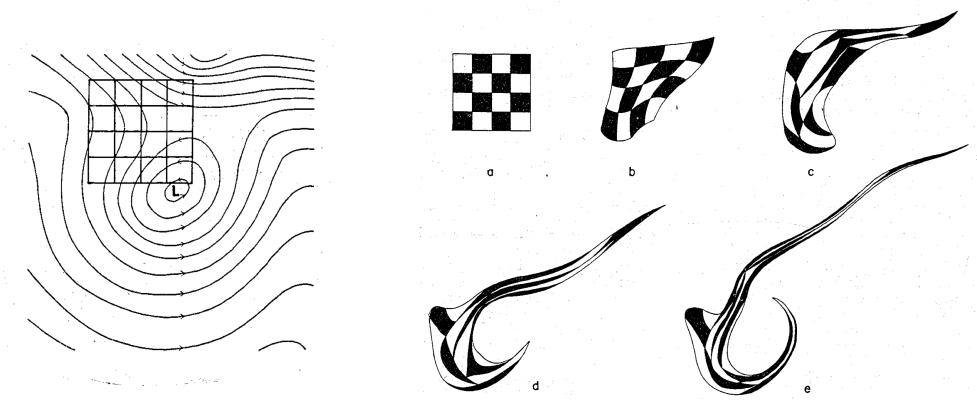


Figure 19.1: This figure is taken from Figure 2 of [Welander \(1955\)](#). It shows the deformation and rotation of a material area in two-dimensional non-divergent flow (left panel) after 6, 12, 24, and 36 hours from a model simulation. As discussed in Section 19.6, the area of a material region remains fixed in two-dimensional non-divergent flow.

19.6 Evolution of material volume and area

As shown by equation (19.2), the volume of a material parcel remains materially constant in an incompressible flow. Correspondingly, a material fluid region maintains a constant volume

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} dV = \int_{\mathcal{R}(\mathbf{v})} \frac{D(\delta V)}{Dt} = \int_{\mathcal{R}(\mathbf{v})} (\nabla \cdot \mathbf{v}) dV = 0. \quad (19.39)$$

Likewise, following from the material area element equation (20.43), the area of a material region in a two-dimensional incompressible flow remains materially constant

$$\frac{d}{dt} \int_{\mathcal{S}(\mathbf{v})} dA = \int_{\mathcal{S}(\mathbf{v})} \frac{D(\delta A)}{Dt} = \int_{\mathcal{S}(\mathbf{v})} (\nabla \cdot \mathbf{u}) dV = 0. \quad (19.40)$$

This area preservation property is illustrated in Figure 19.1, in which a two-dimensional flow is seen to deform a black/white grid, yet to retain a fixed area.

19.7 Meridional-depth overturning circulation

Fluid flow in the atmosphere and ocean is three-dimensional. However, it is sometimes useful to summarize aspects of that flow by integrating the mass transport over one of the directions. A common approach is to integrate over the zonal direction either between two solid-wall boundaries or over a periodic domain. Doing so leaves a two-dimensional transport in the (y, z) plane known as the meridional-depth overturning circulation

$$\mathcal{V}^\rho = \int_{x_1}^{x_2} \rho v dx \quad \mathcal{W}^\rho = \int_{x_1}^{x_2} \rho w dx. \quad (19.41)$$

We can go even further for incompressible flow, or for steady state compressible flow, in which case we introduce a streamfunction for the meridional-depth circulation. For definiteness, consider incompressible flow and the zonal integrated velocity

$$\mathcal{V} = \int_{x_1}^{x_2} v dx \quad \mathcal{W} = \int_{x_1}^{x_2} w dx. \quad (19.42)$$

For a non-periodic basin domain (e.g., North Atlantic Ocean), let the zonal bounds, x_1 and x_2 , be well within rock (where the velocity vanishes) so that we encompass the full basin without having any meridional or depth dependence to the zonal bounds. Taking the meridional derivative of the meridional transport leads to

$$\frac{\partial \mathcal{V}}{\partial y} = \frac{\partial}{\partial y} \left[\int_{x_1}^{x_2} v \, dx \right] \quad (19.43a)$$

$$= \int_{x_1}^{x_2} \frac{\partial v}{\partial y} \, dx \quad (19.43b)$$

$$= - \int_{x_1}^{x_2} \left[\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right] \, dx \quad (19.43c)$$

$$= - \int_{x_1}^{x_2} \frac{\partial w}{\partial z} \, dx \quad (19.43d)$$

$$= - \frac{\partial}{\partial z} \left[\int_{x_1}^{x_2} w \, dx \right] \quad (19.43e)$$

$$= - \frac{\partial \mathcal{W}}{\partial z}. \quad (19.43f)$$

These results show that the two-dimensional zonally integrated transport is non-divergent

$$\frac{\partial \mathcal{V}}{\partial y} + \frac{\partial \mathcal{W}}{\partial z} = 0. \quad (19.44)$$

Consequently, we can introduce a meridional-depth streamfunction

$$\Psi(y, z, t) = - \int_{-H_{\max}}^z \mathcal{V} \, dz' \quad (19.45)$$

to specify the flow. The lower limit of $z = -H_{\max}$ is a constant specified by the maximum depth in the domain, with zero transport for regions below the fluid. To verify Ψ is a streamfunction, we compute

$$\frac{\partial \Psi}{\partial z} = -\mathcal{V} \quad (19.46)$$

and

$$\frac{\partial \Psi}{\partial y} = - \frac{\partial}{\partial y} \left[\int_{-H_{\max}}^z \mathcal{V} \, dz' \right] \quad (19.47a)$$

$$= - \int_{-H_{\max}}^z \frac{\partial \mathcal{V}}{\partial y} \, dz' \quad (19.47b)$$

$$= \int_{-H_{\max}}^z \frac{\partial \mathcal{W}}{\partial z} \, dz' \quad (19.47c)$$

$$= \mathcal{W}. \quad (19.47d)$$

To reach this result, we made use of the non-divergent condition (19.44), and set

$$\mathcal{W}(z = -H_{\max}) = 0. \quad (19.48)$$

An idealized version of the meridional-depth circulation is shown in Figure 19.2.

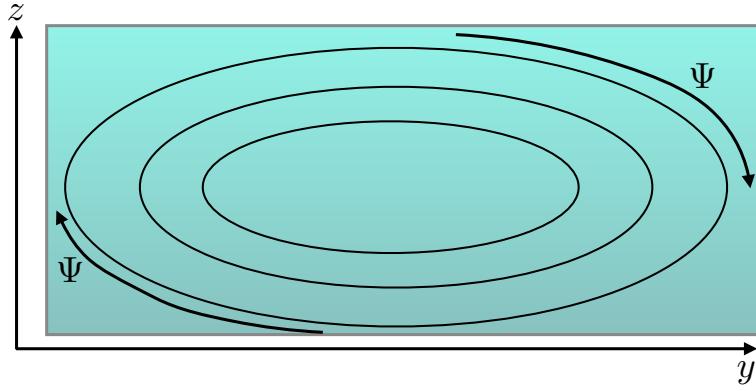


Figure 19.2: An idealized rendition of the meridional-depth overturning circulation found in both the atmosphere and ocean. Shown here are streamlines for the zonally integrated flow between two solid boundaries or over a zonally periodic domain. The flow is assumed to be non-divergent, as per equation (19.44). In the upper reaches of the fluid, flow moves northward (positive y), with downward motion as it reaches the northern boundary, then southward motion at depth and eventual return towards the surface near the southern boundary.

19.8 Exercises

EXERCISE 19.1: STREAMLINES FOR CELLULAR FLOW

Sketch the velocity field for this streamfunction

$$\psi(x, y) = A \sin(k x) \sin(l y), \quad (19.49)$$

where (k, l) are the zonal and meridional wavenumbers.

EXERCISE 19.2: TRANSPORT AND CIRCULATION

Consider a two dimensional velocity field, $\mathbf{u} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}}$, that has both a zero divergence and a zero curl

$$\nabla \cdot \mathbf{u} = \nabla \wedge \mathbf{u} = 0. \quad (19.50)$$

- (a) Show that the circulation of this velocity field through an arbitrary closed loop vanishes.
- (b) Show that the transport of fluid crossing the same closed loop also vanishes.

EXERCISE 19.3: ZERO NET AREA TRANSPORT THROUGH STATIC CLOSED CURVE

For a two-dimensional non-divergent flow, show that there is zero net transport of area crossing an arbitrary static and simply connected closed curve.

EXERCISE 19.4: ZERO NET VOLUME TRANSPORT THROUGH STATIC CLOSED SURFACE

For a three-dimensional non-divergent flow, show that there is zero net transport of volume crossing an arbitrary static and simply connected closed surface within the fluid interior.

EXERCISE 19.5: NET VOLUME TRANSPORT ACROSS AN ARBITRARY SURFACE

Consider a non-divergent flow in a container with static sides/bottom. Draw an arbitrary static surface, \mathcal{S} , within the fluid from one side of the container to the other as in Figure 19.3. Integrate the volume transport over the surface, $\int_{\mathcal{S}} \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S}$. Show that this transport vanishes. That is, the net transport across the surface is zero. Specialize this result to a horizontal surface so that we see there is zero area integrated vertical transport across the surface, $\int_{\mathcal{S}} w dx dy = 0$. Discuss these results. Note: see Section 21.3.7 for the more general case of a non-static surface.

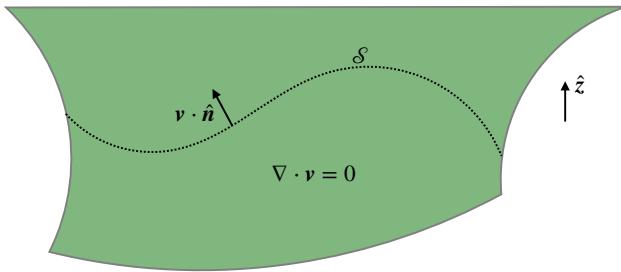


Figure 19.3: Schematic for exercise 19.5, whereby we show that the net flow vanishes across a static surface, \mathcal{S} , that extends from one boundary to the other within an incompressible fluid.

EXERCISE 19.6: SOLID BODY ROTATION

Consider a velocity field corresponding to a time-independent solid-body rotation on a plane

$$\mathbf{u} = \Omega \hat{\mathbf{z}} \wedge \mathbf{x} = \Omega (-y \hat{\mathbf{x}} + x \hat{\mathbf{y}}), \quad (19.51)$$

where Ω is a constant rotation rate.

- (a) Compute the relative vorticity $\nabla \wedge \mathbf{u}$.
- (b) Compute the streamfunction $\mathbf{u} = \nabla \wedge (\hat{\mathbf{z}} \psi)$. Draw streamfunction contours (i.e., lines of constant streamfunction).
- (c) Describe the geometry of material lines. Hint: since the velocity field is time-independent, material parcel trajectories are coincident with streamlines.

EXERCISE 19.7: ALTERNATIVE FORM OF MERIDIONAL-DEPTH STREAMFUNCTION

In equation (19.45), we introduced the meridional-depth overturning streamfunction

$$\Psi(y, z, t) = - \int_{-H_{\max}}^z \mathcal{V} dz'. \quad (19.52)$$

Show that an alternative streamfunction is given by

$$\Gamma(y, z, t) = \int_{y_s}^y \mathcal{W} dy', \quad (19.53)$$

where y_s is a constant latitude just south of the southern-most latitude in the domain.

EXERCISE 19.8: VOLUME TRANSPORT THROUGH STREAMTUBE ENDS

Recall our discussion of streamtubes in Section 16.7.2 (see in particular Figure 16.5). For a steady non-divergent three-dimensional flow, show that the volume transport (volume per time) through the two streamtube ends balances

$$\int_{\mathcal{S}_1} \mathbf{v} \cdot \hat{\mathbf{n}}_1 d\mathcal{S} + \int_{\mathcal{S}_2} \mathbf{v} \cdot \hat{\mathbf{n}}_2 d\mathcal{S} = 0, \quad (19.54)$$

where $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are the outward normals at the two end caps \mathcal{S}_1 and \mathcal{S}_2 . Since the end caps have oppositely directed outward normals, equation (19.54) says that the volume transport entering one streamtube end equals to that leaving the other end. Furthermore, the area of the streamtube is inversely proportional to the local normal velocity, so that flow speeds up when moving through a narrower region of the tube.

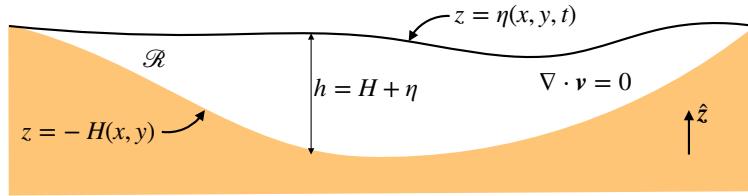


Figure 19.4: Schematic for exercise 19.9 with $z = \eta(x, y, t)$ the free surface at the top of the fluid. This exercise shows that the area integrated time tendency for the free surface vanishes in the absence of mass transport across the free surface.

EXERCISE 19.9: AREA AVERAGE OF FREE SURFACE TIME TENDENCY

Consider an incompressible fluid bounded by a free upper surface and a solid bottom. Let $z = -H(x, y)$ be the vertical position of the static bottom, and $z = \eta(x, y, t)$ be the position of the transient free surface, so that the thickness of the layer is $h = H + \eta$ (see Figure 19.4). The horizontal extent of the layer is a function of time, and is defined by a vanishing thickness $h = H + \eta = 0$ (e.g., ocean water reaching the shoreline). Assume no material crosses either the surface or bottom boundaries, so that both boundaries are material surfaces. Show that the free surface has a time derivative, $\partial\eta/\partial t$, whose area average vanishes. Discuss this result.

20

Material fluid objects

Any surface placed in a continuous fluid separates the matter on the two sides of the surface. What are the kinematic equations describing motion of that surface? In this chapter we refine our kinematic understanding of the movement of material objects such as lines, areas, and volumes moving within a continuous fluid. The discussion moves seamlessly between Lagrangian and Eulerian descriptions, with the two offering complementary insights. Some attention is given to the kinematics of two-dimensional flow due to the relative mathematical ease and the associated intuition useful for general geophysical flows.

READER'S GUIDE FOR THIS CHAPTER

This chapter builds from the kinematics of Chapter 16. It is of fundamental interest to a variety of kinematic aspects of fluid motion, with particular application to the kinematics of mixing and stirring of trace matter in eddying geophysical fluids. The discussion is restricted to Cartesian tensors to reduce the math overhead. Consequently, all tensor labels are downstairs with no distinction between covariant and contravariant. Extension to arbitrary coordinates follow the rules of general covariance detailed in Chapters 8 and 9.

20.1	Increments in material and position space	278
20.1.1	Eulerian/position space differential	278
20.1.2	Lagrangian/material space differential	279
20.1.3	Duality between Eulerian and Lagrangian perspectives	279
20.2	Evolution of a material line element	280
20.2.1	Deformation gradient tensor	280
20.2.2	Cauchy-Green strain tensor	280
20.2.3	Material evolution of a line element	281
20.2.4	Velocity gradient tensor	281
20.2.5	Stretching and tilting of a material line element	282
20.2.6	Rate of strain tensor	282
20.2.7	Rotation tensor	283
20.2.8	Comments and further study	284
20.3	Evolution of a material area element	284
20.3.1	Material area in three-dimensional flow	284
20.3.2	Material area in two-dimensional flow	285
20.4	Volume and the Jacobian of transformation	286
20.4.1	Material parcel volume	286
20.4.2	Evolution of the Jacobian of transformation	287
20.5	Kinematics of two-dimensional flow	287
20.5.1	Diverging flow	288
20.5.2	Rotational flow with nonzero vorticity	288
20.5.3	Flow with nonzero deformation	289
20.5.4	Further study	290
20.6	Exercises	290

20.1 Increments in material and position space

In this section we summarize the mathematics associated with the differential increment of a function, exploring the increment in both position/Eulerian space and in material/Lagrangian space. We make use of the resulting dual expressions throughout this chapter.

20.1.1 Eulerian/position space differential

In Section 16.4.1, we considered the space-time increment of a function. Here we consider just the space increment, as defined by the differential increment of a function evaluated at the same time but at two infinitesimally close points in space

$$d\Phi(\mathbf{x}, t) = \Phi(\mathbf{x} + d\mathbf{x}, t) - \Phi(\mathbf{x}, t) \quad (20.1a)$$

$$= (d\mathbf{x} \cdot \nabla)\Phi. \quad (20.1b)$$

The operator

$$d\mathbf{x} \cdot \nabla = dx_m \frac{\partial}{\partial x_m} \quad (20.2)$$

is a scalar since it remains form invariant when switching to another set of Cartesian position coordinates.¹

¹This form invariance also holds when using curvilinear coordinates if we make use of the general tensor analysis formalism of Chapter 13.

20.1.2 Lagrangian/material space differential

Consider the same function Φ evaluated on a material particle trajectory, and write this “Lagrangian” function as

$$\Phi^L(\mathbf{a}, t) = \Phi[\mathbf{X}(\mathbf{a}, t), t]. \quad (20.3)$$

In words, the Lagrangian version of a function is obtained by evaluating that function on a fluid particle trajectory. We use the notation $\Phi^L(\mathbf{a}, t)$ as a shorthand, which is defined by this equality.

Consider an infinitesimal increment of $\Phi^L(\mathbf{a}, t)$ within material coordinate space. This increment represents the difference of Φ when evaluated on two separate fluid particles labelled by \mathbf{a} and $\mathbf{a} + \delta\mathbf{a}$. Note that we use the δ symbol to signal material increments. Taking a Taylor series and truncating to leading order yields

$$\delta\Phi^L(\mathbf{a}, t) = \Phi[\mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t), t] - \Phi[\mathbf{X}(\mathbf{a}, t), t] \quad (20.4a)$$

$$= \Phi^L(\mathbf{a} + \delta\mathbf{a}, t) - \Phi^L(\mathbf{a}, t) \quad (20.4b)$$

$$= (\delta\mathbf{a} \cdot \nabla_{\mathbf{a}})\Phi^L(\mathbf{a}, t). \quad (20.4c)$$

The operator

$$\delta\mathbf{a} \cdot \nabla_{\mathbf{a}} = \delta a_j \frac{\partial}{\partial a_j} \quad (20.5)$$

is a scalar since it remains form invariant when switching to another set of Cartesian material coordinates.² We use the notation $\nabla_{\mathbf{a}}$ to emphasize that the gradient operator is in material space rather than position space.

20.1.3 Duality between Eulerian and Lagrangian perspectives

By construction, the value of a function at a postion \mathbf{x} (Eulerian perspective) equals to the function evaluated on a moving fluid particle (Lagrangian perspective) when the trajectory passes through \mathbf{x} . Mathematically, this identity takes the form

$$\Phi^L(\mathbf{a}, t) = \Phi(\mathbf{x}, t) \quad \text{if } \mathbf{X}(\mathbf{a}, t) = \mathbf{x}. \quad (20.6)$$

Likewise, if the infinitesimal increment in space, $\delta\mathbf{x}$, equals to the vector increment of the two fluid particles,

$$\delta\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, t), \quad (20.7)$$

then the functional increments are identical

$$\delta\Phi^L(\mathbf{a}, t) = \delta\Phi(\mathbf{x}, t) \quad \text{if } \delta\mathbf{X}(\mathbf{a}, t) = \delta\mathbf{x}, \quad (20.8)$$

where

$$\delta\Phi(\mathbf{x}, t) = \Phi(\mathbf{x} + \delta\mathbf{x}, t) - \Phi(\mathbf{x}, t). \quad (20.9)$$

These identities allow us to develop relations using either a Lagrangian or an Eulerian perspective, and then to interpret them in the complementary perspective. We make routine use of this Eulerian/Lagrangian duality throughout this book.

²As for the position space, this invariance also holds when using curvilinear material coordinates if we make use of the general tensor analysis formalism of Chapter 13.

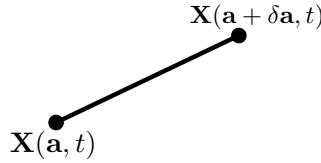


Figure 20.1: The ends of an infinitesimal material line element are defined by the trajectories of two fluid particles, $\mathbf{X}(\mathbf{a}, t)$ and $\mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t)$. All points between the endpoints are part of the material line. Kinematics of the line element are determined by properties of the velocity gradient tensor discussed in Section 20.2.4.

20.2 Evolution of a material line element

Material line elements, and their generalizations to area and volume elements, are geometric objects that follow fluid particles. We initialize a material line element by drawing a line in the fluid and then following the fluid particles contained on the initial line. The material line element is stretched and folded by the fluid flow. We here develop the rudimentary kinematics of material line elements.

20.2.1 Deformation gradient tensor

A material line element is a small line marked in the fluid and whose motion follows that of fluid particles. Assume the line element endpoints are particles with trajectories $\mathbf{X}(\mathbf{a}, t)$ and $\mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t)$ (see Figure 20.1). At time t , the vector displacement between these two particles is written

$$\delta\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, t). \quad (20.10)$$

Expanding this expression to leading order yields

$$\delta\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, t) \approx (\delta\mathbf{a} \cdot \nabla_{\mathbf{a}}) \mathbf{X}(\mathbf{a}, t), \quad (20.11)$$

where again $\nabla_{\mathbf{a}}$ is the gradient operator acting on the material coordinates. Writing this equation in component form leads to

$$\delta X_m = \delta a_j \frac{\partial X_m}{\partial a_j}. \quad (20.12)$$

As in Chapter 16, we assign the labels for m, n, p position/Eulerian coordinates, \mathbf{x} , and the labels i, j, k for material/Lagrangian coordinates, \mathbf{a} .

The components

$$F_{mj} \equiv \frac{\partial X_m}{\partial a_j} \quad (20.13)$$

appearing in equation (20.12) form elements of the transformation matrix linking position space to material space. We already encountered this tensor in Section 16.3.3 (see equation (16.9)). In the continuum mechanics literature, the tensor (20.13) is known as the *displacement gradient tensor* or the *deformation gradient tensor*.

20.2.2 Cauchy-Green strain tensor

The deformation gradient tensor plays a role in measuring the length of material line elements. We see this role by considering the squared length of a line element

$$\delta\mathbf{X} \cdot \delta\mathbf{X} = \frac{\partial \mathbf{X}}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} \delta a_i \delta a_j \equiv C_{ij} \delta a_i \delta a_j. \quad (20.14)$$

Algebraically, this expression is a quadratic form, and the symmetric tensor

$$C_{ij} \equiv \frac{\partial \mathbf{X}}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} = F_{mi} F_{mj} \quad (20.15)$$

is the metric tensor that provides the means to measure distance along an infinitesimal material line element. This metric tensor is called the *Cauchy-Green strain tensor* in the continuum mechanics literature.

20.2.3 Material evolution of a line element

Now consider the material time derivative of the material line element

$$\frac{\partial[\delta\mathbf{X}(\mathbf{a}, t)]}{\partial t} = \frac{\partial\mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t)}{\partial t} - \frac{\partial\mathbf{X}(\mathbf{a}, t)}{\partial t} \quad (20.16a)$$

$$= \mathbf{v}^L(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{v}^L(\mathbf{a}, t) \quad (20.16b)$$

$$\equiv \delta\mathbf{v}^L(\mathbf{a}, t). \quad (20.16c)$$

In these equations, we introduced the Lagrangian velocity

$$\mathbf{v}^L(\mathbf{a}, t) = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (20.17)$$

as per equation (16.31) and the discussion in Section 20.1.2.

As for the line element manipulations in Section 20.2.1, we can massage the expression (20.16c) by performing a Taylor series expansion and truncating to leading order

$$\frac{\partial[\delta\mathbf{X}(\mathbf{a}, t)]}{\partial t} = \delta\mathbf{v}^L(\mathbf{a}, t) = (\delta\mathbf{a} \cdot \nabla_{\mathbf{a}}) \mathbf{v}^L(\mathbf{a}, t). \quad (20.18)$$

Alternatively, we can choose to evaluate this expression using an Eulerian perspective (see Section 20.1.3), in which case

$$\frac{D(\delta\mathbf{x})}{Dt} = \delta\mathbf{v}(\mathbf{x}, t) = (\delta\mathbf{x} \cdot \nabla_{\mathbf{x}}) \mathbf{v}(\mathbf{x}, t). \quad (20.19)$$

20.2.4 Velocity gradient tensor

Writing the Eulerian result (20.19) in component form leads to

$$\frac{D(\delta x_m)}{Dt} = \delta x_n \frac{\partial v_m}{\partial x_n}. \quad (20.20)$$

The derivatives $\partial v_m / \partial x_n$ form components to the second-order *velocity gradient* tensor whose dimensions are inverse time (i.e., a rate). The velocity gradient tensor determines how an infinitesimal line element is deformed as it moves through the fluid.

As with any matrix, a second-order tensor can be decomposed into its symmetric and anti-symmetric components

$$\frac{\partial v_m}{\partial x_n} = \frac{1}{2} \left[\frac{\partial v_m}{\partial x_n} + \frac{\partial v_n}{\partial x_m} \right] + \frac{1}{2} \left[\frac{\partial v_m}{\partial x_n} - \frac{\partial v_n}{\partial x_m} \right] \quad (20.21a)$$

$$\equiv \mathbb{S}_{mn} + \mathbb{A}_{mn}, \quad (20.21b)$$

where³

$$\mathbb{S}_{mn} = \frac{1}{2} \left[\frac{\partial v_m}{\partial x_n} + \frac{\partial v_n}{\partial x_m} \right] = \mathbb{S}_{nm} \quad \text{rate of strain tensor} \quad (20.22a)$$

$$\mathbb{A}_{mn} = \frac{1}{2} \left[\frac{\partial v_m}{\partial x_n} - \frac{\partial v_n}{\partial x_m} \right] = -\mathbb{A}_{nm} \quad \text{rotation tensor.} \quad (20.22b)$$

As seen in the following, these tensors affect the line element evolution in very distinct manners.

20.2.5 Stretching and tilting of a material line element

Consider a line element that is initially aligned with the vertical axis

$$\delta \mathbf{x}_{t=0} = \hat{\mathbf{z}} \delta Z_0. \quad (20.23)$$

Consequently, the initial evolution of this material line element takes on the form

$$\underbrace{\frac{D(\delta x)}{Dt}}_{\text{tilting}} = \delta Z_0 \left[\frac{\partial u}{\partial z} \right] \quad \underbrace{\frac{D(\delta y)}{Dt}}_{\text{tilting}} = \delta Z_0 \left[\frac{\partial v}{\partial z} \right] \quad \underbrace{\frac{D(\delta z)}{Dt}}_{\text{stretching}} = \delta Z_0 \left[\frac{\partial w}{\partial z} \right]. \quad (20.24)$$

In the presence of a vertical derivative in the horizontal velocity field (vertical shear), the first and second terms create a non-zero projection of the line element onto the horizontal plane. That is, these terms *tilt* the line element. Additionally, in the presence of a vertical derivative in the vertical velocity, the line element is expanded or compressed along its initial axis. This term is called *stretching*. We return to the tilting and stretching mechanisms when discussing the dynamics of vorticity in Chapter 48. There, we see that vortex lines in a perfect fluid flow are material lines. Consequently, vortex lines are also affected by tilting and stretching just like a material line.

20.2.6 Rate of strain tensor

Recall the expression (20.14) for the squared length of a line element

$$\delta \mathbf{X} \cdot \delta \mathbf{X} = \frac{\partial \mathbf{X}}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} \delta a_i \delta a_j. \quad (20.25)$$

Its material time derivative is given by

$$\left[\frac{\partial(\delta \mathbf{X} \cdot \delta \mathbf{X})}{\partial t} \right]_a = 2 \frac{\partial \mathbf{v}^L}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} \delta a_i \delta a_j. \quad (20.26)$$

We can express this result using Eulerian \mathbf{x} -coordinates by making use of the duality described in Section 20.1.3, which leads to

$$\frac{\partial \mathbf{v}^L}{\partial a_i} \delta a_i = \frac{\partial \mathbf{v}}{\partial x_n} \delta x_n \quad (20.27a)$$

$$\frac{\partial \mathbf{X}}{\partial a_j} \delta a_j = \delta \mathbf{x}, \quad (20.27b)$$

³The rate of strain tensor, \mathbb{S}_{mn} , is sometimes called the *deformation* tensor in the fluid dynamics literature. We avoid this nomenclature to avoid confusion with the *deformation gradient* tensor defined by equation (20.13).

so that

$$\frac{D(\delta\mathbf{x} \cdot \delta\mathbf{x})}{Dt} = 2 \frac{\partial v_m}{\partial x_n} \delta x_n \delta x_m. \quad (20.28)$$

Since the product $\delta x_n \delta x_m$ is symmetric on the labels m, n , it projects out the symmetric portion of the velocity gradient tensor, thus yielding

$$\frac{1}{2} \frac{D(\delta\mathbf{x} \cdot \delta\mathbf{x})}{Dt} = \mathbb{S}_{mn} \delta x_n \delta x_m. \quad (20.29)$$

Consequently, the rate of strain tensor, \mathbb{S}_{mn} , determines the rate at which a material line element changes its length. When the rate of strain tensor vanishes, then the line element retains a constant length. We can understand the result (20.29) by considering two fluid particles initialized very close together. The distance between the two particles will be modified so long as there are nonzero gradients in the velocity field. The distance between the particles evolves according to the rate of strain tensor as given by equation (20.29).

As a symmetric matrix, the rate of strain tensor can be diagonalized, with the diagonal elements equal to the eigenvalues. Each eigenvalue measures the rate that line elements oriented according to the principle axes (eigenvectors) expand/contract under the impacts of straining motion in the fluid. According to equation (20.29), the expansion/contraction is exponential when aligned along the principle axes, with the exponential rate determined by the strain tensor's eigenvalues. Furthermore, as shown in Section 20.4, the sum of these eigenvectors (trace of the rate of strain tensor) measures the rate that a volume element changes through the divergence of the velocity

$$\mathbb{S}_{mm} = \nabla \cdot \mathbf{v}. \quad (20.30)$$

20.2.7 Rotation tensor

The rotation tensor

$$\mathbb{A}_{mn} = \frac{1}{2} \left[\frac{\partial v_m}{\partial x_n} - \frac{\partial v_n}{\partial x_m} \right] \quad (20.31)$$

is anti-symmetric: $\mathbb{A}_{mn} = -\mathbb{A}_{nm}$. Its components are related to the vorticity vector $\boldsymbol{\omega} = \nabla \wedge \mathbf{v}$ according to

$$\mathbb{A}_{mn} = -\epsilon_{mnp} \omega_p / 2 \iff \mathbb{A} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \quad (20.32)$$

The contribution of the rotation matrix to material evolution of the line element is given by

$$\left[\frac{D(\delta x_m)}{Dt} \right]_{\text{rot}} = \mathbb{A}_{mn} \delta x_n = -(\epsilon_{mnp} \omega_p / 2) \delta x_n \Rightarrow \left[\frac{D(\delta\mathbf{x})}{Dt} \right]_{\text{rot}} = \frac{1}{2} (\boldsymbol{\omega} \wedge \delta\mathbf{x}). \quad (20.33)$$

This relation is in the form of a pure rotation of the vector $\delta\mathbf{x}$ generated by the vector $\boldsymbol{\omega}/2$ (recall the discussion of rotations in Section 13.5). We thus conclude that the anti-symmetric rotation tensor, \mathbb{A}_{mn} , provides a rigid body rotation to a fluid line element (or any infinitesimal fluid region). It rotates the object without altering the size (length, area, volume). We return to this interpretation when discussing vorticity in Section 46.1.

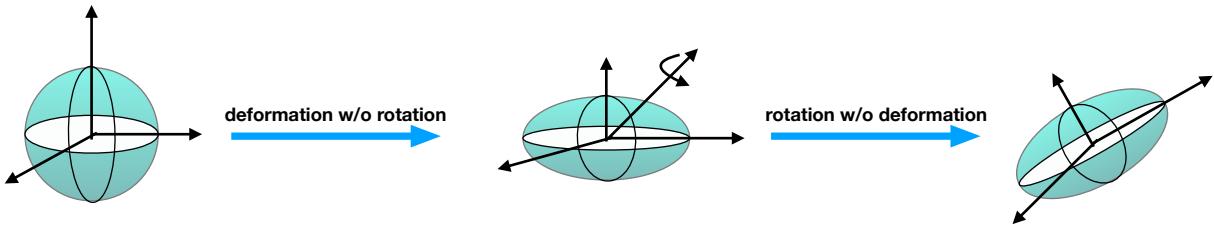


Figure 20.2: Schematic illustrating the decomposition of the manners that fluid flow can modify a material volume. First the sphere can be deformed without rotation, with this process encompassed by the rate of strain tensor, \mathbb{S}_{mn} . Next it can be rigidly rotated without changing its shape, as encompassed by the rotation tensor, \mathbb{A}_{mn} . The axes shown are meant to represent the principle axes of the body, so that deformation of the sphere corresponds to expansion or contraction along the principle axes directions.

20.2.8 Comments and further study

The above discussion of how fluid motion impacts on a material line element falls under the more general insights from the Cauchy-Stokes decomposition theorem. This theorem says that the arbitrary motion of a body can be decomposed into a uniform translation, dilation along three perpendicular axes, plus a rigid body rotation. Mathematically, this decomposition can be written by expanding the equation (20.20) to read

$$v_m(\mathbf{x}, t) = v_m(\mathbf{x}_0, t) + \mathbb{S}_{mn} \delta x_n + \mathbb{A}_{mn} \delta x_n, \quad (20.34)$$

which can be written in the dyadic form

$$\mathbf{v}(\mathbf{x}, t) = \mathbf{v}(\mathbf{x}_0, t) + \mathbb{S} \cdot \delta \mathbf{x} + \mathbb{A} \cdot \delta \mathbf{x}. \quad (20.35)$$

Figure 20.2 illustrates the deformation and rotation portion of this decomposition. A more thorough discussion of these fundamental kinematic notions can be found in Chapter 4 of [Aris \(1962\)](#), with a brief summary in Section 1.1 of [Olbers et al. \(2012\)](#).

20.3 Evolution of a material area element

We extend the discussion of material line elements in Section 20.2 to a material area element such as that shown in Figure 20.3. We consider area elements in both three-dimensional and two-dimensional flows.

20.3.1 Material area in three-dimensional flow

Following from the geometric interpretation of the vector product in Section 3.4.4, we here define a material area element by (see Figure 20.3)

$$\delta \mathbf{S} = \delta \mathbf{A} \wedge \delta \mathbf{B} \Rightarrow \delta \mathcal{S}_m = \epsilon_{mnp} \delta A_n \delta B_p \quad (20.36)$$

where $\delta \mathbf{A}$ and $\delta \mathbf{B}$ are non-parallel infinitesimal material line elements. The area projected onto the direction $\hat{\mathbf{n}}$ is given by

$$\hat{\mathbf{n}} \cdot \delta \mathbf{S} = \hat{\mathbf{n}} \cdot (\delta \mathbf{A} \wedge \delta \mathbf{B}). \quad (20.37)$$

The evolution of the material area element is given by

$$\frac{D(\delta \mathbf{S})}{Dt} = \frac{D(\delta \mathbf{A})}{Dt} \wedge \delta \mathbf{B} + \delta \mathbf{A} \wedge \frac{D(\delta \mathbf{B})}{Dt} \quad (20.38a)$$

$$= [(\delta \mathbf{A} \cdot \nabla) \mathbf{v}] \wedge \delta \mathbf{B} + \delta \mathbf{A} \wedge [(\delta \mathbf{B} \cdot \nabla) \mathbf{v}], \quad (20.38b)$$

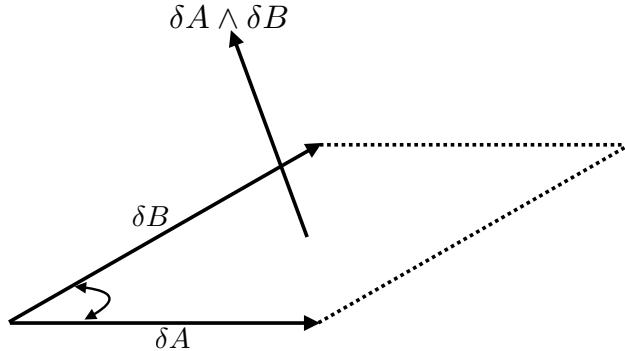


Figure 20.3: A material area defined by the cross product of two material line elements, $\delta\mathcal{S} = \delta\mathbf{A} \wedge \delta\mathbf{B}$. In the special case of $\delta\mathbf{A} = \hat{\mathbf{x}} \delta x$ and $\delta\mathbf{B} = \hat{\mathbf{y}} \delta y$, then $\delta\mathcal{S} = \delta x \delta y \hat{\mathbf{z}}$.

where the second equality made use of the line element evolution equation (20.19). To proceed we expose indices and make use of some tensor identities

$$\frac{D(\delta\mathcal{S}_m)}{Dt} = \epsilon_{mnp} [(\delta A_q \partial_q) v_n] \delta B_p + \epsilon_{mnp} \delta A_n [(\delta B_q \partial_q) v_p] \quad (20.39a)$$

$$= \epsilon_{mnp} [\delta A_q \delta B_p \partial_q v_n + \delta A_n \delta B_q \partial_q v_p] \quad (20.39b)$$

$$= \epsilon_{mnp} \partial_q v_n [\delta A_q \delta B_p - \delta A_p \delta B_q] \quad (20.39c)$$

$$= \epsilon_{mnp} \partial_q v_n \epsilon_{rqp} \delta\mathcal{S}_r \quad (20.39d)$$

$$= (\delta_{mr} \delta_{nq} - \delta_{mq} \delta_{nr}) \partial_q v_n \delta\mathcal{S}_r \quad (20.39e)$$

$$= (\nabla \cdot \mathbf{v}) \delta\mathcal{S}_m - (\partial_m \mathbf{v}) \cdot \delta\mathcal{S}. \quad (20.39f)$$

To reach this result we made use of the following identities

$$\delta A_q \delta B_p - \delta A_p \delta B_q = \epsilon_{rqp} \delta\mathcal{S}_r \quad (20.40a)$$

$$\epsilon_{mnp} \epsilon_{rqp} = \delta_{mr} \delta_{nq} - \delta_{mq} \delta_{nr} \quad (20.40b)$$

where δ_{mn} is the Kronecker (or identity) tensor.

20.3.2 Material area in two-dimensional flow

Now consider a material area element for two-dimensional fluid flow with velocity, $\mathbf{v} = (u, v, 0)$, and $\delta\mathbf{A} = \hat{\mathbf{x}} \delta x$, $\delta\mathbf{B} = \hat{\mathbf{y}} \delta y$, with zero dependence on z . In this case, the area of an infinitesimal material region is

$$\delta\mathcal{S} = (\delta\mathbf{A} \wedge \delta\mathbf{B}) \cdot \hat{\mathbf{z}} = \delta x \delta y, \quad (20.41)$$

and its evolution is given by

$$\frac{D(\delta\mathcal{S})}{Dt} = (\delta\mathbf{B} \wedge \hat{\mathbf{z}}) \cdot (\delta\mathbf{A} \cdot \nabla) \mathbf{u} + (\hat{\mathbf{z}} \wedge \delta\mathbf{A}) \cdot (\delta\mathbf{B} \cdot \nabla) \mathbf{u} \quad (20.42a)$$

$$= \delta x \delta y \nabla \cdot \mathbf{u}, \quad (20.42b)$$

so that

$$\frac{1}{\delta\mathcal{S}} \frac{D(\delta\mathcal{S})}{Dt} = \nabla \cdot \mathbf{u}. \quad (20.43)$$

Hence, the area of the material region evolves according to the divergence of the horizontal velocity. Correspondingly, the area remains constant in a horizontally non-divergent flow. This result follows from specializing the general result (20.39f) to the case of two-dimensional flow with no dependence on the vertical direction.

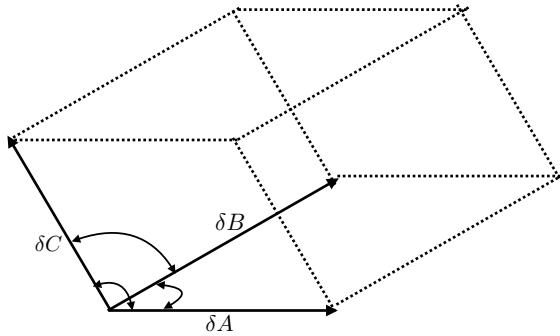


Figure 20.4: A parallelepiped defined by three material line elements, with volume (to within a sign) given by $\delta V = (\delta \mathbf{A} \wedge \delta \mathbf{B}) \cdot \delta \mathbf{C}$. See also the discussion surrounding Figure 3.4.

20.4 Volume and the Jacobian of transformation

The mass of a material parcel is constant. However, the volume is not generally constant, since the fluid density is not generally uniform. We here derive the expression for how volume evolves for a material parcel. We also derive the material evolution equation for the Jacobian of transformation between position space and material space. We will see that the relative change for both the parcel volume and the Jacobian are determined by the divergence of the velocity field.

20.4.1 Material parcel volume

Consider a material region with a volume δV spanned by the infinitesimal material line elements $\delta \mathbf{A}$, $\delta \mathbf{B}$, and $\delta \mathbf{C}$ (see Figure 20.4). To within a sign the volume is given by

$$\delta V = (\delta \mathbf{A} \wedge \delta \mathbf{B}) \cdot \delta \mathbf{C}. \quad (20.44)$$

Making use of the line element evolution equation (20.19) renders

$$\frac{D(\delta V)}{Dt} = (\delta \mathbf{B} \wedge \delta \mathbf{C}) \cdot (\delta \mathbf{A} \cdot \nabla) \mathbf{v} + (\delta \mathbf{C} \wedge \delta \mathbf{A}) \cdot (\delta \mathbf{B} \cdot \nabla) \mathbf{v} + (\delta \mathbf{A} \wedge \delta \mathbf{B}) \cdot (\delta \mathbf{C} \cdot \nabla) \mathbf{v}. \quad (20.45)$$

Now specialize to the case where the parcel is a parallelepiped oriented according to the coordinate axes

$$\delta \mathbf{A} = \hat{\mathbf{x}} \delta x \quad \delta \mathbf{B} = \hat{\mathbf{y}} \delta y \quad \delta \mathbf{C} = \hat{\mathbf{z}} \delta z, \quad (20.46)$$

so that

$$\delta V = \delta x \delta y \delta z. \quad (20.47)$$

Plugging into equation (20.45) leads to

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v}. \quad (20.48)$$

This result is a three-dimensional generalization of the material area equation derived in Section 20.3.2.

We offer an alternative derivation of equation (20.48) in Section 17.1, where no assumptions are made concerning the shape of the material region. That derivation leads us to conclude that the relative volume of a material parcel increases when the parcel moves through a region where the velocity diverges ($\nabla \cdot \mathbf{v} > 0$). We think of a diverging velocity field as “pushing out” the material

parcel boundary, thus increasing its volume. In contrast, the volume of a material parcel decreases where the fluid velocity converges ($\nabla \cdot \mathbf{v} < 0$)

$$\nabla \cdot \mathbf{v} > 0 \Rightarrow \text{material volume increases in diverging flow} \Rightarrow \text{parcel expands} \quad (20.49a)$$

$$\nabla \cdot \mathbf{v} < 0 \Rightarrow \text{material volume decreases in converging flow} \Rightarrow \text{parcel contracts.} \quad (20.49b)$$

20.4.2 Evolution of the Jacobian of transformation

Recall the discussion in Section 16.3.6, where we showed that the Jacobian of transformation between material space (\mathbf{a}, t) and position space (\mathbf{x}, t) is related to the ratio of the volume elements written in the two coordinate systems. In particular, equation (16.21) is given by

$$\delta V(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \delta V(\mathbf{a}) \Rightarrow \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{\delta V(\mathbf{x})}{\delta V(\mathbf{a})}. \quad (20.50)$$

The material coordinate volume element $\delta V(\mathbf{a})$ is time independent when following the flow. Consequently, the material evolution of the Jacobian is given by

$$\frac{D}{Dt} \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{1}{\delta V(\mathbf{a})} \frac{D(\delta V(\mathbf{x}))}{Dt} \quad (20.51a)$$

$$= \frac{\delta V(\mathbf{x})}{\delta V(\mathbf{a})} \nabla \cdot \mathbf{v} \quad (20.51b)$$

$$= \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \nabla \cdot \mathbf{v}. \quad (20.51c)$$

The second equality made use of the result (20.48), which expresses the material time change for the volume of a material fluid parcel, as measured in position space, in terms of the velocity divergence. We thus see that the relative change of the Jacobian is determined by the divergence of the velocity

$$\left[\frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right]^{-1} \frac{D}{Dt} \left[\frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right] = \nabla \cdot \mathbf{v}. \quad (20.52)$$

This equation is identical to the parcel volume equation (20.48), which is expected given the relation between the Jacobian and the parcel volume. In Exercise 20.2, we derive this result using the explicit expression for the Jacobian in terms of the ϵ -tensor.

20.5 Kinematics of two-dimensional flow

In this section we consider the rudiments of two-dimensional flow as a venue to illustrate certain topics presented earlier in this chapter. In so doing we expose kinematic properties commonly used to characterise two-dimensional flow. Generalizations to three-dimensions are available with a bit more mathematical formalism.

Our starting point is Figure 20.5, which shows a square region of fluid exposed to a variety of flow regimes. We can kinematically describe these changes by making use of the velocity gradient tensor introduced in Section 20.2.4, here written for the two-dimensional flow with horizontal velocity components, (u, v)

$$\frac{\partial v_m}{\partial x_n} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \\ \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) & \frac{\partial v}{\partial y} \end{bmatrix} + \frac{\zeta}{2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \mathbb{S}_{mn} + \mathbb{A}_{mn}, \quad (20.53)$$

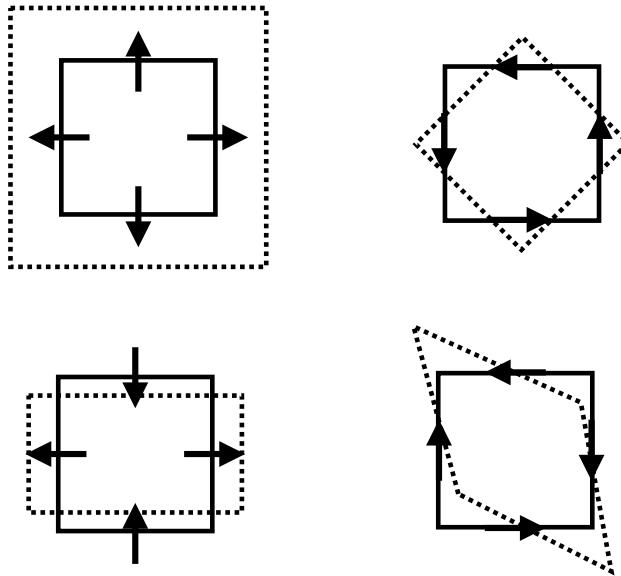


Figure 20.5: Illustrating the varieties of shape changes for a square material fluid region. Upper left: purely divergent flow, whereby $\nabla \cdot \mathbf{u} > 0$ yet with zero vorticity, thus leading to a uniform increase in the area. Upper right: rotational flow with nonzero vorticity, $\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u})$, yet zero divergence, thus leading to a pure rotation of the square patch. Lower left: result of a pure tension/compression straining flow with zero divergence and zero vorticity, leading to compression in one direction and dilation in the orthogonal direction. This flow is realized by the velocity field $\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi$, which is generated by the streamfunction $\psi = -\gamma xy$. Lower right: pure shearing strain flow with zero divergence and zero vorticity, as as realized by the flow $\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi$ generated by the streamfunction $\psi = -(\gamma/2)(x^2 - y^2)$. This figure is based on Figure 2.4 of [Hoskins and James \(2014\)](#).

where

$$\zeta = \hat{\mathbf{z}} \cdot \nabla \wedge \mathbf{v} = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad (20.54)$$

is the vertical component to the vorticity.

20.5.1 Diverging flow

Recall from Section 20.3.2 that the area of a fluid element in two-dimensional flow changes according to the divergence. The upper left panel of Figure 20.5 thus illustrates equation (20.43)

$$\frac{1}{\delta S} \frac{D(\delta S)}{Dt} = \nabla \cdot \mathbf{u} = \mathbb{S}_{mm}, \quad (20.55)$$

where δS is the area and \mathbb{S}_{mm} is the trace of the rate of strain tensor. That is, a diverging flow as depicted by this figure, with $\nabla \cdot \mathbf{u} > 0$, leads to an expansion of the area. The opposite occurs for a converging flow, where the area compresses.

20.5.2 Rotational flow with nonzero vorticity

The upper right panel of Figure 20.5 illustrates the effects from a flow with a non-zero vorticity $\zeta = \partial v / \partial x - \partial u / \partial y$, or a nonzero circulation. We provide extensive discussion of vorticity and circulation in Part VIII of this book. For now we remain satisfied by noting that vorticity measures the spin at a point within a fluid. The nonzero spin imparts a rotation to an area element, in this case bringing about a counter-clockwise rotation. All components of the strain tensor vanish for a purely rotational flow, so that there is no deformation of the square as it rotates.

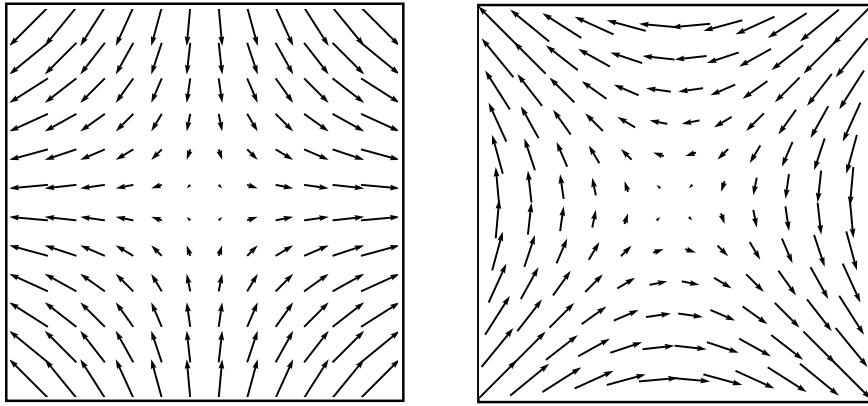


Figure 20.6: Two-dimensional non-divergent and irrotational flow with nonzero deformation/strain. Left panel: pure tension strain as determined by the streamfunction, $\psi = -\gamma xy$, so that the velocity components are $u = -\partial\psi/\partial y = \gamma x$ and $v = \partial\psi/\partial x = -\gamma y$. The vertical axis orients the direction along which flow contracts (compression) whereas the horizontal axis is the dilation axis (tension). Right panel: pure shearing flow as determined by the streamfunction $\psi = -(\gamma/2)(x^2 - y^2)$ so that the velocity components are $u = -\gamma y$ and $v = -\gamma x$. We set $\gamma = 1$ for both examples.

20.5.3 Flow with nonzero deformation

The lower left panel of Figure 20.5 shows the square within a deformational flow whereby it contracts along the y -axis and dilates along the x -axis. This flow is non-divergent, $\nabla \cdot \mathbf{u} = 0$, and has zero vorticity, $\zeta = 0$, so that the area remains constant and the orientation is fixed. However, it has shear that acts to deform the area. More precisely, the non-divergent deformational flow as determined by

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi, \quad (20.56)$$

with the streamfunction, $\psi = -\gamma xy$ where γ is a constant inverse time scale and hence the strength of the strain. The resulting velocity components are $u = -\partial\psi/\partial y = \gamma x$ and $v = \partial\psi/\partial x = -\gamma y$.

There are two combinations of the rate of strain tensor elements that are useful in describing deformational flows:

$$\text{tension strain} = S_T = \mathbb{S}_{11} - \mathbb{S}_{22} = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \quad (20.57a)$$

$$\text{shearing strain} = S_S = 2\mathbb{S}_{12} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}. \quad (20.57b)$$

In the literature, the tension strain and shearing strain are also known as tension and shearing deformation rates. Note that negative tension is known as *compression*. For the deformation flow with streamfunction $\psi = -\gamma xy$, we have

$$S_T = 2\gamma \quad S_S = 0, \quad (20.58)$$

so that this velocity leads to a purely tension straining flow. In contrast, the following non-divergent irrotational flow is a purely shearing strain flow

$$\psi = -(\gamma/2)(x^2 - y^2) \quad u = -\gamma y \quad v = -\gamma x \quad S_T = 0 \quad S_S = -2\gamma, \quad (20.59)$$

as depicted by the right panel of Figure 20.6. This pure shearing flow leads to the deformation of the fluid square shown in the lower right panel of Figure 20.5.

20.5.4 Further study

Elements of this section can be found in Section 2.3 of [Hoskins and James \(2014\)](#). More detailed examinations of two-dimensional flow kinematics are offered by [Weiss \(1991\)](#) and [Lilly \(2018\)](#).

20.6 Exercises

EXERCISE 20.1: VELOCITY FIELD WITH ZERO STRAIN ([Aris \(1962\)](#) EXERCISE 4.41.1)

If the rate of strain tensor vanishes, show that the velocity field can be written

$$\mathbf{v} = \mathbf{U} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (20.60)$$

where $\boldsymbol{\Omega}$ is a constant angular rotation rate and \mathbf{U} is a constant velocity. That is, a fluid velocity equal to a constant rotation plus translation renders zero strain. Hint: if $\mathbb{S}_{mn} = 0$, what does that imply about the velocity field? You may also wish to make use of the general decomposition (20.35).

EXERCISE 20.2: EVOLUTION OF THE JACOBIAN USING ϵ -TENSOR GYMNASTICS

There is another way to derive the identity (20.52) for the evolution of the Jacobian. This other method is somewhat more tedious. However, it exercises some useful methods of index gymnastics involving the ϵ -tensor. It also has a natural generalization to curved spaces. This exercise is only for aficionados of tensor analysis.

An explicit expression for the Jacobian of transformation is given by

$$\frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{1}{3!} \epsilon_{mnp} \epsilon_{ijk} \frac{\partial X_m}{\partial a_i} \frac{\partial X_n}{\partial a_j} \frac{\partial X_p}{\partial a_k}. \quad (20.61)$$

Take the material derivative of this expression and show that we get the same expression as equation (20.52). Hint: make use of the identity

$$\frac{D}{Dt} \frac{\partial X_m}{\partial a_i} = \frac{\partial V_m}{\partial a_i}, \quad (20.62)$$

which holds since the material time derivative is taken with the material coordinates, \mathbf{a} , held fixed.

21

General vertical coordinate kinematics

In providing a mechanistic description of budgets within the ocean or atmosphere, it is often useful to measure the material or momentum transfer through a surface. This transport is termed the *dia-surface transport*. Our discussion in this chapter unifies ideas developed for kinematic boundary conditions in Section 17.4 with transport across an arbitrary surface in the fluid interior. We do so by making use of the generalized vertical coordinates (GVCs) first introduced in Chapter 11. We make use of the dia-surface transport formulation to express the material time derivative operator using GVCs. This form for the material time operator allows us to decompose of the vertical velocity into motion relative to a moving GVC surface. In turn, we are afforded a means to reinterpret the velocity vector and particle trajectories. GVC kinematics also provides a means to express the subduction of fluid into the ocean interior beneath the mixed layer depth. We close the chapter with derivations of the GVC version of mass continuity and the tracer equation. We also introduce the layer integrated version of the continuity and tracer equations, with the layer integrated equations appropriate for discrete numerical fluid models.

READER'S GUIDE TO THIS CHAPTER

We introduced mathematical properties of generalized vertical coordinates (GVCs) in Chapter 11, including the calculus using these non-orthogonal coordinates. It is essential to have a working knowledge of that material to understand material in the present chapter. Later in Chapter 32 we detail the dynamical equations using GVCs, with material in that chapter relying on the kinematics presented here.

21.1	Example generalized vertical coordinates	292
21.1.1	Ocean free surface	292
21.1.2	Ocean bottom	293
21.1.3	Ocean mixed layer base	293
21.1.4	Interior generalized vertical coordinate surfaces	293
21.2	Specific thickness	293
21.3	The dia-surface transport	294
21.3.1	Flow normal to the GVC surface	294
21.3.2	Accounting for movement of the surface	295
21.3.3	Cross GVC transport in terms of GVC material evolution	296
21.3.4	Defining the dia-surface transport	296
21.3.5	Expressions for the dia-surface velocity component	297
21.3.6	An alternative definition of dia-surface velocity component	298
21.3.7	Area integrated dia-surface transport for incompressible fluids	298
21.4	Material time derivative	300
21.5	Vertical velocity and dia-surface velocity	301
21.5.1	Decomposing the vertical velocity	301
21.5.2	Another form of the vertical velocity decomposition	301
21.6	The velocity vector and fluid particle trajectories	303
21.7	Subduction across the mixed layer base	305
21.8	Mass continuity	305
21.8.1	Cartesian coordinates	305
21.8.2	Generalized vertical coordinates	306
21.9	Layer integrated mass continuity	306
21.9.1	Comments about the layer thickness	307
21.9.2	Compressible fluids	307
21.9.3	Mass continuity using pressure coordinates	308
21.9.4	Incompressible fluids	310
21.9.5	Rescaled geopotential coordinates	310
21.10	Layer integrated tracer equation	310

21.1 Example generalized vertical coordinates

We here consider some generalized vertical coordinates that will prove of use for our discussion in this chapter.

21.1.1 Ocean free surface

The first surface is the ocean free surface, whose kinematic boundary conditions were derived in Section 17.4.3. Here, water and tracer penetrate this surface through precipitation, evaporation, river runoff (when applied as an upper ocean boundary condition), and sea ice melt. Momentum exchange arises from stresses between the ocean and atmosphere or ice. The ocean free surface can be represented mathematically by the identity

$$\sigma(x, y, z, t) = z - \eta(x, y, t) = 0 \quad \text{ocean free surface.} \quad (21.1)$$

This identity holds so long as we assume the surface height η is smooth and contains no overturns at the scales of interest. That is, we assume breaking surface waves are filtered from the description.

21.1.2 Ocean bottom

We may describe the solid Earth lower boundary mathematically by using the time independent expression

$$\sigma(x, y, z) = z + H(x, y) = 0 \quad \text{ocean bottom.} \quad (21.2)$$

As detailed in Section 17.4.1, 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 exchange through the action of stresses occur between the solid Earth and ocean fluid.

21.1.3 Ocean mixed layer base

Let

$$\sigma = z + h^{(\text{mld})}(x, y, t) = 0 \quad (21.3)$$

represent the vertical position of the ocean mixed layer base. The corresponding normal vector is given by

$$\hat{\mathbf{n}}^{(\text{mld})} = \frac{\nabla(z + h^{(\text{mld})})}{|\nabla(z + h^{(\text{mld})})|}. \quad (21.4)$$

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 21.7).

21.1.4 Interior generalized vertical coordinate surfaces

Within the ocean 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 (or equivalently constant buoyancy surfaces) as used in isopycnal ocean models as well as theoretical descriptions of adiabatic ocean dynamics.

21.2 Specific thickness

As mentioned in Section 11.9.1, 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

$$\frac{\partial z}{\partial \sigma} = z_\sigma, \quad (21.5)$$

which must then be single signed for useful 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 $z = \eta(x, y, t)$. This restriction places a limitation on the ability of certain GVC models (e.g., isopycnal models) to describe non-hydrostatic processes, such as overturning, common in Kelvin-Helmholz billows or vertical convection. Note that for both the ocean bottom and free surface

$$\frac{\partial z}{\partial \sigma} = 1 \quad \text{ocean free surface and bottom.} \quad (21.6)$$

This relation also holds, trivially, for geopotential coordinates in which $\sigma = z$.

We refer to the Jacobian z_σ as the *specific thickness*, with this name motivated by noting that the vertical thickness of a layer of coordinate thickness $\delta\sigma$ is given by

$$\delta z = (\partial z / \partial \sigma) \delta \sigma. \quad (21.7)$$

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

$$\delta z = (\partial z / \partial b) \delta b = N^{-2} \delta b, \quad (21.8)$$

with

$$N^2 = \frac{\partial b}{\partial z} \quad (21.9)$$

the buoyancy frequency (Section 25.3.4) in a Boussinesq fluid (Chapter 28). For a hydrostatic fluid using pressure as the vertical coordinate, the thickness of a pressure layer is

$$\delta z = (\partial z / \partial p) \delta p = -\frac{1}{\rho g} \delta p \quad (21.10)$$

where we used the hydrostatic relation (Section 27.3)

$$\frac{\partial p}{\partial z} = -\rho g \quad (21.11)$$

with g the acceleration due to gravity. Note that we assume the layer thickness is positive, $\delta z > 0$. For this purpose, with hydrostatic pressure we consider negative pressure increments, $\delta p < 0$, as this corresponds to vertically upward movement in a fluid column.

21.3 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 GVC surface.

21.3.1 Flow normal to the GVC surface

At an arbitrary point on a surface of constant generalized vertical coordinate (see Figure 21.1), the rate at which fluid moves in the direction normal to the surface is given by

$$\text{RATE OF FLUID FLOW IN DIRECTION } \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}}, \quad (21.12)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla \sigma}{|\nabla \sigma|}, \quad (21.13)$$

is the surface unit normal. For the ocean free surface, $\sigma = z - \eta$, the unit normal takes the form

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla \eta}{\sqrt{1 + |\nabla \eta|^2}} \quad \text{ocean free surface}, \quad (21.14)$$

whereas at the solid Earth bottom, $\sigma = z + H$,

$$\hat{\mathbf{n}} = -\frac{\nabla(z + H)}{|\nabla(z + H)|} = -\frac{\hat{\mathbf{z}} + \nabla H}{\sqrt{1 + |\nabla H|^2}} \quad \text{ocean bottom}. \quad (21.15)$$

Introducing the material time derivative

$$\frac{D\sigma}{Dt} = \frac{\partial\sigma}{\partial t} + \mathbf{v} \cdot \nabla\sigma \quad (21.16)$$

to equation (21.12) leads to the identity

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \frac{1}{|\nabla\sigma|} \left[\frac{D\sigma}{Dt} - \frac{\partial\sigma}{\partial t} \right]. \quad (21.17)$$

Hence, the component to the velocity of a fluid particle that is normal to a GVC surface is proportional to the difference between the material time derivative of the surface coordinate and its partial time derivative.

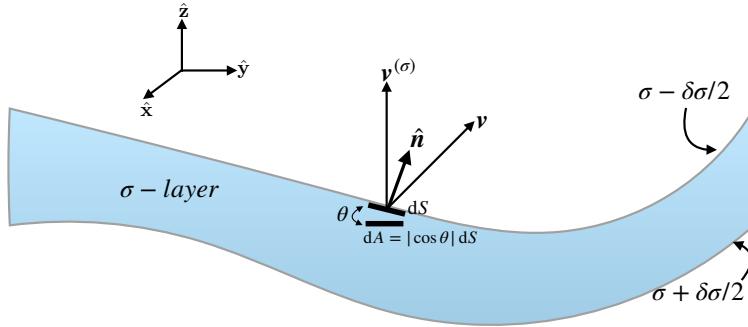


Figure 21.1: Two surfaces of constant generalized vertical coordinate that form the interfaces of a σ -layer. An upward normal direction $\hat{\mathbf{n}}$ is indicated on one of the interfaces. Also shown is an example velocity of a fluid particle \mathbf{v} and velocity $\mathbf{v}^{(\sigma)}$ of a point living 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 (21.23). Following equation (21.28), the horizontal projection of the surface area element is given by $dA = |\cos\theta| dS$, where θ is the angle between the surface and the horizontal.

21.3.2 Accounting for movement of the surface

A GVC 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

$$\text{RATE THAT FLUID CROSSES A MOVING GVC SURFACE} = \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}). \quad (21.18)$$

We next develop a kinematic property of the surface velocity, or more precisely the normal component to that velocity. For that purpose, consider an infinitesimal increment in both space and time under which σ undergoes an infinitesimal change

$$\delta\sigma = \delta\mathbf{x} \cdot \nabla\sigma + \delta t \partial_t\sigma. \quad (21.19)$$

Now restrict attention to a point fixed on a constant σ surface, in which

$$\delta\sigma = \delta\mathbf{x}^{(\sigma)} \cdot \nabla\sigma + \delta t \partial_t\sigma = 0, \quad (21.20)$$

where $\delta\mathbf{x}^{(\sigma)}$ is a differential increment following the moving surface. We define the velocity of that point as

$$\mathbf{v}^{(\sigma)} = \frac{\delta\mathbf{x}^{(\sigma)}}{\delta t}, \quad (21.21)$$

in which case equation (21.20) implies

$$\frac{\partial \sigma}{\partial t} + \mathbf{v}^{(\sigma)} \cdot \nabla \sigma = 0. \quad (21.22)$$

We can likewise write this equation as one for the normal component of the surface velocity

$$\mathbf{v}^{(\sigma)} \cdot \hat{\mathbf{n}} = -\frac{1}{|\nabla \sigma|} \frac{\partial \sigma}{\partial t}. \quad (21.23)$$

Hence, the normal component to the velocity of the surface vanishes when the surface is static. Note that we only need information about $\mathbf{v}^{(\sigma)} \cdot \hat{\mathbf{n}}$ for our kinematic formulation. Specification of the tangential component of $\mathbf{v}^{(\sigma)}$ requires dynamical information specific to the chosen surface.

21.3.3 Cross GVC transport in terms of GVC material evolution

Using expression (21.23) in equation (21.18) leads to the net flux of fluid crossing the GVC surface

$$\hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = \frac{1}{|\nabla \sigma|} \frac{D\sigma}{Dt}. \quad (21.24)$$

The material time derivative of the GVC surface thus vanishes if no fluid crosses the surface. Notably, this result holds for motion of the fluid as defined by the barycentric velocity, \mathbf{v} , of Section 18.1.2. For multi-component fluids, $\dot{\sigma} = 0$ can still, in principle, be associated with trace matter exchange across the surface via diffusion so long as the net matter crossing the surface is zero. However, for the ocean, matter diffusion also occurs with heat diffusion, in which case $\dot{\sigma} = 0$ only occurs in the absence of both matter and heat diffusion.

21.3.4 Defining the dia-surface transport

The area normalizing the volume flux in equation (21.24) 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 11.17 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

$$\frac{dS}{|\nabla \sigma|} = \frac{dS}{\sqrt{(\partial \sigma / \partial x)^2 + (\partial \sigma / \partial y)^2 + (\partial \sigma / \partial z)^2}} \quad (21.25a)$$

$$= \frac{dS}{|\partial \sigma / \partial z| \sqrt{[(\partial \sigma / \partial x) / (\partial \sigma / \partial z)]^2 + [(\partial \sigma / \partial y) / (\partial \sigma / \partial z)]^2 + 1}} \quad (21.25b)$$

$$= \frac{dS}{|\partial \sigma / \partial z| \sqrt{1 + \tan^2 \theta}} \quad (21.25c)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| |\cos \theta| dS \quad (21.25d)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| dA. \quad (21.25e)$$

The equality (21.25c) introduced the angle, θ , between the boundary surface and the horizontal plane. The squared slope of this surface given by (see Section 11.12)

$$\tan^2 \theta = \frac{\nabla_z \sigma \cdot \nabla_z \sigma}{(\partial \sigma / \partial z)^2} = \nabla_\sigma z \cdot \nabla_\sigma z. \quad (21.26)$$

The equality (21.25d) made use of a trigonometric identity so that

$$|\cos \theta|^{-1} = |z_\sigma \nabla \sigma|. \quad (21.27)$$

Furthermore, the equality (21.25e) introduced the horizontal projection of the area,

$$dA = |\cos \theta| dS. \quad (21.28)$$

We now introduce the *dia-surface velocity component* for the GVC coordinate

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = z_\sigma \dot{\sigma}, \quad (21.29)$$

which measures the volume of fluid passing through the surface, per unit horizontal area, per unit time

$$w^{(\dot{\sigma})} \equiv \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \frac{dS}{dA} \quad (21.30)$$

$$= \frac{(\text{VOLUME/TIME}) \text{ FLUID THROUGH SURFACE}}{\text{HORIZONTAL AREA OF SURFACE}}, \quad (21.31)$$

so that

$$w^{(\dot{\sigma})} dA \equiv \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) dS. \quad (21.32)$$

The velocity component $w^{(\dot{\sigma})}$ is referred to 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.

21.3.5 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 11.12, allows us to write the dia-surface velocity component in the following equivalent forms

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} \quad (21.33a)$$

$$= \frac{\partial z}{\partial \sigma} |\nabla \sigma| \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \quad (21.33b)$$

$$= \frac{\partial z}{\partial \sigma} \nabla \sigma \cdot \mathbf{v} - \frac{\partial z}{\partial \sigma} |\nabla \sigma| \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \quad (21.33c)$$

$$= (\hat{\mathbf{z}} - \nabla_\sigma z) \cdot \mathbf{v} + \frac{\partial z}{\partial \sigma} \frac{\partial \sigma}{\partial t} \quad (21.33d)$$

$$= (\hat{\mathbf{z}} - \nabla_\sigma z) \cdot \mathbf{v} - \frac{\partial z}{\partial t} \quad (21.33e)$$

$$= w - (\partial_t + \mathbf{u} \cdot \nabla_\sigma) z, \quad (21.33f)$$

where $\partial z / \partial t = (\partial z / \partial t)_\sigma$ is the time derivative for the depth of the σ surface. We also made use of the identity (see equations (11.25b) and (11.25c))

$$\nabla_\sigma z = -z_\sigma \nabla_z \sigma \quad (21.34)$$

to express the slope of the σ surface as projected onto the horizontal direction plane, as well as the corresponding identity (11.25a) for the time derivative

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} = - \frac{[\partial \sigma / \partial t]_z}{[\partial \sigma / \partial z]} \quad (21.35)$$

The form given by equation (21.33f) directly relates the vertical component to the fluid particle velocity to the dia-surface velocity component

$$w = \frac{Dz}{Dt} \longleftrightarrow w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = w - (\partial_t + \mathbf{u} \cdot \nabla_{\sigma})z. \quad (21.36)$$

When the GVC surface is static, so that it occupies a constant vertical position $\partial z / \partial t = 0$, then the dia-surface velocity component reduces to

$$w^{(\dot{\sigma})} = w - \mathbf{u} \cdot \nabla_{\sigma} z \quad \text{static surface,} \quad (21.37)$$

whereas if the GVC surface is flat, then the dia-surface velocity component measures the flux of fluid moving vertically relative to the motion of the GVC surface. Finally, if the surface is flat and static, the dia-surface velocity component becomes the vertical velocity component

$$w^{(\dot{\sigma})} = w = \frac{Dz}{Dt} \quad \text{GVC surface static and flat,} \quad (21.38)$$

which is the case for the geopotential vertical coordinate. This relation reveals the kinematic distinction between w and $w^{(\dot{\sigma})}$, with the two differing in the presence of GVC transients and horizontal velocities that project onto a non-horizontal GVC surface. Equation (21.33f) thus offers a useful means to distinguish w from $w^{(\dot{\sigma})}$.

21.3.6 An alternative definition of dia-surface velocity component

In some literature presentations, the dia-surface velocity component is taken to be

$$w^{\text{dia}} = \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = \frac{1}{|\nabla \sigma|} \frac{D\sigma}{Dt}. \quad (21.39)$$

For example, [Groeskamp et al. \(2019\)](#) prefer this definition for watermass analysis. As seen in Chapter 41, the reason to prefer expression (21.39) for watermass transformation analysis is that we do not wish to assume vertically stable stratification for surfaces of constant σ . Dropping that assumption allows us to consider transformation between arbitrarily oriented elements of seawater, even those that are gravitationally unstable.

21.3.7 Area integrated dia-surface transport for incompressible fluids

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 an incompressible fluid whereby $\nabla \cdot \mathbf{v} = 0$. Incompressibility means that for any closed surface within the fluid interior, the following identity holds via the divergence theorem

$$0 = \int_{\mathcal{R}} \nabla \cdot \mathbf{v} dV = \oint_{\partial \mathcal{R}} \hat{\mathbf{n}} \cdot \mathbf{v} d\mathcal{S}. \quad (21.40)$$

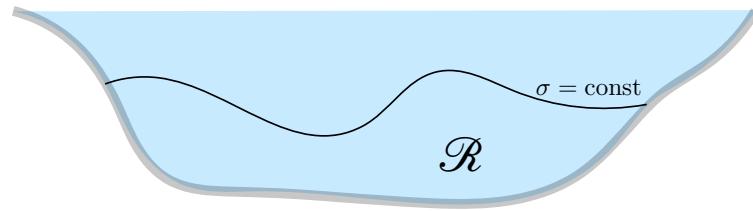


Figure 21.2: A constant GVC surface, $\sigma = \text{constant}$, within an ocean basin that intersects the bottom. The region \mathcal{R} is bounded above by the σ surface and below by the solid-earth. Along the constant σ surface an incompressible fluid satisfies $\int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot \mathbf{v} d\mathcal{S} = 0$.

Notably, only in the case of a static surface do we conclude there is no net flow across the surface (see Exercise 19.5). For surfaces that move, there is generally a nonzero net dia-surface transport. We clarify this rather puzzling statement in the following.

As a specific example, consider a fluid region such as shown in Figure 21.2, which is bounded by the solid-earth bottom and a constant GVC surface. Since the solid-earth bottom is static and there is no-normal flow through the bottom, the identity (21.40) means that the area integrated flow normal to the GVC vanishes

$$\int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot \mathbf{v} d\mathcal{S} = 0. \quad (21.41)$$

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 Exercise 19.5). What about other GVCs?

To address this question consider the general result

$$\int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) d\mathcal{S} = \int_{\sigma=\text{const}} w^{\text{dia}} d\mathcal{S} = \int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA, \quad (21.42)$$

where again $dA = dx dy$. Now make use of the property (21.41) for incompressible flows as well as the identity (21.23) to render

$$\int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA = 0 - \int_{\sigma=\text{const}} \hat{\mathbf{n}} \cdot \mathbf{v}^{(\sigma)} d\mathcal{S} \quad (21.43a)$$

$$= \int_{\sigma=\text{const}} \frac{\partial \sigma / \partial t}{|\nabla \sigma|} d\mathcal{S} \quad (21.43b)$$

$$= \int_{\sigma=\text{const}} \frac{\partial \sigma}{\partial t} \left| \frac{\partial z}{\partial \sigma} \right| dA \quad (21.43c)$$

$$= - \int_{\sigma=\text{const}} \left[\frac{\partial z}{\partial t} \right]_{\sigma} dA. \quad (21.43d)$$

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 σ constant, as is the horizontal area integral. Hence, we can pull the time derivative outside the integral to render

$$\int_{\sigma=\text{const}} w^{(\dot{\sigma})} dA = - \left[\frac{\partial}{\partial t} \right]_{\sigma} \int_{\sigma=\text{const}} z dA. \quad (21.44)$$

This identity means that for an incompressible fluid, the integrated dia-surface transport across the GVC 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 GVC surface so long as there is a volume change for the region beneath the surface.

For the case of an isopycnal surface in an adiabatic fluid, there is no change in the volume beneath any interior isopycnal 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, $\sigma = z^*$ defined by equation (21.109). In this case

$$z^* = H \frac{z - \eta}{H + \eta} \quad (21.45a)$$

$$\frac{\partial z}{\partial z^*} = 1 + H/\eta > 0 \quad (21.45b)$$

$$\left[\frac{\partial z}{\partial t} \right]_{z^*} = \frac{\partial \eta}{\partial t} (1 + z^*/H), \quad (21.45c)$$

so that

$$\int_{z^*=\text{const}} w^{(z^*)} dA = \int_{z^*=\text{const}} (\partial \eta / \partial t) (1 + z^*/H) dA, \quad (21.46)$$

which is generally nonzero. For example, consider a flat bottom so that

$$\int_{z^*=\text{const}} w^{(z^*)} dA = (1 + z^*/H) \int_{z^*=\text{const}} (\partial \eta / \partial t) dA = (1 + z^*/H) \int_{z^*=\text{const}} (Q_m / \rho_0) dA, \quad (21.47)$$

where Q_m is the surface mass flux and we made use of the incompressible free surface equation (19.8). In this case the area integrated dia-surface transport across a z^* surface is proportional to the area integrated surface mass flux.

21.4 Material time derivative

The expression (21.29) for $w^{(\dot{\sigma})}$ brings the material time derivative operator into the following equivalent forms

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (21.48a)$$

$$= \left[\frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma + \frac{D\sigma}{Dt} \frac{\partial}{\partial \sigma} \quad (21.48b)$$

$$= \left[\frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma + w^{(\dot{\sigma})} \frac{\partial}{\partial z}. \quad (21.48c)$$

Note that the chain-rule means that

$$\frac{\partial}{\partial \sigma} = \frac{\partial z}{\partial \sigma} \frac{\partial}{\partial z}, \quad (21.49)$$

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 surface. This situation occurs in the case of adiabatic flows with σ equal to the buoyancy or isopycnal coordinate. For adiabatic flow, the material time derivative in equation (21.48c) only has a horizontal two-dimensional

advective component $\mathbf{u} \cdot \nabla_b$. This result should not be interpreted to mean that the fluid particle velocity in an adiabatic flow is strictly horizontal. Indeed, it generally is not, as the form given by equation (21.48a) 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.

21.5 Vertical velocity and dia-surface velocity

Making use of the material time derivative operator (21.48c) 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

$$w = \frac{Dz}{Dt} = \left[\frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma z + w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \left[-\frac{\partial \sigma}{\partial t} - \mathbf{u} \cdot \nabla_z \sigma + \frac{D\sigma}{Dt} \right], \quad (21.50)$$

and the corresponding expressions for the dia-surface velocity component are given by

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = \frac{\partial z}{\partial \sigma} \left[\frac{\partial \sigma}{\partial t} + \mathbf{u} \cdot \nabla_z \sigma + w \frac{\partial \sigma}{\partial t} \right] = - \left[\frac{\partial z}{\partial t} \right]_\sigma - \mathbf{u} \cdot \nabla_\sigma z + w. \quad (21.51)$$

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 σ surfaces, which are generally moving and sloped. It is notable that the area normalization used in equation (21.31) for the dia-surface velocity component means that it appears only in the expression for the vertical velocity. However, as we will see in the following, its appearance in the w equation does not necessarily mean that it corresponds to vertical particle motion. Instead, when it arises from mixing, $w^{(\dot{\sigma})}$ can lead to vertical motion of the σ surface while maintaining a fixed position for the fluid particle.

21.5.1 Decomposing the vertical velocity

The expression

$$w = \left[\frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma z + w^{(\dot{\sigma})} \quad (21.52)$$

decomposes the vertical velocity of a fluid particle into (A) changes to the vertical position of the σ -surface at a particular horizontal point, (B) lateral particle motion projected onto a sloped σ -surface, (C) motion that crosses a σ -surface. Importantly, the three terms are coupled. For example, consider the case of σ defined by isopycnals, in which case irreversible mixing ($w^{(\dot{\sigma})} \neq 0$) changes the configuration of σ surfaces by changing both their height, $(\partial z / \partial t)_\sigma$, and slope $\nabla_\sigma z$.

21.5.2 Another form of the vertical velocity decomposition

Consider the velocity for the surface itself, $\mathbf{v}^{(\sigma)}$, which satisfies (Section 21.3.2)

$$\frac{\partial \sigma}{\partial t} + \mathbf{v}^{(\sigma)} \cdot \nabla \sigma = 0. \quad (21.53)$$

Making use of the triple product identities from Section 11.5

$$\frac{\partial z}{\partial \sigma} \nabla \sigma = -\nabla_\sigma z + \hat{\mathbf{z}} \quad \text{and} \quad \frac{\partial z}{\partial \sigma} \left[\frac{\partial \sigma}{\partial t} \right]_z = - \left[\frac{\partial z}{\partial t} \right]_\sigma \quad (21.54)$$

brings equation (21.53) into the form

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} = (\hat{z} - \nabla_{\sigma} z) \cdot \mathbf{v}^{(\sigma)} \implies \hat{z} \cdot \mathbf{v}^{(\sigma)} = \left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u}^{(\sigma)} \cdot \nabla_{\sigma} z, \quad (21.55)$$

where $\mathbf{u}^{(\sigma)}$ is the horizontal component to the surface velocity $\mathbf{v}^{(\sigma)}$. This equation shows that the vertical component to the σ -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 σ -surface has no component of velocity in the vertical, the depth of the σ -surface measured at a horizontal point generally changes if the surface is sloped and moves horizontally past that point

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} = -\mathbf{u}^{(\sigma)} \cdot \nabla_{\sigma} z \quad \text{if } \hat{z} \cdot \mathbf{v}^{(\sigma)} = 0. \quad (21.56)$$

Returning to the general result (21.55) allows us to write

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z = \hat{z} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z. \quad (21.57)$$

Furthermore, return to the fundamental definition of the dia-surface velocity component detailed in Section 21.3, in which we showed that

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = \frac{\partial z}{\partial \sigma} \nabla_{\sigma} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = (-\nabla_{\sigma} z + \hat{z}) \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}). \quad (21.58)$$

This expression, along with equation (21.57), leads to the rather elaborate decomposition of the vertical velocity component according to motion of a generalized vertical coordinate surface

$$w = \underbrace{\left[\hat{z} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\sigma}) z} + \underbrace{\left[\hat{z} \cdot \mathbf{v} - \hat{z} \cdot \mathbf{v}^{(\sigma)} - (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z \right]}_{w^{(\dot{\sigma})}}. \quad (21.59)$$

Terms in the first bracket compute vertical particle motion relative to the σ -surface. The dia-surface contribution from the second bracket removes the contribution from σ -surface motion to leave just the vertical motion of the particle. All terms on the right hand side cancel, except for $\hat{z} \cdot \mathbf{v} = w$, thus trivially revealing $w = w$. The decomposition of w is rather pedantic when viewed in the unwrapped form of equation (21.59). 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 σ -surface: $\mathbf{u} = \mathbf{u}^{(\sigma)}$. Alternatively, consider the case with flat σ -surfaces so that $\nabla_{\sigma} z = 0$. In either case the vertical velocity is given by

$$w = \underbrace{\left[\hat{z} \cdot \mathbf{v}^{(\sigma)} \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\sigma}) z} + \underbrace{\left[\hat{z} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \right]}_{w^{(\dot{\sigma})}}. \quad (21.60)$$

The first contribution is from vertical motion of the σ -surface. The second contribution adjusts for the vertical motion of the particle relative to the σ -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

$$w = 0 \quad (21.61a)$$

$$= \left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} \quad (21.61b)$$

$$= \underbrace{\left[\hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\sigma}) z} + \underbrace{\left[-\hat{\mathbf{z}} \cdot \mathbf{v}^{(\sigma)} - (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z \right]}_{w^{(\dot{\sigma})}}. \quad (21.61c)$$

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 σ surface

$$-w^{(\dot{\sigma})} = \left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z \quad \text{if } w = 0. \quad (21.62)$$

A particularly simple realization of this balance holds for σ 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 merely since $w^{(\dot{\sigma})} \neq 0$ can arise from motion of a σ -surface alone.

21.6 The velocity vector and fluid particle trajectories

Recall from Section 21.5 the alternative forms for the vertical velocity component given by equation (21.50). We focus on the form

$$w = \left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} \quad (21.63)$$

so that the velocity vector is written¹

$$\mathbf{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}} \quad (21.64a)$$

$$= u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + \left[(\partial z / \partial t)_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}} \quad (21.64b)$$

$$= u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_{\sigma}] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_{\sigma}] + \left[(\partial z / \partial t)_{\sigma} + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}}. \quad (21.64c)$$

To help further understand these velocity expressions we consider the following three cases, each of which are illustrated in Figure 21.3.

- STEADY AND MATERIAL σ -SURFACE: The velocity vector is aligned with the instantaneous σ -surface ($\mathbf{v} \cdot \nabla \sigma = 0$) when the σ -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

$$w \partial \sigma / \partial z = -\mathbf{u} \cdot \nabla_z \sigma \implies w = \mathbf{u} \cdot \nabla_{\sigma} z, \quad (21.65)$$

¹ As discussed in Section 11.7, we can connect these expressions to the contravariant representation of the velocity vector using GVCs.

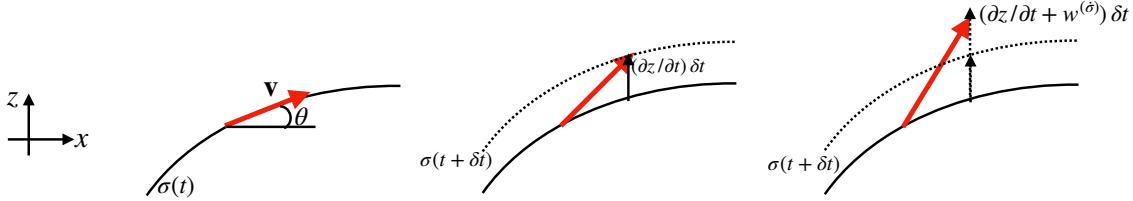


Figure 21.3: This schematic shows the various contributions to the fluid particle velocity (red vector) when written relative to motion of a particular generalized vertical coordinate surface. The fluid particle sits at the tail of the velocity vector at time t and at the head at time $t + \delta t$. The left panel is for the case of a static and material σ -surface so that the particle remains on the σ -surface and has a velocity vector given by equation (21.66). The slope of the σ -surface in the \hat{x} -direction is given by $\tan \theta = (\partial z / \partial x)_\sigma$. The middle panel is for a non-steady material σ -surface whereby the velocity of a particle takes on the form (21.67), with the particle remaining on the moving σ -surface. The right panel shows the case of a non-steady and non-material σ -surface with velocity (21.68). In this final case the particle position departs from the original σ -surface due to the nonzero dia-surface velocity component, $w^{(\dot{\sigma})} \neq 0$. However, it is not known *a priori* whether this departure is due to particle motion or motion of the surface. 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 σ -surface (left panel), motion of the σ -surface (middle panel), and motion crossing the σ -surface (right panel).

where we used the triple product identities (11.25b) and (11.25c) for the final equality. The velocity vector thus takes on the form

$$\mathbf{v} = u [\hat{x} + \hat{z} (\partial z / \partial x)_\sigma] + v [\hat{y} + \hat{z} (\partial z / \partial y)_\sigma] \quad \partial \sigma / \partial t = 0 \text{ and } D\sigma / Dt = 0. \quad (21.66)$$

The velocity vector is determined only by the horizontal velocity plus the slope of the σ surface.

- **NON-STEADY AND MATERIAL σ -SURFACE:** Next consider material σ surfaces ($D\sigma / Dt = 0$) that move ($\partial_t \sigma \neq 0$), in which case the velocity vector is

$$\mathbf{v} = u [\hat{x} + \hat{z} (\partial z / \partial x)_\sigma] + v [\hat{y} + \hat{z} (\partial z / \partial y)_\sigma] + (\partial z / \partial t)_\sigma \hat{z} \quad D\sigma / Dt = 0. \quad (21.67)$$

To remain on the moving surface, the fluid particle must move vertically by the extra amount $(\partial z / \partial t)_\sigma \delta t \hat{z}$ relative to the case of a static σ -surface.

- **NON-STEADY AND NON-MATERIAL σ -SURFACE:** The general case with a non-material and non-steady σ also requires the dia-surface velocity component, $w^{(\dot{\sigma})}$, which is diagnosed based on the material time derivative of σ and the inverse stratification, $w^{(\dot{\sigma})} = (\partial z / \partial \sigma) D\sigma / Dt$:

$$\mathbf{v} = u [\hat{x} + \hat{z} (\partial z / \partial x)_\sigma] + v [\hat{y} + \hat{z} (\partial z / \partial y)_\sigma] + [(\partial z / \partial t)_\sigma + w^{(\dot{\sigma})}] \hat{z}. \quad (21.68)$$

The contribution $w^{(\dot{\sigma})}$ measures the vertical motion of the particle relative to the moving σ -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 21.5, a non-zero $w^{(\dot{\sigma})}$ arises from motion of the fluid particle relative to the σ -surface, and this relative motion does not necessarily mean that the particle moves; e.g., recall the example discussed in Section 21.5.2 with a static particle and moving σ -surface.

21.7 Subduction across the mixed layer base

Consider the GVC (21.3) defined according to the mixed layer base. The dia-surface mass transport across this surface leads us to define the subduction

$$-\mathcal{S}^{(\text{subduction})} \equiv \rho dA \left[\frac{d(z + h^{(\text{mld})})}{dt} \right] \quad \text{at } z = -h^{(\text{mld})}(x, y, t), \quad (21.69)$$

where the mass transport $\mathcal{S}^{(\text{subduction})}$ (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

$$-\left[\frac{\mathcal{S}^{(\text{subduction})}}{\rho dA} \right] = w + [\partial_t + \mathbf{u} \cdot \nabla] h^{(\text{mld})} \quad \text{at } z = -h^{(\text{mld})}(x, y, t), \quad (21.70)$$

where again we define

$$\mathcal{S}^{(\text{subduction})} > 0 \quad \text{subduction} \quad (21.71)$$

$$\mathcal{S}^{(\text{subduction})} < 0 \quad \text{obduction.} \quad (21.72)$$

This definition of subduction corresponds to that given by [Cushmin-Roisin \(1987\)](#).

21.8 Mass continuity

We here derive the Eulerian equation for mass continuity (17.9) using generalized vertical coordinates. We then specialize to incompressible fluids, in which mass conservation is converted to volume conservation. To start, recall that mass conservation for a fluid parcel states that

$$\rho \delta V = \rho \delta x \delta y \delta z = \rho \delta x \delta y z_\sigma \delta \sigma \quad (21.73)$$

is materially constant. To develop the Eulerian expressions we first consider the case of Cartesian coordinates.

21.8.1 Cartesian coordinates

Consider the expression

$$\frac{1}{\rho \delta V} \frac{D(\rho \delta V)}{Dt} = \frac{1}{\rho} \frac{D\rho}{Dt} + \frac{1}{\delta V} \frac{D(\delta V)}{Dt}. \quad (21.74)$$

Now make use of Cartesian coordinates to write for the volume

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \frac{1}{\delta x \delta y \delta z} \frac{D(\delta x \delta y \delta z)}{Dt} \quad (21.75a)$$

$$= \frac{1}{\delta x} \frac{D(\delta x)}{Dt} + \frac{1}{\delta y} \frac{D(\delta y)}{Dt} + \frac{1}{\delta z} \frac{D(\delta z)}{Dt} \quad (21.75b)$$

$$= \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} + \frac{\delta w}{\delta z} \quad (21.75c)$$

$$= \nabla \cdot \mathbf{v}. \quad (21.75d)$$

Setting $D(\rho \delta V)/Dt = 0$ leads to the familiar expression for the continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \quad (21.76)$$

21.8.2 Generalized vertical coordinates

We follow the above procedure but now with generalized vertical coordinates so that

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \frac{1}{\delta x \delta y z_\sigma \delta \sigma} \frac{D(\delta x \delta y z_\sigma \delta \sigma)}{Dt} \quad (21.77a)$$

$$= \frac{1}{\delta x} \frac{D(\delta x)}{Dt} + \frac{1}{\delta y} \frac{D(\delta y)}{Dt} + \frac{1}{z_\sigma} \frac{D(z_\sigma)}{Dt} + \frac{1}{\delta \sigma} \frac{D(\delta \sigma)}{Dt} \quad (21.77b)$$

$$= \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} + \frac{1}{z_\sigma} \frac{D(z_\sigma)}{Dt} + \frac{\delta(\dot{\sigma})}{\delta \sigma} \quad (21.77c)$$

$$= \nabla_\sigma \cdot \mathbf{u} + \frac{1}{z_\sigma} \frac{D(z_\sigma)}{Dt} + \frac{\partial \dot{\sigma}}{\partial \sigma} \quad (21.77d)$$

where we introduced the shorthand

$$\dot{\sigma} = \frac{D\sigma}{Dt}. \quad (21.78)$$

Note that we set

$$\frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} = \nabla_\sigma \cdot \mathbf{u} \quad (21.79)$$

since we are working with generalized vertical coordinates so that we consider infinitesimal displacements occurring on constant σ surfaces. We are thus led to

$$\frac{1}{\rho \delta V} \frac{D(\rho \delta V)}{Dt} = \nabla_\sigma \cdot \mathbf{u} + \frac{1}{z_\sigma} \frac{Dz_\sigma}{Dt} + \frac{\partial \dot{\sigma}}{\partial \sigma} + \frac{1}{\rho} \frac{D\rho}{Dt} = 0. \quad (21.80)$$

Now use the material time derivative in the form (21.48b) to derive the Eulerian expression of mass conservation

$$\frac{\partial(\rho z_\sigma)}{\partial t} + \nabla_\sigma \cdot (\rho z_\sigma \mathbf{u}) + \frac{\partial(\rho z_\sigma \dot{\sigma})}{\partial \sigma} = 0, \quad (21.81)$$

where the time derivative is computed holding σ fixed. We can furthermore introduce the dia-surface velocity component

$$w^{(\dot{\sigma})} = z_\sigma \dot{\sigma} \quad (21.82)$$

so that mass continuity takes the form

$$\frac{\partial(\rho z_\sigma)}{\partial t} + \nabla_\sigma \cdot (\rho z_\sigma \mathbf{u}) + \frac{\partial(\rho z_\sigma \dot{\sigma})}{\partial \sigma} = 0. \quad (21.83)$$

21.9 Layer integrated mass continuity

The formulation thus far has been continuous, with the only assumption made that the specific thickness, $\partial z / \partial \sigma$, is single signed. We here consider a discrete increment in the generalized vertical coordinate,

$$\sigma - \delta\sigma/2 \leq \sigma' \leq \sigma + \delta\sigma/2, \quad (21.84)$$

and formulate the mass budget over this layer whose thickness is given by

$$h \equiv \int_{\sigma - \delta\sigma/2}^{\sigma + \delta\sigma/2} \frac{\partial z}{\partial \sigma} d\sigma = \int_{z(\sigma - \delta\sigma/2)}^{z(\sigma + \delta\sigma/2)} dz, \quad (21.85)$$

and whose mass per horizontal area is

$$\delta m = \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho z_\sigma d\sigma = \bar{\rho} h, \quad (21.86)$$

where $\bar{\rho}$ is the layer averaged density. Note that for Boussinesq fluids the mass per area equals to the layer thickness times the reference density

$$\delta m = \rho_0 h \quad \text{Boussinesq.} \quad (21.87)$$

Note that the following formulation is quite general, even incorporating the trivial case of geopotential coordinates ($\sigma = z$) whereby the specific thickness is unity.

21.9.1 Comments about the layer thickness

As defined by equation (21.85), the thickness of a layer is relatively large in regions where $\partial\sigma/\partial z$ is small; i.e., in weakly stratified regions. 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 σ 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 σ is the hydrostatic pressure, so that $\partial p/\partial z = -\rho g$ (Section 21.9.3).

21.9.2 Compressible fluids

Performing a layer integral of the specific thickness equation (21.83) renders

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \left[\frac{\partial(\rho z_\sigma)}{\partial t} + \nabla_\sigma \cdot (\rho z_\sigma \mathbf{u}) + \frac{\partial(\rho w^{(\dot{\sigma})})}{\partial\sigma} \right] d\sigma = 0. \quad (21.88)$$

The dia-surface term integrates to a finite difference across the layer

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \left[\frac{\partial(\rho z_\sigma)}{\partial t} + \nabla_\sigma \cdot (\rho z_\sigma \mathbf{u}) \right] d\sigma = -\Delta_\sigma(\rho w^{(\dot{\sigma})}), \quad (21.89)$$

where we introduced the finite difference operator for properties defined at the layer interface

$$\Delta_\sigma(\Phi) = \Phi(\sigma + \delta\sigma/2) - \Phi(\sigma - \delta\sigma/2). \quad (21.90)$$

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 σ fixed. Hence, layer mass continuity takes the form

$$\left[\frac{\partial}{\partial t} \right]_\sigma \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho z_\sigma d\sigma + \nabla_\sigma \cdot \int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho \mathbf{u} z_\sigma d\sigma = -\Delta_\sigma(\rho w^{(\dot{\sigma})}). \quad (21.91)$$

The first term involves the layer averaged density times the layer thickness as per equation (21.86). The second term involves the layer averaged density-weighted velocity, which is the layer averaged horizontal mass flux

$$\int_{\sigma-\delta\sigma/2}^{\sigma+\delta\sigma/2} \rho \mathbf{u} z_\sigma d\sigma = h \bar{\rho} \bar{\mathbf{u}}. \quad (21.92)$$

We are thus led to the layer integrated continuity equation

$$\left[\frac{\partial(h\bar{\rho})}{\partial t} \right]_\sigma + \nabla_\sigma \cdot (h\bar{\rho}\bar{\mathbf{u}}) + \Delta_\sigma(\rho w^{(\dot{\sigma})}) = 0. \quad (21.93)$$

When evolving the fields in a discrete numerical model, we have 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 (21.93)? 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 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

$$\delta m = \int_{\sigma - \delta\sigma/2}^{\sigma + \delta\sigma/2} \rho z_\sigma d\sigma = \bar{\rho} h = -g^{-1} \delta p. \quad (21.94)$$

Correspondingly, the layer integrated horizontal mass flux equals to the mass increment times the pressure-layer averaged velocity

$$\int_{\sigma - \delta\sigma/2}^{\sigma + \delta\sigma/2} \rho \mathbf{u} z_\sigma d\sigma = -g^{-1} \int_{p - \delta p/2}^{p + \delta p/2} \mathbf{u} dp = -g^{-1} \bar{\mathbf{u}}^{(p)} \delta p = h \bar{\rho} \bar{\mathbf{u}}^{(p)}. \quad (21.95)$$

With either of the above two methods, we are led to the same layer integrated continuity equation, which we write in the generic form

$$\left[\frac{\partial(h\rho)}{\partial t} \right]_\sigma + \nabla_\sigma \cdot (h\rho\mathbf{u}) + \Delta_\sigma(\rho w^{(\dot{\sigma})}) = 0. \quad (21.96)$$

We illustrate contributions to this layer mass budget in Figure 21.4.

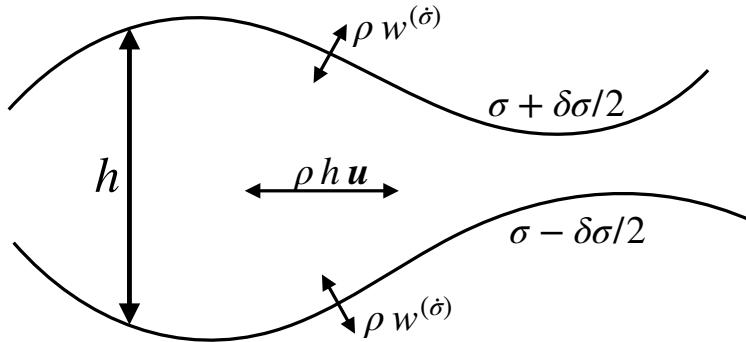


Figure 21.4: Illustrating the terms contributing to changes in layer mass according to the layer integrated continuity equation (21.96). The discrete layer is shown here with bounding interfaces at $\sigma - \delta\sigma/2$ and $\sigma + \delta\sigma/2$. Within a layer there is a horizontal redistribution due to horizontal advective transport. Additionally, matter can cross the layer due to dia-surface transport via $w^{(\dot{\sigma})}$.

21.9.3 Mass continuity using pressure coordinates

Let us here consider in some detail the special case of pressure coordinates in a hydrostatic fluid, and thus derive mass continuity using these coordinates.

Method I

The thickness of a hydrostatic pressure layer (equation (21.85)) takes on the following form

$$h = \int_{p-\delta p/2}^{p+\delta p/2} \frac{\partial z}{\partial p} dp = - \int_{p-\delta p/2}^{p+\delta p/2} \frac{dp}{\rho g}, \quad (21.97)$$

so that its mass per unit area is

$$\int_{p-\delta p/2}^{p+\delta p/2} \rho \frac{\partial z}{\partial p} dp = -\delta p/g. \quad (21.98)$$

The mass continuity equation (21.96) thus becomes

$$\frac{\partial(\delta p)}{\partial t} + \nabla_p \cdot (\mathbf{u} \delta p) + \Delta_p (\dot{p}) = 0. \quad (21.99)$$

The partial time derivative vanishes since it is computed by holding pressure fixed so that the pressure increment has a zero time tendency

$$\left[\frac{\partial(\delta p)}{\partial t} \right]_p = 0. \quad (21.100)$$

Likewise, $\nabla_p(\delta p) = 0$. Thus, we can divide by δp to render the continuity equation

$$\nabla_p \cdot \mathbf{u} + \frac{\partial \dot{p}}{\partial p} = 0 \quad \text{compressible hydrostatic.} \quad (21.101)$$

This equation is isomorphic to the incompressible continuity equation (for both hydrostatic and non-hydrostatic fluids) written using geopotential coordinates (see Chapter 28)

$$\nabla_z \cdot \mathbf{u} + \frac{\partial \dot{z}}{\partial z} = 0 \quad \text{incompressible,} \quad (21.102)$$

where $w = z$ is the vertical component to the velocity vector. In particular, note that for both cases 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 21.8.2, which starts from

$$\frac{D(\rho \delta V)}{Dt} = 0. \quad (21.103)$$

In pressure coordinates the volume of the fluid element takes the form

$$\delta V = \delta x \delta y \delta z = \delta x \delta y \left[\frac{\partial z}{\partial p} \right] \delta p = -(\rho g)^{-1} \delta x \delta y \delta p. \quad (21.104)$$

Consequently,

$$0 = \frac{D(\rho \delta V)}{Dt} = g^{-1} \left(\frac{D(\delta x \delta y \delta p)}{Dt} \right), \quad (21.105)$$

so that

$$0 = \frac{1}{\delta x \delta y \delta p} \left(\frac{D(\delta x \delta y \delta p)}{Dt} \right) = \nabla_p \cdot \mathbf{u} + \frac{\partial \dot{p}}{\partial p}. \quad (21.106)$$

The second step made use of the isomorphism between this result and that for equation (20.48) that holds for a geopotential vertical coordinate.

21.9.4 Incompressible fluids

Specializing to an incompressible volume conserving fluid (see Chapter 28) yields the incompressible layer thickness equation

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) + \Delta_\sigma w^{(\dot{\sigma})} = 0. \quad (21.107)$$

Further specializing to the case of zero dia-surface transport leads to

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) = 0 \quad \text{no dia-surface transport.} \quad (21.108)$$

This case is commonly studied for adiabatic fluids using isopycnal coordinates, in which isopycnal surfaces are material (Section 45.1).

21.9.5 Rescaled geopotential coordinates

The rescaled geopotential coordinate

$$z^* = H \frac{z - \eta}{H + \eta} \quad -H \leq z^* \leq 0, \quad (21.109)$$

is commonly used in Boussinesq ocean models, where $z = \eta(x, y, t)$ is the ocean free surface and $z = -H(x, y)$ is the ocean bottom (see Section 33.3.1). The thickness of a coordinate layer is given by

$$h = dz = \frac{\partial z}{\partial z^*} dz^* = (1 + \eta/H) dz^*. \quad (21.110)$$

The depth integrated column thickness and depth integrated coordinate thickness are given by

$$\int_{-H}^{\eta} dz = H + \eta \quad \int_{-H}^{\eta} dz^* = H. \quad (21.111)$$

Correspondingly, the depth integrated thickness equation is given by the depth integrated volume budget derived in Section 19.3

$$\frac{\partial \eta}{\partial t} + \nabla \cdot \mathbf{U} + [w_{z^*=0}^{(\dot{\sigma})} - w_{z^*=-H}^{(\dot{\sigma})}] = 0. \quad (21.112)$$

We assume no volume flow through the ocean bottom so that $w_{z^*=-H}^{(\dot{\sigma})} = 0$, whereas

$$-\rho_0 w_{z^*=0}^{(\dot{\sigma})} = Q_m \quad (21.113)$$

is the mass flux crossing the ocean free surface.

21.10 Layer integrated tracer equation

The tracer equation from Section 18.1.4 is given by

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}, \quad (21.114)$$

where \mathbf{J} is a subgrid scale flux. Now introduce the material time derivative operator in the form (21.48b) to have

$$\rho \left[\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla_\sigma C + \dot{\sigma} \partial_\sigma C \right] = -\nabla \cdot \mathbf{J}, \quad (21.115)$$

Multiplying by the specific thickness and making use of the mass conservation equation (21.83) renders the flux-form Eulerian equation

$$\frac{\partial(z_\sigma \rho C)}{\partial t} + \nabla_\sigma \cdot (z_\sigma \rho C \mathbf{u}) + \frac{\partial(\rho C w^{(\dot{\sigma})})}{\partial \sigma} = - \left[\nabla_\sigma \cdot (z_\sigma \mathbf{J}^h) + \frac{\partial(z_\sigma \nabla \sigma \cdot \mathbf{J})}{\partial \sigma} \right], \quad (21.116)$$

where we made use of expression (11.83) for the subgrid scale operator. Now perform a layer integral as detailed in Section 21.9 and use the layer mass continuity equation (21.96) to yield the layer integrated tracer equation

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_\sigma \cdot (h \rho C \mathbf{u}) + \Delta_\sigma(\rho C w^{(\dot{\sigma})}) = - \left[\nabla_\sigma \cdot (h \mathbf{J}^h) + \Delta_\sigma(z_\sigma \nabla \sigma \cdot \mathbf{J}) \right]. \quad (21.117)$$

Alternatively, we can bring all terms to the left hand side renders

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_\sigma \cdot (h \rho C \mathbf{u} + h \mathbf{J}^h) + \Delta_\sigma(\rho C w^{(\dot{\sigma})} + J^{(\sigma)}) = 0 \quad (21.118)$$

where we wrote

$$J^{(\sigma)} = z_\sigma \nabla \sigma \cdot \mathbf{J}. \quad (21.119)$$

We illustrate contributions to the layer tracer budget (21.118) in Figure 21.5. Note that we interpret these layer integrated fields and fluxes as per the discussion in Section 21.9.2.

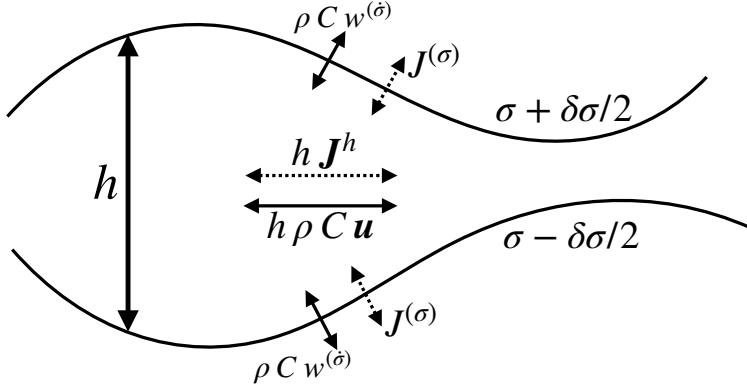


Figure 21.5: Illustrating the terms contributing to changes in layer tracer content according to the layer integrated tracer equation (21.118). The discrete layer is shown here with bounding interfaces at $\sigma - \delta\sigma/2$ and $\sigma + \delta\sigma/2$. Within a layer there is a redistribution 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 $J^{(\sigma)}$.

Part IV

Basics of geophysical fluid dynamics

We now bring together many pieces from earlier chapters to develop the basic equations and concepts of fluid dynamics for rotating and stratified fluids. That is, we here enter the world of geophysical fluid dynamics. These chapters are basic to the remainder of the book.

22

Momentum dynamics

In this chapter, we develop the fluid mechanical equations for linear and angular momentum relevant to the ocean and atmosphere. These equations of geophysical fluid dynamics (GFD) are based on Newton's laws of motion applied to a stratified fluid continuum moving on a rotating spherical planet. We make liberal use of results from classical point particle mechanics detailed in Chapter 14 and from the fluid kinematics discussed in the chapters of Part III. Relative to the point particle, the new dynamical feature afforded to the continuum concerns contact forces between fluid elements. These pressure and frictional forces arise from mechanical interactions among the continuum of fluid elements.

READER'S GUIDE TO THIS CHAPTER

This chapter develops the basic equations of geophysical fluid dynamics, building on our understanding of fluid mechanics and rotating physics. This material is fundamental to nearly all the subsequent chapters in this book.

22.1	Linear momentum equation	316
22.1.1	Body forces	316
22.1.2	Contact forces	316
22.1.3	Equation of motion	317
22.1.4	Summary of the thermo-hydrodynamical equations	318
22.1.5	Further study	320
22.2	Special forms for the momentum equation	320
22.2.1	Spherical coordinates	320
22.2.2	Vector-invariant velocity equation	321
22.3	Conditions for exact hydrostatic balance	322
22.4	Axial angular momentum	323
22.4.1	Axial angular momentum conserving motion of a ring of air	323
22.4.2	Sketching the atmospheric angular momentum budget	324
22.4.3	Further study	325

22.1 Linear momentum equation

We here summarize elements of classical continuum mechanics and in turn apply Newton's second law to derive the equation describing the momentum budget for a fluid continuum, with application to motion on a rotating and gravitating sphere. We present the momentum budget over both a finite volume region of the fluid and for an infinitesimal fluid element.

22.1.1 Body forces

Forces acting on an arbitrary volume, \mathcal{R} , of a continuous media are of two general kinds. The first kind involves *external* or *body* forces, such as gravitation (including tidal forces), Coriolis, and electromagnetic forces. These forces act throughout the extent of the media. Consequently, the total body force acting on a volume of fluid is the integral of the body force per unit mass, \mathbf{f}_{body} , multiplied by the mass of the media

$$\mathbf{F}_{\text{body}} = \int_{\mathcal{R}} \mathbf{f}_{\text{body}} \rho dV. \quad (22.1)$$

For example, the effective gravitational force (central gravity plus centrifugal) acting on a volume of fluid is given by

$$\mathbf{F}_{\text{gravity}} = \int_{\mathcal{R}} \mathbf{g} \rho dV, \quad (22.2)$$

where \mathbf{g} is the acceleration of gravity. Likewise, the Coriolis force is given by

$$\mathbf{F}_{\text{Coriolis}} = -2 \int_{\mathcal{R}} (\boldsymbol{\Omega} \wedge \mathbf{v}) \rho dV. \quad (22.3)$$

22.1.2 Contact forces

The second kind of forces are *internal* or *contact* forces, such as pressure forces and frictional forces. These forces act on a region of continuous media by affecting the boundary enclosing the region. The total contact force exerted on the region through its boundaries is given by

$$\mathbf{F}_{\text{contact}} = \oint_{\partial\mathcal{R}} \mathbb{T} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (22.4)$$

where $\hat{\mathbf{n}}$ is the outward normal direction orienting the domain boundary with $d\mathcal{S}$ the associated area element, and \mathbb{T} is the second order *stress tensor*. Contact forces are present in continuum matter but absent in point particle matter. Hence, they represent a fundamentally new element to the fluid dynamical equations relative to the equations of point particles detailed in Chapter 14.

Stresses from friction and pressure

As discussed in Chapter 26, there are two general types of stress that concern us: diagonal stresses associated with reversible momentum exchange through pressure, and stresses associated with irreversible exchange of momentum through friction. Therefore, we decompose the stress tensor components according to

$$\mathbb{T}^{ab} = \tau^{ab} - p g^{ab}. \quad (22.5)$$

In this equation, p is the pressure, which is a force per unit area acting in a compressive manner on the area of a surface. The second order tensor, g^{ab} , is the metric tensor that summarizes the local geometry, and it equals to the Kronecker or unit tensor for Cartesian coordinates in Euclidean space (Section 8.1). The frictional stress tensor is written τ^{ab} . It is also known as the *deviatoric* stress tensor as it represents deviations from the static case when stress is due solely to pressure. The friction stress tensor generally has zero trace, with pressure comprising the trace portion of the full stress tensor.

Substitution of the stress tensor (22.5) into the contact force expression (22.4) leads to

$$\mathbf{F}_{\text{contact}} = \int (\tau \cdot \hat{\mathbf{n}} - p \hat{\mathbf{n}}) d\mathcal{S}, \quad (22.6)$$

where the integral is taken over the bounding surface of the domain whose outward normal is $\hat{\mathbf{n}}$. Given this expression for contact forces acting on the boundary of a fluid domain, it is seen that positive pressure acts in the direction opposite to the surface's outward normal. That is, it acts in a compressive manner. Deviatoric stresses create more general forces on the bounding surface, which can have compressive, expansive, shearing, and/or rotational characteristics.

Exchange of momentum between fluid elements

We mathematically represent the exchange of momentum between fluid elements via a symmetric stress tensor. The divergence of the stress tensor then leads to a force acting on the fluid element boundaries. The forces arising from molecular viscosity provide an irreversible exchange of momentum that acts to reduce the kinetic energy of fluid elements. This process is dissipative and thus referred to as friction. Furthermore, when averaging over turbulent realizations of a fluid, the impacts on the mean flow are generally far larger than those associated with molecular viscosity. Nonetheless, these exchanges are also commonly parameterized via a symmetric stress tensor.

22.1.3 Equation of motion

The linear momentum of a fluid region is given by

$$\mathbf{P} = \int_{\mathcal{R}} \mathbf{v} \rho dV. \quad (22.7)$$

Applying Newton's law of motion to the continuum leads to the finite volume equation of motion

$$\frac{d}{dt} \int_{\mathcal{R}} \mathbf{v} \rho dV = \int_{\mathcal{R}} \rho \mathbf{f}_{\text{body}} dV + \oint_{\partial\mathcal{R}} \mathbb{T} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (22.8)$$

The time derivative can be either material, as for a constant mass fluid region moving with the barycentric velocity, or Eulerian, as for a fixed region in space (see Section 18.3). Applying Gauss's law (Section 4.7.2) to the area integral yields

$$\frac{d}{dt} \int_{\mathcal{R}} \rho \mathbf{v} dV = \int_{\mathcal{R}} (\rho \mathbf{f}_{\text{body}} + \nabla \cdot \mathbb{T}) dV. \quad (22.9)$$

General form of the equation of motion for a fluid element

Since the volume under consideration is arbitrary, the integral relation (22.9) is satisfied for an arbitrary region. We apply the result to an infinitesimal fluid element moving with the flow so that

$$\frac{D(\rho \mathbf{v} \delta V)}{Dt} = \delta V (\rho \mathbf{f}_{\text{body}} + \nabla \cdot \mathbb{T}). \quad (22.10)$$

Assuming the mass for the fluid element is constant then leads to the equation of motion

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f}_{\text{body}} + \nabla \cdot \mathbb{T}. \quad (22.11)$$

Momentum equation for a rotating fluid in a gravitational field

Now specialize the momentum equation (22.11) to suite our needs. We first write the stress tensor in terms of the deviatoric component from friction and a diagonal component from pressure (equation (22.6))

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f}_{\text{body}} - \nabla p + \nabla \cdot \boldsymbol{\tau}. \quad (22.12)$$

Next, move to a rotating terrestrial reference frame and thus expose the Coriolis acceleration and the effective gravitational force (Section 14.2)

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -\rho \nabla \Phi - \nabla p + \nabla \cdot \boldsymbol{\tau}. \quad (22.13)$$

This form of the equation of motion arises from extracting the solid-body motion of the basis vectors to define the Coriolis acceleration (see Section 13.11). Any remaining changes to the basis vectors arise from motion of the fluid relative to the solid-body rotating reference frame, and thus appear when expanding the material time derivative. The form (22.13) for the equation of motion offers a suitable starting point for studies of geophysical fluid dynamics. It often goes by the name of *Navier-Stokes* equation. However, that name is more correctly applied to the non-rotating case with a specific form for the friction operator. We thus refer to equation (22.13) as Newton's law of motion for a rotating continuous fluid.

22.1.4 Summary of the thermo-hydrodynamical equations

The full suite of equations describing rotating and stratified fluids consists of the dynamical equations of motion (Newton's second law), along with mass continuity, the potential temperature equation, material tracer equations, and an equation of state for density. We term these the

thermo-hydrodynamical equations of motion and write them in the form

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -\rho \nabla \Phi - \nabla p + \nabla \cdot \tau \quad \text{momentum} \quad (22.14)$$

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v} \quad \text{mass continuity} \quad (22.15)$$

$$\rho \frac{D\theta}{Dt} = -\nabla \cdot \mathbf{J}(\theta) \quad \text{heat conservation} \quad (22.16)$$

$$\rho \frac{DS}{Dt} = -\nabla \cdot \mathbf{J}(S) \quad \text{matter conservation} \quad (22.17)$$

$$\rho = \rho(S, \theta, p) \quad \text{equation of state.} \quad (22.18)$$

It is a testament to the success of classical continuum mechanics that these equations are of use for describing fluid phenomena from the millimetre scale to the astrophysical scale. We summarize the following terms in these equations.

- **VELOCITY:** The velocity field, \mathbf{v} , contains three prognostic components. Each velocity component evolves according to its respective dynamical equation (22.14). As noted at the end of Section 22.1.3, we write the momentum equation in the form (22.14) by separating the time dependence of the basis vectors into a term arising from solid-body rotation (which leads to Coriolis and centrifugal accelerations) and a term arising from the motion of the fluid relative to the rotating sphere (which leads to the metric acceleration when using spherical coordinates).
- **TRACERS:** The potential temperature and matter concentration have corresponding prognostic equations that evolve the fields forward in time. Furthermore, they have corresponding fluxes, \mathbf{J} . The flux has a form specified by molecular diffusion as discussed in Section 37.2 (or by other parameterized processes when the theory has a space cutoff larger than millimetres).
- **DENSITY:** Density can be updated in time via mass continuity (equation (22.15)). We discussed the many forms of density for the ocean and atmosphere in Section 25.2.
- **PRESSURE:** There is no prognostic equation for pressure. Rather, pressure is diagnosed based on knowledge of other fields. Here are sketches of how that diagnostic calculation is performed.
 - For an ideal gas, pressure is diagnosed from the ideal gas relation (23.54) using the density and temperature.
 - For a hydrostatic fluid (Section 27.3), pressure is diagnosed at a point through knowledge of the weight per area above the point (i.e., the mass density).
 - For an incompressible liquid, pressure is diagnosed by solving a Poisson equation derived from taking the divergence of the momentum equation (see Exercise 28.2).
- **GEOPOTENTIAL:** The geopotential is specified once the height above an arbitrary reference level is known, as well as the effective gravitational acceleration (Section 14.1.2). For geophysical fluid studies, the reference level is generally taken at the level of a resting sea surface. We thus often write the radial coordinate as

$$r = R + z \quad (22.19)$$

where $R = 6.371 \times 10^6$ m is the earth radius (equation (14.9)), and z is the geopotential coordinate measuring the height above sea level.

- ANGULAR ROTATION: The earth's angular velocity, Ω , is constant for geophysical fluid studies of concern here. Its value is discussed in Section 13.1.
- FRICTION: The friction vector,

$$\rho \mathbf{F} = \nabla \cdot \tau, \quad (22.20)$$

is the divergence of a symmetric and trace-free deviatoric stress tensor, τ (Section 22.1.2). It is specified in Section 37.2.5 for molecular viscosity. More general forms for the friction vector can be considered for purposes of subgrid-scale modeling (e.g., see Part 5 of *Griffies (2004)*).

- BOUNDARY CONDITIONS: Boundary conditions consist of the exchange of matter, momentum, and enthalpy with the surrounding media, such as the solid earth or another fluid component (e.g., atmosphere-ocean exchange). We discuss the boundary conditions for matter in Section 17.4 and for momentum in Section 26.9.

22.1.5 Further study

Chapter 5 of *Aris (1962)* offers an insightful discussion of continuum mechanics as applied to a fluid. Section 2.2 *Vallis (2017)* provides a thorough derivation of the dynamical equations of motion for the atmosphere and ocean. We offer further discussion of the mathematics and physics of stress in fluids (including pressure) in Chapter 26. Fundamentals of geopotential coordinates and sea level are summarized in *Gregory et al. (2019)*.

22.2 Special forms for the momentum equation

We here display some special forms of the momentum equation commonly found in our study of geophysical fluid mechanics.

22.2.1 Spherical coordinates

Geophysical fluids move on a rotating planet with the planet commonly assumed to have a spherical geometry. We thus find it useful to display the spherical coordinate form for the equations of motion. For this purpose, make use of the acceleration as derived in Section 14.2.3 for the point particle. In particular, use the geopotential coordinate to measure radial distances from the center of the sphere (equation (22.19)), as well as the longitude and latitude angular coordinates defined by Figure 10.1. Notably, the point particle time derivative translates into a material time derivative for fluid elements. We are thus led to

$$\frac{Du}{Dt} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) = -\frac{1}{\rho r \cos \phi} \frac{\partial p}{\partial \lambda} + F^\lambda \quad (22.21)$$

$$\frac{Dv}{Dt} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi = -\frac{1}{\rho r} \frac{\partial p}{\partial \phi} + F^\phi \quad (22.22)$$

$$\frac{Dw}{Dt} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi = -g - \frac{1}{\rho} \frac{\partial p}{\partial r} + F^r, \quad (22.23)$$

where we introduced the spherical components to the friction acceleration

$$\mathbf{F} = F^\lambda \hat{\lambda} + F^\phi \hat{\phi} + F^r \hat{r}. \quad (22.24)$$

We also note the spherical coordinate form for the gradient operator

$$\nabla = \frac{\hat{\lambda}}{r \cos \phi} \frac{\partial}{\partial \lambda} + \frac{\hat{\phi}}{r} \frac{\partial}{\partial \phi} + \hat{r} \frac{\partial}{\partial r}, \quad (22.25)$$

as well as the material time derivative operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla = \frac{\partial}{\partial t} + \frac{u}{r \cos \phi} \frac{\partial}{\partial \lambda} + \frac{v}{r} \frac{\partial}{\partial \phi} + w \frac{\partial}{\partial r}. \quad (22.26)$$

We can write the spherical momentum equations in a bit more compact form by introducing the spherical coordinate velocity field (see equation (13.41b))

$$\mathbf{v} = \mathbf{u} + \hat{r} w = u \hat{\lambda} + v \hat{\phi} + w \hat{r} \quad (22.27)$$

and the corresponding spherical coordinate acceleration

$$\mathbf{A}_{\text{sphere}} = \frac{Du}{Dt} \hat{\lambda} + \frac{Dv}{Dt} \hat{\phi} + \frac{Dw}{Dt} \hat{r}. \quad (22.28)$$

We also introduce the expression (13.71c) for the metric acceleration to render

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -\rho \nabla \Phi - \nabla p + \rho \mathbf{F}, \quad (22.29)$$

where we have the acceleration relative to the rotating frame

$$\frac{D\mathbf{v}}{Dt} = \mathbf{A}_{\text{sphere}} + \frac{1}{r} [u \tan \phi (\hat{r} \wedge \mathbf{v}) + w \mathbf{u} - \hat{r} \mathbf{u} \cdot \mathbf{u}]. \quad (22.30)$$

For some purposes it is convenient to combine one piece of the metric acceleration to the Coriolis acceleration to yield

$$\mathbf{A}_{\text{sphere}} + \frac{1}{r} [w \mathbf{u} - \hat{r} \mathbf{u} \cdot \mathbf{u}] + \left(2\boldsymbol{\Omega} + \frac{u \tan \phi \hat{r}}{r} \right) \wedge \mathbf{v} = -\left(\frac{\rho \nabla \Phi + \nabla p}{\rho} \right) + \mathbf{F}. \quad (22.31)$$

22.2.2 Vector-invariant velocity equation

The metric terms appearing in the momentum equation (22.31) are those terms proportional to r^{-1} that arise from spatial dependence of the spherical unit vectors. An alternative formulation removes these terms in favor of the vorticity and kinetic energy. For that purpose we make use of the identity (equation (4.37)) for the nonlinear self-advection term

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = \boldsymbol{\omega} \wedge \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2, \quad (22.32)$$

where

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (22.33)$$

is the vorticity discussed in Part VIII of the book. We derive the corresponding *vector-invariant* form of the velocity equation using Cartesian coordinates and then invoke general covariance (Section 7.2) to extend the result to arbitrary coordinates.¹ Making use of equation (22.32) thus leads to the material acceleration

$$\frac{D\mathbf{v}}{Dt} = \frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega} \wedge \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2 \quad (22.34)$$

so that the momentum equation (22.13) becomes the vector-invariant velocity equation

$$\frac{\partial \mathbf{v}}{\partial t} + (2\boldsymbol{\Omega} + \boldsymbol{\omega}) \wedge \mathbf{v} = -\nabla(\Phi + \mathbf{v} \cdot \mathbf{v}/2) + (1/\rho)(-\nabla p + \nabla \cdot \tau). \quad (22.35)$$

This equation forms the basis for Bernoulli's theorem in Section 24.2.3 as well as vorticity equation in Chapter 48.

¹See Section 4.4.4 of [Griffies \(2004\)](#) for a detailed derivation using general coordinates.

22.3 Conditions for exact hydrostatic balance

We are mostly interested in moving fluids within this book. Even so, it is useful to expose the signature of a static fluid supporting the trivial solution, $\mathbf{v} = 0$. The equations of motion (22.14)-(22.18) possess an exact static solution so long as the pressure gradient force balances the gravitational force

$$\nabla p = -\rho \nabla \Phi, \quad (22.36)$$

and where the frictional stress tensor has zero divergence when the flow is static. Equation (22.36) constitutes the *hydrostatic fluid* solution. As justified in Chapter 27, the hydrostatic balance is a very good approximation for the vertical momentum equation in large-scale geophysical fluids, even when those fluids are moving. For the current considerations, we are interested in a purely static flow so that the hydrostatic equation (22.36) is exact.

We make the following observations of the exact hydrostatic balance (22.36).

- Since ∇p is directly proportional to $\nabla \Phi$, surfaces of constant pressure in a static fluid correspond to surfaces of constant geopotential.
- Since the curl of the pressure gradient vanishes, a static fluid maintains its density gradients parallel to geopotential gradients

$$\nabla \rho \wedge \nabla \Phi = 0, \quad (22.37)$$

which in turn means that density surfaces are parallel to geopotentials so that

$$\rho = \rho(\Phi) \quad \text{static fluid.} \quad (22.38)$$

For the geopotential $\Phi = g z$, a static fluid is realized if the density has no horizontal gradients; i.e., it is everywhere flat so that

$$\rho = \rho(z) \quad \text{static fluid with } \Phi = g z. \quad (22.39)$$

If the density gradient has any component perpendicular to $\nabla \Phi$, then pressure forces will affect fluid flow thus implying that the fluid is not in exact hydrostatic balance.

- Projecting both sides of equation (22.36) onto an infinitesimal space increment, $d\mathbf{x}$, renders

$$d\mathbf{x} \cdot \nabla p = -\rho d\mathbf{x} \cdot \nabla \Phi \implies \frac{dp}{d\Phi} = -\rho. \quad (22.40)$$

Hence, the difference in pressure between any two geopotential is given by the integral

$$p(\Phi_2) - p(\Phi_1) = - \int_{\Phi_1}^{\Phi_2} \rho(\Phi) d\Phi. \quad (22.41)$$

Again, if $\Phi = g z$, then we recover the familiar expression for the hydrostatic pressure field

$$p(z_2) - p(z_1) = -g \int_{z_1}^{z_2} \rho(z) dz, \quad (22.42)$$

so that the difference in hydrostatic pressure is given by the weight per horizontal area of fluid between the two geopotentials.

22.4 Axial angular momentum

As in our discussion of a point particle in Section 15.7, the axial angular momentum of a fluid element is given by

$$L^z = (\rho \delta V) r_{\perp} (u + r_{\perp} \Omega) \equiv (\rho \delta V) l^z \quad (22.43)$$

where

$$l^z = r_{\perp} (u + r_{\perp} \Omega) \quad (22.44)$$

is the axial angular momentum per unit mass, and the distance to the polar rotation axis,

$$r_{\perp} = r \cos \phi \quad (22.45)$$

is the moment-arm for determining the torques acting on a fluid element. Making use of the zonal momentum equation (22.21) leads to the material time change

$$\frac{Dl^z}{Dt} = (u + 2\Omega r_{\perp}) \frac{Dr_{\perp}}{Dt} + r_{\perp} \frac{Du}{Dt} \quad (22.46a)$$

$$= (u + 2\Omega r_{\perp}) \frac{Dr_{\perp}}{Dt} + (u + 2\Omega r_{\perp})(v \sin \phi - w \cos \phi) - \frac{1}{\rho} \frac{\partial p}{\partial \lambda} \quad (22.46b)$$

$$= (u + 2\Omega r_{\perp}) \left[\frac{Dr_{\perp}}{Dt} + v \sin \phi - w \cos \phi \right] - \frac{1}{\rho} \frac{\partial p}{\partial \lambda} \quad (22.46c)$$

$$= -\frac{1}{\rho} \frac{\partial p}{\partial \lambda}, \quad (22.46d)$$

so that

$$\rho \frac{Dl^z}{Dt} = -\frac{\partial p}{\partial \lambda} \Rightarrow \frac{\partial(\rho l^z)}{\partial t} + \nabla \cdot (\rho l^z \mathbf{v}) = -\frac{\partial p}{\partial \lambda}. \quad (22.47)$$

Comparing to the result for a point particle in Section 15.7 reveals that the axial angular momentum of a fluid element is affected by the zonal pressure gradient. Hence, in the absence of a zonal pressure gradient the axial angular momentum is materially invariant just like for the point particle. In particular, the physical results described in Section 15.8 hold whereby we can equate the zonal Coriolis acceleration to the zonal acceleration induced by axial angular momentum conservation. For example, a fluid element initially at rest in a fluid with zero zonal pressure gradient will zonally accelerate when moved meridionally (e.g., as from a meridional pressure gradient) according to the needs of axial angular momentum conservation.

22.4.1 Axial angular momentum conserving motion of a ring of air

Atmospheric and oceanic flows rarely experience a zero zonal pressure gradient along the trajectory of a fluid element. However, on a smooth spherical planet without meridional boundaries there is a zero zonally integrated zonal pressure gradient

$$\int_{\text{ring}} \frac{\partial p}{\partial \lambda} d\lambda = 0. \quad (22.48)$$

Hence, a constant mass ring of fluid circling the planet (Figure 22.1) will preserve its axial angular momentum in the absence of friction

$$\frac{d}{dt} \oint_{\text{ring}} \rho l^z dV = \oint_{\text{ring}} \rho \frac{Dl^z}{Dt} dV = 0. \quad (22.49)$$

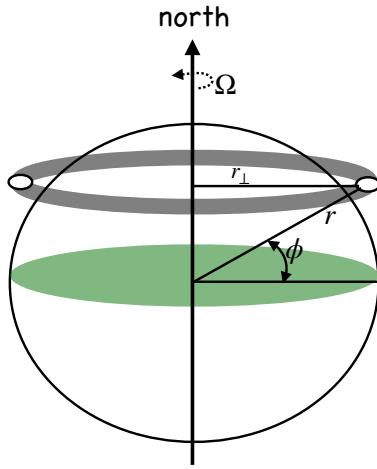


Figure 22.1: A ring of inviscid air circulating around a constant latitude circle over a smooth planet. This ring conserves its axial angular momentum. Consequently, axial angular momentum conserving motion of the ring induces a zonal acceleration if the ring alters its distance from the rotation axis, r_{\perp} , by moving meridionally or vertically.

For example, consider a latitudinal ring of constant mass inviscid fluid circling around the equator at radius R_0 , and assume it is at rest relative to the rotating earth. The angular momentum per mass for this ring is $l^z = \Omega R_0^2$. Moving the ring vertically to $r \neq R_0$ while maintaining constant axial angular momentum induces a zonal velocity given by

$$u = \frac{\Omega (R_0^2 - r^2)}{r}. \quad (22.50)$$

Movement downward to $r < R_0$ leads to an eastward flow, $u > 0$, (westerly winds), whereas upward motion leads to the opposite. Likewise, moving the ring latitudinally while keeping $r = R_0$ leads to the zonal velocity

$$u = \frac{\Omega R_0 \sin^2 \phi}{\cos \phi}. \quad (22.51)$$

Since $\cos \phi \geq 0$ on the sphere, latitudinal motion away from the equator while preserving axial angular momentum leads to eastward flow ($u > 0$) whether the ring is moved northward or southward.

22.4.2 Sketching the atmospheric angular momentum budget

How realistic is it to have coherent rings of inviscid air circulating around the planet at all latitudes? To help answer this question, consider a ring of radius $R_0 = R_e$ that starts with zero relative velocity at the equator. Equation (22.51) says that the westerly winds induced by axial angular momentum conserving motion have the following speeds at a selection of latitudes

$$u(10^\circ) = 14 \text{ m s}^{-1} \quad u(20^\circ) = 58 \text{ m s}^{-1} \quad u(30^\circ) = 134 \text{ m s}^{-1}. \quad (22.52)$$

The values at higher latitudes are unbounded since $\cos \phi \rightarrow 0$ as the poles are approached. So there is a problem with an idealized theory of atmospheric circulation based on axial angular momentum conserving rings of air. In fact, these ideas extend only so far as the Hadley circulation is concerned, with the latitude extents of the Hadley Cell extending only to the middle latitudes. Essential missing elements include frictional dissipation between the atmosphere and land, and baroclinic eddies that contribute to poleward transport of angular momentum. It is outside of our scope to detail these

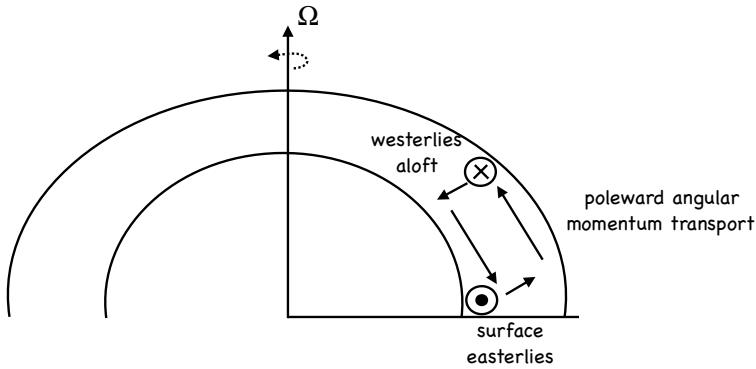


Figure 22.2: A sketch of the Hadley circulation. At the low latitudes, low level winds are predominantly easterlies and so the atmosphere gains axial angular momentum from the earth via frictional boundary layer processes. As air rises and moves poleward it zonally accelerates according to conservation of axial angular momentum, thus becoming the mid-latitude westerlies. The westerlies are generally unstable thus producing baroclinic eddies that vertically transfer axial angular momentum via form stress, thus depositing axial angular momentum back to the earth to close the axial angular momentum cycle.

physical processes and the corresponding atmospheric circulation. Instead, we merely sketch the scene following Section 10.3 of [Holton \(1992\)](#) and Section 8.2 of [Marshall and Plumb \(2008\)](#).

Given that the earth has a near constant rotation rate, we can examine a zonally integrated axial momentum budget for the atmosphere and assume that the angular momentum of the earth is fixed.² Rather than an inviscid fluid, we consider the dissipation due to boundary layer friction and/or vertical eddy form stress from baroclinic eddies (see Section 26.8 for a discussion of form stress). In the tropical atmosphere, the atmosphere generally has less angular momentum than the earth given that the winds are predominantly easterly (lower branch of the Hadley Cell). In this region, boundary layer interactions transfer angular momentum from the earth to the atmosphere. In contrast, the presence of middle latitude westerlies aloft (upper branch of the Hadley Cell) signal that the atmosphere has more angular momentum than the earth at these latitudes. There is an implied transfer of angular momentum from the atmosphere to the earth, largely mediated through vertical form stresses due to baroclinic eddies.

As sketched in Figure 22.2, a steady state axial angular momentum budget is realized through a meridional transport of axial angular momentum from the lower latitude atmosphere to the middle latitudes. This transport is mediated partly through an atmospheric overturning circulation (Hadley Cell) and partly through the vertical transfer of momentum via synoptic scale baroclinic eddies (see Section 39.2.4). Furthermore, dissipation by eddies and boundary layer interactions is crucial to realize the observed zonal wind profile. Namely, as seen above, a non-dissipative angular momentum conserving atmosphere exhibits winds that are far larger in magnitude than observed.

22.4.3 Further study

In Section 26.10 we consider the zonally integrated axial angular momentum budget for the ocean with sloping solid-earth bottom boundary as well as the upper surface (ocean) boundary. In particular, we see how boundary form stresses (Section 26.8) affect the angular momentum in addition to boundary frictional stresses. That analysis is analogous to that for the atmosphere given in Section 10.3 of [Holton \(1992\)](#).

²In fact, the length of a day fluctuates by roughly 10^{-3} s over a seasonal cycle due to transfer of angular momentum between the atmosphere and land.

23

Thermodynamics

Thermodynamics is a phenomenological discipline whose fundamentals lie in statistical mechanics. When applied to continuum fluid mechanics, we assume fluid elements to be in *local thermodynamic equilibrium*. This assumption is based on the quasi-static nature of macroscopic motions (Section 23.2.2) whereby microscopic motions have a much shorter equilibration time as compared to the relatively slow macroscopic processes of interest for fluid mechanics such as advection, waves, turbulence, and mixing. Hence, we make use of phenomenological thermodynamic laws to develop evolution equations for thermodynamic properties of continuum fluid elements. Consistent with the huge time scale separation between microscopic and macroscopic processes, we can accurately assume that the macroscopic motion of a fluid element does not alter its entropy. That is, advective transport is a reversible process. In contrast, mixing of properties between fluid elements is irreversible and thus increases entropy.

Use of equilibrium thermodynamics for time dependent phenomena falls under the discipline of *quasi-equilibrium thermodynamics*, also called *linear irreversible thermodynamics*. In this chapter, we make use of this discipline as applied to a continuum fluid. In particular, we explore various thermodynamic properties of the fluid, including the specific heat, lapse rate, and potential temperature. We present relations for an ideal gas, which well approximates the dry atmosphere, and further relations for the more general case of a binary liquid such as seawater.

READER'S GUIDE FOR THIS CHAPTER

One aim for this chapter is to expose the basics of equilibrium and quasi-equilibrium thermodynamics without diving too deeply into details. Alas, this goal is elusive since a deductive approach involves many concepts and mathematical manipulations. Although aiming to offer a self-contained presentation, the reader is expected to have some prior exposure to thermodynamics thus allowing for shortcuts to be exploited. Thermodynamics is treated in many areas of science, such as physics, chemistry, biology, and in most areas of engineering. A treatment based on the postulates of thermodynamics is given by [Callen \(1985\)](#). [Reif \(1965\)](#) offers an elementary connection between thermodynamics and statistical mechanics, with [Huang \(1987\)](#) and [Reichl \(1987\)](#) offering even more details. Linear irreversible thermodynamics has been formulated for fluid mechanics by [DeGroot and Mazur \(1984\)](#) and [Landau and Lifshitz \(1987\)](#).

23.1	Open threads	328
23.2	The First Law of thermodynamics	329
23.2.1	The First Law in its extensive form	329
23.2.2	Quasi-static processes	329
23.2.3	Mechanical work, heat, and chemical work	330
23.2.4	Fundamental thermodynamic relation	331
23.2.5	A note on partial derivatives	332
23.2.6	Internal energy, homogeneous functions and Gibbs-Duhem	332
23.2.7	Fundamental thermodynamic relation per unit mass	332
23.2.8	The special case of seawater as a binary fluid	333
23.3	Thermodynamic potentials	333
23.3.1	Internal energy	333
23.3.2	Entropy	334
23.3.3	Enthalpy	334
23.3.4	Gibbs potential	335
23.3.5	Combined fundamental thermodynamic relation	335
23.3.6	Alternate functional dependencies	335
23.3.7	Further study	336
23.4	Specific heat capacity	336
23.5	A simple ideal gas atmosphere	337
23.5.1	Equation of state	337
23.5.2	Internal energy	337
23.5.3	Heat capacity from statistical mechanics	338
23.5.4	Enthalpy	339
23.5.5	Thermal expansion coefficient	339
23.5.6	Fundamental thermodynamic relations	339
23.5.7	Further study	340
23.6	Adiabatic lapse rate	340
23.6.1	Isentropic rearrangement	340
23.6.2	Thermodynamic formulation	340
23.6.3	Adiabatic lapse rate for pressure changes	341
23.6.4	Adiabatic lapse rate for height changes	341
23.6.5	Adiabatic lapse rate for an ideal gas atmosphere	342
23.6.6	Further study	342
23.7	Potential temperature	343
23.7.1	Adiabatic temperature changes	343
23.7.2	Defining the potential temperature	344
23.7.3	Potential temperature and specific entropy	344
23.7.4	Potential temperature for an ideal gas	345
23.8	Thermodynamics of a moving fluid	346
23.8.1	First Law for a moving fluid element	346
23.8.2	The First Law in terms of potential temperature	348
23.8.3	Potential temperature and numerical models	348
23.8.4	Further study	349
23.9	Exercises	349

23.1 Open threads

- potential temp versus in situ for the ocean and their evolution
- potential enthalpy

- Available internal energy
- Available potential energy
- Moist static energy as per David Romps' notes

23.2 The First Law of thermodynamics

As discussed in Chapter 2, there are a huge number of microscopic (molecular) degrees of freedom that are averaged over when describing a fluid as a continuous media. Internal energy embodies the energy of microscopic degrees of freedom not explicitly considered in a macroscopic continuum treatment. Internal energy arises from the translational kinetic energy of molecular motion, together with their internal degrees of freedom associated with rotation and vibration as well as intermolecular forces between molecules. When accounting for the total energy of a fluid system, we must include its internal energy due to microscopic degrees of freedom as well as the mechanical energy (kinetic and gravitational potential) arising from macroscopic degrees of freedom. The First Law of thermodynamics offers a means to account for changes to the internal energy of a thermodynamic system.

23.2.1 The First Law in its extensive form

The First Law of thermodynamics establishes a relationship between infinitesimal changes of internal energy within a thermodynamic system, the work done to the internal (i.e., molecular) degrees of freedom of a system, the heat (thermal energy) transferred to the system, and changes in the matter composition. The First Law takes on the mathematical form

$$d\mathcal{I}^e = dQ + dW + dC \quad \text{SI units Joule} = \text{kg m}^2 \text{s}^{-2}. \quad (23.1)$$

In this equation, $d\mathcal{I}^e$ is the infinitesimal increment (or the differential) of the system's internal energy; dQ is the internal energy change due to thermal energy (heat) transferred to the system; dW is the change in internal energy due to work applied to the system; and dC is the change in internal energy due to changes in the matter content (chemical work). Each term in the First Law (23.1) has units of energy: Joule = kg m² s⁻². Furthermore, the internal energy is proportional to the size of the system, with systems having more volume and mass having more internal energy (see Section 23.2.6). As discussed in Section 18.3.1, properties whose value changes when the system size changes are termed *extensive*. Extensive properties are labeled with the superscript *e* in the following (except for the mass and volume), with this label *not* representing a tensor index.

23.2.2 Quasi-static processes

In thermodynamics we find it useful to make use of a huge time scale separation when considering the changes considered for a macroscopic system. We are led to define a quasi-static process as one that proceeds via an infinite number of intermediate equilibrium states. This definition is an idealization of the actual situation in which macroscopic processes generally proceed in time via macroscopic time increments, $\Delta\tau_{\text{macro}}$, that are much larger than the time required for molecular equilibration. We estimate the molecular equilibration time by the time between molecular collisions. As seen in Section 2.2.5, the time between molecular collisions is less than 10^{-9} s, which is incredibly tiny. Hence, any macroscopic process moves through time increments that are much larger than the molecular collision time

$$\Delta\tau_{\text{macro}} \gg 10^{-9} \text{ s}. \quad (23.2)$$

This time scale separation means that the molecular motions have no problem equilibrating between each of the $\Delta\tau_{\text{macro}}$ time increments. In fluid mechanics, we assume that all fluid motions occur in a quasi-static manner, again since any of the relevant time scales for macroscopic fluid motions are much larger than the molecular time scales.

23.2.3 Mechanical work, heat, and chemical work

We here identify features of mechanical work, heat, and chemical work and their role in geophysical fluid mechanics. The forms appear for quasi-static processes.

Mechanical work

There are many ways that mechanical forces can do work to a system. For fluid mechanics we generally focus on changes to the volume of a fluid element through the action of pressure. In this case, the pressure-work term takes the form

$$dW = -p dV \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}, \quad (23.3)$$

where p is the pressure applied to the boundaries of the fluid element. The negative sign in the pressure-work relation (23.3) arises since compressing the fluid into a smaller volume ($dV < 0$) requires positive mechanical work be applied to the fluid ($dW > 0$). In fluid mechanics, this mechanical work term is referred to as *pressure-work*.

Pressure is an intensive variable that measures the *intensity* of a force conjugate to the extensive variable V . Pressure is also the integrating factor connecting the inexact differential dW to the exact differential dV . Finally, we note that incompressible fluid elements do not alter their volume (Chapter 19), so that there is no pressure-work applied to incompressible fluids.

Heat

If we assume that the heat transferred to a fluid occurs in a quasi-static manner, then it can be related to changes in entropy via

$$dQ = T dS^e \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}. \quad (23.4)$$

In this equation, T is the absolute temperature (measured relative to absolute zero), which is an intensive variable, whereas S^e is the entropy, which is an extensive variable. Entropy is also a state function, with T providing the integrating factor connecting the inexact differential dQ to the exact differential dS^e .

Chemical work

Work due to changes in the chemical composition of a thermodynamic system are written as

$$dC = \sum_n \mu_n dM_n \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}, \quad (23.5)$$

where dM_n are changes to the matter content for constituent n , and μ_n is the corresponding chemical potential.

Work and heat are processes rather than state properties

Work and heat refer to path-dependent thermodynamic processes that transform a system from one thermodynamic state to another. Work and heat are not state properties, which is the fundamental reason for the inexact nature of their infinitesimal changes. Correspondingly, work and heat denote actions (verbs) rather than properties (nouns). This distinction prompts some to refer to “working” and “heating” rather than “work” and “heat”. Correspondingly, it is not relevant to seek information about the “work content” or “heat content” of a fluid state.

This point is particularly relevant when asking questions about the heat transported by a fluid (with units energy per time: Watt = Joule/s). In heat budget analyses, it is tempting to define the “heat content” of a fluid element according to its temperature, mass, and heat capacity. But the notion of heat content spuriously conflates a thermodynamic process (heat) with a thermodynamic state property (internal energy or enthalpy). Furthermore, in practice any definition of heat content is ambiguous due to the arbitrariness of the temperature scale, such as Celsius or Kelvin. Therefore, when working with heat transport, care should be exercised if also including the notion of heat content. One way to detect an error in the analysis is to ask whether a particular conclusion is modified by changing the temperature scale. If so, then one should carefully revisit assumptions of the analysis.

23.2.4 Fundamental thermodynamic relation

Assuming quasi-static processes, substitution of relations (23.3), (23.4), and (23.5) into the First Law (23.1) leads to

$$d\mathcal{I}^e = T dS^e - p dV + \sum_n \mu_n dM_n. \quad (23.6)$$

Notably, there are no inexact differentials in this equation. Rather, it provides a relation between infinitesimal changes in thermodynamic state functions. Although derived for quasi-static processes from the First Law using connections to work and heat, equation (23.6) holds for arbitrary infinitesimal changes. It therefore offers great utility in the formalism of thermodynamics, even though its connection to the First Law holds only for quasi-static processes.

Equation (23.6) is known as the *fundamental thermodynamic relation* in terms of internal energy. It is the starting point for many mathematical manipulations in thermodynamics. In particular, it leads to the following identities

$$\left[\frac{\partial \mathcal{I}^e}{\partial S^e} \right]_{V, M_n} = T \quad (23.7)$$

$$\left[\frac{\partial \mathcal{I}^e}{\partial V} \right]_{S^e, M_n} = -p \quad (23.8)$$

$$\left[\frac{\partial \mathcal{I}^e}{\partial M_n} \right]_{S^e, V} = \mu_n. \quad (23.9)$$

In these expressions, partial derivatives are taken with the noted variables held constant. Each equation relates an intensive property (right hand side) to the partial derivative of internal energy with respect to an extensive property. In thermodynamics, these equations are known as *equations of state*. However, in the geophysical fluid literature, the equation of state generally refers to the expression for mass density in terms of fluid properties (see Section 25.2).

23.2.5 A note on partial derivatives

Thermodynamics contains a plethora of partial derivatives. Recall that partial derivatives are defined with the complement variables held fixed during the differentiation. Hence, so long as we are clear about functional dependence, extra subscripts such as those exposed in the equations of state (23.7)–(23.9) are not needed for the partial derivatives. Nonetheless, traditional thermodynamic notation exposes all of the subscripts in order to remain explicit about the dependent and independent variables. Such notation, though clumsy, can be essential when in the midst of manipulations with thermodynamic potentials and their derivatives.

23.2.6 Internal energy, homogeneous functions and Gibbs-Duhem

The fundamental thermodynamic relation (23.6) indicates that internal energy is naturally considered a function of entropy, volume, and constituent mass

$$\mathcal{I}^e = \mathcal{I}^e(\mathcal{S}^e, V, M_n). \quad (23.10)$$

Now scale the system by an arbitrary parameter λ . Under this operation, the extensive variables entropy, volume, and mass scale by the same scale factor. Through the fundamental thermodynamic relation (23.6), the internal energy scales likewise, giving

$$\mathcal{I}^e(\lambda \mathcal{S}^e, \lambda V, \lambda M_n) = \lambda \mathcal{I}^e(\mathcal{S}^e, V, M_n). \quad (23.11)$$

A function that scales in this way is termed a *homogeneous function of degree one*. Differentiating both sides of this identity with respect to λ , setting λ to unity, and using the partial derivative identities (23.7)–(23.9) yields

$$\mathcal{I}^e = T \mathcal{S}^e - p V + \sum_n \mu_n M_n. \quad (23.12)$$

This result represents a special case of Euler's theorem of homogeneous functions. Taking the differential of this equation and using the fundamental thermodynamic relation (23.6) leads to the *Gibbs-Duhem* relation

$$\mathcal{S}^e dT - V dp + \sum_n M_n d\mu_n = 0. \quad (23.13)$$

23.2.7 Fundamental thermodynamic relation per unit mass

For many purposes in fluid mechanics, it proves convenient to consider thermodynamic relations for a system of unit mass; i.e., “per unit mass”. For this purpose, one scales away the mass of the system by setting the scale factor $\lambda = M^{-1}$ and introducing the *specific* quantities

$$\mathcal{I}^e = M \mathcal{J} \quad (23.14)$$

$$\mathcal{S}^e = M \mathcal{S} \quad (23.15)$$

$$V^e = M \nu \quad (23.16)$$

$$M_n = M C_n, \quad (23.17)$$

where

$$\nu = 1/\rho \quad (23.18)$$

is the specific volume. In the last equality, C_n is the mass fraction or concentration of species n in the fluid, with this *tracer concentration* also introduced in Section 18.1. Substituting the specific

quantities (23.14)-(23.17) into the first Gibbs relation (23.6) and using expression (23.12) for the internal energy leads to the fundamental thermodynamic relation in terms of specific thermodynamic quantities

$$d\mathcal{J} = T dS - p d\nu + \sum_n \mu_n dC_n. \quad (23.19)$$

We make use of this form of the fundamental thermodynamic relation in the following.

23.2.8 The special case of seawater as a binary fluid

Considering seawater to be a binary system of salt and fresh water, we have

$$C_{\text{salt}} + C_{\text{water}} = 1. \quad (23.20)$$

Introducing this constraint into the first Gibbs relation (23.19) leads to

$$d\mathcal{J} = T dS - p d\nu + \mu dC, \quad (23.21)$$

where $C = C_{\text{salt}}$ is the concentration of salt, and

$$\mu = \mu_{\text{salt}} - \mu_{\text{water}} \quad (23.22)$$

is the relative chemical potential. The *absolute salinity* S , with units parts per thousand, is related to C_{salt} via

$$S = 1000 C_{\text{salt}}. \quad (23.23)$$

The range of salinity in the ocean (roughly, $0 \leq S \leq 40$) is more convenient than the range of C_{salt} , making salinity more commonly used in oceanography.

23.3 Thermodynamic potentials

The state functions internal energy and entropy are also known as thermodynamic potentials. Each thermodynamic potential is a natural function of certain other thermodynamic properties, as defined by the fundamental thermodynamic relation. When written in terms of their natural functional dependencies, the expressions for the thermodynamic potentials are known as *fundamental equations of state*.

It is useful to have access to a suite of thermodynamic potentials (internal energy, entropy, enthalpy, Gibbs potential, Helmboltz free energy) that have different functional dependencies, which in turn yield distinct expressions for the fundamental equation of state. Thermodynamic potentials are related mathematically through a *Legendre transformation*. Motivation for their introduction comes from the distinct laboratory and environmental conditions whereby the controlling parameters may differ.

23.3.1 Internal energy

Recall the fundamental thermodynamic relation (23.21) written for a binary fluid

$$d\mathcal{J} = T dS - p d\nu + \mu dC. \quad (23.24)$$

Equation (23.24) identifies the specific internal energy, \mathcal{J} , as a function of specific entropy, S , specific volume, ν , and matter concentration, C

$$\mathcal{J} = \mathcal{J}(S, \nu, C). \quad (23.25)$$

This equation is the *fundamental equation of state* written in terms of internal energy. This equation of state is more general than the *thermal equation of state* used to express density as a function of temperature, pressure, and matter concentration (Section 25.2).

Knowledge of the fundamental thermodynamic relation and use of the fundamental equation of state allows one to derive the plethora of thermodynamic relations. For example, we can identify the partial derivatives

$$\left[\frac{\partial \mathcal{I}}{\partial S} \right]_{\nu, C} = T \quad (23.26)$$

$$\left[\frac{\partial \mathcal{I}}{\partial \nu} \right]_{S, C} = -p \quad (23.27)$$

$$\left[\frac{\partial \mathcal{I}}{\partial C} \right]_{S, \nu} = \mu. \quad (23.28)$$

These equations are the intensive form of the extensive equations of state (23.7)-(23.9)

23.3.2 Entropy

Rearrangement of the fundamental thermodynamic relation (23.24) leads to the exact differential for specific entropy

$$dS = \frac{1}{T} d\mathcal{I} + \frac{p}{T} d\nu - \frac{\mu}{T} dC. \quad (23.29)$$

In this form, specific entropy has the functional dependence

$$S = S(\mathcal{I}, \nu, C), \quad (23.30)$$

whose knowledge provides yet another form of the fundamental equation of state. This functional dependence, along with equation (23.29), lead to the following thermodynamic equations of state

$$\left[\frac{\partial S}{\partial \mathcal{I}} \right]_{\nu, C} = \frac{1}{T} \quad (23.31)$$

$$\left[\frac{\partial S}{\partial \nu} \right]_{\mathcal{I}, C} = \frac{p}{T} \quad (23.32)$$

$$\left[\frac{\partial S}{\partial C} \right]_{\mathcal{I}, \nu} = -\frac{\mu}{T}. \quad (23.33)$$

23.3.3 Enthalpy

Thus far we have worked only with the fundamental thermodynamic relation (23.24). We now introduce the specific enthalpy

$$\mathcal{H} = \mathcal{I} + p\nu. \quad (23.34)$$

Use of equation (23.24) leads to the exact differential for enthalpy

$$d\mathcal{H} = d\mathcal{I} + d(p\nu) \quad (23.35a)$$

$$= T dS - p d\nu + \mu dC + p d\nu + \nu dp \quad (23.35b)$$

$$= T dS + \nu dp + \mu dC. \quad (23.35c)$$

This equation provides the fundamental thermodynamic relation with enthalpy rather than internal energy. Consequently, the *Legendre transformation* (23.34) renders a functional dependence for enthalpy

$$\mathcal{H} = \mathcal{H}(\mathcal{S}, p, C), \quad (23.36)$$

which is yet another fundamental equation of state. This functional dependence is more convenient than that for internal energy, $\mathcal{I}(\mathcal{S}, \nu, C)$, or for entropy $\mathcal{S}(\mathcal{I}, \nu, C)$. The reason is that we have direct mechanical means of measuring pressure in a fluid, whereas specific volume requires indirect methods involving the equation of state for density discussed in Section 25.2. Additionally, specific entropy remains constant on a fluid element in the absence of mixing or other irreversible effects. Correspondingly, enthalpy remains constant for constant pressure motion without mixing. Conversely, in the presence of mixing at constant pressure, fluid elements mix their specific enthalpy, specific entropy, and tracer concentration. Finally, as discussed in Section 23.7, potential temperature is related to specific entropy, making the functional dependence equivalently written as $\mathcal{H}(\theta, p, C)$.

23.3.4 Gibbs potential

The Gibbs potential is defined by the Legendre transformation

$$\mathcal{G} = \mathcal{I} + p\nu - T\mathcal{S} = \mathcal{H} - T\mathcal{S}. \quad (23.37)$$

Its exact differential is given by

$$d\mathcal{G} = d\mathcal{H} - d(T\mathcal{S}) \quad (23.38a)$$

$$= Td\mathcal{S} + \nu dp + \mu dC - Td\mathcal{S} - \mathcal{S}dT \quad (23.38b)$$

$$= -\mathcal{S}dT + \nu dp + \mu dC, \quad (23.38c)$$

where we made use of the fundamental thermodynamic relation (23.35c) written in terms of enthalpy. The Gibbs potential has the functional dependence

$$\mathcal{G} = \mathcal{G}(C, T, p). \quad (23.39)$$

This form of the fundamental equation of state is used quite often in fluid mechanics and physical chemistry. The reason is that temperature, pressure, and concentration are readily measured in the laboratory or the environment, thus making the partial derivatives of \mathcal{G} readily measured.

23.3.5 Combined fundamental thermodynamic relation

Bringing together the various forms for the fundamental thermodynamic relations renders

$$d\mathcal{I} + \nu dp = Td\mathcal{S} + \mu dC = d\mathcal{H} - \nu dp. \quad (23.40)$$

This relation is useful for fluid mechanics applications.

23.3.6 Alternate functional dependencies

The natural functional dependence for entropy is given by its fundamental equation of state (23.30)

$$\mathcal{S} = \mathcal{S}(\mathcal{I}, \nu, C). \quad (23.41)$$

Likewise, the internal energy has a corresponding fundamental equation of state (23.25)

$$\mathcal{I} = \mathcal{I}(\mathcal{S}, \nu, C). \quad (23.42)$$

However, there are other “un-natural” functional expressions for these, and other, thermodynamic potentials. For example, the entropy can be expressed as a function of temperature, pressure, and concentration

$$\mathcal{S} = \mathcal{S}(C, T, p). \quad (23.43)$$

We do not need to include specific volume, since it is determined through the thermal equation of state discussed in Section 25.2, which yields a relation for the density as a function of temperature, pressure, and concentration (salinity in the ocean, humidity in the atmosphere)

$$\nu^{-1} = \rho = \rho(C, T, p). \quad (23.44)$$

We make use of the functional dependence (23.43) in Section 23.6 when discussing the lapse rate.

23.3.7 Further study

Much of this section follows Section 1.5 of [Vallis \(2017\)](#). The Gibbs potential plays a central role in establishing the thermodynamics of the ocean as formulated by [Feistel \(1993\)](#) and codified by [IOC et al. \(2010\)](#).

23.4 Specific heat capacity

The specific heat capacity is a thermodynamic *response function* that measures the change in heat associated with a change in temperature at constant matter composition. There are two distinct heat capacities generally considered: one with specific volume held fixed and the other with pressure held fixed

$$c_v \equiv \frac{1}{M} \left[\frac{d\mathcal{Q}}{dT} \right]_{\nu, C} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1}. \quad (23.45)$$

$$c_p \equiv \frac{1}{M} \left[\frac{d\mathcal{Q}}{dT} \right]_{p, C} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1}. \quad (23.46)$$

If heating occurs quasi-statically, we can make use of the equation (23.4) relating heat and entropy, applied here in its specific form $M^{-1} d\mathcal{Q} = T d\mathcal{S}$. The result is a state-function form of the specific heat capacities

$$c_v = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{\nu, C} \quad (23.47)$$

$$c_p = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{p, C}. \quad (23.48)$$

We can furthermore make use of the fundamental thermodynamic relation (23.19) with specific volume and matter concentration held fixed to write

$$c_v = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{\nu} = \left[\frac{\partial \mathcal{I}}{\partial T} \right]_{\nu, C}. \quad (23.49)$$

Likewise, making use of the fundamental thermodynamic relation (23.35c) written in terms of enthalpy leads to the constant pressure heat capacity

$$c_p = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{p, C} = \left[\frac{\partial \mathcal{H}}{\partial T} \right]_{p, C}. \quad (23.50)$$

23.5 A simple ideal gas atmosphere

A dry atmosphere is well approximated by an ideal gas, so it is useful to develop various thermodynamic relations for an ideal gas. In an ideal gas we ignore intermolecular forces between molecules. Also, the molecules in an ideal gas are assumed to occupy zero volume. So the internal energy of an ideal gas is just due to translation, rotation, and vibration of molecules. We refer to a *simple ideal gas*, in which the internal energy is a linear function of temperature. A simple ideal gas offers a remarkably accurate basis for studying the thermodynamics of a dry atmosphere.

23.5.1 Equation of state

An ideal gas satisfies the following equation (see Section 2.2.2)

$$PV = nRT, \quad (23.51)$$

where p is the pressure, V is the volume, n is the number of moles,

$$R = 8.314 \text{ J mole}^{-1} \text{ K}^{-1} = 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1} \quad (23.52)$$

is the *universal gas constant*, and T is the absolute temperature (see Section 2.2.2). The number of moles equals to the mass, M , of the gas divided by the mass per mole, M_{mole}

$$n = \frac{M}{M_{\text{mole}}}. \quad (23.53)$$

The mass density, $\rho = M/V$, is thus given by

$$\rho = \frac{p M_{\text{mole}}}{T R} \equiv \frac{p}{T R^M}, \quad (23.54)$$

where

$$R^M = \frac{R}{M_{\text{mole}}} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1} \quad (23.55)$$

is the *specific gas constant* as defined by the universal gas constant normalized by the molar mass for the constituent. The relation (23.54) is known as a *thermal equation of state*, or more succinctly just an equation of state (see Section 25.2 for more discussion). It shows that the mass density of an ideal gas is directly proportional to the pressure: increasing pressure in turn increases density. In contrast, mass density is inversely proportional to the temperature: increases in temperature lead to lower mass density. This behavior for the ideal gas density is reflected in certain real gases and liquids. However, a notable counter-example is water near its freezing point, which becomes more dense as temperature rises. This anomalous behavior is why a body of water freezes from the top down rather than from the bottom up.

23.5.2 Internal energy

An ideal gas is comprised of molecules that interact only through elastic collisions. There are no inter-molecular forces. Furthermore, the volume of the individual molecules is ignored in comparison to the volume of empty space between the molecules, so they are approximated as point masses. Consequently, the internal energy for an ideal gas is independent of density and of the matter concentration. It is hence a function only of the temperature (i.e., kinetic energy of the elastic point molecules)

$$\mathcal{I} = \mathcal{I}(T) \quad \text{ideal gas.} \quad (23.56)$$

Consequently, the exact differential of internal energy for an ideal gas is

$$d\mathcal{I} = c_v dT. \quad (23.57)$$

The appearance of c_v , the constant volume specific heat capacity discussed in Section 23.4, arises in order for the ideal gas internal energy to satisfy the general equation (23.49). The heat capacity for an ideal gas is generally a function of temperature. However, for many applications it is sufficient to consider a *simple ideal gas*, in which c_v is a constant so that

$$\mathcal{I} = c_v T + \text{constant} \quad \text{simple ideal gas.} \quad (23.58)$$

The arbitrary constant of integration is generally set to zero so that the internal energy vanishes at absolute zero.

23.5.3 Heat capacity from statistical mechanics

Recall again that for an ideal gas, the internal energy is just a function of temperature, $\mathcal{I} = c_v T$. A *simple ideal gas* is an ideal gas for which the heat capacity is a constant (equation 23.58). Results from statistical mechanics (outside our scope) show that the thermal/internal energy per molecule equals to $k_B T/2$ per excited molecular degree of freedom, where

$$k_B = 1.3806 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1} \quad (23.59)$$

is the Boltzmann constant. Dry air is mostly comprised of the diatomic molecules N_2 and O_2 . Diatomic molecules at temperatures of the lower atmosphere have two rotational and three translational degrees of freedom,¹ so that $\mathcal{I}_{\text{molecule}} = 5 k_B T/2$.

We convert this energy per molecule to an energy per mole of diatomic molecules by multiplying by Avogadro's number (equation (2.10))

$$\mathcal{I}_{\text{mole diatomic}} = \frac{5 A_v k_B T}{2} = \frac{5 R T}{2}, \quad (23.60)$$

where the gas constant is given by

$$R = A_v k_B \quad (23.61a)$$

$$= (6.022 \times 10^{23} \text{ mole}^{-1}) (1.3806 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}) \quad (23.61b)$$

$$= 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1}. \quad (23.61c)$$

Finally, dividing by the molar mass for dry air (equation (2.11))

$$M_{\text{air}} = 0.028 \text{ kg mole}^{-1} \quad (23.62)$$

leads to the *simple ideal gas* approximation to the dry air heat capacity

$$c_v = \frac{5 R}{2 M_{\text{air}}} = 742 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1}. \quad (23.63)$$

The measured heat capacity for dry air at standard temperature (300 K) is $718 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1}$, so the simple ideal gas estimate is only $(742 - 718)/718 = 3.3\%$ too large.

¹At high temperatures, two vibrational degrees of freedom are also excited so that $\mathcal{I}_{\text{molecule}} = 7 k_B T/2$.

23.5.4 Enthalpy

The enthalpy is generally given by equation (23.34)

$$\mathcal{H} = \mathcal{I} + p/\rho. \quad (23.64)$$

For a *simple ideal gas* this expression takes the form

$$\mathcal{H} = \mathcal{I} + p/\rho \quad (23.65a)$$

$$= c_v T + \frac{T R}{M_{\text{mole}}} \quad (23.65b)$$

$$= T [c_v + R^M] \quad (23.65c)$$

where $R^M = R/M_{\text{mole}}$ (equation (23.55)) is the universal gas constant divided by the molar mass for the gas. Recall that the constant pressure heat capacity is given by equation (23.50)

$$c_p = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{p,C} = \left[\frac{\partial \mathcal{H}}{\partial T} \right]_{p,C}. \quad (23.66)$$

Consequently, for a *simple ideal gas* we have

$$c_p = c_v + R^M \quad (23.67)$$

$$\mathcal{H} = c_p T. \quad (23.68)$$

23.5.5 Thermal expansion coefficient

The thermal expansion coefficient measures the relative changes in density as temperature changes at constant pressure and concentration

$$\beta_T = -\frac{1}{\rho} \left[\frac{\partial \rho}{\partial T} \right]_{p,C} = \frac{1}{\nu} \left[\frac{\partial \nu}{\partial T} \right]_{p,C}. \quad (23.69)$$

The thermal expansion coefficient for an ideal gas is given by

$$\beta_T = \frac{1}{T}, \quad (23.70)$$

so that as temperature rises the thermal expansion reduces.

23.5.6 Fundamental thermodynamic relations

The fundamental thermodynamic relation in terms of internal energy (equation (23.21)) and enthalpy (equation (23.34)) are given by

$$d\mathcal{I} = T d\mathcal{S} - p d\nu + \mu dC \quad (23.71)$$

$$d\mathcal{H} = T d\mathcal{S} + \nu dp + \mu dC. \quad (23.72)$$

For a simple ideal gas these relations take the form

$$c_v dT = T d\mathcal{S} - p d\nu + \mu dC \quad (23.73)$$

$$c_p dT = T d\mathcal{S} + \nu dp + \mu dC. \quad (23.74)$$

23.5.7 Further study

Sections 1.5 and 1.7 of [Vallis \(2017\)](#) offer more details and results for a simple ideal gas atmosphere.

23.6 Adiabatic lapse rate

The temperature of a fluid can change without the transfer of heat. This *adiabatic* temperature change arises when the fluid pressure changes. We here introduce the *adiabatic lapse rate*, which refers to the changes in temperature arising from a static fluid placed in a gravity field. There are two lapse rates commonly considered: one related to height changes and one related to pressure changes. We then introduce some manipulations commonly performed with thermodynamic state functions and their partial derivatives, with the goal of expressing the lapse rate in terms of commonly measured response functions.

23.6.1 Isentropic rearrangement

Consider a finite region of a static fluid in a gravitational field. Assume the fluid is initially in a thin horizontal layer with a uniform temperature. Now rearrange the fluid into a vertical column, and do so without changing the entropy; i.e., without the transfer of heat across the fluid boundary (adiabatically) and without mixing. Performing this rearrangement raises the center of mass of the fluid system and thus increases the gravitational potential energy. This process thus requires mechanical work against the gravitational field.

Gravity makes pressure at the bottom of the vertical fluid column greater than at the top. This pressure difference affects the temperature in the column. We seek a thermodynamic expression for how changes in pressure affects changes in temperature for a static fluid, with the pressure changes imparted reversibly and adiabatically so that entropy does not change. Mathematically, we seek an expression for the partial derivative

$$\hat{\Gamma} \equiv \left[\frac{\partial T}{\partial p} \right]_{C,S}, \quad (23.75)$$

which is known as the *adiabatic lapse rate*. The adiabatic lapse rate can be measured directly, with empirical expressions fit to laboratory measurements. Additionally, it is convenient to express it in terms of other thermodynamic response functions in order to garner further physical insight. The necessary manipulations form the bulk of this section.

23.6.2 Thermodynamic formulation

When the matter concentration is held fixed, we can write the exact differential of entropy as a function of temperature and pressure (see equation (23.43)). Hence, infinitesimal changes in entropy are given by

$$dS = \left[\frac{\partial S}{\partial T} \right]_p dT + \left[\frac{\partial S}{\partial p} \right]_T dp. \quad (23.76)$$

Substituting the definition of heat capacity (23.48) leads to

$$T dS = c_p dT + T \left[\frac{\partial S}{\partial p} \right]_T dp. \quad (23.77)$$

It is useful to eliminate $(\partial S/\partial p)_T$ in favor of a more easily measurable quantity. For this purpose, note that use of the fundamental thermodynamic relation (23.21) (with $dC = 0$) leads to

$$T \left[\frac{\partial S}{\partial T} \right]_p = \left[\frac{\partial J}{\partial T} \right]_p + p \left[\frac{\partial \nu}{\partial T} \right]_p \quad (23.78)$$

as well as

$$T \left[\frac{\partial S}{\partial p} \right]_T = \left[\frac{\partial J}{\partial p} \right]_T + p \left[\frac{\partial \nu}{\partial p} \right]_T. \quad (23.79)$$

Applying $(\partial/\partial p)_T$ to equation (23.78) and $(\partial/\partial T)_p$ to equation (23.79), and then subtracting, leads to the identity

$$\left[\frac{\partial S}{\partial p} \right]_T = - \left[\frac{\partial \nu}{\partial T} \right]_p. \quad (23.80)$$

Introducing the thermal expansion coefficient (23.69) yields an expression for changes in entropy in terms of changes in temperature and pressure

$$T dS = c_p dT - T \left[\frac{\partial \rho^{-1}}{\partial T} \right]_p dp \quad (23.81a)$$

$$= c_p dT - \left[\frac{T \beta_T}{\rho} \right] dp. \quad (23.81b)$$

c_p and β_T are readily measurable *response functions*, thus making equation (23.81b) a useful expression for infinitesimal entropy changes when matter concentration is held constant.

23.6.3 Adiabatic lapse rate for pressure changes

Equation (23.81b) means that the change in temperature associated with motion through a pressure field, with $dS = 0$ and $dC = 0$ can be written

$$\hat{\Gamma} = \left[\frac{\partial T}{\partial p} \right]_{C,S} = \frac{T \beta_T}{\rho c_p}. \quad (23.82)$$

Temperature indeed changes when pressure changes, even though there has been no heat exchanged with the environment.

23.6.4 Adiabatic lapse rate for height changes

A static fluid in a gravity field is in hydrostatic balance, whereby the pressure at a point equals to the weight per area above that point (Section 27.3). Hydrostatic balance in a constant gravity field maintains the following relation between the pressure differential increment and the vertical differential increment

$$dp = -g \rho dz. \quad (23.83)$$

Use of the chain rule within the lapse rate expression (23.82) leads to

$$\Gamma = \left[\frac{\partial T}{\partial z} \right]_{C,S} \quad (23.84a)$$

$$= \left[\frac{\partial T}{\partial p} \right]_{C,S} \left[\frac{\partial p}{\partial z} \right] \quad (23.84b)$$

$$= -\rho g \left[\frac{T \beta_T}{\rho c_p} \right] \quad (23.84c)$$

$$= - \left[\frac{g T \beta_T}{c_p} \right]. \quad (23.84d)$$

This form for the lapse rates measures the change in temperature (the *lapse*) within a fluid element as it is isentropically moved vertically through a hydrostatic pressure field.

23.6.5 Adiabatic lapse rate for an ideal gas atmosphere

For an ideal gas (Section 23.5), the thermal expansion coefficient is given by (equation (23.70))

$$\beta_T = \frac{1}{T} \quad (23.85)$$

so that the lapse rates are given by

$$\hat{\Gamma} = \frac{1}{\rho c_p} \quad (23.86)$$

$$\Gamma = -\frac{g}{c_p}. \quad (23.87)$$

The measured specific heat capacity for a dry atmosphere at standard temperature (300 K) is

$$c_p = 1005 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1} \quad (23.88)$$

so that the adiabatic lapse rate for a dry atmosphere is roughly

$$\Gamma_d = -9.8 \text{ K/(1000 m)}. \quad (23.89)$$

Hence, temperature decreases by nearly 10 K when rising 1000 m in a dry and ideal gas atmosphere.

23.6.6 Further study

[McDougall and Feistel \(2003\)](#) provide a discussion of the lapse rate in terms of molecular dynamics. In particular, they note that the lapse rate, being proportional to the thermal expansion coefficient, can be negative when the thermal expansion is negative. A negative thermal expansion coefficient occurs in cool fresh water, such as the Baltic Sea. Hence, although work is done on the fluid element under increasing pressure, its temperature decreases in these cases.

The addition of water to the atmosphere modifies the lapse rate, as the air is then no longer well approximated by an ideal gas. Chapter 18 of [Vallis \(2017\)](#) offers a pedagogical discussion of the thermodynamics of a moist tropical atmosphere.

23.7 Potential temperature

Heating and cooling of the ocean predominantly occur near the ocean surface. In contrast, transport in the ocean interior is nearly adiabatic and isohaline (i.e., nearly isentropic). The physical picture is suggested whereby the surface ocean boundary layer experiences irreversible processes that set characteristics of water masses that move quasi-reversibly within the ocean interior. Oceanographers thus prefer to mark or label water masses using properties that maintain their values when moving within the quasi-isentropic ocean interior. Salinity is a good label for this purpose since it is only altered by mixing between waters of varying concentrations, and in turn it is materially constant in the absence of mixing. This behavior constitutes a basic property of material tracers (tracers that measure the mass per mass of a constituent as discussed in Section 18.1). However, it is *not* a property of the *in situ* temperature, T , which changes even in the absence of mixing due to pressure effects (see the adiabatic lapse rate discussion in Section 23.6).

We are thus motivated to seek a thermodynamic tracer that evolves analogously to material tracers, so that it can be used as a second material label for fluid elements. A similar motivation stems from the analysis of atmospheric motions. These considerations lead to the introduction of potential temperature.

23.7.1 Adiabatic temperature changes

Vertical motion of a fluid element, without exchange of heat (adiabatic) or matter (constant concentration) changes the pressure of the fluid element. In turn, this motion causes the *in situ* temperature to change in proportion to the adiabatic lapse rate given by (Section 23.6)

$$dT = \hat{\Gamma} dp. \quad (23.90)$$

Consequently, *in situ* temperature is not a useful thermodynamic variable to label fluid elements, since it changes even in the absence of irreversible mixing processes. Instead, it is more useful to remove the adiabatic pressure effects. This is the reason to introduce potential temperature.

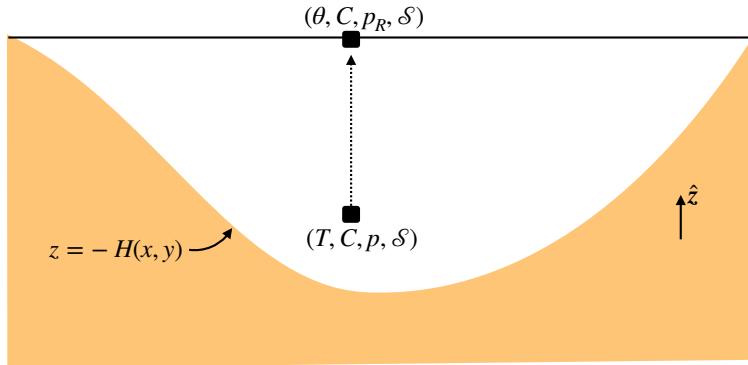


Figure 23.1: Potential temperature is the *in situ* temperature that a fluid element of fixed material composition would have if isentropically displaced from its *in situ* pressure to a reference pressure p_R . The schematic here depicts that displacement for a seawater fluid element with *in situ* temperature T , salinity $S = 1000 C$, pressure p , and specific entropy S . The element is moved to the ocean surface with the standard sea level atmospheric pressure providing the reference pressure.

23.7.2 Defining the potential temperature

Removing adiabatic pressure effects from *in situ* temperature leads to the concept of *potential temperature*. Potential temperature is defined as the *in situ* temperature that a fluid element of fixed material composition would have if it were isentropically transported from its *in situ* pressure to a reference pressure p_R , with the reference pressure typically taken at the ocean/land surface (see Figure 23.1).

Mathematically, the potential temperature, θ , is the reference temperature obtained via integration of $dT = \hat{\Gamma} dp$ for an isentropic *in situ* temperature change with respect to pressure

$$\int_{\theta}^T dT' = \int_{p_R}^p \hat{\Gamma}(C, T, p') dp' \implies T = \theta(C, T, p_R) + \int_{p_R}^p \hat{\Gamma}(C, T, p') dp', \quad (23.91)$$

with $\hat{\Gamma}$ the lapse rate defined in terms of pressure changes (equation (23.82)). By definition, the *in situ* temperature T equals the potential temperature θ at the reference pressure $p = p_R$. Elsewhere, they differ by an amount determined by the adiabatic lapse rate. Furthermore, we see that

$$\left[\frac{\partial T}{\partial p} \right]_{C,S} = \left[\frac{\partial \theta}{\partial p} \right]_{C,S} + \hat{\Gamma}. \quad (23.92)$$

However, by definition

$$\left[\frac{\partial T}{\partial p} \right]_{C,S} = \hat{\Gamma}, \quad (23.93)$$

so that

$$\left[\frac{\partial \theta}{\partial p} \right]_{C,S} = 0. \quad (23.94)$$

That is, by definition, the potential temperature has a parametric dependence on the reference pressure (it depends explicitly on the concentration, C and *in situ* temperature, T) yet it has no explicit dependence on the *in situ* pressure when holding concentration and entropy fixed. Finally, note that potential temperature is a function of concentration, C . Hence, the potential temperature generally changes if the material concentration changes. For example, potential temperature in the ocean changes if the salinity changes.

23.7.3 Potential temperature and specific entropy

An alternative definition of the potential temperature follows by noting that the entropy of a fluid element remains unchanged as it is reversibly moved to the reference pressure. Consequently, writing entropy as a function of temperature, pressure, and matter concentration (equation (23.43))

$$\mathcal{S} = \mathcal{S}(C, T, p) \quad (23.95)$$

leads to the defining identity for potential temperature

$$\mathcal{S}(C, T, p) = \mathcal{S}(C, \theta, p_R). \quad (23.96)$$

This relation directly connects changes in entropy to changes in potential temperature

$$d\mathcal{S} = \left[\frac{\partial \mathcal{S}(C, \theta, p_R)}{\partial \theta} \right]_C d\theta. \quad (23.97)$$

Consequently, the reversible transport of a fluid element with constant matter concentration ($dC = 0$) occurs with both a constant entropy and constant potential temperature.

We can go even further than the relation (23.97) by recalling that equation (23.81b) relates increments in specific entropy to temperature and pressure

$$T dS = c_p (dT - \hat{\Gamma} dp), \quad (23.98)$$

where $\hat{\Gamma}$ is the adiabatic lapse rate defined in terms of pressure changes (equation (23.82)), and we set $dC = 0$. Taking the infinitesimal increment (i.e., the differential) of the potential temperature (23.91) leads to

$$dT = d\theta + \hat{\Gamma}(C, T, p) dp + \int_{p_R}^p [d\hat{\Gamma}(C, T, p')] dp'. \quad (23.99)$$

Evaluate this increment at the reference pressure, $p = p_R$, so that the integral vanishes, thus leaving

$$d\theta = dT - \hat{\Gamma}(C, T, p_R) dp. \quad (23.100)$$

We make use of this relation in equation (23.98) to render an expression for the entropy increment in terms of the potential temperature increment

$$dS = c_p \theta^{-1} d\theta \quad p = p_R \text{ and } dC = 0. \quad (23.101)$$

As part of exercise 23.6, we show that this relation holds for an ideal gas at all pressures, and as part of exercise 23.7, we see that the relation also holds for all pressures in certain liquids.

23.7.4 Potential temperature for an ideal gas

The fundamental thermodynamic relation for a *simple ideal gas* (23.74) takes on the following form for an isentropic change

$$c_p dT = \nu dp. \quad (23.102)$$

Dividing both sides by temperature and using the ideal gas relation

$$\frac{\nu}{T} = \frac{R^M}{p} \quad (23.103)$$

leads to

$$c_p d \ln T = R^M d \ln p. \quad (23.104)$$

Since c_p and R^M are constants, we can integrate this relation from the reference pressure to an arbitrary pressure

$$c_p \int_{\theta}^T d \ln T = R^M \int_{p_R}^p d \ln p, \quad (23.105)$$

which renders the explicit expression for the simple ideal gas potential temperature

$$\theta = T \left[\frac{p_R}{p} \right]^{R^M/c_p} \quad (23.106)$$

where

$$c_p = \frac{7 R^M}{2} \quad (23.107)$$

is the constant pressure heat capacity of a simple ideal gas of diatomic molecules (Section 23.5.3). In Exercise 23.5 we show that $\partial\theta/\partial p = 0$ for the ideal gas, thus exemplifying the removal of explicit pressure effects from the potential temperature.

23.8 Thermodynamics of a moving fluid

Recall the fundamental thermodynamic relation (23.21) for a binary fluid such as seawater

$$d\mathcal{J} = T dS - p d\nu + \mu dC. \quad (23.108)$$

This relation is an expression of the First Law of thermodynamics for a quasi-static process. It provides an expression for the exact differential of internal energy for a thermodynamic system infinitesimally close to equilibrium.

Now assume that the thermodynamic system of interest is a finite region of fluid comprised of infinitesimal fluid elements. The finite fluid may be out of equilibrium, in that it experiences mechanical and thermal forces that support macroscopic motion. However, we assume that each fluid element is in local thermodynamic equilibrium. This assumption is supported by noting that the equilibrium time scale for an individual fluid element is tiny (i.e., fluid elements equilibrate very fast) compared to the equilibrium time scales of the macroscopic motion as well as the time scales for changes in the macroscopic forces. We are therefore justified in making use of quasi-equilibrium (also called linear irreversible) thermodynamics.² In this approach, we make use of equilibrium thermodynamics locally, yet we allow for macroscopic gradients in fluid properties. That is, the fluid is locally in thermodynamic equilibrium while it is macroscopically out of equilibrium. Furthermore, we assume that advective transport is reversible so that it does not modify the fluid entropy.

23.8.1 First Law for a moving fluid element

For a fluid, the key operational feature of quasi-equilibrium thermodynamics is that we extend equilibrium thermodynamic relations to moving and evolving fluid elements. Consequently, the equilibrium relation (23.108), which is the First Law for a quasi-static process, takes the material form

$$\frac{D\mathcal{J}}{Dt} = T \frac{DS}{Dt} - p \frac{D\nu}{Dt} + \mu \frac{DC}{Dt}. \quad (23.109)$$

This equation is the First Law of thermodynamics for a moving fluid element.

Concerning the transition to a continuum

The expression (23.109) for the First Law in a continuum fluid was not derived rigorously, and as such it can appear somewhat mysterious on first encounter. We thus offer further discussion to further expose why it is a rather obvious result of moving to the continuum while assuming local thermodynamic equilibrium.

For a continuum fluid, each of the thermodynamic properties in the equilibrium First Law expression (23.108) are continuous functions of space and time. Furthermore, equation (23.108) provides a relation between exact differentials as detailed in Section 4.8. As exact differentials of continuous fields, we can make use of the space-time increments detailed in Section 16.4.1 to write

$$d\Psi = \Psi(\mathbf{x} + d\mathbf{x}, t + dt) - \Psi(\mathbf{x}, t) \quad (23.110a)$$

$$= dt \partial_t \Psi + d\mathbf{x} \cdot \nabla \Psi, \quad (23.110b)$$

²The *linear* in the name linear irreversible thermodynamics refers to an assumption that the system is close to thermodynamic equilibrium. Dissipative thermodynamic fluxes are thus linear functions of the gradients of the thermodynamic state variables. Nonlinear effects are present from advective transport, nonlinear source terms, a nonlinear equation of state, and nonlinear dependence of the transport coefficients.

where Ψ is one of the thermodynamic properties, dt is the infinitesimal time increment, and $d\mathbf{x}$ is the vector of infinitesimal space increments. Following the discussion in Section 16.4.2, we are thus led to the total time derivative for a property following a trajectory $\mathbf{x} = \mathbf{x}(t)$

$$\frac{d\Psi}{dt} = \frac{\partial\Psi}{\partial t} + \frac{d\mathbf{x}}{dt} \cdot \nabla\Psi. \quad (23.111)$$

Restricting the trajectory to that defined by a fluid particle, so that $\mathbf{v} = d\mathbf{x}/dt$, then renders the material time derivative as in Section 16.4.4

$$\frac{D\Psi}{Dt} = \frac{\partial\Psi}{\partial t} + \mathbf{v} \cdot \nabla\Psi. \quad (23.112)$$

We are thus led to conclude that the expression (23.109) for the First Law in a continuum fluid is an inevitable result of transitioning the equilibrium relation (23.108) to the continuum.

Massaging the First Law

We now massage the result (23.109) to further reveal its connection to the First Law written in the form (23.6). For this purpose, recall that mass conservation as discussed in Section 17.1 means that changes in the volume of a fluid element are related to density changes via

$$\frac{1}{\delta V} \frac{D\delta V}{Dt} = \frac{1}{\nu} \frac{D\nu}{Dt}, \quad (23.113)$$

where again $\nu = \rho^{-1}$ is the specific volume. Hence, equation (23.109) can be written

$$\delta M \frac{D\mathcal{J}}{Dt} = T \delta M \frac{DS}{Dt} - p \frac{D\delta V}{Dt} + \mu \delta M \frac{DC}{Dt}, \quad (23.114)$$

where

$$\delta M = \rho \delta V \quad (23.115)$$

is the mass of the fluid element. Since the mass of the fluid element is constant, equation (23.114) is the fluid element extension of the First Law given by equation (23.6). Alternatively, we can use the further result from mass conservation

$$\frac{1}{\delta V} \frac{D\delta V}{Dt} = \nabla \cdot \mathbf{v} \quad (23.116)$$

to write

$$\rho \frac{D\mathcal{J}}{Dt} = T \rho \frac{DS}{Dt} - p \nabla \cdot \mathbf{v} + \mu \rho \frac{DC}{Dt}. \quad (23.117)$$

We now introduce a commonly used notation for the heating rate

$$T \frac{DS}{Dt} \equiv \dot{Q} \quad (23.118)$$

in which case the First Law for a moving fluid element (23.117) takes the form

$$\frac{D\mathcal{J}}{Dt} = -\nu p \nabla \cdot \mathbf{v} + \dot{Q} + \mu \frac{DC}{Dt}. \quad (23.119)$$

Comments on the steady state First Law

Following from the material expression (23.109) of the first law, we have the corresponding steady state expression

$$\mathbf{v} \cdot (\nabla J - T \nabla S + p \nabla \nu - \mu \nabla C) = 0. \quad (23.120)$$

Note that it is generally *incorrect* to set the bracketed term to zero. Rather, this identity only holds for the projection of the bracketed term onto the flow direction.

23.8.2 The First Law in terms of potential temperature

Recall the expression of the First Law for a moving fluid element given by equation (23.119)

$$\frac{D\mathcal{J}}{Dt} = -p \nabla \cdot \mathbf{v} + \dot{Q} + \mu \frac{DC}{Dt}, \quad (23.121)$$

where

$$T \frac{DS}{Dt} = \dot{Q} \quad (23.122)$$

is the heating rate. From equation (23.101) we see that the change in entropy for an element moving with constant matter concentration and at the reference pressure is given in terms of the potential temperature

$$\frac{DS}{Dt} = \frac{c_p}{\theta} \frac{D\theta}{Dt} \quad (23.123)$$

Since the potential temperature equals to the temperature when $p = p_R$, we have

$$c_p \frac{D\theta}{Dt} = \theta \frac{DS}{Dt} = \dot{Q} \quad \text{at } p = p_R \text{ and } dC = 0. \quad (23.124)$$

In Exercise 23.6 we consider these relations for an ideal gas.

23.8.3 Potential temperature and numerical models

The *in situ* temperature, T , does not make for a convenient prognostic variable for numerical models of either the ocean or atmosphere. Its awkward nature relates to the adiabatic changes in T experienced when pressure changes. Namely, since *in situ* temperature changes when pressure changes, even when there is no heat transport (see adiabatic lapse rate discussion in Section 23.6), the *in situ* temperature T is modified even when the flow is adiabatic. In contrast, modelers prefer to work with scalar fields that remain materially constant in the absence of mixing. Hence, potential temperature has historically been used for both atmosphere and ocean models rather than *in situ* temperature.³

In numerical and theoretical models, potential temperature satisfies an advection/diffusion equation analogous to material tracers such as salinity (see Chapter 37). We will make use of potential temperature as the primary field measuring the buoyancy of a fluid in various theoretical models (Chapter 25).

³Further advances ([McDougall, 2003](#); [IOC et al., 2010](#)) show that Conservative Temperature is an even more suitable thermodynamic tracer for the ocean than potential temperature.

23.8.4 Further study

DeGroot and Mazur (1984) provide a full accounting of quasi-equilibrium thermodynamics as applied to continuum matter such as a fluid. *Gregg* (1984) and *Davis* (1994) apply these methods to small-scale mixing in the ocean. Slightly different formulations can be found in *Landau and Lifshitz* (1987) and *Batchelor* (1967). The presentation here is an abbreviation of that given in Section 1.6 of *Vallis* (2017).

23.9 Exercises

EXERCISE 23.1: DERIVATION OF THE GIBBS-DUHEM RELATION

Starting from the scaling (23.11) for the internal energy, work through the steps leading to the Gibbs-Duhem relation (23.13).

EXERCISE 23.2: HELMHOLTZ FREE ENERGY

The Helmholtz free energy is defined by the Legendre transformation

$$\mathcal{F} = \mathcal{I} - T\mathcal{S}. \quad (23.125)$$

Show that the fundamental thermodynamic relation in terms of \mathcal{F} is given by

$$d\mathcal{F} = -\mathcal{S}dT - pd\nu + \mu dC. \quad (23.126)$$

EXERCISE 23.3: CONSTANT OF MOTION FOR ADIABATIC FLOW

Show that for a simple ideal gas, isentropic flow (i.e., both adiabatic and of constant matter concentration) maintains

$$p\nu^{c_p/c_v} = \text{constant}, \quad (23.127)$$

where $\nu = \rho^{-1}$ is the specific volume.

EXERCISE 23.4: GEOPOTENTIAL HEIGHT

The *geopotential height* is the height above the earth of a chosen pressure surface.

- Show that an ideal gas atmosphere in hydrostatic balance with a uniform lapse rate

$$\frac{\partial T}{\partial z} = -|\Gamma| = \text{constant} \quad (23.128)$$

has a geopotential height at a pressure p given by

$$z = \frac{T_0}{|\Gamma|} \left[1 - \left(\frac{p_0}{p} \right)^{-R^M |\Gamma| / g} \right], \quad (23.129)$$

where T_0 is the temperature at $z = 0$.

- For an isothermal atmosphere, obtain an expression for the geopotential height as a function of pressure, and show that this result is consistent with the expression (23.129) in the appropriate limit.

EXERCISE 23.5: POTENTIAL TEMPERATURE FOR AN IDEAL GAS

Show that $\partial\theta/\partial p = 0$ for the potential temperature of an ideal gas given by equation (23.106)

$$\theta = T \left[\frac{p_R}{p} \right]^{R^M / c_p}. \quad (23.130)$$

Hint: remember that $\partial T/\partial p \neq 0$.

EXERCISE 23.6: THERMODYNAMIC MANIPULATIONS FOR IDEAL GASES

This question develops some manipulations with the potential temperature.

- (a) Beginning with the expression (23.106) for potential temperature of an ideal gas, show that

$$d\theta = \frac{\theta}{T} \left[dT - \frac{\nu}{c_p} dp \right]. \quad (23.131)$$

- (b) Given the result (23.131), show that an ideal gas satisfies the following relation

$$T dS = \frac{c_p T}{\theta} d\theta. \quad (23.132)$$

Whereas the relation (23.101) holds for a general fluid only at the reference pressure, this exercise shows that it holds for an ideal gas at all pressures. As a result, a moving fluid of ideal gas satisfies the material time relation

$$T \frac{DS}{Dt} = \frac{c_p T}{\theta} \frac{D\theta}{Dt} \Rightarrow \frac{c_p T}{\theta} \frac{D\theta}{Dt} = \dot{Q}. \quad (23.133)$$

EXERCISE 23.7: THERMODYNAMIC MANIPULATIONS FOR A LIQUID

Consider seawater with specific entropy given by (see Section 1.7.2 of [Vallis \(2017\)](#))

$$S(S, T, p) = S_0 + c_{p0} \ln(T/T_o) [1 + \beta_s^*(S - S_o)] - \alpha_o p \left[\beta_T + \beta_T \gamma^* \frac{p}{2} + \beta_T^* (T - T_o) \right], \quad (23.134)$$

and corresponding specific heat capacity at constant pressure

$$c_p(S, T, p) = c_{p0} [1 + \beta_s^*(S - S_o)] - \alpha_o p \beta_T^* T. \quad (23.135)$$

In these equations, T is the *in situ* temperature, S is the salinity, and p is the *in situ* pressure. All other terms on the right hand side to these expressions are empirical constants. Verify that an infinitesimal change in entropy for a fluid element with constant composition is given by

$$\theta dS = c_p(S, \theta, p_R) d\theta, \quad (23.136)$$

where θ is the potential temperature and p_R is the corresponding reference pressure. Consequently, we can write for a moving fluid element

$$\dot{Q} = \frac{c_p T}{\theta} \frac{D\theta}{Dt}, \quad (23.137)$$

where we evaluate the heat capacity at $c_p(S, \theta, p_R)$. We see that certain liquids have an expression for heating that is analogous to that for an ideal gas, with the ideal gas case discussed in Exercise 23.6. Hint: Make use of the identity (23.96).

24

Energy dynamics

In this chapter, we develop the fluid mechanical equations for mechanical energy and total energy relevant to the ocean and atmosphere.

READER'S GUIDE TO THIS CHAPTER

This chapter builds from the momentum dynamics of Chapter 22 and the thermodynamics of Chapter 23.

24.1	Mechanical energy budget	352
24.1.1	Kinetic energy	352
24.1.2	Gravitational potential energy	353
24.1.3	Mechanical energy	354
24.2	Total energy budget	354
24.2.1	Mechanical energy plus internal energy	354
24.2.2	Bernoulli function and mechanical injection work	355
24.2.3	Bernoulli's theorem for a perfect fluid	356
24.2.4	Further study	357
24.3	Comments on gauge symmetry	357
24.4	Materially constant specific entropy for a perfect fluid	358
24.5	Moist static energy	359
24.5.1	Further study	359
24.6	Exercises	359

24.1 Mechanical energy budget

Momentum is but one dynamical property of a fluid. Mechanical energy is another and it is formed by adding the energy due to motion of fluid elements (kinetic energy) to the energy arising from the position of a fluid element within the gravitational field (gravitational potential energy). Energy is a scalar field, which often means that its budget is simpler to understand than that for the vector momentum field. In this section we develop the evolution equations for mechanical energy of a fluid on a rotating sphere, making use of the momentum equations (22.21)-(22.23) written in spherical coordinates with a geopotential $\Phi = g z$.

24.1.1 Kinetic energy

Multiplying the zonal momentum equation (22.21) by u , the meridional equation (22.22) by v , and the radial equation (22.23) by w reveals that both the metric acceleration and the Coriolis acceleration drop from the kinetic energy equation. This result is expected since neither the metric acceleration nor the Coriolis acceleration perform work on a fluid element and thus do not affect changes to the mechanical energy.

Following from our discussion of mechanical energy of a point particle in Section 15.5, we define the kinetic energy per mass as

$$\mathcal{K} = \frac{\mathbf{v} \cdot \mathbf{v}}{2}. \quad (24.1)$$

The material evolution of the kinetic energy per mass is given by

$$\rho \frac{D\mathcal{K}}{Dt} = -\mathbf{v} \cdot \nabla p - w g \rho + \rho \mathbf{v} \cdot \mathbf{F}, \quad (24.2)$$

where we made use of the momentum equation (22.14) and wrote $\mathbf{F} = \rho^{-1} \nabla \cdot \tau$ for the friction acceleration. Equation (24.2) says that the kinetic energy of a fluid element is affected by three physical processes.

- Kinetic energy increases in regions where the velocity projects down the pressure gradient,

$$\mathbf{v} \cdot \nabla p < 0 \implies \text{increase kinetic energy}, \quad (24.3)$$

thus resulting in an increase in fluid speed imparted by the pressure gradient force. Conversely, kinetic energy is reduced in regions where the flow is directed up the pressure gradient. It is notable that horizontal geostrophic flows (Section 29.3) with

$$\mathbf{v}_g = \frac{\hat{\mathbf{z}} \wedge \nabla p}{f\rho} \quad (24.4)$$

have the velocity oriented perpendicular to the horizontal pressure gradient ($\mathbf{v}_g \cdot \nabla p = 0$). Consequently, pressure has no impact on the horizontal kinetic energy of a geostrophic fluid.

- Kinetic energy of a fluid element decreases in regions where the vertical velocity is positive,

$$w > 0 \implies -w g \rho < 0 \Leftrightarrow \text{upward motion decreases } \mathcal{K}. \quad (24.5)$$

Hence, motion against the gravitational field (upward) reduces kinetic energy. As seen in Section 24.1.2, this decrease in kinetic energy due to vertical motion is exactly balanced by an increase in gravitational potential energy: raising fluid elements increases their gravitational potential energy.

- Kinetic energy is reduced in regions where the velocity has a negative projection onto the direction of the friction vector, $\rho \mathbf{v} \cdot \mathbf{F} < 0$. Rayleigh drag is a particularly simple form of frictional acceleration, in which

$$\mathbf{F}_{\text{Rayleigh}} = -\gamma \mathbf{v}, \quad (24.6)$$

with $\gamma > 0$ having dimensions of inverse time. In this case

$$\mathbf{v} \cdot \mathbf{F}_{\text{Rayleigh}} = -\gamma \mathbf{v} \cdot \mathbf{v} \leq 0, \quad (24.7)$$

so that Rayleigh drag dissipates kinetic energy at each point where velocity is nonzero. Furthermore, the e-folding time for dissipation is set by γ^{-1} . We discuss other forms of frictional dissipation, and its impacts on kinetic energy, in Section 26.6.3.

24.1.2 Gravitational potential energy

The gravitational potential energy per mass; i.e., the geopotential, for a fluid element is given by

$$\Phi = g z \quad (24.8)$$

so that its material evolution is

$$\rho \frac{D\Phi}{Dt} = w g \rho. \quad (24.9)$$

As already noted above in our discussion of kinetic energy, we see that just as for the point particle, there is an exchange of mechanical energy between the kinetic energy and potential energy conveyed through vertical motion in the gravitational field. Vertically upward motion ($w > 0$) increases potential energy whereas downward motion ($w < 0$) decreases it. Each of these potential energy changes are exactly balanced by kinetic energy changes.

24.1.3 Mechanical energy

Adding the kinetic and potential energy leads to the material form of the mechanical energy equation

$$\rho \frac{Dm}{Dt} = -\mathbf{v} \cdot \nabla p + \rho \mathbf{v} \cdot \mathbf{F}, \quad (24.10)$$

where

$$m = \mathcal{K} + \Phi \quad (24.11)$$

is the mechanical energy per mass for a fluid element. As already anticipated, there is an exact cancellation of the mechanical energy exchange due to vertical motion through the gravitational field. In the absence of pressure and friction, equation (24.10) reduces to the mechanical energy budget for a point particle detailed in Section 15.6. That limiting case is expected since pressure and friction arise from the continuum nature of a fluid and are absent from the dynamics of a point particle.

24.2 Total energy budget

Recall from Section 15.6 that a point particle, in the absence of friction, conserves its mechanical energy. In contrast, the mechanical energy for a fluid element is not materially constant. The reason is there is a conversion between mechanical energy and internal energy as pressure does work to alter the volume of fluid elements, as friction dissipates mechanical energy and converts it to heat, and as mixing of tracer matter acts to increase entropy. We explore these points here by combining the mechanical energy budget from Section 24.1 to the internal energy budget from Section 23.8, thus rendering the budget for total energy of a fluid element. To simplify the discussion we restrict attention to the case of a homogeneous fluid so that there is no contribution from entropy of mixing.

24.2.1 Mechanical energy plus internal energy

The internal energy per mass, \mathcal{I} , for a constant composition fluid element changes through pressure work and heating (equation (23.119))

$$\rho \frac{D\mathcal{I}}{Dt} = -p \nabla \cdot \mathbf{v} + \rho \dot{Q}, \quad (24.12)$$

where

$$\dot{Q} = T \frac{D\mathcal{S}}{Dt} \quad (24.13)$$

is the heating rate acting on the fluid element and \mathcal{S} is the specific entropy. We define the total energy per mass as the sum of the internal energy plus the mechanical energy,

$$\mathcal{E} = \mathcal{I} + m = \mathcal{I} + \mathcal{K} + \Phi. \quad (24.14)$$

Adding the internal energy equation (24.12) to the mechanical energy equation (24.10) leads to the material evolution equation for the total energy

$$\rho \frac{D\mathcal{E}}{Dt} = -\nabla \cdot (p \mathbf{v}) + \rho (\mathbf{v} \cdot \mathbf{F} + \dot{Q}). \quad (24.15)$$

We now discuss the variety of physical processes that affect changes to \mathcal{E} .

Equation (24.15) reveals that the material time change for the total energy of a fluid element is affected by the convergence of pressure times velocity. Hence, even when incorporating the internal

energy and in the absence of dissipation or heating, the fluid element's total energy is not materially constant. The energy source term $p \mathbf{v}$ is fundamental to energy within the continuum. Namely, there is pressure work required for the fluid element to mechanically exist within the continuum of adjacent fluid elements. We further describe this mechanical *injection work* in the context of the Bernoulli function in Section 24.2.2.

24.2.2 Bernoulli function and mechanical injection work

Converting the material budget (24.15) into its Eulerian form renders

$$\frac{\partial(\rho \mathcal{E})}{\partial t} + \nabla \cdot [\rho \mathbf{v} (\mathcal{E} + p/\rho)] = \rho (\mathbf{v} \cdot \mathbf{F} + \dot{Q}). \quad (24.16)$$

The left hand side indicates that total energy of a fluid element is locally modified by the advective transport of the quantity

$$\mathcal{E} + p/\rho = \mathcal{K} + \mathcal{P} + \mathcal{I} + p \nu \quad (24.17a)$$

$$= \mathcal{K} + \mathcal{P} + \mathcal{H} \quad (24.17b)$$

$$\equiv \mathcal{B} \quad (24.17c)$$

where

$$\mathcal{H} = \mathcal{I} + p/\rho = \mathcal{I} + p \nu \quad (24.18)$$

is the enthalpy per mass of the fluid element (Section 23.3.3), and we introduced the Bernoulli function, which is the sum of the enthalpy per mass plus the mechanical energy per mass

$$\mathcal{B} = \mathcal{H} + \mathcal{K} + \mathcal{P} = \mathcal{H} + \mathcal{M}. \quad (24.19)$$

Integration over a region with zero boundary transfer of $\mathbf{v} \mathcal{B}$, and with zero mechanical dissipation or heating, leads to the conservation of total energy for the region.

Energy equation in terms of the Bernoulli function

In terms of the Bernoulli function, the total energy equation (24.16) takes the form

$$\frac{\partial(\rho \mathcal{E})}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathcal{B}) = \rho (\mathbf{v} \cdot \mathbf{F} + \dot{Q}). \quad (24.20)$$

We compare this conservation law to that for trace matter in a fluid with a diffusive flux \mathbf{J} , which satisfies the conservation law (see Section 18.1.4)

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho \mathbf{v} C + \mathbf{J}) = 0. \quad (24.21)$$

Note how the advective flux equals to the tracer concentration times the velocity, $\mathbf{v} C$. In contrast, the total energy per volume, $\rho \mathcal{E}$, changes at a point through the advective convergence of $\rho \mathcal{B}$ onto that point along with source/sink terms due to friction and heating

$$\frac{\partial(\rho \mathcal{E})}{\partial t} = -\nabla \cdot (\rho \mathbf{v} \mathcal{B}) + \rho (\mathbf{v} \cdot \mathbf{F} + \dot{Q}) \quad (24.22)$$

Why is $\rho \mathcal{E}$ affected by the convergence of $\rho \mathbf{v} \mathcal{B}$ rather than the convergence of $\rho \mathbf{v} \mathcal{E}$, even in the perfect fluid case without friction and heating?

The mechanical injection of work

To answer the above question,¹ again note that the Bernoulli function is the sum of the total energy per mass of a fluid parcel, \mathcal{E} , plus the term $p/\rho = p\nu$. So what is $p\nu$? Imagine carving out a unit mass from within a continuous fluid with pressure p and specific volume ν , leaving behind a “hole”. The mechanical work required to carve out this hole is precisely equal to $p\nu$. Correspondingly, we interpret $p\nu$ as the mechanical work required to inject a unit mass of fluid with specific volume ν into a region with pressure p . We thus refer to $p\nu$ as the *injection work*.

Recall that the specific enthalpy is the sum of the specific internal energy plus the injection work (equation (24.18)). Hence, specific enthalpy provides a measure of the energy (internal energy plus mechanical energy) required for a fluid element to exist within a continuum fluid. Adding this energy to the mechanical energy yields the Bernoulli function.

24.2.3 Bernoulli's theorem for a perfect fluid

Consider a perfect fluid flow in steady state (vanishing Eulerian time derivatives). Steady state mass conservation means that

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}) = 0. \quad (24.23)$$

This relation, along with a steady state energy in equation (24.22) (absent friction and heating), means that the steady state velocity field is parallel to contours of constant Bernoulli function

$$\mathbf{v} \cdot \nabla B = 0. \quad (24.24)$$

We thus see that for the perfect fluid to be in a steady state, the Bernoulli function is constant along streamlines. Hence, as the fluid moves along a streamline, there is an exchange between the total energy per mass, \mathcal{E} , and the injection work, $p\nu$, such that their sum remains constant.

A constant Bernoulli function for steady flow is used frequently in classical fluid dynamics to interpret flow around objects, such as for flow around a wing. It leads to a realization of *Bernoulli's Principle*, whereby in regions of low pressure the energy per mass is relatively large, whereas the converse holds in regions of high pressure. The change in energy is largely due to a change in the kinetic energy, so that flow is fast in regions of low pressure (e.g., top of the wing) and slow in regions of high pressure (e.g., bottom of the wing). We revisit Bernoulli's Principle in Section 36.1.2 when studying inviscid and irrotational flow.

Traditional derivation of Bernoulli's theorem

For completeness we offer a second derivation of Bernoulli's theorem that follows a more traditional route and reveals some useful manipulations. For this purpose, convert the advective-form momentum equation (22.13) into its vector-invariant form by making use of the vector identity (see Section 4.3.4)²

$$\boldsymbol{\omega} \wedge \mathbf{v} = -\mathcal{K} + (\mathbf{v} \cdot \nabla) \mathbf{v}. \quad (24.25)$$

This identity allows us to eliminate velocity self-advection in favor of the vorticity and kinetic energy per mass

$$\frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega}_a \wedge \mathbf{v} = -\frac{1}{\rho} \nabla p - \nabla m, \quad (24.26)$$

¹This argument follows Section 13.5.4 of [Thorne and Blandford \(2017\)](#).

²We pursue the same manipulations in Section 48.4.1 when deriving the vorticity equation.

where

$$\boldsymbol{\omega}_a = \boldsymbol{\omega} + 2\boldsymbol{\Omega} \quad (24.27)$$

is the absolute vorticity (see Chapter 48) and we set the friction and heating to zero since we are assuming a perfect fluid. The Eulerian time evolution for the kinetic energy per mass is therefore given by

$$\frac{\partial \mathcal{K}}{\partial t} = -\frac{1}{\rho} \mathbf{v} \cdot \nabla p - \mathbf{v} \cdot \nabla m, \quad (24.28)$$

where we set $\mathbf{v} \cdot (\boldsymbol{\omega}_a \wedge \mathbf{v}) = 0$. Next, we note that the specific entropy is materially constant for a fluid parcel in a perfect fluid, in which case material changes in the specific enthalpy are related to changes in pressure via (see Section 24.4)

$$\frac{D\mathcal{H}}{Dt} = \frac{1}{\rho} \frac{Dp}{Dt}. \quad (24.29)$$

Hence, a steady state perfect fluid maintains the balance

$$\rho(\mathbf{v} \cdot \nabla) \mathcal{H} = (\mathbf{v} \cdot \nabla) p, \quad (24.30)$$

so that the time tendency for the specific kinetic energy is given by

$$\frac{\partial \mathcal{K}}{\partial t} = -\mathbf{v} \cdot \nabla \mathcal{B}. \quad (24.31)$$

We thus see that for a steady state perfect fluid, the Bernoulli function is materially constant since

$$\frac{D\mathcal{B}}{Dt} = \mathbf{v} \cdot \nabla \mathcal{B} = 0 \quad \text{steady state perfect fluid.} \quad (24.32)$$

24.2.4 Further study

In our approach to the conservation of total energy, we postulated the existence of a conserved total energy and then examined the implications. This approach follows that of [Landau and Lifshitz \(1987\)](#), [DeGroot and Mazur \(1984\)](#) (see their Section II.4), Appendix A.13 of [IOC et al. \(2010\)](#), Appendix B in Chapter 1 of [Vallis \(2017\)](#), and Section 13.5.5 of [Thorne and Blandford \(2017\)](#). Although we exposed friction and heating in this section, our physical considerations mostly focused on the perfect fluid case. We revisit frictional impacts to energy in Section 26.6.3 after studying further properties of the frictional stress tensor. That discussion shows that friction provides a global sink of kinetic energy that is converted to heat through the process of Joule heating.

For an examination of Bernoulli's theorem for non-rotating flows, such next to laminar boundary layers, see [this video](#) produced by the National Committee for Fluid Mechanics Films featuring Prof. Ascher Shapiro.

24.3 Comments on gauge symmetry

Consider again the total energy equation as written in the Eulerian form (24.20) with the Bernoulli function

$$\frac{\partial(\rho \mathcal{E})}{\partial t} = -\nabla \cdot (\rho \mathbf{v} \mathcal{B}) + \rho(\mathbf{v} \cdot \mathbf{F} + \dot{Q}). \quad (24.33)$$

We identify the advective flux of energy according to

$$\text{advective flux of energy} = \rho \mathbf{v} \mathcal{B} \quad \text{dimensions } M T^{-3}. \quad (24.34)$$

However, it is notable that the time tendency for the total energy at a point remains unchanged if we shift the energy flux by a total curl

$$\text{modified advective flux of energy} = \rho \mathbf{v} \cdot \mathcal{B} + \nabla \wedge \mathbf{G}, \quad (24.35)$$

where \mathbf{G} is an arbitrary *gauge function*. This arbitrariness in the definition of energy flux is ubiquitous in physics; e.g., see the discussion of the electromagnetic field energy flux in Section 27-4 in Volume II of the [Feynman Lectures](#). It means that the energy flux itself has no unique local physical meaning; only its convergence has a meaning as it gives rise to time changes in the field energy at a point in space.

We also encounter such *gauge symmetry* in the potential vorticity flux discussed in Sections 50.4 and 52.1.3, as well as the vector streamfunction for an incompressible fluid in Section 19.5.2. In some cases we can exploit the symmetry to our advantage, as exemplified by the discussion of potential vorticity in Section 52.1.5. In other cases we simply choose to set the gauge function to zero.

These ambiguities with fluxes of derived quantities such as energy and potential vorticity are contrasted by fluxes of the material tracers described in Chapter 18. The flux of a material tracer has a physical meaning as the mass of matter transported across a unit area by advective plus non-advective fluxes. Even so, there remains an ambiguity in tracer fluxes since the convergence of the advective flux of a tracer is identical to the skew diffusive flux, with the two fluxes differing by a curl (Section 38.3). Measurements often sample the skew flux rather than the advective flux, thus furthering the need to acknowledge the “do-nothing” flux arising from a curl.

24.4 Materially constant specific entropy for a perfect fluid

In our discussion of thermodynamics in Chapter 23, we asserted that a fluid element maintains a constant specific entropy if it experiences no dissipation (no friction), maintains a constant composition (no mixing), and encounters no heat sources. Entropy for the perfect fluid is reversibly stirred through advection as it remains materially unchanged. It is useful here to verify that the equations developed thus far manifest this physical property.

For that purpose, make use of the fundamental thermodynamic relation in the form of equation (23.35c)

$$d\mathcal{H} = T dS + \nu dp, \quad (24.36)$$

where we assume a uniform matter composition so to set $dC = 0$. Applying this relation to a material fluid element leads to

$$T\rho \frac{D\mathcal{S}}{Dt} = \rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt}. \quad (24.37)$$

The Legendre transformation for specific enthalpy (equation (23.34)) leads to

$$\rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt} = \rho \frac{DJ}{Dt} - \frac{p}{\rho} \frac{D\rho}{Dt}. \quad (24.38)$$

Use of the continuity equation (17.8c) renders

$$\rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt} = \rho \frac{DJ}{Dt} + p \nabla \cdot \mathbf{v}, \quad (24.39)$$

and further use of the First Law in the form of equation (24.12) yields

$$\rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt} = 0. \quad (24.40)$$

Making use of this result in equation (24.37) implies that specific entropy is indeed materially constant

$$\frac{D\mathcal{S}}{Dt} = 0. \quad (24.41)$$

The key step in deriving this result sits with equation (24.12) for the First Law, whereby we removed contributions from heating and matter composition changes, again since we are making the perfect fluid assumption. Although a rather circular argument, it is important to verify that the circle indeed closes.

24.5 Moist static energy

Follow the discussion of Section 4.8 in [Gill \(1982\)](#). Then derive the lapse rate as per David Romps' discussion at the convection workshop 8-10 Feb 2018 in Princeton.

Let us consider the enthalpy equation (23.74) for the special case in which entropy changes occur through heating so that

$$T d\mathcal{S} = dQ. \quad (24.42)$$

Now assume heating occurs through radiation plus another term whose

$$T d\mathcal{S} = dQ \quad (24.43a)$$

$$= L dT + dQ_{rad}, \quad (24.43b)$$

where L is the latent heat of vaporization.

24.5.1 Further study

This section is incomplete. The intent is to merge material from Section 1.10 of [Vallis \(2017\)](#) to lecture notes from David Romps.

24.6 Exercises

EXERCISE 24.1: CROCCO'S THEOREM

Show that the spatial gradient of the Bernoulli function for a single-component steady perfect fluid can be written

$$\nabla \mathcal{B} = T \nabla \mathcal{S} + \mathbf{v} \wedge \boldsymbol{\omega}_a. \quad (24.44)$$

This result is known as Crocco's Theorem.

Hint: study the discussion in Section 24.2.3 where we showed that the Bernoulli function is constant along a steady flow streamline in a perfect fluid. Also recall from Section 23.3.3 that each differential in the fundamental thermodynamic relation $d\mathcal{H} = T d\mathcal{S} + \nu dp$ is an exact differential. Consequently, when considering an infinitesimal increment in space for a continuum fluid in steady state, then the fundamental thermodynamic relation implies

$$d\mathcal{H} = T d\mathcal{S} + \nu dp \implies \nabla \mathcal{H} = T \nabla \mathcal{S} + \nu \nabla p. \quad (24.45)$$

25

Buoyancy

Buoyancy is the acceleration felt by a massive body within a fluid due to the gravitational acceleration of the body relative to the gravitational acceleration of the fluid displaced by the body. As shown in Section 25.1, this statement of *Archimedes' Principle* is a basic consequence of hydrostatics. A body's buoyancy is measured by its density relative to that of the displaced fluid. Although the concept of buoyancy has wide applications for marine engineering (e.g., design of ships), we are interested in the buoyancy of fluid itself. That is, the “body” of interest for geophysical fluid mechanics is a test fluid element, one that does not disturb the environment into which it is placed.

We are thus interested in the buoyancy of a fluid element relative to the fluid that surrounds the element. A fluid element that is less dense than its local environment rises, whereas it sinks if less dense. We are furthermore interested in how density is measured and how it changes, with Section 25.2 just the beginning of our investigations into how density is modified by physical processes. In Section 25.3 we examine the resistance to motion within a gravitating fluid as measured by the field of buoyancy and its spatial gradients. As part of that study, we conceive of *neutral directions* where infinitesimal displacements of a fluid element lead to no buoyancy forces; i.e., directions where the fluid element is neutrally buoyant.

READER'S GUIDE TO THIS CHAPTER

This chapter builds from our ideas of pressure in Chapters 22 and 26. Buoyancy is a feature of many of the subsequent chapters in this book given its central role in the motion of geophysical fluids.

25.1	Buoyancy force and Archimedes' Principle	362
25.1.1	Hydrostatic pressure force	362
25.1.2	Vertical buoyancy force	363
25.2	Mass density and its flavors	364
25.2.1	Realistic fluids	364
25.2.2	Equation of state in terms of potential temperature	364
25.2.3	Equation of state in terms of Conservative Temperature	365
25.2.4	Infinitesimal density increments and material time changes	365
25.2.5	Potential density	366
25.2.6	Linear equation of state for the ocean	367
25.2.7	Further study	367
25.3	Buoyancy stratification	367
25.3.1	Buoyancy as a field	367
25.3.2	Physical ideas underlying neutral directions	368
25.3.3	Comparing density under two forms of displacement	369
25.3.4	Buoyancy frequency	370
25.3.5	Buoyancy frequency and locally referenced potential density	371
25.4	Gravitational stability of a dry ideal gas atmosphere	371
25.5	Neutral directions	372
25.5.1	Neutral directions and the neutrality condition	373
25.5.2	Comments on neutral directions	373
25.6	Neutral helicity	373
25.6.1	Mathematical preliminaries	374
25.6.2	Helical nature of neutral displacements	374
25.6.3	Comments and further study	375
25.7	Exercises	375

25.1 Buoyancy force and Archimedes' Principle

Consider an arbitrary massive body or region contained within a fluid at rest, such as that shown in Figure 25.1. The body can be a rigid solid or perhaps a human swimmer. For our studies, we are most interested in arbitrary region of the fluid itself, such as an infinitesimal fluid element or a finite volume. In general, the massive body displaces a volume of fluid, with the mass of the displaced fluid given by the integral of the density over the region

$$m_{\text{fluid}} = \int_{\mathcal{R}} \rho \, dV. \quad (25.1)$$

We now seek an understanding of the forces acting on the massive body within the fluid.

25.1.1 Hydrostatic pressure force

As we will discuss in Chapter 26, any surface (even an imaginary surface) contained in a fluid experiences a contact stress due to interactions between the fluid and the surface. For a fluid at rest in a gravitational field, the only contact stress arises from pressure. Pressure is a compressive stress, acting in the direction determined by minus the outward normal along the surface. Integrating the pressure over the closed surface of \mathcal{R} leads to the pressure force acting on the region

$$\mathbf{F}_{\text{pressure}} = - \oint_{\partial\mathcal{R}} p \hat{\mathbf{n}} \, dS, \quad (25.2)$$

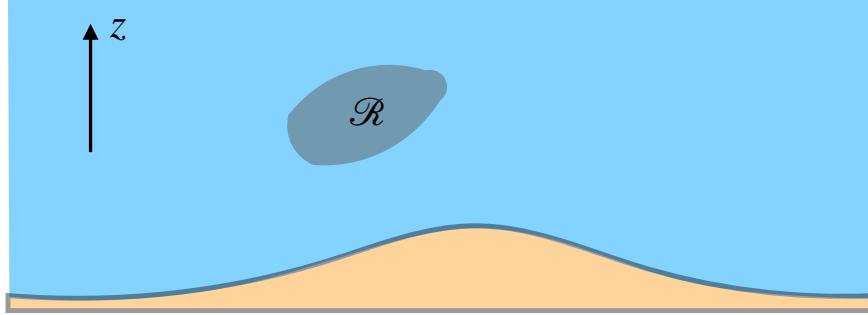


Figure 25.1: An arbitrary massive body, \mathcal{R} , within a fluid experiences a gravitational force acting down and a buoyant force acting up. The body is arbitrary, such as a rigid solid, a swimming creature, a fluid element, or a finite region of fluid. When the fluid is at rest, hydrostatics says that the fluid displaced by the body imparts a force equal to the weight of the displaced fluid. This result is Archimedes' Principle. The buoyancy force acting on the body is determined by the difference in weight between the body and the displaced fluid. A positive buoyancy force acts on a body less dense than the displaced fluid, and a negative buoyancy force acts on a body more dense.

where p is the hydrostatic pressure, $\hat{\mathbf{n}}$ is the outward normal on the boundary, and dS is the area element. Making use of Gauss's Law in the form of equation (4.63) leads to the equivalent expression in terms of the volume integral of the pressure gradient

$$\mathbf{F}_{\text{pressure}} = - \int_{\mathcal{R}} \nabla p \, dV. \quad (25.3)$$

This expression holds regardless the composition of the massive body. However, it is most useful when the body is a region of fluid so that the pressure gradient can be computed within the volume.

Since the fluid is at rest, the pressure field is hydrostatic and only has a dependence on the vertical position within the fluid, $p = p(z)$. The hydrostatic pressure equals to the weight per area of fluid sitting above any point in the fluid so that its vertical derivative given by

$$\frac{\partial p}{\partial z} = -\rho g. \quad (25.4)$$

Hence, the force acting on the region is

$$\mathbf{F}_{\text{pressure}} = - \int_{\mathcal{R}} \nabla p \, dV = - \int_{\mathcal{R}} \hat{\mathbf{z}} (\partial p / \partial z) \, dV = \hat{\mathbf{z}} g \int_{\mathcal{R}} \rho \, dV = \hat{\mathbf{z}} g m_{\text{fluid}}, \quad (25.5)$$

where we assumed a constant gravitational acceleration over \mathcal{R} . The hydrostatic pressure thus imparts a vertical upward force equal to the weight of the fluid displaced by the body. This result is a mathematical expression of Archimedes' Principle.

Note that there are no net horizontal forces acting on the body. The reason is that there are no horizontal pressure gradients within the fluid since, by assumption, the fluid is at rest and thus experiences no horizontal acceleration. So although the immersed body experiences horizontal compressive pressure forces, these forces balance when integrated over the body, thus leaving zero net horizontal acceleration.

25.1.2 Vertical buoyancy force

The vertical pressure force (25.6) acts to push the body vertically upward against gravity. It therefore appears to reduce the vertical gravitational force when the body is within the fluid. We

say that the massive body has a *buoyancy* imparted by the displaced fluid. If the body has a density less than the displaced fluid, then the body experiences a positive buoyancy force relative to the displaced fluid, in which case the body experiences an upward acceleration. The converse holds for a body more dense than the displaced fluid, in which case the body sinks downward.

We are thus motivated to define a *buoyancy force* acting on a massive body relative to the displaced fluid according to

$$\mathbf{F}_{\text{buoyancy}} = \hat{\mathbf{z}} g \int_{\mathcal{R}} (\rho_{\text{fluid}} - \rho_{\text{body}}) dV = \hat{\mathbf{z}} g (m_{\text{fluid}} - m_{\text{body}}). \quad (25.6)$$

The vertical buoyancy force is negative if the mass of the body is larger than that of the fluid it displaces, and conversely if the body is less massive. If the densities are equal, then the body is *neutrally* buoyant and thus experiences no net vertical force; i.e., it floats.

25.2 Mass density and its flavors

The density of a fluid element plays a central role in the study of buoyancy in geophysical fluids. More precisely, it is the density of the fluid element relative to that of its surroundings. As commonly referred to in geophysical fluid mechanics, the equation of state provides an expression for the mass density as a function of pressure, temperature, and material tracer concentration (salinity in the ocean and humidity in the atmosphere). This equation is called the *thermal equation of state* in the thermodynamics literature. We already encountered it for an ideal gas in Section 23.5.1. We here discuss the equation of state as well as the related flavors of mass density used to study stratified fluids.

25.2.1 Realistic fluids

To allow for general situations of multi-component fluids, we write the *in situ* density as a general function

$$\rho = \rho(S, T, p). \quad (25.7)$$

We made use of the standard oceanographic notation where $S = 1000 C$ is the salinity as defined by equation (23.23). This equation of state can be derived from one of the more general *fundamental equations of state* discussed in Section 23.3.

The thermal equation of state for fluid density is generally not as simple as that for an ideal gas discussed in Section 23.5.1. Indeed, liquids such as seawater have rather complex empirical expressions obtained from statistical fits to data. Part of the complexity arises from the multi-component nature of seawater (salt plus freshwater) as well as the nontrivial inter-molecular forces. In contrast, the equation of state for moist air can be written much like that for an ideal gas, thus making the equation of state for air far less complex than for seawater.

The equation of state (25.7) is a function of the *in situ* temperature, T , the *in situ* pressure, p , and the *in situ* salinity, S (ocean) or humidity (atmosphere). The term *in situ* means that a property is measured locally at a point in the fluid, with the resulting density also that measured at that point. Such *in situ* properties contrast to *potential* properties, which are based on referencing to a chosen pressure (e.g., potential temperature described in Section 23.7).

25.2.2 Equation of state in terms of potential temperature

Equation (23.91) provides a unique relation between potential temperature θ and *in situ* temperature, T , salinity and pressure. Furthermore, as discussed in Section 23.8.3, potential temperature

is more convenient for modeling than *in situ* temperature. Hence, it is useful to express the *in situ* density as a function of salinity, potential temperature, and pressure

$$\rho = \rho(S, \theta, p). \quad (25.8)$$

One comment on notation is key here. Namely, the functions $\rho(S, T, p)$ and $\rho(S, \theta, p)$ are distinct, so that their respective polynomial coefficients are different. Hence, one may wish to introduce distinct notation to distinguish the two functions; e.g., $\rho = F(S, T, p) = G(S, \theta, p)$. However, we choose brevity in notation by allowing the functional dependence to make the distinction. Doing so is standard in the oceanography literature.

25.2.3 Equation of state in terms of Conservative Temperature

For seawater, potential enthalpy is a more accurate means to measure the heat transfer through the ocean than potential temperature [McDougall \(2003\)](#). Consequently, since the advent of the Thermodynamic Equation of State 2010 ([IOC et al., 2010](#)), oceanographers make use of Conservative Temperature, Θ , rather than potential temperature for modeling and observations. For this reason we often compute the density as a function

$$\rho = \rho(S, \Theta, p). \quad (25.9)$$

For our purposes it is not important to make the distinction between potential temperature and Conservative Temperature. We use θ for brevity throughout this chapter.

25.2.4 Infinitesimal density increments and material time changes

Given the functional dependence for the equation of state written in terms of S, θ, p (equation (25.8)), an infinitesimal density increment is given by

$$d\rho = \left[\frac{\partial \rho}{\partial S} \right] dS + \left[\frac{\partial \rho}{\partial \theta} \right] d\theta + \left[\frac{\partial \rho}{\partial p} \right] dp \quad (25.10a)$$

$$\equiv \rho \beta dS - \rho \alpha d\theta + c_{\text{sound}}^{-2} dp. \quad (25.10b)$$

The second line introduced the following thermodynamic properties of the fluid

$$\beta = \frac{1}{\rho} \left[\frac{\partial \rho}{\partial S} \right]_{\theta, p} \quad \text{haline contraction coefficient} \quad (25.11)$$

$$\alpha = -\frac{1}{\rho} \left[\frac{\partial \rho}{\partial \theta} \right]_{S, p} \quad \text{thermal expansion coefficient} \quad (25.12)$$

$$c_{\text{sound}}^2 = \left[\frac{\partial \rho}{\partial p} \right]_{S, \theta} \quad \text{squared sound speed.} \quad (25.13)$$

The haline contraction coefficient, β , is considered for the ocean, where *haline* refers to salinity.¹

The infinitesimal density increment (25.10b) leads to the expression for the material change

$$\frac{1}{\rho} \frac{D\rho}{Dt} = \beta \frac{DS}{Dt} - \alpha \frac{D\theta}{Dt} + \frac{1}{\rho c_{\text{sound}}^2} \frac{Dp}{Dt}. \quad (25.14)$$

In the absence of mixing, the potential temperature and salinity are materially constant. In this case, the *in situ* density changes only through adiabatic processes that lead to pressure changes

$$\frac{D\rho}{Dt} = \frac{1}{c_{\text{sound}}^2} \frac{Dp}{Dt} \quad \text{adiabatic and isohaline changes.} \quad (25.15)$$

¹Note that in many chapters of this book, $\beta = \partial f / \partial y$ is the meridional derivative of the Coriolis parameter. We are careful to keep the two usages for β distinct.

25.2.5 Potential density

Isentropic motion of a frictionless fluid element generally occurs at materially constant potential temperature and materially constant tracer concentration (e.g., salinity or humidity) (see Section 23.7.3). We find it convenient to combine the evolution of salinity and potential temperature into the evolution of a single variable. *Potential density* is one such combination.

Defining the potential density

We start by recalling the definition of potential temperature, which is the temperature of a fluid element that is isentropically moved to the reference pressure, p_R . This isentropic displacement leads to the implicit expression for the potential temperature (equation (23.96))

$$\mathcal{S}(S, T, p) = \mathcal{S}(S, \theta, p_R). \quad (25.16)$$

We define potential density similarly, by computing the density a fluid element would have if isentropically moved to the same reference pressure as the potential temperature

$$\rho_{\text{pot}} = \rho(S, \theta, p_R). \quad (25.17)$$

As for the potential temperature, the reference pressure is often taken as that at sea level, though this is not necessary.

Material evolution of potential density

With the definition (25.17), the material evolution of potential density is given by

$$\frac{1}{\rho_R} \frac{D\rho_{\text{pot}}}{Dt} = \beta_R \frac{DS}{Dt} - \alpha_R \frac{D\theta}{Dt}, \quad (25.18)$$

where

$$\beta_R = \frac{1}{\rho_R} \left[\frac{\partial \rho(S, \theta, p_R)}{\partial S} \right]_\theta \quad \text{haline contraction at } p = p_R \quad (25.19)$$

$$\alpha_R = -\frac{1}{\rho_R} \left[\frac{\partial \rho(S, \theta, p_R)}{\partial \theta} \right]_S \quad \text{thermal expansion at } p = p_R \quad (25.20)$$

are the haline contraction and thermal expansion coefficients evaluated at the reference pressure $p = p_R$. Potential temperature and salinity are materially constant for adiabatic motion that also maintains constant matter content (e.g., isohaline) for fluid elements. By construction, potential density is also materially constant for this motion. This behavior is in contrast to *in situ* density, whose evolution is affected by pressure changes as seen by equations (25.14) and (25.15).

Reference pressures for ρ_{pot} and θ

As defined by equation (25.17), the reference pressure for the potential density is assumed to be the same as for the potential temperature. This assumption is common for the atmosphere, where the reference pressure is generally taken at the sea level. Likewise for the ocean, the potential temperature is generally computed using a sea level reference pressure. However, it is common to consider potential density with larger reference pressures, such as found in the ocean interior. Doing so is motivated by the rather strong nonlinear effects associated with the seawater equation

of state. In this case, pressure effects prompt one to choose a reference pressure closer to the *in situ* pressure near to the region of analysis.

Even though it is common to choose a potential density reference pressure distinct from the surface pressure, the potential temperature reference pressure generally remains at the surface. There is no fundamental problem with the use of distinct reference pressures for ρ_{pot} and θ . In particular, all of the above properties of potential density remain unchanged.

25.2.6 Linear equation of state for the ocean

For certain purposes, it is useful to approximate the equation of state used to study ocean fluid mechanics. One common idealization is to compute density as a linear function of potential temperature and salinity

$$\rho_{\text{pot}} = \rho_0 [1 - \alpha (\theta - \theta_0) + \beta (S - S_0)], \quad (25.21)$$

where α , β , θ_0 , and S_0 are positive constants. An even further simplification is to set salinity to a constant, so that density is just a linear function of potential temperature.

25.2.7 Further study

Chapter 1 of [Vallis \(2017\)](#) provides a pedagogical discussion of the equation of state for both a moist atmosphere and for seawater, as well as a discussion of the various flavors of density. The seawater equation of state is far more complex than the atmosphere, with [IOC et al. \(2010\)](#) providing the authoritative treatment.

25.3 Buoyancy stratification

We now return to the notions of buoyancy as introduced in Section 25.1, with a particular focus on how buoyancy is stratified.

25.3.1 Buoyancy as a field

Buoyancy is the gravitational acceleration that acts on a massive body due to the difference between the density of the body and the density of the local fluid environment. For geophysical fluid mechanics, we consider the massive body to be a fluid element whose presence does not alter the flow field; i.e., a “test” fluid element.

Locally defined environment

Consider a local definition of fluid buoyancy according to

$$b_{\text{local}} = g(\rho_{\text{environ}} - \rho_{\text{element}})/\rho_{\text{environ}} = g(1 - \rho_{\text{element}}/\rho_{\text{environ}}), \quad (25.22)$$

where ρ_{environ} is the local density of the fluid environment, and ρ_{element} is the density of the test fluid element within that environment. If the fluid element has a density greater than the environment, then it has a negative buoyancy, and vice versa.

We determine ρ_{element} by specifying its point of origin and specifying how it is moved (e.g., with or without mixing?) to the environment point. Conventional approaches are specified later in this section. The key point is that buoyancy as defined by equation (25.22) is a function of the path that the test fluid element takes to reach the environment point. This subjectivity lends ambiguity in the definition of buoyancy. We remove this ambiguity by asking specific questions

about local buoyancy. For example, if the fluid element moves an infinitesimal distance while mixing its temperature and salinity with the environment, what direction maintains a neutrally buoyant state for the fluid element? This question forms the basis for defining *neutral directions*.

Globally defined environment

The definition (25.22) accepts that buoyancy is a relative field in which at each point in the fluid one must redefine the environment to which to compare the density of the test fluid element. However, there are cases in which it is sufficient to define a globally constant environment with a constant density, ρ_{ref} . In this case we consider the global buoyancy as

$$b_{\text{global}} = g(1 - \rho/\rho_{\text{ref}}), \quad (25.23)$$

where we compute ρ according to the local environmental density. This definition is useful for idealized cases where the *in situ* density is not a function of pressure. In this case buoyancy is a function only of potential temperature and salinity so that we can make use of potential density to measure buoyancy (as explained below).

Although numerical value of b_{global} is a function of the reference density, what is more relevant is the buoyancy of one fluid element relative to another, in which case the reference density is irrelevant since it is a global constant. Furthermore, with a globally constant environmental density, the buoyancy becomes a local function of space. That is, we no longer compare the fluid element density to a changing local density. Instead, we compute the local density and compare it to the reference density. We can thus determine b_{global} at a point through information just at that point. Correspondingly, we can map b_{global} and determine the relative buoyancy of fluid elements anywhere in the fluid.

Comments

Although appealing, the use of a globally constant reference density has its limitations for real fluids whose density is a function of pressure. Hence, for the remainder of this section we retain the local definition of buoyancy given by equation (25.22) and make use of the local definition to determine neutral directions as defined below.

25.3.2 Physical ideas underlying neutral directions

As a constant mass fluid element moves through the ocean and atmosphere, it is exposed to a suite of physical processes that modify its thermal, material, and mechanical properties; i.e., its θ , S , and p . Modification of its pressure occurs through contact stresses with other fluid elements (Chapter 26). Modification of the thermal and material properties occurs through the exchange of heat and matter with adjacent fluid elements. The exchange of heat and matter occurs only in the presence of irreversible processes such as mixing, whereas mechanical exchanges occur either reversibly (pressure exchange) or irreversibly (viscous exchange; Section 26.2).

The exchange of heat and matter generally alters the density of a fluid element relative to its local environment, thus affecting the buoyancy of the fluid element. However, it is possible to modify θ and S without modifying its buoyancy defined local to the environment surrounding the fluid element. To do so requires changes in θ to precisely compensate changes in S . Directions in space determined by such compensated changes define *neutral directions*.

The *in situ* density generally changes when a fluid element is displaced an infinitesimal distance, $d\mathbf{x}$. How the density changes is determined by how the element interacts with the surrounding environment. We examine two cases.

- Displace the fluid element allowing for θ , S , and p to equilibrate with the local environment; i.e., full mixing of the element with the environment.
- Displace the fluid element without changing θ and S yet allowing p to equilibrate with the local environment; i.e., no mixing of the element and the environment.

Notably, the pressure of the fluid element is modified the same amount under both displacements since in both cases the element reaches the same mechanical equilibrium with the local environment. Hence, subtracting the *in situ* density of the above two displaced elements removes the effects from pressure changes, leaving only the effects on density from changes to θ and S . Directions where the *in situ* density is the same for the two forms of parcel displacements determine neutral directions.

The above thought experiment is identical to that used to determine the gravitational stability of a vertical fluid column (e.g., Section 3.6 of [Gill, 1982](#)), whereby $d\mathbf{x} = \hat{\mathbf{z}} dz$ is a vertical displacement. The only difference is that neutral directions consider arbitrary three dimensional displacements rather than just vertical displacements.

25.3.3 Comparing density under two forms of displacement

Consider an infinitesimal displacement, $d\mathbf{x}$, of a fluid element and examine how its *in situ* density $\rho = \rho(S, \theta, p)$ changes under two different displacements. First, assume that the element exchanges heat and salt with the surroundings as it moves, and that it feels the local pressure. In equilibrium, density at the new location agrees with that of the local environment, $\rho(\mathbf{x} + d\mathbf{x})$, since the salinity, potential temperature, and pressure are those of the local environment. To leading order, the density difference between the two positions is computed according to

$$d\rho = \rho(\mathbf{x} + d\mathbf{x}) - \rho(\mathbf{x}) \quad (25.24a)$$

$$= \rho[S(\mathbf{x} + d\mathbf{x}), \theta(\mathbf{x} + d\mathbf{x}), p(\mathbf{x} + d\mathbf{x})] - \rho[S(\mathbf{x}), \theta(\mathbf{x}), p(\mathbf{x})] \quad (25.24b)$$

$$= d\mathbf{x} \cdot \left[\frac{\partial \rho}{\partial \theta} \nabla \theta + \frac{\partial \rho}{\partial S} \nabla S + \frac{\partial \rho}{\partial p} \nabla p \right] \quad (25.24c)$$

$$= \rho d\mathbf{x} \cdot \left[-\alpha \nabla \theta + \beta \nabla S + \frac{1}{\rho c_{\text{sound}}^2} \nabla p \right]. \quad (25.24d)$$

For the second displacement, do not allow the fluid element to exchange (mix) heat or salt with the environment, thus undergoing an adiabatic and isohaline motion. In this case, the element's density change is associated just with pressure changes

$$(d\rho)_{(\text{no mix})} = \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} - \rho(\mathbf{x}) \quad (25.25a)$$

$$= \rho[S(\mathbf{x}), \theta(\mathbf{x}), p(\mathbf{x} + d\mathbf{x})] - \rho[S(\mathbf{x}), \theta(\mathbf{x}), p(\mathbf{x})] \quad (25.25b)$$

$$= \rho d\mathbf{x} \cdot \left[\frac{1}{\rho c_{\text{sound}}^2} \nabla p \right]. \quad (25.25c)$$

That is, when the fluid element moves through the fluid without exchange of heat or salt, then the only way for the *in situ* density to change is via pressure changes. Comparing the two densities renders

$$\rho(\mathbf{x} + d\mathbf{x}) - \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} = d\rho - (d\rho)_{(\text{no mix})} = \rho d\mathbf{x} \cdot [-\alpha \nabla \theta + \beta \nabla S]. \quad (25.26)$$

Figure 25.2 offers a schematic of the calculation used to determine neutral directions for the special case of a vertical displacement.

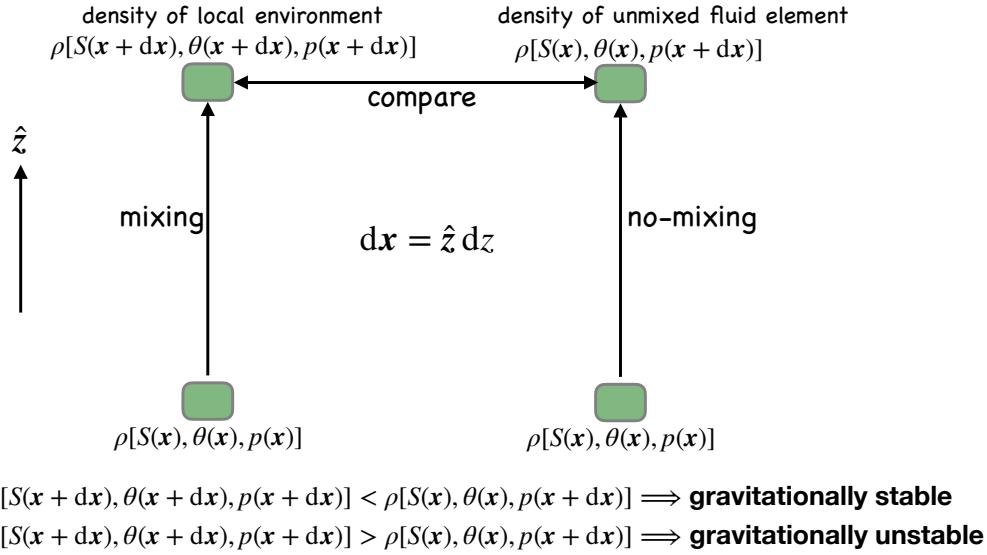


Figure 25.2: Schematic of the calculation used to examine whether a fluid column is gravitationally stable under a vertical displacement of a fluid element. A fluid element is displaced from its original location at a position x to a close position $x + dx$, with $dx = \hat{z} dz$ and $dz > 0$ in this figure. If the displacement of an element occurs without mixing, it carries its potential temperature and matter concentration with it unchanged, whereas the pressure acting on the element generally changes to that of the local environment (right path). Comparing this “no mixing” displaced fluid element to the density of the local environment (left path) determines whether the density stratification of the fluid environment is gravitationally stable or unstable.

25.3.4 Buoyancy frequency

The special case of a vertical displacement (Figure 25.2) yields

$$\rho(z + dz) - \rho(z + dz)_{(\text{no mix})} = \rho dz \left[-\alpha \frac{\partial \theta}{\partial z} + \beta \frac{\partial S}{\partial z} \right]. \quad (25.27)$$

Consider a vertically upward displacement so that $dz > 0$. If the surrounding environment has a lower density than the adiabatic and isohaline displaced fluid element, $\rho(z + dz) < \rho(z + dz)_{(\text{no mix})}$, then the element will feel a buoyancy force returning it to the original depth. The restorative buoyancy force per volume is written

$$g [\rho(z + dz) - \rho(z + dz)_{(\text{no mix})}] = g \rho dz \left[-\alpha \frac{\partial \theta}{\partial z} + \beta \frac{\partial S}{\partial z} \right] \equiv -N^2 \rho dz, \quad (25.28)$$

where we defined the squared buoyancy frequency

$$N^2 = g \left[\alpha \frac{\partial \theta}{\partial z} - \beta \frac{\partial S}{\partial z} \right]. \quad (25.29)$$

Stable vertical motion results from a background density profile with $N^2 > 0$. An unstable profile occurs when $N^2 < 0$, in which case motion of the fluid element results in an exponential growth associated with a *gravitational instability*. That is, when the fluid column is unstably stratified, an initially tiny vertical displacement will lead to an even larger displacement, thus causing the perturbation to grow unbounded, which is the sign of an instability.

25.3.5 Buoyancy frequency and locally referenced potential density

Equation (25.29) defines the squared buoyancy frequency in terms of the vertical temperature and salinity gradients. This expression is identical to the vertical gradient of the potential density (25.17), when the reference pressure for density is taken local to the point where the buoyancy frequency is computed. That is, the vertical gradient of the *locally referenced potential density* provides a measure of the vertical stratification

$$N^2 = -g \left[\frac{\partial \ln \rho_{\text{pot}}}{\partial z} \right]_{p_{\text{R}}=p} = g \left[\alpha \frac{\partial \theta}{\partial z} - \beta \frac{\partial S}{\partial z} \right]. \quad (25.30)$$

Note that at a point in the fluid, the locally referenced potential density equals to the *in situ* density. However, when probing nearby points, and thus in taking spatial gradients, the two have distinct gradients. Namely, the *in situ* density is modified by pressure gradients, whereas spatial gradients of the locally referenced potential density remove pressure effects.

25.4 Gravitational stability of a dry ideal gas atmosphere

We introduced the adiabatic lapse rate in Section 23.6 as a measure of how temperature varies as a function of pressure or depth. For an ideal gas atmosphere, the squared buoyancy frequency can be written (see exercise 25.1)

$$N^2 = \frac{g}{\theta} \frac{\partial \theta}{\partial z}. \quad (25.31)$$

The potential temperature for an ideal gas is given by equation (23.106)

$$\theta = T \left[\frac{p_{\text{R}}}{p} \right]^{\varphi} \quad (25.32)$$

where

$$\varphi = \frac{R^{\text{M}}}{c_p} \quad (25.33)$$

is a constant for a simple ideal gas. Consequently, the squared buoyancy frequency takes the form

$$g^{-1} N^2 = \frac{\partial \ln \theta}{\partial z} \quad (25.34a)$$

$$= \frac{\partial \ln T}{\partial z} - \varphi \frac{\partial \ln p}{\partial z}. \quad (25.34b)$$

For a hydrostatic fluid with a constant gravitational acceleration, the vertical derivative of pressure is given by

$$\frac{\partial p}{\partial z} = -\rho g, \quad (25.35)$$

so that pressure at a point in the fluid equals to the weight per area above that point. Using this result leads to the squared buoyancy frequency

$$g^{-1} N^2 = \frac{\partial \ln T}{\partial z} + \frac{\varphi g \rho}{p} \quad (25.36a)$$

$$= \frac{1}{T} \frac{\partial T}{\partial z} + \frac{g}{c_p T}, \quad (25.36b)$$

where we used the ideal gas relation $p = \rho T R^{\text{M}}$ for the final step.

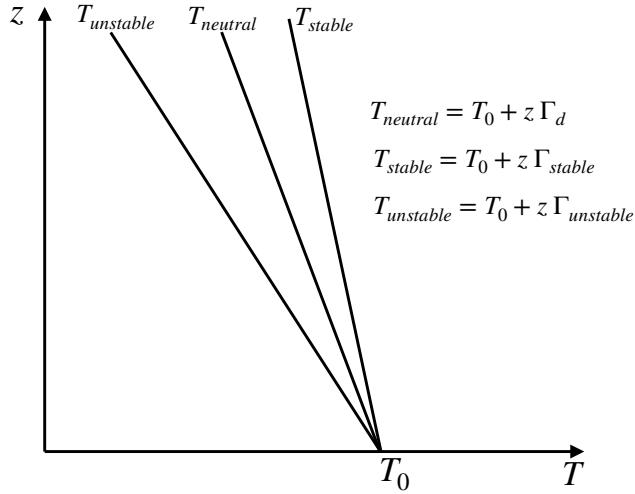


Figure 25.3: Three vertical profiles of *in situ* temperature in a dry ideal gas atmosphere. The neutrally stable profile has $T_{\text{neutral}} = T_0 + z \Gamma_d$, where $\Gamma_d = -g/c_p \approx -9.8 \text{ K}/(1000 \text{ m})$. A vertically unstable profile (heavy over light) has $T_{\text{unstable}} = T_0 + z \Gamma_{\text{unstable}}$ where $\Gamma_{\text{unstable}} < \Gamma_d = -g/c_p$. In contrast, the gravitationally stable atmosphere has $T_{\text{stable}} = T_0 + z \Gamma_{\text{stable}}$ where $\Gamma_{\text{stable}} > \Gamma_d = -g/c_p$.

A vanishing buoyancy frequency, or equivalently a vanishing vertical derivative of potential temperature, occurs when the vertical temperature gradient equals to the dry adiabatic lapse rate

$$N^2 = 0 \iff \frac{\partial T}{\partial z} = \Gamma_d, \quad (25.37)$$

where (see equation (23.89))

$$\Gamma_d = -\frac{g}{c_p} \approx -9.8 \text{ K}/(1000 \text{ m}). \quad (25.38)$$

That is, if the *in situ* temperature decreases upon ascent more strongly than the dry adiabatic lapse rate, then the vertical column is gravitationally unstable. In effect, the column becomes top heavy and subject to overturning. We summarize this stability criteria as

$$\text{stable} \quad N^2 > 0 \iff -\frac{\partial T}{\partial z} < \frac{g}{c_p} \quad (25.39)$$

$$\text{unstable} \quad N^2 < 0 \iff -\frac{\partial T}{\partial z} > \frac{g}{c_p}, \quad (25.40)$$

with Figure 25.3 providing an illustration for three linear temperature profiles.

25.5 Neutral directions

Rather than specializing to a vertical displacement as for the buoyancy frequency in Section 25.3.4, consider an arbitrary three-dimensional displacement. The resulting directions are known as *neutral directions*.

25.5.1 Neutral directions and the neutrality condition

To determine neutral directions, return to equation (25.26) to write

$$\rho(\mathbf{x} + d\mathbf{x}) - \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} = \rho(\mathbf{x}) d\mathbf{x} \cdot [-\alpha \nabla \theta + \beta \nabla S] \quad (25.41a)$$

$$= \rho(\mathbf{x}) d\mathbf{x} \cdot \hat{\gamma} | -\alpha \nabla \theta + \beta \nabla S|. \quad (25.41b)$$

The second expression introduced the *dianeutral unit vector*

$$\hat{\gamma} = \frac{\rho_\theta \nabla \theta + \rho_S \nabla S}{|\rho_\theta \nabla \theta + \rho_S \nabla S|} = \frac{-\alpha \nabla \theta + \beta \nabla S}{|-\alpha \nabla \theta + \beta \nabla S|}. \quad (25.42)$$

Displacements, $d\mathbf{x}$, orthogonal to $\hat{\gamma}$ lead to no difference in the density between the environment and the unmixed fluid element

$$d\mathbf{x} \cdot \hat{\gamma} = 0 \Rightarrow \rho(\mathbf{x} + d\mathbf{x}) = \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} \iff \text{neutral displacements.} \quad (25.43)$$

Such displacements are said to occur along a *neutral direction*. Neutral directions generalize the notion of buoyancy surfaces or statification surfaces within the ocean. Motion perpendicular to such surfaces is suppressed through the restoring force from buoyancy.

For a neutral displacement let us write

$$d\mathbf{x} = \hat{\mathbf{t}} ds, \quad (25.44)$$

where ds is the arc-distance along the neutral direction and $\hat{\mathbf{t}}$ is the local tangent parallel to the neutral direction. The neutral displacement condition (25.43) leads to the *neutrality condition*

$$\alpha \nabla_{\text{neut}} \theta = \beta \nabla_{\text{neut}} S, \quad (25.45)$$

where

$$\nabla_{\text{neut}} = \hat{\mathbf{t}} \cdot \nabla \quad (25.46)$$

is the gradient operator oriented parallel to the neutral direction.

25.5.2 Comments on neutral directions

As defined, neutral displacements generally occur via the irreversible mixing of θ and S . To be displaced along a neutral direction requires the mixing of θ to precisely balance that of S so that the fluid element's *in situ* density remains identical to that of the local environment. That is, a fluid element displaced along a neutral direction leaves the *in situ* density of the element equal to that of the local environment, thus engendering no local buoyancy force on the element.

25.6 Neutral helicity

As discussed in Section 25.5, movement of a fluid element along a neutral direction requires the mixing of θ and S , with θ mixing precisely balanced by S mixing so that the fluid element's *in situ* density remains identical to that of the local environment. In so doing, the fluid element encounters no locally defined buoyancy force, thus prompting the name "neutral direction". We make the balance of θ mixing and S mixing precise when presenting the neutrality condition in Section 39.3.4. What we ask here concerns the path taken when undergoing a suite of neutral displacements. In particular, if a suite of neutral displacements close in latitude/longitude space, will they also close in depth? As we show here, neutral displacements generally do not close due to a property of seawater known as *neutral helicity*, thus revealing a nontrivial helical topology.

25.6.1 Mathematical preliminaries

Consider a simply connected smooth surface with outward normal written in the form

$$\hat{\mathbf{n}} = |\mathbf{N}|^{-1} \mathbf{N}. \quad (25.47)$$

Likewise, consider a unit vector $\hat{\mathbf{t}}$ that lives within the surface and is directed tangent to an arbitrary closed loop. Since $\hat{\mathbf{n}} \cdot \hat{\mathbf{t}} = 0$ by construction, we can integrate around an arbitrary closed loop within the surface and still maintain the trivial result

$$\oint_{\partial\mathcal{S}} \mathbf{N} \cdot \hat{\mathbf{t}} \, dl = 0. \quad (25.48)$$

Note that we chose a counterclockwise orientation of the loop around the boundary, with $\partial\mathcal{S}$ denoting the boundary of the area \mathcal{S} within the surface.

Now apply Stokes' Theorem to the loop integral (25.48) to yield

$$\int_{\mathcal{S}} (\nabla \wedge \mathbf{N}) \cdot \hat{\mathbf{n}} \, d\mathcal{S} = 0, \quad (25.49)$$

where $d\mathcal{S}$ is the area element in the surface \mathcal{S} with outward normal $\hat{\mathbf{n}}$. Since the closed path is arbitrary, the area integral (25.49) vanishes only if the integrand is identically zero. We conclude that for the surface to be simply connected requires that the helicity must vanish

$$\mathcal{H} = \mathbf{N} \cdot (\nabla \wedge \mathbf{N}) = 0 \implies \text{simply connected surface.} \quad (25.50)$$

25.6.2 Helical nature of neutral displacements

Now apply the above mathematical results towards the question of whether a neutral surface is simply connected. For that purpose we set

$$\mathbf{N} = -\alpha \nabla \theta + \beta \nabla S \quad (25.51)$$

rendering the neutral helicity

$$\mathcal{H}_{\text{neutral}} = (-\alpha \nabla \theta + \beta \nabla S) \cdot \nabla \wedge (-\alpha \nabla \theta + \beta \nabla S), \quad (25.52)$$

which can be written

$$\mathcal{H}_{\text{neutral}} = (-\alpha \nabla \theta + \beta \nabla S) \cdot \nabla \wedge (-\alpha \nabla \theta + \beta \nabla S) \quad (25.53a)$$

$$= -\alpha \nabla \theta \cdot (\nabla \wedge \beta \nabla S) - \beta \nabla S \cdot (\nabla \wedge \alpha \nabla \theta) \quad (25.53b)$$

$$= -\alpha \nabla \theta \cdot (\nabla \beta \wedge \nabla S) - \beta \nabla S \cdot (\nabla \alpha \wedge \nabla \theta). \quad (25.53c)$$

Expand the gradients of α and β according to

$$\nabla \alpha = \alpha_\theta \nabla \theta + \alpha_S \nabla S + \alpha_p \nabla p \quad (25.54a)$$

$$\nabla \beta = \beta_\theta \nabla \theta + \beta_S \nabla S + \beta_p \nabla p, \quad (25.54b)$$

so that

$$-\mathcal{H}_{\text{neutral}} = \alpha \nabla \theta \cdot (\beta_\theta \nabla \theta + \beta_p \nabla p) + \beta \nabla S \cdot (\alpha_S \nabla S + \alpha_p \nabla p) \quad (25.55a)$$

$$= \alpha \nabla \theta \cdot (\nabla p \wedge \nabla S) \beta_p + \beta \nabla S \cdot (\nabla p \wedge \nabla \theta) \alpha_p \quad (25.55b)$$

$$= \nabla p \cdot (\nabla S \wedge \nabla \theta) (\alpha \beta_p - \beta \alpha_p). \quad (25.55c)$$

Introducing the thermobaricity parameter from Section 39.4

$$\mathcal{T} = \beta \partial_p(\alpha/\beta) \quad (25.56)$$

renders the tidy result

$$\mathcal{H}_{\text{neutral}} = \beta \mathcal{T} \nabla p \cdot (\nabla S \wedge \nabla \theta). \quad (25.57)$$

Consequently, a nonzero neutral helicity is fundamentally related to a nonzero thermobaricity parameter \mathcal{T} . It is also associated with the non-zero volume for a parallelopiped in (θ, S, p) space (see Section 3.5)

$$\nabla p \cdot (\nabla S \wedge \nabla \theta) = \nabla \theta \cdot (\nabla p \wedge \nabla S) = \nabla S \cdot (\nabla \theta \wedge \nabla p). \quad (25.58)$$

A nonzero $\mathcal{H}_{\text{neutral}}$ means that a path defined by the accumulation of neutral directions does not close. Rather, they possess a helical structure.

25.6.3 Comments and further study

Neutral directions were introduced to oceanography by [McDougall \(1987a\)](#) and [McDougall \(1987b\)](#), and they are the basis for how oceanographers think about stratification. [McDougall et al. \(2014\)](#) offer another presentation of why neutral directions are relevant for the ocean. Section 2.7.2 of [Olbers et al. \(2012\)](#) offers a concise and pedagogical summary of neutral directions.

Neutral helicity is a property of seawater that is revealed through the neutrally compensated mixing of θ and S . That is, neutral helicity is not a direct property of mixing though mixing is required to determine neutral directions.

[McDougall and Jackett \(1988\)](#) were the first to make note of the helical nature of neutral displacements, and [McDougall and Jackett \(2007\)](#) provide more elaboration and analysis from observational based measurements. This property of the seawater equation of state, though somewhat exotic, has some very practical implications on the choice for vertical coordinate used in realistic numerical ocean climate models. [Stanley \(2019\)](#) offers recent insights into the mathematics of neutral directions.

25.7 Exercises

EXERCISE 25.1: BUOYANCY FREQUENCY FOR AN IDEAL GAS

Write the expression for the squared buoyancy frequency of an ideal gas. Hint: first derive the expression for the potential density and then take its vertical derivative as per equation (25.30).

EXERCISE 25.2: EXAMPLES OF BUOYANCY PERIOD

Using approximate but realistic values for the observed stratification, determine the buoyancy period ($T_b = 2\pi/N$) for

- mid-latitude troposphere
- stratosphere
- ocean thermocline
- ocean abyss.

Provide references for where you obtained the observed stratification. Hint: for both the atmosphere and ocean, it is sufficient to assume stratification is dominated by potential temperature.

EXERCISE 25.3: VERTICAL INTEGRAL OF N^2

The expression for squared buoyancy frequency

$$N^2 = -g \left[\frac{\partial \ln \rho_{\text{pot}}}{\partial z} \right]_{p_R=p} \quad (25.59)$$

makes it tempting to consider its vertical integral according to

$$-g^{-1} \int_{-H}^{\eta} N^2 dz \stackrel{?}{=} [\ln \rho_{\text{pot}}]_{\eta} - [\ln \rho_{\text{pot}}]_{-H}. \quad (25.60)$$

Discuss what is wrong with this equation. Under what conditions is it correct?

26

Stress

A fluid element experiences two kinds of forces. *External* or *body* forces act throughout the fluid element. Their accumulated effect over a finite region results from integrating the body force at each point within the region. In geophysical fluid mechanics, we are concerned with body forces from the effective gravitational acceleration (central gravity plus centrifugal) and the Coriolis acceleration. These body forces share physical characteristics with point particles as described in Chapter 14.

The second kind of forces are called *internal* or *contact* forces. These forces act on a fluid element due to its local interactions with the surrounding fluid media through the exchange of dynamical properties (e.g., momentum, kinetic energy, vorticity) matter, and thermodynamical properties such as temperature and entropy. The exchange of dynamical properties is mediated by mechanical interactions, with this exchange mathematically formulated as a stress (force per unit area) acting on the boundary of the fluid element. Contact forces are in local mechanical equilibrium at each point in the interior of a fluid region as per Newton's 3rd law (action/reaction). Consequently, the net contact forces acting on a finite fluid region arise just from those contact forces acting at the region's boundary. Contact forces communicated through stresses embody the fundamental distinction between the dynamics of a fluid element and the dynamics of a point particle.

We introduced contact forces in Chapters 22 and 27 when deriving the fluid equations of motion. In this chapter we further pursue our study of stress by examining physical properties and deriving their mathematical expressions. We also discuss conditions placed on stress and velocity at boundaries.

READER'S GUIDE TO THIS CHAPTER

This chapter builds from our study of forces, acceleration, and Newton's laws as applied to a fluid continuum in Chapters 22 and 27. Fluid stresses can be organized into a second order tensor and further decomposed into pressure stresses and viscous stresses. Understanding the mathematical and physical aspects of stress is important for the suite of fluid models studied in this book. Because the material involves vectors and tensors it can require a bit more patience from the reader than analogous chapters that discuss scalar fields. To make the formalism less onerous we make use of Cartesian tensors as discussed in Chapters 3 and 4. Results can be readily generalized for arbitrary coordinates through the rules of general covariance detailed in Chapter 7.

26.1	Cauchy's stress principle and Newton's laws	378
26.1.1	Cauchy's stress principle	378
26.1.2	Local equilibrium of contact stresses	379
26.1.3	Comments on the local equilibrium relation	379
26.2	The stress tensor	380
26.3	Angular momentum and stress tensor symmetry	382
26.4	Flux-form Eulerian momentum equation	383
26.5	Linear momentum for arbitrary regions	384
26.6	Relating stress to pressure and strain	385
26.6.1	Hydrostatic pressure	385
26.6.2	Friction tensor for a Newtonian fluid	385
26.6.3	Kinetic energy dissipation and Joule heating	387
26.6.4	Comments and further study	389
26.7	The Reynolds number and turbulence	389
26.8	Form stress	390
26.9	Boundary conditions	391
26.9.1	Formulation	392
26.9.2	Matter transport boundary condition	392
26.9.3	Momentum transport arising from matter transport	393
26.9.4	Net boundary condition	393
26.9.5	No-slip boundary condition	394
26.10	Regional ocean angular momentum budget	394
26.10.1	Depth integrated angular momentum budget	395
26.10.2	Surface and bottom form stresses	396
26.10.3	Surface and bottom frictional stresses	396
26.10.4	Summary budget of angular momentum for a fluid column	397
26.10.5	Steady state balances	397
26.10.6	Further study	398
26.11	Exercises	398

26.1 Cauchy's stress principle and Newton's laws

We here develop some general properties of contact forces and the associated contact stresses. For that purpose, consider an arbitrary smooth closed region, \mathcal{R} , of fluid with volume $V = \int_{\mathcal{R}} dV$ and mass $M = \int_{\mathcal{R}} \rho dV$ (Figure 26.1). Furthermore, let $\partial\mathcal{R}$ be the bounding surface for the region, and let $\hat{\mathbf{n}}$ be the outward normal at a point on the boundary.

26.1.1 Cauchy's stress principle

The bounding surface of the region experiences mechanical interactions with the surrounding fluid continuum and these interactions lead to contact forces acting on the boundary. Let $\boldsymbol{\tau}$ be the contact stress vector (force per unit area) acting at a point on $\partial\mathcal{R}$. Cauchy's stress principle asserts that the contact stress vector is a function of the position, time, and boundary normal

$$\boldsymbol{\tau} = \boldsymbol{\tau}(\mathbf{x}, t, \hat{\mathbf{n}}). \quad (26.1)$$

For example, a static fluid in hydrostatic balance has the contact stress vector equal to just the pressure

$$\boldsymbol{\tau} = -p(\mathbf{x}, t) \hat{\mathbf{n}} \quad \text{static fluid in hydrostatic balance.} \quad (26.2)$$

We return to this form of the stress vector in Section 26.6.1.

Cauchy's stress principle is sensible for points within the fluid media, and its relevance has been supported by experimental studies over the time since Cauchy made this assertion in the year 1823. Furthermore, it holds for pressure and viscous stresses at the interface between fluid media or at solid-earth boundaries. However, Cauchy's stress principle does not hold for surface tension, which is proportional to the curvature of the surface separating two fluid media (e.g., atmosphere and ocean), where curvature involves spatial gradients of the normal vector (see Chapter 35).

26.1.2 Local equilibrium of contact stresses

Newton's 2nd law (unbalanced forces alter linear momentum) provides a statement for the conservation of linear momentum. When applied to the fluid region it says that the change in momentum for the region equals to the net force acting on the region. Assuming we follow the barycentric velocity of the region, Newton's law states that the material time evolution of the region's momentum is given by

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV = \int_{\mathcal{R}(\mathbf{v})} \mathbf{f} \rho dV + \oint_{\partial\mathcal{R}(\mathbf{v})} \boldsymbol{\tau} d\mathcal{S}, \quad (26.3)$$

where $\int_{\mathcal{R}(\mathbf{v})} \mathbf{f} \rho dV$ is the domain integrated body force (from gravity, centrifugal, and Coriolis), and $\mathcal{R}(\mathbf{v})$ is the region that follows the flow. To develop a general property for the contact forces, consider this balance for a region whose size gets infinitesimally small. The volume integrals in this balance are proportional to L^3 , where L is a length scale measuring the size of the region (e.g., radius for a spherical region). However, the integral of the contact forces goes to zero as L^2 . Self-consistency for this balance requires the contact forces to satisfy the limiting behavior

$$\lim_{L \rightarrow 0} \frac{1}{L^2} \oint_{\partial\mathcal{R}(\mathbf{v})} \boldsymbol{\tau} d\mathcal{S} = 0. \quad (26.4)$$

Hence, contact forces at a point in the fluid are locally in equilibrium. As a corollary, contact stress vectors respecting Cauchy's principle satisfy

$$\boldsymbol{\tau}(\mathbf{x}, t, \hat{\mathbf{n}}) = -\boldsymbol{\tau}(\mathbf{x}, t, -\hat{\mathbf{n}}), \quad (26.5)$$

which is an expression of Newton's 3rd law (action/reaction law) in terms of the contact stress vectors.

26.1.3 Comments on the local equilibrium relation

The local equilibrium relation (26.4), and the corresponding expression of Newton's 3rd law, (26.5), might suggest that contact forces, such as pressure, cannot lead to motion. However, that suggestion is incorrect since contact stresses integrated over a finite region can lead to a net force that causes motion. Since contact forces within the domain interior cancel pointwise, the local equilibrium relation (26.4) says that the net contact force acting on the region arises only from the area integrated contact forces acting on the region boundary. Local or pointwise mechanical equilibrium does not imply mechanical equilibrium for finite regions.

To further emphasize the above point, consider an ocean region bounded at its bottom by the solid earth and its upper surface by a massive atmosphere. Variations in contact stresses over finite regions within the ocean fluid lead to accelerations; e.g., ocean circulation. However, when integrated over the full ocean domain, all contact stresses cancel pointwise. Consequently, the net contact forces acting on the full ocean domain reduce to those just on the ocean boundaries. The

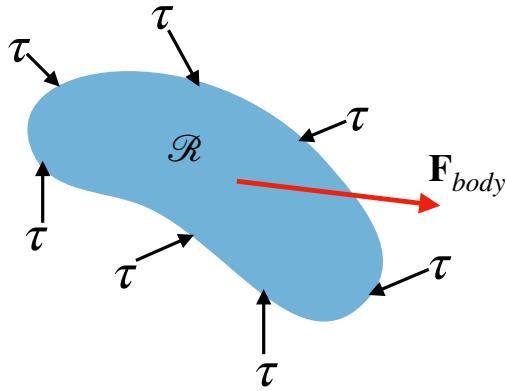


Figure 26.1: Schematic of the net external or body force, \mathbf{F}_{body} , acting on a finite region of fluid, plus the accumulation of internal or contact stress vectors, τ , acting on the region boundaries. The net body force is determined by a volume integral of the body force at each point within the volume. In contrast, since the contact stresses are in local equilibrium, the volume integral of the contact stresses reduces to an area integral of the contact stress over the region boundary. The area integrated contribution to horizontal accelerations is often referred to as *form stress*, with that coming from the bottom boundary referred to as *topographic form stress* and the surface (air-sea) boundary as *atmospheric form stress* or *oceanic form stress*, depending on the media considered. We have more to say about form stress in Section 26.8 for a general fluid, and in Section 43.2 for the shallow water fluid.

boundary contact forces arise from mechanical interactions with the solid-earth and the overlying atmosphere. The center of mass for the ocean basin remains static if the accumulation of contact forces acting over its boundaries plus the volume integrated body forces from gravity and Coriolis sum to zero.

In Figure 26.2 we illustrate the net pressure force acting on an arbitrary fluid domain. Pressure acts solely in a compressive manner as directed along the inward normal to the domain. Area integration over a domain boundary renders the net pressure force acting on the domain

$$\mathbf{F}^{\text{pressure}} = - \oint_{\partial\mathcal{R}} p \hat{\mathbf{n}} d\mathcal{S} = - \int_{\mathcal{R}} \nabla p dV, \quad (26.6)$$

where the second equality follows from application of Gauss's divergence theorem for a scalar field (Section 4.7.2). When decomposed according to coordinate axes, the pressure force acting on the boundary has a component in both the vertical and horizontal directions, thus contributing to both vertical and horizontal accelerations. The vertical accelerations are closely balanced by the weight of fluid, with exact balance in the case of a hydrostatic fluid. The area integrated contribution to horizontal accelerations is referred to as *form stress*.

26.2 The stress tensor

Cauchy's stress principle reduces the mathematical complexity of describing contact stresses. To further pursue that description, consider the tetrahedron fluid region shown in Figure 26.3, where three of the four sides are aligned according to the Cartesian coordinate axes and the fourth side has an outward normal, $\hat{\mathbf{n}} = (\hat{n}_1, \hat{n}_2, \hat{n}_3)$, projecting into all three directions. The results developed for this rather contrived region hold for an arbitrary region using arbitrary coordinates. The reason for this generality is that we make use of tensor analysis, thus offering the means to move from the specific to the general.

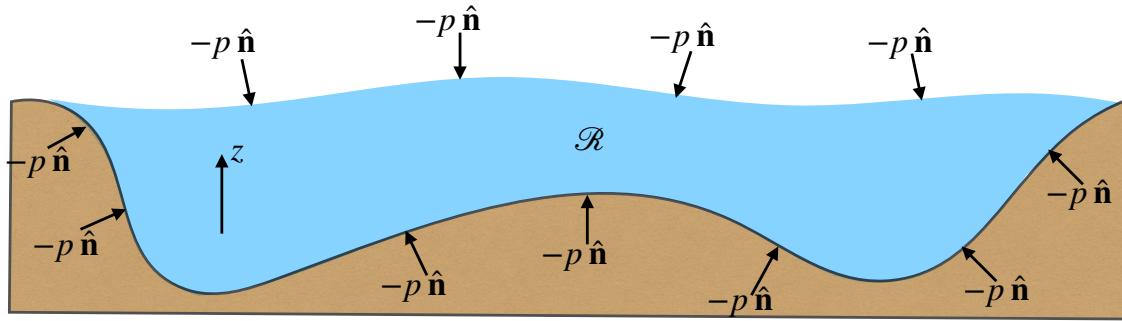


Figure 26.2: Schematic of contact forces from pressure acting on the boundaries to an ocean domain. Pressure forces are directed according to minus the local normal (compressive force). As with all contact forces, the pressure forces acting in the interior of the ocean are locally in equilibrium. Hence, when integrated over the global domain the net pressure forces only arise at the domain boundaries. That is, the net pressure force acting on the full ocean domain arises only at the interface between the solid-earth and the ocean, plus the interface between the atmosphere and the ocean. Note that the pressure force has a component in both the vertical and horizontal directions as per the orientation of the local normal vector. Further boundary stresses arise from viscous exchange, which generally have components perpendicular to the boundary normal.

In the limit that the tetrahedron size goes to zero, local equilibrium of the contact forces means that

$$-\sum_{m=1}^3 \boldsymbol{\tau}_{(m)} dA_m + \boldsymbol{\tau}_{\hat{\mathbf{n}}} dA = 0. \quad (26.7)$$

In this equation, $\boldsymbol{\tau}_{(m)} dA_m$ (no implied summation) is the contact force vector acting on the face with outward normal parallel to the corresponding coordinate axis and $\boldsymbol{\tau}_{\hat{\mathbf{n}}} dA$ is the contact force acting on the slanted face with outward normal $\hat{\mathbf{n}}$. The minus sign arises for the summation term since the outward normals for these three faces point in the negative directions, and our convention is for $\boldsymbol{\tau}_{(m)}$ to align with the positive coordinate directions. The areas for each face are related to the slanted face area through

$$dA_m = \hat{\mathbf{n}}_m dA, \quad (26.8)$$

so that the local equilibrium relation (26.7) becomes

$$-\sum_{m=1}^3 \boldsymbol{\tau}_{(m)} \hat{\mathbf{n}}_m + \boldsymbol{\tau}_{\hat{\mathbf{n}}} = 0. \quad (26.9)$$

Now organize the stress components into a 3×3 matrix, \mathbb{T}_{mn} , measuring the $n'th$ component of the $m'th$ contact stress $\boldsymbol{\tau}_{(m)}$. That is, let \mathbb{T}_{mn} measure the force per area in the n -direction along a surface whose outward normal points in the m -direction, as in Figure 26.4. Making use of \mathbb{T}_{mn} in the expression (26.9) of local equilibrium leads to

$$\sum_{m=1}^3 \hat{\mathbf{n}}_m \mathbb{T}_{mn} = (\boldsymbol{\tau}_{\hat{\mathbf{n}}})_n. \quad (26.10)$$

This relation can be written as

$$\hat{\mathbf{n}} \cdot \mathbb{T} = \boldsymbol{\tau}_{\hat{\mathbf{n}}}, \quad (26.11)$$

so that the stress vector acting on a surface oriented according to a normal vector, $\hat{\mathbf{n}}$, equals to the projection of the stress tensor onto the normal vector.

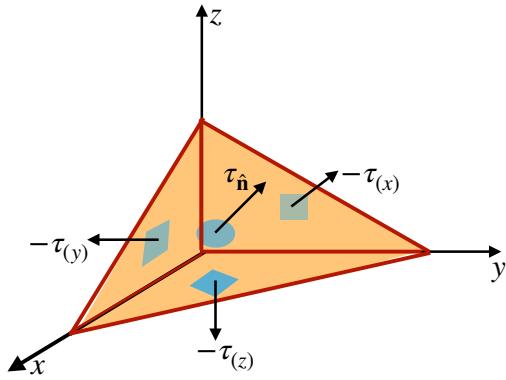


Figure 26.3: Tetrahedron region of fluid with contact stresses acting on the four faces. Note that the stresses are not necessarily directed normal to the faces. Local equilibrium of contact stresses means that the accumulation of these four stresses around the region adds to zero as the region volume goes to zero.

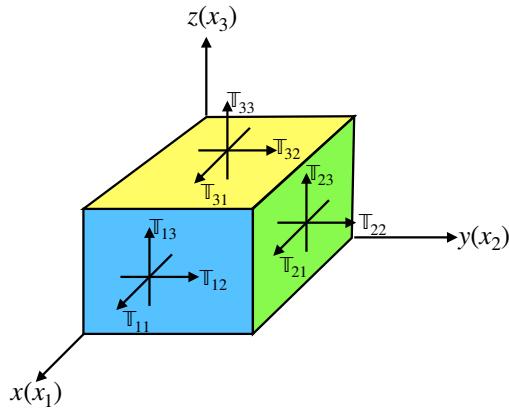


Figure 26.4: Illustrating the components to the stress tensor, \mathbb{T}_{mn} and how they are organized according to the coordinate axes. The component \mathbb{T}_{mn} is the stress that points in the n 'th direction along the face with outward normal in the m 'th direction.

How do we know that \mathbb{T}_{mn} are components to a tensor? The normal direction, \hat{n} , and the contact stress vector, $\tau_{\hat{n}}$, are components to first order tensors (i.e., vectors). The quotient rule from tensor analysis means that equation (26.10) yields stresses, \mathbb{T}_{mn} , that are components to a second order *stress tensor*. As components to a second order tensor, the \mathbb{T}_{mn} transform according to the rules developed in Chapter 3 for Cartesian tensors and Chapter 8 for general tensors.

26.3 Angular momentum and stress tensor symmetry

We made use of linear momentum conservation to deduce the local equilibrium property (26.4) of the contact stress. We here consider the constraint on the contact stresses imposed by angular momentum conservation. For this purpose, consider a constant mass fluid element that has a Cartesian position \mathbf{x} relative to an arbitrary origin. The angular momentum of the fluid element with respect to the origin is

$$\mathbf{L} = \rho \delta V (\mathbf{x} \wedge \mathbf{v}), \quad (26.12)$$

and its material time evolution is

$$\frac{D\mathbf{L}}{Dt} = \rho \delta V \mathbf{x} \wedge \frac{D\mathbf{v}}{Dt}, \quad (26.13)$$

which follows since $D(\rho \delta V)/Dt = 0$, $D\mathbf{x}/Dt = \mathbf{v}$, and $\mathbf{v} \wedge \mathbf{v} = 0$. Making use of Cauchy's form for the equation of motion (22.11)

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} + \nabla \cdot \mathbb{T} \quad (26.14)$$

allows us to write the angular momentum evolution as

$$\frac{D\mathbf{L}}{Dt} = \delta V \mathbf{x} \wedge (\rho \mathbf{f} + \nabla \cdot \mathbb{T}). \quad (26.15)$$

The first term arises from body forces (e.g., gravity and Coriolis) and the second term arises from contact stresses. Focusing on the contact stresses and expanding the divergence operation renders

$$\left[\frac{D\mathbf{L}_m}{Dt} \right]_{\text{contact stress}} = \delta V \epsilon_{mnp} x_n (\nabla \cdot \mathbb{T})_p \quad (26.16a)$$

$$= \delta V \epsilon_{mnp} x_n \partial_q \mathbb{T}_{qp} \quad (26.16b)$$

$$= \delta V \epsilon_{mnp} [\partial_q (x_n \mathbb{T}_{qp}) - (\partial_q x_n) \mathbb{T}_{qp}] \quad (26.16c)$$

$$= \delta V \epsilon_{mnp} [\partial_q (x_n \mathbb{T}_{qp}) - \mathbb{T}_{np}], \quad (26.16d)$$

where the final equality follows since $\partial_q x_n = \delta_{qn}$. The first term on the right hand side accounts for movement of torques within the fluid from one region to another. When integrating over a finite region, this term accounts for torques that arise from stresses acting on the region boundaries. For example, stresses at the boundary between the ocean and atmosphere create torques on both media that alter their respective angular momenta.

The second term in equation (26.16d) contributes to angular momentum throughout the fluid region. It is a source of torque that is *not* associated with body forces (gravity, Coriolis, centrifugal). What are these torque sources? Given that they do not arise from body forces, they must be associated with internal molecular degrees of freedom that introduce force couplets that render a torque acting at each point. However, geophysical fluids (air and water) do not spontaneously generate angular momentum at each point within the fluid region. They have no internally generated torques. Rather, the only means for an element of water or air to modify its angular momentum is through boundary torques, transfer of torques between fluid elements, and through torques due to body forces. So to remove internal torque sources we only consider symmetric stress tensors

$$\mathbb{T}_{mn} = \mathbb{T}_{nm} \implies \epsilon_{mnp} \mathbb{T}_{np} = 0. \quad (26.17)$$

This is a central property of the stresses acting on geophysical fluids.

26.4 Flux-form Eulerian momentum equation

We often find it useful to consider Cauchy's form of the momentum equation (26.14) in its flux-form Eulerian expression. Making use of Cartesian tensors, we expand the material time derivative acting on the velocity and introduce the mass conservation equation (17.9) so that

$$\rho \frac{D\mathbf{v}}{Dt} = \rho [\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v}] \quad (26.18a)$$

$$= \rho [\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v}] + \mathbf{v} (\partial_t \rho + \nabla \cdot (\rho \mathbf{v})) \quad (26.18b)$$

$$= \partial_t (\rho \mathbf{v}) + \nabla \cdot [\rho \mathbf{v} \otimes \mathbf{v}], \quad (26.18c)$$

where $\mathbf{v} \otimes \mathbf{v}$ is the outer product of the velocity vector with components

$$(\mathbf{v} \otimes \mathbf{v})_{mn} = v_m v_n. \quad (26.19)$$

Consequently, the momentum equation (26.14) takes on the flux-form Eulerian expression

$$\partial_t(\rho \mathbf{v}) + \nabla \cdot [\rho \mathbf{v} \otimes \mathbf{v}] = \rho \mathbf{f} + \nabla \cdot \mathbb{T}. \quad (26.20)$$

Alternatively, we can move the advection of momentum term onto the right hand side so that

$$\partial_t(\rho \mathbf{v}) = \rho \mathbf{f} + \nabla \cdot [\mathbb{T} - \rho \mathbf{v} \otimes \mathbf{v}], \quad (26.21)$$

which takes on the component form

$$\partial_t(\rho v_m) = \rho f_m + \partial_n [\mathbb{T}_{mn} - \rho v_m v_n]. \quad (26.22)$$

In this form we see that momentum advection can be interpreted as a stress that modifies the linear momentum per volume at a point in space. We refer to the stress,

$$\mathbb{T}_{mn}^{\text{kinetic}} = -\rho (\mathbf{v} \otimes \mathbf{v})_{mn} = -\rho v_m v_n, \quad (26.23)$$

as the *mechanical stress* or *kinetic stress*, which arises from the mechanical interactions between moving fluid elements in addition to that from pressure and viscosity. The turbulent contribution to the mechanical stress is known as the *Reynolds stress*.

26.5 Linear momentum for arbitrary regions

We now consider the budget of linear momentum for an arbitrary region, \mathcal{R} , moving in an arbitrary manner within the fluid. For this purpose we make use of the Leibniz-Reynolds Transport Theorem (18.30)

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \varphi dV \right] = \int_{\mathcal{R}} \frac{\partial \varphi}{\partial t} dV + \oint_{\partial \mathcal{R}} \varphi \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (26.24)$$

where $\mathbf{v}^{(b)}$ is the velocity of the region boundary, $\partial \mathcal{R}$, with $\hat{\mathbf{n}}$ the outward normal along the boundary. Applying this result to a component of the linear momentum per volume, $\varphi = \rho v_m$, and making use of the flux-form Eulerian momentum equation (26.22) leads to

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho v_m dV \right] = \int_{\mathcal{R}} \partial_t(\rho v_m) dV + \oint_{\partial \mathcal{R}} (\rho v_m) \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (26.25a)$$

$$= \int_{\mathcal{R}} [\rho f_m + \partial_n (\mathbb{T}_{mn} - \rho v_m v_n)] dV + \oint_{\partial \mathcal{R}} (\rho v_m) \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (26.25b)$$

$$= \int_{\mathcal{R}} \rho f_m dV + \oint_{\partial \mathcal{R}} (\mathbb{T}_{mn} - \rho v_m v_n) \hat{\mathbf{n}}_n d\mathcal{S} + \oint_{\partial \mathcal{R}} (\rho v_m) \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (26.25c)$$

$$= \int_{\mathcal{R}} \rho f_m dV + \oint_{\partial \mathcal{R}} [\mathbb{T}_{mn} + \rho v_m (v_n^{(b)} - v_n)] \hat{\mathbf{n}}_n d\mathcal{S}. \quad (26.25d)$$

We can write this relation in a geometric form as

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}} \rho \mathbf{f} dV + \oint_{\partial \mathcal{R}} [\mathbb{T} + \rho \mathbf{v} \otimes (\mathbf{v}^{(b)} - \mathbf{v})] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (26.26)$$

As a relation between geometric objects (vectors and tensors), the momentum budget (26.26) is independent of coordinate representation. We conclude that the evolution of linear momentum over an arbitrary region is affected by the volume integrated body force acting over the region, plus the impacts from contact stresses acting on the region boundary. Notably, the contact stresses have a contribution from the advection of linear momentum across the region boundary, with advection computed relative to motion of the boundary.

We refer to a *Lagrangian region* as one that moves with the barycentric velocity, $\mathbf{v}^{(b)} = \mathbf{v}$, in which case the mechanical stress is eliminated from the finite volume momentum budget (26.26). In fact, to eliminate the mechanical stress only requires the normal components of the boundary velocity to equal that of the fluid element, $(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0$. In either case we distinguish a Lagrangian region by writing $\mathcal{R}(\mathbf{v})$ to emphasize that the region moves with the barycentric fluid velocity, \mathbf{v} . For this case the linear momentum is only affected by body forces as well as contact stresses contained in the stress tensor

$$(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0 \implies \frac{d}{dt} \left[\int_{\mathcal{R}(\mathbf{v})} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}(\mathbf{v})} \rho \mathbf{f} dV + \oint_{\partial \mathcal{R}(\mathbf{v})} \mathbb{T} \cdot \hat{\mathbf{n}} dS. \quad (26.27)$$

This relation is Reynold's transport theorem (Section 18.3.5) as applied to linear momentum.

26.6 Relating stress to pressure and strain

Thus far we have offered a rather general treatment of stress, developing its properties according to the conservation of linear momentum and angular momentum. We now develop *constitutive relations*. In the present context, a constitutive relation connects stress acting between fluid elements to fluid properties and to the pressure and rate of strain that cause fluid elements to deform.

26.6.1 Hydrostatic pressure

Consider a fluid in which the stress on an area element is always normal to the area element and is independent of the orientation. This fluid is said to be in *hydrostatic balance* and the corresponding stress tensor and stress vector are written

$$\mathbb{T}_{mn} = -p \delta_{mn} \iff \mathbb{T} \cdot \hat{\mathbf{n}} = -p \hat{\mathbf{n}}, \quad (26.28)$$

where p is the hydrostatic pressure field. For a compressible fluid at rest, we can identify p with the thermodynamic pressure discussed in Chapter 23. If we furthermore assume that local thermodynamic equilibrium is maintained for fluid elements within a moving fluid, then we can continue making this identification. See Section 4.5 of [Kundu et al. \(2012\)](#) for more on the relation between mechanical pressure and thermodynamical pressure.

26.6.2 Friction tensor for a Newtonian fluid

A moving fluid has a more complex stress relation than a static fluid. To account for such complexity we introduce an additional contribution to the stress tensor

$$\mathbb{T}_{mn} = -p \delta_{mn} + \tau_{mn} \iff \mathbb{T} = -p \mathbb{I} + \tau. \quad (26.29)$$

We refer to τ as the *friction tensor*. In general, it captures the irreversible exchanges of momentum between moving fluid elements arising from the presence of viscosity. The friction tensor vanishes

in a perfect fluid where viscosity is assumed to be zero. It also vanishes when the fluid is static since there are no strains between fluid elements that can be transferred, through viscosity, into a stress. When it has a zero trace, $\tau_{mm} = 0$, then it is called the *deviatoric* stress tensor, which we further discuss below.

Mechanical equilibrium separately for pressure and friction?

As seen by equation (26.5), Newton's 3rd law says that the net contact stress is in mechanical equilibrium at each point within a fluid. Consequently, it is only the boundary integrated contact stress that contributes a contact force on a finite domain. When decomposing stress into contributions from pressure and friction, the net contact force is given by the boundary integral

$$\mathbf{F}^{\text{contact}} = \oint (-p \hat{\mathbf{n}} + \boldsymbol{\tau} \cdot \hat{\mathbf{n}}) d\mathcal{S}. \quad (26.30)$$

Are pressure and friction separately in mechanical equilibrium within the fluid interior? We have no deductive argument supporting the affirmative answer. Even so, we conjecture that mechanical equilibrium does hold separately for pressure and friction, if not exactly then very close for most purposes. If mechanical equilibrium is not separately maintained, then pressure forces (which are thermodynamically reversible) must be intimately aligned with friction forces (which are irreversible). Given that molecular friction is tiny in geophysical fluid interiors, it is plausible that pressure is indeed very close to mechanical equilibrium within the fluid interior. Furthermore, mechanical equilibrium holds for pressure when the vertical momentum equation is well approximated by the hydrostatic balance (Section 27.1).

General properties of the friction tensor

To help develop some of the general properties of the friction tensor, consider a fluid in uniform motion. Boosting the reference frame allows us to move to a reference frame where the fluid is static. Through Galilean invariance (Section 16.5) we expect the dynamics to remain unchanged. Since the friction tensor vanishes when the fluid is static, Galilean invariance implies that the friction tensor vanishes when the fluid undergoes uniform motion in any direction. We can extend this result to finite regions of the fluid where a group of fluid elements is moving coherently in a single direction. Moving to a local rest frame again renders a zero friction tensor when the fluid locally has uniform motion; i.e., zero gradients in the flow.

Galilean invariance thus leads us to consider a friction tensor that is a function of gradients in the velocity field, $\partial_m v_n$. Furthermore, as the stress tensor must be symmetric, the simplest expression for the friction tensor is one that is linearly proportional to the rate of strain tensor introduced in Section 20.2.4. Such fluids are known as *Newtonian fluids*. The linear stress-rate of strain relation takes the same mathematical form as *Hooke's Law* used in the study of elastic materials.

Friction tensor related to the rate of strain tensor

The simplest form of this relation is when there is a single viscosity so that the stress-strain relation takes the form

$$\tau_{mn} = \rho \mu (\partial_m v_n + \partial_n v_m) = 2 \rho \mu \mathbb{S}_{mn}, \quad (26.31)$$

with $\mu > 0$ a kinematic viscosity (dimensions of squared length per time) (see Section 37.2.5) and \mathbb{S}_{mn} components to the rate of strain tensor introduced in Section 20.2.6. Note that with this form

of the friction tensor the trace of the full stress tensor takes the form

$$\mathbb{T}_{mm} = -p + 2\rho\mu\nabla \cdot \mathbf{v}. \quad (26.32)$$

The trace deviates from pressure in the presence of velocity divergence. However, for an incompressible fluid in which $\nabla \cdot \mathbf{v} = 0$ (Chapter 19), the trace equals to minus the pressure, in which case τ_{mn} is known as the *deviatoric* stress tensor.¹

The viscous force per volume is given by the divergence of the stress tensor

$$\rho F_m = 2\partial_n(\rho\mu\mathbb{S}_{mn}). \quad (26.33)$$

Assuming the dynamic viscosity, $\rho\mu$, is constant leads to

$$\rho \mathbf{F} = \rho\mu[\nabla(\nabla \cdot \mathbf{v}) + \nabla^2 \mathbf{v}] \quad \text{if } \rho\mu \text{ is constant.} \quad (26.34)$$

Furthermore, if the flow is incompressible then

$$\rho \mathbf{F} = \rho\mu\nabla^2 \mathbf{v} \quad \text{if } \rho\mu \text{ is constant and } \nabla \cdot \mathbf{v} = 0. \quad (26.35)$$

Frictional stress for a vertically sheared zonal flow

Consider the case of an incompressible flow field that only has a zonal component with a vertical structure (Figure 26.5)

$$\mathbf{v} = u(z)\hat{\mathbf{x}}. \quad (26.36)$$

In this case the only non-zero components to the rate of strain tensor are due to the vertical shear, $\mathbb{S}_{13} = \mathbb{S}_{31} = \partial_z u/2$. Now consider a horizontal area whose outward normal is parallel to the $\hat{\mathbf{z}}$ direction. The frictional force acting on that area is given by the area integral of the frictional stress

$$\mathbf{F}_{\text{area}} = \int \tau \cdot \hat{\mathbf{n}} d\mathcal{S} = \int \tau \cdot \hat{\mathbf{z}} dx dy = \frac{\hat{\mathbf{x}}}{2} \rho\mu A \frac{\partial u}{\partial z}, \quad (26.37)$$

where $A = \int dx dy$ is the horizontal area. Hence, the zonal stress arises from the nonzero vertical shear.

Momentum is deposited in regions where there is a divergence in the stress, in which case momentum is transferred from regions of high vertical shear to low vertical shear. At a point, the momentum is affected by the divergence of the viscous stress at that point. For $\mathbf{v} = u(z)\hat{\mathbf{x}}$ we have

$$\left[\frac{\partial(\rho v_m)}{\partial t} \right]_{\text{viscous}} = \partial_n \tau_{nm} \implies \left[\frac{\partial(\rho u)}{\partial t} \right]_{\text{viscous}} = \rho\mu \frac{\partial^2 u}{\partial z^2}, \quad (26.38)$$

so that zonal momentum is preferentially deposited to or removed from regions with high vertical curvature in the zonal velocity.

26.6.3 Kinetic energy dissipation and Joule heating

In Section 24.1 we developed the budget for kinetic energy per mass of a fluid element (equation (24.2))

$$\rho \frac{D\mathcal{R}}{Dt} = -\mathbf{v} \cdot \nabla p - w g \rho + \rho \mathbf{v} \cdot \mathbf{F}. \quad (26.39)$$

¹A 2nd order tensor, \mathbb{A} , has a *deviator* with components given by $\mathbb{A}'_{mn} = \mathbb{A}_{mn} - (1/3) \delta_{mn} \mathbb{A}_{mm}$.

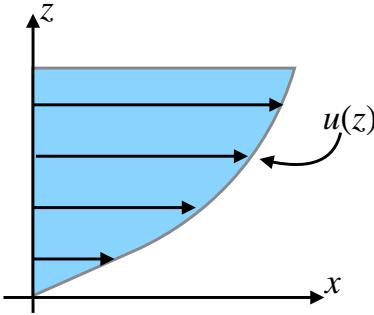


Figure 26.5: Sample profile of zonal velocity possessing a vertical shear: $\mathbf{v} = u(z) \hat{\mathbf{x}}$. The resulting zonal frictional stress arises from the nonzero vertical shear and viscosity.

We here examine the role of the friction term

$$\text{friction power per volume} = \rho \mathbf{v} \cdot \mathbf{F}, \quad (26.40)$$

with this term interpreted as a power (energy per time) per volume acting to alter the kinetic energy per volume of a fluid element.

To get started, recall from Section 24.1.1 that for the special case of Rayleigh drag,

$$\mathbf{F} = \mathbf{F}_{\text{Rayleigh}} = -\gamma \mathbf{v}, \quad (26.41)$$

the kinetic energy is dissipated locally at each point

$$\rho \mathbf{v} \cdot \mathbf{F}_{\text{Rayleigh}} = -\gamma \rho \mathbf{v} \cdot \mathbf{v} \leq 0. \quad (26.42)$$

What happens with a more general stress-rate of strain relation (26.31) applicable to a Newtonian fluid? To answer this question, expose Cartesian tensor labels to render

$$\rho \mathbf{v} \cdot \mathbf{F} = v_m \rho F_m \quad (26.43a)$$

$$= 2 v_m \partial_n (\rho \mu S_{mn}) \quad (26.43b)$$

$$= 2 \partial_n (v_m \rho \mu S_{mn}) - 2 \partial_n v_m \rho \mu S_{mn} \quad (26.43c)$$

$$= 2 \nabla \cdot (\rho \mu \mathbf{v} \cdot \mathbb{S}) - 2 \rho \mu \mathbb{S} \cdot \mathbb{S}, \quad (26.43d)$$

where

$$\mathbb{S} \cdot \mathbb{S} = S_{mn} S_{mn} \geq 0 \quad (26.44)$$

is the contraction of the strain tensor with itself. We interpret the two contributions in equation (26.43d) as

$$\rho \mathbf{v} \cdot \mathbf{F} = \text{viscous flux divergence} - \text{kinetic energy dissipation}. \quad (26.45)$$

When integrated over the full domain, the flux divergence reduces to a contribution of boundary stresses, which are typically quite turbulent. In contrast, the dissipation term provides a sign-definite sink to the kinetic energy at each point in the fluid interior, with this term commonly written

$$\epsilon \equiv 2 \mu \mathbb{S} \cdot \mathbb{S} \geq 0 \quad \text{mechanical energy dissipation}. \quad (26.46)$$

The dimensions of ϵ are $L^2 T^{-3}$, which in SI units are $m^2 s^{-3} = W \text{ kg}^{-1}$. We thus commonly refer to ϵ as the mechanical energy dissipation per mass due to viscous effects; i.e., the dissipated mechanical power per mass.

Making use of equation (26.43d) in the total energy budget (24.16)

$$\frac{\partial(\rho \mathcal{E})}{\partial t} + \nabla \cdot [\rho \mathbf{v} (\mathcal{E} + p/\rho)] = \rho (\mathbf{v} \cdot \mathbf{F} + \dot{Q}). \quad (26.47)$$

leads to

$$\frac{\partial(\rho \mathcal{E})}{\partial t} + \nabla \cdot [\rho \mathbf{v} (\mathcal{E} + p/\rho) + \rho \mu \mathbf{v} \cdot \mathbb{S}] = \rho (-\epsilon + \dot{Q}). \quad (26.48)$$

When integrated over the full domain, total energy conservation implies that in the absence of external sources of heating, the integral of the right hand side of the total energy budget vanishes. For that to hold locally, we see that as mechanical energy is dissipated by friction, the internal energy is increased through *Joule heating*²

$$\dot{Q}_{\text{Joule}} \equiv \epsilon = 2 \mu \mathbb{S} \cdot \mathbb{S}. \quad (26.49)$$

The Joule heating by molecular viscosity is larger in regions where the fluid strains are larger, signalling a more efficient transfer of power to the microscales where molecular viscosity can act on the flow. In the ocean interior, measurements indicate that $\epsilon \approx 10^{-9} \text{ W kg}^{-1}$. Dividing by $c_p = 3900 \text{ J kg}^{-1} \text{ K}^{-1}$ leads to a heating rate of less than $10^{-3} \text{ K century}^{-1}$, which is a very small rate of ocean heating. Consequently, ocean Joule heating has a negligible role in the ocean heat budget and as such is generally ignored.³

26.6.4 Comments and further study

There are far more elaborate stress-strain relations than that from molecular viscosity in equation (26.31). The most general form for a Newtonian fluid introduces a fourth-order viscosity tensor as in Section 4.5 of [Kundu et al. \(2012\)](#) and Chapter 17 in [Griffies \(2004\)](#). Nonlinear stress-strain relations can arise in Newtonian fluids when the viscous tensor is a function of the flow, such as with the Smagorinsky scheme commonly used for Large Eddy Simulations (LES) ([Smagorinsky, 1993](#)). Further nonlinearities arise in non-Newtonian fluids, which make use of products of the strain tensor for computing stress.

The physics of mechanical energy dissipation forms a central role in the study of ocean mixing and its measurement using instruments. [MacKinnon et al. \(2013\)](#) is a useful place to begin studying this vast and active field of physical oceanography.

Geophysical fluids such as air and water are generally well treated using Newtonian constitutive relations. However, there are some geophysical turbulence theories that propose a non-Newtonian constitutive relation for part of their closures. [Anstey and Zanna \(2017\)](#) offer a compelling approach with a subgrid scale stress tensor containing a non-zero trace, thus resulting in a modification to pressure.

26.7 The Reynolds number and turbulence

How important is friction relative to other terms in the momentum equation? In particular, how does it compare to the material acceleration? When determining friction as per the Laplacian

²Joule heating commonly refers to the process whereby an electric current flowing through a resistor generates heat. It is used here to denote the heat generated by mechanical dissipation of kinetic energy through viscous effects.

³SMG: discuss the atmospheric Joule heating and why it is not negligible. Provide some numbers.

operator in equation (26.35), the ratio for the scales of material acceleration to frictional acceleration is called the Reynolds number

$$Re = \frac{\text{inertial accelerations}}{\text{frictional accelerations}} = \frac{U/T}{\mu U/L^2} = \frac{L^2/T}{\mu} = \frac{LU}{\mu}, \quad (26.50)$$

where U is a typical velocity scale, L is a typical length scale, and we assumed that $U = L/T$ so that the time scale is determined by advection.

The Reynolds number is dependent on the chosen velocity and length scales. For large-scale oceanic flows, such as Gulf Stream rings (see Figure 29.1) with $L \approx 10^5$ m, $U \approx 0.1 - 1.0$ m s⁻¹ and (page 75 of [Gill \(1982\)](#))

$$\mu_{\text{water}} = 10^{-6} \text{ m}^2 \text{ s}^{-1}, \quad (26.51)$$

then

$$Re_{\text{Gulf Stream}} = 10^{10} - 10^{11}. \quad (26.52)$$

For the atmosphere, we take $L = 10^6$ m for a typical atmospheric weather system, $U = 10$ m s⁻¹ for the speed, and

$$\mu_{\text{air}} = 1.4 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}, \quad (26.53)$$

for the kinematic viscosity of air at standard pressure (page 75 of [Gill \(1982\)](#)). Given the larger length and velocity scales, the Reynolds number for large-scale atmospheric circulation features is

$$Re_{\text{weather system}} = 10^{12}. \quad (26.54)$$

These values for the Reynolds number are huge relative to typical values found in engineering flows. They signal the minor role that molecular friction plays in large-scale geophysical fluid flows. Even so, molecular friction is the process leading to mechanical energy dissipation at the small scales. A fundamental feature of large Reynolds number flow is the presence of turbulent motions. Turbulent flows are highly nonlinear and affect a transfer of energy across length and time scales. This cascade leads to the dissipation of mechanical energy at the small scales. It is at the small scales that flow curvature can be large enough for the relatively tiny values of molecular viscosity to dissipate the energy, thus preventing an ultraviolet catastrophe; i.e., preventing the unbounded pile up of mechanical energy at the smallest scales.

The ocean and atmosphere exhibit a huge variety of turbulent regimes, from the macroturbulence of geostrophic eddies to the microturbulence of boundary layers. Turbulence is not directly considered in this book. However, certain of its implications are identified in various places given that it is so basic to the ocean and atmosphere flows. [Vallis \(2017\)](#) offers a pedagogical entry point for the physics and maths of geophysical turbulence.

26.8 Form stress

As noted in Figure 26.1, form stress refers to the horizontal stress arising from pressure acting on a sloped surface. To expose the mathematics of form stress, consider an arbitrary surface, \mathcal{S} , as in Figure 26.6 whose depth is written

$$z = \eta(x, y, t), \quad (26.55)$$

with this expression available so long as the surface is nowhere vertical.⁴ The outward normal pointing away from the top side of the surface is given by

$$\hat{\mathbf{n}}_{\text{top}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla\eta}{\sqrt{1 + |\nabla\eta|^2}}. \quad (26.56)$$

Multiplying the pressure times the horizontal area element on the surface, dS , leads to the net pressure force acting at a point on the top side of the surface

$$\mathbf{F}^{\text{press}} = -p \hat{\mathbf{n}}_{\text{top}} dS = -p (\hat{\mathbf{z}} - \nabla\eta) dA, \quad (26.57)$$

where we used the identity⁵

$$dS = |\nabla(z - \eta)| dA, \quad (26.58)$$

with

$$dA = dx dy \quad (26.59)$$

the horizontal projection of the surface area element (see Figure 26.6). We identify the form stress acting on the top side of this surface as

$$\Sigma^{\text{form}} \equiv p \nabla\eta. \quad (26.60)$$

The name follows since the stress is determined by the “form” of the surface as measured by its slope, $\nabla\eta$. It is larger in magnitude when the slope increases. We can thus write the pressure force acting on the top side of the surface as

$$\mathbf{F}^{\text{press}} = (-p \hat{\mathbf{z}} + \Sigma^{\text{form}}) dA. \quad (26.61)$$

Local mechanical equilibrium of pressure forces within a fluid implies that the pressure acting on the bottom side is equal in magnitude but oppositely directed (see Section 26.6.2).

We further our understanding of form stress when discussing the regional angular momentum budget in Section 26.10, geostrophic motions in Section 29.5, and for the shallow water model in Section 43.2.

26.9 Boundary conditions

Throughout this chapter we have considered domain boundaries within a single fluid body. Here we formalize the resulting boundary conditions and extend the considerations to boundaries between a liquid and a gas or between a fluid and a rigid boundary. We ignore the effects from surface tension discussed in Chapter 35. We do so since we are interested in length scales larger than the capillary-gravity waves where surface tension is important, $\lambda_{\text{cap-grav}} \approx 17$ cm (equation (35.16)).

⁴For more on the maths of such surfaces, see the geometry discussion in Chapter 6. See also the discussion of generalized vertical coordinates in Chapters 11 and 21.

⁵The identity (26.58) follows from trigonometry as summarized in Figure 26.6, and further detailed in the discussion of kinematic boundary conditions in Section 17.4 and the analogous discussion of dia-surface transport in Section 21.3.

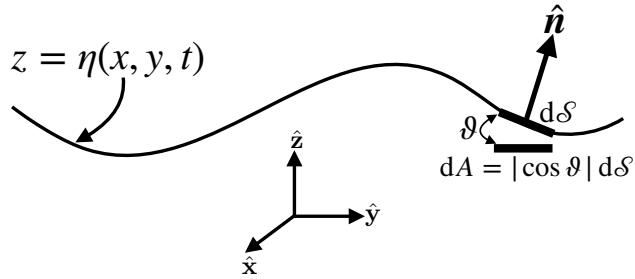


Figure 26.6: Form stress is the horizontal pressure force per area acting on a sloped surface. The area integral of the pressure leads to the net pressure force acting on the surface, $\int \mathbf{F}^{\text{press}} = -\int p \hat{\mathbf{n}} d\mathcal{S}$, with the horizontal component of this force arising from the slope, or the “form”, of the surface. The area element on the surface, $d\mathcal{S}$, has a horizontal projection given by $dA = \cos \vartheta d\mathcal{S}$, with the angle assumed to be within the range $-\pi/2 < \vartheta < \pi/2$ so that the surface is nowhere vertical.

26.9.1 Formulation

Formulation of the stress boundary conditions follows from applying the momentum equation (26.26) to a pillbox region such as that shown for the air-sea interface in Figure 26.7. The sides of the box are fixed in space with thickness h , whereas the top and bottom have an area $\delta\mathcal{S}$ that straddles the moving interface. In the limit that the pillbox thickness goes to zero, the volume integrals in equation (26.26) vanish under the assumption of a smooth velocity field on both sides of the interface as well as smooth body forces. We are thus left with the constraint that the area integrated contact forces must vanish when integrated around the pillbox boundary

$$\oint_{\partial\text{pillbox}} [\mathbb{T} + \rho \mathbf{v} \otimes (\mathbf{v}^{(b)} - \mathbf{v})] \cdot \hat{\mathbf{n}} d\mathcal{S} = 0. \quad (26.62)$$

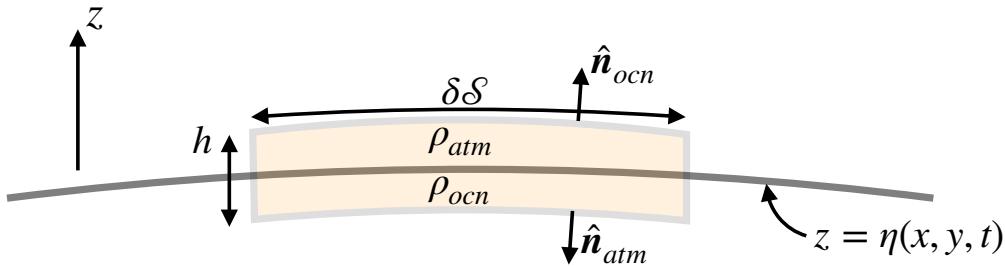


Figure 26.7: Infinitesimal pillbox region used in formulating the air-sea stress boundary condition, with fluid-1 the atmosphere and fluid-2 the ocean. The pillbox straddles the moving air-sea interface with a thickness h and top and bottom faces having an area $\delta\mathcal{S}$. We orient the fluid media through the outward normals according to $\hat{\mathbf{n}} = \hat{\mathbf{n}}_{ocn} = -\hat{\mathbf{n}}_{atm}$. That is, the outward normal for the ocean points into the atmosphere, whereas the outward normal for the atmosphere points into the ocean.

26.9.2 Matter transport boundary condition

In Section 17.4.3 we formulated the kinematic boundary condition at the air-sea interface. That discussion focused on the ocean side of the interface, thus leading to the definition of the mass transport given by equation (17.61). Identical considerations for the atmosphere side of the interface

lead to

$$\rho_{\text{ocn}} (\mathbf{v}^{(b)} - \mathbf{v}_{\text{ocn}}) \cdot \hat{\mathbf{n}}_{\text{ocn}} = \mathcal{Q}_m \quad (26.63a)$$

$$\rho_{\text{atm}} (\mathbf{v}^{(b)} - \mathbf{v}_{\text{atm}}) \cdot \hat{\mathbf{n}}_{\text{atm}} = -\mathcal{Q}_m, \quad (26.63b)$$

where \mathcal{Q}_m measures the matter transport crossing the air-sea interface. The sign is chosen so that $\mathcal{Q}_m > 0$ measures matter entering the ocean and thus leaving the atmosphere or land (see Section 17.4.3). Setting

$$\hat{\mathbf{n}} = \hat{\mathbf{n}}_{\text{ocn}} = -\hat{\mathbf{n}}_{\text{atm}} \quad (26.64)$$

for the outward normals shown in Figure 26.7 then leads to

$$\rho_{\text{ocn}} (\mathbf{v}^{(b)} - \mathbf{v}_{\text{ocn}}) \cdot \hat{\mathbf{n}} = \rho_{\text{atm}} (\mathbf{v}^{(b)} - \mathbf{v}_{\text{atm}}) \cdot \hat{\mathbf{n}} = \mathcal{Q}_m. \quad (26.65)$$

26.9.3 Momentum transport arising from matter transport

Following the above discussion of boundary matter transport, we see that the kinetic term $\rho \mathbf{v} \otimes (\mathbf{v}^{(b)} - \mathbf{v})$ gives rise to the net momentum transport across the interface

$$\oint_{\partial\text{pillbox}} \rho \mathbf{v} \otimes (\mathbf{v}^{(b)} - \mathbf{v}) \cdot \hat{\mathbf{n}} d\mathcal{S} = (\mathbf{v}_{\text{ocn}} - \mathbf{v}_{\text{atm}}) \mathcal{Q}_m d\mathcal{S}. \quad (26.66)$$

Since the velocity of the ocean and atmosphere are generally distinct, there is a net transfer of momentum between the fluid media due to the transfer of matter.

26.9.4 Net boundary condition

Following equation (26.29), we decompose the stress tensor into a contribution from pressure and from friction

$$\mathbb{T}_{mn} = -p \delta_{mn} + \tau_{mn}, \quad (26.67)$$

in which case the infinitesimal version of the pillbox boundary condition (26.62) takes the form

$$(\mathbf{v}_{\text{ocn}} - \mathbf{v}_{\text{atm}}) \mathcal{Q}_m + (-p_{\text{ocn}} + p_{\text{atm}}) \hat{\mathbf{n}} + (\tau_{\text{ocn}} - \tau_{\text{atm}}) \cdot \hat{\mathbf{n}} = 0, \quad (26.68)$$

where the vertical sides of the pillbox do not contribute since we are ignoring surface tension. Rearrangement leads to

$$(\mathbf{v}_{\text{ocn}} - \mathbf{v}_{\text{atm}}) \mathcal{Q}_m = (p_{\text{ocn}} - p_{\text{atm}}) \hat{\mathbf{n}} - (\tau_{\text{ocn}} - \tau_{\text{atm}}) \cdot \hat{\mathbf{n}}. \quad (26.69)$$

The absence of matter transport leads to the simpler condition

$$-p_{\text{ocn}} \hat{\mathbf{n}} + \tau_{\text{ocn}} \cdot \hat{\mathbf{n}} = -p_{\text{atm}} \hat{\mathbf{n}} + \tau_{\text{atm}} \cdot \hat{\mathbf{n}} \quad \text{if } \mathcal{Q}_m = 0, \quad (26.70)$$

which is an expression of Newton's 3rd law (action/reaction).

Although we focused on the air-sea interface, these boundary conditions also hold for interfaces within a single fluid, such as for isopycnal layers within the ocean or isentropic layers in the atmosphere. In particular, within the interior of geophysical fluids, the effects from molecular viscosity are generally tiny relative to pressure. In that case the boundary condition is satisfied by matching pressure across interior layers (labelled A and B)

$$p_A = p_B. \quad (26.71)$$

In Section 26.6.2, we offer further comment on the question of mechanical equilibrium for pressure and friction.

26.9.5 No-slip boundary condition

At solid boundaries, the kinematic boundary condition from Section 17.4.1 sets the normal component of the velocity to zero

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{no-flux condition on static material boundary.} \quad (26.72)$$

However, nowhere in the present discussion have we specified the tangential component of the velocity along a solid boundary. What is it?

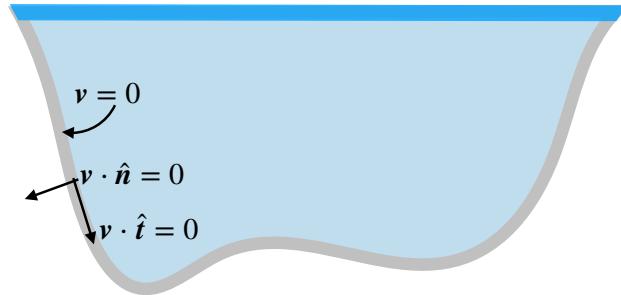


Figure 26.8: The no-slip boundary condition means that fluid exhibits zero relative motion at the solid-fluid boundary. That is, the fluid sticks to the solid boundary. The no-slip condition means that both the normal and tangential components of the velocity vanish at solid boundaries.

Careful laboratory experiments over the 19th and 20th centuries indicate that there is no relative motion of molecules at solid-fluid interfaces. That is, a fluid at the fluid-solid interface has a velocity matching that of the solid so that the fluid sticks to the solid boundary as depicted in Figure 26.8. The no-slip boundary condition means that both the normal and tangential components of the fluid velocity vanish next to static solid boundaries

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{t}} = 0 \quad \text{no-slip condition on static solid boundaries.} \quad (26.73)$$

The no-slip boundary condition gives rise to an exchange of momentum between the solid and fluid, with this exchange mediated by friction. Hence, in the absence of friction, as per an inviscid fluid, the no-slip boundary condition cannot be imposed since doing so would mathematically overspecify the flow. Consequently, for inviscid fluids the tangential component of the velocity remains unspecified at solid boundaries.

26.10 Regional ocean angular momentum budget

We here develop the depth and zonally integrated budget for axial angular momentum over a finite region of the ocean, such as shown in Figure 26.9. The analysis exposes how a variety of physical processes affect torques that change the angular momentum of the ocean fluid, in particular those processes associated with form stress and boundary friction. The presentation requires manipulations that are typical of those encountered with finite regional budgets, and as such it offers useful technical points beyond the issues specific to the study of angular momentum.

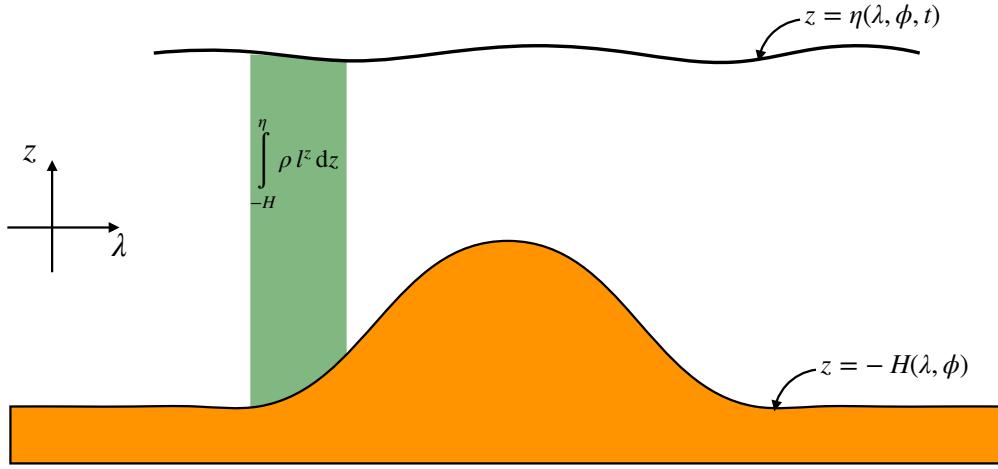


Figure 26.9: Schematic of the depth integrated angular momentum, here depicted moving in an ocean with a topographic bump. In Section 26.10 we develop the budget for the depth and zonal integrated angular momentum in the ocean, where we see that the angular momentum is affected by a variety of boundary processes as well as interior transports and pressures. Under certain circumstances, such as for the large-scale Southern Ocean circulation, the steady state balance becomes quite simple, as indicated in Figure 26.10.

26.10.1 Depth integrated angular momentum budget

The axial angular momentum budget for a fluid element follows that developed in Section 22.4, here written with zonal friction

$$\rho \frac{Dl^z}{Dt} = -\frac{\partial p}{\partial \lambda} + r_{\perp} \rho F^{\lambda}, \quad (26.74)$$

where

$$l^z = r_{\perp} (u + r_{\perp} \Omega) \quad (26.75)$$

is the angular momentum per unit mass, and $r_{\perp} = r \cos \phi$ is the distance to the polar rotation axis. Use of the Eulerian form of mass conservation (equation (17.9)) leads to the Eulerian flux-form expression

$$\frac{\partial(\rho l^z)}{\partial t} + \nabla \cdot (\rho \mathbf{v} l^z) = -\frac{\partial p}{\partial \lambda} + r_{\perp} \rho F^{\lambda}, \quad (26.76)$$

with $l^z \rho dz$ the angular momentum per unit horizontal area. Vertically integrating this budget over a column of ocean fluid renders a budget for the column-integrated angular momentum. This derivation requires Leibniz's Rule (e.g., Section 17.5) to reach the following identities that expose boundary contributions

$$\int_{-H}^{\eta} \frac{\partial(\rho l^z)}{\partial t} dz = \frac{\partial}{\partial t} \left[\int_{-H}^{\eta} \rho l^z dz \right] - \left[\rho l^z \frac{\partial \eta}{\partial t} \right]_{z=\eta} \quad (26.77a)$$

$$\int_{-H}^{\eta} \nabla_z \cdot (\rho \mathbf{u} l^z) dz = \nabla_z \cdot \left[\int_{-H}^{\eta} \rho \mathbf{u} l^z dz \right] - [\rho l^z \mathbf{u} \cdot \nabla \eta]_{z=\eta} - [\rho l^z \mathbf{u} \cdot \nabla H]_{z=-H} \quad (26.77b)$$

$$\int_{-H}^{\eta} \frac{\partial(\rho w l^z)}{\partial z} dz = [w \rho l^z]_{z=\eta} - [w \rho l^z]_{z=-H}. \quad (26.77c)$$

The surface kinematic boundary condition (17.74) and bottom kinematic boundary condition (17.38) lead to the more tidy form

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} l^z \rho dz \right] + \nabla_z \cdot \left[\int_{-H}^{\eta} l^z \mathbf{u} \rho dz \right] = [l^z Q_m]_{z=\eta} + \int_{-H}^{\eta} \left[-\frac{\partial p}{\partial \lambda} + r_{\perp} \rho F^{\lambda} \right] dz. \quad (26.78)$$

The budget (26.78) says that the depth integrated angular momentum per unit horizontal area has a time tendency arising from the convergence of horizontal advection of angular momentum plus torques due to surface boundary mass fluxes, depth integrated zonal pressure gradients, and depth integrated irreversible stresses.

26.10.2 Surface and bottom form stresses

We can further unpack the contribution from pressure in the budget (26.78) by making use of Leibniz's rule to write

$$\int_{-H}^{\eta} \frac{\partial p}{\partial \lambda} dz = \frac{\partial P}{\partial \lambda} - p_a \frac{\partial \eta}{\partial \lambda} - p_b \frac{\partial H}{\partial \lambda}, \quad (26.79a)$$

where

$$P = \int_{-H}^{\eta} p dz \quad (26.80)$$

is the depth-integrated pressure field, p_a is the pressure applied to the ocean at its surface, $z = \eta$, and p_b is the pressure at the ocean bottom, $z = -H$. The pressure contribution is thus decomposed into three terms, the first of which is the contribution from zonal pressure differences integrated across the depth of the column, whereas the second and third contributions arise from form stresses acting at the surface and bottom.

26.10.3 Surface and bottom frictional stresses

Now extract that portion of the friction given by the vertical transfer of zonal momentum according to

$$\rho F^\lambda = \frac{\partial \tau^\lambda}{\partial z}, \quad (26.81)$$

where τ^λ is the zonal component to the stress vector. As seen in Section 39.2.4, this vertical transfer in the ocean interior is dominated by the exchange of form stress associated with baroclinic eddies. In contrast, at the ocean surface and bottom the stress arises from turbulent motions that transfer momentum across the boundaries. Assume that the moment-arm is well approximated by its value at the ocean surface

$$r_\perp = r \cos \phi = (z + R) \cos \phi \approx R \cos \phi = R_\perp, \quad (26.82)$$

with this assumption holding for the shallow fluid approximation built into the hydrostatic primitive equations discussed in Section 27.1. This assumption allows us to write the frictional contribution to the angular momentum budget (26.78) in the form

$$\int_{-H}^{\eta} r_\perp \rho F^\lambda dz \approx R_\perp \int_{-H}^{\eta} \rho F^\lambda dz = R_\perp (\tau_a^\lambda - \tau_b^\lambda). \quad (26.83)$$

The final expression introduced τ_a^λ , which is the zonal component to the stress acting on the ocean surface imparted through irreversible interactions between the ocean and the overlying atmosphere and/or ice. The signs are such that $\tau_a^\lambda > 0$ transfers an eastward momentum to the ocean (e.g., from westerly winds). Likewise, the stress τ_b^λ is the zonal stress at the ocean bottom imparted through irreversible interactions between the ocean and the solid-earth. The signs are such that $\tau_b^\lambda > 0$ reflects the transfer of eastward momentum from the ocean to the solid-earth, or conversely the transfer of westward momentum from the earth to the ocean. The net contribution from vertical friction is thus given by the moment arm multiplied by the difference in boundary stresses. In the unusual case where the eastward surface stress equals to the westward bottom stress, then these boundary terms have no net impact on axial angular momentum.

26.10.4 Summary budget of angular momentum for a fluid column

Bringing all the pieces together leads to the depth integrated axial angular momentum budget

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} l^z \rho dz \right] + \nabla_z \cdot \left[\int_{-H}^{\eta} l^z \mathbf{u} \rho dz \right] = -\frac{\partial P}{\partial \lambda} + [l^z Q_m]_{z=\eta} + p_a \frac{\partial \eta}{\partial \lambda} + p_b \frac{\partial H}{\partial \lambda} + R_{\perp} (\tau_a^{\lambda} - \tau_b^{\lambda}). \quad (26.84)$$

Other than assuming the approximate form (26.82) for the moment-arm, this result is an exact budget for the axial angular momentum of a column of ocean fluid.

26.10.5 Steady state balances

Steady state balances are of particular interest when studying the large-scale low frequency behavior of the circulation. A steady state holds for the angular momentum so long as following balance is maintained

$$\nabla_z \cdot \left[\int_{-H}^{\eta} l^z \mathbf{u} \rho dz \right] = -\frac{\partial P}{\partial \lambda} + [l^z Q_m]_{z=\eta} + p_a \frac{\partial \eta}{\partial \lambda} + p_b \frac{\partial H}{\partial \lambda} + R_{\perp} (\tau_a^{\lambda} - \tau_b^{\lambda}). \quad (26.85)$$

Consequently, the horizontal divergence of depth integrated angular momentum advection (left hand side) is balanced by torques created by the variety of physical processes on the right hand side.

Zonal and depth integrated steady angular momentum budget

To further refine the balance we integrate zonally around a periodic region or a region with zonal boundaries, in which case the depth and zonally integrated steady balance results

$$\frac{1}{R} \int d\lambda \frac{\partial}{\partial \phi} \left[\int_{-H}^{\eta} l^z v \rho dz \right] = \int d\lambda \left[[l^z Q_m]_{z=\eta} + p_a \frac{\partial \eta}{\partial \lambda} + p_b \frac{\partial H}{\partial \lambda} + R_{\perp} (\tau_a^{\lambda} - \tau_b^{\lambda}) \right]. \quad (26.86)$$

In this case the meridional divergence of the advective transport of angular momentum is balanced, in the steady state, only through the boundary terms on the right hand side. To further expose the physics of the form stresses, we find it useful to introduce the anomalous zonal pressures

$$p'_a = p_a - \bar{p}_a \quad \text{with} \quad \bar{p}_a = \frac{1}{L_{\lambda}} \int p_a d\lambda \quad (26.87a)$$

$$p'_b = p_b - \bar{p}_b \quad \text{with} \quad \bar{p}_b = \frac{1}{L_{\lambda}} \int p_b d\lambda \quad (26.87b)$$

where L_{λ} is the zonal length of the domain. The angular momentum balance (26.86) thus takes the form

$$\frac{1}{R} \int d\lambda \frac{\partial}{\partial \phi} \left[\int_{-H}^{\eta} l^z v \rho dz \right] = \int d\lambda \left[[l^z Q_m]_{z=\eta} + p'_a \frac{\partial \eta}{\partial \lambda} + p'_b \frac{\partial H}{\partial \lambda} + R_{\perp} (\tau_a^{\lambda} - \tau_b^{\lambda}) \right], \quad (26.88)$$

thus revealing that the zonal mean surface and bottom pressures have no impact on the zonally integrated balance. This balance embodies a number of physical processes whose importance depends on the particular region under consideration. In the following we consider some features of the Southern Ocean.

Wind stress balanced by bottom form stress

Under certain cases the primary balance in equation (26.86) is between the form stress and friction stress, whereby

$$\int \left[p'_a \frac{\partial \eta}{\partial \lambda} + R_{\perp} \tau_a^{\lambda} \right] d\lambda \approx \int \left[-p'_b \frac{\partial H}{\partial \lambda} + R_{\perp} \tau_b^{\lambda} \right] d\lambda. \quad (26.89)$$

For much of the large-scale Southern Ocean circulation, the primary balance is even simpler: it is a balance between surface wind stress and bottom form stress

$$\int \tau_a^{\lambda} d\lambda \approx - \int p'_b \frac{1}{R_{\perp}} \frac{\partial H}{\partial \lambda} d\lambda, \quad (26.90)$$

with this balance exemplified in Figure (26.10). For the balance (26.90) to hold requires the upwind side of a topographic bump to generally realize an anomalously high bottom pressure with the opposite on the downwind side. This balance in turn means there is an anomalously high mass on the upwind side and low mass anomaly on the downwind side.

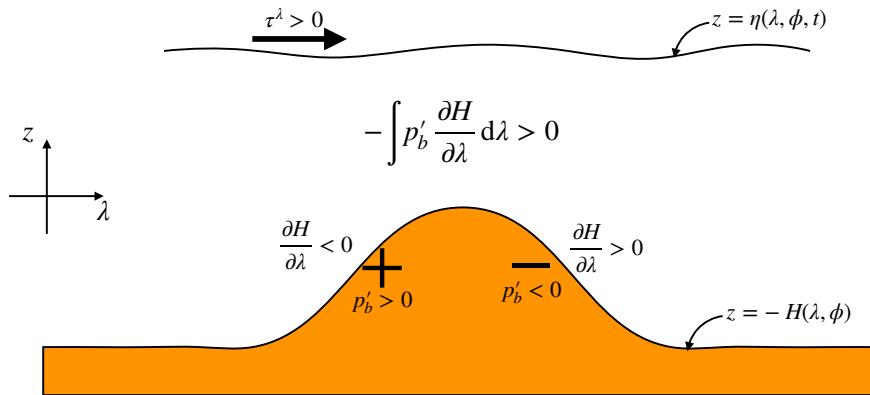


Figure 26.10: Balance between zonal wind stress and bottom form stress. For an eastward zonal wind stress, there is a steady balance if the upwind side of a topographic bump sees an anomalously high bottom pressure whereas on the downwind side it is anomalously low. We come to the same conclusion in Figure 43.4 for the shallow water fluid.

26.10.6 Further study

Elements of this section are based on [Hughes and de Cueves \(2001\)](#) and the analogous discussion of the global atmospheric angular momentum budget developed in Section 10.3 of [Holton \(1992\)](#). The physical processes establishing the balances noted for the Southern Ocean remain under investigation, with further resources to the literature including Section 21.7 of [Vallis \(2017\)](#), and the reviews by [Rintoul et al. \(2001\)](#), [Rintoul and Naveira Garabato \(2013\)](#), and [Rintoul \(2018\)](#). We further revisit this balance within the shallow water system in Section 43.2.6.

26.11 Exercises

EXERCISE 26.1: STEADY STATE OCEAN FORCE BALANCE

Consider a hydrostatic fluid in an ocean basin, \mathcal{R} . The bottom interface separates the ocean from the solid-earth, and the upper interface separates the ocean from the atmosphere with a nonzero mass. Apply a horizontal stress over the ocean surface with a stress vector τ^{surf} , and allow for

the ocean bottom to exchange momentum with the solid-earth through a horizontal bottom stress, τ^{bott} .

- (a) What is the force balance for the full ocean domain at steady state? Express this balance in words and in equations. The answer should be generally stated; no need for specific details.
- (b) What is the vertical component of the force balance, again computed over the full ocean domain? Hint: recall the fluid is assumed to be in hydrostatic balance.
- (c) Assume an *f*-plane (i.e., the Coriolis parameter *f* is a constant; see Section 27.2) and that the center of mass motion vanishes, $\int \rho \mathbf{v} dV = 0$. Discuss the zonal and meridional steady state force balance.

EXERCISE 26.2: FRICTIONAL DISSIPATION FROM VISCOSITY

This exercise works through a simple case of the more general considerations from Section 26.6.3. Here, assume the friction in the momentum equation takes the form

$$\rho \mathbf{F} = \nabla \cdot (\rho \kappa \nabla \mathbf{v}) = \partial_n (\rho \kappa \partial_n \mathbf{v}), \quad (26.91)$$

with $\kappa > 0$ a scalar kinematic viscosity (generally non-constant). Show that when integrated over the full domain

$$\int \mathbf{F} \cdot \mathbf{v} \rho dV < 0, \quad (26.92)$$

where boundary terms are dropped. Hence, the global integrated kinetic energy is dissipated (reduced) through the impacts of viscosity. This dissipation of mechanical energy is converted to an increase in internal energy through Joule heating. Hint: for this exercise, it is sufficient to assume Cartesian tensors so that

$$\rho \mathbf{F} \cdot \mathbf{v} = \rho F_m v_m = \partial_n (\rho \kappa \partial_n v_m) v_m. \quad (26.93)$$

27

Filtered equations

The thermo-hydrodynamical equations (22.14)-(22.18) are suitable to explain a huge range of phenomena. Unfortunately, this generality comes at a cost. Namely, by encapsulating so many physical scales of motion and associated dynamical processes, the equations are difficult to manage when aiming to study a focused dynamical regime. They offer us a tool whose power is overwhelming. Therefore, it is common to seek systematic methods to filter the equations to remove scales of little interest, thus enabling a more telescopic view of the dynamics. In this chapter, we develop certain of the approximations commonly used in geophysical fluid mechanics, in particular we develop the hydrostatic primitive equations as well as the tangent plane approximations.

READER'S GUIDE TO THIS CHAPTER

The equations developed in this chapter, and their associated approximations, will be used extensively throughout the remainder of this book.

27.1	The primitive equations	402
27.1.1	Hydrostatic balance	402
27.1.2	Shallow fluid approximation	402
27.1.3	Traditional approximation	403
27.1.4	Summary of the primitive equations	403
27.1.5	Comments and further study	404
27.2	The f -plane and β -plane approximations	404
27.2.1	Tangent plane approximation	404
27.2.2	Traditional approximation and the f -plane	405
27.2.3	β -plane approximation	405
27.3	Hydrostatic scaling	406
27.3.1	Preliminaries	406
27.3.2	Scaling relations	407
27.3.3	Computing the horizontal hydrostatic pressure gradient	408
27.3.4	Further study	410
27.4	Exercises	410

27.1 The primitive equations

The *primitive equations* provide a set of filtered equations for use in studying large-scale atmospheric and oceanic phenomena. Indeed, nearly all numerical models of the large-scale atmospheric and oceanic circulation are based on the primitive equations. They follow from the following three approximations to the unfiltered equations.

27.1.1 Hydrostatic balance

As discussed in Section 22.3, a static fluid in a gravity field maintains the hydrostatic balance, whereby the pressure at a point equals to the weight per area of fluid above that point. As shown in Section 27.3, the hydrostatic balance is very closely maintained for the larger scales in a moving geophysical fluid. Hence, the hydrostatic balance is central to the study of geophysical fluid dynamics.

Mathematically, the hydrostatic balance represents a balance in the vertical momentum equation (22.23) between the vertical pressure gradient and the effective gravitational force

$$\frac{\partial p}{\partial r} = -\rho g. \quad (27.1)$$

Vertical integration of this equation, while assuming g is constant, renders a diagnostic expression for the hydrostatic pressure at a point as a function of the weight per horizontal area above the point

$$p(r) = p(r_0) + g \int_r^{r_0} \rho dr'. \quad (27.2)$$

27.1.2 Shallow fluid approximation

The ocean and atmosphere each form a fluid shell that envelopes the outer portion of the planet. The thickness of these fluids is small relative to the earth's radius. The shallow fluid approximation¹

¹The shallow fluid approximation is distinct from the *shallow water approximation* treated in Part VII.

builds this scale separation into the equations of motion by setting the radial coordinate equal to the earth's radius

$$r = R + z \approx R. \quad (27.3)$$

This approximation is made where r appears as a multiplier, but not as a derivative operator. For example, the spherical coordinate gradient operator takes the form

$$\nabla \approx \frac{\hat{\lambda}}{R \cos \phi} \frac{\partial}{\partial \lambda} + \frac{\hat{\phi}}{R} \frac{\partial}{\partial \phi} + \hat{r} \frac{\partial}{\partial r}. \quad (27.4)$$

27.1.3 Traditional approximation

The Traditional approximation comprises three approximations that come as a package.

Coriolis acceleration

The Traditional approximation sets to zero the Coriolis terms in the horizontal momentum equations involving the vertical velocity. The earth's angular rotation vector thus takes the form discussed in Section 13.11.6

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{Z}} \quad (27.5a)$$

$$= \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi) \quad (27.5b)$$

$$\approx \Omega \sin \phi \hat{r} \quad (27.5c)$$

$$= \mathbf{f}/2, \quad (27.5d)$$

where

$$\mathbf{f} = (2 \Omega \sin \phi) \hat{r} \quad (27.6)$$

is the Coriolis parameter. Hence, the Traditional approximation is concerned only with the local vertical component of the Earth's angular velocity.

Metric terms

The Traditional approximation also drops the metric terms, uw/r and vw/r , associated with the vertical velocity as they appear in the horizontal momentum equations.

Self consistency

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. Taking one but not the other leads to an inconsistent set of equations (see Exercise 27.2).

27.1.4 Summary of the primitive equations

The above approximations lead to the primitive equations written in spherical coordinates

$$\frac{Du}{Dt} - \frac{uv \tan \phi}{R} - fv = -\frac{1}{\rho R \cos \phi} \frac{\partial p}{\partial \lambda} + F^\lambda \quad (27.7)$$

$$\frac{Dv}{Dt} + \frac{u^2 \tan \phi}{R} + fu = -\frac{1}{\rho R} \frac{\partial p}{\partial \phi} + F^\phi \quad (27.8)$$

$$\frac{\partial p}{\partial z} = -g \rho, \quad (27.9)$$

where the gradient operator is given by

$$\nabla = \frac{\hat{\lambda}}{R \cos \phi} \frac{\partial}{\partial \lambda} + \frac{\hat{\phi}}{R} \frac{\partial}{\partial \phi} + \hat{z} \frac{\partial}{\partial z}. \quad (27.10)$$

We can write these equations in the succinct form

$$\left(\frac{D}{Dt} + (f + u \tan \phi / R) \hat{z} \wedge \right) \mathbf{u} = -\rho \nabla \Phi - \nabla p + \mathbf{F}, \quad (27.11)$$

where

$$\mathbf{F} = \hat{\lambda} F^\lambda + \hat{\phi} F^\phi \quad (27.12)$$

is the horizontal friction vector. Furthermore, the material time derivative in this equation signifies the relative acceleration

$$\frac{Du}{Dt} = \hat{\lambda} \frac{Du}{Dt} + \hat{\phi} \frac{Dv}{Dt}. \quad (27.13)$$

27.1.5 Comments and further study

The primitive equations make use of the momentum equations, which contrasts to “non-primitive” equation methods that develop evolution equations for the vorticity and divergence. [Smagorinsky \(1963\)](#) was among the earliest proponents of the primitive equations for study of the large-scale ocean and atmospheric circulation. These equations form the basis for many of the general circulation models used to simulate the ocean and atmosphere. However, it is notable that a number of finer resolution simulations are making use of the non-hydrostatic fluids, particularly when studying clouds. These models are no longer considered “primitive equation” models. However, computers remain unable to run such “cloud resolving” models for more than a few years, whereas climate models generally require simulations of a century or longer.

27.2 The f -plane and β -plane approximations

Spherical coordinates are ideally suited for the study of planetary fluid dynamics for cases where the fluid samples the earth’s sphericity. However, spherical coordinates remain more complicated to work with than Cartesian coordinates. We are thus led to consider the utility of an idealized tangent plane configuration as part of a hierarchy of theoretical models to help understand the full geophysical system. This motivation leads to the f -plane and β -plane approximations.

27.2.1 Tangent plane approximation

Consider a tangent plane located at latitude $\phi = \phi_0$ and introduce a Cartesian set of coordinates according to

$$(x, y, z) = (R \lambda \cos \phi_0, R (\phi - \phi_0), z) \quad (27.14)$$

$$(\hat{x}, \hat{y}, \hat{z}) = (\hat{\lambda}, \hat{\phi}, \hat{r}). \quad (27.15)$$

Use of these Cartesian coordinates leads to the following inviscid (i.e., no friction) equations of motion

$$\frac{Du}{Dt} + 2(\Omega^y w - \Omega^z v) = -\frac{1}{\rho} \frac{\partial p}{\partial x} \quad (27.16)$$

$$\frac{Dv}{Dt} + 2(\Omega^z u - \Omega^x w) = -\frac{1}{\rho} \frac{\partial p}{\partial y} \quad (27.17)$$

$$\frac{Dw}{Dt} + 2(\Omega^x v - \Omega^y u) = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g, \quad (27.18)$$

with rotational vector components

$$\boldsymbol{\Omega} = \Omega (\cos \phi_0 \hat{\mathbf{y}} + \sin \phi_0 \hat{\mathbf{z}}). \quad (27.19)$$

Note the absence of metric terms is due to the use of Cartesian coordinates on a flat planar geometry.

It is important to remind ourselves that the tangent plane approximation originates from the geopotential vertical coordinate system used for the sphere. In that coordinate system, the effective gravitational acceleration (gravity plus centrifugal) is aligned with the local vertical direction. Correspondingly, the resulting tangent plane equations have the effective gravitational force aligned just in the $\hat{\mathbf{z}}$ direction. These equations are thus slightly different from those describing a fluid in a rotating tank, in which the effective gravity is not aligned with the vertical (see Exercise 27.5 and Section 42.3).

27.2.2 Traditional approximation and the f -plane

The Traditional approximation retains only the local vertical component of the rotation vector, resulting in

$$\frac{Du}{Dt} - fv = -\frac{1}{\rho} \frac{\partial p}{\partial x} \quad (27.20)$$

$$\frac{Dv}{Dt} + fu = -\frac{1}{\rho} \frac{\partial p}{\partial y} \quad (27.21)$$

$$\frac{Dw}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g, \quad (27.22)$$

where we introduced the constant Coriolis parameter

$$f = 2\Omega \sin \phi_0 \equiv f_0. \quad (27.23)$$

The f -plane approximation is the simplest model for a rotating fluid. It provides a useful member in a hierarchy of theoretical models of geophysical flows.

27.2.3 β -plane approximation

Rossby waves are planetary scale waves that sample the earth's spherical nature. The essential ingredient in their existence is the latitudinal dependence of the Coriolis parameter. To capture Rossby waves on a tangent plane requires the meridional gradient of the Coriolis parameter while retaining the flat plane geometry. We thus write

$$f = f_0 + R^{-1} (2\Omega \cos \phi_0) (y - y_0), \quad (27.24)$$

or more simply

$$f = f_0 + \beta y \quad (27.25)$$

$$\beta = \frac{\partial f}{\partial y} = \frac{2 \Omega \cos \phi_0}{R}, \quad (27.26)$$

which constitutes the β -plane approximation.

27.3 Hydrostatic scaling

The hydrostatic balance consists of the vertical momentum equation for a static fluid in a gravitational field

$$\frac{\partial p}{\partial z} = -\rho g. \quad (27.27)$$

Vertically integrating this equation over the depth of the ocean leads to

$$p(z) = p(\eta) + g \int_z^\eta \rho dz, \quad (27.28)$$

where $p(\eta)$ is the pressure at the top surface of the ocean, $z = \eta(x, y, t)$, arising from the weight of the overlying atmosphere. A similar integration applies to the atmosphere

$$p(z) = g \int_z^{z_{\text{top}}} \rho dz, \quad (27.29)$$

where we assume g is a constant over the vertical extent of the atmosphere. In either the ocean or atmosphere, the hydrostatic pressure at a geopotential z equals to the weight per horizontal area of material above that depth.

In addition to static flows, the hydrostatic balance is accurately maintained for a wide range of scales in a moving atmosphere and ocean. We here present a scale analysis for the hydrostatic balance in both unstratified and stratified fluids. This analysis serves to introduce a common method used in fluid mechanics to identify those processes that may be dominant for a particular flow regime. In particular, the flow regime of interest here occurs with a small vertical to horizontal aspect ratio

$$\alpha_{\text{aspect}} \equiv \frac{H}{L} \ll 1, \quad (27.30)$$

with H a typical length scale for vertical motion and L the horizontal length scale. This regime is fundamental to the large-scale circulation of the ocean and atmosphere. As the hydrostatic balance is concerned with balances over a fluid column, it is sufficient to ignore rotation when performing a scale analysis.

27.3.1 Preliminaries

Consider the vertical momentum equation (27.18) from the tangent plane and Traditional approximations (Section 27.2), along with the associated scales for the various terms

$$\frac{Dw}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g \quad (27.31)$$

$$\frac{W}{T} + \frac{U W}{L} + \frac{W W}{H} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g. \quad (27.32)$$

In the second equation we introduced the following scales for the terms appearing on the left hand side of the first equation.

- L is the horizontal scale of the motion.
- H is the vertical scale of the motion.
- W is the vertical velocity scale.
- U is the horizontal velocity scale. For this analysis we do not distinguish between the zonal and meridional velocity scales, writing U for both. This assumption is not always valid, such as when scaling for jet stream or equatorial flows, both of which have larger zonal speeds than meridional.
- T is the time scale of the motion. We assume that the time scale is determined by horizontal advection, so that $T \sim L/U$. For studies of waves, we may alternatively consider time to scale according to a wave speed and wave length, $T \sim \lambda/c$.

To get a sense for the numbers, consider the atmospheric case of $W = 10^{-2} \text{ m s}^{-1}$, $L = 10^5 \text{ m}$, $H = 10^3 \text{ m}$, $U = 10 \text{ m s}^{-1}$. These numbers lead to $T = L/U = 10^4 \text{ s}$ and to the values for the vertical momentum equation

$$10^{-6} \text{ m s}^{-2} \sim -\frac{1}{\rho} \frac{\partial p}{\partial z} - g. \quad (27.33)$$

With $g \sim 10 \text{ m s}^{-2}$, the only term that can balance the gravitational acceleration is the vertical pressure gradient. We thus conclude that large scale motion maintains a hydrostatic balance whereby $\partial p/\partial z = -\rho g$.

27.3.2 Scaling relations

We now proceed more formally by non-dimensionalizing the non-rotating inviscid oceanic Boussinesq momentum equations

$$\frac{D\mathbf{u}}{Dt} = -\nabla_z \psi \quad (27.34)$$

$$\frac{Dw}{Dt} = -\frac{\partial \psi}{\partial z} + b \quad (27.35)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (27.36)$$

where

$$\nabla_z = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} \quad (27.37)$$

is the horizontal gradient operator. For the present analysis, ignore the potential impacts from vertical stratification (we include stratification in Section 28.3). To non-dimensionalize, introduce the dimensional scales (capital letters) and corresponding non-dimensional variables (hat variables) according to

$$(x, y) = L(\hat{x}, \hat{y}) \quad z = H \hat{z} \quad \mathbf{u} = U \hat{\mathbf{u}} \quad w = W \hat{w} \quad (27.38)$$

$$t = T \hat{t} \quad \psi = \Psi \hat{\psi} \quad b = B \hat{b}. \quad (27.39)$$

Importantly, the dimensional scales are constrained through the equations of motion. To expose the constraints requires us to impose our subjective input based on the regimes of interest. To start, assume, as before, that the time scale is determined by the horizontal velocity and the length scale

$$T = \frac{L}{U}. \quad (27.40)$$

Secondly, scale the vertical velocity according to the continuity equation

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \Rightarrow W \sim \frac{UH}{L} = U\alpha_{\text{aspect}}. \quad (27.41)$$

Third, scale the pressure according to the non-rotating balance of the horizontal advection and the horizontal pressure gradient

$$\frac{UU}{L} \sim \frac{\Psi}{L} \Rightarrow \Psi \sim U^2. \quad (27.42)$$

Fourth, scale buoyancy according to the hydrostatic balance

$$B \sim \frac{\Psi}{H} = \frac{U^2}{H}. \quad (27.43)$$

Bringing these results together leads to

$$\frac{U}{T} \frac{\partial \hat{u}}{\partial \hat{t}} + \frac{U^2}{L} \hat{u} \frac{\partial \hat{u}}{\partial \hat{x}} + \frac{U^2}{L} \hat{v} \frac{\partial \hat{u}}{\partial \hat{y}} + \frac{UW}{H} \hat{w} \frac{\partial \hat{u}}{\partial \hat{z}} = -\frac{\Psi}{L} \frac{\partial \hat{\psi}}{\partial \hat{x}} \quad (27.44)$$

$$\frac{U}{T} \frac{\partial \hat{v}}{\partial \hat{t}} + \frac{U^2}{L} \hat{u} \frac{\partial \hat{v}}{\partial \hat{x}} + \frac{U^2}{L} \hat{v} \frac{\partial \hat{v}}{\partial \hat{y}} + \frac{UW}{H} \hat{w} \frac{\partial \hat{v}}{\partial \hat{z}} = -\frac{\Psi}{L} \frac{\partial \hat{\psi}}{\partial \hat{y}} \quad (27.45)$$

$$\frac{W}{T} \frac{\partial \hat{w}}{\partial \hat{t}} + \frac{UW}{L} \hat{u} \frac{\partial \hat{w}}{\partial \hat{x}} + \frac{UW}{L} \hat{v} \frac{\partial \hat{w}}{\partial \hat{y}} + \frac{WW}{H} \hat{w} \frac{\partial \hat{w}}{\partial \hat{z}} = -\frac{\Psi}{H} \frac{\partial \hat{\psi}}{\partial \hat{z}} + B \hat{b}. \quad (27.46)$$

Tidying up these relations leads to the non-dimensional oceanic Boussinesq equations

$$\frac{D\hat{u}}{Dt} = -\hat{\nabla} \hat{\psi} \quad (27.47)$$

$$\alpha_{\text{aspect}}^2 \frac{D\hat{w}}{Dt} = -\frac{\partial \hat{\psi}}{\partial \hat{z}} + \hat{b}, \quad (27.48)$$

where we introduced the non-dimensional material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \hat{\nabla} \quad (27.49)$$

and the non-dimensional gradient operator

$$\hat{\nabla} = \frac{\partial}{\partial \hat{x}} + \frac{\partial}{\partial \hat{y}} + \frac{\partial}{\partial \hat{z}}. \quad (27.50)$$

The non-dimensional vertical momentum equation (27.48) reveals that for small aspect ratio flow, the vertical momentum equation reduces to the hydrostatic balance

$$\frac{\partial \hat{\psi}}{\partial \hat{z}} = \hat{b} \quad \text{for } \alpha_{\text{aspect}}^2 \ll 1. \quad (27.51)$$

27.3.3 Computing the horizontal hydrostatic pressure gradient

We offer an example of how horizontal differences in density lead to horizontal pressure gradients through the hydrostatic relation. This example is emblematic of how one can determine the sign for horizontal gradients in a hydrostatic fluid, and thus to determine the direction for the pressure force.

Two columns with equal mass yet different densities

For this purpose consider two adjacent columns of seawater with equal mass but with distinct density, and assume the density in each column is constant throughout the respective columns. Figure 27.1 offers a schematic, where we make the additional assumption that the two columns sit on a flat surface. We can imagine setting up this configuration by starting with uniform density water, then warming the water in column B more than column A while maintaining constant mass in the two columns. This process sets up a horizontal density gradient with an associated horizontal gradient in the hydrostatic pressure. Furthermore, the less dense water in column B occupies more volume so that its free surface sits higher

$$\eta_B > \eta_A. \quad (27.52)$$

So, what is the sign of the horizontal pressure gradient? As we show in the following, column B (the low density column) has larger hydrostatic pressure than column A (the high density column) for every point in the column, except at the bottom where the two pressures are identical since the two columns have equal mass.

Computing $p(z)$ starting from the equal bottom pressures

Since the two columns have equal mass and equal cross-sectional area, the hydrostatic pressures (weight per unit area) at the bottom of the two columns are equal and given by

$$p_{\text{bot}} = g \rho_A (\eta_A + H) = g \rho_B (\eta_B + H), \quad (27.53)$$

where $z = -H$ is the vertical position at the bottom and $z = \eta$ is the surface. Hence, there is no horizontal pressure gradient at the bottom so that all pressure gradients exist above the bottom.

The hydrostatic pressure at an arbitrary position within column A is given by

$$p_A(z) = g \rho_A (\eta_A - z) = p_{\text{bot}} - g \rho_A (H + z). \quad (27.54)$$

The second equality arose from substituting the bottom pressure from equation (27.53) to eliminate the surface height η_A . Likewise, the pressure in column B is given by

$$p_B(z) = g \rho_B (\eta_B - z) = p_{\text{bot}} - g \rho_B (H + z). \quad (27.55)$$

We can now take the difference between the two hydrostatic pressures to find

$$p_B(z) - p_A(z) = g (H + z) (\rho_A - \rho_B) > 0. \quad (27.56)$$

Hence, at any point above the bottom, the hydrostatic pressure in column B is greater than that in column A. This horizontal difference in the hydrostatic pressure renders a force pointing from column B to column A. Vertically integrating this pressure difference over the thickness of column A leads to the net force per horizontal length

$$F_{\text{pressure B to A}} = \int_{-H}^{\eta_A} [p_B(z) - p_A(z)] dz = (g/2) (\rho_A - \rho_B) (\eta_A + H)^2. \quad (27.57)$$

Inferring pressure gradients starting from the top

Another way to understand why the pressure force points from column B to A is to note that at the top of both columns the pressures are the same (and equal to the uniform atmospheric pressure). However, since column B sits higher than column A , as we move down from $z = \eta_B$ the pressure increases in column B immediately, whereas the pressure in column A remains zero until entering the water at $z = \eta_A < \eta_B$. So it is clear that the pressure in column B is greater than A starting from the surface and moving down. And since the two bottom pressures are equal, then one can infer the pressure isolines as drawn in Figure 27.1.

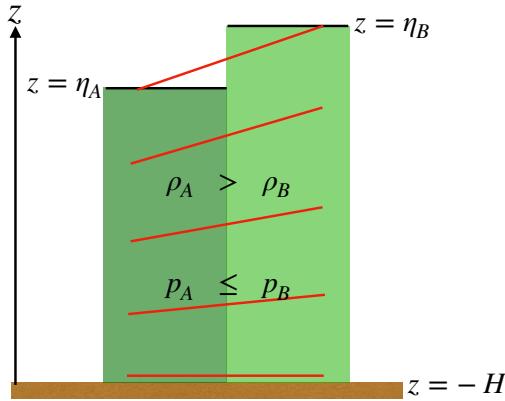


Figure 27.1: Two seawater columns on a flat bottom with equal mass but different densities with $\rho_A > \rho_B$. We assume the atmosphere above the columns has the same pressure over both columns, thus offering zero horizontal pressure force. Furthermore, the cross-sectional area of the two columns are the same so that the less dense water in column B has more volume and thus a greater thickness: $\eta_B > \eta_A$. Since the column masses are the same, the hydrostatic pressures (weight per horizontal area) at the bottom of the two columns are equal: $p_A(z = -H) = p_B(z = -H) = p_{bot}$. In oceanographic parlance, the bottom offers a “level of no motion” from which to reference the pressure field. However, at any position z above the bottom, equation (27.56) shows that the hydrostatic pressure in column B is greater than A : $p_B(z) - p_A(z) = g(H + z)(\rho_A - \rho_B) > 0$. The horizontal gradient in hydrostatic pressure thus points from column B towards column A . The red lines show lines of constant pressure (isobars), which are horizontal next to the bottom but which slope upward to the right moving towards the surface. This configuration provides salient points about hydrostatic pressure relevant for the slightly more complex reduced gravity example in Figure 42.5. Compare this schematic to Figure 29.3, which discusses the depth dependence of the horizontal gradient in hydrostatic pressure as per $\partial(\nabla_z p)/\partial z = -g \nabla_z \rho$.

27.3.4 Further study

Section 2.7.4 in [Vallis \(2017\)](#) provides examples of scales over which the hydrostatic relation remains a useful approximation in geophysical fluids. Our discussion of the hydrostatic pressure field in Section 27.3.3 is motivated by similar considerations presented in Chapter 2 of the descriptive oceanography text from [Tomczak and Godfrey \(1994\)](#).

27.4 Exercises

EXERCISE 27.1: HYDROSTATIC PRESSURE IN AN ISOTHERMAL IDEAL GAS ATMOSPHERE

Consider an isothermal ideal gas atmosphere that is in hydrostatic balance. Derive the following expression for the pressure as a function of geopotential height

$$p(z) = p(0) \exp[-g z / (T R^M)], \quad (27.58)$$

where $p(0)$ is the pressure at the earth's surface, and

$$R^M = \frac{R}{M_{\text{mole}}} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1} \quad (27.59)$$

is the *specific gas constant* as defined by the universal gas constant normalized by the molar mass for the constituent (see Section 23.5.1).

EXERCISE 27.2: PRIMITIVE EQUATIONS AND AXIAL ANGULAR MOMENTUM

The axial angular momentum of a fluid element satisfying the primitive equations is given by

$$L^z = (\rho \delta V) R_\perp (u + R_\perp \Omega) \equiv (\rho \delta V) l^z \quad (27.60)$$

where

$$R_\perp = R \cos \phi \quad (27.61)$$

is the distance from the polar rotation axis to a point on the sphere with radius R , and

$$l^z = R_\perp (u + R_\perp \Omega) \quad (27.62)$$

is the angular momentum per unit mass.

- (a) Consider a constant mass fluid element in the absence of friction. Show that the primitive equation zonal momentum equation (27.7) implies that the material evolution of axial angular momentum per mass is given by

$$\frac{Dl^z}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial \lambda}. \quad (27.63)$$

- (b) Assume the zonal pressure gradient vanishes. Move the fluid element vertically while maintaining a fixed latitude. What happens to the zonal momentum of this primitive equation fluid element? Hint: be sure to remain within the “world” of the primitive equations.
- (c) Give a very brief symmetry argument for why the axial angular momentum is materially conserved when $\partial p / \partial \lambda = 0$. Hint: recall the discussion of Noether’s Theorem in Section 15.1.1.
- (d) Consider the material evolution of primitive equation axial angular momentum per mass in the case where the zonal momentum equation retains the unapproximated form of the Coriolis acceleration. Discuss the resulting material evolution equation. Does this equation make sense based on your symmetry argument given in the previous part of this exercise?

EXERCISE 27.3: RELATIONS FOR AN ATMOSPHERE

In this exercise, we establish some relations for an atmosphere. Assume for this problem that the gravitational acceleration is constant throughout the full depth of the atmosphere. This assumption becomes questionable when integrating to the top of the atmosphere. We furthermore ignore differences in the horizontal cross-sectional area of a fluid column at the bottom and top of the atmosphere arising from the spherical nature of the planet. These two assumptions are sufficient for our purposes. Answering the parts to this question will require reviewing elements of thermodynamics detailed in Chapter 23.

- (a) For an ideal gas atmosphere in hydrostatic balance, show that the integral of the gravitational potential energy plus internal energy from the surface to the top of the atmosphere is equal to the integral of the enthalpy of the atmosphere

$$\int_0^{z_{\text{top}}} (\Phi + \mathcal{I}) \rho dz = \int_0^{z_{\text{top}}} \mathcal{H} \rho dz, \quad (27.64)$$

where

$$\mathcal{H} = p \alpha + \mathcal{I} \quad (27.65)$$

is the enthalpy per mass,

$$\Phi = g z \quad (27.66)$$

is the gravitational potential energy per mass (also known as the *geopotential*) (Section 14.1.2), and \mathcal{I} is the internal energy per mass. The height integral extends from the surface where $z = 0$, to the top of the atmosphere where $z = z_{\text{top}}$.

- (b) For an ideal gas atmosphere in hydrostatic balance, show that

$$\frac{d\sigma}{dz} = \Pi \frac{d\theta}{dz}, \quad (27.67)$$

where

$$\sigma = \mathcal{H} + \Phi \quad (27.68)$$

is the dry static energy and

$$\Pi = \frac{c_p T}{\theta} \quad (27.69)$$

is the Exner function.

- (c) For an ideal gas atmosphere (either hydrostatic or non-hydrostatic), derive the following expression for the pressure gradient force

$$-\frac{1}{\rho} \nabla p = -\theta \nabla \Pi. \quad (27.70)$$

- (d) For an ideal gas atmosphere (either hydrostatic or non-hydrostatic), derive the following expression for the pressure gradient force

$$-\frac{1}{\rho} \nabla p = -\frac{c_s^2}{\rho \theta} \nabla(\rho \theta), \quad (27.71)$$

where c_s is the sound speed.

- (e) Show that for a hydrostatic atmosphere with an arbitrary equation of state

$$\int_0^{p_s} (\Phi + \mathcal{I}) dp = \int_0^{p_s} \mathcal{H} dp. \quad (27.72)$$

That is, show that the relation in the first part of this problem holds even without making the ideal gas assumption.

EXERCISE 27.4: MASS BALANCE FOR A COLUMN OF HYDROSTATIC FLUID

Return to Exercise 17.2. Show that for a hydrostatic fluid the mass balance for a fluid column (equation (17.89)) takes the form

$$\frac{1}{g} \frac{\partial (p_b - p_a)}{\partial t} = -\nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} \rho dz \right] + Q_m, \quad (27.73)$$

where

$$p_b = p_a + g \int_{-H}^{\eta} \rho dz \quad (27.74)$$

is the hydrostatic pressure at the ocean bottom, and p_a is the pressure applied to the ocean surface from the overlying atmosphere or sea ice. Hint: this exercise simply involves the substitution of equation (27.74) into (17.89).

EXERCISE 27.5: ROTATING TANK FLUID DYNAMICS

Consider a rotating circular flat bottom laboratory-scale tank filled with uniform density water. Let the rotational axis be through the center of the circle with angular velocity $\Omega = \Omega \hat{z}$. Orient the vertical coordinate so that the bottom is at $z = 0$ and water surface at $z = \eta(x, y, t)$. Let the gravitational acceleration be uniform $-g_e \hat{z}$ (Section 14.1). Ignore friction throughout.

- (a) A laboratory frame observer is not inertial, since he/she is on the rotating planet. But for the purpose of describing fluid motion in a laboratory scale rotating tank, it seems intuitive that one can ignore the Coriolis force associated with the rotating earth reference frame. Discuss why this intuition is sensible, given a typical rotational speed for a rotating tank (e.g., the speed of a 45 rpm record player) versus that of the planet. We offer scaling analysis in Section 29.2 to further support this intuition.
- (b) Derive the equations of motion for the fluid in the reference frame of an observer in the laboratory, assumed to be in an inertial reference frame.
- (c) Derive the equations of motion in the frame rotating with the tank.
- (d) What is the geometric shape of the free surface when the fluid undergoes solid-body rotation. Neglect any variations in the atmospheric pressure applied to the upper surface.

28

Oceanic Boussinesq fluid

Ocean density deviates no more than a few percent relative to the mean density. Although small, the density deviations are crucial for driving large-scale circulation patterns derived from variations in temperature and salinity (*thermohaline circulation*). The oceanic Boussinesq approximation offers a systematic means to exploit the small density deviations where they can be exploited, while retaining the variations where they are critical. We focus in this chapter with deriving the oceanic Boussinesq equations and studying their properties.

28.1	The oceanic Boussinesq approximation	415
28.1.1	Density decomposition	415
28.1.2	Momentum equation	416
28.1.3	Mass continuity	416
28.1.4	A non-divergent velocity with density evolution	417
28.1.5	Thermodynamic equation	417
28.1.6	Summary of the oceanic Boussinesq equations	418
28.1.7	Atmospheric analog	418
28.1.8	Comments and further study	418
28.2	Hydrostatic Boussinesq ocean model equations	419
28.2.1	Governing equations	419
28.2.2	Material evolution of buoyancy	419
28.2.3	Mechanical forcing and dissipation	420
28.3	Stratification effects on hydrostatic scaling	421
28.4	Exercises	423

28.1 The oceanic Boussinesq approximation

In this section we derive the oceanic Boussinesq approximation, making use of basic scaling analysis.

28.1.1 Density decomposition

Decompose density according to

$$\rho(x, y, z, t) = \rho_0 + \delta\rho(x, y, z, t) \quad (28.1)$$

where the deviation density is much smaller than the reference density

$$\delta\rho \ll \rho_0. \quad (28.2)$$

Also write pressure as the sum of a reference pressure, $p_0(z)$, and a deviation $\delta p(x, y, z, t)$,

$$p(x, y, z, t) = p_0(z) + \delta p(x, y, z, t), \quad (28.3)$$

with the reference pressure in hydrostatic balance with the reference density

$$\frac{dp_0}{dz} = -\rho_0 g. \quad (28.4)$$

28.1.2 Momentum equation

With the above density and pressure decompositions, the momentum equation

$$\rho \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] = -\nabla p - \hat{\mathbf{z}} g \rho \quad (28.5)$$

takes the form

$$(\rho_0 + \delta\rho) \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] = -\nabla \delta p - g \delta\rho \hat{\mathbf{z}} - \left[\frac{dp_0}{dz} + \rho_0 g \right] \quad (28.6a)$$

$$= -\nabla \delta p - g \delta\rho \hat{\mathbf{z}}, \quad (28.6b)$$

where we used the hydrostatic balance (28.4) for the second equality. Consequently, the background hydrostatic pressure, p_0 , has no contribution to the dynamics. It is only δp that has dynamical implications. This decomposition thus offers a useful means to isolate the dynamically relevant portion of the pressure field.

Dividing the momentum equation (28.6b) by the reference density, and dropping the small term $\delta\rho/\rho_0$ on the left hand, yields the Boussinesq momentum equation

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} = -\nabla \psi + b \hat{\mathbf{z}}. \quad (28.7)$$

We introduced the deviation pressure normalized by the reference density

$$\psi = \frac{\delta p}{\rho_0}, \quad (28.8)$$

as well as the buoyancy

$$b = -\frac{g \delta\rho}{\rho_0} = g \left[1 - \frac{\rho}{\rho_0} \right]. \quad (28.9)$$

As defined, buoyancy is positive when the density is less than the reference density. That is, $b > 0$ when the fluid element is lighter (more buoyant) than the reference density.

Buoyancy is the product of the gravitational acceleration, which is a relatively large term, and the small number $\delta\rho/\rho_0$. Their product is not small, so it cannot be neglected from the momentum equation. In effect, we see that the Boussinesq momentum equation ignores all density variations *except* when multiplied by gravity.

28.1.3 Mass continuity

The mass continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}, \quad (28.10)$$

takes the form

$$\frac{D\delta\rho}{Dt} = -(\rho_0 + \delta\rho)(\nabla_z \cdot \mathbf{u} + \partial_z w). \quad (28.11)$$

The material time derivative on the left hand side is much smaller than either of the two terms appearing on the right hand side. Consequently, to leading order, the divergence of the horizontal velocity balances the vertical convergence of the vertical velocity

$$\nabla_z \cdot \mathbf{u} = -\partial_z w. \quad (28.12)$$

That is, the velocity field for a Boussinesq fluid is non-divergent

$$\nabla \cdot \mathbf{v} = 0, \quad (28.13)$$

so that the flow is incompressible.

28.1.4 A non-divergent velocity with density evolution

We here offer some comments regarding the use of a non-divergent velocity for the Boussinesq system, while still having a non-zero material evolution of density.

Comments about density evolution and the thermohaline circulation

Use of a non-divergent velocity for the Boussinesq equations does not mean that the material time evolution of ρ vanishes identically. Instead, the scaling in Section 28.1.3 focuses just on the mass continuity equation. We must additionally acknowledge that as temperature and salinity evolve, so too does density as determined through the equation of state. Such changes in density translate into changes in pressure, which in turn drive the large-scale *thermohaline circulation*.

Divergent and non-divergent velocity components

As noted above, the velocity that results from the Boussinesq momentum equation (i.e., the prognostic Boussinesq velocity) is non-divergent. This is the velocity used for transport as per the material time derivative operator. Additionally, there is a divergent velocity component, \mathbf{v}^d , that balances the material evolution of density

$$\frac{1}{\delta\rho} \frac{D\delta\rho}{Dt} = -\nabla \cdot \mathbf{v}^d \neq 0. \quad (28.14)$$

The divergent velocity \mathbf{v}^d is not used for any of the Boussinesq dynamical equations. Nonetheless, $\mathbf{v}^d \neq 0$, as its divergence is required to balance the material evolution of density according to equation (28.14).

28.1.5 Thermodynamic equation

The thermodynamic equation provides a prognostic relation for the buoyancy. There are various forms for this relation, depending on assumptions made in determining the density $\delta\rho$. We discussed the flavors for density in Section 25.2. For purposes of realistic ocean modeling, the most accurate expression for density is critical. For idealized modeling, it is common to assume buoyancy equals to a constant times the potential temperature.

28.1.6 Summary of the oceanic Boussinesq equations

The oceanic Boussinesq equations are given by

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} = -\nabla\psi + b\hat{\mathbf{z}} + \mathbf{F} \quad (28.15a)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (28.15b)$$

$$\frac{Db}{Dt} = \dot{b} \quad (28.15c)$$

$$b = -\frac{g\delta\rho}{\rho_0} = -\frac{g(\rho - \rho_0)}{\rho_0} \quad (28.15d)$$

$$\psi = \frac{\delta p}{\rho_0} = \frac{p - p_0(z)}{\rho_0} \quad (28.15e)$$

$$\rho = \rho_0 + \delta\rho(S, \theta, z) \quad (28.15f)$$

$$\frac{dp_0}{dz} = -\rho_0 g. \quad (28.15g)$$

The term \dot{b} is a shorthand for diffusion of buoyancy, a boundary flux of buoyancy, or an internal source/sink of buoyancy. Also note that the equation of state is written as a function of salinity, potential temperature, and depth, z . The more accurate expression discussed in Section 25.2 has density as a function of pressure rather than depth. However, a self-consistent energetic balance for the Boussinesq system requires pressure in the equation of state to be approximated by

$$\rho(S, \theta, p) = \rho(S, \theta, p = -g\rho_0 z) \quad \text{Boussinesq density.} \quad (28.16)$$

28.1.7 Atmospheric analog

The atmosphere is far more compressible than the ocean, so that density variations cannot be neglected and the divergent nature of the velocity is important. However, there are some cases in which an atmospheric analog to the Boussinesq approximation can be useful. This analog is known as the anelastic approximation. It is mathematically isomorphic to the oceanic Boussinesq approximation. This approximation has been found to be less useful for the atmosphere than the Boussinesq approximation is for the ocean.

28.1.8 Comments and further study

This section is a summary of Section 2.4 of [Vallis \(2017\)](#), where more details can be found to show that density variations are small within the ocean. Section 2.4.3 and Appendix 2A of [Vallis \(2017\)](#) discuss energetics of the Boussinesq system with a general equation of state, thus showing that energetic consistency requires density to be taken as a function of depth rather than pressure. Further discussion of the Boussinesq approximation can be found in Section 9.3 of [Griffies and Adcroft \(2008\)](#). Section 2.5 of [Vallis \(2017\)](#) discusses the anelastic approximation for the atmosphere.

The oceanic Boussinesq approximation is slightly more general than the Boussinesq approximation considered in other areas of fluid mechanics (e.g., [Chandrasekhar, 1961](#)). In particular, the oceanic Boussinesq approximation does not necessarily assume a linear equation of state (though often that is assumed for idealized studies). Rather, the oceanic Boussinesq approximation is quite commonly used for realistic ocean circulation studies, where accuracy of the equation of state is important for determining gravitational stability and pressure gradients.

28.2 Hydrostatic Boussinesq ocean model equations

In this section we examine some properties of the hydrostatic Boussinesq fluid in the presence of subgrid scale (SGS) processes. We also consider boundary fluxes of buoyancy and momentum. This physical system encapsulates many elements of the primitive equations used for studying the large-scale ocean circulation with numerical ocean models. In this section we merely state the equations, with Exercises 28.5, 28.6, 28.7, and 28.8 developing certain properties of this system.

28.2.1 Governing equations

The governing equations for a hydrostatic Boussinesq fluid in a rotating reference frame are given by

$$\frac{Du}{Dt} = f v - \frac{\partial \psi}{\partial x} + \frac{\partial}{\partial z} \left[A \frac{\partial u}{\partial z} \right] \quad (28.17a)$$

$$\frac{Dv}{Dt} = -f u - \frac{\partial \psi}{\partial y} + \frac{\partial}{\partial z} \left[A \frac{\partial v}{\partial z} \right] \quad (28.17b)$$

$$\frac{\partial \psi}{\partial z} = b \quad (28.17c)$$

$$\frac{Db}{Dt} = -\nabla \cdot \mathbf{F} \quad (28.17d)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (28.17e)$$

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (28.17f)$$

Notation corresponds to the oceanic Boussinesq system discussed in Section 28.1. In particular, $b = -g \delta \rho / \rho_o = -g (\rho - \rho_o) / \rho_o$ is the buoyancy relative to the reference state, and $\psi = \delta p / \rho_o = (p - p_o) / \rho_o$ is the pressure anomaly divided by the reference density ρ_o . Note that for simplicity we ignore effects from viscous dissipation due to horizontal shears, although nearly all ocean models have this friction in addition to the vertical friction considered here. We also assume the global volume of the ocean remains unchanged, so that there are no boundary fluxes of precipitation, evaporation, or river runoff. Finally, we make use of Cartesian coordinates though note that ocean models are generally posed on the sphere.

28.2.2 Material evolution of buoyancy

The material evolution of buoyancy given by equation (28.17d) is affected by the convergence of a subgrid scale (SGS) flux, \mathbf{F} , which we assume takes the form

$$\mathbf{F} = -\kappa \frac{\partial b}{\partial z} \hat{\mathbf{z}} + \mathbf{v}^* b. \quad (28.18)$$

The first term is a downgradient vertical diffusive flux. The vertical eddy diffusivity, $\kappa > 0$, is a function of the flow state, which means that it is a function of space and time

$$\kappa = \kappa(x, y, z, t). \quad (28.19)$$

The second term is an advective flux, where the advective velocity, $\mathbf{v}^* = (\mathbf{u}^*, w^*)$, is assumed to be non-divergent

$$\nabla \cdot \mathbf{v}^* = \nabla_z \cdot \mathbf{u}^* + \frac{\partial w^*}{\partial z} = 0. \quad (28.20)$$

The velocity \mathbf{v}^* is commonly termed the *eddy-induced* velocity, with a particular choice for its parameterization discussed in Exercise 28.8.

The boundary conditions for the buoyancy are given by

$$\kappa \frac{\partial b}{\partial z} = Q_b \quad \text{at } z = \eta \quad (28.21a)$$

$$\kappa \frac{\partial b}{\partial z} = 0 \quad \text{at } z = -H \quad (28.21b)$$

$$\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0 \quad \text{all boundaries,} \quad (28.21c)$$

where Q_b is the surface buoyancy flux, and we assume no buoyancy flux at the ocean bottom (i.e., no geothermal heating).

28.2.3 Mechanical forcing and dissipation

The ocean is a forced-dissipative system, with mechanical and buoyant forcing predominantly at the surface and bottom boundaries and mechanical dissipation via molecular viscosity. We must parameterize mechanical dissipation in the ocean interior arising from subgrid scale processes. A common form for this parameterization is via a vertical transfer of momentum through vertical shears that is weighted by a vertical viscosity $A > 0$. The vertical viscosity is assumed to be a function of the flow state, meaning it is a function of space and time

$$A = A(x, y, z, t). \quad (28.22)$$

We parameterize the mechanical forcing at the ocean boundaries via a boundary stress. This stress is introduced to the governing equations through the following surface and bottom boundary conditions placed on the viscous flux

$$\rho_o A \frac{\partial u}{\partial z} = \tau^x \quad \text{at } z = \eta \quad (28.23a)$$

$$\rho_o A \frac{\partial v}{\partial z} = \tau^y \quad \text{at } z = \eta \quad (28.23b)$$

$$A \frac{\partial u}{\partial z} = C_D u |\mathbf{u}| \quad \text{at } z = -H(x, y) \quad (28.23c)$$

$$A \frac{\partial v}{\partial z} = C_D v |\mathbf{u}| \quad \text{at } z = -H(x, y). \quad (28.23d)$$

At the ocean surface, $z = \eta$, we introduce the stress components, τ^x and τ^y , arising from the transfer of momentum between the ocean and atmosphere. In practice, the stress is computed by a boundary layer parameterization that ingests the wind speed from the atmosphere and computes a stress that is transferred to the ocean through these boundary conditions. Note that the stress imparted to the ocean is equal and opposite the stress felt by the atmosphere at its lower boundary. The same occurs at the ice-ocean boundary.

At the ocean bottom, $z = -H$, we parameterize subgrid scale interactions with bottom topography via a quadratic bottom drag, where $C_D > 0$ is a dimensionless drag coefficient that is sometimes assumed to be a function of the bottom topographic roughness. This bottom stress acts to slow down the ocean bottom currents. It is equal and opposite to the stress transferred to the solid earth.

28.3 Stratification effects on hydrostatic scaling

Stratification impacts on how the hydrostatic equation scales. To include stratification requires us to include the thermodynamics equation, which for the Boussinesq system is the evolution of buoyancy

$$\frac{D\mathbf{u}}{Dt} = -\nabla_z \psi \quad (28.24)$$

$$\frac{Dw}{Dt} = -\frac{\partial \psi}{\partial z} + b \quad (28.25)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (28.26)$$

$$\frac{Db}{Dt} = 0, \quad (28.27)$$

where we ignore irreversible effects on buoyancy. To help isolate the dynamically important portion of stratification, write the buoyancy as

$$b = b'(x, y, z, t) + \tilde{b}(z), \quad (28.28)$$

where $\tilde{b}(z)$, is a static background stratification that is in hydrostatic balance with a corresponding portion of the pressure field

$$\frac{\partial \tilde{\psi}}{\partial z} = \tilde{b}(z). \quad (28.29)$$

The Boussinesq equations thus take the form

$$\frac{D\mathbf{u}}{Dt} = -\nabla_z \psi' \quad (28.30)$$

$$\frac{Dw}{Dt} = -\frac{\partial \psi'}{\partial z} + b' \quad (28.31)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (28.32)$$

$$\frac{Db'}{Dt} = -w N^2, \quad (28.33)$$

where

$$N^2 = \frac{d\tilde{b}}{dz} \quad (28.34)$$

defines the background vertical stratification. The decomposition into a background stratification helps to isolate the dynamical portion of the horizontal pressure gradient by removing a static depth dependent background. It also allows us to consider the dynamically interesting, but simpler, case in which the background stratification dominates those perturbations around it.

Now introduce the dimensional scales and corresponding non-dimensional quantities

$$(x, y) = L(\hat{x}, \hat{y}) \quad z = H\hat{z} \quad \mathbf{u} = U\hat{\mathbf{u}} \quad w = W\hat{w} \quad (28.35)$$

$$t = T\hat{t} \quad \psi' = \Psi\hat{\psi}' \quad b' = B\hat{b}' \quad N^2 = \bar{N}^2 \hat{N}^2. \quad (28.36)$$

As for the previous case in Section 27.3.2, we impose our subjective regime choices to help constrain the dimensional variables. In addition to those choices considered in the absence of stratification, we acknowledge that the vertical velocity will likely be reduced in the presence of stratification, given

that vertical stratification acts to suppress vertical motion. We thus introduce a non-dimensional number ϵ so that

$$w = W \hat{w} = \epsilon \left[\frac{HU}{L} \right] \hat{w}. \quad (28.37)$$

Inserting into the stratified Boussinesq equations then leads to the non-dimensional system

$$\frac{D\hat{u}}{Dt} = -\hat{\nabla}\hat{\psi}' \quad (28.38)$$

$$\epsilon \alpha_{\text{aspect}}^2 \frac{D\hat{w}}{Dt} = -\frac{\partial\hat{\psi}'}{\partial\hat{z}} + \hat{b}' \quad (28.39)$$

$$\left[\frac{U^2}{N^2 H^2} \right] \frac{D\hat{b}'}{Dt} + \epsilon \hat{N}^2 \hat{w} = 0 \quad (28.40)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}} + \epsilon \frac{\partial\hat{w}}{\partial\hat{z}} = 0 \quad (28.41)$$

where we introduced the non-dimensional material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial\hat{t}} + \hat{\mathbf{u}} \cdot \hat{\nabla}_z + \epsilon \hat{w} \frac{\partial}{\partial\hat{z}}. \quad (28.42)$$

At this point we make a choice for the parameter ϵ . There are many choices that one could consider. For our interests it is suitable to set ϵ equal to the squared Froude number

$$\epsilon = \text{Fr}^2 = \frac{U^2}{N^2 H^2}. \quad (28.43)$$

The Froude number measures the relative strength of vertical shears of the horizontal velocity, U/H , versus the buoyancy stratification, N . Alternatively, it measures the ratio of the horizontal speed for a fluid particle, U , to an internal gravity wave speed, NH . Large Froude numbers indicate large particle speeds relative to wave speeds, with $Fr > 1$ a common indicator of hydraulic instability (see Exercise 42.3 for a brief example). In contrast, a relatively strong stratification (N^2 large) corresponds to a small Froude number and thus flow that is stabilized by vertical stratification. Note that the squared Froude number is the inverse of the Richardson number

$$\text{Ri} = \text{Fr}^{-2} = \frac{N^2 H^2}{U^2}. \quad (28.44)$$

It is a matter of taste whether one works with Fr or Ri .

With this choice the vertical velocity scale is given by

$$W = \text{Fr}^2 \left[\frac{HU}{L} \right]. \quad (28.45)$$

We see that for $Fr < 1$ that the stratification reduces the scale for the vertical velocity. With this choice for ϵ , the non-dimensional Boussinesq equations now take the form

$$\frac{D\hat{u}}{Dt} = -\hat{\nabla}_z \hat{\psi}' \quad (28.46)$$

$$\text{Fr}^2 \alpha_{\text{aspect}}^2 \frac{D\hat{w}}{Dt} = -\frac{\partial\hat{\psi}'}{\partial\hat{z}} + \hat{b}' \quad (28.47)$$

$$\frac{D\hat{b}'}{Dt} + \hat{N}^2 \hat{w} = 0 \quad (28.48)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}} + \text{Fr}^2 \frac{\partial\hat{w}}{\partial\hat{z}} = 0. \quad (28.49)$$

The condition for hydrostatic balance in a stratified fluid thus takes the form

$$\text{Fr}^2 \alpha_{\text{aspect}}^2 \ll 1. \quad (28.50)$$

This result supports our initial suspicion that stratification suppresses vertical motion, thus reducing the vertical acceleration terms that break hydrostatic balance. That is, a stratified flow is more likely to be in hydrostatic balance than an unstratified flow. Note also that the horizontal divergence of the horizontal flow is reduced by the presence of stratification, which thus leads to a nearly horizontally non-divergent flow

$$\left| \hat{\nabla} \cdot \hat{\mathbf{u}} \right| = \left| \text{Fr}^2 \frac{\partial \hat{w}}{\partial \hat{z}} \right| \ll \left| \frac{\partial \hat{w}}{\partial \hat{z}} \right|. \quad (28.51)$$

28.4 Exercises

EXERCISE 28.1: GLOBAL MEAN SEA LEVEL

In this problem, we consider some basic features of global mean sea level by making use of the mass budget of liquid seawater. Elements of this problem are discussed in [Griffies and Greatbatch \(2012\)](#) and [Griffies et al. \(2014\)](#). Note that most of this question involves fully compressible non-Boussinesq notions, yet the final part asks about global mean sea level in a Boussinesq fluid.

Background

The total liquid seawater mass, \mathcal{M} , changes via boundary mass fluxes

$$\frac{d\mathcal{M}}{dt} = \mathcal{A} \overline{Q_m}, \quad (28.52)$$

where \mathcal{A} is the ocean surface area and $\overline{Q_m}$ is the area averaged surface mass flux. Global volume of liquid seawater

$$\mathcal{V} = \frac{\mathcal{M}}{\langle \rho \rangle} \quad (28.53)$$

changes due to mass changes *and* changes to the global mean density, $\langle \rho \rangle$. Furthermore, assuming the surface area of the ocean, \mathcal{A} , is constant, and a constant area averaged ocean bottom depth, \overline{H} , then changes in ocean volume arise just from changes in 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

$$\left[\frac{d\bar{\eta}}{dt} \right]_{\text{observed}} \approx 3 \text{ mm yr}^{-1}. \quad (28.54)$$

There are many processes contributing to this rise. We consider just two in this exercise.

The questions

- (a) If one-half of the observed sea level rise is associated with surface mass flux (e.g., melting land ice), what would be the required area averaged surface ocean mass flux $\overline{Q_m}$ (mass per time per horizontal ocean area)? Compare this mass flux to the net mass transport associated with all the rivers in the world, which is roughly

$$\mathcal{T}_{\text{river}}^{\text{river}} \approx 1.2 \times 10^9 \text{ kg s}^{-1}. \quad (28.55)$$

- (b) If another half of the observed sea level rise is associated with thermal expansion due to ocean warming, what would be the required rate of global mean ocean potential temperature increase (expressed as Kelvin per century)?
- (c) For part (b), what would be the corresponding area averaged surface ocean heat flux \overline{Q}_H (expressed as Watts per square meter of ocean surface area) required to induce this ocean warming? Assume an area averaged ocean depth of $\overline{H} = 4000$ m.
- (d) The heat released by one atomic bomb detonated during World War II is roughly

$$\mathcal{E}_{\text{bomb}} \approx 6.3 \times 10^{13} \text{ J.} \quad (28.56)$$

Assume $\mathcal{E}_{\text{bomb}}$ of energy is evenly distributed over the surface area of the ocean every second. Roughly how many bombs per second does the heat flux from part (c) correspond to? That is, convert the rate of surface ocean heating into units of bombs per second of energy entering the ocean.

- (e) To derive the global mean sea level budget, we started from the mass budget for the global ocean, (28.52). However, we know that for many purposes it is suitable to assume the ocean satisfies Boussinesq kinematics, which are based on volume budgets. So consider a Boussinesq ocean in the absence of boundary mass fluxes, $\overline{Q}_m = 0$. Discuss what happens to the volume of the Boussinesq ocean when there is a surface boundary heat flux, $\overline{Q}_H > 0$?

To help answer these questions, note the following.

- Ignore salinity and pressure effects on density, so that changes in global mean density arise just from changes in global mean potential temperature (this is a reasonable assumption).
- Assume a constant thermal expansion coefficient ([Vallis \(2017\)](#) page 14)

$$\beta_\theta = -\frac{1}{\rho} \left[\frac{\partial \rho(S, \theta, p)}{\partial \theta} \right]_{S,p} = 2 \times 10^{-4} \text{ K}^{-1}. \quad (28.57)$$

This is not a great approximation, since the thermal expansion coefficient ranges over the ocean by a factor of 10. Nonetheless, for this exercise it is not horribly wrong.

- You will need information about further ocean properties. Provide sources for your numbers.
- This exercise is seeking rough calculations requiring just a bit of elementary calculus and attention to physical dimensions. No more than two significant digits are warranted on any numerical result.

EXERCISE 28.2: POISSON EQUATION FOR PRESSURE

Consider a perfect Boussinesq fluid that is rotating on a β -plane and thus satisfies the governing equations

$$\frac{D\mathbf{v}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{v} = -\nabla \psi + \hat{\mathbf{z}} b \quad (28.58a)$$

$$\frac{Db}{Dt} = 0 \quad (28.58b)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (28.58c)$$

$$f = f_o + \beta y. \quad (28.58d)$$

Show that the dynamic pressure ψ satisfies the following elliptic boundary value problem (see Section 5.4 for discussion of elliptic partial differential equations)

$$-\nabla^2 \psi = \nabla \cdot \mathbf{G} \quad \text{within the domain} \quad (28.59a)$$

$$-\hat{\mathbf{n}} \cdot \nabla \psi = \hat{\mathbf{n}} \cdot \mathbf{G} \quad \text{on the domain boundaries.} \quad (28.59b)$$

Determine the vector \mathbf{G} .

EXERCISE 28.3: ENERGETICS FOR A PERFECT BOUSSINESQ FLUID

Consider the energy budget for a perfect unforced Boussinesq fluid in a rotating frame under the traditional approximation. The momentum and buoyancy equations are given by

$$\frac{D\mathbf{v}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{v} = -\nabla \psi + \hat{\mathbf{z}} b \quad (28.60a)$$

$$\frac{Db}{Dt} = 0 \quad (28.60b)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (28.60c)$$

Assume a linear equation of state so that buoyancy is linearly proportional to temperature with a constant thermal expansion coefficient. Assume zero boundary fluxes of mass, heat, and momentum.

Hint: elements of this exercise are discussed in Section 2.4.3 of [Vallis \(2017\)](#). Furthermore, we are uninterested in the potential and kinetic energies of the background state of a constant density fluid with density ρ_0 . That is the reason for working with the buoyancy rather than the full density. Finally, note that it is sufficient to use Cartesian tensors throughout this exercise.

- (a) Derive the material evolution equation for kinetic energy per mass.
- (b) Consider the potential function $\Phi = -z$ and derive the material evolution equation for Φb . Interpret the product Φb .
- (c) Derive the material evolution equation for mechanical energy per mass. Discuss the reversible transfer of energy between potential energy and kinetic energy associated with vertical motion.

EXERCISE 28.4: ENERGETICS FOR A DISSIPATIVE BOUSSINESQ FLUID

We here consider the energy budget for a Boussinesq fluid in a rotating frame under the Traditional approximation, here including diabatic forcing, momentum mixing from molecular viscosity, and buoyancy mixing from molecular diffusion. The equations for this system are

$$\frac{D\mathbf{v}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{v} = -\nabla \psi + \hat{\mathbf{z}} b + \nu \nabla^2 \mathbf{v} \quad (28.61a)$$

$$\frac{Db}{Dt} = Q + \kappa \nabla^2 b \quad (28.61b)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (28.61c)$$

In the above, ν is the viscosity leading to irreversible mixing of velocity, and κ is the diffusivity leading to irreversible mixing of buoyancy. Assume both are constant for this exercise; e.g., they are molecular values. Assume a linear equation of state so that buoyancy is linearly related to temperature with a constant thermal expansion coefficient. Finally, Q is a buoyancy forcing associated with irreversible changes in heat (e.g., boundary fluxes and/or penetrative shortwave radiation).

Hint: elements of this exercise are discussed in Section 2.4.3 of [Vallis \(2017\)](#) for the perfect fluid case (no irreversible processes) and Chapter 21 for the case with dissipation. It is sufficient to use Cartesian tensors throughout this exercise.

- (a) Derive the material evolution of kinetic energy per mass.
- (b) Derive the equation for domain integrated evolution of kinetic energy per mass. Make use of the identity
- $$\nabla^2 \mathbf{v} = -\nabla \wedge \boldsymbol{\omega} \quad \text{if } \nabla \cdot \mathbf{v} = 0 \quad (28.62)$$
- to express the viscosity contributions in terms of the vorticity
- $$\boldsymbol{\omega} = \nabla \wedge \mathbf{v}. \quad (28.63)$$
- (c) Show that the domain integrated kinetic energy per mass is reduced (dissipated) by the viscosity in the presence of a nonzero vorticity. Discuss this result.
- (d) Consider the potential function $\Phi = -z$ and derive the material evolution equation for Φb . Interpret the product Φb .
- (e) Derive the equation for domain integrated evolution of gravitational potential energy per mass.
- Discuss how downgradient vertical diffusion of buoyancy impacts on the domain integrated gravitational potential energy.
 - Discuss how diabatic heating impacts on the domain integrated gravitational potential energy.
- (f) Discuss the reversible conversion between kinetic and mechanical energy.

- (g) Derive the equation for material evolution of mechanical energy per mass.
- (h) Derive the domain integrated evolution of mechanical energy per mass.
- (i) Over a closed volume (no boundary fluxes), show that the dissipation of domain integrated kinetic energy is balanced by the buoyancy source.
- (j) In a steady state (Eulerian time derivative vanishes) and absent buoyancy diffusion, show that diabatic heating from Q must occur at a lower level (lower gravitational geopotential) than cooling if a kinetic energy dissipating circulation is to be maintained.

EXERCISE 28.5: KINETIC ENERGY AND THE OCEAN MODEL EQUATIONS

In this exercise we develop some properties of the kinetic energy for the hydrostatic ocean model equations stated in Section 28.2.

- (a) Why does the kinetic energy only have contributions from the horizontal velocity components?
- (b) Derive the Eulerian flux-form expression for the kinetic energy budget.
- (c) Discuss the role of vertical viscosity in transporting kinetic energy in the vertical.
- (d) Discuss the role of vertical viscosity in dissipating kinetic energy.
- (e) Discuss how wind stress and bottom drag impact the globally integrated kinetic energy.

EXERCISE 28.6: GRAVITATIONAL POTENTIAL ENERGY AND THE OCEAN MODEL EQUATIONS

In this exercise we develop some properties of the gravitational potential energy for the hydrostatic ocean model equations stated in Section 28.2.

- (a) Derive the Eulerian flux-form budget for gravitational potential energy.
- (b) Discuss the role of SGS advection in this budget. In particular, discuss its impact on the center of mass of the fluid.
- (c) Discuss the role of vertical diffusion in this budget. In particular, discuss its impact on the center of mass of the fluid.
- (d) Integrate the gravitational potential energy budget over the global ocean. Discuss how the boundary buoyancy flux, Q_b , impacts on the global potential energy budget through impacts on the center of mass.

EXERCISE 28.7: BUOYANCY AND THE OCEAN MODEL EQUATIONS

In this exercise we develop some properties of the squared buoyancy for the hydrostatic ocean model equations stated in Section 28.2.

- (a) Write the Eulerian flux-form budget describing the evolution of b^2 , the squared buoyancy.
- (b) Discuss the impacts from vertical diffusion on the b^2 budget.

EXERCISE 28.8: 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 ocean model equations stated in Section 28.2. Namely, we consider the specific form for the parameterized eddy-induced velocity proposed by [Gent et al. \(1995\)](#)

$$\mathbf{u}^* = -\partial_z(B \mathbf{S}) \quad (28.64a)$$

$$w^* = \nabla_z \cdot (B \mathbf{S}) \quad (28.64b)$$

$$\mathbf{S} = -\frac{\nabla_z b}{N^2} \quad (28.64c)$$

$$\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0 \quad \text{at all ocean boundaries.} \quad (28.64d)$$

In this expression, $B > 0$ is an eddy diffusivity. In order to ensure $\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0$ at all ocean boundaries, we must have $B = 0$ at 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 $N^2 > 0$.

- (a) Determine the vector streamfunction Ψ^* such that

$$\mathbf{v}^* = \nabla \wedge \Psi^*. \quad (28.65)$$

- (b) Show that

$$\int_{-H}^{\eta} \mathbf{u}^* dz = 0. \quad (28.66)$$

That is, the depth integrated parameterized horizontal flow vanishes.

- (c) At any chosen meridional position y , the meridional buoyancy transport from advection (resolved and parameterized) is computed by

$$\mathcal{B}^{(y)}(y, t) = \int_{x1}^{x2} dx \int_{-H}^{\eta} b(v + v^*) dz. \quad (28.67)$$

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 v^* to the buoyancy equation (28.17d) affect the global integrated gravitational potential energy? Discuss.
- (e) How does the introduction of v^* to the buoyancy equation (28.17d) affect the global integrated available potential energy? Discuss.

29

Geostrophy and thermal wind

Fluid motion dominated by rotation is characterized by a small Rossby number. Inviscid flow with a small Rossby number maintains the geostrophic balance, which is a balance between the Coriolis acceleration and the pressure gradient acceleration. In this chapter, we introduce salient features of geostrophically balanced flow and the associated thermal wind shear. The associated diagnostic relations involve no time derivatives, and so cannot be used to predict the evolution of the fluid flow. However, their power for diagnostics is unquestioned as they provide a robust framework for interpreting the large-scale circulation of the atmosphere and ocean.

READER'S GUIDE TO THIS CHAPTER

This chapter assumes an understanding of the primitive equations from Chapter 27 and the Coriolis acceleration from Chapter 14. The material in this chapter is fundamental to understanding the nature and mechanisms of large-scale flow in the atmosphere and ocean. We make much use of this chapter in the remainder of this book. Throughout this discussion we are not explicitly concerned with sphericity, thus enabling the use of Cartesian coordinates.

29.1	Primitive equations	430
29.2	The Rossby number	431
29.2.1	Scaling for the Rossby number	431
29.2.2	Ratio of material acceleration to Coriolis acceleration	431
29.2.3	Ratio of local time tendency to Coriolis acceleration	432
29.2.4	Rossby number for a kitchen sink	432
29.2.5	Rossby number for a Gulf Stream ring	432
29.3	Geostrophic balance	433
29.3.1	Geostrophic balance is distinctly fluid mechanical	433
29.3.2	Geostrophic relation in geopotential coordinates	433
29.3.3	Cyclonic and anti-cyclonic orientation	433
29.3.4	Density gradients and thermal wind shear	434
29.3.5	Geostrophic relation in pressure coordinates	435
29.3.6	Further study	435
29.4	Introducing planetary geostrophy	436
29.4.1	Planetary geostrophic equations	436
29.4.2	Planetary geostrophic vorticity equation	437
29.4.3	Taylor-Proudman and vertical stiffening	438
29.4.4	Vorticity balance	439
29.4.5	Thermal wind balance for the ocean	439
29.4.6	Thermal wind balance for the atmosphere	441
29.4.7	Further study	443
29.5	Isopycnal form stress from geostrophic eddies	443
29.5.1	Zonal mean zonal form stress on an isopycnal surface	444
29.5.2	Zonal mean zonal form stress acting on an isopycnal layer	445
29.5.3	Comments and further study	447
29.6	Exercises	448

29.1 Primitive equations

Throughout this chapter, we make use of the hydrostatic primitive equations derived in Section 27.1

$$\rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla + f \hat{z} \wedge \right] \mathbf{u} = -\nabla_z p \quad (29.1a)$$

$$\frac{\partial p}{\partial z} = -g \rho \quad (29.1b)$$

$$\frac{D\rho}{Dt} = \rho \nabla \cdot \mathbf{v}, \quad (29.1c)$$

where the velocity vector is written

$$\mathbf{v} = \mathbf{u} + \hat{z} w = \hat{x} u + \hat{y} v + \hat{z} w, \quad (29.2)$$

and the horizontal gradient operator is

$$\nabla_z = \hat{x} \partial_x + \hat{y} \partial_y. \quad (29.3)$$

For some of the scaling analysis we assume an incompressible fluid as per the Boussinesq approximation (Section 28.1). In this case, the mass continuity equation (29.1c) becomes the non-divergent condition on the velocity

$$\nabla \cdot \mathbf{v} = 0. \quad (29.4)$$

Furthermore, ρ in the horizontal momentum equation (29.1a) is converted to a constant reference density ρ_0 . However, it remains the full density when appearing in the hydrostatic equation due to being multiplied by the gravitational acceleration.

29.2 The Rossby number

Large-scale geophysical fluid flows are strongly influenced by the earth's rotation. Indeed, the earth can be considered a rapidly rotating planet for much of the observed motion of the ocean and atmosphere. There are two points to emphasize in this regard. First, much of the atmospheric motion is close to solid-body rotation, in which weather patterns are best viewed relative to the rotating sphere rather than relative to the "fixed" stars. Second, human scale horizontal length scales are generally far too small to take notice of the planetary rotation. This point is quantified by considering the Rossby number, which includes a horizontal length scale, a velocity scale, and angular rotation speed.

29.2.1 Scaling for the Rossby number

The *Rossby number* measures the ratio of the Coriolis acceleration to the material acceleration (acceleration of a fluid particle). The material acceleration has two contributions; one from local time tendencies and one from advection. We expose typical scales for the acceleration by writing

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{u} \sim \frac{U}{T} + \frac{U^2}{L} + \frac{WU}{H}, \quad (29.5)$$

where U, W are typical horizontal and vertical velocity scales, L, H are typical horizontal and vertical length scales, and T is a typical time scale (recall a similar scale analysis for the hydrostatic balance in Section 27.3). Likewise, the Coriolis acceleration scales as

$$f \hat{\mathbf{z}} \wedge \mathbf{u} \sim f_0 U, \quad (29.6)$$

where $f_0 > 0$ is the scale for the Coriolis parameter. From the continuity equation for incompressible flow ($\nabla \cdot \mathbf{v} = 0$) we see that the vertical and horizontal velocity scales are related by

$$\frac{W}{H} \sim \frac{U}{L} \Rightarrow W \sim U \frac{H}{L}. \quad (29.7)$$

For compressible hydrostatic flows, we replace W with the scale for motion across pressure surfaces. In either the incompressible or compressible case, we assume the vertical to horizontal grid aspect ratio is small

$$\frac{H}{L} \ll 1, \quad (29.8)$$

as per the hydrostatic balance discussed in Section 27.3. Consequently, the vertical velocity scale is much less than the horizontal

$$W \ll U. \quad (29.9)$$

29.2.2 Ratio of material acceleration to Coriolis acceleration

Taking the ratio of the advection scale to the Coriolis scale leads to our first expression for the Rossby number

$$Ro = \frac{U^2/L}{f_0 U} = \frac{U}{f_0 L}. \quad (29.10)$$

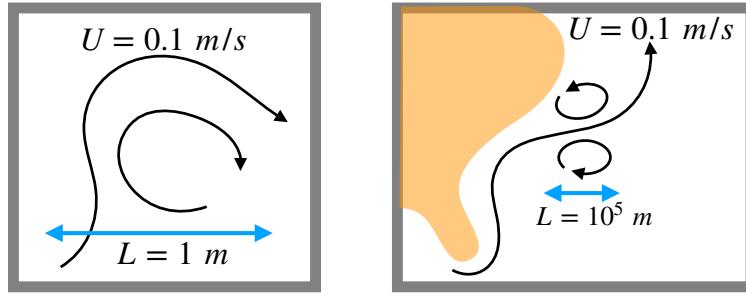


Figure 29.1: Estimating the Rossby number for flow in a kitchen sink (left panel) and rings spawned from the Gulf Stream (right panel). The kitchen sink has velocity scales on the order of $U \sim 0.01 - 0.1 \text{ m s}^{-1}$ whereas Gulf Stream rings have velocity scales on the order $U \sim 0.1 - 1.0 \text{ m s}^{-1}$. However, their length scales are very different, with the scale for a sink $L \sim 1\text{m}$ yet the Gulf Stream rings $L \sim 10^5 \text{m}$. Taking the Coriolis parameter to be at 30° for both flows leads to $Ro_{\text{sink}} \sim 10^2 - 10^3$ and $Ro_{\text{ring}} \sim 10^{-2} - 10^{-1}$. The Coriolis acceleration is clearly important for the Gulf Stream rings whereas it is utterly negligible for the kitchen sink.

As the length and velocity scales are not constant, the Rossby number is a function of more than just the latitude. Nonetheless, due to the latitudinal variation of the Coriolis parameter, the Rossby number is generally small near the poles and large in the tropics.

29.2.3 Ratio of local time tendency to Coriolis acceleration

A complementary way to understand the Rossby number is to consider it as the ratio of the local time tendency for the horizontal velocity to the Coriolis acceleration

$$Ro = \frac{U/T}{Uf_0} = \frac{1/T}{f_0}. \quad (29.11)$$

Thus, for motions that have a low frequency T^{-1} compared to the *rotational inertial frequency* f_0^{-1} , the Rossby number is small. In both ways of writing the Rossby number, we associate small Ro with regimes of flow where the earth's rotation plays a crucial role on the dynamics.

29.2.4 Rossby number for a kitchen sink

Consider flow in a kitchen sink (left panel of Figure 29.1). Here, the length scale is $L = 1 \text{ m}$ (sink size) and the velocity scale is $U = 0.01 - 0.1 \text{ m s}^{-1}$, thus giving a typical time scale for sink motion of $L/U \approx 10 \text{ s} - 100 \text{ s}$. Hence, at 30° latitude, where $f = 2\Omega \sin \phi = \Omega$, the Rossby number for fluid motion in a sink is

$$Ro_{\text{sink}} \approx 10^2 - 10^3. \quad (29.12)$$

The Coriolis force is therefore negligible for kitchen sink fluid dynamics. This result explains the difficulty of experimentally correlating the hemisphere to a preferred rotational direction of water leaving a sink drain.

29.2.5 Rossby number for a Gulf Stream ring

For a Gulf Stream ring (right panel of Figure 29.1), the typical length scale is $L = 10^5 \text{ m}$ and velocity scale is $U = 0.1 - 1.0 \text{ m s}^{-1}$, thus leading to a time scale $L/U \approx 10^5 - 10^6 \text{ s}$. At 30° latitude the Rossby number is

$$Ro_{\text{ring}} \approx 10^{-2} - 10^{-1}, \quad (29.13)$$

thus indicating the importance of the Coriolis acceleration for dynamics of Gulf Stream rings.

29.3 Geostrophic balance

Under the influence of horizontal pressure forces, a fluid accelerates down the pressure gradient (movement from high pressure to low pressure). In the presence of rotation, a nonzero horizontal velocity couples to the Coriolis parameter f , thus giving rise to a nonzero horizontally oriented Coriolis acceleration $-f \hat{\mathbf{z}} \wedge \mathbf{u}$. In a manner directly analogous to the Lorentz force in electrodynamics, the Coriolis acceleration acts perpendicular to the fluid motion

$$\mathbf{u} \cdot (\hat{\mathbf{z}} \wedge \mathbf{u}) = 0. \quad (29.14)$$

Hence, the Coriolis force effects the fluid motion but does not alter its kinetic energy; i.e., it does zero work on the fluid. In the northern hemisphere where $f > 0$, the Coriolis force acts to the right of the parcel motion, thus causing counter-clockwise motion around low pressure centers and clockwise motion around high pressure centers (Figure 29.2). In the southern hemisphere, where $f < 0$, it acts in the opposite direction.

29.3.1 Geostrophic balance is distinctly fluid mechanical

When pressure and Coriolis forces balance, parcel motion is said to be in *geostrophic* balance, whereby large-scale winds and currents generally follow isobars (lines of constant pressure). Recall from Chapter 13 that point particles also experience a Coriolis acceleration when viewed in a rotating reference frame. However, geostrophic balance is not afforded to particles since particles do not experience a pressure force that can balance the Coriolis force. Hence, the geostrophic balance is a distinctly fluid mechanical phenomena.

29.3.2 Geostrophic relation in geopotential coordinates

Mathematically, the geostrophic balance becomes important when the Rossby number is small. When the Rossby number is small and friction is negligible, the leading order dynamical balance in the horizontal momentum equation (29.1a) is between the Coriolis acceleration and pressure gradient acceleration

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = -\rho^{-1} \nabla_z p, \quad (29.15)$$

or equivalently, we have the expression for the geostrophic velocity¹

$$\mathbf{u}_g = \frac{\hat{\mathbf{z}} \wedge \nabla p}{f \rho} \implies u_g = -\frac{1}{f \rho} \frac{\partial p}{\partial y} \quad \text{and} \quad v_g = \frac{1}{f \rho} \frac{\partial p}{\partial x}. \quad (29.16)$$

Note that the equator is special since the Coriolis parameter, $f = 2\Omega \sin \phi$, vanishes, thus precluding the relevance of geostrophy near the equator.

29.3.3 Cyclonic and anti-cyclonic orientation

When oriented in the same sense as the earth's rotation (i.e., same sign of the Coriolis parameter) rotational motion is said to be in a *cyclonic* sense. Oppositely oriented motion is *anti-cyclonic*. For example, geostrophic motion around a low pressure center in the northern hemisphere is counter-clockwise (Figure 29.2). Using the right hand rule, this motion represents a positively oriented rotation. Hence, with $f > 0$ in the north, counter-clockwise motion is cyclonic. Similarly in the

¹We can write either ∇ or ∇_z in equation (29.16). The reason is that the $\hat{\mathbf{z}} \wedge$ operator selects only the horizontal portion of the gradient.

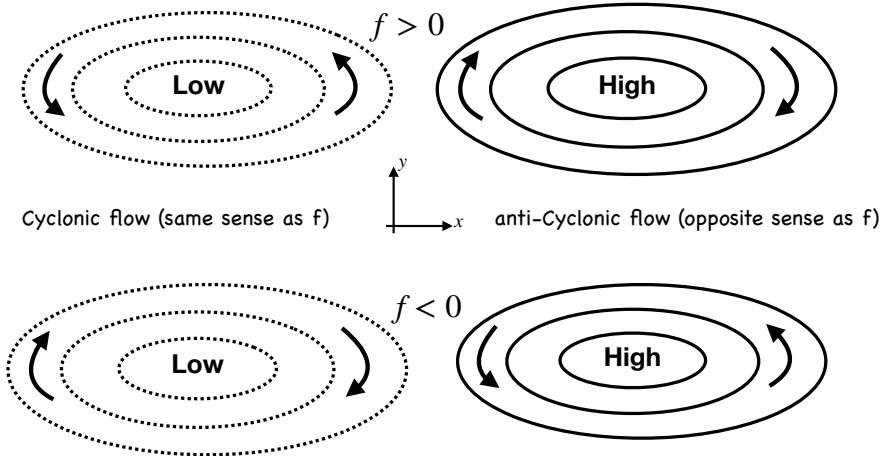


Figure 29.2: Geostrophic motion around low and high pressure centers in the northern hemisphere and southern hemispheres ($f = 2\Omega \sin \phi > 0$ in the north and $f < 0$ in the south). Upper panel: the counter-clockwise motion around the low pressure center is in the same sense as the planetary rotation, and is thus termed cyclonic. Cyclonic motion in the Southern Hemisphere occurs in a clockwise direction, again corresponding to the planetary rotation direction as viewed from the south. Geostrophic motion around a high pressure center is counter to the planetary rotation in both hemispheres, and is thus termed anti-cyclonic.

south, geostrophic motion around a low pressure center is clockwise, which is a negatively oriented rotational motion (again, recall the right hand rule). In the south where $f < 0$, clockwise motion around a low pressure center also represents cyclonic motion (Figure 29.2).

29.3.4 Density gradients and thermal wind shear

The horizontal momentum is affected by horizontal pressure gradient forces. Furthermore, the hydrostatic balance says that the vertical derivative of the horizontal pressure gradient is determined by horizontal density gradients

$$\frac{\partial(\nabla_z p)}{\partial z} = -g \nabla_z \rho. \quad (29.17)$$

Hence, in the presence of horizontal density gradients, the horizontal pressure gradient forces are depth dependent. Correspondingly, the horizontal velocity field experiences a depth dependent pressure force.

We illustrate this depth dependence in Figure 29.3 with a depth independent horizontal density gradient, $\partial \rho / \partial x = \text{constant} < 0$ (see also Figure 27.1), thus leading to a depth dependent horizontal gradient in the hydrostatic pressure. This figure also serves to illustrate how the horizontal pressure gradient can change sign in the vertical, depending on the value of the gradient at depth. It also illustrates how horizontal density gradients lead to a nonzero baroclinicity vector

$$\mathbf{B} = \frac{\nabla \rho \wedge \nabla p}{\rho^2}. \quad (29.18)$$

As shown in Section 48.3, a nonzero baroclinicity imparts a torque on fluid elements that acts as a source of vorticity.

The depth dependence to the horizontal pressure gradient imparts a vertical shear to the horizontal geostrophic velocity

$$\frac{\partial(\rho f \mathbf{u}_g)}{\partial z} = \hat{z} \wedge \nabla(\partial p / \partial z) = -g \hat{z} \wedge \nabla \rho. \quad (29.19)$$

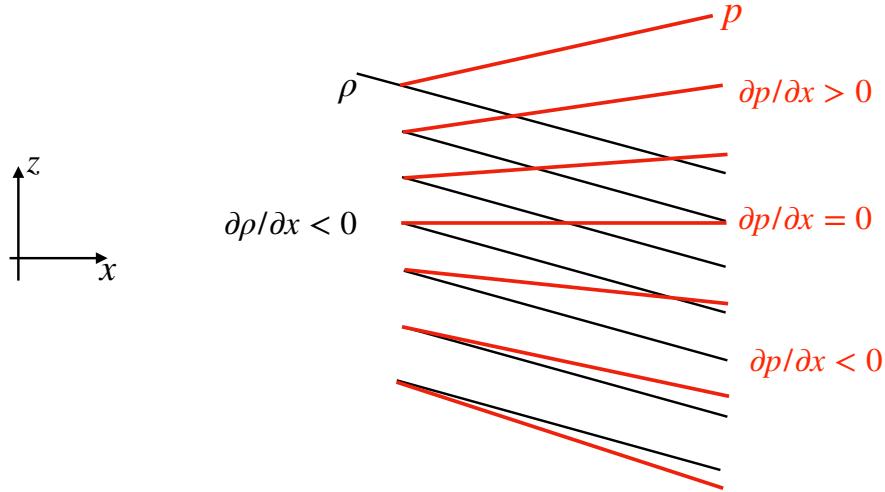


Figure 29.3: Horizontal density gradients support a vertical dependence to the horizontal gradient of the hydrostatic pressure via $\partial(\nabla_z p)/\partial z = -g\nabla_z \rho$. This figure depicts a constant horizontal density gradient with $\partial\rho/\partial x < 0$, thus leading to an increase in the zonal pressure gradient with height, $\partial(\partial p/\partial x)/\partial z > 0$. Depending on the thickness of the fluid column, the horizontal pressure gradient can change sign when moving up in the column, as shown here. Compare this figure to Figure 27.1, which discusses how to compute horizontal pressure gradients in a hydrostatic fluid.

This connection between horizontal density gradients and vertical shears in the geostrophic velocity is known as the *thermal wind relation*, which we return to in Section 29.4.5.

29.3.5 Geostrophic relation in pressure coordinates

The hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho g \quad (29.20)$$

can be used to eliminate density from the geostrophic balance (29.15) to render

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = \frac{g \nabla_z p}{\partial p / \partial z}. \quad (29.21)$$

The right hand side is minus the gradient of the geopotential

$$\Phi = g z \quad (29.22)$$

along surfaces of constant pressure (see Appendix 11.12.2)

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = -\nabla_p \Phi \Rightarrow f \mathbf{u}_g = \hat{\mathbf{z}} \wedge \nabla_p \Phi. \quad (29.23)$$

This is a useful expression of geostrophy for the compressible atmosphere.

29.3.6 Further study

Visualizations from rotating tank experiments provide very useful illustrations of the Coriolis acceleration and geostrophic balance, such as the experiments shown near the 10 minute mark in [this video from Prof. Dave Fultz of the University of Chicago](#).

29.4 Introducing planetary geostrophy

We here introduce the planetary geostrophic (PG) equations, which have found great use in describing elements of the large-scale laminar ocean circulation. We state the equations here and discuss their implications, deferring a more systematic derivation for later, with shallow water PG derived in Section 54.3 and continuously stratified PG derived in Section 55.3.

29.4.1 Planetary geostrophic equations

The governing equations for PG are based on the Boussinesq equations stated in Section 28.1.6, with the assumption of a steady state linear and frictional/geostrophic balance for the horizontal momentum

$$\rho_o f (\hat{z} \wedge \mathbf{u}) = -\nabla p - \rho g \hat{z} + \frac{\partial \boldsymbol{\tau}}{\partial z} \quad (29.24a)$$

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (29.24b)$$

$$\frac{Db}{Dt} = b. \quad (29.24c)$$

Note that the stress $\boldsymbol{\tau}$ is assumed to act just in the horizontal directions. Furthermore, the vertical component of the momentum equation (29.24a) is the hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho g. \quad (29.25)$$

Velocity is slaved to buoyancy

The only time derivative appearing in the PG equations appears in the buoyancy equation (29.24c). All other equations are diagnostic. As the buoyancy evolves, the hydrostatic pressure changes and so too does the geostrophic velocity. In effect, the velocity is a slave to the buoyancy field.

Planetary geostrophy admits no turbulence

The momentum equation is linear since PG drops the nonlinear advection of momentum. Hence, there is no turbulence phenomena in the planetary geostrophic fluid. Instead, planetary geostrophy is used to describe laminar ocean circulation features at the large-scales.

Vertical transfer of horizontal stress and subgrid scale parameterizations

We introduced a horizontal stress (dimensions of force per area) into the momentum equation

$$\boldsymbol{\tau} = (\tau^x, \tau^y, 0). \quad (29.26)$$

This stress is associated with vertical transfer of momentum in the ocean interior through vertical viscosity, as well as vertical transport of momentum from the atmosphere to the ocean.

The vertical stress transport is enhanced by waves and turbulent features such as mesoscale eddies. However, such transient processes are not represented by planetary geostrophy. Hence, they must be parameterized, which generally leads to an enhanced vertical viscosity relative to the molecular value. In general, all models (analytic or numerical) of planetary scale circulations are too coarse to resolve the molecular scales. Consequently, it is necessary to provide rational *subgrid-scale (SGS) parameterizations* of the variety of physical processes that are unresolved by the model.

We have more to say about the parameterization of vertical transfer of horizontal momentum in Section 39.2.4.

29.4.2 Planetary geostrophic vorticity equation

The vertical component of relative vorticity is given by

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \quad (29.27)$$

with a thorough discussion of vorticity given in Chapter 48. Here, we form the relative vorticity budget for the planetary geostrophic system by taking the curl of the momentum equation.

Curl of the PG momentum equation

Taking the curl of the momentum equation (29.24a), and rearranging terms, leads to the planetary geostrophic vorticity equation

$$-\rho_o f \frac{\partial \mathbf{u}}{\partial z} + \hat{z} \rho_o \nabla \cdot (f \mathbf{u}) = -g \nabla \wedge (\hat{z} \rho) + \frac{\partial (\nabla \wedge \boldsymbol{\tau})}{\partial z}. \quad (29.28)$$

Note that $\nabla \cdot (f \mathbf{u}) = \nabla_z \cdot (f \mathbf{u})$ since \mathbf{u} is the horizontal velocity vector. Introducing buoyancy (Section 28.1.2)

$$b = -g \left[\frac{\rho - \rho_o}{\rho_o} \right] \quad (29.29)$$

leads to

$$-f \frac{\partial \mathbf{u}}{\partial z} + \hat{z} \nabla_z \cdot (f \mathbf{u}) = \nabla \wedge (\hat{z} b) + \frac{1}{\rho_o} \frac{\partial (\nabla \wedge \boldsymbol{\tau})}{\partial z}. \quad (29.30)$$

Pressure gradients are removed from the vorticity equation

One of the key reasons to study vorticity is that its evolution equation is not explicitly affected by pressure gradients, since the curl of the pressure gradient vanishes. Eliminating pressure gradients affords a simpler evolution equation for vorticity, and we pursue that evolution more thoroughly in Chapter 48. However, it is important to note that the effects on vorticity from pressure gradients appear indirectly through their effects on velocity, given that vorticity is the curl of velocity. So although pressure is not directly present in the vorticity equation, its effects are nonetheless felt.

Relative vorticity is absent from the PG vorticity equation

It is notable that there is no explicit appearance of the relative vorticity, ζ , in the planetary geostrophic vorticity equation (29.24a). The reason is that we dropped the material time derivative when forming the planetary geostrophic momentum equation (29.24a). By doing so, we drop all expressions of ζ in the vorticity equation. Planetary geostrophy is valid for those cases where

$$|\zeta| \ll |f|, \quad (29.31)$$

which means vorticity is dominated by the planetary vorticity. We encounter more complete versions of the vorticity equation in Chapter 48, where we do not make the planetary geostrophic assumption.

Rather than taking the curl of the planetary geostrophic momentum equation, we could have also derived the vorticity equation (29.30) by taking planetary geostrophic scaling in the full vorticity equation. We choose here the path through the PG momentum equation since we have yet to discuss the full vorticity equation (see Chapter 48).

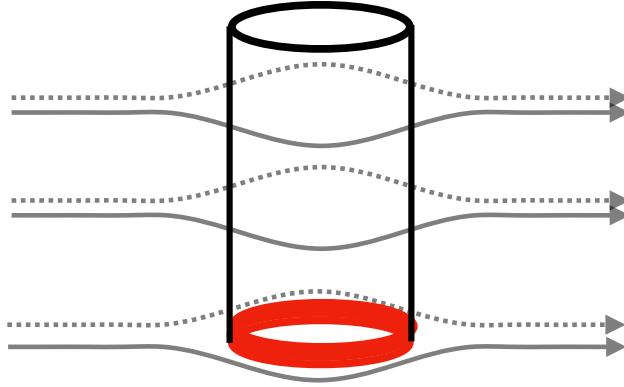


Figure 29.4: The Taylor-Proudman result (29.34) says that horizontal flow in a homogeneous rapidly rotating fluid is depth-independent. Hence, when flow encounters an obstacle anywhere in the column, such as the red ring shown here at the bottom, then flow throughout the full depth must coherently move around the obstacle. The result is a vertically stiffened motion known as *Taylor columns*.

29.4.3 Taylor-Proudman and vertical stiffening

Consider the vorticity equation (29.30) on an f -plane with zero friction, in which the horizontal geostrophic motion is horizontally non-divergent

$$\nabla_z \cdot \mathbf{u} = 0 \quad f\text{-plane geostrophy.} \quad (29.32)$$

Use of continuity (equation (29.24b)) means there is no vertical stretching of a vertical material line element (Section 20.2.5)

$$\frac{\partial w}{\partial z} = 0. \quad (29.33)$$

As shown in Chapter 48, a vortex tube exhibits the same kinematics as a material line element described in Section 20.2. Hence, $\partial w / \partial z = 0$ means there is no vertical stretching of a vortex tube in the planetary geostrophic fluid.

Flat bottom boundary and columnar motion

If there is a solid flat bottom to the domain, then the vertical velocity vanishes at that surface. With $\partial_z w = 0$ in the interior as well, w vanishes throughout the domain. Hence, the fluid has zero vertical velocity, and motion occurs on horizontal planes perpendicular to the rotation axis; i.e., the flow is two-dimensional. We furthermore assume zero horizontal buoyancy gradients, so that the vorticity equation (29.30) implies that the horizontal velocity has zero vertical shear

$$\frac{\partial \mathbf{u}}{\partial z} = 0 \quad f\text{-plane and homogeneous density.} \quad (29.34)$$

This result is known as the Taylor-Proudman theorem.

Relevance to the ocean and atmosphere

In the ocean and atmosphere, the assumptions leading to the Taylor-Proudman theorem are rarely satisfied due to the presence of stratification (i.e., vertical density variations). Nonetheless, there is a tendency for vertical velocities to be quite small due to the effects of rotation; even smaller than

the incompressible scaling $W/H \sim U/L$ would indicate.² Additionally, for unstratified or linearly stratified fluids, there is a tendency for geostrophic turbulence to cascade energy into the gravest (i.e., the largest scale) vertical mode. This largest vertical scale mode is termed the *barotropic* mode, and motion of this mode is predominantly horizontal and depth independent. Smaller vertical scales of variation are captured by an infinite hierarchy of *baroclinic* modes.

29.4.4 Vorticity balance

The vertical component to the vorticity balance (29.30) leads to

$$\nabla_z \cdot (f \mathbf{u}) = \frac{1}{\rho_o} \frac{\hat{z} \cdot \partial(\nabla \wedge \boldsymbol{\tau})}{\partial z} \quad (29.35)$$

which can be written

$$\beta v = -f \nabla_z \cdot \mathbf{u} + \frac{1}{\rho_o} \frac{\hat{z} \cdot \partial(\nabla \wedge \boldsymbol{\tau})}{\partial z} \quad (29.36)$$

where

$$\beta = \frac{\partial f}{\partial y} \quad (29.37)$$

is the gradient of planetary vorticity. The continuity equation (29.24b) can be used to remove the horizontal divergence, which yields the vorticity balance

$$\beta v = f \frac{\partial w}{\partial z} + \frac{1}{\rho_o} \frac{\hat{z} \cdot \partial(\nabla \wedge \boldsymbol{\tau})}{\partial z}. \quad (29.38)$$

The left hand side represents the meridional advection of planetary vorticity. The first term on the right hand size represents the vortex stretching by planetary vorticity; i.e., planetary induction or the β -effect discussed in Section 48.6.2. The second term is the vertical divergence of the curl of the stress. Hence, meridional motion of planetary geostrophic flow is associated with vortex stretching and with the curl of vertical friction. We return to equation (29.38) in Section 55.5 where we consider its depth integrated form. This relatively simple vorticity balance lends great insight into the large-scale ocean circulation, and as such is part of all theories for ocean general circulation.

29.4.5 Thermal wind balance for the ocean

Horizontal components to the inviscid vorticity equation (equation (29.30) with $\boldsymbol{\tau}$ set to zero) form the *thermal wind balance*

$$f \frac{\partial \mathbf{u}}{\partial z} = -\nabla \wedge (\hat{z} b) = \hat{z} \wedge \nabla b, \quad (29.39)$$

which takes on the component form

$$f \frac{\partial u}{\partial z} = -\frac{\partial b}{\partial y} \quad \text{and} \quad f \frac{\partial v}{\partial z} = \frac{\partial b}{\partial x}. \quad (29.40)$$

As seen already in Section 29.3.4, these relations can also be derived direction from taking the vertical derivative of the horizontal momentum equation (29.24a) and then using the horizontal gradient of the hydrostatic balance (29.25). In either case, the thermal wind balance (29.39) says that the horizontal geostrophic velocity possesses a vertical shear where the buoyancy field has a horizontal gradient. Buoyancy with a horizontal gradient is termed *baroclinic*. Correspondingly, it is only the baroclinic (depth dependent) piece of geostrophic velocity that is related to horizontal buoyancy gradients. The depth-independent flow is not constrained by horizontal buoyancy gradients.

²Incompressibility in the form $\partial_x u + \partial_y v + \partial_z w = 0$ leads to the relation $U/L \sim W/H$, where W is a typical vertical velocity scale, H is a typical vertical length scale, and U and L are the corresponding horizontal scales. See Section 27.3.2 for a discussion in the context of the hydrostatic balance.

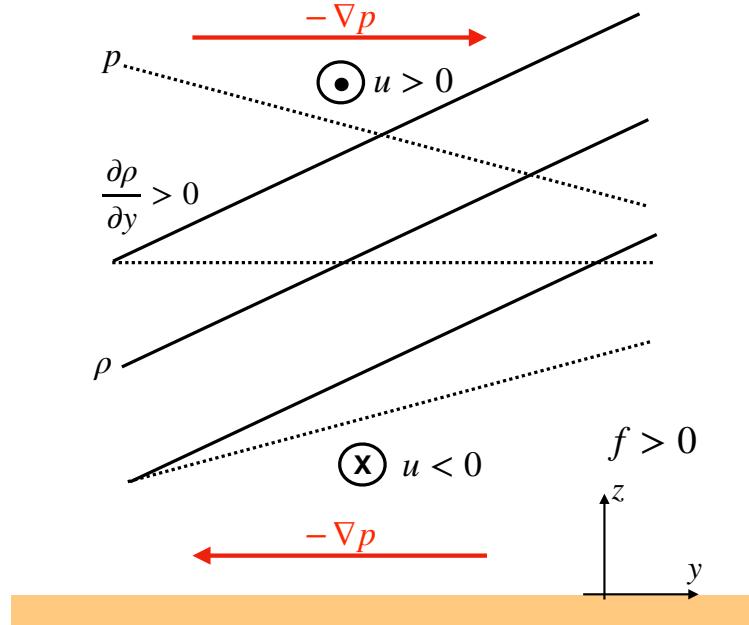


Figure 29.5: Schematic of the density and hydrostatic pressure fields and the associated thermal wind balanced flow in the northern hemisphere ($f > 0$) with north to the right and east out of the page. We show surfaces of constant density (solid lines) and constant pressure (isobars; dashed lines). Density increases poleward ($\partial \rho / \partial y > 0$) so that, according to the discussion surrounding Figure 29.3, the meridional pressure gradient decreases when moving upward, $\partial(\partial p / \partial y) / \partial z < 0$. We illustrate isobars with an equatorward directed down gradient pressure force at lower elevations ($\partial p / \partial y > 0$) and poleward directed pressure force at higher elevations ($\partial p / \partial y < 0$). The zonal geostrophic wind is in geostrophic balance with these pressure gradients, with a westward zonal flow at lower elevations (easterly winds) and eastward flow at higher elevations (westerly winds). This flow configuration creates an eastward vertical shear of the zonal geostrophic winds, $\partial u_g / \partial z > 0$.

Thermal wind, the atmospheric jet stream and the Antarctic Circumpolar Current

Due to the increased solar radiation reaching the equator relative to the poles, the zonal averaged temperature generally reduces poleward. This poleward reduction in temperature corresponds to a poleward reduction in buoyancy. Also, for a stably stratified fluid, density increases with depth. Figure 29.5 illustrates this situation.

The zonal average removes all zonal variations, thus putting $\partial_x \rho = 0$ and so rendering the zonally averaged thermal wind relation

$$f \frac{\partial \bar{u}}{\partial z} = \frac{g}{\rho_o} \frac{\partial \bar{\rho}}{\partial y} = - \frac{\partial \bar{b}}{\partial y} > 0, \quad (29.41)$$

where $(\bar{\ })$ is the zonal mean operator. In the northern hemisphere, $\partial_y \bar{b} < 0$, so that the zonal averaged thermal wind shear is positive, $\partial_z \bar{u} > 0$. For example, a westerly zonal wind (blowing to the east) strengthens with height (easterly thermal wind shear). In the Southern Hemisphere, $f < 0$ with poleward decreasing buoyancy, $\partial_y \bar{b} > 0$, means there is also an eastward thermal wind shear. Note that movement towards the poles, where $|f|$ increases, leads to a smaller thermal wind shear given the same buoyancy gradient.

Diagnosing geostrophic velocity from the buoyancy field

Vertical integration of the thermal wind relation (29.39) between two constant depth surfaces leads to

$$\mathbf{u}(z) = \mathbf{u}(z_{\text{ref}}) - f^{-1} \nabla \wedge \hat{\mathbf{z}} \int_{z_{\text{ref}}}^z b \, dz. \quad (29.42)$$

Hence, knowledge of the buoyancy field (e.g., through hydrographic measurements of temperature and salinity in the ocean), along with knowledge of the geostrophic velocity at a single point along the integration path, allows for determination of the full geostrophic velocity in terms of density. Unfortunately, specification of the unknown reference velocity is unavailable just from hydrographic measurements. This is the origin of the “depth of no motion” problem in diagnostic oceanography.

29.4.6 Thermal wind balance for the atmosphere

The large-scale atmosphere is compressible and predominantly in hydrostatic balance. The expression for geostrophic balance (29.23) in pressure coordinates is a suitable starting point to derive thermal wind for the atmosphere. For this purpose, we take the pressure derivative, $\partial/\partial p$, of (29.23) to render

$$f \frac{\partial \mathbf{u}}{\partial p} = \hat{\mathbf{z}} \wedge \nabla_p \left[\frac{\partial \Phi}{\partial p} \right]. \quad (29.43)$$

The hydrostatic relation $\partial p/\partial z = -\rho g$ takes the form

$$\frac{\partial p}{\partial \Phi} = -\rho \Rightarrow \frac{\partial \Phi}{\partial p} = -1/\rho \quad (29.44)$$

in which case

$$f \frac{\partial \mathbf{u}}{\partial p} = -\hat{\mathbf{z}} \wedge \nabla_p (1/\rho). \quad (29.45)$$

Ideal gas atmosphere

The specific volume takes the following form for an ideal gas atmosphere (see Section 23.5.1)

$$\rho^{-1} = \frac{R^M T}{p}. \quad (29.46)$$

Since the horizontal derivative in the thermal wind relation (29.45) is along pressure surfaces, we have

$$f \frac{\partial \mathbf{u}}{\partial p} = -\frac{R^M}{p} [\hat{\mathbf{z}} \wedge \nabla_p T]. \quad (29.47)$$

This expression gives rise to the name “thermal wind”, with vertical shears of the horizontal velocity generated by horizontal temperature gradients.

As for the ocean in equation (29.42), we vertically integrate the thermal wind expression (29.47), only now do so between two pressure levels

$$\mathbf{u}(p_a) - \mathbf{u}_g(p_b) = f^{-1} R^M \hat{\mathbf{z}} \wedge \nabla_p \left[\int_{p_a}^{p_b} \frac{T \, dp}{p} \right], \quad (29.48)$$

where $p_a < p_b$, so that p_a has a higher altitude than p_b . We define the thermal wind shear as the difference between the wind aloft (higher altitude and lower pressure) from that at a lower altitude (greater pressure)

$$\mathbf{u}_T = \mathbf{u}(p_a) - \mathbf{u}(p_b) \quad \text{with } p_a < p_b \quad (29.49)$$

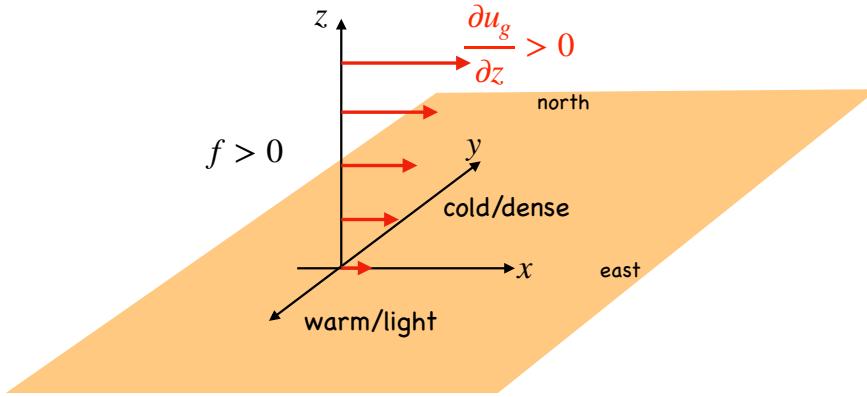


Figure 29.6: Thermal wind shear in the northern hemisphere ($f > 0$) middle latitude atmosphere, whereby cold/dense air sits to the north and warm/light air to the south. The zonal geostrophic winds, u_g , increase to the east when rising in elevation, $\partial u_g / \partial z > 0$. We say that the zonal winds have an eastward thermal wind shear. In general, a geostrophic wind in the northern hemisphere atmosphere has cold/dense air to the left when facing downwind, whereas the opposite orientation holds for the southern hemisphere where $f < 0$.

so that

$$\mathbf{u}_T = \frac{R^M}{f} \hat{\mathbf{z}} \wedge \nabla_p \bar{T}^{\ln p}, \quad (29.50)$$

where we introduced the log-pressure weighted temperature between the two pressure surfaces

$$\bar{T}^{\ln p} = \int_{p_a}^{p_b} \frac{T dp}{p}. \quad (29.51)$$

The relation (29.50) means that on the f -plane, R^M/f times the log-pressure weighted temperature serves as a streamfunction for the thermal wind shear. Reconsider the previous example where the polar regions are colder than tropics, so that in the northern hemisphere on pressure surfaces, $\partial \bar{T}^{\ln p} / \partial y < 0$. Hence, the zonal westerly winds increase in magnitude with height. Furthermore, the thermal wind shear points to the east. In general for the northern hemisphere, cold air sits on the left side of the thermal wind shear and warm air on the right. The opposite orientation holds for the Southern Hemisphere where the Coriolis parameter is negative, $f < 0$

Barotropic flow

Return to the thermal wind equation (29.45)

$$f \frac{\partial \mathbf{u}}{\partial p} = -\hat{\mathbf{z}} \wedge \nabla_p \alpha = \frac{\hat{\mathbf{z}} \wedge \nabla_p \rho}{\rho^2}. \quad (29.52)$$

For the special case of density a function just of the pressure, $\rho = \rho(p)$, then $\nabla_p \rho = 0$. This situation defines a *barotropic* flow, which is characterized here by a horizontal geostrophic velocity with zero vertical variations. Note that we are here only concerned with the geostrophic flow. A density related to pressure through $\rho = \rho(p)$ can still support vertical variations of the ageostrophic flow.

We further discuss barotropic flow in Section 48.2 as part of our study of vorticity. As detailed in that discussion, the general definition of a barotropic flow is one whereby the *baroclinicity* vector vanishes, $\mathbf{B} = \nabla p \wedge \nabla \alpha = 0$. The functional relation $\rho = \rho(p)$ (equivalently $p = p(\rho)$) is a sufficient condition for vanishing baroclinicity. As seen in Section 48.3, for a barotropic flow there is no generation of vorticity through the torques created when density isosurfaces (isopycnals) deviate from pressure isosurfaces (isobars).

29.4.7 Further study

Much of the material in this section forms the basis for laminar theories of the large-scale ocean circulation. Many of the concepts are detailed in Chapter 7 of [Marshall and Plumb \(2008\)](#). Chapters 19-22 of [Vallis \(2017\)](#) as well as [Samelson \(2011\)](#) present ocean circulation theory making use of fundamental concepts of geophysical fluid dynamics. A compelling discussion of the cascade of energy from the baroclinic modes to barotropic mode is offered by [Smith and Vallis \(2001\)](#). [Gill \(1982\)](#) provides a discussion of the depth of no motion problem in dynamic oceanography.

Rotating tank laboratory experiments offer a powerful means to explore the variety of rotating fluid mechanics relevant to the atmosphere and oceans. The following resources exemplify the Taylor-Proudman result (29.34) and the associated vertical stiffening of rotating fluids.

- One means to test Taylor-Proudman is to insert a dye into a rapidly rotating tank of unstratified water. After a few rotation periods the dye forms vertical sheets known as “Taylor curtains” whose center is along the rotation axis. The fluid is said to have a “vertical stiffness” due to the effects of rotation. Vertical stiffening in turn means that flow over a small obstacle is deflected throughout the column rather than just near the bump. This situation is depicted in Figure 29.4 and more vividly illustrated in [this video from the UCLA SpinLab](#).
- Near the 20 minute mark of [this video, also from UCLA](#), we see how vortices in a rotating fluid maintain their coherency much more than in a non-rotating fluid.
- Another laboratory test shown in [this video from Prof. Dave Fultz of the University of Chicago](#) shows that a buoyant object (a ping pong ball) placed into a rotating tank rises much slower than in a non-rotating tank. The reason for the slower rise is that the ball must push up against the vertically stiffened fluid column when rotating, thus slowing its rise relative to when in a non-rotating fluid. Later in the same video, near the 16 minute mark, shows how Taylor curtains arise in rotating fluids.

29.5 Isopycnal form stress from geostrophic eddies

As introduced in Section 26.8, form stress is the horizontal stress arising from pressure acting on a sloped surface. The mathematical expression for the form stress acting on the top side of a surface is given by equation (26.60)

$$\Sigma^{\text{form}} = p \nabla \eta, \quad (29.53)$$

with the opposite sign for the form stress on the bottom side of a surface. Here, $z = \eta(x, y, t)$ is the depth of the surface (see Figure 26.6 or Figure 29.7). The net horizontal force from form stress is the area integral over the surface.

In this section we examine the zonal mean zonal form stress acting on an isopycnal surface (Section 29.5.1) and on an isopycnal layer (Section 29.5.2), each for an adiabatic, Boussinesq, hydrostatic fluid in geostrophic balance and within a zonally periodic channel of length L . As we show, the zonal mean zonal form stress arising from deviations from the zonal mean; i.e., zonal fluctuations associated with waves and geostrophic turbulence, provides an eastward acceleration to the fluid while transporting buoyancy and thickness/volume meridionally. Although the channel geometry is rather simple, it has applications to the middle latitude atmospheric circulation as well as for ocean circulation, particularly in the Southern Ocean where there is circumpolar channel-like flow within the Antarctic Circumpolar Current. Furthermore, the discussion exposes key elements of eddy-mean flow interactions, sharing points with the leading order generalized Lagrangian mean of Section 38.2 and the quasi-Stokes transport discussed in Section 39.2.

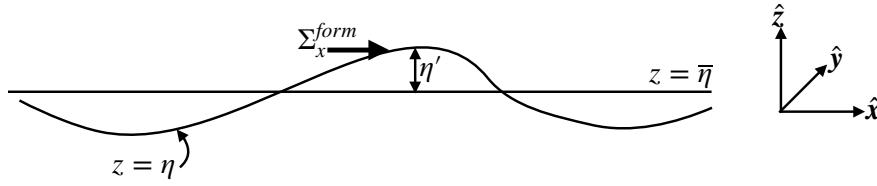


Figure 29.7: Schematic of the zonal form stress, Σ_x^{form} , acting on a surface whose zonal mean vertical position is $z = \bar{\eta}(y, t)$ and whose vertical position relative to the zonal mean is $z = \bar{\eta}(y, t) + \eta'(x, y, t)$.

29.5.1 Zonal mean zonal form stress on an isopycnal surface

We are here interested in the form stress acting on an isopycnal surface. Before specializing to an isopycnal, we decompose the form stress according to the zonal mean depth and its deviation from zonal mean (see Figure 29.7). Thereafter, specialization to an isopycnal surface in an adiabatic fluid connects the zonal mean form stress to the meridional eddy flux of buoyancy.

Zonal form stress on an arbitrary surface in a channel

The zonal mean vertical position of the surface is written

$$\bar{\eta} = \frac{1}{L} \int_0^L \eta \, dx \quad (29.54)$$

and its corresponding zonal fluctuation is

$$\eta' = \eta - \bar{\eta}. \quad (29.55)$$

The zonal component of the form stress is thus given by

$$p \partial_x \eta = p(x, \bar{\eta} + \eta') \partial_x (\bar{\eta} + \eta') \quad (29.56a)$$

$$= p(x, \bar{\eta} + \eta') \partial_x \eta' \quad (29.56b)$$

$$\approx [p(x, \bar{\eta}) + \partial_z p(x, \bar{\eta}) \eta'] \partial_x \eta' \quad (29.56c)$$

$$= p(x, \bar{\eta}) \partial_x \eta' + \mathcal{O}(\eta')^2. \quad (29.56d)$$

Hence, to second order in fluctuations, η' , the zonal form stress acting on the surface equals to $p(x, \bar{\eta}) \partial_x \eta'$, where it is important to note that pressure is evaluated at the zonal mean depth, $z = \bar{\eta}$.

To within the same accuracy, the zonal integral of the zonal form stress is given by

$$\int_0^L \Sigma_x^{form} \, dx \approx \int_0^L p(\bar{\eta}) (\partial \eta' / \partial x) \, dx = - \int_0^L \eta' [\partial p(\bar{\eta}) / \partial x] \, dx, \quad (29.57)$$

where the final equality follows from zonal periodicity. Now assume the zonal pressure gradient at $\bar{\eta}$ is balanced by a meridional geostrophic velocity

$$\partial p(\bar{\eta}) / \partial x = f \rho_0 v(\bar{\eta}), \quad (29.58)$$

where

$$v(\bar{\eta}) = \bar{v}(\bar{\eta}) + v'(\bar{\eta}) \quad (29.59)$$

so that

$$\int_0^L \Sigma_x^{form} \, dx = -\rho_0 f \int_0^L \eta' v' \, dx, \quad (29.60)$$

where we noted that the Coriolis parameter is independent of zonal position. Hence, there is a nonzero zonal mean zonal form stress when there is a nonzero zonal correlation between fluctuations in the meridional velocity and the depth of the surface

$$\bar{\Sigma}_x^{\text{form}} = -\rho_0 f \bar{v}' \bar{\eta}'. \quad (29.61)$$

Zonal mean zonal form stress acting on an isopycnal surface

To further unpack the correlation appearing in equation (29.61), specialize to the case of an isopycnal surface in an adiabatic fluid. As shown in Sections 38.2.6 and 38.4.7, vertical fluctuations in the position of the isopycnal surfaces, relative to the zonal mean $\bar{\eta}$, are related to zonal fluctuations in the density

$$\eta' \approx -\frac{\rho'}{\partial \bar{\rho} / \partial z} = -\frac{b'}{N^2}, \quad (29.62)$$

where we introduced the squared buoyancy frequency of the zonal mean state as well as the fluctuating buoyancy

$$N^2 = -\frac{g}{\rho_0} \frac{\partial \bar{\rho}}{\partial z} \quad \text{and} \quad b' = -\frac{g \rho'}{\rho_0}. \quad (29.63)$$

The zonally averaged zonal form stress thus takes the form

$$\bar{\Sigma}_x^{\text{form}} = \frac{\rho_0 f}{N^2} \bar{v}' \bar{b}'. \quad (29.64)$$

Again, the assumptions rendering the result (29.64) are (i) zonal periodicity, (ii) adiabatic and Boussinesq fluid, (iii) geostrophically balanced flow. Under these assumptions, the zonal mean zonal form stress acting on an isopycnal surface is proportional to the zonal correlation between fluctuations in the meridional velocity and the buoyancy. It is a general statistical property of geostrophic eddies in the atmosphere and ocean to transport positive buoyancy (e.g., warm air/water) poleward and negative buoyancy (e.g., cold air/water) equatorward, thus ameliorating the equator-to-pole buoyancy difference setup by solar radiation that preferentially warms the tropics. In turn, this property of geostrophic eddies leads to a positive zonal mean zonal form stress

$$\bar{\Sigma}_x^{\text{form}} > 0. \quad (29.65)$$

Hence, in addition to transporting buoyancy poleward, geostrophic eddies provide a positive zonal mean force through zonal integrated form stress that accelerates the fluid in the eastward direction. These two properties of geostrophic eddies (poleward flux of positive buoyancy anomalies along with an eastward acceleration from form stress) are fundamental to the middle latitude atmospheric circulation as well as for ocean circulation, particularly within the channel-like Antarctic Circumpolar Current.

29.5.2 Zonal mean zonal form stress acting on an isopycnal layer

We offer yet another means to understand the zonal mean zonal form stress by here examining the form stress acting on a constant density layer of adiabatic Boussinesq fluid such as shown in Figure 29.8. This layered/isopycnal analysis anticipates some of the development considered for the stacked shallow water model in Chapters 42 and 43 as well as for isopycnal models in Chapter 45.

The net form stress acting on the upper and lower layer interfaces in Figure 29.8 is given by

$$\Sigma^{\text{layer form}} = p_1 \nabla \eta_1 - p_2 \nabla \eta_2 \quad (29.66\text{a})$$

$$= p(\eta + h/2) \nabla(\eta + h/2) - p(\eta - h/2) \nabla(\eta - h/2) \quad (29.66\text{b})$$

$$\approx [p(\eta) - \rho g h/2] \nabla(\eta + h/2) - [p(\eta) + \rho g h/2] \nabla(\eta - h/2) \quad (29.66\text{c})$$

$$= p \nabla h - \rho g h \nabla \eta \quad (29.66\text{d})$$

$$= \nabla(p h) - h \nabla(p + \rho g \eta) \quad (29.66\text{e})$$

$$= \nabla(p h) - \rho_0 h \nabla M. \quad (29.66\text{f})$$

In this relation we set $z = \eta$ for the vertical position at the center of the layer, introduced the Montgomery potential from Section 45.1.1

$$M \rho_0 = p + \rho g \eta, \quad (29.67)$$

and noted that ρ is a uniform constant layer density so that it commutes with the horizontal gradient operator computed along ρ surfaces. We also made use of the hydrostatic balance to approximate the interface pressures as

$$p(\eta + h/2) \approx p(\eta) + \frac{\partial p}{\partial z} \frac{h}{2} = p(\eta) - \rho g h/2 \quad (29.68\text{a})$$

$$p(\eta - h/2) \approx p(\eta) - \frac{\partial p}{\partial z} \frac{h}{2} = p(\eta) + \rho g h/2. \quad (29.68\text{b})$$

The zonal mean of the zonal layer form stress is thus given by the correlation between the layer thickness fluctuations and fluctuations in the zonal derivative of the Montgomery potential

$$\bar{\Sigma}_x^{\text{layer form}} = -\rho_0 \bar{h'} \overline{\partial M' / \partial x}, \quad (29.69)$$

where we set $\overline{\partial M / \partial x} = 0$ due to zonal periodicity. As seen in Section 45.1.1, the Montgomery potential is the geostrophic streamfunction in isopycnal coordinates, so that the fluctuating meridional geostrophic velocity is given by

$$f v' = \partial_x M'. \quad (29.70)$$

Consequently, the zonal mean zonal form stress acting on the layer equals to the correlation between the thickness fluctuations and fluctuations in the meridional geostrophic velocity

$$\bar{\Sigma}_x^{\text{layer form}} = -\rho_0 f \overline{v' h'}. \quad (29.71)$$

Hence, as the geostrophic eddies provide a net eastward acceleration to the layer (equation (29.65)), they also move volume around within isopycnal layers meridionally, moving positive thickness fluctuations equatorward.

To further understand the physics of the form stress in equation (29.71), parameterize the velocity-thickness eddy correlation, $\overline{v' h'}$, by downgradient diffusion of thickness

$$\overline{v' h'} = -\kappa \partial_y \bar{h}, \quad (29.72)$$

where $\kappa > 0$ is a nonzero kinematic diffusivity (dimensions of squared length per time). This parameterization is suggested by the work of [Gent and McWilliams \(1990\)](#) as discussed in Section 39.2.5. As noted there, thickness diffusion as a parameterization reflects the general tendency of

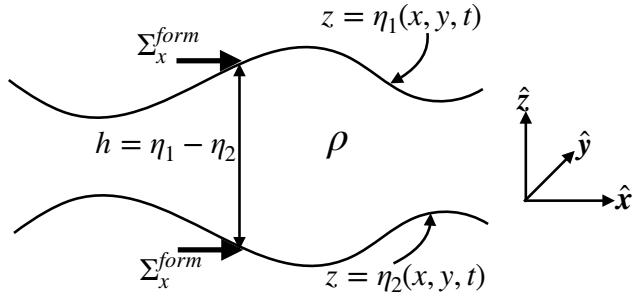


Figure 29.8: Schematic of a constant density layer of an adiabatic, hydrostatic, Boussinesq fluid with thickness $h(x, y, t) = \eta_1(x, y, t) - \eta_2(x, y, t) = (\eta + h/2) - (\eta - h/2)$, and uniform density $\rho = \text{constant}$. East points to the right and north is oriented into the page. The zonal form stress, Σ_x^{form} , acting on the upper and lower interfaces at a horizontal position (x, y) are shown by the thick horizontal vectors. The zonal form stress is the horizontal component of the compressive pressure force per area acting on the layer interfaces, with the sign of the form stress determined by the slope of the layer interface. For a zonally periodic fluid layer, the net zonal pressure force acting on the layer arises from the zonal form stress integrated over the layer interfaces.

geostrophic eddies to reduce horizontal gradients in layer thickness as they reduce the available potential energy of the flow. In this case the zonal mean zonal form stress is

$$\bar{\Sigma}_x^{\text{layer form}} = \rho_0 f \kappa \partial_y \bar{h}. \quad (29.73)$$

So in the northern hemisphere in regions where the zonal mean layer thickness increases to the north, $\partial_y \bar{h} > 0$, there is a corresponding eastward zonal mean zonal form stress arising from parameterized geostrophic eddies acting on layer thickness. This situation corresponds to the case in Section 29.5.1, where we saw that geostrophic eddies preferentially transport positive buoyancy anomalies poleward and negative buoyancy anomalies equatorward. In the present analysis, meridional changes to the layer thickness correspond to a nonzero thermal wind shear. If layer thickness increases poleward, as for the case of weaker vertical stratification in the high latitudes, then the zonal velocity has a positive vertical shear, thus contributing an eastward zonal mean form stress.

29.5.3 Comments and further study

A key feature of geostrophic eddies exposed by this discussion concerns the connection between zonal form stress (providing an eastward force on the zonally periodic channel flow) and meridional eddy transport of buoyancy (positive buoyancy anomalies are transported poleward) and thickness (positive thickness anomalies are transported equatorward). The periodic channel domain is highly idealized. Nonetheless, the basic ideas form the roots for much of how we think about geostrophic eddies in the middle latitude atmosphere and the Southern Ocean. Further generalizations lead to the generalized Lagrangian mean, whose kinematic rudiments are discussed in Section 38.2.

The fundamental role of form stress in geostrophic turbulent flows is pedagogically treated by [Vallis \(2017\)](#). See, in particular, his Chapter 21 for a thorough and insightful discussion of circulation in the Southern Ocean. We also return to form stress within the shallow water fluid in Section 43.2. That discussion complements the presentation given here, there focusing exclusively on the case of a layer of shallow water fluid. We also touch on the notions of form stress when discussing the [Gent and McWilliams \(1990\)](#) mesoscale eddy parameterization in Section 39.2.

29.6 Exercises

EXERCISE 29.1: SMALL ROSSBY NUMBER AT HUMAN SCALES

Consider motion of a car at a speed $U \sim 10^5$ m hour $^{-1}$ and a length scale of $L \sim 10$ m.

- (a) What is the rotation period required render a unit Rossby number for the given scales? Give result in units of seconds.
- (b) If the earth rotated at the angular speed Ω_{human} , what would be the solid-body speed for a point at rest on the earth's surface? Give result in units of meter per second.
- (c) How does the solid-body speed compare to the speed of sound at standard atmospheric conditions? What about the root-mean-square speed for air molecules? Hint: read Section 2.2.4.
- (d) Discuss one or two astronomical objects that have very large rotational speeds.

EXERCISE 29.2: THE BETA SPIRAL

Consider a steady state Boussinesq planetary geostrophic fluid in the absence of mixing. Write the geostrophic velocity as

$$u = |\mathbf{u}| \cos \Delta \quad v = |\mathbf{u}| \sin \Delta, \quad (29.74)$$

where Δ is the angle measured counter-clockwise from east. Use thermal wind and the steady state perfect fluid buoyancy equation to determine an expression for $\partial \Delta / \partial z$. Show that for $f > 0$ (northern hemisphere) and $\partial b / \partial z = N^2 > 0$ (gravitationally stable fluid column), then $\partial \Delta / \partial z$ has opposite sign from the vertical velocity, w . This spiralling of the geostrophic velocity is known as the *beta spiral* in oceanography.

EXERCISE 29.3: ALTERNATIVE FORM OF THERMAL WIND

Consider a fluid with density a function of pressure and potential temperature

$$\rho = \rho(p, \theta). \quad (29.75)$$

A physical realization of this equation of state is a lake. Show that the thermal wind shear for a hydrostatic and *compressible* fluid with this equation of state can be written in the form

$$\frac{\partial \mathbf{u}}{\partial z} = \left[\frac{N^2}{f \rho g} \right] (\hat{\mathbf{z}} \wedge \nabla_\theta p), \quad (29.76)$$

where

$$N^2 = -\frac{g}{\rho} \frac{\partial \rho}{\partial \theta} \frac{\partial \theta}{\partial z} = g \beta_\theta \frac{\partial \theta}{\partial z} > 0 \quad (29.77)$$

is the squared buoyancy frequency, assumed positive so that the fluid is gravitationally stable in the vertical (see Section 25.3.4). The term β_θ is the thermal expansion coefficient written in terms of potential temperature (Section 25.2.5). Finally, the horizontal gradient projected onto constant θ surfaces is given by (see Section 11.12.2)

$$\nabla_\theta = \hat{\mathbf{x}} \left[\frac{\partial}{\partial x} \right]_{y,\theta} + \hat{\mathbf{y}} \left[\frac{\partial}{\partial y} \right]_{x,\theta} \quad (29.78a)$$

$$= \nabla_z - \left[\frac{\nabla_z \theta}{\partial \theta / \partial z} \right] \frac{\partial}{\partial z}. \quad (29.78b)$$

Hint: This exercise requires careful use of the chain rule and the hydrostatic relation, along with the equations given in the problem statement. Furthermore, assume the fluid is fully compressible.

Hint: Some may wish to “warm-up” by showing that the result holds for the simpler equation of state $\rho = \rho(\theta)$. Some of the steps used for the simpler case are relevant for the case with $\rho = \rho(\theta, p)$.

30

Balanced inviscid horizontal flows

We here consider a variety of inviscid horizontal flow regimes characterized by a balance between a subset of terms appearing in the horizontal momentum equation. This discussion allows us to compare the geostrophic flow of Chapter 29 to a variety of *ageostrophic flows* such as gradient wind, inertial motion, and cyclostrophic balance. We offer a categorization of the flow following the *natural coordinates* used by [Holton \(1992\)](#). Natural coordinates provide a lucid means to compare the relative magnitudes of the Coriolis, pressure, and centrifugal accelerations acting on a fluid particle moving horizontally.

READER'S GUIDE TO THIS CHAPTER

This chapter assumes an understanding of geostrophic flow from Chapter 29. Some of this material is used in subsequent chapters, in particular Chapter 31 on Ekman layer dynamics as well as Chapters 54 and 56 on quasi-geostrophy. We make use of some geometry discussed in Chapter 6, though most of the salient points are revisited here so that Chapter 6 is an option rather than a requirement.

30.1	Horizontal flow described by natural coordinates	452
30.1.1	Natural coordinates	452
30.1.2	Material acceleration	453
30.1.3	Centripetal and centrifugal accelerations	454
30.1.4	Coriolis and pressure gradient	454
30.1.5	Horizontal momentum equation	455
30.1.6	An Eulerian decomposition of the acceleration	455
30.1.7	Further study	456
30.2	Exact geostrophic balance	456
30.2.1	Steady f -plane flow	457
30.2.2	Steady flow on a sphere	457
30.2.3	What about geostrophic balance with curvature?	458
30.3	Inertial motion of fluid particles	458
30.3.1	Anti-cyclonic circular motion	458
30.3.2	Balance between Coriolis and centrifugal	459
30.3.3	Period for inertial motion	459
30.3.4	Observing inertial motion	459
30.3.5	Inertial motion is Lagrangian	460
30.3.6	“Inertial” motion is not motion viewed from an inertial frame	460
30.4	Cyclostrophic balance	460
30.5	Gradient wind balance	461
30.5.1	Constraints on gradient wind flow	461
30.5.2	The variety of gradient wind flows	462
30.5.3	Comments	463

30.1 Horizontal flow described by natural coordinates

In this section we decompose the horizontal Boussinesq momentum equation into motion parallel to and motion perpendicular to the instantaneous trajectory of a fluid particle. That is, we characterize the velocity and acceleration according to the local flow direction. Furthermore, we are only concerned with horizontal motion (on a constant geopotential). Using the “natural coordinates” arising from this description we encounter the centripetal/centrifugal acceleration that arises from curvature in the trajectory. Note that this acceleration is distinct from the centrifugal acceleration that arises from planetary rotation and is contained within the effective gravitational acceleration acting in the local vertical (see Section 13.10.1). We also decompose the accelerations from pressure, friction, and Coriolis into natural coordinate directions. The kinematics afforded by natural coordinates succinctly summarizes the horizontal motion and suggests a fruitful categorization of balances arising in limiting cases.

30.1.1 Natural coordinates

Natural coordinates for horizontal motion are defined by a locally orthogonal set of unit vectors (see Figure 30.1)

$$\hat{z} = \hat{u} \wedge \hat{n} = \text{vertical direction} \tag{30.1a}$$

$$\hat{s} = \hat{n} \wedge \hat{z} = \text{tangent to velocity} \tag{30.1b}$$

$$\hat{n} = \hat{z} \wedge \hat{u} = \text{normal direction.} \tag{30.1c}$$

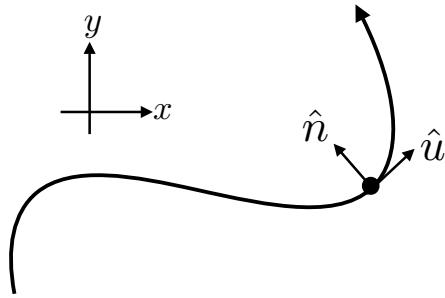


Figure 30.1: Illustrating the decomposition of horizontal motion of a fluid particle into natural coordinate directions. These directions are defined by a unit tangent vector, \hat{u} , pointing in the direction of the fluid particle motion, and a unit normal vector, \hat{n} , pointing to the left of the motion facing downstream.

The unit vector \hat{u} is tangent to the velocity vector (which is horizontal), so that

$$\mathbf{u} = |\mathbf{u}| \hat{\mathbf{u}} = \frac{D\mathbf{s}}{Dt} \hat{\mathbf{u}}, \quad (30.2)$$

where s is the arc-length measured along the trajectory as introduced in Section 4.4. The unit vector \hat{n} is perpendicular to the velocity and points to the left of the trajectory facing downstream.

30.1.2 Material acceleration

When writing the velocity according to equation (30.2), we decompose the acceleration into the change in speed and change in direction

$$\frac{D\mathbf{u}}{Dt} = \frac{D|\mathbf{u}|}{Dt} \hat{\mathbf{u}} + |\mathbf{u}| \frac{D\hat{\mathbf{u}}}{Dt}. \quad (30.3)$$

Following our discussion of rotation in Section 13.5 (see Figure 13.3), the magnitude of the direction change can be written in terms of the infinitesimal angle swept out by the motion as the fluid particle moves along a trajectory

$$|\delta\hat{\mathbf{u}}| = \delta\vartheta. \quad (30.4)$$

The infinitesimal angle swept out by the trajectory is related to the radius of curvature, R (Figure 30.2), and the arc-length increment, δs , traversed by the trajectory

$$\delta\vartheta = \frac{\delta s}{R}. \quad (30.5)$$

Finally, the infinitesimal change in tangent, $\delta\hat{\mathbf{u}}$, is directed normal to the motion, and is thus directed along the \hat{n} direction. We see this property by noting that

$$\hat{\mathbf{u}} \cdot \hat{\mathbf{u}} = 1 \Rightarrow \delta\hat{\mathbf{u}} \cdot \hat{\mathbf{u}} = 0. \quad (30.6)$$

That is, $\delta\hat{\mathbf{u}}$ is orthogonal to $\hat{\mathbf{u}}$, so that it points parallel or anti-parallel to \hat{n} . We detailed this property of rotating unit vectors in Section 4.1.4 (see Figure 4.2). Our convention is that \hat{n} points to the left of $\hat{\mathbf{u}}$, so that if the trajectory turns to the left, then $\delta\hat{\mathbf{u}}$ points parallel to \hat{n} , whereas if the trajectory turns to the right then $\delta\hat{\mathbf{u}}$ points anti-parallel to \hat{n} . That is, $\delta\hat{\mathbf{u}}$ always points towards the center of the circle used to compute the radius of curvature as in Figure 30.2.

Bringing these results together leads to the expression for the infinitesimal unit vector change

$$\delta\hat{\mathbf{u}} = \hat{\mathbf{n}} \frac{\delta s}{R}. \quad (30.7)$$

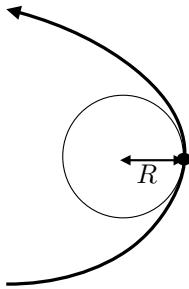


Figure 30.2: Illustrating the radius of curvature associated with turning motion of a fluid particle. The radius of curvature equals to the radius of a tangent circle (the curvature circle) that approximates, to second order accuracy, the trajectory at a particular point. The radius is smaller in magnitude when the trajectory is highly curved, and $|R| = \infty$ when the trajectory is straight. The radius is positive when the trajectory turns into the normal direction as depicted here (to the left; concave as defined by $\hat{\mathbf{n}}$) and negative when turning in the opposite direction (to the right; convex as defined by $\hat{\mathbf{n}}$). See Section 6.2 for more details on curvature, with Figure 6.4 offering more details for how to determine the radius of curvature.

Again, our sign convention takes $R > 0$ for a particle turning in the direction of $\hat{\mathbf{n}}$ (to the left facing downstream) and $R < 0$ for a particle turning opposite to $\hat{\mathbf{n}}$ (to the right facing downstream). Hence, the material time change is

$$\frac{D\hat{\mathbf{u}}}{Dt} = \frac{D\hat{\mathbf{u}}}{Ds} \frac{Ds}{Dt} = \frac{\hat{\mathbf{n}}}{R} \frac{Ds}{Dt} = \frac{\hat{\mathbf{n}}}{R} |\mathbf{u}|, \quad (30.8)$$

where the speed is given by the time change of the arc-length along the trajectory

$$|\mathbf{u}| = \frac{Ds}{Dt}. \quad (30.9)$$

Combining these results renders the acceleration

$$\frac{D\mathbf{u}}{Dt} = \frac{D|\mathbf{u}|}{Dt} \hat{\mathbf{u}} + \frac{|\mathbf{u}|^2}{R} \hat{\mathbf{n}}. \quad (30.10)$$

The acceleration has thus been decomposed into the change in speed of the fluid particle along the direction of the motion, plus the centripetal acceleration due to curvature of the trajectory. In Section 30.1.3 we justify referring to $(|\mathbf{u}|^2/R) \hat{\mathbf{n}}$ as the *centripetal* acceleration.

30.1.3 Centripetal and centrifugal accelerations

The centripetal acceleration points towards the concave side of a turning particle trajectory; “centripetal” means “towards the center.” Its opposing partner, the centrifugal (“away from center”) acceleration points to the convex side (see Figure 30.3). So how do we interpret $\hat{\mathbf{n}} |\mathbf{u}|^2/R$? For motion turning to the left, towards $\hat{\mathbf{n}}$, the radius of curvature is positive, $R > 0$, so that $\hat{\mathbf{n}} |\mathbf{u}|^2/R$ points to the concave side of the trajectory (left side). If the particle is turning to the right then $R < 0$, which again means that $\hat{\mathbf{n}} |\mathbf{u}|^2/R$ points to the concave side (now on the right). We conclude that the acceleration $\hat{\mathbf{n}} |\mathbf{u}|^2/R$ indeed represents a centripetal acceleration and $-\hat{\mathbf{n}} |\mathbf{u}|^2/R$ is the centrifugal acceleration.

30.1.4 Coriolis and pressure gradient

The Coriolis acceleration takes the following form in natural coordinates

$$-f \hat{\mathbf{z}} \wedge \mathbf{u} = -(\hat{\mathbf{z}} \wedge \hat{\mathbf{u}}) f |\mathbf{u}| \quad (30.11a)$$

$$= -\hat{\mathbf{n}} f |\mathbf{u}|, \quad (30.11b)$$

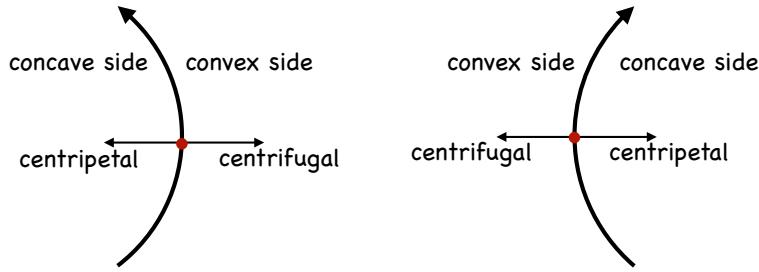


Figure 30.3: Centripetal acceleration of a turning fluid particle, $\hat{n}|\mathbf{u}|^2/R$, points to the concave side of the curve (towards the center) whereas the centrifugal acceleration, $-\hat{n}|\mathbf{u}|^2/R$, points to the convex side (away from the center). The centripetal and centrifugal accelerations are paired action/reaction accelerations. For a left turning trajectory, the concave side is on the left and has positive radius of curvature, $R > 0$, whereas for the right turning trajectory the concave side is to the right with $R < 0$. To help remember the signs, note that centrifugal means “away from the center” whereas centripetal means “towards the center”. 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.

so that the Coriolis acceleration always points to the right of the flow direction. In contrast, the pressure gradient has two components

$$\nabla p = \hat{\mathbf{u}} (\hat{\mathbf{u}} \cdot \nabla p) + \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \nabla p) \quad (30.12a)$$

$$= \hat{\mathbf{u}} \frac{\partial p}{\partial s} + \hat{\mathbf{n}} \frac{\partial p}{\partial n}. \quad (30.12b)$$

30.1.5 Horizontal momentum equation

Bringing the above results together leads to the horizontal momentum equation decomposed into natural coordinates

$$\frac{D|\mathbf{u}|}{Dt} = -\frac{1}{\rho_0} \frac{\partial p}{\partial s} + \mathbf{F} \cdot \hat{\mathbf{u}} \quad (30.13a)$$

$$\frac{|\mathbf{u}|^2}{R} + f |\mathbf{u}| = -\frac{1}{\rho_0} \frac{\partial p}{\partial n} + \mathbf{F} \cdot \hat{\mathbf{n}}, \quad (30.13b)$$

where \mathbf{F} is the friction force per mass and ρ_0 is the reference density for the Boussinesq fluid. These equations decompose the accelerations into those acting parallel to and normal to the trajectory. It is notable that the normal component is purely diagnostic; there is no time derivative in equation (30.13b). Instead, it is a balance between centrifugal, Coriolis, normal pressure gradient, and normal component of friction. In the next few sections we consider certain limiting cases as revealed by the equations of motion (30.13a) and (30.13b). Friction remains zero in this chapter but is nonzero for the discussion of Ekman layers in Chapter 31.

30.1.6 An Eulerian decomposition of the acceleration

The horizontal equations of motion (30.13a) and (30.13b) are relatively simple since we are focused on a Lagrangian description. In contrast, an Eulerian description requires us to expand the material time derivative to render

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \frac{D|\mathbf{u}|}{Dt} \hat{\mathbf{u}} + \frac{|\mathbf{u}|^2}{R} \hat{\mathbf{n}}, \quad (30.14)$$

so that

$$\hat{\mathbf{u}} \cdot \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right] = \frac{D|\mathbf{u}|}{Dt} = -\frac{1}{\rho_0} \frac{\partial p}{\partial s} + \mathbf{F} \cdot \hat{\mathbf{u}} \quad (30.15a)$$

$$\hat{\mathbf{n}} \cdot \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right] = \frac{|\mathbf{u}|^2}{R} = -f |\mathbf{u}| - \frac{1}{\rho_0} \frac{\partial p}{\partial n} + \mathbf{F} \cdot \hat{\mathbf{n}}. \quad (30.15b)$$

Depending on the information provided by a field measurement or numerical simulation, one might more readily diagnose the kinematic expressions on the left side of these equations or the force balances on the right side. For example, the inverse radius of curvature for the fluid flow can be diagnosed in either of the following ways

$$R^{-1} = \frac{\hat{\mathbf{n}} \cdot [\partial \mathbf{u} / \partial t + (\mathbf{u} \cdot \nabla) \mathbf{u}]}{|\mathbf{u}|^2} = \frac{-f |\mathbf{u}| - \rho_0^{-1} \partial p / \partial n + \mathbf{F} \cdot \hat{\mathbf{n}}}{|\mathbf{u}|^2}. \quad (30.16)$$

We focus in the remainder of this chapter on the Lagrangian perspective given its much more concise formulation.

30.1.7 Further study

Section 3.2 of [Holton \(1992\)](#) details the use of natural coordinates for geophysical flows, with a similar decomposition provided in Section 7.10 of [Gill \(1982\)](#) and Section 2.9 of [Vallis \(2017\)](#). Natural coordinates are also used in describing non-rotating flows as illustrated in [this video](#) featuring Prof. A. Shapiro.

30.2 Exact geostrophic balance

Frictionless flow parallel to pressure contours experiences no pressure gradient ($\partial p / \partial s = 0$), so that the speed of a fluid particle remains constant. If this motion furthermore occurs with an infinite radius of curvature (straight line motion parallel to pressure contours), then the force balance is between the normal pressure gradient and Coriolis. More precisely, exact geostrophic balance occurs under the following conditions:

- fluid particles move on a straight line so that the radius of curvature is infinite, $|R| = \infty$;
- fluid particles move along lines of constant pressure so that $\partial p / \partial s = 0$;
- friction vanishes.

In this case the equations of motion (30.13a) and (30.13b) take the form

$$\frac{D|\mathbf{u}|}{Dt} = 0 \quad (30.17a)$$

$$f |\mathbf{u}| = -\frac{1}{\rho_0} \frac{\partial p}{\partial n}. \quad (30.17b)$$

Equation (30.17a) says that the speed of a fluid particle is constant. Equation (30.17b) says that the pressure gradient normal to the motion balances the Coriolis acceleration. We refer to this flow, depicted in Figure 30.4, as *exact geostrophic balance* since it is an exact solution under the above assumptions.

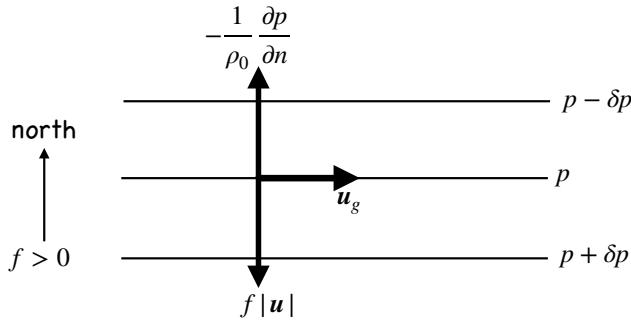


Figure 30.4: Exact geostrophic balance occurs when the flow is horizontal, frictionless, straight, and follows contours of constant pressure. For this case the pressure gradient exactly balances the Coriolis acceleration so that the motion is perpendicular to both of these accelerations. We here depict motion assuming $f > 0$ as for the northern hemisphere. If flow is on an f -plane then the exact geostrophic balance is steady for any arbitrary flow direction. On a sphere, however, steady exact geostrophic balance holds only for zonal flow.

Writing the horizontal advection of speed in the form

$$\mathbf{u} \cdot \nabla |\mathbf{u}| = |\mathbf{u}| \hat{\mathbf{u}} \cdot \nabla |\mathbf{u}| = |\mathbf{u}| \frac{\partial |\mathbf{u}|}{\partial s}, \quad (30.18)$$

allows us to write the material constancy of the flow speed as

$$\frac{\partial |\mathbf{u}|}{\partial t} + |\mathbf{u}| \frac{\partial |\mathbf{u}|}{\partial s} = 0. \quad (30.19)$$

Hence, a steady flow speed, with $\partial |\mathbf{u}| / \partial t = 0$, only holds for the exact geostrophic balance if the flow speed is fixed along each trajectory path

$$\frac{\partial |\mathbf{u}|}{\partial s} = 0 \implies \frac{\partial |\mathbf{u}|}{\partial t} = 0. \quad (30.20)$$

What is required for this condition to hold?

30.2.1 Steady f -plane flow

Geostrophic motion on an f -plane is horizontally non-divergent (Section 29.3)

$$\nabla \cdot \mathbf{u} = \nabla \cdot (\hat{\mathbf{u}} |\mathbf{u}|) = 0. \quad (30.21)$$

Flow in a straight line, with each trajectory parallel to one another, has the particle trajectory direction independent of space. Hence, the non-divergent condition means that

$$\nabla \cdot (\hat{\mathbf{u}} |\mathbf{u}|) = (\hat{\mathbf{u}} \cdot \nabla) |\mathbf{u}| = \frac{\partial |\mathbf{u}|}{\partial s} = 0, \quad (30.22)$$

which then proves that exact geostrophic flow on an f -plane is steady.

30.2.2 Steady flow on a sphere

The geostrophic velocity in the presence of a meridional gradient of the Coriolis parameter, $f = f(y)$, satisfies (Section 29.3)

$$\nabla \cdot (f \mathbf{u}) = 0. \quad (30.23)$$

Making use of $\nabla \cdot \hat{\mathbf{u}} = 0$ for straight line trajectories leads to

$$\nabla \cdot (f \mathbf{u}) = \frac{\partial(f|\mathbf{u}|)}{\partial s} = 0. \quad (30.24)$$

We conclude that $\partial|\mathbf{u}|/\partial s = 0$ holds only for trajectories that are parallel to latitude lines, in which case $\partial f/\partial s = \partial f/\partial x = 0$. Therefore, exact geostrophic motion on the sphere is steady only for trajectories that follow constant latitude lines; i.e., zonal trajectories as depicted in Figure 30.4.

30.2.3 What about geostrophic balance with curvature?

The geostrophically balanced flows discussed in Chapter 29 generally have curvature, such as for the geostrophic motion around a pressure center as shown in Figure 29.2. But as emphasized by the natural coordinate decomposition as per equations (30.13a) and (30.13b), curved motion has an associated centrifugal acceleration. So when speaking of geostrophic balance for curved flow we are ignoring the centrifugal acceleration arising from curvature. This is a sensible approximation for large-scale flows. Even so, it is an approximation and it is identically zero only for straight line motion in which there is an *exact* balance between pressure and Coriolis.

30.3 Inertial motion of fluid particles

Inertial motion occurs under the following conditions:

- vanishing pressure gradient
- vanishing friction,

so that the equations of motion (30.13a) and (30.13b) take the form

$$\frac{D|\mathbf{u}|}{Dt} = 0 \quad (30.25a)$$

$$\frac{|\mathbf{u}|^2}{R} + f|\mathbf{u}| = 0. \quad (30.25b)$$

30.3.1 Anti-cyclonic circular motion

Equation (30.25a) says that inertial motion occurs with constant speed. Equation (30.25b) says that the motion maintains the balance between Coriolis and centrifugal accelerations

$$f|\mathbf{u}| = -\frac{|\mathbf{u}|^2}{R}. \quad (30.26)$$

To further understand the implications of this result, divide by the speed to render

$$f = -\frac{|\mathbf{u}|}{R} \quad (30.27)$$

so that the radius for the inertial circle is

$$R = -|\mathbf{u}|/f. \quad (30.28)$$

Equation (30.27) can be satisfied in the northern hemisphere ($f > 0$) only for motion turning to the right (in which $R < 0$). The opposite orientation occurs in the southern hemisphere, where inertial motion turns to the left. Hence, inertial motion is oriented anti-cyclonically (opposite orientation to the earth's rotation). If the Coriolis parameter is constant, then the motion is circular, as depicted in Figure 30.5.

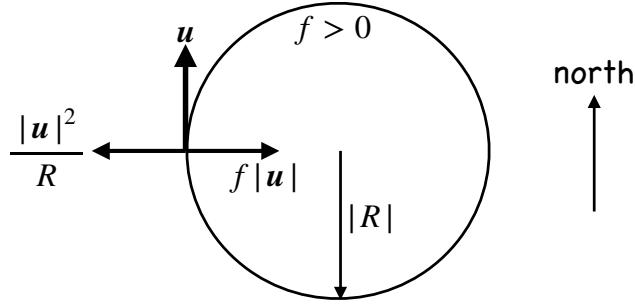


Figure 30.5: Inertial motion of a fluid particle on a plane occurs when the flow is horizontal, frictionless, and the centrifugal acceleration balances the Coriolis in the presence of zero pressure gradient. We here depict motion assuming $f > 0$ as for the northern hemisphere, revealing that inertial motion is an anti-cyclonic circular motion with radius $|R| = |\mathbf{u}|/|f|$.

30.3.2 Balance between Coriolis and centrifugal

Return to the balance given by equation (30.26) and recall that the Coriolis acceleration in the northern hemisphere points to the right when facing downstream, as per equation (30.11b). Hence, we interpret the balance (30.26) as between the Coriolis acceleration pointing to the right and the centrifugal acceleration pointing to the left. The same balance occurs in the southern hemisphere between Coriolis and centrifugal, though with the opposite orientation of the motion.

30.3.3 Period for inertial motion

Equation (30.28) says that the speed of the fluid particle is given by the radius of curvature times the magnitude of the Coriolis parameter

$$|\mathbf{u}| = R|f|. \quad (30.29)$$

The time for a fluid particle to traverse an inertial circle is given by the circumference of the circle, $2\pi R$, divided by the constant speed, thus yielding the inertial period

$$T_{\text{inertial}} = \frac{2\pi R}{|\mathbf{u}|} = \frac{2\pi}{|f|}. \quad (30.30)$$

We encountered this inertial period in Section 15.3 when considering inertial oscillations for a point particle.

30.3.4 Observing inertial motion

Inertial motion is rarely observed in the atmosphere since fluid motion nearly always occurs in the presence of a pressure gradient. In contrast, surface ocean flow is commonly generated by wind stresses that setup motion even in the absence of pressure gradients. The moving fluid then engenders a Coriolis acceleration so that there can readily be a balance between centrifugal and Coriolis for the moving seawater fluid particle. As a result, the observed surface ocean currents have nontrivial power within the inertial frequency band, rivaling energy contained in frequencies associated with astronomical tides (e.g., see Figure 3.3 of Holton (1992)).

How large is an inertial circle? Consider a surface current speed of $|\mathbf{u}| \sim 0.1 \text{ m s}^{-1}$, which is not atypical of current speeds outside of strong boundary currents or mesoscale eddies, and assume the Coriolis parameter $f = 10^{-4} \text{ s}^{-1}$. In this case the inertial radius is

$$R_{\text{inertial}} \approx 10^3 \text{ m}. \quad (30.31)$$

Observations of inertial motion, such as that reproduced in Figure 8.3 of [Gill \(1982\)](#), confirm that the radii are indeed on the order of a few kilometers.

30.3.5 Inertial motion is Lagrangian

The analysis in the current section concerns a fluid particle moving without feeling the impacts from pressure forces. The fluid particle thus exhibits an identical force balance to the point particle discussed in Section 15.3. So although we can measure inertial oscillations at a fixed point in space, the present considerations are Lagrangian or material in nature, focusing on motion of a fluid particle. Furthermore, the inertial period refers to the time it takes for a fluid particle to move around the inertial circle at its constant speed. It does not refer to the period of a wave, for example, and yet there are inertia-gravity waves that have periods close to the inertial period (see Section 44.3).

30.3.6 “Inertial” motion is not motion viewed from an inertial frame

We make use of the term “inertial” when referring to inertial motion since both the Coriolis and centrifugal accelerations are nonzero only in the presence of motion; i.e., they require inertia. Hence, as noted in Section 15.3.4, “inertial motion” in this context does *not* refer to motion as viewed in an inertial reference frame, such as that discussed in Section 13.2.

30.4 Cyclostrophic balance

Cyclostrophic balance occurs under the following conditions:

- fluid particles move along lines of constant pressure so that $\partial p/\partial s = 0$;
- vanishing Coriolis acceleration;
- vanishing friction.

The resulting equations of motion (30.13a) and (30.13b) take the form

$$\frac{D|\mathbf{u}|}{Dt} = 0 \quad (30.32a)$$

$$\frac{|\mathbf{u}|^2}{R} = -\frac{1}{\rho_0} \frac{\partial p}{\partial n}. \quad (30.32b)$$

Again, equation (30.32a) says that the speed is constant following a material fluid particle. Equation (30.32b) says that cyclostrophic flow occurs when the centrifugal acceleration balances the pressure gradient, with the squared speed given by

$$|\mathbf{u}|^2 = -\frac{R}{\rho_0} \frac{\partial p}{\partial n}. \quad (30.33)$$

This equation can be satisfied for either clockwise or counter-clockwise motion around a low pressure center, as shown in Figure 30.6. For clockwise flow, the radius of curvature is negative, $R < 0$, whereas $\partial p/\partial n > 0$. The signs are swapped for counter-clockwise flow. Cyclostrophic balance cannot be maintained around a high pressure center. The reason is that if both the pressure and centrifugal accelerations point away from the circle’s center, then they are unable to balance one another.

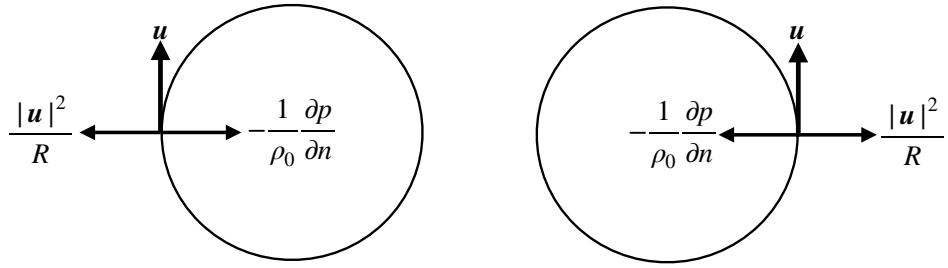


Figure 30.6: Cyclostrophic motion of a fluid particle on a plane occurs when the flow is horizontal, frictionless, with constant speed, and where the centrifugal acceleration balances the pressure gradient normal to the flow direction. We here depict motion for clockwise and counter-clockwise cyclostrophic flow, both around a low pressure. Cyclostrophic balance does not occur for flow around a high pressure center. The reason is that if both the pressure and centrifugal accelerations point away from the center, then they are unable to balance one another.

Cyclostrophic balance is relevant for scales on the order of a tornado, with radii on the order of 300 m where tangential speeds are on the order of 30 m s^{-1} (see Section 3.2.4 of Holton (1992)). For this flow scale, the Rossby number is on the order of 1000 at middle latitudes, thus justifying our neglect of Coriolis acceleration. Although tornadoes in cyclostrophic balance can rotate either clockwise or counter-clockwise, they are more often observed rotating cyclonically given that they are generally embedded within cyclonic storm systems. In contrast, smaller motions such as dust devils and water spouts are quite often seen rotating in either direction.

30.5 Gradient wind balance

Gradient wind balance occurs under the following conditions:

- fluid particles move along lines of constant pressure so that $\partial p / \partial s = 0$;
- vanishing friction.

The resulting equations of motion (30.13a) and (30.13b) take the form

$$\frac{D|u|}{Dt} = 0 \quad (30.34a)$$

$$\frac{|u|^2}{R} + f|u| = -\frac{1}{\rho_0} \frac{\partial p}{\partial n}, \quad (30.34b)$$

Again, equation (30.34a) says that the speed is constant following a material fluid particle. Equation (30.34b) says that gradient wind balanced flow occurs when the centrifugal and Coriolis accelerations balance the pressure gradient acting normal to the motion.

30.5.1 Constraints on gradient wind flow

The quadratic formula leads to the following expression for the speed of gradient wind flow

$$|u| = \frac{R}{2} \left[f \pm \sqrt{f^2 - \frac{4}{R} \frac{1}{\rho_0} \frac{\partial p}{\partial n}} \right]. \quad (30.35)$$

The speed is a real number if the pressure gradient, Coriolis parameter, and radius of curvature satisfy

$$f^2 > \frac{4}{R} \frac{1}{\rho_0} \frac{\partial p}{\partial n} \implies \frac{1}{\rho_0} \frac{\partial p}{\partial n} < \frac{f^2 R}{4}. \quad (30.36)$$

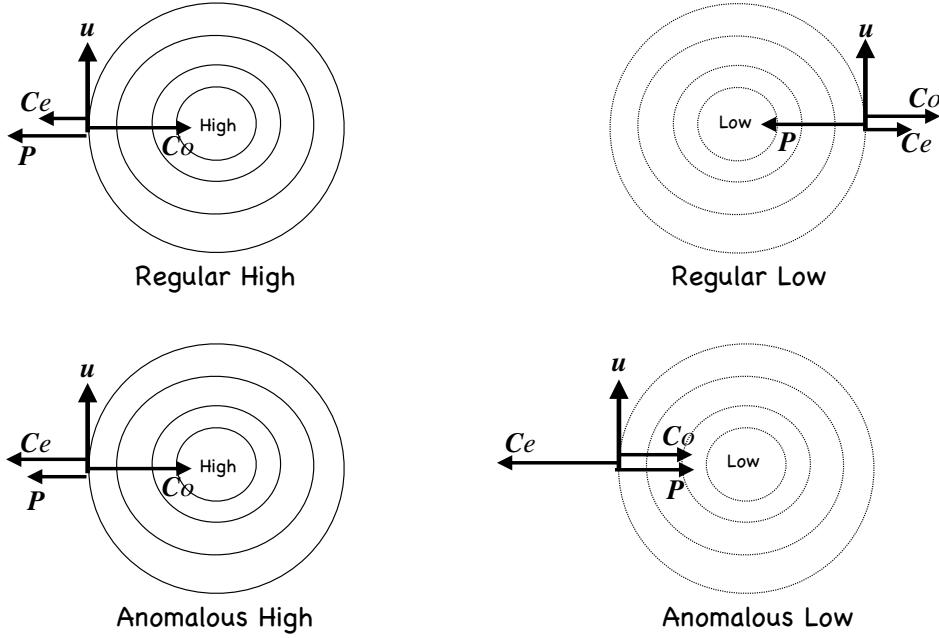


Figure 30.7: The variety of gradient wind balances available in the northern hemisphere ($f > 0$). Gradient wind motion occurs when the flow is horizontal, frictionless, with constant speed, and where the centrifugal, pressure, and Coriolis accelerations balance under a variety of magnitudes. To reduce clutter, we use the following shorthand for the accelerations: $P = -\rho_0^{-1} \partial p / \partial n$, $C_o = f |\mathbf{u}|$, and $C_e = |\mathbf{u}|^2 / R$. Upper left panel: motion around a regular high pressure center, whereby the centrifugal acceleration helps the pressure acceleration to balance the Coriolis acceleration. The pressure acceleration is larger in magnitude than the centrifugal. This flow is termed “regular” as it directly corresponds to geostrophic flow around a high pressure center. Lower left panel: motion around an anomalous high pressure center, whereby the centrifugal acceleration helps the pressure acceleration to balance the Coriolis acceleration, with the pressure acceleration smaller in magnitude than the centrifugal. This flow is termed “anomalous” as the pressure acceleration is subdominant to the centrifugal, in contrast to the case of geostrophic flow. Upper right panel: motion around a regular low pressure center, whereby the Coriolis and centrifugal accelerations balance the pressure acceleration. Lower right panel: motion around an anomalous low pressure center, whereby the Coriolis and pressure accelerations balance the centrifugal acceleration. Note the opposite flow orientation between the regular and anomalous lows.

This constraint means that the pressure gradient at the center of a high pressure region must go to zero as the radius of curvature vanishes, which renders the pressure field relatively flat near the center of highs. In contrast, there is no analogous limit for the magnitude of the pressure gradient approaching a low pressure center. This asymmetry between high and low pressures manifests in atmospheric flow with low pressure centers (cyclonic lows) having stronger magnitude than high pressure centers (anti-cyclonic highs).

30.5.2 The variety of gradient wind flows

Section 3.2 of [Holton \(1992\)](#) identifies the following force balances available with a gradient wind balance. We illustrate these cases in Figure 30.7.

- **REGULAR HIGH PRESSURE CENTER (RIGHT TURN WITH HIGH PRESSURE ON RIGHT):** This flow occurs with $R < 0$ and $\partial p / \partial n < 0$. This case occurs with the centrifugal and pressure accelerations pointing away from the center, and these balance the Coriolis acceleration pointing to the high pressure center (upper left panel of Figure 30.7). Furthermore, the inequality

(30.36) provides a bound to the size of the pressure gradient so that

$$\frac{1}{\rho_0} \frac{\partial p}{\partial n} \leq \frac{R f^2}{4}. \quad (30.37)$$

That is, the high pressure center cannot be larger than this bound in order for there to be a real solution. Note the [Holton \(1992\)](#) identifies two subcases for this balance depending on the relative size of the pressure and centrifugal accelerations, with the anomalous high the case where the pressure gradient acceleration is weaker than the centrifugal (lower left panel of Figure 30.7).

- REGULAR LOW (LEFT TURN WITH LOW PRESSURE ON LEFT): This flow occurs with $R > 0$ and $\partial p / \partial n > 0$. This case occurs with the Coriolis and centrifugal accelerations pointing away from the low pressure center, and these two accelerations balance the pressure acceleration pointing toward the center (upper right panel of Figure 30.7). The inequality (30.36) provides no bound to the magnitude of the low pressure.
- ANOMALOUS LOW (RIGHT TURN WITH LOW PRESSURE ON RIGHT): This flow occurs with $R < 0$ and $\partial p / \partial n > 0$. This case occurs with the Coriolis and pressure accelerations pointing toward the low pressure center, and these two accelerations balance the centrifugal acceleration pointing away from the center (lower right panel of Figure 30.7). The inequality (30.36) provides no bound to the magnitude of the low pressure. Note the opposite orientation for the flow around an anomalous low relative to the regular low.
- LEFT TURN WITH HIGH PRESSURE ON LEFT: In this case $R > 0$ and $\partial p / \partial n > 0$. There is no solution for the northern hemisphere since all accelerations point to the right of the motion thus disallowing any balance.

30.5.3 Comments

As noted in Section 3.2 of [Holton \(1992\)](#), the difference between gradient wind speeds and geostrophic wind speeds is no more than 10% to 20% in middle latitude synoptic atmosphere flow. However, in the tropics, where geostrophy becomes less relevant, it is important to apply the gradient wind relation to capture the balanced flow states. Furthermore, [van Heijst \(2010\)](#) makes use of the analysis in this section for the study of vortices in ocean flows. The deviations from geostrophy become important when considering relatively small vortices and/or vortices in the tropics.

31

Ekman mechanics

Ekman mechanics explains key features of how momentum and matter are transferred across the turbulent boundary layers into the quasi-adiabatic geostrophic interior of the ocean and atmosphere. The Ekman boundary layer is affected by rotation and as such it exhibits many behaviors that are distinct relative to non-rotating boundary layers. For example, the Ekman layer imparts a stretching and squeezing of interior fluid columns that strongly couples to vorticity and circulation of the interior fluid regime. Ekman layers thus play a fundamental role in oceanic and atmospheric general circulation.

In this chapter we explore the rudimentary mechanics of Ekman layers. In Ekman layers we find a balance between the horizontal pressure gradient, Coriolis acceleration , and vertical frictional acceleration. These layers realize a horizontally divergent/convergent mass transport that leads to the vertical exchange of mass, momentum, and vorticity with the fluid interior.

READER'S GUIDE TO THIS CHAPTER

We here assume an understanding of geostrophic flow from Chapter 29 and the description of flow using natural coordinates in Chapter 30. Ekman mechanics is a primary topic of geophysical fluid mechanics and applications to general circulation of the ocean and atmosphere. However, we do not make much use of Ekman layer mechanics in subsequent chapters, thus making this chapter largely self-contained.

31.1	Dynamical balance within Ekman layers	466
31.2	Natural coordinates according to isobars	467
31.3	Ekman balance in natural coordinates	467
31.4	Cross isobar flow driven by Rayleigh drag	468
31.4.1	What is Rayleigh drag?	468
31.4.2	Ekman velocity driven by Rayleigh drag	469
31.5	Horizontal spiral plus vertical rising/sinking	469
31.6	Laplacian vertical friction	469
31.7	Ekman number and Ekman layer thickness	470
31.7.1	Non-dimensionalization	471
31.7.2	Defining the Ekman number	471
31.7.3	Ekman layer thickness	472
31.7.4	Estimates for the vertical eddy viscosity	472
31.7.5	Ekman layer thickness with Rayleigh drag	472
31.8	Ekman layer integrated mass transport	472
31.8.1	Horizontal Ekman mass transport	473
31.8.2	Vertical transport into or out of the Ekman layer	474
31.8.3	Ekman and geostrophic mass transports	475
31.9	Further study	476

31.1 Dynamical balance within Ekman layers

To understand the essential features of Ekman layers it is sufficient to study a Boussinesq hydrostatic fluid within a turbulent boundary layer. The balance of accelerations in the momentum equation is between Coriolis, pressure, and friction

$$f \hat{\mathbf{z}} \wedge \mathbf{u} = -\frac{1}{\rho_0} \nabla_z p + \mathbf{F}. \quad (31.1)$$

In this chapter we focus on a frictional acceleration, \mathbf{F} , arising from the vertical exchange of horizontal momentum between fluid layers. We ignore the centrifugal acceleration described in Section 30.1.3, even though the motion generally has curvature.

It is useful to separate velocity into a geostrophic contribution defined by a balance between pressure gradient and Coriolis accelerations

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = -\frac{1}{\rho_0} \nabla_z p \implies \mathbf{u}_g = \frac{1}{f \rho_0} \hat{\mathbf{z}} \wedge \nabla p \quad (31.2)$$

plus an ageostrophic or Ekman contribution defined by a balance between the frictional and Coriolis accelerations

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_e = \mathbf{F} \implies \mathbf{u}_e = -f^{-1} \hat{\mathbf{z}} \wedge \mathbf{F}. \quad (31.3)$$

This decomposition has the appearance of superposing linearly independent flows, one geostrophic and one ageostrophic. And yet the flows are quite coupled. Namely, ageostrophic motions alter the pressure field which in turn affects the geostrophic flow. So the presence of friction and the associated ageostrophic flows leads to a geostrophic flow that differs from the inviscid case. Hence, the above decomposition is motivated by mathematical convenience and does not reflect a physical decoupling of the geostrophic and ageostrophic flows.

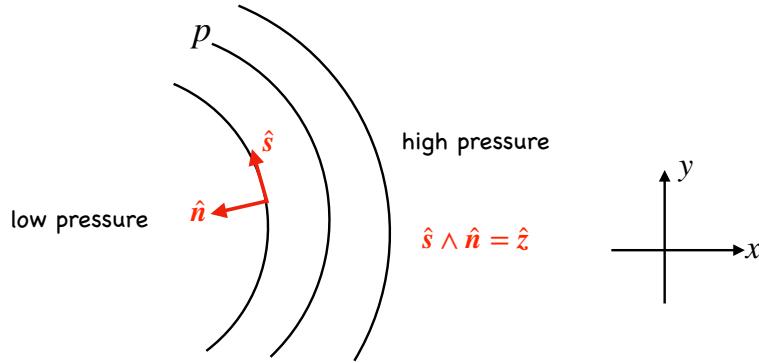


Figure 31.1: Natural coordinates defined along an arbitrary geopotential surface according to isobars in the horizontal plane. The normal direction, $\hat{n} = -\nabla_z p / |\nabla_z p|$, is oriented down the horizontal pressure gradient so that it points to the left of the geostrophic wind (facing downwind) in the northern hemisphere and to the right in the south. The tangent direction, \hat{s} , points along the isobar so that $\hat{s} \wedge \hat{n} = \hat{z}$.

31.2 Natural coordinates according to isobars

As per the definition (31.2), geostrophic motion occurs along lines of constant pressure. We are here interested in deviations from geostrophic motion realized by adding friction: how does friction support horizontal motion across isobars? To help answer this question it is useful to decompose the horizontal flow into natural coordinates along an arbitrary geopotential surface. But instead of defining the natural coordinates according to the flow direction, as done in Section 30.1, we here define the directions according to pressure contours (isobars)

$$\hat{z} = \hat{s} \wedge \hat{n} = \text{vertical direction} \quad (31.4a)$$

$$\hat{n} = \hat{z} \wedge \hat{s} = \text{down pressure gradient direction} \quad (31.4b)$$

$$\hat{s} = \hat{n} \wedge \hat{z} = \text{tangent to isobar}. \quad (31.4c)$$

The unit vector \hat{s} is tangent to isobars in the horizontal plane, whereas \hat{n} is perpendicular to isobars and oriented down the horizontal pressure gradient

$$\hat{n} = -\frac{\nabla_z p}{|\nabla_z p|}. \quad (31.5)$$

These natural coordinates are illustrated in Figure 31.1.

31.3 Ekman balance in natural coordinates

We now represent the geostrophic and Ekman velocities in natural coordinates. As found in studying the geostrophic balance in Section 30.2, the geostrophic velocity flows along isobars and so only has a component in the \hat{s} direction

$$\hat{s} \cdot \mathbf{u}_g = -\frac{1}{f\rho_0} \frac{\partial p}{\partial n} \quad (31.6a)$$

$$\hat{n} \cdot \mathbf{u}_g = 0. \quad (31.6b)$$

In contrast, the Ekman velocity has a component both along and across isobars

$$\hat{s} \cdot \mathbf{u}_e = f^{-1} \hat{n} \cdot \mathbf{F} \quad (31.7a)$$

$$\hat{n} \cdot \mathbf{u}_e = -f^{-1} \hat{s} \cdot \mathbf{F}. \quad (31.7b)$$

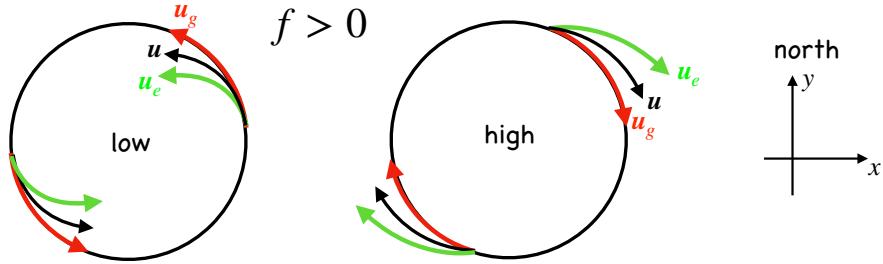


Figure 31.2: Illustrating the flow associated with Ekman dynamics in the northern hemisphere ($f > 0$). Left panel: geostrophic flow, u_g , around a low pressure center is counter-clockwise and aligned with pressure isosurfaces. Friction aligned drives Ekman flow down the pressure gradient, thus causing the fluid to spiral into the low pressure center. Right panel: the opposite oriented flow occurs around high pressure centers, where fluid spirals away from the high due to the cross-isobar flow driven by friction.

As expected, the Ekman velocity vanishes when the frictional acceleration vanishes. These equations show that friction aligned along and across isobars drive Ekman velocities across and along isobars, respectively. If there is no friction component along the isobar, $\hat{s} \cdot \mathbf{F} = 0$, then the Ekman flow follows isobars just as the geostrophic flow. More commonly, friction has a nonzero component along isobars, which leads to Ekman flow crossing isobars.

31.4 Cross isobar flow driven by Rayleigh drag

We here assume the frictional acceleration takes the form of Rayleigh drag. As noted in Section 7.4 of [Marshall and Plumb \(2008\)](#), the relative simplicity of Rayleigh drag affords analytical expressions for the Ekman velocity in terms of the geostrophic velocity, offering an explicit illustration of how the Ekman flow provides a cross-isobar component to the flow in the direction down the pressure gradient.

31.4.1 What is Rayleigh drag?

We take the frictional acceleration in the form of a Rayleigh drag acting on the velocity field

$$\mathbf{F}^{\text{drag}} = -\frac{U_{\text{fric}} \mathbf{u}}{\delta} = -\gamma \mathbf{u}, \quad (31.8)$$

where δ is a depth scale and U_{fric} is a friction velocity scale with dimensions L/T . The ratio

$$\gamma = \frac{U_{\text{fric}}}{\delta} \quad (31.9)$$

has dimensions T^{-1} and can be thought of as an inverse spin-down time. That is, if only Rayleigh drag affected changes to the horizontal momentum, $\partial_t \mathbf{u} = -\gamma \mathbf{u}$, then the flow would exponentially come to a halt with an e-folding time γ^{-1} . The drag is relatively large over rough surfaces, thus leading to a small e-folding time. In particular, drag on the lower atmospheric winds is larger over land than over the ocean. The reason is that trees, cities, and mountains dissipate more of the atmosphere's mechanical energy than interactions with the relatively smooth ocean surface.

Rayleigh drag dissipates all flow features regardless of their structure; i.e., mathematically it does not prefer any particular length scales in the fluid flow. This lack of scale selectivity contrasts to the Laplacian friction discussed in Section 31.6, with Laplacian friction dissipating small scales more rapidly than the large scales.

31.4.2 Ekman velocity driven by Rayleigh drag

Making use of the Rayleigh drag (31.8) brings the expressions (31.7a) and (31.7b) for the Ekman velocity into the form

$$\hat{s} \cdot \mathbf{u}_e = -(\gamma/f) \hat{n} \cdot \mathbf{u}_e \quad (31.10a)$$

$$\hat{n} \cdot \mathbf{u}_e = (\gamma/f) \hat{s} \cdot (\mathbf{u}_e + \mathbf{u}_g). \quad (31.10b)$$

These equations allow us to express the Ekman velocity in terms of the geostrophic velocity

$$\hat{s} \cdot \mathbf{u}_e = -(\gamma/f) \hat{n} \cdot \mathbf{u}_e \quad (31.11a)$$

$$\hat{n} \cdot \mathbf{u}_e = \hat{s} \cdot \mathbf{u}_g \left[\frac{(\gamma/f)}{1 + (\gamma/f)^2} \right], \quad (31.11b)$$

so that the velocity components for the total velocity, $\mathbf{u} = \mathbf{u}_g + \mathbf{u}_e$, and its squared magnitude, are

$$\hat{s} \cdot \mathbf{u} = \hat{s} \cdot \mathbf{u}_g \left[\frac{1}{1 + (\gamma/f)^2} \right] \quad (31.12a)$$

$$\hat{n} \cdot \mathbf{u} = (\gamma/f) \hat{s} \cdot \mathbf{u} \quad (31.12b)$$

$$(\hat{s} \cdot \mathbf{u})^2 + (\hat{n} \cdot \mathbf{u})^2 = \frac{(\hat{s} \cdot \mathbf{u}_g)^2}{1 + (\gamma/f)^2}. \quad (31.12c)$$

As expected, the cross-isobar flow is directly driven by the Rayleigh drag, and it is directed down the pressure gradient so long as the flow has a positive projection onto the tangent direction, $\hat{s} \cdot \mathbf{u} > 0$. Rayleigh drag thus causes the fluid to spiral into low pressure centers and away from a high pressure centers, such as that depicted in Figure 31.2. Furthermore, the magnitude of the total flow is reduced relative to the geostrophic flow, thus reflecting the dissipation of kinetic energy from drag.

31.5 Horizontal spiral plus vertical rising/sinking

A robust feature of Ekman flow is the spiraling motion of the fluid in the direction down the pressure gradient. This feature holds for other forms of friction, though it is more difficult to show analytically than in the case of Rayleigh drag. So far, we have focused on the horizontal spiral motion as shown in Figure 31.2. However, there is a corresponding vertical motion as well, with the vertical motion induced by the convergence of mass into the low pressure center and the divergence of mass away from the high pressure center. Figure 31.3 illustrates the vertical motion in an atmospheric boundary layer whereby mass rises above a low pressure center Ekman layer in response to the horizontal convergence. Conversely, mass diverges from the high pressure Ekman layer, thus induces a sinking motion over the high to replace the diverging mass.

In the remaining portion of this section we develop the formalism needed to compute the mass transport into and out of the Ekman layer. This mass transport plays an especially important role in how mechanical forcing drives the ocean general circulation.

31.6 Laplacian vertical friction

Frictional acceleration is generally not well defined since it depends on details of the turbulence and how it is parameterized. We made use of Rayleigh drag in Section 31.4 for mathematical

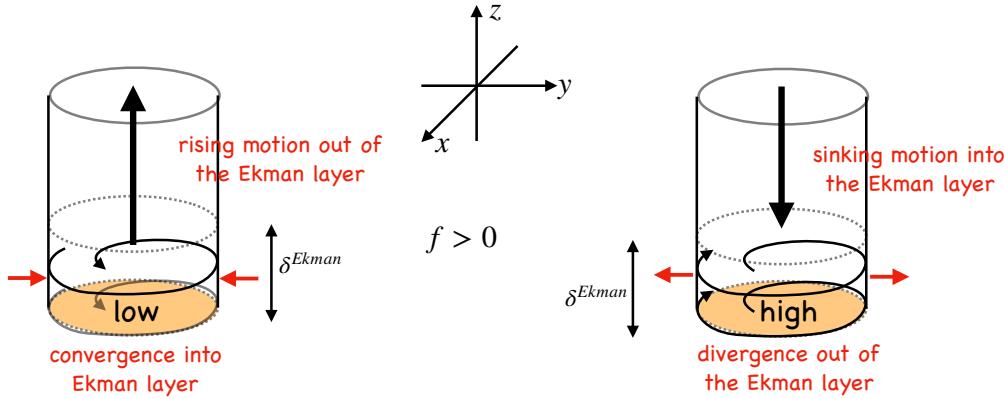


Figure 31.3: Illustrating the three-dimensional flow associated with Ekman layers in the northern hemisphere ($f > 0$) atmosphere. Left panel: flow spiralling into a low pressure center creates convergence of mass into the Ekman layer. Mass continuity means that flow must vertically leave the Ekman layer entering the interior atmosphere above. Right panel: flow spiralling away from a high pressure center creates divergence of mass out from the Ekman layer. Mass continuity means that flow must vertically sink into the Ekman layer from above. The Ekman layer thickness is denoted by δ^{Ekman} . An analogous picture holds for the surface Ekman layer in the ocean, yet with the Ekman layer at the top of the column.

expediency. Even so, it is useful to here offer a few words about the more conventional Laplacian friction operator given by

$$\mathbf{F}^{\text{viscous}} = \frac{1}{\rho_o} \frac{\partial \boldsymbol{\tau}}{\partial z} = \frac{\partial}{\partial z} \left[A \frac{\partial \mathbf{u}}{\partial z} \right], \quad (31.13)$$

with $A > 0$ a vertical turbulent kinematic viscosity with dimensions L^2/T . This form of the friction operator emulates the Laplacian operator representing molecular viscous friction (equation (26.35)). It is also that form most commonly used in numerical models and it was used by Ekman in his investigations leading to the development of Ekman theory.

Expanding the derivative reveals that the Laplacian friction operator is nonzero where there is curvature in the vertical profile of the horizontal velocity, and where there is depth dependence to the viscosity and velocity

$$\mathbf{F}^{\text{viscous}} = \frac{\partial A}{\partial z} \frac{\partial \mathbf{u}}{\partial z} + A \frac{\partial^2 \mathbf{u}}{\partial z^2}. \quad (31.14)$$

The turbulent viscosity generally has a depth dependence, with larger values in the boundary layer where turbulence is most energetic. This form of the friction preferentially acts on velocity exhibiting nontrivial vertical structure, thus acting to smooth any depth dependent behaviour. In Section 37.3 we discuss the mathematical properties of Laplacian friction.

31.7 Ekman number and Ekman layer thickness

The Ekman layer thickness is a function of the vertical friction acting within the boundary layer. The more friction the thicker the boundary layer. However, there is no deductive theory for the friction arising from turbulence nor even for its appropriate mathematical form. Even so, we find it useful to here consider the Rayleigh drag and Laplacian viscosity forms introduced earlier and develop expressions for the boundary layer thickness based on these operators. Doing so offers intuition and some level of quantitative accuracy.

As per the usual process in fluid mechanics, we find it useful to non-dimensionalize the equations to isolate important non-dimensional numbers affecting the flow regime. We here isolate the Ekman

number for boundary layers affected by friction and rotation. Friction is important where the Ekman number is order unity. Note that flows unaffected by rotation, it is the Reynolds number that proves to be the key non-dimensional number (Section 26.7) for boundary layers, rather than the Ekman number arising when rotation is important.

31.7.1 Non-dimensionalization

We make use of the following scaling

$$(x, y) = L(\hat{x}, \hat{y}) \quad z = L_z \hat{z} \quad (u, v) = U(\hat{u}, \hat{v}) \quad f = f_0 \hat{f} \quad p = P \hat{p} \quad (31.15)$$

where the variables with hats are non-dimensional,¹ and we introduced typical scales for length, velocity, Coriolis, and pressure. As defined below, the vertical scale, L_z , is the thickness over which friction is nontrivial, thus offering a scale for the Ekman layer thickness. The pressure scale, P , is assumed to follow geostrophic scaling

$$P = f_0 \rho_0 U L. \quad (31.16)$$

Inserting the relations (31.15) into equation (31.1) leads to the non-dimensional frictional geostrophic equation

$$\hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla} \hat{p} + \frac{\mathbf{F}}{f_0 U}. \quad (31.17)$$

31.7.2 Defining the Ekman number

The ratio of scales for the frictional acceleration to the Coriolis acceleration defines the Ekman number

$$Ek = \frac{\text{frictional acceleration}}{\text{Coriolis acceleration}}. \quad (31.18)$$

For the viscous stress form of Laplacian frictional acceleration (equation (31.13))

$$\mathbf{F}^{\text{viscous}} = \frac{A U}{L_z^2} \frac{\partial^2 \hat{\mathbf{u}}}{\partial \hat{z}^2}, \quad (31.19)$$

so that the Ekman number is

$$Ek = \frac{A}{f_0 L_z^2}. \quad (31.20)$$

If we take the vertical scale, L_z , equal to the depth scale over which interior flow processes occur, then the Ekman number will generally be very small, in which case friction is negligible. However, this reasoning is flawed when closer to the boundary since the Ekman number multiplies the highest derivative in equation (31.17). Setting the Ekman number to zero thus represents a singular limit in the sense of perturbation theory. Hence, we expect a boundary layer to form, inside of which the Ekman number is order unity and the role of Laplacian friction is nontrivial.

¹Be careful not to confuse $\hat{\mathbf{u}}$ used here for non-dimensional velocity with the $\hat{\mathbf{u}}$ unit vector pointing along the fluid particle trajectory as defined in Section 30.1.1.

31.7.3 Ekman layer thickness

The Ekman number increases when there is more boundary layer turbulence, in which case the eddy viscosity, A , is large relative to its small values in the interior. Additionally, the Ekman number increases when moving towards the equator, where the Coriolis parameter reduces.² The viscous friction acceleration, (31.19), is order unity over a depth scale where the Ekman number is order unity, which occurs within a boundary layer. This vertical scale defines the viscous Ekman boundary layer thickness

$$Ek = 1 \Rightarrow \delta^{\text{viscous}} = \left[\frac{A}{f_0} \right]^{1/2}. \quad (31.21)$$

31.7.4 Estimates for the vertical eddy viscosity

The eddy viscosity is not readily available from measurements or theory. However, if we can observe the boundary layer thickness, then we have a means to infer the scale for the viscosity

$$A = f_0 (\delta^{\text{viscous}})^2. \quad (31.22)$$

In the atmosphere, the boundary layer thickness is order 1000 m, so that at mid-latitudes, with $f_0 = 10^{-4}$ s⁻¹, we expect

$$A^{\text{atmos}} \sim 10^2 \text{ m}^2 \text{ s}^{-1}. \quad (31.23)$$

In the ocean, the upper ocean boundary layer depth, outside of the deep convection regions, is roughly 50 m, in which case

$$A^{\text{ocean}} \sim 0.25 \text{ m}^2 \text{ s}^{-1}. \quad (31.24)$$

31.7.5 Ekman layer thickness with Rayleigh drag

For the Rayleigh drag form of friction given by

$$\mathbf{F}^{\text{drag}} = - \left[\frac{U_{\text{fric}}}{\delta} \right] U \hat{\mathbf{u}}, \quad (31.25)$$

the Ekman number takes the form

$$Ek = \frac{U_{\text{fric}}}{\delta f_0}. \quad (31.26)$$

The Ekman number is unity, and hence friction important, if the thickness δ takes on the Ekman value

$$\delta^{\text{drag}} = \frac{U_{\text{fric}}}{f_0}. \quad (31.27)$$

The Laplacian and Rayleigh expressions for the Ekman layer thicknesses are the same if the frictional velocity scale and viscosity are related by

$$\delta^{\text{drag}} = \delta^{\text{viscous}} \iff U_{\text{fric}} = \sqrt{A f_0}. \quad (31.28)$$

31.8 Ekman layer integrated mass transport

It is possible to establish some integral properties of the Ekman layer even without knowing details of the friction. A particularly important quantity is the integrated mass transport crossing the

²When getting too close to the equator, our assumption of a frictional geostrophic balance breaks down, so that other terms in the momentum equation, such as advection, become important.

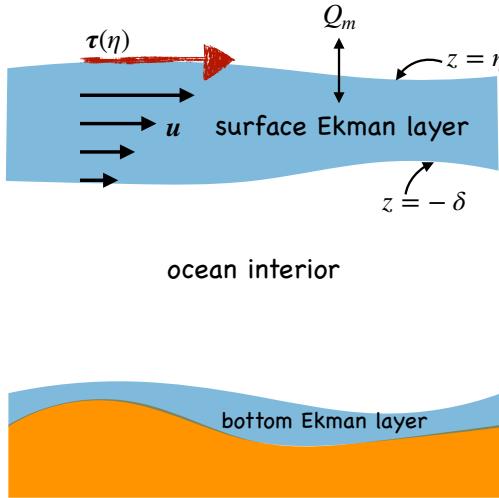


Figure 31.4: Ekman layer at the ocean surface, defined for vertical position $-\delta(x, y, t) \leq z \leq \eta(x, y, t)$, with δ specifying the Ekman layer bottom and η the free surface vertical position. Wind stress imparts horizontal momentum to the upper ocean that is transmitted throughout the Ekman layer via turbulent stresses. In addition, mass flux, Q_m , crosses the ocean free surface thus affecting the mass budget in the Ekman layer ($Q_m > 0$ means mass enters the ocean). There is also a bottom Ekman layer created by bottom boundary layer turbulence.

base of the ocean surface Ekman layer sitting between the Ekman layer bottom and the ocean free surface

$$-\delta(x, y, t) \leq z \leq \eta(x, y, t), \quad (31.29)$$

as depicted in Figure 31.4. Knowledge of this mass transport has important implications for how mechanical energy imparted to the boundary layer drives circulation within the interior of the fluid.

31.8.1 Horizontal Ekman mass transport

Integrating the horizontal Ekman balance (31.3) over the depth of the Ekman layer leads to

$$\mathbf{f} \wedge \mathbf{M}_e = \int_{-\delta}^{\eta} \rho_o \mathbf{F} dz, \quad (31.30)$$

where

$$\mathbf{M}_e = \int_{-\delta}^{\eta} \rho_o \mathbf{u}_e dz \quad (31.31)$$

is the depth integrated ageostrophic mass transport within the Ekman boundary layer. Assume friction in the form of a vertical viscous stress (equation (31.13)), so that

$$\mathbf{M}_e = -f^{-1} \hat{\mathbf{z}} \wedge [\boldsymbol{\tau}(\eta) - \boldsymbol{\tau}(-\delta)]. \quad (31.32)$$

Stress at the bottom of the Ekman layer, $\boldsymbol{\tau}(-\delta)$, matches to the stress in the ocean interior, which is much smaller than stress at the ocean surface, $\boldsymbol{\tau}(\eta)$. Neglecting this interior stress then leads to the ageostrophic Ekman layer horizontal mass transport

$$\mathbf{M}_e = -f^{-1} \hat{\mathbf{z}} \wedge \boldsymbol{\tau}(\eta). \quad (31.33)$$

Mass transport in the Ekman layer is thus at right angles to the surface stress. This orientation is a consequence of the Coriolis force. It is also a major result from Ekman's original studies of ice motion in the Arctic.

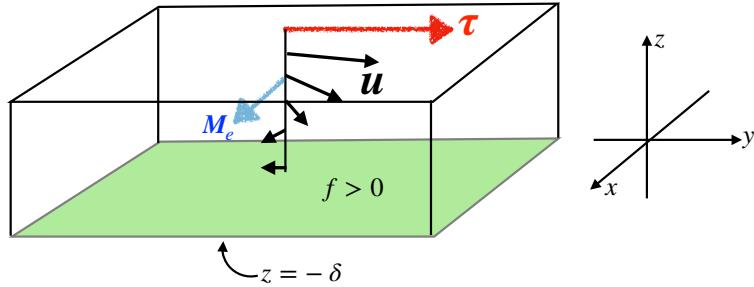


Figure 31.5: Horizontal transport integrated over the depth of the surface ocean Ekman layer is directed at right angles to the wind stress in the northern hemisphere (to the left in the southern). Here, the wind stress, τ , is shown blowing to the north so that in the northern hemisphere ($f > 0$), the depth integrated horizontal Ekman transport, M_e , is to the east. This result from Ekman dynamics is independent of the assumptions made about friction within the boundary layer.

31.8.2 Vertical transport into or out of the Ekman layer

As seen in Figure 31.3, the horizontal Ekman transport into or out of the Ekman layer induces a vertical transport into or out of the Ekman layer. To obtain an expression for the vertical transport, integrate the continuity equation $\nabla \cdot \mathbf{v} = 0$ over the vertical extent of the Ekman layer

$$\frac{\partial}{\partial x} \left[\int_{-\delta}^{\eta} u \, dz \right] + \frac{\partial}{\partial y} \left[\int_{-\delta}^{\eta} v \, dz \right] + [w(\eta) - \mathbf{u}(\eta) \cdot \nabla \eta] - [w(-\delta) + \mathbf{u}(-\delta) \cdot \nabla \delta] = 0. \quad (31.34)$$

For an incompressible fluid, the kinematic boundary condition at the ocean free surface is given by equation (19.5)

$$w + \frac{Q_m}{\rho_0} = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta \quad z = \eta(x, y, t). \quad (31.35)$$

Similarly, at the bottom of the Ekman layer we measure the volume transport through this layer by computing the dia-surface transport, $w^{(\dot{\delta})}$, according to equation (21.36)

$$w^{(\dot{\delta})} = w - (\partial_t + \mathbf{u} \cdot \nabla) z \quad z = -\delta(x, y, t). \quad (31.36)$$

The sign convention is such that $w^{(\dot{\delta})} > 0$ means that mass enters the Ekman layer through its base, whereas mass leaves through the base when $w^{(\dot{\delta})} < 0$.

Using the kinematic boundary conditions (31.35) and (31.36) in the depth integrated volume budget (31.34), and rearranging, leads to the expression for the volume transport crossing the surface and bottom of the Ekman layer

$$w^{(\dot{\delta})} = \frac{\partial h_e}{\partial t} + \frac{-Q_m + \nabla \cdot \mathbf{M}}{\rho_0} \quad (31.37)$$

where

$$h_e = \eta + \delta \quad (31.38)$$

is the thickness of the Ekman layer and

$$\mathbf{M} = \int_{-\delta}^{\eta} \rho_0 \mathbf{u} \, dz \quad (31.39)$$

is the Ekman layer integrated horizontal mass transport. It is useful as a sanity check to note that letting the Ekman layer go to the ocean bottom (so that $w^{(\dot{\delta})} = 0$) reduces equation (31.37) to the

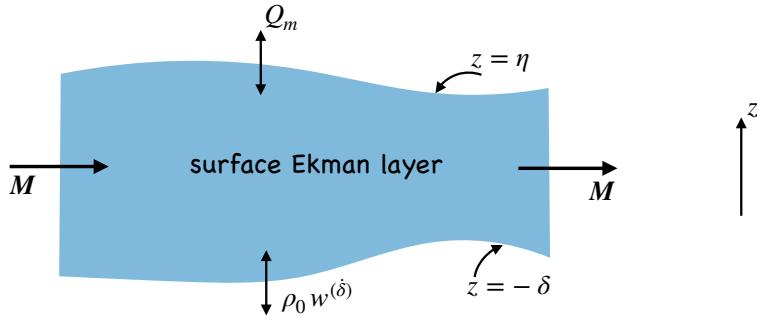


Figure 31.6: Mass budget over the surface Ekman layer of the ocean, with impacts from surface mass transport Q_m , transport through the bottom of the layer, $w^{(\delta)}$, and horizontal transport, \mathbf{M} . If these transports static. But if there are any imbalances then the layer thickness will have a nonzero time tendency, $\partial h_e / \partial t \neq 0$.

kinematic free surface equation for the full ocean column (equation (19.8))

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}. \quad (31.40)$$

Bringing the surface and bottom boundary terms together in equation (31.37) helps with its interpretation

$$\rho_0 w^{(\delta)} + Q_m = \rho_0 \frac{\partial h_e}{\partial t} + \nabla \cdot \mathbf{M}. \quad (31.41)$$

The left hand side measures the mass transport crossing the bottom of the Ekman layer plus the free surface. This transport balances a time change in the Ekman layer thickness plus the divergence of mass from the layer. A steady state layer thickness ($\partial h_e / \partial t = 0$) is realized if the divergence of mass through horizontal transport is exactly balanced by mass entering the boundary layer through the top and/or bottom of the layer. We illustrate this budget in Figure 31.6.

31.8.3 Ekman and geostrophic mass transports

The horizontal mass transport given by equation (31.39) has a contribution from both the geostrophic and Ekman velocities

$$\mathbf{M} = \mathbf{M}_g + \mathbf{M}_e. \quad (31.42)$$

The geostrophic velocity is horizontally non-divergent on an f -plane (Section 29.4.3), yet generally has a nonzero divergence on the sphere. The horizontal Ekman transport is determined largely by the wind stress according to equation (31.33), with its divergence given by

$$\nabla \cdot \mathbf{M}_e = \hat{z} \cdot [\nabla \wedge (\boldsymbol{\tau}/f)]. \quad (31.43)$$

This result brings the Ekman layer mass budget (31.41) into the form

$$\rho_0 w^{(\delta)} + Q_m = \rho_0 \frac{\partial h_e}{\partial t} + \nabla \cdot \mathbf{M}_g + \hat{z} \cdot [\nabla \wedge (\boldsymbol{\tau}/f)]. \quad (31.44)$$

The effects from wind stress in equation 31.44 warrant particular attention. The wind stress curl, as well as changes in f on the sphere, drive vertical motion through the base of the Ekman layer. The flow crossing the Ekman layer boundary acts to stretch or compress vertical fluid columns in the adjoining nearly inviscid interior. These interior fluid columns are stiffened through the effects of Taylor-Proudman (Section 29.4.3), thus coupling the Ekman layer to a vast portion

of the interior flow. From our understanding of vorticity (studied in Chapter 48), particularly the notions of vortex stretching, we see that the Ekman induced stretching/compression of interior fluid columns leads to a change in vorticity of the fluid interior, thus driving motion that affects a broad scale circulation. This coupling between the Ekman layer to the interior fluid is a fundamental feature of rotating fluids that is missing from non-rotating boundary layer flows. These general ideas of Ekman layer mechanics form the basis for how mechanical forcing from the winds drives ocean vorticity and the large-scale ocean general circulation.

31.9 Further study

This chapter borrows from Section 5.7 of [Vallis \(2017\)](#) and Section 7.4 of [Marshall and Plumb \(2008\)](#). The student is also encouraged to read Chapter 8 of [Cushman-Roisin and Beckers \(2011\)](#) for further insights.

For visualization, the following videos are recommended:

- [This video from MIT Earth, Atmospheric, and Planetary Sciences](#) illustrates the spiral flow found within an Ekman layer as realized in a rotating tank experiment.
- [This video from the UCLA SpinLab](#), near the 18 minute mark, shows how Ekman transport helps to explain the garbage patches found near the center of the ocean's sub-tropical gyres.
- [This video from the University of Chicago](#), starting near the 23 minute mark, provides examples of Ekman layers in a rotating tank. The other portions of this video exhibit many other novel aspects of rotating fluids and is highly recommended.

Part V

Selected topics in ocean dynamics

In this part of the book presents a suite of special topics in relevant to studies of ocean dynamics.

32

General vertical coordinate dynamics

In this chapter we derive the dynamical equations using generalized vertical coordinates. These equations provide the foundations for many treatments of ocean and atmosphere modeling using either Eulerian or Lagrangian algorithms for the vertical. Consequently, we provide certain details appropriate to formulate the equations of a numerical ocean model. In particular, we discuss two foundational algorithms used to time step the ocean equations. The first is the split-explicit method used for the hydrostatic momentum equation. The second is the vertical Lagrangian-remap algorithm used with hybrid vertical coordinate models. We close the chapter by outlining how to diagnose fluid particle trajectories by making use of the vertical Lagrangian-remap algorithm.

READER'S GUIDE TO THIS CHAPTER

We assume a working knowledge of the mathematics of generalized vertical coordinates as detailed in Chapter 11 and the corresponding kinematics in Chapter 21. We make particular use of the layer integrated notions introduced for mass continuity and the tracer equations in Sections 21.9 and 21.10. We also make use of the dynamical equations derived in Chapter 22. For most purposes in this chapter we find Cartesian horizontal coordinates sufficient. However, we note some places where spherical coordinates warrant special consideration. Sections 33.3 and 33.4 offer rather specialized treatments of algorithms commonly used in ocean models. Although exposing some of the details of these algorithms, the treatment is terse relative to the full details provided in the specialized literature.

32.1	Layer integrated equations of motion	480
32.1.1	Mass and tracer equations	480
32.1.2	Momentum equation	480
32.1.3	Eulerian flux-form horizontal momentum equation	481
32.1.4	Vector-invariant horizontal momentum equation	482
32.1.5	Hydrostatic balance with constant gravitational acceleration	482
32.1.6	Difficulties computing the horizontal pressure gradient	483
32.2	The pressure force acting on a grid cell	483
32.2.1	Integrated pressure force on the cell faces	484
32.2.2	Net vertical pressure force	485
32.2.3	Net horizontal pressure force	485
32.2.4	Comments	485

32.1 Layer integrated equations of motion

We here derive the equations of motion based on generalized vertical coordinates. The scalar equations (mass and tracer) were already discussed in Sections 21.8, 21.9, and 21.10. We thus focus on the momentum equations.

32.1.1 Mass and tracer equations

The layer integrated Eulerian flux-form equations for mass and tracer are given by

$$\frac{\partial(\rho h)}{\partial t} + \nabla_\sigma \cdot (\rho h \mathbf{u}) + \Delta_\sigma(\rho w^{(\dot{\sigma})}) = 0 \quad (32.1a)$$

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_\sigma \cdot (h \rho C \mathbf{u} + h \mathbf{J}^h) + \Delta_\sigma(\rho C w^{(\dot{\sigma})} + z_\sigma \nabla \sigma \cdot \mathbf{J}) = 0. \quad (32.1b)$$

Compatibility is maintained between the mass continuity equation and the tracer equation 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 uniform. 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 Section 18.1.

32.1.2 Momentum equation

From Section 22.1.4, the horizontal and vertical components to the momentum equation are

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{u} = -\rho \nabla_z \Phi - \nabla_z p + \rho \mathbf{F}^h \quad \text{horizontal} \quad (32.2a)$$

$$\rho \frac{Dw}{Dt} = -\rho \frac{\partial \Phi}{\partial z} - \frac{\partial p}{\partial z} + \rho F^z \quad \text{vertical.} \quad (32.2b)$$

Note that for the traditional form of the geopotential (Section 14.1), $\Phi = g z$ so that the horizontal gradient of the geopotential vanishes

$$\nabla_z \Phi = 0 \quad \text{if } \Phi = g z. \quad (32.3)$$

This gradient is nonzero for more general cases such as in the presence of astronomical tide forcing (Chapter 34).

Horizontal momentum equation

We transform the horizontal derivatives from geopotential coordinates to GVCs according to (see equation (11.64))

$$\nabla_z = \nabla_\sigma - (\nabla_\sigma z) \partial_z \quad (32.4)$$

thus leading to the horizontal momentum equation

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{u} = -\rho [\nabla_\sigma - (\nabla_\sigma z) \partial_z] \Phi - [\nabla_\sigma - (\nabla_\sigma z) \partial_z] p + \rho \mathbf{F}^h. \quad (32.5)$$

In Section 32.1.5 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

$$\rho \frac{Dw}{Dt} = -\frac{\partial \sigma}{\partial z} \left[\rho \frac{\partial \Phi}{\partial \sigma} - \frac{\partial p}{\partial \sigma} \right] + \rho F^z. \quad (32.6)$$

However, the vertical velocity component, $w = Dz/Dt$, is not conveniently computed using GVCs. Rather, it is the dia-surface velocity, $w^{(\dot{\sigma})}$, from Section 21.3 that concerns us, with the two velocities related by (see equation (21.50))

$$w^{(\dot{\sigma})} = w - \left[\frac{\partial z}{\partial t} \right]_\sigma - \mathbf{u} \cdot \nabla_\sigma z. \quad (32.7)$$

For some GVC algorithms that follow the approach used in isopycnal coordinate models, it is convenient to compute $w^{(\dot{\sigma})}$ according to its connection to material changes in σ

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt}. \quad (32.8)$$

For other algorithms that follow the approach in geopotential coordinate models, we diagnose $w^{(\dot{\sigma})}$ using mass continuity (see [Adcroft and Hallberg \(2006\)](#) for a discussion of these two methods for computing $w^{(\dot{\sigma})}$). Furthermore, as described in Section 33.4, the vertical Lagrangian-remap method dispenses with the explicit computation of $w^{(\dot{\sigma})}$. Instead, the method performs a vertical remapping that is fundamentally the same as dia-surface transport but numerically quite distinct.

32.1.3 Eulerian flux-form horizontal momentum equation

In Cartesian horizontal coordinates, the horizontal momentum equation includes a contribution from the acceleration that has a form as for a tracer (Section 21.10)

$$h \rho \frac{Du}{Dt} = \frac{\partial(h \rho u)}{\partial t} + \nabla_\sigma \cdot (h \rho u \mathbf{u}) + \Delta_\sigma(\rho u w^{(\dot{\sigma})}) \quad (32.9a)$$

$$h \rho \frac{Dv}{Dt} = \frac{\partial(h \rho v)}{\partial t} + \nabla_\sigma \cdot (h \rho v \mathbf{u}) + \Delta_\sigma(\rho v w^{(\dot{\sigma})}). \quad (32.9b)$$

As for the continuity equation and tracer equation discussed in Section 32.1.1, we interpret the terms in these equations as layer averaged quantities.

With spherical coordinates there are additional metric terms as detailed in Section 22.2.1. 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 GVCs. We thus favor the vector-invariant form for GVCs derived in Section 32.1.3 as these equations dispense with the usual metric terms.

32.1.4 Vector-invariant horizontal momentum equation

As noted in Section 22.2.2, the *vector-invariant* form of the momentum 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 (see Section 48.4). Here, we start with the material time derivative in the GVC form (21.48c) so that the horizontal acceleration is given by

$$\frac{D\mathbf{u}}{Dt} = \left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + (\mathbf{u} \cdot \nabla_\sigma) \mathbf{u} + (w^{(\dot{\sigma})} \partial_z) \mathbf{u}. \quad (32.10)$$

Now make use of the identity (see Section 4.3.4)

$$(\mathbf{u} \cdot \nabla_\sigma) \mathbf{u} = \nabla_\sigma K + (\nabla_\sigma \wedge \mathbf{u}) \wedge \mathbf{u}, \quad (32.11)$$

where

$$K = \mathbf{u} \cdot \mathbf{u}/2 \quad (32.12)$$

is the kinetic energy per mass of the horizontal flow. Introducing the GVC version of the relative vorticity (see Section 52.2.1)

$$\tilde{\zeta} \equiv \hat{\mathbf{z}} \cdot (\nabla_\sigma \wedge \mathbf{u}) = \left[\frac{\partial v}{\partial x} \right]_\sigma - \left[\frac{\partial u}{\partial y} \right]_\sigma \quad (32.13)$$

renders

$$\frac{D\mathbf{u}}{Dt} = \left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + \nabla_\sigma K + \hat{\mathbf{z}} \tilde{\zeta} \wedge \mathbf{u} + (w^{(\dot{\sigma})} \partial_z) \mathbf{u}, \quad (32.14)$$

so that the horizontal momentum equation takes the form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + w^{(\dot{\sigma})} \frac{\partial \mathbf{u}}{\partial z} + (2\boldsymbol{\Omega} + \hat{\mathbf{z}} \tilde{\zeta}) \wedge \mathbf{u} = -\nabla_\sigma K - \nabla_z \Phi - (1/\rho) \nabla_z p + \mathbf{F}^h, \quad (32.15)$$

where again $\nabla_z = \nabla_\sigma - (\nabla_\sigma z) \partial_z$ as per equation (32.4). This equation is form-invariant regardless the horizontal coordinates, thus motivating its name.¹

32.1.5 Hydrostatic balance with constant gravitational acceleration

There are many special cases that simplify various terms in the dynamical equations. For example, when considering a geopotential in the form $\Phi = g z$ (Section 14.1) with g a constant gravitational acceleration, then the horizontal momentum equation (32.5) becomes

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{u} = -[\nabla_\sigma - (\nabla_\sigma z) \partial_z] p + \rho \mathbf{F}^h \quad \text{if } \Phi = g z. \quad (32.16)$$

Furthermore, assuming a hydrostatic balance (and corresponding simplification of the Coriolis acceleration as per Section 27.1.3) allows us to write $\partial p / \partial z = -g \rho$ so that

$$\rho \frac{D\mathbf{u}}{Dt} + \rho f \hat{\mathbf{z}} \wedge \mathbf{u} = -[\nabla_\sigma p + \rho \nabla_\sigma \Phi] + \rho \mathbf{F}^h, \quad (32.17)$$

which also takes on the vector-invariant form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + w^{(\dot{\sigma})} \frac{\partial \mathbf{u}}{\partial z} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_\sigma(K + \Phi) - (1/\rho) \nabla_\sigma p + \mathbf{F}^h. \quad (32.18)$$

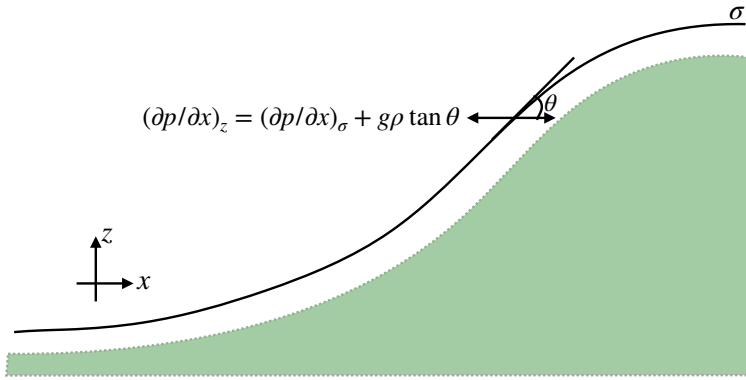


Figure 32.1: Illustrating how the horizontal pressure gradient is decomposed into two terms, one aligned with the surface of constant σ , and another associated with the slope of the σ -surface relative to the horizontal, $\tan \theta = (\partial z / \partial x)_{\sigma}$. We here consider the decomposition using terrain following vertical coordinates, where the vertical coordinate is aligned according to the solid-earth bottom (shaded region). Specifically, for terrain following coordinates we have $\sigma = (z - \eta) / (H + \eta)$, where $z = \eta(x, y, t)$ is the ocean free surface and $z = -H(x, y)$ is the ocean bottom.

32.1.6 Difficulties computing the horizontal pressure gradient

The horizontal pressure gradient is aligned perpendicular to the local gravitational direction. 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 (32.4) can lead to numerical difficulties. For example, with a simple geopotential and a hydrostatic fluid, equation (32.17) shows that the horizontal pressure gradient takes the form

$$\nabla_z p = \nabla_{\sigma} p + \rho \nabla_{\sigma} \Phi = \nabla_{\sigma} p + g \rho (\nabla_{\sigma} z), \quad (32.19)$$

with this decomposition illustrated in Figure 32.1 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. Their sum is thus exposed to potentially nontrivial numerical truncation errors that 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. \(2000a\)](#), with advances offered by [Lin \(1997\)](#), [Shchepetkin and McWilliams \(2002\)](#), and [Adcroft et al. \(2008\)](#). In Section 32.2 we outline the finite volume method for computing the pressure force as proposed by [Lin \(1997\)](#) for atmospheric models and [Adcroft et al. \(2008\)](#) for ocean models.

32.2 The pressure force acting on a grid cell

As seen in Section 26.1.3, the pressure force acting on a finite region of fluid is given by the integral

$$\mathbf{F}^{\text{pressure}} = - \oint_{\partial\mathcal{R}} p \hat{\mathbf{n}} dS = - \int_{\mathcal{R}} \nabla p dV, \quad (32.20)$$

where the second equality follows from Gauss's divergence theorem applied to a scalar field (Section 4.7.2). We refer to the right-most expression as the pressure gradient body force, and this expression

¹See Section 4.4.4 of [Griffies \(2004\)](#) for a detailed derivation using general coordinates.

is the basis for the discussion in Sections 32.1.5 and 32.1.6. The alternative is referred to as the area integrated contact force, with the integration taken over the area bounding the region. In this section we focus on the contact force expression, using it as a basis for computing the net pressure force acting on a finite region such as a model grid cell as shown in Figure 32.2. The contact force perspective was taken by [Lin \(1997\)](#) and [Adcroft et al. \(2008\)](#) in their finite volume approach to computing the pressure force acting on a grid cell.

32.2.1 Integrated pressure force on the cell faces

The outward normal on the grid cell vertical side boundaries points in the horizontal direction. For example, on the left side of Figure 32.2 the pressure force acts in the positive $\hat{\mathbf{y}}$ direction

$$\mathbf{F}_{\text{left}}^{\text{press}} = \hat{\mathbf{y}} \int_{\text{left}} p \, dx \, dz \quad (32.21)$$

whereas pressure on the right wall acts in the opposite direction

$$\mathbf{F}_{\text{right}}^{\text{press}} = -\hat{\mathbf{y}} \int_{\text{right}} p \, dx \, dz. \quad (32.22)$$

Similar expressions appear for the front and back vertical boundaries acting in the $\hat{\mathbf{x}}$ direction.

Since the top and bottom boundaries of the grid cell are sloped, there is a pressure force acting on this surface directed in both the horizontal and vertical directions. To unpack the form of 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

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla\eta}{\sqrt{1 + |\nabla\eta|^2}}. \quad (32.23)$$

Following our discussion of dia-surface transport in Section 21.3.4, we know that the product of the normal direction and the area element can be written

$$\hat{\mathbf{n}} \, d\mathcal{S} = (\hat{\mathbf{z}} - \nabla\eta) \, dA, \quad (32.24)$$

where $dA = dx \, dy$ is the horizontal projection of the area element (see Figure 21.1). Hence, the net pressure force acting on the top face is given by

$$\mathbf{F}_{\text{top}}^{\text{press}} = -\hat{\mathbf{z}} \int_{\text{top}} p \, dx \, dy + \hat{\mathbf{x}} \int_{\text{top}} p (\partial z / \partial x)_\sigma \, dx \, dy + \hat{\mathbf{y}} \int_{\text{top}} p (\partial z / \partial y)_\sigma \, dx \, dy, \quad (32.25)$$

where we set $z = \eta$ in the second and third terms and placed a σ subscript to emphasize that the horizontal derivative is taken with σ held constant. Notice that the pressure acts in the positive horizontal direction if the top surface slopes upward (surface shoaling) when moving in either of the two horizontal directions. Pressure acting on the bottom face has the same appearance yet with opposite signs

$$\mathbf{F}_{\text{bott}}^{\text{press}} = \hat{\mathbf{z}} \int_{\text{bott}} p \, dx \, dy - \hat{\mathbf{x}} \int_{\text{bott}} p (\partial z / \partial x)_\sigma \, dx \, dy - \hat{\mathbf{y}} \int_{\text{bott}} p (\partial z / \partial y)_\sigma \, dx \, dy. \quad (32.26)$$

The pressure acts in the positive horizontal direction if the bottom surface slopes downward (surface deepens) when moving in either of the two horizontal directions. As discussed in Section 26.1.3, 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. In Section 43.2 we consider the form stress acting on the interfaces of shallow water layers.

32.2.2 Net vertical pressure force

Bringing the pieces together leads to the net vertical pressure force acting on the grid cell

$$\mathbf{F}_{\text{vertical}}^{\text{press}} = -\hat{\mathbf{z}} \left[\int_{\text{top}} p \, dx \, dy - \int_{\text{bott}} p \, dx \, dy \right]. \quad (32.27)$$

If the fluid is in hydrostatic balance, then this vertical force is given by the weight of fluid within the cell

$$\mathbf{F}_{\text{vertical}}^{\text{press}} = \hat{\mathbf{z}} M g \quad \text{hydrostatic}, \quad (32.28)$$

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.

32.2.3 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

$$\mathbf{F}_{\text{merid}}^{\text{press}} = \left[\int_{\text{left}} p \, dx \, dz - \int_{\text{right}} p \, dx \, dz \right] + \left[\int_{\text{top}} p (\partial z / \partial y)_{\sigma} \, dx \, dy - \int_{\text{bott}} p (\partial z / \partial y)_{\sigma} \, dx \, dy \right]. \quad (32.29)$$

We can write this expression in a more compact form by orienting our integration in a counter-clockwise manner around the cell, and making use of the identity $(\partial z / \partial y)_{\sigma} dy = dz$ on the top and bottom faces, so that

$$\mathbf{F}_{\text{merid}}^{\text{press}} = -\oint p \, dx \, dz. \quad (32.30)$$

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 meridional pressure force as

$$\mathbf{F}_{\text{merid}}^{\text{press}} = -\oint p \, dx \, dz = -\oint dx [d(pz) - z dp] = g^{-1} \oint \Phi \, dx \, dp, \quad (32.31)$$

where $\oint dx d(pz) = 0$. This form is useful with non-Boussinesq models, in which pressure is a natural vertical coordinate.

32.2.4 Comments

A numerical realization of the integrated contact pressure force requires a representation of pressure 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.

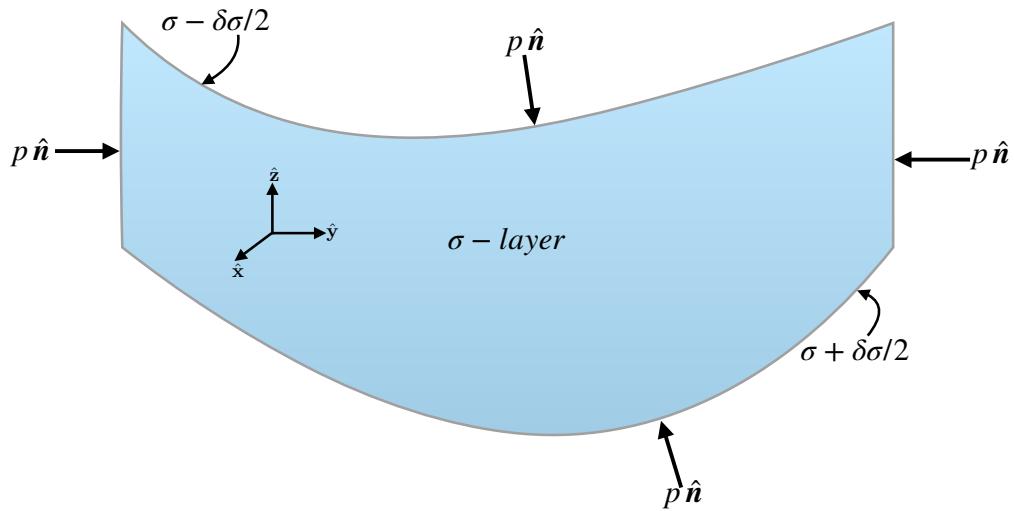


Figure 32.2: Schematic of pressure forces acting on the boundaries of a finite region such as a discrete model grid cell. In generalized vertical coordinate models, the side faces are vertical, so that pressure acts only in the horizontal directions. The top and bottom faces are defined by surfaces of constant generalized vertical coordinates with depth $\sigma(x, y, z, t) = \text{constant}$. We assume that these surfaces have an outward normal that has a nonzero projection into the vertical so that we can write the depth of a point on these surfaces as $z = \eta(x, y, t)$. Because of the slope of the top and bottom surfaces, pressure has both a horizontal and vertical component when acting on these surfaces. The net pressure acting on the grid cell is given by the area integral of the pressures around the grid cell boundary.

33

Time stepping the ocean equations

In this chapter we present elements of how to time step the fluid state using a numerical model. In particular, we discuss two foundational methods used to time step the ocean equations. The first is the split-explicit method used for the hydrostatic momentum equation. The second is the vertical Lagrangian-remap method used with hybrid vertical coordinate models. We close the chapter by outlining how to diagnose fluid particle trajectories by making use of the vertical Lagrangian-remap method.

READER'S GUIDE TO THIS CHAPTER

We require a working knowledge of the material in Chapter 32. For most purposes Cartesian horizontal coordinates are sufficient. Although exposing some of the details of numerical methods, the treatment is terse relative to the full details provided in the specialized literature.

33.1	Time stepping the non-hydrostatic fluid	488
33.1.1	Compressible perfect fluid equations	488
33.1.2	Comments about non-hydrostatic ocean models	490
33.1.3	Further study	491
33.2	Evolving the Boussinesq fluid forward in time	491
33.3	A split-explicit algorithm for hydrostatic models	491
33.3.1	Two common vertical coordinates	492
33.3.2	Depth integrated mass and volume budgets	493
33.3.3	Fast and slow motions	493
33.3.4	The pressure gradient for Boussinesq fluids	494
33.3.5	The pressure gradient for non-Boussinesq fluids	495
33.3.6	Depth integrated momentum equation	496
33.3.7	Summary of the split-explicit algorithm	497
33.3.8	Comments	498
33.4	The vertical Lagrangian-remap algorithm	498
33.4.1	Ocean model equations	499
33.4.2	Outlining the algorithm	499
33.4.3	Two specific examples	502
33.4.4	Connection to operator splitting	503
33.4.5	Comments	504
33.5	Numerically diagnosing fluid particle trajectories	504
33.5.1	Basics of estimating fluid particle trajectories	504
33.5.2	High wave number power in the vertical velocity	504
33.5.3	Trajectories from the vertical Lagrangian-remap algorithm	505
33.5.4	Interpolation versus extrapolation	505

33.1 Time stepping the non-hydrostatic fluid

A fundamental aim of numerical modeling of the atmosphere and ocean is the development of efficient and accurate methods for evolving the fluid state forward in time. As an introduction to these notions, we here consider elements for time stepping equations (22.14)-(22.18). For simplicity, we assume Cartesian coordinates and drop rotation and subgrid scale processes (i.e., no diffusion of momentum or tracers). We also ignore boundary conditions, allowing us to focus exclusively on questions related to evolving (i.e., time stepping) the fluid state. Although highly simplified, the presentation serves to introduce the basics of numerical modeling while offering insight into the physics and maths of the governing equations.

33.1.1 Compressible perfect fluid equations

Most numerical models are based on the Eulerian form of the governing equations since most spatial grids are static rather than moving with the fluid. The one notable exception concerns a moving vertical grid position, which takes its most general form in quasi-Lagrangian vertical coordinate models. We discuss such generalized vertical coordinate approaches in Chapter 32, and in particular the vertical Lagrangian-remap algorithm in Section 33.4. However, for the present discussion we assume the spatial grid is fixed. We thus consider the governing equations for a compressible perfect

fluid in a gravitational field as expressed in their Eulerian flux-form

$$\partial_t(\rho u) + \nabla \cdot [(\rho u) \mathbf{v}] = -\partial_x p \quad (33.1)$$

$$\partial_t(\rho v) + \nabla \cdot [(\rho v) \mathbf{v}] = -\partial_y p \quad (33.2)$$

$$\partial_t(\rho w) + \nabla \cdot [(\rho w) \mathbf{v}] = -\partial_z p - \rho g \quad (33.3)$$

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (33.4)$$

$$\partial_t(\rho \theta) + \nabla \cdot (\rho \theta \mathbf{v}) = 0 \quad (33.5)$$

$$\partial_t(\rho S) + \nabla \cdot (\rho S \mathbf{v}) = 0 \quad (33.6)$$

$$\rho = \rho(S, \theta, p). \quad (33.7)$$

As seen in our discussion of scaling for the hydrostatic balance in Section 27.3, large-scale ocean and atmosphere motion is very close to a hydrostatic balance. Even if considering small scale motion, it proves useful to split the pressure field into its hydrostatic component and a non-hydrostatic perturbation¹

$$p = p_h + p_{nh} \quad (33.8)$$

where

$$\frac{\partial p}{\partial z} = -\rho g + \frac{\partial p_{nh}}{\partial z}. \quad (33.9)$$

Making use of this split allows us to write the momentum equations (33.1)-(33.3) in the form

$$\partial_t(\rho u) + \partial_x[(\rho u) u] + \partial_y[(\rho u) v] + \partial_z[(\rho u) w] = -\partial_x p \quad (33.10)$$

$$\partial_t(\rho v) + \partial_x[(\rho v) u] + \partial_y[(\rho v) v] + \partial_z[(\rho v) w] = -\partial_y p \quad (33.11)$$

$$\partial_t(\rho w) + \partial_x[(\rho w) u] + \partial_y[(\rho w) v] + \partial_z[(\rho w) w] = -\partial_z p_{nh}. \quad (33.12)$$

With this decomposition, we see that the vertical acceleration is driven by the vertical derivative of the non-hydrostatic pressure field, since the hydrostatic portion of the pressure balances the gravitational force due to the weight per area of the fluid. In contrast, the horizontal acceleration is driven by the horizontal gradient of the full pressure field.

Isolating the time tendency for the momentum per mass, $\rho \mathbf{v}$, the mass density, ρ , and the density weighted tracers $\rho \theta$, ρS renders

$$\partial_t(\rho u) = -\partial_x p - \partial_x[(\rho u) u] - \partial_y[(\rho u) v] - \partial_z[(\rho u) w] \quad (33.13)$$

$$\partial_t(\rho v) = -\partial_y p - \partial_x[(\rho v) u] - \partial_y[(\rho v) v] - \partial_z[(\rho v) w] \quad (33.14)$$

$$\partial_t(\rho w) = -\partial_z p_{nh} - \partial_x[(\rho w) u] - \partial_y[(\rho w) v] - \partial_z[(\rho w) w] \quad (33.15)$$

$$\partial_t(\rho \theta) = -\nabla \cdot (\rho \theta \mathbf{v}) \quad (33.16)$$

$$\partial_t(\rho S) = -\nabla \cdot (\rho S \mathbf{v}) \quad (33.17)$$

$$\partial_t \rho = -\nabla \cdot (\rho \mathbf{v}) \quad (33.18)$$

$$\rho = \rho(S, \theta, p). \quad (33.19)$$

To develop the rudiments of a time stepping algorithm, assume knowledge at an arbitrary time τ of the fluid state as determined by the fields $\rho, \mathbf{v}, \theta, S, p$. Additionally, a knowledge of density allows us to diagnose the hydrostatic pressure through vertically integrating $\partial p_h / \partial z = -\rho g$. This initial condition information is sufficient to compute all the terms on the right hand side of the tendency

¹We perform a similar decomposition of pressure when writing the Boussinesq momentum equation in Section 28.1.2.

equations (33.13)-(33.18), evaluated at time τ , thus offering us the ability to time step the fluid state to time $\tau + \Delta\tau$.

To be specific, we start by noting that knowledge of the full pressure, hydrostatic pressure, and velocity at time τ allows us to time step the momentum equations (33.13)-(33.15) to determine the linear momentum per volume, $\rho \mathbf{v}$, at the new time $\tau + \Delta\tau$. Likewise, knowledge of the velocity, density, and tracer at time τ allows us to time step the tracer equations (33.16)-(33.17) to estimate the density-weighted tracers $\rho \theta$, and ρS at time $\tau + \Delta\tau$. To determine the unweighted fields requires us to update the density by time stepping the continuity equation (33.18), and then dividing to get

$$\mathbf{v}(\tau + \Delta\tau) = \frac{(\rho \mathbf{v})(\tau + \Delta\tau)}{\rho(\tau + \Delta\tau)} \quad (33.20)$$

$$\theta(\tau + \Delta\tau) = \frac{(\rho \theta)(\tau + \Delta\tau)}{\rho(\tau + \Delta\tau)} \quad (33.21)$$

$$S(\tau + \Delta\tau) = \frac{(\rho S)(\tau + \Delta\tau)}{\rho(\tau + \Delta\tau)}. \quad (33.22)$$

The updated density also allows us to update the hydrostatic pressure. Finally, the equation of state (33.19) is used to update the full pressure field through

$$\rho(\tau + \Delta\tau) = \rho[S(\tau + \Delta\tau), \theta(\tau + \Delta\tau), p(\tau + \Delta\tau)]. \quad (33.23)$$

For an ideal gas, this expression is trivial to invert for the updated pressure. In contrast, a realistic ocean equation of state requires an algebraic iterative solver. At this point we have updated the full fluid state to the new time, $\tau + \Delta\tau$, in which case we can again move forward in time to $\tau + 2\Delta\tau$.

33.1.2 Comments about non-hydrostatic ocean models

Many non-hydrostatic ocean circulation models are based on the Boussinesq equations, so that the velocity field is non-divergent. The main motivation for making the Boussinesq approximation is to filter out the acoustic modes, thus allowing for larger time steps. However, the simplicity of the compressible fluid algorithm warrants some attention. Furthermore, the computational cost of inverting the elliptic operator in the incompressible fluid can be nontrivial, especially for a complex ocean domain geometry. Consequently, the cost for the elliptic operator inversion can bring the incompressible case more in line with the compressible. Additionally, there may be means to artificially slow down the acoustic modes to enable a larger time step for the compressible case, thus making the compressible case more efficient.

One advantage of the oceanic Boussinesq non-hydrostatic equations concerns the needs for process modeling. For many purposes, it is useful to consider an idealized equation of state in which density is a linear function of potential temperature and/or salinity and is independent of pressure. However, it is notable that density must be a function of pressure in a compressible fluid. The fundamental reason is that a compressible fluid experiences pressure work as part of its internal energy evolution (see Section 23.2). Pressure work changes the volume of a constant mass fluid element, which means that it changes the density of the fluid element. Hence, pressure dependence to density is a basic feature of a compressible fluid. The ideal gas is the canonical example. We therefore must maintain a pressure dependence to density when using the compressible fluid equations.

33.1.3 Further study

The comments here concerning non-Boussinesq and non-hydrostatic ocean modeling remain untested in practice. For the atmosphere, [Chen et al. \(2013\)](#) present an algorithm for numerically solving the equations for a non-hydrostatic compressible atmosphere, with non-hydrostatic atmospheric models becoming quite common due to interests in resolving cloud processes.

33.2 Evolving the Boussinesq fluid forward in time

In Section 33.1 we presented the rudiments of how to evolve the fluid state forward in time. These ideas form the foundation for developing a numerical prediction algorithm. We here expose some of the unique features of a Boussinesq fluid.

A key element of the Boussinesq system is that when evaluating density, we must determine pressure according to (see equation (28.16))

$$\rho = \rho(S, \theta, p = -g \rho_0 z) \quad \text{Boussinesq density.} \quad (33.24)$$

This form of the density is required to maintain energetic consistency of the Boussinesq system (see Section 2.4.3 of [Vallis \(2017\)](#) for details). Consequently, we can no longer make use of the equation of state to diagnose pressure as we did for the compressible case in equation (33.23). The alternative required for the incompressible case is found by taking the divergence of the momentum equation. Since $\nabla \cdot \mathbf{v} = 0$ (incompressible), we can eliminate the time derivative, thus leading to a diagnostic (elliptic) equation for the pressure field. This equation is derived in Exercise (28.2) for a rotating Boussinesq fluid.

The key point of this discussion is the fundamental distinction between how we diagnose pressure in a compressible fluid (through the equation of state) versus an incompressible fluid (inverting an elliptic operator). Notably, the inversion of a 3d elliptic operator is generally difficult in a complex geometry such as the ocean. Hence, the compressible case is algorithmically far simpler. However, the price to pay for simplicity in the compressible case is the need to take time steps small enough to resolve acoustic modes present in a compressible fluid.

33.3 A split-explicit algorithm for hydrostatic models

In this section we describe an algorithm used to time step the hydrostatic momentum equations for the ocean. Motivation for the algorithm is the observation that the most rapid dynamical motions in a hydrostatic fluid correspond to the shallow water gravity waves of Chapter 44. Depending on the ocean depth, these *external* or *barotropic* gravity waves move at 50 to 100 times faster than the first baroclinic waves and particle motion due to advection.

We thus seek an algorithm that sufficiently splits the motion into slow and fast components, ideally with the two motions at most weakly interacting. The fast motion is associated with motion of the full fluid column (e.g., single layer shallow water gravity waves), and the slow motion is associated with deviations from the depth integrated motion. This split leads to the *split-explicit* method used for hydrostatic ocean models. Our presentation of this method represents a summary of the more detailed discussion offered in Chapter 12 of [Griffies \(2004\)](#), and that discussion is itself a mere summary of many details provided by the specialized literature.

A fundamental element of the split is to decompose the pressure force into a slow component and a fast component. We follow the traditional approach whereby the pressure force is represented as a

gradient. The alternative approach described in Section 32.2, which focuses on the area integrated contact pressure force, can also be split. *Hallberg and Adcroft* (2009) provide a working example.

To keep our discussion reasonably streamlined we make use of Cartesian horizontal coordinates. Furthermore, we use the Eulerian flux-form momentum equation derived in Section 32.1.3

$$\frac{\partial(h \rho \mathbf{u})}{\partial t} + \nabla_\sigma \cdot [\mathbf{u}(h \rho \mathbf{u})] + \delta_\sigma(\rho \mathbf{u} w^{(\dot{\sigma})}) + f h \rho \hat{\mathbf{z}} \wedge \mathbf{u} = -h [\nabla_\sigma p + \rho \nabla_\sigma \Phi] + \rho h \mathbf{F}. \quad (33.25)$$

33.3.1 Two common vertical coordinates

As part of our formulation we introduce two vertical coordinates commonly used in ocean modeling. One is suitable for a compressible non-Boussinesq fluid whereas the other is used for incompressible Boussinesq fluids. These vertical coordinates are often used in tandem with the isopycnal coordinate of Section 45.1 as part of a hybrid vertical coordinate algorithm. In Section 33.4 we describe the vertical Lagrangian-remap algorithm used to time step a Lagrangian hybrid vertical coordinate model. The split-explicit method described in the present section is a necessary component to the vertical Lagrangian-remap algorithm.

A vertical coordinate for compressible/non-Boussinesq fluids

Consider the vertical coordinate defined by

$$\sigma = p^* = p_b^0 \left[\frac{p - p_a}{p_b - p_a} \right] \implies 0 \leq p^* \leq p_b^0. \quad (33.26)$$

In this equation we made use of

$$p_b - p_a = g \int_{-H}^{\eta} \rho dz = g \sum_k h \rho, \quad (33.27)$$

which is the hydrostatic pressure at the ocean bottom due to the overlying water, and where p_a is the pressure applied to the ocean surface from the overlying atmosphere or ice. The second equality in equation (33.27) introduced the discrete vertical sum over the k grid cells in a numerical ocean model. Finally, the pressure

$$p_b^0 = \rho_0 g H, \quad (33.28)$$

which is the hydrostatic pressure at the ocean bottom in a uniform density ocean with a resting ocean surface. Notably, the density times layer thickness takes the form

$$\rho h = \rho \frac{\partial z}{\partial \sigma} d\sigma = \rho \frac{p_b - p_a}{p_b^0} (-\rho g)^{-1} d\sigma = -(1/g) \frac{p_b - p_a}{p_b^0} dp^*. \quad (33.29)$$

The pressure increment, dp^* , is negative since pressure decreases moving upward in the fluid column. Consequently, the layer thickness is a positive number

$$dp^* < 0 \implies h > 0, \quad (33.30)$$

and the depth integrated layer thickness times density yields the expected result

$$\sum_k h \rho = -(1/g) \frac{p_b - p_a}{p_b^0} \sum_k dp^* = (1/g) (p_b - p_a). \quad (33.31)$$

The integral was performed as a sum over a full column of fluid from the ocean bottom, where $p^* = p_b^0$, to the surface where $p^* = 0$.

A vertical coordinate for incompressible/Boussinesq fluids

An analogous choice for vertical coordinate appropriate for an incompressible/Boussinesq fluid was suggested by [Stacey et al. \(1995\)](#) and [Adcroft and Campin \(2004\)](#), in which

$$\sigma = z^* = H \left[\frac{z - \eta}{H + \eta} \right] \implies -H \leq z^* \leq 0, \quad (33.32)$$

so that the layer thickness is given by

$$h = \frac{\partial z}{\partial z^*} dz^* = (1 + \eta/H) dz^*. \quad (33.33)$$

Consequently, the depth integrated layer thickness is the expected result

$$\sum_k h = (1 + \eta/H) \sum_k dz^* = H + \eta. \quad (33.34)$$

33.3.2 Depth integrated mass and volume budgets

The mass per horizontal area of fluid contained within a full column of ocean water is given by $\sum_k \rho h$. From Exercise 17.2 we know that the time change for this column mass is determined by equation (17.89)

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} \rho dz \right] = -\nabla \cdot \mathbf{U}^\rho + Q_m, \quad (33.35)$$

where Q_m is the mass per time per area of matter that crosses the ocean surface. Hence, the mass per area of a fluid column changes due to the horizontal convergence of mass moving with the flow, plus the mass entering or leaving through the surface boundary. Making the hydrostatic approximation for the mass per area in the fluid column leads to

$$g^{-1} \partial_t (p_b - p_a) = -\nabla \cdot \mathbf{U}^\rho + Q_m. \quad (33.36)$$

We see that the difference between the bottom pressure and the applied surface pressure changes in time depending on the mass converging into the column. Similarly, in Section 19.3 we determine the free surface equation, which expresses the conservation of volume for a column of Boussinesq fluid

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}, \quad (33.37)$$

where $\mathbf{U} = \int_{-H}^{\eta} \mathbf{u} dz$ is the depth-integrated horizontal velocity. Hence, as volume enters the column either through the surface or through horizontal convergence, it causes the sea level to rise.

The equations for bottom pressure (for non-Boussinesq fluids) and sea surface (for Boussinesq fluids) form the kinematic portion of the split-explicit algorithm. Both the bottom pressure and the sea level fluctuate as external gravity waves move through the fluid. Hence, these kinematic equations must be time stepped using a time step sufficiently small to resolve these fast wave motions.

33.3.3 Fast and slow motions

An algorithm of practical utility for large-scale ocean modeling must split the fast wave motions from the slower waves and advection. When doing so, the slow dynamics can be updated with a much longer time step than the fast dynamics. Indeed, for large-scale ocean simulations, the time

step available for an time update of the slow dynamics is much larger (50 to 100 times larger) than the fast dynamics. It is this large time split, and the attendant improved model efficiency, that motivate the complexities associated with splitting the flow into fast and slow components.

For a hydrostatic fluid, the fast motions can be approximated by the vertically integrated dynamics. The remainder constitutes an approximation to the slow dynamics. Motions constituting the fast dynamics are often referred to as the *barotropic* or *external* mode, and the slower motions are embodied by advection as well as the *baroclinic* or *internal* mode. This terminology originates from an orthogonal eigenmode decomposition of the linearized momentum equations for a flat bottom hydrostatic fluid (e.g., Chapter 6 in [Gill \(1982\)](#)). The depth integrated and depth deviation “modes” computed in an ocean model only approximate these eigenmodes. Even so, the terminology remains.

We formulate the split between the fast and slow modes using density weighting, as appropriate for a compressible non-Boussinesq fluid. For a Boussinesq fluid the density weighting reduces to an extra ρ_0 factor that trivially cancels. We thus consider the following split of the horizontal velocity

$$\mathbf{u} = \underbrace{\left[\mathbf{u} - \frac{\sum_k \mathbf{u} \rho h}{\sum_k \rho h} \right]}_{\text{slow}} + \underbrace{\left[\frac{\sum_k \mathbf{u} \rho h}{\sum_k \rho h} \right]}_{\text{fast}} \equiv \hat{\mathbf{u}} + \bar{\mathbf{u}}^z. \quad (33.38)$$

The depth-averaged or barotropic velocity is given by

$$\bar{\mathbf{u}}^z = \frac{\mathbf{U}^\rho}{\sum_k \rho h} \quad (33.39)$$

and it is time stepped according to the vertically integrated momentum equation derived later in this section.

By construction, the baroclinic velocity $\hat{\mathbf{u}}$ has zero density weighted vertical integral,

$$\sum_k \rho h \hat{\mathbf{u}} = 0. \quad (33.40)$$

Consequently, its time tendency is independent of any depth independent forcing, such as from the fast fluctuations in the surface height associated with external gravity waves. Therefore, we can choose to update the slow dynamics using all pieces of the momentum equation forcing except for those contributions from the rapid fluctuating pressure forces. Performing an update as such produces a velocity field that we write as \mathbf{u}' . This velocity is related to the baroclinic velocity via

$$\hat{\mathbf{u}} = \mathbf{u}' - \frac{\sum_k \mathbf{u}' \rho h}{\sum_k \rho h}. \quad (33.41)$$

Completing the updates of \mathbf{u}' and \mathbf{U}^ρ allows for an update of the full horizontal velocity via

$$\mathbf{u} = \left[\mathbf{u}' - \frac{\sum_k \mathbf{u}' \rho h}{\sum_k \rho h} \right] + \frac{\mathbf{U}^\rho}{\sum_k \rho h}. \quad (33.42)$$

33.3.4 The pressure gradient for Boussinesq fluids

The *in situ* density in the bulk of the ocean deviates less than 3% from the constant density (see page 47 of [Gill \(1982\)](#))

$$\rho_0 = 1035 \text{ kg m}^{-3}. \quad (33.43)$$

The hydrostatic pressure associated with this constant density has no horizontal gradients, and so it does not contribute to horizontal pressure gradient forces. Hence, for increased accuracy computing the horizontal pressure gradient, it is useful to remove this contribution from calculation of hydrostatic pressure. We can do so by writing the hydrostatic balance as

$$\frac{\partial p}{\partial z} = -g \rho = -g (\rho_o + \rho'), \quad (33.44)$$

which has an associated split in the hydrostatic pressure field

$$p = p_a(x, y, t) + p_o(x, y, t) + p'(x, y, z, t), \quad (33.45)$$

where $p_a(x, y, t)$ is the pressure applied to the surface of the ocean from the overlying atmosphere or ice. We can solve for the pressures by assuming

$$p_o(z = \eta) = 0 \quad \text{and} \quad p'(z = \eta) = 0, \quad (33.46)$$

which leads to

$$p_o = -g \rho_o (z - \eta) = -\rho_o (\Phi - g \eta), \quad (33.47)$$

and

$$p' = g \int_z^\eta \rho' dz, \quad (33.48)$$

so that

$$p = p_a + g \rho_o \eta - \rho_o \Phi + p'. \quad (33.49)$$

Isolating the free surface height allows for a split of the pressure gradient into its fast two dimensional contributions, associated with undulations in the free surface, and slow three dimensional contributions, associated with undulations of interior density interfaces. We perform that split as follows

$$\nabla_z p = \nabla_\sigma p + \rho \nabla_\sigma \Phi \quad (33.50a)$$

$$= \nabla_\sigma (p_a + g \rho_0 \eta - \rho_0 \Phi + p') + \rho \nabla_\sigma \Phi \quad (33.50b)$$

$$= \underbrace{\nabla (p_a + g \rho_0 \eta)}_{\text{fast}} + \underbrace{\nabla_\sigma p' + \rho' \nabla_\sigma \Phi}_{\text{slow}}. \quad (33.50c)$$

We have *assumed* that the geopotential falls inside the slow portion of the pressure gradient. The utility of this assumption can be assessed by the integrity and stability of the simulation.

33.3.5 The pressure gradient for non-Boussinesq fluids

We now offer the complementary discussion for pressure based vertical coordinates. As seen in Section 33.3.2, for pressure based vertical coordinates used in hydrostatic non-Boussinesq fluids, we solve for the bottom pressure rather than the free surface height. It is therefore useful to formulate the geopotential in terms of the bottom pressure rather than the atmospheric pressure.

For this purpose, consider the following identities

$$\Phi + g H = g \int_{-H}^z dz \quad (33.51a)$$

$$= g \int_{p_b}^p \frac{\partial z}{\partial p} dp \quad (33.51b)$$

$$= - \int_{p_b}^p \rho^{-1} dp \quad (33.51c)$$

$$= - \int_{p_b}^p (\rho_o^{-1} + \rho^{-1} - \rho_o^{-1}) dp \quad (33.51d)$$

$$= (p_b - p)/\rho_o + \rho_o^{-1} \int_{p_b}^p (\rho'/\rho) dp \quad (33.51e)$$

$$= (p_b - p)/\rho_o - (g/\rho_o) \int_{-H}^z \rho' dz. \quad (33.51f)$$

We are thus led to the identity

$$\rho_o \Phi = p_b - p + \rho_o (\Phi_b + \Phi'), \quad (33.52)$$

where

$$\rho_o \Phi' = -g \int_{-H}^z \rho' dz \quad (33.53)$$

is an anomalous geopotential similar to the anomalous hydrostatic pressure introduced in Section 33.3.4, and

$$\Phi_b = -g H \quad (33.54)$$

is the geopotential at the ocean bottom. The horizontal pressure gradient is therefore written

$$\nabla_\sigma p + \rho \nabla_\sigma \Phi = \nabla_\sigma p + (\rho/\rho_o) \nabla (p_b + \rho_o \Phi_b) - (\rho/\rho_o) \nabla_\sigma p + \rho \nabla_\sigma \Phi' \quad (33.55a)$$

$$= \underbrace{(\rho/\rho_o) \nabla (p_b + \rho_o \Phi_b)}_{\text{fast}} + \underbrace{\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p}_{\text{slow}}. \quad (33.55b)$$

33.3.6 Depth integrated momentum equation

The thickness weighted horizontal momentum equation (33.25) is already integrated over the thickness of a grid cell. Hence, to produce the depth integrated momentum equations requires a sum over all cells from the ocean bottom to the surface

$$\begin{aligned} (\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho &= - \sum_k \nabla_\sigma \cdot [\mathbf{u} (h \rho \mathbf{u})] - \sum_k h [\nabla_\sigma p + \rho \nabla_\sigma \Phi] + \sum_k \rho h \mathbf{F} \\ &\quad + [\boldsymbol{\tau}^{\text{surf}} - \boldsymbol{\tau}^{\text{bott}} + Q_m \mathbf{u}_m]. \end{aligned} \quad (33.56)$$

On the left hand side we wrote

$$(\partial_t + f \hat{z} \wedge) \sum_k h \rho \mathbf{u} = (\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho. \quad (33.57)$$

On the right hand side of equation (33.56) we identified the mechanical boundary fluxes from surface air-sea and ice-sea interactions, $\boldsymbol{\tau}^{\text{surf}}$, as well as bottom stresses, $\boldsymbol{\tau}^{\text{bott}}$. Likewise, the momentum input through the surface due to mass exchanges appears in the term $Q_m \mathbf{u}_m$, with Q_m the mass flux crossing the ocean surface (mass per time per area), and \mathbf{u}_m the horizontal velocity of that mass flux.

Pressure-based vertical coordinates

We now make use of the pressure gradient decomposition (33.55b) to write

$$\sum_k h [\nabla_\sigma p + \rho \nabla_\sigma \Phi] = \sum_k (\rho/\rho_0) h \nabla (p_b + \rho_o \Phi_b) + \sum_k h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p] \quad (33.58a)$$

$$= \frac{p_b - p_a}{g \rho_0} \nabla (p_b + \rho_o \Phi_b) + \sum_k h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p], \quad (33.58b)$$

so that the depth-integrated momentum equation (33.56) takes on the form

$$\begin{aligned} (\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho + \frac{p_b - p_a}{g \rho_0} \nabla (p_b + \rho_o \Phi_b) &= - \sum_k h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p] \\ &\quad - \sum_k \nabla_\sigma \cdot [\mathbf{u} (h \rho \mathbf{u})] + \sum_k \rho h \mathbf{F} + [\boldsymbol{\tau}^{\text{surf}} - \boldsymbol{\tau}^{\text{bott}} + Q_m \mathbf{u}_m]. \end{aligned} \quad (33.59)$$

The split-explicit method assumes that all terms on the right hand side are slow whereas those terms on the left hand side are fast. Indeed, linearizing the left hand side results in the linear shallow water equations whose fluctuations are linear waves moving at speed \sqrt{gH} (see Chapter 44). Therefore, we sub-cycle these fast wave motions by taking small time steps for the left hand side while keeping the right hand side terms constant over the small time steps. Interleaving the fast and slow motions requires careful meshing of the time stepping schemes, with details outside of our scope.

Depth-based vertical coordinates

For the Boussinesq fluid, we set density factors to the reference value, ρ_0 , and make use of the decomposition (33.50c) of the pressure gradient. The resulting depth-integrated horizontal momentum equation is given by

$$\begin{aligned} (\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho + (H + \eta) \nabla (p_a + g \rho_0 \eta) &= - \sum_k h [\nabla_\sigma p' + \rho' \nabla_\sigma \Phi] \\ &\quad - \sum_k \nabla_\sigma \cdot [\mathbf{u} (h \rho_0 \mathbf{u})] + \sum_k \rho_0 h \mathbf{F} + [\boldsymbol{\tau}^{\text{surf}} - \boldsymbol{\tau}^{\text{bott}} + Q_m \mathbf{u}_m]. \end{aligned} \quad (33.60)$$

Again, the left hand side terms constitute the fast wave dynamics that must be sub-cycled with the slower terms on the right hand side.

33.3.7 Summary of the split-explicit algorithm

The split-explicit algorithm can be summarized by the following steps.

Fast motions

The fast motions are captured by the depth integrated kinematics and dynamics. For pressure-based non-Boussinesq models these equations are

$$g^{-1} \partial_t (p_b - p_a) = -\nabla \cdot \mathbf{U}^\rho + Q_m \quad (33.61)$$

$$(\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho + \frac{p_b - p_a}{g \rho_0} \nabla (p_b + \rho_o \Phi_b) = \mathbf{G}, \quad (33.62)$$

where \mathbf{G} symbolizes those terms on the right hand side of equation (33.59). All terms in \mathbf{G} are typically held constant while time stepping the fast dynamics. Likewise, for depth-based Boussinesq models we have the depth integrated equations

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}, \quad (33.63)$$

$$(\partial_t + f \hat{\mathbf{z}} \wedge) \mathbf{U}^\rho + (H + \eta) \nabla (p_a + g \rho_0 \eta) = \mathbf{G}. \quad (33.64)$$

Slow motions

For the slow dynamics we time step the momentum equation (and the tracer equations) using the relatively large time step. To ensure the momentum equation remains stable with the long time steps requires us to remove, or approximately remove, the forcing that contributes to fast barotropic wave motions. We do so by dropping the fast portion from the pressure gradient. Doing so then leads us to determine the \mathbf{u}' velocity discussed in Section 33.3.3 according to

$$\frac{\partial(h \rho \mathbf{u}')}{\partial t} + \nabla_\sigma \cdot [\mathbf{u} (h \rho \mathbf{u})] + \delta_\sigma(\rho \mathbf{u} w^{(\dot{\sigma})}) + f h \rho \hat{\mathbf{z}} \wedge \mathbf{u} = -h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p] + \rho h \mathbf{F}. \quad (33.65)$$

Once \mathbf{u}' is updated we compute the baroclinic velocity according to equation (33.41)

$$\hat{\mathbf{u}} = \mathbf{u}' - \frac{\sum_k \mathbf{u}' \rho h}{\sum_k \rho h}. \quad (33.66)$$

The full velocity is then updated by

$$\mathbf{u} = \hat{\mathbf{u}} + \frac{\mathbf{U}^\rho}{\sum_k \rho h}. \quad (33.67)$$

33.3.8 Comments

As already noted, further details of the split-explicit algorithm take us deep into the world of numerical ocean modeling, well beyond that intended for this book. The interested reader is encouraged to start penetrating that literature by reading Chapter 12 of [Griffies \(2004\)](#), which iteself is a pedagogical summary of the paper by [Griffies et al. \(2001\)](#) (split-explicit free surface method for the Modular Ocean Model versions 3,4, and 5). Thereafter, the specialized literature should be a bit more accessible, with notable papers those by [Killworth et al. \(1991\)](#) (the precursor to the [Griffies et al. \(2001\)](#) method); [Shchepetkin and McWilliams \(2005\)](#) (time stepping algorithm for the Regional Ocean Modeling System (ROMS)); and [Hallberg \(1997\)](#) and [Hallberg and Adcroft \(2009\)](#) (time stepping for the Modular Ocean Model version 6).

33.4 The vertical Lagrangian-remap algorithm

[Adcroft and Hallberg \(2006\)](#) identified two algorithm classes that frame how ocean models are formulated, with the two classes differing in how they treat the vertical coordinate. A vertical Eulerian algorithm is traditionally used in geopotential and terrain following ocean models, whereby vertical fluid motion is diagnosed via the continuity equation. The second class employs a vertical Lagrangian approach as traditionally used by layered isopycnal models. This method specifies fluid motion that crosses coordinate surfaces. Indeed, such dia-surface flow can be set identically to zero for studies of adiabatic geophysical fluid dynamics.

As suggested by [Griffies et al. \(2000a\)](#), the use of hybrid vertical coordinates has emerged as a promising method to reduce certain biases, such as the spurious diapycnal mixing identified by [Griffies et al. \(2000b\)](#). A common choice combines isopycnal coordinates in the ocean interior with the z^* or p^* coordinates in unstratified regions (see Section 33.3.1). In principle, either an Eulerian or Lagrangian vertical algorithm is available with hybrid vertical coordinates. However, in practice the Lagrangian algorithm is favored given its precise control over dia-surface transport. In this section we outline a vertical Lagrangian-remap algorithm as implemented in the Modular Ocean Model version 6 (MOM6) ([Adcroft et al., 2019](#)).

33.4.1 Ocean model equations

We consider the Boussinesq hydrostatic ocean primitive equations formulated in their generalized vertical coordinate form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} + w^{(\dot{\sigma})} \frac{\partial \mathbf{u}}{\partial z} = -\nabla_\sigma (K + p/\rho_0) - (\rho/\rho_0) \nabla_\sigma \Phi + \mathbf{F}^h \quad \text{momentum (33.68)}$$

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad \text{hydrostatic (33.69)}$$

$$\left[\frac{\partial h}{\partial t} \right]_\sigma + \nabla_\sigma \cdot (h \mathbf{u}^\dagger) + \delta_\sigma w^{(\dot{\sigma})} = 0 \quad \text{thickness (33.70)}$$

$$\left[\frac{\partial (h C)}{\partial t} \right]_\sigma + \nabla_\sigma \cdot (h C \mathbf{u}^\dagger + h \mathbf{J}^h) + \delta_\sigma (C w^{(\dot{\sigma})} + z_\sigma \nabla \sigma \cdot \mathbf{J}) = 0 \quad \text{tracer equation (33.71)}$$

$$\rho = \rho(S, \Theta, -g\rho_0 z) \quad \text{eqn of state. (33.72)}$$

Here are some characteristics of these equations and how they are handled in light of the vertical Lagrangian-remap algorithm.

- The horizontal momentum equation (33.68) is written using the advective vector-invariant form. The geopotential is $\Phi = g z$ and the kinetic energy per mass is $K = \mathbf{u} \cdot \mathbf{u}/2$.
- The thickness and tracer equations are written in flux-form to ensure conservation of volume and tracer content.
- The tracer flux \mathbf{J} parameterizes subgrid scale diffusive processes as well as boundary fluxes. The acceleration \mathbf{F}^h parameterizes subgrid momentum transport as well as boundary stresses.
- The parameterized subgrid scale eddy-induced advection is combined with the lateral advection of thickness and tracer, thus providing a residual mean advective transport, \mathbf{u}^\dagger , for the thickness and tracers (see Section 39.2). Furthermore, there is only a horizontal parameterized subgrid advective velocity, which we interpret as a layer bolus velocity as appropriate for vertical Lagrangian models. This interpretation contrasts to the three-dimensional eddy-induced quasi-Stokes velocity required for vertical Eulerian models (see [McDougall and McIntosh \(2001\)](#) for details). Figure 33.1 offers a schematic of these points.

33.4.2 Outlining the algorithm

Fluid particle motion is three-dimensional. However, the vertical Lagrangian-remap algorithm focuses on the vertical since a one-dimensional Lagrangian-remap algorithm is more tractable than a three-dimensional algorithm. Additionally, the generalized vertical coordinate equations (33.68)-(33.72) are ideally setup for a vertical Lagrangian-remap algorithm. The key to this suitability is

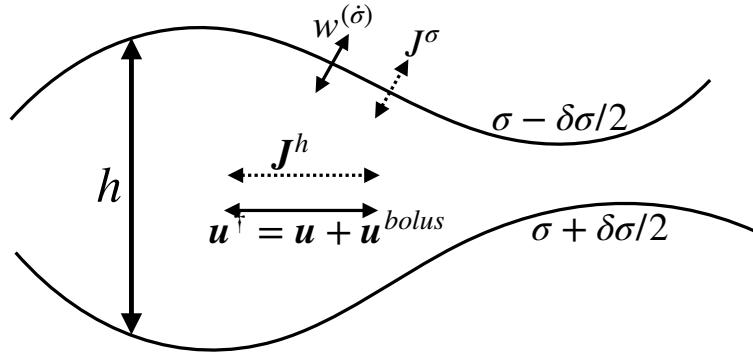


Figure 33.1: Illustrating how thickness and tracer concentration evolve in a generalized vertical coordinate ocean model using a vertical Lagrangian-remap algorithm. Thickness evolves according to the layer convergence of the horizontal thickness flux, $-\nabla_\sigma \cdot (h \mathbf{u}^\dagger)$, as well as the dia-surface convergence, $-\delta_\sigma w^{(\dot{\sigma})}$. Tracer evolves similarly along with horizontal and dia-surface flux convergences from diffusion and boundary fluxes, $-\nabla_\sigma \mathbf{J}^h - \delta_\sigma (z_\sigma \nabla_\sigma \cdot \mathbf{J})$. Air-sea, land-sea, and ice-sea boundary mass fluxes are incorporated through $w^{(\dot{\sigma})} \neq 0$, and further tracer boundary fluxes are incorporated through J^σ . Note that \mathbf{u}^\dagger is the horizontal residual mean velocity built from the resolved horizontal velocity plus a parameterized eddy-induced bolus velocity. There is no dia-surface component to the bolus velocity (see [McDougall and McIntosh \(2001\)](#) for details). In the Lagrangian portion of the Lagrangian-remap algorithm, the thickness and tracer concentration evolve with $w^{(\dot{\sigma})} = 0$, even while maintaining the full subgrid scale fluxes in the tracer equation. The remap step then rearranges the σ layers according to the prescribed target values. The remap step is equivalent to transport using $w^{(\dot{\sigma})} \neq 0$.

the dia-surface velocity component, $w^{(\dot{\sigma})}$, which measures transport of fluid crossing the σ surface (see Section 21.3). We here outline the vertical Lagrangian-remap algorithm, with Figure 33.2 providing a schematic of the main steps.

The vertical Lagrangian step

We introduce the Lagrangian vertical coordinate, σ , as a means to mark isosurfaces by following fluid elements. What does that mean? As discussed in Section 21.3, a surface maintaining $w^{(\dot{\sigma})} = 0$ means that no net fluid crosses the surface. If the fluid is homogeneous, then $w^{(\dot{\sigma})} = 0$ means the surface is material so that no matter crosses it. However, for ocean modeling we are interested in real fluids with multiple constituents and irreversible mixing. Mixing acts to exchange trace matter between fluid elements while keeping the volume (mass if considering a compressible non-Boussinesq fluid) of each fluid element constant (see Section 18.1). We thus consider a σ -surface in a real fluid that follows the fluid velocity (the barycentric velocity of Section 18.1) even as trace matter is irreversibly exchanged across the surface.

For the Lagrangian step in the algorithm, we ensure that the σ -surfaces vertically follow fluid elements by time stepping the fluid state with $w^{(\dot{\sigma})} = 0$. To allow trace matter and momentum to irreversibly cross the σ -surfaces, we retain diffusion and friction in the tracer and momentum equations. Consequently, the ocean equations time stepped during this Lagrangian portion of the

algorithm are the following

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_{\sigma} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_{\sigma} (K + p/\rho_0) - (\rho/\rho_0) \nabla_{\sigma} \Phi + \mathbf{F}^h \quad (33.73)$$

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad (33.74)$$

$$\left[\frac{\partial h}{\partial t} \right]_{\sigma} + \nabla_{\sigma} \cdot [h \mathbf{u}^{\dagger}] = 0 \quad (33.75)$$

$$\left[\frac{\partial (h C)}{\partial t} \right]_{\sigma} + \nabla_{\sigma} \cdot (h C \mathbf{u}^{\dagger} + h \mathbf{J}^h) + \delta_{\sigma}(z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}) = 0 \quad (33.76)$$

$$\rho = \rho(S, \Theta, -g\rho_0 z). \quad (33.77)$$

The second panel of Figure 33.2 illustrates this step of the algorithm.

Vertical regridding step

As emphasized above, the Lagrangian step in the algorithm allows for both reversible and irreversible changes in the ocean state, all while having the σ coordinate surfaces vertically follow fluid elements. If continued indefinitely, the σ -surfaces will generally drift far from their initial positions. If they drift too far, then they will be unable to accurately represent the ocean state. Hence, for the second step in the algorithm, we reinitialize the vertical grid.

We reinitialize the vertical coordinates by laying down a new σ -grid according to a prescribed regridding method and with a target coordinate layout. For example, we may choose a grid according to surfaces of constant geopotential placed at prescribed vertical positions, as illustrated in the third panel of Figure 33.2. Alternatively, we may choose σ according to prescribed potential density classes. Indeed, as emphasized by [Bleck \(2002\)](#), there is no need to provide an analytical expression for the vertical grid; one only needs a prescription. Furthermore, the vertical resolution of the grid is not specified so that it can change between time steps.

Vertical remapping step

Once the new grid is prescribed, the value of the ocean state is estimated on the new grid via interpolation (or extrapolation if needed). We say that this step remaps the ocean state from its Lagrangian-displaced grid onto the target grid. However, that language can be somewhat confusing since the ocean state *does not change* during the remapping. No fluid elements move nor does any trace matter. Rather, the vertical location of the σ -grid changes during regridding, and correspondingly our information regarding the fluid state changes as remapping estimates properties on the new grid. As seen in Section 33.4.4, vertical remapping corresponds operationally to advection though realized in a Lagrangian manner rather than Eulerian.

In practice, there is some spurious evolution during the remapping step since a discrete remapping will incur discretization error. Reducing these errors while maintaining scalar conservation (i.e., integrated scalar properties must remain unchanged by remapping) is a primary concern for vertical remapping methods, such as those developed by [White and Adcroft \(2008\)](#) and [White et al. \(2009\)](#). The fourth panel of Figure 33.2 illustrates this final step in the vertical Lagrangian-remap algorithm.

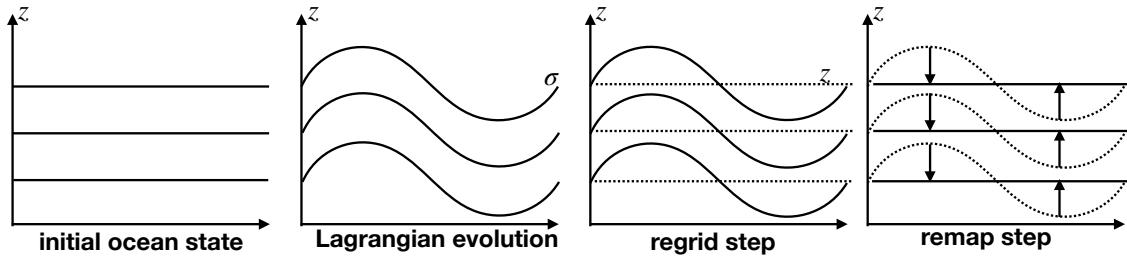


Figure 33.2: Illustrating the various steps for the vertical Lagrangian-remap algorithm used to time step the fluid state. The first panel shows an initial state where the σ -coordinate surfaces are horizontal. The second panel shows the vertical Lagrangian evolution of the σ -surfaces as determined by the immiscible thickness equation, $(\partial h / \partial t)_{\sigma} + \nabla_{\sigma} \cdot [h \mathbf{u}^{\dagger}] = 0$. The third panel shows the regrid step, whereby we identify a target σ -grid according to a user defined prescription. In this example the target grid is determined by constant geopotentials, but any other choices are possible. The final panel shows the remap step, whereby the ocean state is estimated at the vertical position of the target grid. Under a perfect remapping the ocean state does not evolve. Rather, remapping merely adjusts the location of the σ -coordinate surfaces, thus offering a new estimate of the ocean state.

33.4.3 Two specific examples

We further illustrate the algorithm by considering some specific examples. Although rather trivial, it is nonetheless useful to work through the time stepping as per the vertical Lagrangian-remap algorithm.

Remapping to geopotentials

Figure 33.2 illustrates the vertical Lagrangian-remap algorithm for the specific case of geopotential vertical coordinates. Hence, the σ -surfaces are initially horizontal. The Lagrangian portion of the algorithm allows for these σ surfaces to evolve without any mixing so that the volume within each of the layers remains fixed. The figure illustrates this step via a wave (e.g., a gravity wave) that causes the σ -surfaces to undulate. What is not shown in the figure is the evolution of other properties as affected by the wave as well as irreversible mixing processes. As the σ -surfaces are displaced by the wave, the remapping step reinitializes the σ -surfaces to their target geopotentials and estimates the ocean state at these target positions.

A trivial example occurs when there is zero motion of fluid elements, such as for a horizontally homogeneous ocean state without boundary forcing. If there are vertical tracer gradients and vertical mixing, then fluid elements remain stationary while tracer isolines evolve. The Lagrangian step of the algorithm retains σ -isolines unchanged since there is no horizontal motion. Hence, there is no need for a remapping step. All that happens is the tracers evolve according to vertical diffusion.

Remapping to isopycnals

Now consider the case in which σ is remapped according to the Conservative Temperarture, Θ , and assume potential density is linearly related to Θ . The Lagrangian step in the algorithm is identical to the case of σ -remapping to geopotentials. However, the remapping step differs. Here the remapping returns the σ -surfaces to their target Θ values in response to mixing that causes σ -surfaces to drift away from the targets. Without mixing, then σ -surfaces remain on their initial Θ surfaces and no remapping is needed.

As for the geopotential case, consider a horizontally homogeneous thought experiment with zero

motion yet with nonzero mixing. Again, the σ layer thickness remains static during the Lagrangian step while the Θ isolines evolve due to diffusion. As there is mixing, the σ coordinates lose their initial Θ values, thus requiring a remapping step to return the vertical σ -grid to its target values.

Another way to understand the need for remapping is to note that in the absence of vertical fluid motion, setting $w = 0$ in equation (21.63) means that the vertical position of a Θ -surface evolves according to

$$\left[\frac{\partial z}{\partial t} \right]_{\Theta} = -w^{(\dot{\Theta})}. \quad (33.78)$$

Vertical diffusion causes a nonzero $w^{(\dot{\Theta})}$ so that Θ -surfaces move during the Lagrangian portion of the algorithm. However, during this part of the algorithm the σ -surfaces remain fixed since there is no change in the σ -layer thicknesses. It is only during the remapping step that we bring the σ -surfaces back to their target Θ -surfaces. We do so by moving the σ -surfaces vertically by an amount determined by $\int_0^{\Delta t} w^{(\dot{\Theta})} dt$, with Δt the time step between remapping.

33.4.4 Connection to operator splitting

As noted by [Adcroft and Hallberg \(2006\)](#), the vertical Lagrangian-remap algorithm can be related to operator splitting. In this view, we split the dia-surface terms in equations (33.68)-(33.72) from the remaining terms. That is, we update the full equations in two parts. To exemplify this operator splitting perspective, consider just the thickness and tracer equations and assume a forward Euler time stepping scheme. All of this discussion generalizes to the momentum equation as well.

For the thickness and tracer equations, the time-discrete vertical Lagrangian-remap algorithm takes on the form

$$h^* = h^{(n)} - \Delta t \nabla_{\sigma} \cdot (h^{(n)} \mathbf{u}^\dagger) \quad \text{thickness} \quad (33.79a)$$

$$h^* C^* = h^{(n)} C^{(n)} - \Delta t \left[\nabla_{\sigma} \cdot (h C \mathbf{u}^\dagger + h \mathbf{J}^h) + \delta_{\sigma}(z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}) \right] \quad \text{tracer} \quad (33.79b)$$

$$h^{(n+1)} = h^* - \Delta t \delta_{\sigma} w^{(\dot{\sigma})} \quad \text{regrid} \quad (33.79c)$$

$$h^{(n+1)} C^{(n+1)} = h^* C^* - \Delta t \delta_{\sigma} (w^{(\dot{\sigma})} C^*) \quad \text{remap tracer.} \quad (33.79d)$$

Equation (33.79a) provides an intermediate thickness, h^* , resulting from the vertical Lagrangian update realized by dropping $w^{(\dot{\sigma})}$. Similarly, equation (33.79b) updates the tracer concentration to an intermediate value, C^* , by vertically following the fluid particles so that the $w^{(\dot{\sigma})}$ contribution to the tracer equation (33.71) is dropped. Equation 33.79c is the regrid step, and this is the key step in the algorithm. In this step the new thicknesses are *prescribed* by the new grid. That is, the new thickness $h^{(n+1)}$ is prescribed rather than prognosed, and it is prescribed by the pre-defined target values for the vertical grid. We then use these prescribed values for $h^{(n+1)}$ to *diagnose* the dia-surface transport, $w^{(\dot{\sigma})}$, according to the regrid equation (33.79c). The diagnosed dia-surface transport is then used to remap the tracer concentration in equation (33.79d). For example, if the prescribed coordinate surfaces are geopotentials, then $w^{(\dot{\sigma})} = w$, in which case the remap step for the tracer is vertical advection. This example clearly exposes the general connection between remapping and advection.

In closing this example, note that substituting equation (33.79a) into equation (33.79c) leads to the time discretized form of the full thickness equation (33.70). Similarly, substituting equation (33.79b) into equation (33.79d) recovers the time discrete form of the tracer equation (33.71). This sanity check verifies that the operator split algorithm offers a consistent discretization of the thickness and tracer equations.

33.4.5 Comments

The vertical Lagrangian-remap algorithm follows [Bleck \(2002\)](#). It complements the horizontal Lagrangian-remap algorithm of [Dukowicz and Baumgardner \(2000\)](#) used for horizontal advective transport. Each approach shares elements with the three-dimensional Arbitrary Lagrangian-Eulerian (ALE) algorithm of [Hirt et al. \(1997\)](#). In addition to ocean applications, the vertical Lagrangian-remap algorithm is used for atmospheric modeling by [Lin \(2004\)](#). Indeed, the algorithm is becoming widespread in numerical ocean and atmospheric modeling.

33.5 Numerically diagnosing fluid particle trajectories

We introduced fluid particle trajectories in Section 16.3 as part of our discussion of fluid kinematics. In this section we revisit that discussion with a focus on the diagnostic calculation of trajectories in a numerical model, making use of the vertical Lagrangian-remap algorithm of Section 33.4. This discussion serves to outline a method to compute trajectories and to emphasize key aspects of the vertical Lagrangian-remap algorithm.

33.5.1 Basics of estimating fluid particle trajectories

The particle trajectory calculation requires us to time integrate the velocity following a fluid particle marked with the material coordinate \mathbf{a}

$$\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, t_0) + \int_{t_0}^t \mathbf{V}(\mathbf{a}, t') dt'. \quad (33.80)$$

In this equation, we made use of capital letters to denote the position and corresponding velocity of a fluid particle, whereas the lower case \mathbf{x} and \mathbf{v} represent the Eulerian field position and Eulerian velocity field, respectively. We most commonly wish to represent the velocity vector in terms of Cartesian unit vectors, in which case

$$\mathbf{X}(t) = \mathbf{X}(t_0) + \int_{t_0}^t [\hat{\mathbf{x}} U(t') + \hat{\mathbf{y}} V(t') + \hat{\mathbf{z}} W(t')] dt', \quad (33.81)$$

where we dropped the material coordinate \mathbf{a} for brevity. Representations for the trajectory making use of alternative basis vectors and basis one-forms are discussed in Section 11.6.

In a numerical fluid model, the velocity field is generally known only over grid cells with a fixed horizontal position. In contrast, particle trajectories can traverse any point within a grid cell. We thus need to interpolate the velocity vector to the point of the particle prior to time integrating the particle position in equation (33.80). There are a variety of methods used for this interpolation, with [van Sebille et al. \(2018\)](#) offering a review.

33.5.2 High wave number power in the vertical velocity

The vertical component to the velocity field, w , can exhibit high wave number features associated with gravity waves and flow near topography. If the numerical grid is not sufficiently resolved to represent these features, then the simulation can lose physical integrity by producing excessive power at the grid scale; i.e., it can become “noisy”. One means to understand the origin for the numerical grid noise is to consider an incompressible fluid whereby

$$\partial_z w = -\nabla_z \cdot \mathbf{u} \quad \text{incompressible.} \quad (33.82)$$

Hence, knowledge of the horizontal velocity, \mathbf{u} , allows us to diagnose w through vertical integration of the horizontal convergence. Now the convergence of the horizontal velocity is the small difference of the relatively large horizontal velocity, thus exposing the numerical calculation of the convergence to truncation errors. Noise is exacerbated near boundaries since the horizontal velocity rapidly changes there, thus contributing to a larger convergence and potentially larger amplitude noise. Noise is also exacerbated with strong gravity waves since gravity waves have a nonzero horizontal convergence (see Sections 44.2 and 44.3). Now w results from vertically integrating the horizontal convergence, with integration acting to smooth. Even so, it does not remove all noise, particularly that found near boundaries and within gravity wave fields. Furthermore, vertical integration acts to transfer noise found at one depth to subsequent depths.

Is there an alternative approach that is subject to less noise? Notably, when seeking alternatives it is important to ensure that the resulting trajectories do not allow the particles to leave the ocean domain. We here propose an alternative approach that offers a possible means to produce more accurate trajectories with less noise while remaining within the ocean domain.

33.5.3 Trajectories from the vertical Lagrangian-remap algorithm

van Sebille et al. (2018) summarize methods used to diagnose fluid particle trajectories making use of output from ocean models. All of the extant methods use the vertical velocity as per equation (33.81), even ocean models such as HYCOM that are based on the vertical Lagrangian-remap algorithm (e.g., see the HYCOM particle trajectory algorithm in Section A2.4 of *van Sebille et al. (2018)*).

The vertical Lagrangian-remap algorithm of Section 33.4 allows for a rather trivial and elegant computation of the vertical position of fluid particles. The reason is that it is only the first part of the vertical Lagrangian-remap algorithm that leads to particle motion. Furthermore, since this portion of the algorithm is vertically Lagrangian, with the σ -surface following fluid particles, then the vertical position of the particle is computed as part of the vertical Lagrangian-remap algorithm. We thus only need to diagnose the horizontal movement associated with the trajectory equation (33.81). Importantly, the remap part of the algorithm merely reorganizes the σ -coordinate surfaces without altering the fluid state. Hence, the spatial position of a fluid particle is unaffected by the remapping step.

Here is a summary of the trajectory method using the vertical Lagrangian-remap algorithm.

- VERTICAL LAGRANGIAN STEP: Horizontal displacements are determined by time integrating the horizontal velocity as per equation (33.81). Their vertical position is determined according to the vertical position of its σ -surface. The middle panel of Figure 33.3 provides a schematic for this step.
- VERTICAL REMAPPING STEP: The remap step returns the σ -surfaces to their target values. However, the fluid particles do not move in space. Instead, the σ -coordinate values associated with their positions change under remapping. The right panel of Figure 33.3 provides a schematic for this step.

33.5.4 Interpolation versus extrapolation

The proposed fluid particle trajectory calculation described in this section makes use of interpolation to the extent possible given the grid layout. For example, consider particle trajectories in a perfect fluid in which the fluid particle remains on an isopycnal surface, and assume that σ -coordinate surfaces are defined by isopycnals. The traditional trajectory approach uses the vertical velocity

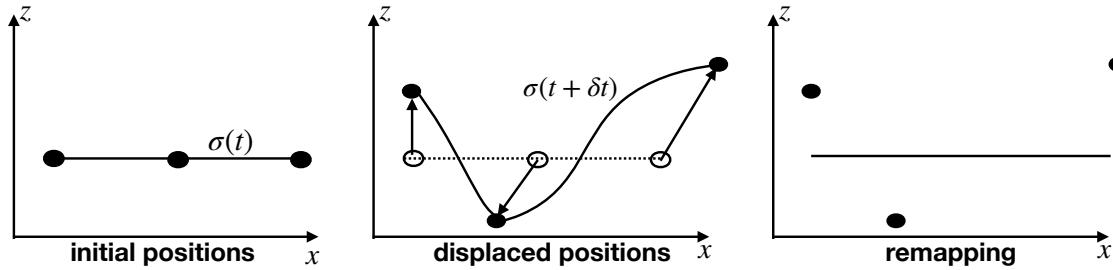


Figure 33.3: Illustrating particle trajectories with the vertical Lagrangian-remap algorithm. The left panel shows an initial horizontal line of fluid particles situated on a constant σ -surface, assumed here to be a geopotential. The middle panel shows the result of the Lagrangian step, with the displaced particles at time $t + \delta t$ and the σ -coordinate surface attached to the particles. Note that the particles move both horizontally and vertically during this step. The right panel shows the remapped σ surface having returned to its target geopotential. The fluid particle positions in space are unchanged during the remapping step. Rather, it is only their values of σ that changes under remapping.

diagnosed by vertically integrating the convergence of the horizontal velocity over a grid cell. This local grid-cell computed w is used to determine vertical motion of the particle. But if the particle stays on a σ -surface that moves beyond the grid cell in a single time step, the w -computed vertical motion is less accurate since it relies on extrapolation.² Interpolation to the vertical position of the displaced σ -surface offers an alternative means to determine the particle's vertical position. Interpolation is more reliable than extrapolation, and there are a variety of high order accurate interpolation schemes from which to choose (e.g., [White and Adcroft \(2008\)](#)).

We thus conjecture that the new trajectory approach provides a more accurate framework to track fluid particles. Furthermore, we suggest that the increased accuracy afforded by interpolation will reduce noise in the particle trajectories. These considerations are especially important for studies that make use of particles in high Rossby number flows (e.g., gravity waves, submesoscale flows) and their correspondingly strong vertical motion.

²Time steps here refer to the time steps of the sampled velocity field. If the trajectory calculation is performed online as the model is time stepping the ocean state, then sampling occurs each model time step. More commonly, the trajectory calculation is performed offline with the velocity sampled less frequently than online. In such cases there can be even larger excursions of σ -surfaces over the sampled time steps, thus leading to an even greater degree of extrapolation.

34

Space-time dependent gravity

We formulate the dynamical equations for a geophysical fluid in the presence of a space and time dependent gravitational acceleration. This formulation has application to the study of astronomical tides, thus motivating a brief discussion of the astronomical tidal forcing that follows the treatment given in Chapter 3 of [Pugh \(1987\)](#) and Section 5.15 of [Apel \(1987\)](#), with Chapter 2 of [Brown \(1999\)](#) and Section 17.4 of [Stewart \(2008\)](#) useful pedagogical supplements. Besides tides, a topic of increasing interest to climate science concerns the study of how the ocean sea level responds to changes in mass distributions associated with melting land ice. The nontrivial impact that melting land glaciers has on the earth's geoid ([Farrell and Clark \(1976\)](#) and [Mitrovica et al. \(2001\)](#)) further motivates developing the dynamical equations of a liquid ocean in the presence of a space-time dependent gravity.

READER'S GUIDE TO THIS CHAPTER

This chapter assumes an understanding of the equations of motion derived in Chapter 22 as well as the gravitational and centrifugal accelerations from Section 14.1. We dispense with tensor notation in this chapter, with subscripts used here as descriptive labels rather than tensor indices. No other chapter depends on the material in this chapter.

34.1	Gravitational potential	508
34.1.1	Simple geopotential	508
34.1.2	General geopotential	508
34.1.3	Comments	509
34.2	Momentum equation	509
34.3	Primitive equations	509
34.4	Depth independent perturbed geopotential	510
34.5	Forces contributing to ocean tides	510
34.5.1	Tidal acceleration in a spherically symmetric gravity field	510
34.5.2	Heuristics of tidal acceleration on the surface of a sphere	511
34.5.3	Gravitational potential for an idealized earth-moon system	514
34.5.4	Concerning realistic tides	517
34.5.5	Comments	518

34.1 Gravitational potential

In this section we summarize elements of the gravitational force, including the case with a non-constant gravitational acceleration such as occurs from astronomical tidal forcing and changes to the mass distribution of the planet.

34.1.1 Simple geopotential

As detailed in Section 14.1, the effective gravitational field incorporates the effects from the centrifugal force. The effective gravitational field is conservative, so that the gravitational acceleration of a fluid parcel can be represented as the gradient of a scalar (see Section 14.1.2),

$$\mathbf{g} = -\nabla \Phi, \quad (34.1)$$

with Φ the geopotential. In most applications of this book, the local vertical direction is denoted by

$$z = r - R, \quad (34.2)$$

with $z = 0$ the geopotential surface corresponding to a resting ocean. The geopotential in this case is given by

$$\Phi \approx \Phi_0 = g z, \quad (34.3)$$

with $g \approx 9.8 \text{ m s}^{-2}$ the typical value used for the acceleration due to gravity at the earth's surface.

34.1.2 General geopotential

Consider a generalized geopotential written in the form

$$\Phi = \Phi_0(r) + \Phi_1(r, \lambda, \phi, t), \quad (34.4)$$

where $\Phi_0(r)$ is the unperturbed geopotential given by equation (34.3), and Φ_1 incorporates perturbations to the geopotential associated with changes in land ice cover. Within the ocean fluid, the radial dependence of Φ_1 is generally quite weak, though it can be large for regions near the melting land ice. We thus maintain this dependence for purposes of generality, though it will be

dropped for certain specialized examples. The calculation of ocean tides arising from astronomical forcing is formulated with a space-time dependent geopotential as in equation (34.4), with the radial dependence of Φ_1 neglected (e.g., Section 9.8 in [Gill, 1982](#)). [Arbic et al. \(2004\)](#) provide a recent discussion of global tide modelling.

34.1.3 Comments

For the study of ocean tides, variations in Φ_1 arise from astronomical perturbations to the earth's gravity field. Nontrivial Φ_1 variations also arise from perturbations in terrestrial masses, such as the melting of land ice such as that occurring on Greenland or Antarctica due to global warming. In contrast to ocean tides, geoid perturbations associated with melting land ice are not periodic. Furthermore, as evidenced by Figure 1 in [Mitrovica et al. \(2001\)](#), the amplitude of geoid perturbations can be far greater than typical open ocean tide fluctuations. Such changes to the gravitational field can furthermore lead to perturbations that are a function of $\Phi_1(\lambda, \phi, r, t)$. Such perturbations modify the hydrostatic balance as seen by equation (34.6b).

34.2 Momentum equation

As detailed in Section 22.1.3, the inviscid momentum equation for a rotating fluid in a gravitational field is given by

$$\rho \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \rho \mathbf{v} = -\nabla p - \rho \nabla \Phi. \quad (34.5)$$

In writing the momentum equation in the form (34.5), we have chosen to retain an orientation afforded by the unperturbed geopotential surfaces, which correspond to surfaces of constant depth z . This approach reflects that commonly used to study ocean tides. In the presence of a perturbed geopotential Φ_1 , the "horizontal" directions defined by surfaces of constant z are no longer parallel to geopotential surfaces. We thus may interpret the sum $\nabla_z p + \rho \nabla_z \Phi$ as an orientation of the pressure gradient along surfaces of constant geopotential, where the geopotential is determined by $\Phi = \Phi_0 + \Phi_1$, rather than just the unperturbed geopotential Φ_0 .

34.3 Primitive equations

As detailed in Section 27.1, the primitive equations reduce the vertical momentum equation to its static inviscid form, which is the hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad (34.6a)$$

$$= -\rho(g + \partial_z \Phi_1). \quad (34.6b)$$

The hydrostatic balance is modified from its traditional form for cases where the perturbation geopotential Φ_1 exhibits nontrivial depth dependence. Correspondingly, the horizontal momentum equation (making the Traditional Approximation from Section 27.1) takes the form

$$\rho \frac{D\mathbf{u}}{Dt} + \hat{z} f \wedge \rho \mathbf{u} = -(\rho \nabla_z \Phi_1 + \nabla_z p) \quad (34.7)$$

where ∇_z is the horizontal gradient taken on surfaces of constant z . In their oceanic Boussinesq form (Chapter 28), the inviscid horizontal momentum equation becomes

$$\frac{D\mathbf{u}}{Dt} + \hat{z} f \wedge \mathbf{u} = -(1/\rho_0)(\rho_0 \nabla_z \Phi_1 + \nabla_z p) \quad (34.8)$$

where ρ_0 is the constant reference density for a Boussinesq fluid. The Boussinesq form makes the addition of a perturbed geopotential quite straightforward, in which it is gradients in $\rho_0 \Phi_1 + p$ that take the place of gradients in pressure p .

34.4 Depth independent perturbed geopotential

A particularly simple form of Φ_1 occurs when it is depth independent,

$$\Phi_1 = \Phi_1(\lambda, \phi, t), \quad (34.9)$$

in which case the hydrostatic balance (34.6b) returns to its traditional form $\partial_z p = -\rho g$. This perturbed geopotential is generally sufficient for the study of ocean tides. In this case it is convenient to write the geopotential as

$$\Phi_1 = -g h, \quad (34.10)$$

with $h = h(\lambda, \phi, t)$ the perturbed geopotential height field. The full geopotential is thus written

$$\Phi = g(z - h), \quad (34.11)$$

with this form revealing that the zero of the geopotential is now set by $z = h$ rather than $z = 0$. In the study of ocean tides, h is referred to as the *equilibrium tide*. In geodesy, h is referred to as the *static equilibrium sea level*.

Since the perturbed geopotential is depth independent, it only affects the depth integrated horizontal momentum, and it does so through the term

$$-\int_{-H}^{\eta} \nabla_z \Phi_1 dz = g \int_{-H}^{\eta} \nabla_z h dz = g(H + \eta) \nabla_z h. \quad (34.12)$$

Hence, modifications to the geopotential as embodied by the perturbed geopotential height field, $h = h(\lambda, \phi, t)$, are isolated to their impacts on the horizontal pressure gradients acting on the depth integrated horizontal momentum.

34.5 Forces contributing to ocean tides

We here describe the rudiments of forces that contribute to ocean tides as well as solid-earth tides. For simplicity we focus just on the earth-moon system, though note that the sun also plays an analogous role for observed tidal motion.

34.5.1 Tidal acceleration in a spherically symmetric gravity field

Before considering the earth-moon system, we introduce the notion of *tidal acceleration*, which arises on a finite sized body placed within a non-uniform gravitational field. Figure 34.1 depicts this situation where the finite sized body is a narrow rod whose axis points towards the center of a spherically symmetric massive body. One end of the rod experiences a different gravitational acceleration than the other since the gravitational field falls off as the inverse squared distance from the center of the sphere. It is this differential gravitational acceleration that we refer to as the tidal acceleration. As we will see, its key property is that the tidal acceleration falls off as the inverse cube of the distance rather than the more familiar inverse square.

To develop a mathematical expression for the tidal acceleration, focus on the spherically symmetric gravitational field in which the gravitational acceleration at a point is given by (Section 14.1.1)

$$\mathbf{g} = -\frac{GM}{r^2} \hat{\mathbf{r}}, \quad (34.13)$$

where r is the distance from the sphere's center, G is Newton's gravitational constant, M is the mass of the sphere, and $\hat{\mathbf{r}}$ is the radial unit vector. The minus sign indicates that the gravitational acceleration points toward the center of the sphere. For the rod in Figure 34.1, the difference between the gravitational acceleration acting at a point nearest to the sphere (point B) and a point furthest from the sphere (point A) is given by

$$\mathbf{g}(r_B) - \mathbf{g}(r_A) = \mathbf{g}(r_0 - L/2) - \mathbf{g}(r_0 + L/2), \quad (34.14)$$

where r_0 is the distance from the sphere's center to the center of the rod. Assuming the rod is not long, we can expand this difference in a Taylor series about the rod center at r_0 , thus leading to an expression for the tidal acceleration

$$\mathbf{g}(r_B) - \mathbf{g}(r_A) \approx -L \frac{\partial \mathbf{g}}{\partial r} = -2L \frac{GM}{r_0^3} \hat{\mathbf{r}} = (2L/r_0) \mathbf{g}(r_0). \quad (34.15)$$

The key point to conclude from this example is that the tidal acceleration is proportional to the inverse cube of the distance to the center of the sphere. We see this property again when considering in Section 34.5.3 the gravitational acceleration generated from a remote body (e.g., the moon) acting on the surface of a sphere (e.g., the earth).

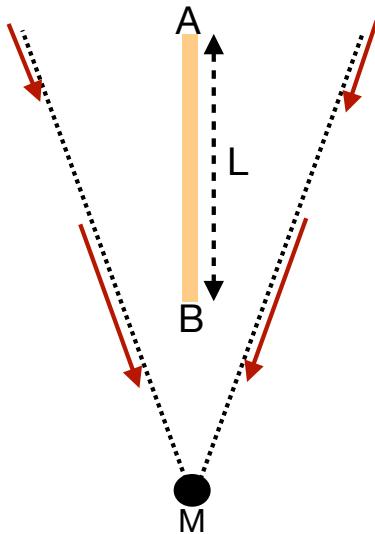


Figure 34.1: Tidal acceleration is the acceleration that acts on a finite sized object placed in a non-uniform gravitational field. The finite object is here depicted as a narrow rod of length L placed in the gravity field of a spherically symmetric body of mass M . That portion of the rod closer to the gravitating sphere (end B) experiences a stronger gravitational acceleration than the end that is further away (end A). The gradient in the gravitational acceleration constitutes the tidal acceleration acting on the rod.

34.5.2 Heuristics of tidal acceleration on the surface of a sphere

We now consider the tidal acceleration acting on the surface of a smooth massive sphere due to a spherically symmetric gravitational field generated by a neighboring massive body. Figure

34.2 depicts this system, which we consider an idealized earth-moon system where each body is assumed homogeneous and spherical. Given that they gravitationally attract one another, it is not astronomically possible for the two bodies to remain spatially fixed. Instead, they orbit around their common center of mass while conserving their angular momentum.

A central question of tidal studies is why there are generally two ocean tides per day (semi-diurnal tides) rather than just one (diurnal tides). We here offer two complementary arguments. The first is based on extending the tidal acceleration discussion of Section 34.5.1, whereas the second follows the more traditional account by considering a balance between gravitational and centrifugal accelerations.

General ideas

Every point on the surface of the earth is attracted to the earth's center by the earth's gravitational field. For a spherical earth, this attractive force is purely radial, so that it cannot lead to lateral motion on the surface of the perfect sphere. We thus conclude that the radial gravitational field is not the cause of tidal motion. Instead, tidal motion arises from a non-radial gravitational field.

The earth-moon gravitational field accelerates the earth and moon toward one another along the axis connecting their centers. Additionally, the spatial dependence of the moon's gravitational field over the earth leads to lateral forces along the earth's surface, thus providing the ingredient for ocean tidal motion. To capture the essence of this force, we examine how the moon's gravitational field acts on a point on the earth relative to its action at the center of the earth.

Sample tidal accelerations on the sphere

Again, we are tasked with computing the tidal acceleration from the moon's gravitational field for selected points on the earth, computing these accelerations relative to the earth center. As for the rod in Figure 34.1, the tidal acceleration at point *B* relative to the center of the earth is given by

$$\mathbf{g}(r_B) - \mathbf{g}(R_{\text{em}}) = (2R_e/R_{\text{em}}) \mathbf{g}(R_{\text{em}}). \quad (34.16)$$

This acceleration points towards the moon. In contrast, the tidal acceleration at point *A* relative to the center of the earth is given by

$$\mathbf{g}(r_A) - \mathbf{g}(R_{\text{em}}) = -(2R_e/R_{\text{em}}) \mathbf{g}(R_{\text{em}}), \quad (34.17)$$

which is of equal magnitude but points away from the moon.

The tidal accelerations at points *A* and *B* act radially away from the earth's center. Hence, as noted above, these radial forces do not directly lead to tidal motion at those points. However, through symmetry of the configuration, points between *A* and *B* have a tidal acceleration from the moon's gravitational field with a nonzero lateral component. These lateral forces lead to the accumulation of water at points *A* and *B*. We can compute the gravitational acceleration at intermediate points. However, the trigonometry is somewhat complex and we prefer to compute the forces in Section 34.5.3 through use of the gravitational potential. For the current discussion we appeal to symmetry to conclude that the lateral tidal accelerations act to pile up water at both points *A* and *B* as depicted in the second panel of Figure 34.2. This argument, though heuristic, provides the means to understand how a water covered spherical planet has two bulges, rather than one, due to spatial gradients in the moon's gravitational field. We confirm this argument in Section 34.5.3 by explicitly computing the gravitational potential for this idealized earth-moon system and then taking the gradient to compute the gravitational acceleration (see Figure 34.4).

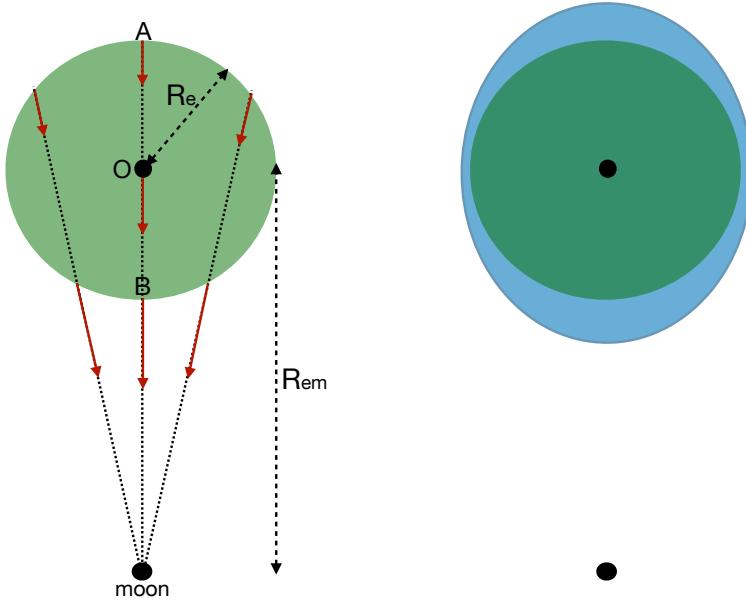


Figure 34.2: Illustrating the tidal force on the surface of a sphere. The sphere is an ideal depiction of the earth and the smaller massive object is the moon. The distance between the center of the earth and moon is R_{em} , and the radius of the earth is R_e . The left panel shows representative moon-generated gravitational field lines. Two points along these field lines on the surface of the earth represent the two ends of an imaginary rod as depicted in Figure 34.1. The tidal acceleration acting at point B, relative to the earth's center, points toward the moon. In contrast, the tidal acceleration at point A, relative to the earth's center, points in the opposite direction. Points on the earth surface between A and B have tidal accelerations with a non-zero component directed along the surface of the earth. Symmetry of the configuration allows us to conclude that a layer of water on the surface of the sphere will accumulate to produce two bulges as shown in the right panel. It is the lateral component of the gravitational acceleration that causes the water to accumulate to produce tidal bulges at points A and B. In contrast, the radial component to the moon's gravitational field has no contribution to the tides. Note that as shown in Section 34.5.3, the bulge shown in the right panel is greatly exaggerated.

Including orbital motion

Thus far we have ignored the orbital motion of the earth-moon system around their common center of mass. As we will see, the above arguments lead to the same results as when orbital motion is considered.

In the absence of dissipation, as assumed here, the earth-moon distance remains constant due to their angular momentum conserving orbital motion. From a force-balance perspective, the two spherical bodies remain in a fixed orbit since the gravitational acceleration acting at their centers is balanced by their respective centrifugal accelerations, where the centrifugal acceleration is computed relative to the center of mass of the two-body system. The gravitational acceleration from the moon, acting at the center of the earth, is given by the *free fall* value $\mathbf{g}(R_{\text{em}})$, which has magnitude GM_m/R_{em}^2 and is directed along the axis connecting the earth and moon centers.

Furthermore, when a body exhibits orbital motion, each point on the body exhibits the same orbital motion and has the same linear velocity. Consequently, each point on the earth possess the same centrifugal acceleration

$$\mathbf{a}_{\text{orbital centrifugal}} = -\mathbf{g}(R_{\text{em}}). \quad (34.18)$$

This property of orbital motion is distinct from the spinning motion of a planet rotating about its axis, whereby points further from the rotational axis have larger centrifugal acceleration (see Section 14.1). To help understand orbital motion, move your hand in a circle while maintaining

the arm in a single direction so that the hand exhibits an orbital motion rather than a spinning motion. Notice that all parts of the hand move with the same linear velocity and exhibit the same orbital motion. Hence, each point on the hand has the same centrifugal acceleration.

We can now ask about the acceleration felt by a point on the surface of the earth. The acceleration giving rise to tidal motions is the sum of the gravitational acceleration from the moon plus the centrifugal acceleration due to orbital motion. However, this calculation is identical to that considered previously, which led, for example, to the tidal accelerations for points B and A as given by equations (34.16) and (34.17). We are thus led to the same result as before.

34.5.3 Gravitational potential for an idealized earth-moon system

We now perform a more thorough calculation of the gravitational acceleration by computing the gradient of the gravitational potential. First recall the discussion of Newton's gravitational law in Section 14.1.1, whereby the gravitational potential for a point at distance r from the center of a spherical earth is given by

$$\Phi_e(r) = -\frac{GM_e}{r}, \quad (34.19)$$

where M_e is the mass of the earth. The corresponding radial gravitational acceleration is given by

$$\mathbf{g}_e = -\nabla\Phi_e = -\frac{GM_e \hat{\mathbf{r}}}{r^2}. \quad (34.20)$$

The same considerations hold for the moon's gravitational potential. Hence, referring to Figure 34.3, the moon's gravitational potential evaluated at a distance L from the moon's center is given by

$$\Phi_m(L) = -\frac{GM_m}{L}. \quad (34.21)$$

Trigonometry leads to the law of cosines relation

$$L^2 = (R_{em} - r \cos \psi)^2 + (r \sin \psi)^2 = R_{em}^2 + r^2 - 2r R_{em} \cos \psi, \quad (34.22)$$

where again r is the distance to the earth's center and ψ is the polar angle relative to the $\hat{\mathbf{x}}$ axis pointing between the earth and moon centers (see Figure 34.3).

Identifying the leading order contributions

Assuming the test point in Figure 34.3 is closer to the earth than to the moon, we can perform a Taylor series expansion in the small parameter r/R_{em} to render

$$\Phi_m(L) = -\frac{GM_m}{L} = -\frac{GM_m}{R_{em}} \left[1 + \frac{r \cos \psi}{R_{em}} + \frac{r^2}{2R_{em}^2} (3 \cos^2 \psi - 1) + \mathcal{O}(r/R_{em})^3 \right]. \quad (34.23)$$

We thus identify the leading three terms to the geopotential

$$\Phi_m^{(0)} = -\frac{GM_m}{R_{em}} \quad (34.24)$$

$$\Phi_m^{(1)} = -\frac{GM_m}{R_{em}^2} r \cos \psi \quad (34.25)$$

$$\Phi_m^{(2)} = -\frac{GM_m}{2R_{em}^3} r^2 (3 \cos^2 \psi - 1). \quad (34.26)$$

Assuming the distance between the earth and moon remains fixed, the zeroth order term $\Phi_m^{(0)}$ is a spatial constant and thus leads to no gravitational acceleration. We now examine the gravitational accelerations from the other two terms.

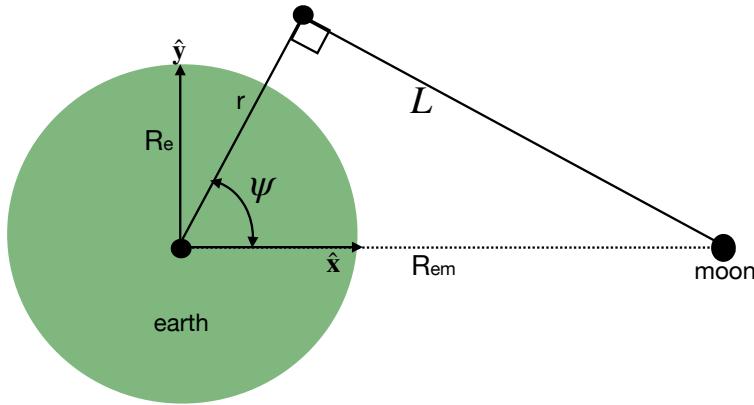


Figure 34.3: Geometry of an idealized earth-moon system. The center of the earth is a distance R_{em}^2 from the center of the moon; the moon has a mass M_m ; and the earth has a radius R_e . An arbitrary test point is shown a distance L from the center of the moon, r from the center of the earth, and with a polar angle ψ relative to the \hat{x} axis, where the \hat{x} axis points from the earth center to the moon center. Relative to the earth's center, the test point has Cartesian coordinates $(x, y) = r(\cos \psi, \sin \psi)$. See Section 10.3 for details on relating polar and Cartesian coordinates.

Acceleration maintaining the orbiting earth-moon system

For the first order term, $\Phi_m^{(1)}$, we introduce the Cartesian coordinate as in Figure 34.3 to write

$$\Phi_m^{(1)} = -\frac{GM_m x}{R_{\text{em}}^2}, \quad (34.27)$$

where $x = r \cos \psi$ is the distance along \hat{x} . Hence, the gradient of $\Phi_m^{(1)}$ leads to the gravitational acceleration

$$\mathbf{g}_m^{(1)} = -\nabla \Phi_m^{(1)} = \hat{x} \frac{GM_m}{R_{\text{em}}^2}. \quad (34.28)$$

This gravitational acceleration has a constant magnitude at every point in space and it everywhere points in a direction parallel to the earth-moon axis. Furthermore, the magnitude of $\mathbf{g}_m^{(1)}$ equals to that of the moon's gravitational acceleration, \mathbf{g}_m , when evaluated at the earth's center. As seen in Section 34.5.2, the acceleration $\mathbf{g}_m^{(1)}$ maintains the earth in orbit about the center of mass for the earth-moon system; i.e., this is the free fall acceleration towards the moon. Notably, at the earth's surface, the magnitude of $\mathbf{g}_m^{(1)}$ is tiny relative to the gravitational acceleration from the earth itself, with their ratios given by

$$\frac{M_m/R_{\text{em}}^2}{M_e/R_e^2} \approx 3.4 \times 10^{-6}, \quad (34.29)$$

where we set

$$M_e = 5.97 \times 10^{24} \text{ kg} \quad M_m = 7.35 \times 10^{22} \text{ kg} = (1/81.2) M_e \quad (34.30a)$$

$$R_e = 6378 \text{ km} \quad R_{\text{em}} = 384 \times 10^3 \text{ km} = 60.2 R_e. \quad (34.30b)$$

Tide producing geopotential

The main tide producing acceleration results from $\Phi_m^{(2)}$. Introducing the second Cartesian coordinate, $y = r \sin \psi$, leads to

$$\Phi_m^{(2)} = -\frac{GM_m}{2R_{\text{em}}^3} r^2 (3 \cos^2 \psi - 1) = -\frac{GM_m}{2R_{\text{em}}^3} (2x^2 - y^2). \quad (34.31)$$

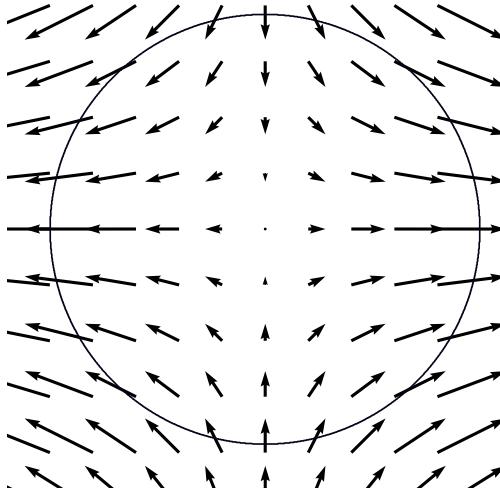


Figure 34.4: The tide producing gravitational acceleration $\mathbf{g}_m^{(2)}$ given by equation (34.34). The moon is assumed to be positioned in the equatorial plane of the earth.

The corresponding perturbed geopotential height field (see equation (34.11)) is given by

$$h = -\frac{\Phi_m^{(2)}}{g} = \frac{R_e^2}{2 R_{em}^3} \frac{M_m}{M_e} r^2 (3 \cos^2 \psi - 1). \quad (34.32)$$

Letting $r = R_e$ renders

$$h = \frac{R_e^4}{2 R_{em}^3} \frac{M_m}{M_e} (3 \cos^2 \psi - 1) \approx 2.8 \times 10^{-8} R_e (3 \cos^2 \psi - 1). \quad (34.33)$$

Plugging in numbers for the earth-moon system suggests that the maximum perturbation to the geopotential height arising from the moon's gravity field is roughly 36 cm. Correspondingly, the bulge shown in Figure 34.2 is greatly exaggerated. Note that ocean tidal amplitudes can get much larger (order meters) than this "equilibrium tide" amplitude due to resonances from ocean geometry, with the Bay of Fundy in Nova Scotia a particularly striking example.

Tide producing acceleration

The gravitational acceleration arising from the tidal potential is determined by the gradient of the tidal geopotential

$$\mathbf{g}_m^{(2)} = -\nabla \Phi_m^{(2)} = \frac{GM_m}{R_{em}^3} (2x \hat{\mathbf{x}} - y \hat{\mathbf{y}}). \quad (34.34)$$

We illustrate the vector field $\mathbf{g}_m^{(2)}$ in Figure 34.4. Note how the accelerations lead to two bulges on opposite sides of the planet. We can write this acceleration using polar coordinates by introducing the polar unit vectors $\hat{\mathbf{r}}$ and $\hat{\psi}$ according to Section 10.3.2

$$\hat{\mathbf{r}} = \hat{\mathbf{x}} \cos \psi + \hat{\mathbf{y}} \sin \psi \quad (34.35a)$$

$$\hat{\psi} = -\hat{\mathbf{x}} \sin \psi + \hat{\mathbf{y}} \cos \psi \quad (34.35b)$$

thus rendering

$$\mathbf{g}_m^{(2)} = \frac{GM_m R_e}{R_{em}^3} \left[\hat{\mathbf{r}} (3 \cos^2 \psi - 1) - (3/2) \hat{\psi} \sin 2\psi \right], \quad (34.36)$$

where we evaluated the acceleration at the earth surface so that $r = R_e$. Evaluating the acceleration at $\psi = 0, \pi$ verifies the heuristic calculation performed in Section 34.5.2 for points on the earth surface nearest and furthest from the moon. We can further gauge the magnitude of the tidal acceleration by introducing the acceleration due to the earth's gravity field

$$\mathbf{g}_m^{(2)} = g_e \frac{M_m}{M_e} \frac{R_e^3}{R_{em}^3} \left[\hat{\mathbf{r}} (3 \cos^2 \psi - 1) - (3/2) \hat{\boldsymbol{\psi}} \sin 2\psi \right], \quad (34.37)$$

where $g_e = G M_e / R_e^2$ is the acceleration at the earth's surface from the earth's gravity field. The dimensional prefactor has magnitude $\approx 5.6 \times 10^{-8} g_e$, so that the tidal acceleration is tiny relative to that from the earth's gravity field. It is for this reason that the radial component of the tidal acceleration is largely irrelevant since it is dominated by the far larger radial component of the earth's gravity field. However, the angular component of the tidal acceleration, although small relative to the earth's radial gravitational acceleration, is able to move water along the surface of the planet as indicated by Figure 34.4, thus leading to tidal motion.

34.5.4 Concerning realistic tides

Our discussion of tides has been rather terse, aiming to identify key aspects of the tidal accelerations but giving little attention to details that impact real ocean tides. Here are a few points that must be considered for these purposes.

- As the earth spins under the tidal bulges, there are two high and two low tides per day. Additional orbital motion of the moon adds roughly 50 minutes per day to the diurnal (daily) tide and 25 minutes to the semi-diurnal (twice daily).
- The moon orbits the earth at a latitude of roughly $28.5^\circ N$ rather than within the equatorial plane, so that the tidal bulges are offset from the equator. As the earth spins under the bulges, one of the high tides is generally larger than the other due to the offset. This offset in turn introduces a diurnal component to the tides in addition to the semi-diurnal.
- The sun contributes to tides in a manner similar to the moon. The sun is more massive than the moon, yet it is further away, so that the ratio of the magnitudes for the tidal producing accelerations is given by

$$\frac{\text{moon tidal acceleration}}{\text{sun tidal acceleration}} = \frac{M_m/R_{em}^3}{M_s/R_{es}^3} \approx 2.2 \quad (34.38)$$

where we set

$$M_s = 1.99 \times 10^{30} \text{ kg} \quad R_{es} = 23460 R_e. \quad (34.39)$$

Hence, the moon has an impact on tides that is somewhat more than double that of the sun.

- The gravitational acceleration that leads to the tidal bulge moves around the mid-latitudes at roughly 330 m s^{-1} , which is faster than the $\approx 200 \text{ m s}^{-1}$ wave speed for shallow water gravity waves. Hence, the ocean tidal motion is never equilibrated to the *equilibrium tides* defined by the tidal acceleration. In contrast, solid-earth waves are much faster and so the solid-earth tidal motions are mostly equilibrated with the equilibrium tidal acceleration. Solid-earth tides have an amplitude on the order of 10 cm with wavelengths spanning the planet. Hence, an accurate treatment of ocean tides must take into account the solid-earth tides.

- The movement of ocean mass modifies the earth's gravity field, and this modification is referred to as *self-attraction*. Additionally, movement of the ocean mass alters the *loading* felt by the solid-earth and thus causes the crust to compress and expand. These two terms are referred to as the *self attraction and loading* (SAL) terms.
- Geometry of the ocean plays a leading role in determining tides at a particular location. Since we have incomplete information about that geometry, the best predictions for tides are generally based on the analysis of past tides, with that information used to fit sinusoidal waves to the measured time series for use in projecting forward in time.

34.5.5 Comments

A key feature of the tidal producing forces is that it is the lateral (along-earth) component of the moon's tidal gravitational force that produces the earth's tides. These lateral forces cause water to accumulate at the point nearest to and furthest from the moon (points *A* and *B* in Figure 34.2), thus producing the characteristic double-bulge pattern. Notably, many common literature presentations make it appear that it is the radial (i.e., pointing to the earth's center) component of the moon's gravitational force, and its gradient across the earth, that leads to the earth's tidal bulges. But as discussed in Section 34.5.2, radial gravitational forces cannot lead to tidal motions; what is needed is a force that leads to lateral motion. These key notions are nicely emphasized in [this Space Time video](#).

35

Surface tension

Surface tension is present on surfaces that separate two immiscible liquids or between a liquid and gas. It has many consequences familiar from nature, such as allowing certain insects to walk on water even though their body density is greater than water, and for the predominantly spherical shape of rain drops. A molecular dynamics understanding of surface tension involves tools from physical chemistry that are well outside of our scope. Instead, we develop some heuristics sufficient to determine when one needs to be concerned with surface tension in the study of geophysical fluid mechanics.

Surface tension is generally negligible for length scales larger than a few centimeters. It is for this reason that surface tension is commonly absent from books on geophysical fluid dynamics, where most of the focus concerns much larger length scales. Nevertheless, the effects are important if studying physical processes associated with air-sea interactions, such as tracer, heat, and momentum exchange through bubbles, droplets, and capillary-gravity waves.

READER'S GUIDE TO THIS CHAPTER

This chapter makes use of the notions of stress detailed in Chapter 26, as well as curvature described in Chapter 6. We make use of Cartesian tensors as discussed in Chapters 3 and 4. Results can be readily generalized for arbitrary coordinates through the rules of general covariance detailed in Chapter 7. The material in this chapter has no direct impact on other chapters.

35.1 A container of water	520
35.2 Force balance on an air-water interface	520
35.3 Young-Laplace formula	522
35.4 Some oceanographic examples	524
35.4.1 Soluble gas bubbles inside water	524
35.4.2 Length scale for capillary waves	524
35.5 Further study	525

35.1 A container of water

Atmospheric pressure at the earth's surface is roughly $p_{\text{atm}} = 10^5 \text{ N m}^{-2}$. As we saw in Section 26.6, pressure acts normal to a surface regardless the surface orientation. So fill a container of water whose weight per horizontal area is less than the atmospheric pressure, $\rho g h < p_{\text{atm}}$ and turn the container upside-down as in Figure 35.1. Does the water spill from the container? Common experience with drinking glasses indicate that water will spill. But what about containers with a very small cross-sectional area such as the pipettes used in chemistry laboratories? Pipettes, or more generally capillary tubes, hold the liquid regardless the orientation. They do so since their cross-sectional area is small enough to allow forces from surface tension to overcome gravitational instabilities acting at the liquid-gas interface.

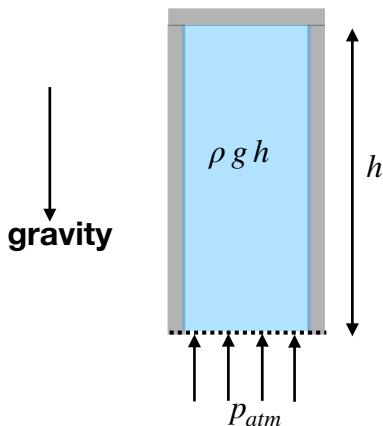


Figure 35.1: A container of water with density ρ and height h is placed upside-down. Atmospheric pressure, p_{atm} , will support water with thickness $h < p_{\text{atm}}/(\rho g) \approx 10 \text{ m}$ if the cross-sectional area of the container is small enough to allow for surface tension to overcome the gravitationally unstable waves that otherwise allow water to spill from the container. The liquid-gas interface supports both gravity waves (as in Section 44.2) and capillary waves. If the wavelength is small enough then surface tension suppresses the growth of unstable gravity waves so that the liquid remains within the “capillary tube”. However, for longer waves allowed by increasing the cross-sectional area, then any fluctuation will allow the gravitational instability to overcome surface tension, thus breaking the interface and releasing water onto the floor.

35.2 Force balance on an air-water interface

Consider two immiscible fluids with distinct densities. Air and water provide one example of special importance to understanding physics at the ocean-atmosphere boundary. Another example

concerns two immiscible layers of water within the ocean or two layers of air within the atmosphere. For molecules well within either of the fluid regions, the intermolecular forces are statistically isotropic. In contrast, intermolecular forces are not isotropic for molecules within a mean free path distance from the interface.¹ Attractive (cohesive) intermolecular (van der Waals) forces dominate within a liquid whereas gas molecules generally feel more repulsive forces. Hence, a liquid molecule within the liquid-gas interface preferentially experiences an attractive force towards the liquid side of the interface, as depicted in Figure 35.2. Surface tension arises from the cohesive force per area acting between molecules in a direction that parallels the interface, with surface tension acting to resist perturbations to the interface shape.

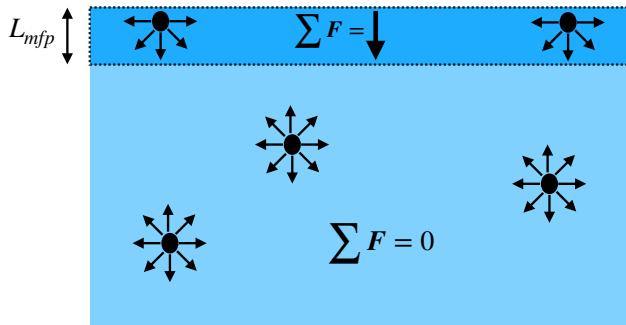


Figure 35.2: Surface tension at a liquid-gas interface arises from the anisotropic cohesive forces acting on liquid molecules within a mean-free-path distance, L_{mfp} , from the interface, which contrasts to the isotropic cohesive forces acting away from the interface. The net intermolecular force vanishes for interior molecules, whereas the net force acts inward on molecules at the interface. Surface tension refers to the cohesive force per area acting between molecules in a direction that is parallel to the interface.

Anisotropic attractive intermolecular forces cause the interface between the two fluids to behave as a stretched membrane that experiences a tensile force resisting any stretching of the interface. The magnitude of the tensile force per unit length is the *surface tension*, γ (units N m^{-1}). The surface tension is a property of the two fluids, including their temperature, as well as any impurities that might be included on the interface; e.g., oil on the surface of water effects properties of the capillary waves found on the air-sea interface. In the following we focus on the liquid-gas example to be specific and to expose issues that arise in studies of the air-sea interface. For a liquid-gas interface surrounding a liquid drop, the tensile force acts to curve the interface towards the liquid into a spherical shape.

The tensile force along a line segment is directed normal to the line and tangent to the interface

$$\mathbf{f}_{\text{interface}} = -\gamma \hat{\mathbf{n}} \wedge \delta \mathbf{x}, \quad (35.1)$$

where $\hat{\mathbf{n}}$ is a normal vector pointing towards the center of the curved interface, and $\delta \mathbf{x}$ is a line element oriented so that the normal $\hat{\mathbf{n}}$ points to the left facing in the direction of the line increment. Figure 35.3 depicts the surface tensile forces acting on the surface of a spherical bubble of water. Note that it is sometimes useful to consider the product γdS as the work (units of $\text{N m}^{-1} = \text{Joule}$) required to create an area, dS , on the interface. We make use of this energetic perspective in Section 35.3.

To develop an expression for the pressure jump across the liquid-gas interface, consider a spherical droplet of radius R shown in Figure 35.3 and focus on the circular cross-section cut through

¹ As discussed in Section 2.1, the mean free path is a statistical measure of the distance a molecule moves before hitting another molecule.

the center of the sphere. The net tensile force acting on the circumference of the circle is

$$\mathbf{F}_{\text{circle}} = \oint_{\text{circle}} \mathbf{f}_{\text{interface}} = - \oint_{\text{circle}} \gamma \hat{\mathbf{n}} \wedge \delta \mathbf{x} = -2\pi R \gamma \hat{\mathbf{z}}. \quad (35.2)$$

Equilibrium of the spherical droplet is realized by a pressure jump across the circular cross-sectional area

$$\pi R^2 (p_{\text{in}} - p_{\text{out}}) = 2\pi R \gamma \implies (p_{\text{in}} - p_{\text{out}}) = 2\gamma/R. \quad (35.3)$$

Hence, the pressure jump is determined by the surface tension (a property of the two fluids) and the curvature of the sphere, R , which is also the radius of curvature for the sphere. Pressure is higher inside of the sphere, with this pressure required to balance the pressure outside the sphere plus the surface tension. Notably, equilibrium for smaller bubbles requires a larger pressure difference than for larger bubbles.

The pressure jump is known as the *capillary pressure*. It arises from surface tension and curvature of the interface. The relation (35.3) is a special case of the Young-Laplace formula, specialized here to a sphere.

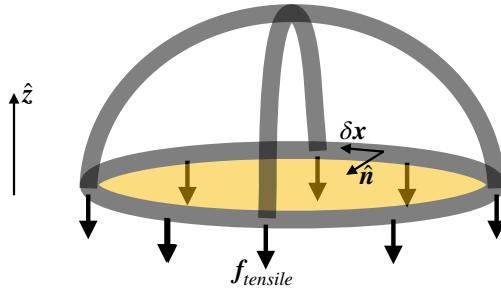


Figure 35.3: Surface tension on a spherical water droplet, with water on the inside of the sphere and air on the outside. The tensile forces act parallel to the spherical interface between the water and air. When cutting a circular cross-section as shown here, the surface tensile force acts downward. In equilibrium, the net tensile forces acting downward along the circumference of the hemisphere ($2\pi R \gamma$) are balanced by a pressure jump across the droplet, with the interior pressure larger than the exterior pressure. Focusing on the circular cross-section, this area remains static so long as $2\pi R \gamma = \pi R^2 (p_{\text{in}} - p_{\text{out}})$, leading to a pressure jump across the droplet interface $p_{\text{in}} - p_{\text{out}} = 2\gamma/R$.

35.3 Young-Laplace formula

We have added insight into the physics of surface tension by considering the energetics required to enable a virtual displacement of a surface through a pressure field along with the work required to change the area of the surface. The resulting equation for the pressure jump across the surface is referred to as the *Young-Laplace formula*, which expresses the pressure jump in terms of the surface tension and the principle radii of curvature for the surface.

Consider a horizontal surface depicted in Figure 35.4 that represents the interface separating fluid-A from fluid-B, with $\hat{\mathbf{n}}$ a unit normal vector oriented from fluid-A to fluid-B. Now consider a virtual displacement of each point along the interface by an infinitesimal distance, δh , with $\hat{\mathbf{n}} \delta h$ connecting points on the initial position of the interface to the displaced position, where $\delta h > 0$ if the displacement is directed towards fluid-B and $\delta h < 0$ if directed towards fluid-A. The (signed) volume swept out by an infinitesimal area dA is given by $\delta h dA$, with this volume realized by applying the pressure work to the surface

$$W_{\text{volume}} = (p_B - p_A) \delta h dA. \quad (35.4)$$

For example, if $p_B > p_A$ and the displacement is into fluid-B ($\delta h > 0$), then $W_{\text{volume}} > 0$, thus indicating the need to apply positive work to the surface to move it into the fluid region with higher pressure. Conversely, the required pressure work is negative if displacing the interface into a region with lower pressure.

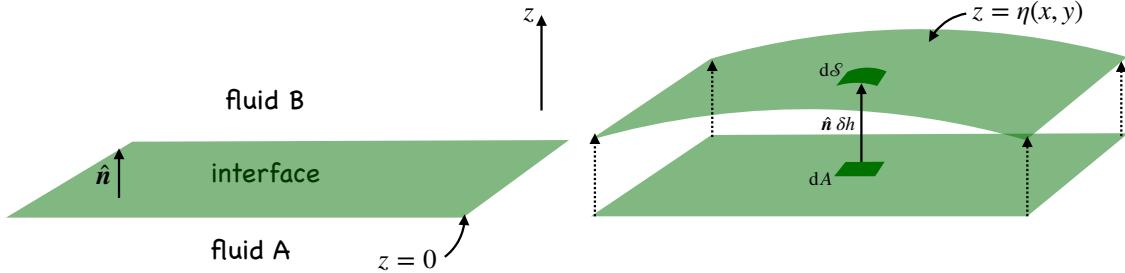


Figure 35.4: Left panel: initial position of an interface separating two fluid regions, fluid-A and fluid-B. Right panel: infinitesimal displacement of the interface sweeps out a volume in space. To determine the volume, extend a unit normal vector, \hat{n} , from the initial interface position and pointing towards fluid-B. Let δh be the distance along that normal to the new position, with $\delta h > 0$ if the displacement moves towards fluid-B and $\delta h < 0$ for displacements pointing to fluid-A. We assume that displacements at each interface point can move independently of adjacent points, so that the interface area generally changes.

In the presence of surface tension, work must overcome the surface area energy in order to change to the interface area

$$W_{\text{area}} = \gamma \delta A, \quad (35.5)$$

where δA is the change in area of an infinitesimal element on the interface

$$\delta A = dS - dA \quad (35.6a)$$

$$= dA \left[\sqrt{1 + (\nabla \delta h)^2} - 1 \right] \quad (35.6b)$$

$$\approx dA (\nabla \delta h)^2 / 2. \quad (35.6c)$$

To reach this result we made use of equation (6.29) that relates the area of an infinitesimal element on a curved surface to the area of its horizontal projection (see Section 6.3.1). We next make use of the surface curvature detailed in Section 6.3.2, where equation (6.33) shows that the vertical displacement is given, for small displacements, by

$$\delta h \approx \frac{1}{2} R_1^{-1} (\mathbf{x} \cdot \mathbf{e}_1)^2 + \frac{1}{2} R_2^{-1} (\mathbf{x} \cdot \mathbf{e}_2)^2. \quad (35.7)$$

R_1^{-1}, R_2^{-1} are the eigenvalues and $\mathbf{e}_1, \mathbf{e}_2$ are the corresponding eigenvectors of the matrix of second partial derivatives of $\delta h(x, y)$, whereas the inverse eigenvalues, R_1, R_2 , are the radii of curvature of the displaced surface. Orienting the Cartesian axes along the eigenvector directions renders

$$(\nabla \delta h)^2 \approx (x/R_1)^2 + (y/R_2)^2 = \delta h \left[\frac{1}{R_1} + \frac{1}{R_2} \right], \quad (35.8)$$

where we set

$$\delta h/R_1 = (x/R_1)^2 \quad \text{and} \quad \delta h/R_2 = (y/R_2)^2. \quad (35.9)$$

We are thus led to the area difference

$$\delta A \approx dA \delta h \left[\frac{1}{R_1} + \frac{1}{R_2} \right]. \quad (35.10)$$

Note that $\delta A > 0$ whether displacing the surface into a concave or convex direction, since the sign of δh accounts for the sign of the radii of curvature.

The total work for the interface displacement is given by the sum of the area work and volume work

$$W_{\text{area}} + W_{\text{volume}} = dA \delta h [\gamma (R_1^{-1} + R_2^{-2}) + p_B - p_A] \quad (35.11)$$

and equilibrium results if the work vanishes

$$p_A - p_B = \gamma (R_1^{-1} + R_2^{-2}). \quad (35.12)$$

This equation is the Young-Laplace formula, which reduces to equation (35.3) if $R_1 = R_2$ as for a sphere. It says that there is a pressure jump, known as the *capillary pressure*, across an interface as given by the surface tension times the sum of the inverse principle radii of curvature. Pressure is higher on the concave side of the interface, such as fluid-A depicted in Figure 35.4 or the inside of a bubble/droplet.

35.4 Some oceanographic examples

We close this chapter with examples relevant to the ocean.

35.4.1 Soluble gas bubbles inside water

The previous considerations hold whether there is liquid or gas inside the spherical droplet/bubble. As an example, consider a spherical gas bubble of radius $R = 10^{-6}$ m inside water and make use of the air-water surface tension $\gamma = 0.072$ N m⁻¹

$$p_{\text{in}} - p_{\text{out}} = 2\gamma/R \approx 144 \times 10^3 \text{ N m}^{-2} = 1.42 p_{\text{atm}}, \quad (35.13)$$

where $p_{\text{atm}} = 101 \times 10^3$ N m⁻² is standard atmospheric pressure. If the gas inside the bubble is water soluble, then the enhanced pressure inside the bubble will induce more gas to dissolve in the water, which in turn will cause the bubble to shrink and thus increase the pressure inside the bubble. Small bubbles of soluble gases can thus be squeezed towards zero radius by the effects of surface tension induced pressure.

35.4.2 Length scale for capillary waves

Capillary waves arise along the air-sea interface due to the restorative effects from surface tension. When present within a gravity field, the capillary waves appear along with gravity waves. We see capillary waves when there is a very slight breeze on the ocean surface. Capillary waves also arise when a tiny stone is thrown into a still pond, whereas gravity waves dominate when a larger stone is used. This phenomenology arises from the following considerations of the dispersion relation.

The dispersion relation for capillary-gravity waves (e.g., Section 54 of [Fetter and Walecka \(1980\)](#)) is given by

$$\omega^2 = k g \left[1 + \frac{k^2 \gamma}{g \rho} \right], \quad (35.14)$$

where ρ is the density of water, $k = 2\pi/\lambda$ is the wave number, and ω is the radial frequency. The non-dimensional parameter $k^2 \gamma / (g \rho)$ provides a regime boundary where capillary waves are

important ($k^2 \gamma > g \rho$) and negligible ($k^2 \gamma < g \rho$). To deduce a correspondingly length scale we introduce the wavelength

$$\lambda_{\text{cap-grav}} = 2\pi \sqrt{\frac{\gamma}{\rho g}}, \quad (35.15)$$

with $\lambda < \lambda_{\text{cap-grav}}$ the capillary wave regime and $\lambda > \lambda_{\text{cap-grav}}$ the gravity wave regime. Using $\gamma = 0.072 \text{ N m}^{-1}$ and $\rho = 1000 \text{ kg m}^{-3}$ leads to

$$\lambda_{\text{cap-grav}} = 2\pi \sqrt{\frac{\gamma}{\rho g}} \approx 0.017 \text{ m} = 17 \text{ cm}. \quad (35.16)$$

Since this book is mostly concerned with length scales larger than $\lambda_{\text{cap-grav}}$, we generally ignore the dynamics of capillary waves for our study of geophysical fluid mechanics.

35.5 Further study

Although we have no further concern for surface tension in this book, its study forms an important aspect of air-sea interaction physics. There are many places to continue its study, with the following offering treatments similar to the physical ideas given here.

This [30-minute video](#) from Prof. L. Trefethen provides a pedagogical summary of surface tension. The upside-down container of water in Section 35.1 is based on a discussion of capillary-gravity waves in Section 3.1.3 of [Falkovich \(2011\)](#). Section 1.9 of [Batchelor \(1967\)](#) discusses how surface tension acts between two fluid media, with that discussion extended into his Section 3.3 to develop boundary conditions for velocity and stress. The bubble example in Section 35.4.1 is taken from Section 1.3 of [Kundu et al. \(2012\)](#). Section 4.10 of [Kundu et al. \(2012\)](#) provides a detailed accounting of the force balance at an interface, offering more details than found in [Batchelor \(1967\)](#). The energetic arguments used to derive the Young-Laplace formula follows Section 61 of [Landau and Lifshitz \(1987\)](#). Section 46 of [Fetter and Walecka \(1980\)](#) discuss the dynamics of membranes under tension, and Section 54 considers surface capillary-gravity waves.

36

Surface gravity waves and Stokes drift

In this chapter we consider the ocean surface to be a material interface and examine its small amplitude fluctuations. These fluctuations manifest as linear surface gravity waves whose existence relies on the restoring force from gravity. They provide the prototypical wave-like response that offer an ideal physical system to introduce linear wave behavior. In formulating the equations for surface gravity waves, we bring together a number of fluid mechanical principles considered earlier in this book. In so doing, this analysis supports our understanding of how basic principles are used to describe motion. Notably, we here ignore rotation, although many properties illustrated here well approximate surface waves found in the ocean.

Our treatment of surface waves is relatively brief, offering only sufficient details to expose their characteristic exponential decay with depth when moving into the ocean interior. This depth decay leads to a general class of fluid particle motion known as *Stokes drift*, where Stokes drift arises for all waves whose amplitude changes in those directions not parallel to the wave phase direction. So although the surface gravity waves are linear, the depth decay in their amplitude leads to the a net drift motion of fluid particles and hence to the transport of trace matter. The Stokes drift exemplified by surface gravity waves provides an example of how averaging at a fixed point in space (Eulerian average) yields distinct behaviors from averaging on a fixed fluid particle (Lagrangian average). The kinematics of Stokes drift has wide application throughout geophysical fluid mechanics.

READER'S GUIDE TO THIS CHAPTER

In deriving the basic equations of surface gravity waves we make use of dynamical ideas from Chapter 22 and elements of the filtered equations from Chapter 27. We also make use of ideas from partial differential equations introduced in Chapter 5. The mathematical description of Stokes drift requires an understanding of Eulerian and Lagrangian kinematic descriptions from Chapter 16. Generalizations of Stokes drift appear in Chapter 38 in our study of wave-mean flow interactions, isopycnal averaging, and the corresponding eddy-induced tracer transport. Finally, we ignore surface tension in this chapter given that our interests are for wave motions with lengths far larger than the capillary wavelengths (a few centimeters) discussed in Section 35.4.2.

36.1	Surface gravity waves	528
36.1.1	Harmonic scalar potential	528
36.1.2	Equation of motion and Bernoulli's Principle	529
36.1.3	Non-hydrostatic pressure and the shallow water limit	530
36.1.4	Dynamic boundary condition at the free surface	531
36.1.5	Kinematic boundary conditions	531
36.1.6	Summary of the linear equation set	532
36.2	Surface gravity waves in a flat bottom channel	532
36.2.1	Monochromatic wave solution	533
36.2.2	Dispersion relation for surface gravity waves	533
36.2.3	Heuristics of waves in the shortwave/deep water limit	534
36.2.4	Heuristics of waves in the longwave/shallow water limit	534
36.2.5	Further study	535
36.3	Stokes drift	535
36.3.1	General formulation of Stokes drift	535
36.3.2	Stokes drift in surface gravity waves	537
36.3.3	Comments and further study	540
36.4	Exercises	541

36.1 Surface gravity waves

We study linear fluctuations of the ocean free surface on a non-rotating plane, with the free surface assumed to be a material interface separating a homogeneous ocean from a homogeneous atmosphere (Figure 36.1). The mass of the atmosphere is horizontally uniform and static so that surface atmosphere pressure does not contribute to oceanic motion. We develop the boundary value problem describing linear wave motions of the free surface, and characterize physical aspects of the waves.

36.1.1 Harmonic scalar potential

We are interested in fluctuations that have zero vorticity so that the velocity field has zero curl¹

$$\nabla \wedge \mathbf{v} = 0. \quad (36.1)$$

Recall from Section 4.3.2 that vector fields with zero curl can be determined from a scalar potential

$$\mathbf{v} = -\nabla\psi. \quad (36.2)$$

Note that the scalar potential is unspecified up to an arbitrary function of time, since ψ and $\psi + F(t)$ yield the same velocity field. We make use of this *gauge* freedom in Section 36.1.2.

Since the ocean is assumed to have uniform density, the velocity field is non-divergent. Consequently, the scalar potential satisfies Laplace's equation (Section 4.3.3)

$$\nabla \cdot \mathbf{v} = -\nabla \cdot \nabla\psi = \nabla^2\psi = 0. \quad (36.3)$$

ψ is termed a *harmonic function*, with salient properties noted in Section 5.4. To fully specify the scalar potential requires boundary conditions, which enter our development via the equation of motion.

¹We study fluids with nonzero vorticity in Part VIII of this book.

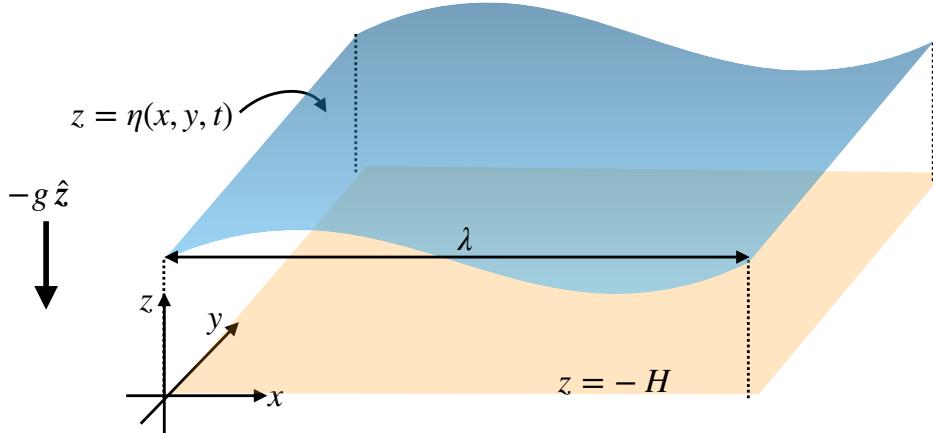


Figure 36.1: A depiction of the ocean free surface as a mathematical surface whose linear fluctuations exhibit gravity wave motion due to the restoring effects from a uniform gravitational field, $\mathbf{g} = -g \hat{\mathbf{z}}$. The atmosphere applies a pressure to the ocean due to its mass; however, that mass is assumed to be uniform and static so that it does not affect surface motion. Our formulation of surface waves makes no restrictions on the depth of the fluid relative to the wavelength of the waves, with the resulting waves exhibiting dispersion (i.e., the wave speed is a function of the wavelength). However, assuming the depth to be far smaller than the wavelength, $H \ll \lambda$ recovers the dispersion relation for linear non-dispersive shallow water gravity waves discussed in Section 44.1.

36.1.2 Equation of motion and Bernoulli's Principle

The vector-invariant equation of motion (22.35) for a non-rotating, irrotational, inviscid, uniform density fluid is given by

$$\frac{\partial \mathbf{v}}{\partial t} = -\nabla(\Phi + \mathcal{K} + p/\rho), \quad (36.4)$$

with

$$\mathcal{K} = \frac{\mathbf{v} \cdot \mathbf{v}}{2} \quad (36.5)$$

the kinetic energy per mass of a fluid element. In the following we assume the simple form of the geopotential (Section 14.1.2)

$$\Phi = g z \quad (36.6)$$

with g the constant gravitational acceleration. Inserting the scalar potential, $\mathbf{v} = -\nabla\psi$, brings the equation of motion (36.4) to the form

$$\nabla(\Phi + \mathcal{K} + p/\rho - \partial_t\psi) = 0. \quad (36.7)$$

This equation means that everywhere in the fluid the dynamical fields satisfy

$$\Phi + \mathcal{K} + p/\rho - \partial_t\psi = C(t), \quad (36.8)$$

for some arbitrary time dependent function $C(t)$. This equation is a particular expression of Bernoulli's theorem encountered in Section 24.2.3.

Removing the time dependent function via a gauge transformation

We ascribe no physical meaning to the arbitrary function $C(t)$ appearing in equation (36.8). In fact, it can be completely removed by exploiting the gauge degree of freedom in the scalar potential as noted following equation (36.2). We do so by introducing a modified scalar potential

$$\Psi(\mathbf{x}, t) = \psi(\mathbf{x}, t) + \int^t C(t') dt'. \quad (36.9)$$

Both ψ and Ψ lead to the same velocity vector

$$\mathbf{v} = -\nabla\psi = -\nabla\Psi \quad (36.10)$$

and as such they are physically indistinguishable. However, Ψ has the mathematical advantage of absorbing the arbitrary time dependent function $C(t)$, thus rendering the simpler expression for the equation of motion

$$\Phi + \mathcal{K} + p/\rho - \partial_t\Psi = 0. \quad (36.11)$$

This equation provides a balance condition that constrains the flow properties at every point within the fluid.

A brief diversion into Bernouilli's Principle for steady flow

The central focus of this analysis is with transient fluctuations of the free surface. Even so, we here comment on the case of steady flow, in which the constraint (36.11) leads to

$$\Phi + \mathcal{K} + p/\rho = 0. \quad (36.12)$$

Furthermore, assuming motion occurs on a constant geopotential then leads to

$$p + \rho\mathcal{K} = \text{constant}. \quad (36.13)$$

Motion on a geopotential means that all fluid particles feel the same gravitational acceleration. Hence, spatial changes in pressure provide the only force that can lead to spatial changes in fluid speed. That is the physical content of this special form of Bernouilli's Principle given by equation (36.13). It says that in regions of relatively high fluid speed the pressure is relatively low, and vice versa. Correspondingly, if the fluid moves into a region of relatively high pressure then it must slow down, whereas the speed increases when moving to a region of low pressure. This principle helps to explain the basic features of steady state flow around obstacles such as airplane wings.

36.1.3 Non-hydrostatic pressure and the shallow water limit

The fluid is non-hydrostatic since

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} + \frac{\partial(\partial_t\Psi - \mathcal{K})}{\partial z} = \frac{\partial p_{\text{hydrostatic}}}{\partial z} + \frac{\partial p_{\text{non-hydrostatic}}}{\partial z}. \quad (36.14)$$

Depth variations in the kinetic energy and depth-time variations in the velocity potential lead to deviations from hydrostatic balance. We do not generally expect the flow to be hydrostatically balanced for two reasons: (i) the fluid layer has a uniform density so there is no stratification to suppress vertical accelerations that contribute to non-hydrostatic pressures; (ii) the fluid is nonrotating and so there is no vertical stiffening via the Taylor-Proudman result (Section 29.4.3), with vertical stiffening acting to suppress vertical accelerations that cause deviations from hydrostatic balance.

For the shallow water model we also consider a homogenous density layer. However, as shown in Chapter 42, the hydrostatic balance is fundamental to the dynamics. In particular, in Section 42.1 we see that the hydrostatic balance over a single homogeneous layer leads to horizontal motion that is depth independent throughout the layer. Hence, $\partial_z(\partial_t\Psi) = 0$ and the kinetic energy in the horizontal motions is depth independent, $\partial_z(u^2 + v^2) = 0$. Furthermore, the vertical motion has a linear depth dependence across the shallow water layer (Section 42.1.6) and its magnitude is far

smaller than horizontal motions. Therefore, we can drop all contributions to $\partial_z \mathcal{K}$ for the shallow water layer, in which case equation (36.14) reduces to the hydrostatic limit

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} = -\rho g \quad \text{hydrostatic limit.} \quad (36.15)$$

So in summary, a homogeneous layer of fluid can have a depth dependence to its horizontal flow. But that depth dependence is driven only through non-hydrostatic pressure forces. The reason is the hydrostatic pressure has a depth-independent horizontal gradient within a homogeneous layer

$$\partial_z (\nabla_z p_{\text{hydrostatic}}) = -g \nabla_z \rho = 0. \quad (36.16)$$

36.1.4 Dynamic boundary condition at the free surface

Return now to the time dependent expression of the equation of motion (36.11). It is a rather remarkable result that applies to any point within the fluid and at any time. In particular, it applies at the free surface, $z = \eta(x, y, t)$ where pressure equals to the atmospheric pressure. As stated earlier, we assume that the atmospheric pressure is constant in space and time so that

$$g \eta + \mathcal{K} - \partial_t \Psi = -p_a / \rho = \text{constant}. \quad (36.17)$$

Without loss of generality we can set this constant to zero.² We are thus left with the boundary condition

$$g \eta + \mathcal{K} - \partial_t \Psi = 0 \quad \text{at } z = \eta. \quad (36.18)$$

At this point we linearize relative to a state of rest with $\eta = 0$, $\mathbf{v} = 0$, and $\partial_t \Psi = 0$. Linear fluctuations about this rest state have small velocities and as such the kinetic energy, which is second order in velocity, can be dropped to leave the linearized boundary condition

$$g \eta = \frac{\partial \Psi}{\partial t} \quad \text{linearized dynamic b.c. at } z = \eta. \quad (36.19)$$

This boundary condition directly connects the free surface to time tendencies of the velocity potential. The free surface rises when the velocity potential has a positive tendency, and vice versa.

36.1.5 Kinematic boundary conditions

The free surface is assumed to be a material interface, meaning that we ignore effects from matter transport across this surface. Consequently, following the discussion of kinematic boundary conditions in Section 17.4.2, we have

$$\frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta = w \quad \text{at } z = \eta. \quad (36.20)$$

Linearizing this boundary condition about the state of rest, and introducing the scalar potential, leads to

$$\frac{\partial \eta}{\partial t} = -\frac{\partial \Psi}{\partial z} \quad \text{linearized kinematic b.c. at } z = \eta. \quad (36.21)$$

This is yet another constraint that links the free surface to the velocity potential.

One final kinematic boundary condition applies just to the velocity potential when it intersects with the rigid solid boundaries. Namely, the no normal flow condition from Section 17.4.1 means that

$$\hat{\mathbf{n}} \cdot \nabla \Psi = 0 \quad \text{on rigid solid boundaries,} \quad (36.22)$$

where $\hat{\mathbf{n}}$ is the outward normal on the solid boundaries.

²Alternatively, make use of a further gauge transformation $\Psi' = \Psi - t(p_a/\rho)$ to eliminate the constant.

36.1.6 Summary of the linear equation set

The boundary value problem for the velocity potential and free surface is given by

$$\nabla^2 \Psi = 0 \quad \text{irrotational and non-divergent velocity throughout domain} \quad (36.23a)$$

$$\frac{\partial \Psi}{\partial t} = g \eta \quad \text{linearized dynamic b.c. at } z = \eta \quad (36.23b)$$

$$\frac{\partial \Psi}{\partial z} = -\frac{\partial \eta}{\partial t} \quad \text{linearized kinematic b.c. at } z = \eta \quad (36.23c)$$

$$\hat{n} \cdot \nabla \Psi = 0 \quad \text{no-normal flow kinematic b.c. on rigid boundaries.} \quad (36.23d)$$

The first equation holds throughout the fluid whereas the remaining equations hold only at the boundaries.

Although the equations (36.23a)-(36.23d) were derived through linearization, there is an additional nonlinear term to remove in order to allow for analytical treatment of the wave solutions. Namely, when combining the boundary conditions into a single equation we compute the time derivative of equation (36.23b) according to

$$g \frac{\partial \eta}{\partial t} = \left[\frac{\partial}{\partial t} + \frac{\partial \eta}{\partial t} \frac{\partial}{\partial z} \right] \frac{\partial \Psi}{\partial t} \quad (36.24)$$

and then combine with equation (36.23c) to render

$$\left[\frac{\partial}{\partial t} + \frac{\partial \eta}{\partial t} \frac{\partial}{\partial z} \right] \frac{\partial \Psi}{\partial t} = -g \frac{\partial \Psi}{\partial z}. \quad (36.25)$$

With $w \approx \partial \eta / \partial t$ at the free surface, we identify $(\partial \eta / \partial t) \partial_z$ as a vertical advection operator. The corresponding term $(\partial \eta / \partial t) \partial_{zt} \Psi$ is nonlinear and second order in perturbation terms. It is therefore dropped when fully linearizing the system. An equivalent means to realize this linearization is to evaluate the free surface boundary condition at $z = 0$ rather than at $z = \eta(x, y, t)$. For this approximation to be self-consistent requires the amplitude of free surface undulations to be much smaller than the typical wavelengths of the fluctuations

$$|\eta|/\lambda \ll 1. \quad (36.26)$$

In summary, the fully linearized equation set takes the form

$$\nabla^2 \Psi = 0 \quad \text{irrotational and non-divergent velocity throughout domain} \quad (36.27a)$$

$$g \eta = \frac{\partial \Psi}{\partial t} \quad \text{linearized dynamic b.c. at } z = 0 \quad (36.27b)$$

$$\frac{\partial \eta}{\partial t} = -\frac{\partial \Psi}{\partial z} \quad \text{linearized kinematic b.c. at } z = 0 \quad (36.27c)$$

$$\hat{n} \cdot \nabla \Psi = 0 \quad \text{no-normal flow kinematic b.c. on rigid boundaries.} \quad (36.27d)$$

36.2 Surface gravity waves in a flat bottom channel

We now derive a wave-like solution to the equations (36.27a)-(36.27d) posed in a flat bottom channel domain as illustrated in Figure 36.1. This analysis provides experience with the *separation of variables* method of use for deriving solutions to certain partial differential equations. We are not interested in the most general wave solution. Instead, we aim to determine a particular solution

of sufficient generality to expose the underlying physics of the linear wave fluctuations, and in particular to expose the exponential decay of the wave amplitude with depth. Furthermore, given linearity, the superposition principle holds whereby the linear sum of particular solutions is also a solution.

36.2.1 Monochromatic wave solution

We seek a monochromatic wave solution with the wave moving parallel to the channel with radial frequency $\omega = 2\pi/\tau$ and wavenumber $k = 2\pi/\lambda$, with τ the period and λ the wavelength. We furthermore assume the waves appear in the velocity potential in the shape of a cosine with an undetermined vertical structure

$$\Psi(x, y, z, t) = \Gamma(z) \cos(kx - \omega t). \quad (36.28)$$

Note the assumed absence of y dependence, as motivated by our choice to examine waves moving parallel to the channel and with symmetry across the channel. Furthermore, the cosine and sine functions form a complete set of basis functions for the channel domain, so that they can be superposed to generate any arbitrary wave pattern. This property is the basis for Fourier analysis commonly employed to study waves.

Plugging the ansatz (36.28) into Laplace's equation $\nabla^2\Psi = \partial_{xx}\Psi + \partial_{zz}\Psi = 0$ leads to the ordinary differential equation satisfied by the vertical structure function

$$\frac{d^2\Gamma}{dz^2} = k^2 \Gamma \quad -H \leq z \leq 0 \quad (36.29a)$$

$$\frac{d\Gamma}{dz} = 0 \quad \text{at } z = -H, \quad (36.29b)$$

where the bottom boundary condition is required to satisfy the no-normal flow condition (36.27d). We write the solution in the form

$$\Psi = \Psi_o \cosh[k(z + H)] \cos(kx - \omega t), \quad (36.30a)$$

$$\Psi_o = \frac{g \eta_o / \omega}{\cosh(kH)} \quad (36.30b)$$

so that the boundary condition (36.27b) renders the free surface height

$$\eta(x, t) = -\eta_o \sin(kx - \omega t). \quad (36.31)$$

The corresponding velocity field, $\mathbf{v} = -\nabla\Psi$, is given by

$$u = -\Psi_o k \cosh[k(z + H)] \sin(kx - \omega t) \quad (36.32a)$$

$$w = -\Psi_o k \sinh[k(z + H)] \cos(kx - \omega t). \quad (36.32b)$$

36.2.2 Dispersion relation for surface gravity waves

Combining the two $z = 0$ boundary conditions (36.27b) and (36.27c) yields

$$\frac{\partial^2\Psi}{\partial t^2} + g \frac{\partial\Psi}{\partial z} = 0 \quad z = 0. \quad (36.33)$$

Substituting the wave solution (36.30a) into this relation leads to the constraint

$$\omega^2 = g k \tanh(kH). \quad (36.34)$$

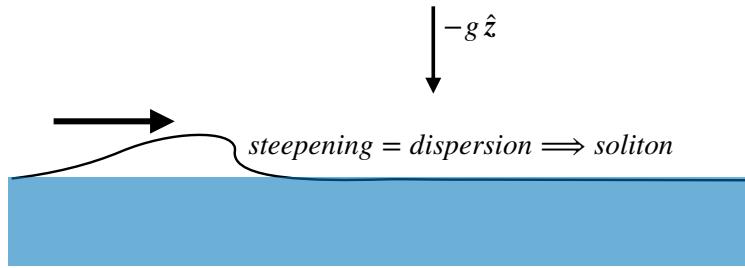


Figure 36.2: A soliton in the deep water limit results when the wave dispersion (long waves travel faster than short waves) balances the nonlinear steepening. The result is a soliton, which has an exact analytic expression following from the KdV equation (see [Drazin and Johnson \(1989\)](#)).

This is the *dispersion relation* that constrains values available for the radial frequency, ω , and wavenumber, k . That is, the surface gravity waves only exist if their frequency and wavenumber are related according to the dispersion relation (36.34). The corresponding phase speed for the wave is given by

$$c = \frac{\omega}{k} = \sqrt{(g/k) \tanh(kH)}. \quad (36.35)$$

We emphasize the two limits: $kH \gg 1$ (shortwaves/deep water) and $kH \ll 1$ (longwave/shallow water), in which the phase speed satisfies

$$\omega \approx \sqrt{gk} \quad c \approx \sqrt{g/k} \quad kH \gg 1 \quad \text{shortwave/deep water limit} \quad (36.36a)$$

$$\omega \approx k\sqrt{gH} \quad c \approx \sqrt{gH} \quad kH \ll 1 \quad \text{longwave/shallow water limit.} \quad (36.36b)$$

36.2.3 Heuristics of waves in the shortwave/deep water limit

The shortwave/deep water waves are notable for having shorter waves travel slower than longer waves. In the event of a perturbation to the fluid, such as from a stone dropped into a pond or a storm on a lake or the ocean, deep water waves are energized. The dispersion relation means that longer waves spread away from the storm center faster than the shorter waves, leading to a self-organization of the wavelengths and corresponding wave packets.

Now imagine a deepwater wave packet (group of waves traveling together) that somehow steepens and takes on a nonlinear form. Fourier decomposing this nonlinear wave into linear deepwater modes requires more shortwave linear modes in the steep region, whereas the less steep portion of the wave requires longer deep water Fourier modes, which travel faster. If the nonlinear steepening on the wave face is exactly balanced by the faster dispersion of the long waves near the wave base and back-side, then the wave pattern remains stable; it does not break. This balance of steepening and dispersion describes the fundamental features of a soliton.

36.2.4 Heuristics of waves in the longwave/shallow water limit

The longwave/shallow water limit is notable for the absence of wave dispersion; i.e., shallow water gravity waves of all wavelengths travel at speed \sqrt{gH} . Tsunamis are the prototypical shallow water waves that travel at speeds well approximated by \sqrt{gH} .

The dispersion relation also means that shallow water gravity waves slow down when the depth shoals, as when approaching a beach. Consequently, as waves reach the shoreline there is a tendency to accumulate wave energy as the deeper waves pile up behind the shallower waves. Furthermore, the steeper part of the wave, being part of a thicker region of the fluid and thus a larger effective H , travels slightly faster than the wave trough. As such, the steeper part of the wave overtakes the

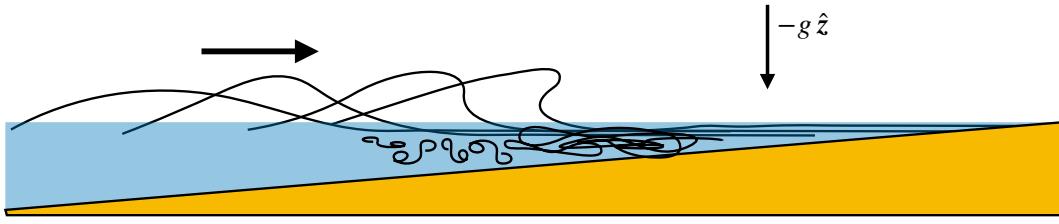


Figure 36.3: Shallow water waves approaching a shoreline steepen and eventually break. We can infer this behavior from the dispersion relation, \sqrt{gH} , whereby waves in deeper water move slightly faster than those in shallower water, so that the wave energy accumulates near the shore. Furthermore, water on the steeper part of the wave moves slightly faster than water in the trough, due to the difference in thickness of the water. This process causes water on the steeper portion of the wave to travel slightly faster than in the trough, leading to steepening of the waves. Nonlinearities eventually invalidate the assumptions made in deriving the linear wave dispersion relation $c = \sqrt{gH}$. Even so, the qualitative characterization based on the linear analysis allows for a useful heuristic understanding of shallow water wave breaking on the beach.

trough and, at some point, the assumptions of linearity breakdown and the shallow water waves break on the shore as depicted in Figure 36.3.

36.2.5 Further study

Elements from this section follow Section 54 of the [Fetter and Walecka \(1980\)](#), with their discussion of surface gravity waves also probing topics such as energetics, group velocity, and the inclusion of surface tension as appropriate for deep water waves with wavelengths on the order of centimeters. An accessible development of the theory of solitons can be found in the book by [Drazin and Johnson \(1989\)](#). Ocean surface waves are generally energized by atmospheric winds, in particular large storms. The dynamics of gravity wave breaking in the open ocean and on the beach remains a decades-long problem in fluid mechanics that stretches the limits of theory, experiment, and simulation. [Bühler \(2014\)](#) offers a thorough treatment of waves and their effects on mean flows

This video from Prof. A. Hogg provides a pedagogical introduction to shallow water wave breaking as well as deep water solitons realized in a laboratory.

36.3 Stokes drift

We here consider a fluid particle moving as part of a wave field. In homogenous linear wave fields, the particle periodically returns to its original position. However, in the presence of wave inhomogeneities, such as the surface gravity waves considered in this chapter, fluid particles generally oscillate between regions where the undulation in one direction does not match that in the other direction. In this case there is a net particle drift or transport known as *Stokes drift*. This drift occurs even though the phase average of the wave at a fixed spatial point vanishes. Formulating the mathematics of Stokes drift offers a means to explore the differences between averages formed at a fixed space point (Eulerian mean) versus averages following fluid particles (Lagrangian mean). We here introduce these notions, which form part of the rudiments for wave-mean flow interaction theory further pursued in Chapter 38.

36.3.1 General formulation of Stokes drift

Consider a three-dimensional particle trajectory written in Cartesian coordinates,

$$\mathbf{X}(\mathbf{a}, t) = X(\mathbf{a}, t) \hat{\mathbf{x}} + Y(\mathbf{a}, t) \hat{\mathbf{y}} + Z(\mathbf{a}, t) \hat{\mathbf{z}}. \quad (36.37)$$

In the analysis of waves, it is common to assume the material coordinate, \mathbf{a} , is the initial position of a fluid particle so we make that assumption here. As discussed in Section 16.7.1, the particle trajectory is determined by time integrating the particle velocity (also known as the “flow map”)

$$\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (36.38)$$

so that

$$\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, 0) + \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t'), t'] dt'. \quad (36.39)$$

This equation is a trivial result of time integrating the flow map. Nonetheless, it is useful to express the content of this equation in words. It says that the position at time t of a fluid particle labelled by the material coordinate \mathbf{a} is given by the initial position of the particle, $\mathbf{X}(\mathbf{a}, 0)$, plus the time integrated movement of the particle following the fluid flow.

We now form a Taylor series computed relative to the initial position of the particle, so that the particle velocity at time t takes on the approximate form

$$v^n[\mathbf{X}(\mathbf{a}, t), t] \approx v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot [\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)] \quad (36.40a)$$

$$= v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \frac{d\mathbf{X}(\mathbf{a}, t')}{dt'} dt' \quad (36.40b)$$

$$= v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t'), t'] dt', \quad (36.40c)$$

where the Taylor series was truncated after terms linear in the particle displacement $\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)$. We emphasize two points regarding equation (36.40c).

- How do we interpret $v^n[\mathbf{X}(\mathbf{a}, 0), t]$? This is the n 'th component of the velocity field evaluated at the initial point of the trajectory, $\mathbf{X}(\mathbf{a}, 0)$, at time t . That is, it is the Eulerian velocity evaluated at the fixed Eulerian point $\mathbf{X}(\mathbf{a}, 0)$.
- What determines the accuracy of the Taylor series? A suitable non-dimensional expansion coefficient for the Taylor expansion is the ratio of the particle displacement to the scale, Λ , of inhomogeneities in flow properties

$$\epsilon = \frac{|\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)|}{\Lambda}. \quad (36.41)$$

This ratio is small for the small amplitude waves considered here, whereby the particle displacements are far smaller than inhomogeneities in flow properties.

The integrand on the right hand side of equation (36.40c) is the Lagrangian velocity integrated over the time interval. To within the same order of accuracy as maintained thus far, we can use the Eulerian velocity evaluated at the initial position, thus rendering

$$v^n[\mathbf{X}(\mathbf{a}, t), t] \approx v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt', \quad (36.42)$$

with rearrangement leading to

$$v^n[\mathbf{X}(\mathbf{a}, t), t] - v^n[\mathbf{X}(\mathbf{a}, 0), t] \approx \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt'. \quad (36.43)$$

The left hand side is the difference between the velocity following a fluid particle (the Lagrangian velocity for the moving fluid particle) from the velocity at the initial particle point (the Eulerian velocity at the initial point of the trajectory). The right hand side terms are all evaluated at the initial position, $\mathbf{X}(\mathbf{a}, 0)$. Furthermore, the right hand side is non-zero where the velocity at the initial position has a nonzero gradient (i.e., it is spatially inhomogeneous), with its inhomogeneity projecting onto the time integrated velocity at that point. Equation (36.43) says that the velocity following a fluid particle is modified from the velocity at its initial position if the particle moves through an inhomogeneous velocity field.

The Stokes drift is defined as the difference of the velocities in equation (36.43) when time averaged over a wave period, which we write as

$$v_{(S)}^n[\mathbf{X}(\mathbf{a}, 0), t] = \overline{v^n[\mathbf{X}(\mathbf{a}, t), t] - v^n[\mathbf{X}(\mathbf{a}, 0), t]} \quad (36.44a)$$

$$\approx \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt'. \quad (36.44b)$$

This expression holds for any arbitrary initial point in the fluid, so that we can write it in a concise Eulerian form that dispenses with trajectories

$$v_{(S)}^n(\mathbf{x}, t) \approx \overline{\nabla v^n(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt'}. \quad (36.45)$$

We liken Stokes drift to surfing: the more a particle samples larger amplitude variations in the velocity field (the gradient term), the further it drifts (the integral term). Note that for the case of a transverse wave disturbance, where the particle disturbance is orthogonal to the wavevector, then the Stokes drift vanishes to leading order.

36.3.2 Stokes drift in surface gravity waves

The canonical example of Stokes drift occurs in the near surface ocean, where the surface gravity waves discussed in Section 36.1 create particle motion that is larger near the surface than at depth. For this example, consider a monochromatic wave field in the zonal-vertical directions written in the form

$$\frac{dX}{dt} = U \sin(\omega t) \quad (36.46a)$$

$$\frac{dZ}{dt} = U \cos(\omega t), \quad (36.46b)$$

where $U > 0$ is the speed of the circular particle motion and we use the capital X and Z to denote Cartesian components of the particle trajectory. To simplify the mathematics we perform the analysis in a frame where the waves are stationary, hence the $k x - \omega t$ phase from Section 36.2. Figure 36.4 shows a schematic of the particle trajectories as derived in the following.

Homogeneous flow field

If the background flow is homogeneous, then the speed is a constant, $U = U_o > 0$. Particle trajectories in this case are clockwise in the x-z plane around a circle with radius U_o/ω

$$X(t) - X_o = -\frac{U_o}{\omega} [\cos(\omega t) - 1] \quad (36.47a)$$

$$Z(t) - Z_o = \frac{U_o}{\omega} \sin(\omega t), \quad (36.47b)$$



Figure 36.4: Sketch of Stokes drift in the upper ocean with the wave vector in the horizontal direction and clockwise time integrated fluid particle motion induced by the traveling wave. For the case of a wave amplitude that decreases with depth, lateral motion of the particle is larger when the particle is closer to the surface thus leading to a Stokes drift in the direction of the wave. Note that there is zero Stokes drift for the case of a homogeneous wave, in which the wave amplitude is independent of depth. Additionally, and even more trivially, if the particle motion is purely transverse to the wave vector, in this case purely vertical, then the particle merely retraces its motion along a vertical line and does not undergo any lateral Stokes drift.

where the initial position at time $t = 0$ is

$$\mathbf{X}(t = t_o) = \mathbf{X}_o, \quad (36.48)$$

and the center of the circle is

$$\mathbf{X}_{\text{center}} = \left[X_o + \frac{U_o}{\omega} \right] \hat{\mathbf{x}} + Z_o \hat{\mathbf{z}}. \quad (36.49)$$

Inhomogeneous flow field

In the presence of vertical inhomogeneities of the flow field, the wave amplitude becomes a function of depth, $U = U(z)$. The canonical example is where the wave amplitude decreases with depth, as for the surface gravity waves detailed in Section 36.2. In turn, we expect there to be a fluid particle drift in the zonal direction introduced by the vertical wave inhomogeneity. This drift is a particular realization of Stokes drift.

To compute the leading order expression for the Stokes drift, expand U in a Taylor series about the initial position

$$U \approx U_o + R(Z - Z_o) \quad (36.50)$$

where R has units of inverse time and is given by

$$R = \left[\frac{dU}{dZ} \right]_{Z=Z_o}. \quad (36.51)$$

For the surface gravity waves with velocity components (36.32a) and (36.32b), we have

$$U = \Psi_o k \sqrt{\cosh[2k(z + H)]} \quad (36.52)$$

so that with $Z_o = 0$,

$$R = \Psi_o k^2 \frac{\sinh(2kH)}{\sqrt{\cosh(2kH)}}. \quad (36.53)$$

In general, the Taylor series is valid so long as the vertical trajectories maintain the inequality

$$|R| |Z - Z_o| \ll U_o, \quad (36.54)$$

which says that the vertical shear is small

$$|R| \ll \frac{U_o}{|Z - Z_o|}. \quad (36.55)$$

We use the Taylor series expansion (36.50) to solve for the vertical trajectory as determined by

$$\frac{d(Z - Z_o)}{dt} = [U_o + R(Z - Z_o)] \cos(\omega t). \quad (36.56)$$

Rearrangement leads to

$$\int_{Z_o}^Z \frac{d(Z - Z_o)}{U_o + R(Z - Z_o)} = \int_0^t \cos(\omega t) dt. \quad (36.57)$$

The left hand side integral can be computed by changing variables

$$\Sigma = U_o + R(Z - Z_o) \quad (36.58a)$$

$$d\Sigma = R d(Z - Z_o), \quad (36.58b)$$

so that equation (36.57) becomes

$$\int_{U_o}^{\Sigma} \frac{d\Sigma}{\Sigma} = R \int_0^t \cos(\omega t) dt. \quad (36.59)$$

Performing the integrals and evaluating the end points renders

$$\ln \left[1 + \frac{R}{U_o} (Z - Z_o) \right] = \frac{R \sin(\omega t)}{\omega}, \quad (36.60)$$

which yields the exponential solution

$$1 + \frac{R}{U_o} (Z - Z_o) = e^{(R/\omega) \sin(\omega t)} \Rightarrow Z - Z_o = \frac{U_o}{R} \left(-1 + e^{(R/\omega) \sin(\omega t)} \right). \quad (36.61)$$

The vertical particle position is seen to oscillate around its initial position Z_o .

We next consider the zonal particle position, in which case

$$\frac{d(X - X_o)}{dt} = U_o \left[1 + \frac{R}{U_o} (Z - Z_o) \right] \sin(\omega t) \quad (36.62a)$$

$$= U_o e^{(R/\omega) \sin(\omega t)} \sin(\omega t) \quad (36.62b)$$

where we used equation (36.61) for the vertical trajectory. To make progress, we expand the exponential assuming the ratio of inverse time scales, R/ω , is small

$$|R/\omega| \ll 1. \quad (36.63)$$

In this limit, the vertical trajectory retains its unperturbed form (36.47b), and the zonal trajectory satisfies

$$\frac{d(X - X_o)}{dt} \approx U_o \sin(\omega t) \left[1 + \frac{R}{\omega} \sin(\omega t) \right], \quad (36.64)$$

where we dropped terms of order $(R/\omega)^2$. We can understand the scaling in equation (36.63) by noting that the period for the circular motion is given by

$$\tau_{\text{circle}} = \frac{2\pi}{\omega}. \quad (36.65)$$

The inverse time R introduces a time scale for the drift, defined according to

$$\tau_{\text{drift}} = \frac{2\pi}{|R|}. \quad (36.66)$$

A small ratio $|R/\omega|$ thus implies

$$|R/\omega| = \tau_{\text{circle}}/\tau_{\text{drift}} \ll 1. \quad (36.67)$$

Hence, we are solving for the zonal trajectory in the limit where the time scale for the circular motion is small (i.e., fast oscillations around the circle) relative to the time scale for the drift (i.e., slow drift).

Returning now to the approximate zonal trajectory equation (36.64) yields

$$\frac{d(X - X_o)}{dt} = U_o \sin(\omega t) \left[1 + \frac{R}{\omega} \sin(\omega t) \right] \quad (36.68a)$$

$$= U_o \sin(\omega t) + \frac{U_o R}{2\omega} [1 - \cos(2\omega t)], \quad (36.68b)$$

which integrates to

$$X - X_o = \left(\frac{U_o}{\omega} \right) \left[1 - \cos(\omega t) - \frac{R \sin(2\omega t)}{4\omega} + \frac{R t}{2} \right] \quad (36.69a)$$

$$= \underbrace{\left(\frac{U_o}{\omega} \right) [1 - \cos(\omega t)]}_{\text{homogeneous}} + \underbrace{\frac{U_o R t}{2\omega}}_{\text{Stokes drift}} - \underbrace{\frac{U_o R \sin(2\omega t)}{4\omega^2}}_{\text{higher harmonic}} + \mathcal{O}(R/\omega)^2. \quad (36.69b)$$

The leading order term is the homogeneous motion given by equation (36.47a). The next term is the Stokes drift, followed by a higher order harmonic and then further terms on the order of $(R/\omega)^2$. There is no vertical Stokes drift to this order in (R/ω) , so that the Stokes drift velocity is given by

$$\left[\frac{\mathbf{X} - \mathbf{X}_o}{t} \right]^{\text{drift}} = \frac{R U_o}{2\omega} \hat{\mathbf{x}}. \quad (36.70)$$

The circular motion of the parcels is therefore deformed by the zonal Stokes drift. The drift increases with larger wave amplitude (U_o large); with larger vertical shear (R large); and with longer period waves (ω small). See Figure 36.5 for an illustration based on a particular choice for the dimensional parameters.

36.3.3 Comments and further study

Stokes drift occurs in many guises when studying the motion of fluid particles within wave fields. We revisit elements of Stokes drift in Chapter 38 when studying the rudiments of eddy-induced tracer transport. This video from [Prof. Andy Hogg at Australian National University](#) provides an overview of the discussion in this section along with some laboratory experiments to illustrate Stokes drift.

Is Stokes drift a nonlinear wave phenomena? In answering this question we note that Stokes drift occurs with particle motion in linear waves, but the waves must be inhomogeneous such as the surface gravity waves studies in this chapter. However, nonlinearity appears in the form of the particle-following averaging, as can be seen by the expression of Stokes drift given by equation (36.45)

$$\overline{v_{(S)}^n(\mathbf{x}, t)} \approx \nabla v^n(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt'. \quad (36.71)$$

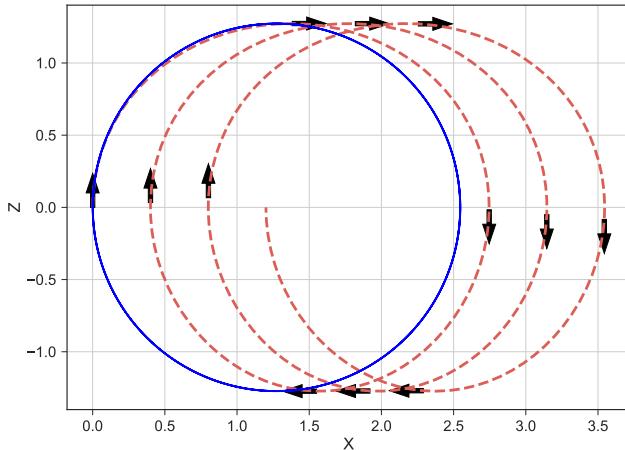


Figure 36.5: Example trajectories of fluid particles undergoing Stokes drift for short surface ocean waves. Particle motion is clockwise in the x - z plane. For homogeneous waves, there is zero Stokes drift with circular trajectories given by equations (36.47a) and (36.47b), as depicted here by the blue trajectory. There is a Stokes drift in the presence of vertical shear in the wave amplitude and thus in the particle velocity, with the trajectories for this example given by equation (36.61) for the vertical component and equation (36.69b) for the horizontal component. We set the parameters as follows: $T = 2\pi/\omega = 60$ s, $U_o = 0.1$ m s $^{-1}$, and $R = \omega/10$ and exhibit trajectories over four minutes.

The dot product of the velocity gradient with the time integrated velocity (to give the time integrated position) is nonlinear. So although the waves are linear, the Lagrangian kinematics of particle trajectories introduces nonlinearities.

36.4 Exercises

EXERCISE 36.1: STOKES DRIFT FOR ONE-DIMENSIONAL MONOCHROMATIC WAVE

Consider a one-dimensional monochromatic longitudinal wave with velocity

$$u = u_o \sin(k x - \omega t), \quad (36.72)$$

where u_o is the wave amplitude, $k = 2\pi/\lambda$ the wave number, λ the wavelength, $\omega = 2\pi/T$ the radial frequency, T the wave period, and $c = \omega/k = \lambda/T$ the wave speed. A longitudinal wave is one whose particle motions are parallel to the wave vector, which in this exercise are both in the \hat{x} direction. Determine the wave period averaged Stokes velocity to first order accuracy in the small parameter

$$\epsilon = \frac{u_o}{c} = \frac{u_o k}{\omega} = \frac{u_o T}{\lambda} \quad (36.73)$$

with this parameter the ratio of the wave amplitude to wave speed, or equivalently the ratio of the length scale of particle displacements to the wavelength. Hint: make use of the general result given by equation (36.45).

Part VI

Tracer mechanics

In this part of the book we develop the physics and mathematics of scalar tracer fields, including thermodynamic tracers such as potential temperature and material tracers such as salinity and humidity. This *tracer mechanics* offers a distinct view on the fluid flows and thus provides a useful complement to the mechanics of momentum, vorticity, and energy.

We devote Chapter 37 to exploring the physical and mathematical properties of advection and diffusion when acting on scalar fields. On the large-scales, the advection and diffusion felt by a tracer generally arises from more than just the mean flow and molecular diffusion. In Chapter 38 we introduce the notions of wave-mean flow interactions that give rise to eddy-induced advection (or skew diffusion) and diffusion. Throughout this part of the book, in particular in Chapter 37, we encounter some of the canonical partial differential equations appearing in fluid mechanics. Mathematical facets of these equations are introduced in Chapter 5, which offers a synopsis of the linear partial differential equations of mathematical physics.

37

Advection and diffusion

In this chapter we discuss physical and mathematical aspects of advection and diffusion, which are the two primary processes that affect tracer concentration at a point in the fluid. Advection and diffusion have complementary physical and mathematical properties. In the absence of diffusion, advection imparts a reversible stirring of fluid elements that increases the magnitude of tracer gradients. Diffusion, in contrast, provides an irreversible mixing of fluid elements that reduces the magnitude of tracer gradients. [Eckart \(1948\)](#) articulated what has become the standard conceptual paradigm for stirring and mixing in geophysical fluids, with elements of that paradigm supported by the discussion in this chapter.

For most of this chapter we consider a compressible/non-Boussinesq fluid. For a Boussinesq fluid (Chapter 28), the density factor, ρ , appearing in the compressible formulation is set to a constant and trivially cancels from all formula. Although the incompressible/Boussinesq fluid is commonly assumed for the ocean, we retain the compressible/non-Boussinesq formulation as it adds generality without much complexity.

37.1	Introduction	546
37.2	Diffusion physics	547
37.2.1	Diffusion of matter by random molecular motions	547
37.2.2	Diffusion of matter by random turbulent motions	547
37.2.3	Fick's law for matter diffusion	548
37.2.4	Fourier's law for heat diffusion	549
37.2.5	Newtonian frictional stress and momentum diffusion	549
37.2.6	Further study	550
37.3	Diffusion maths	550
37.3.1	Sample diffusion tensors	551
37.3.2	Evolution of tracer concentration powers	551
37.3.3	Moments of tracer concentration	551
37.3.4	Connecting tracer variance to the diffusion operator	553
37.4	Advection physics	554
37.4.1	The advection equation	555
37.4.2	Eulerian time tendencies from advection	555
37.5	Advection maths	556
37.5.1	Material constancy of C^γ	556
37.5.2	Mass transport	556
37.5.3	Advective tracer fluxes and skew tracer fluxes	557
37.5.4	Skew diffusion	558
37.5.5	Further study	559
37.6	Advection and skewness	559
37.6.1	Choosing a gauge	559
37.6.2	Boundary conditions	561
37.7	Active tracers and dia-surface flow	561
37.7.1	Adiabatic flow	561
37.7.2	Diabatic processes generating dia-surface transport	562
37.8	Exercises	563

37.1 Introduction

As derived in Section 18.1, the tracer equation takes on the general form

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}(C), \quad (37.1)$$

where \mathbf{J} is a flux that embodies molecular diffusion as well as subgrid scale advection and subgrid scale diffusion (Chapter 39). Advective transport appears when transforming to an Eulerian or laboratory reference frame, in which case

$$\rho \frac{DC}{Dt} = \frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\mathbf{v} \rho C), \quad (37.2)$$

with $\mathbf{v} \rho C$ the advective flux. In the absence of diffusion, advection renders a reversible stirring and stretching of fluid elements that generally increases the magnitude of concentration gradients. Advection does so while maintaining, for each fluid element, a fixed mass for all matter constituents and fixed specific entropy.¹ In contrast, diffusion affects an irreversible exchange, or mixing, of matter, thermodynamic, and mechanical properties between fluid elements. Correspondingly, diffusion reduces the magnitude of property gradients as it irreversibly exchanges properties between fluid elements.

¹Recall from Chapter 23 that specific entropy remains materially constant on fluid parcels in the absence of mixing.

37.2 Diffusion physics

The continuum hypothesis from Chapter 2 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 motion through the process of *molecular diffusion* of matter. Analogously, the random motion of fluid elements within a turbulent fluid give rise to *turbulent diffusive transport*.² In this section, we explore the basic physical nature of molecular and turbulent diffusion.

37.2.1 Diffusion of matter by random molecular motions

Consider a fluid comprised of a single matter constituent, such as a lake of pure H_2O . As discussed in Section 16.1, for a macroscopic description of this single-component fluid, a constant mass fluid element is identical to a constant mass material fluid parcel. Now place a passive dye tracer (Section 18.2) into a corner of the lake so that the lake is comprised of two material components (H_2O 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.

We introduced the notion of matter exchange between fluid elements when discussing the tracer equation in Section 18.1. 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 hypothesis 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. This transport of matter by random molecular motions is known as *molecular diffusion*. A statistical description of molecular diffusion was first given by Einstein through his investigations of Brownian Motion ([Einstein, 1905](#)).

Diffusion of matter is a familiar process. For example, the odor from an open perfume bottle will spread throughout a room, even in the absence of macroscopic motion of air in the room. When the ambient macroscopic motion is zero, the spread of the perfume arises from random molecular motions whose properties depend on details of the molecules (e.g., their size, speed, inter-molecular forces). The time scale for molecular diffusion is generally much longer than the analogous *turbulent diffusion* that results if there is random motion in the macroscopic fluid, such as occurs by placing a fan next to the perfume bottle.

37.2.2 Diffusion of matter by random turbulent motions

It is common for geophysical fluid systems to exhibit some form of turbulent motion. In these systems, the spread of matter by macroscopic turbulent motion is many times more efficient than the spread of matter from molecular motion. In such cases, we are justified in ignoring molecular diffusion since the efficiency of the turbulent diffusive transport is far greater than that from molecular diffusion.

[Taylor \(1921\)](#) described the statistical properties of turbulent diffusion, with many of his insights forming the basis for theories of how turbulent motion impacts on matter concentrations. In Taylor's theory, turbulent diffusion is not concerned with details of the molecular properties of the fluid. Rather, the properties of turbulent diffusion (e.g., the efficiency of the turbulent diffusion) depend just on the nature of the turbulent motion of fluid elements. In this way, turbulent diffusion as described by Taylor is a phenomena that sits fully in the realm of continuum mechanics.

²For our purposes, turbulence is characterized by a quasi-random motion of fluid elements.

Correspondingly, each type of turbulent motion gives rise to a distinct form of turbulent diffusion. For example, in a geophysical context, turbulent diffusion associated with the breaking of internal gravity waves is distinct from turbulent diffusion by geostrophic eddies.

37.2.3 Fick's law for matter diffusion

Consider a fluid with a non-uniform matter concentration such as that drawn for a one-dimensional case in Figure 37.1. Random motion, due either to molecular motion or turbulent fluctuations, will transfer matter across an arbitrary point, line, or plane. Random motion preferentially moves matter 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 matter is linearly proportional to the concentration gradient, and thus can be written in the form

$$\mathbf{J} = -\kappa \rho \nabla C. \quad (37.3)$$

In this equation, we introduced the positive proportionality factor $\kappa > 0$, known as the *kinematic diffusivity*, whereas the product $\kappa \rho$ is known as the *dynamic diffusivity*

$$\kappa \quad \text{kinematic diffusivity with SI units } \text{m s}^{-2} \quad (37.4)$$

$$\rho \kappa \quad \text{dynamic diffusivity with SI units } \text{kg m}^{-2} \text{ s}^{-2}. \quad (37.5)$$

The kinematic diffusivity has dimensions of squared length per time and it sets the efficiency or strength of the diffusion. The diffusive flux (37.3) is known as Fick's law of matter diffusion. It is the most common mathematical form used to represent the mixing of matter through diffusion. Note that the minus sign in the diffusive flux arises since the flux is directed down the concentration gradient.

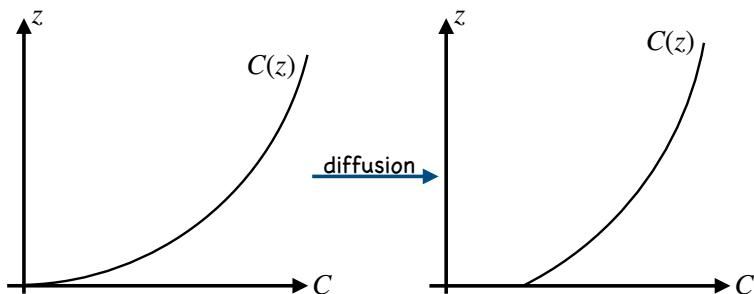


Figure 37.1: Shown here is a line graph illustrating the concentration, C , of a tracer drawn as a function of the space coordinate z , with the left panel showing the concentration at an earlier time than the right panel. Across any arbitrary point, transport of matter through random motions generally reduces the magnitude of the concentration gradient; i.e., the diffusive transport is down the concentration gradient. For example, where the concentration is relatively high, random motion mixes this high concentration with adjacent lower concentration, 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/deeper regions and low concentration fluid into higher/shallow regions. The concentration is uniform in equilibrium, leading to a flat concentration profile.

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_{mfp} (see Section 2.2.3), and the root-mean-square molecular speed, v_{rms} (see Section 2.2.4). Each of these properties is a function of the molecules comprising the matter. For air, the mean free path is

roughly 2×10^{-7} m and the RMS speed is 500 m s $^{-1}$, so that $L_{\text{mfp}} v_{\text{rms}} \approx 10^{-4}$ m 2 s $^{-1}$. The precise value for the molecular diffusivity depends on the molecular properties of the matter diffusing through air. For turbulent diffusion, Prandtl suggested that we consider a characteristic length and velocity scale determined by properties of the turbulent flow. The turbulent length scale (also called the *mixing length*) is generally much larger than the molecular mean free path, whereas the turbulent velocity scale is much smaller than molecular speeds. Determination of these turbulent length and velocity scales is subject to large uncertainties and is the topic of much research.

In regions where the diffusive flux is not a constant, there will be a net transport of matter that leads to the reduction of the tracer concentration gradient. At a particular point in space, the concentration changes in time according to the convergence of the diffusive flux

$$\rho \frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{J} = \nabla \cdot (\kappa \rho \nabla C). \quad (37.6)$$

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

$$\frac{\partial C}{\partial t} = \rho^{-1} \nabla(\kappa \rho) \cdot \nabla C + \kappa \nabla^2 C. \quad (37.7)$$

The first term is nonzero in regions where the dynamic diffusivity, $\kappa \rho$, spatially varies. The second term is nonzero in regions where the curvature of the concentration is nonzero. Correspondingly, when the tracer concentration is uniform in space then both terms vanish, whereas the Laplacian term also vanishes when the concentration is linear in space.

37.2.4 Fourier's law for heat diffusion

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, with the diffusive flux given by

$$\mathbf{J} = -\gamma \rho \nabla T, \quad (37.8)$$

where $\gamma > 0$ is the temperature 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 matter diffusivity. In general, matter diffuses by molecular processes slower than heat, so that the matter molecular diffusivity is smaller than the heat molecular diffusivity. In contrast, the turbulent thermal diffusivity is roughly the same as the matter diffusivity, since the turbulent diffusion of matter and heat are both mediated by the same turbulent fluctuations of fluid elements.

37.2.5 Newtonian frictional stress and momentum diffusion

In the same way that matter concentration and temperature gradients lead to diffusion by random 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 37.2, with shear (i.e., nonzero velocity gradient) in a single direction, Newtonian frictional stress takes the form

$$\tau = \rho \mu \frac{\partial u}{\partial z}, \quad (37.9)$$

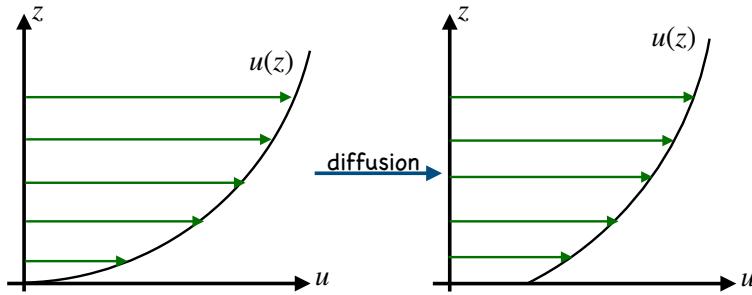


Figure 37.2: Shown here is a line graph illustrating the velocity, u , as a function of the space coordinate z , with the left panel showing the velocity at an earlier time than the right panel. Across any arbitrary point, transport of momentum through random motions generally reduces the magnitude of the velocity gradient; i.e., the diffusive transport leads to a viscous stress that acts to reduce the velocity shear.

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 when exploring the fluid dynamical equations in Chapter 22 and the nature of stress in Chapter 26.

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 is proportional to the scalar diffusivity. In general, the non-dimensional ratio of the viscosity to the diffusivity is known as the *Prandtl* number

$$\text{Pr} = \frac{\mu}{\kappa}. \quad (37.10)$$

Theories for the turbulent Prandtl number are largely empirical in nature, with first principles arguments elusive.

37.2.6 Further study

More thorough treatments of molecular diffusion for ideal gases can be found in books that describe the kinetic theory of gases, such as [Reif \(1965\)](#) and [Huang \(1987\)](#). The more terse treatment given in this section largely follows that from Section 1.5 of [Kundu et al. \(2012\)](#). A lucid treatment of Brownian motion in the context of turbulent diffusion is given by [Vallis \(2017\)](#).

37.3 Diffusion maths

We now explore various mathematical properties of the diffusion equation, here generalized to allow for distinct behavior in the different directions. Such distinctions are relevant especially in stratified fluids, where turbulent mixing across stratification surfaces is suppressed relative to turbulent mixing parallel to these surfaces (see Section 25.3). For this purpose, introduce the second order symmetric and positive definite diffusion tensor $\mathbb{K}_{mn} = \mathbb{K}_{nm}$. The resulting diffusive tracer flux takes the form

$$J_m = -\rho \mathbb{K}_{mn} \frac{\partial C}{\partial x^n}, \quad (37.11)$$

and the corresponding diffusion equation is

$$\rho \frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{J} = \frac{\partial(\rho \mathbb{K}_{mn} \partial_n C)}{\partial x_m}. \quad (37.12)$$

37.3.1 Sample diffusion tensors

For the isotropic case of molecular diffusion considered in equation (37.3), the diffusion tensor takes on the form

$$\mathbb{K}_{mn} = \kappa \delta_{mn} \quad \text{isotropic diffusion.} \quad (37.13)$$

If we rotate the diffusive fluxes to be along surfaces of constant γ , then

$$\mathbb{K}_{mn} = \kappa (\delta_{mn} - \hat{\gamma}_m \hat{\gamma}_n) \quad \text{rotated diffusion,} \quad (37.14)$$

where

$$\hat{\gamma} = \frac{\nabla \gamma}{|\nabla \gamma|} \quad (37.15)$$

is the normal to the surface. The most common case in oceanography is to set γ equal to a measure of the vertical stratification, in which case we have *neutral diffusion* (see Section 39.3).

37.3.2 Evolution of tracer concentration powers

For certain applications it is of interest to determine how diffusion acts on powers of the tracer concentration. For this purpose consider the material evolution of C^γ , where $\gamma \geq 1$

$$\rho \frac{DC^\gamma}{Dt} = \gamma C^{\gamma-1} \rho \frac{DC}{Dt} \quad (37.16a)$$

$$= -\gamma C^{\gamma-1} \nabla \cdot \mathbf{J} \quad (37.16b)$$

$$= -\gamma \nabla \cdot (C^{\gamma-1} \mathbf{J}) + \gamma (\gamma - 1) C^{\gamma-2} \nabla C \cdot \mathbf{J} \quad (37.16c)$$

$$= \partial_m \left[\rho \mathbb{K}_{mn} \frac{\partial C^\gamma}{\partial x^n} \right] - \rho \gamma (\gamma - 1) C^{\gamma-2} \frac{\partial C}{\partial x^n} \mathbb{K}_{mn} \frac{\partial C}{\partial x^m} \quad (37.16d)$$

$$= -\nabla \cdot \mathbf{J}(C^\gamma) + \gamma (\gamma - 1) C^{\gamma-2} \mathbf{J} \cdot \nabla C. \quad (37.16e)$$

The first term in equation (37.16e) is the convergence of the diffusive flux defined in terms of C^γ . It therefore acts to diffuse C^γ just like diffusion acts on C . The second term in equation (37.16e) is negative since the diffusion tensor is symmetric and positive-definite

$$\mathbf{J} \cdot \nabla C = -\rho \mathbb{K}_{mn} \frac{\partial C}{\partial x^m} \frac{\partial C}{\partial x^n} < 0. \quad (37.17)$$

That is, the diffusive flux corresponding to any positive definite and symmetric tensor is oriented down the tracer concentration gradient. Consequently, the second term in equation (37.16e) always acts to reduce the magnitude of C^γ towards zero. We make use of this result in Section 37.3.3 when considering moments of tracer concentration.

37.3.3 Moments of tracer concentration

Consider the evolution of domain integrated tracer concentration and its powers, and focus on impacts just from diffusion. For that purpose, assume the boundaries are insulating so that $\mathbf{J} \cdot \hat{\mathbf{n}} = 0$

with $\hat{\mathbf{n}}$ the outward normal at the boundary. Also assume the total fluid mass in the domain remains fixed

$$m = \int \rho dV \iff \frac{dm}{dt} = 0. \quad (37.18)$$

We can thus treat the domain as material given that there is no exchange of mass or tracer across the boundaries.

Domain average tracer concentration

The domain averaged tracer concentration is given by

$$\bar{C} = \frac{\int C \rho dV}{m}, \quad (37.19)$$

and it follows that its time derivative vanishes since

$$m \frac{d\bar{C}}{dt} = \frac{d}{dt} \int C \rho dV = \int \frac{DC}{Dt} \rho dV = - \int \nabla \cdot \mathbf{J} \rho dV = - \oint \mathbf{J} \cdot \hat{\mathbf{n}} d\mathcal{S} = 0, \quad (37.20)$$

where we set $\mathbf{J} \cdot \hat{\mathbf{n}} = 0$ to reach the last equality. 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 Reynold's transport theorem from Section 18.3.5. The result (37.20) 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.

Variance of tracer within the domain

The variance of the tracer concentration is defined by

$$\text{var}(C) \equiv \frac{\int (C - \bar{C})^2 \rho dV}{m} = \bar{C^2} - \bar{C}^2 \geq 0. \quad (37.21)$$

The tracer variance measures the deviation of the tracer concentration relative to the domain mean concentration. Since the domain averaged remains fixed in time, the time change of the variance is given by

$$\frac{d[\text{var}(C)]}{dt} = \frac{d\bar{C^2}}{dt}. \quad (37.22)$$

Thus, it is common to refer to $\bar{C^2}$ as the tracer variance, though strictly speaking only their time derivatives are equal. Performing the time derivative, and again noting that the domain is material thus allowing us to use Reynolds transport theorem, renders

$$\frac{d\bar{C^2}}{dt} = \frac{d}{dt} \int C^2 \rho dV = 2 \int C \frac{DC}{Dt} \rho dV = -2 \int C \nabla \cdot \mathbf{J} \rho dV = 2 \int \nabla C \cdot \mathbf{J} dV. \quad (37.23)$$

The final equality again made use of the assumed 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 (37.17) that diffusive fluxes are oriented down the tracer gradient. Consequently, diffusion of the tracer concentration results in a reduction in tracer variance

$$\frac{d[\text{var}(C)]}{dt} = \frac{d\bar{C^2}}{dt} \leq 0. \quad (37.24)$$

This result further supports our preconceived notions of diffusion as a process that removes differences (i.e., gradients) within the tracer field.

Evolution of arbitrary moments

Proceeding as before, and dropping boundary contributions, it is straightforward to show that the time derivative of an arbitrary tracer moment is given by

$$\frac{d\overline{C^\gamma}}{dt} = \gamma(\gamma - 1) \int C^{\gamma-2} \nabla C \cdot \mathbf{J} dV \leq 0. \quad (37.25)$$

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 (37.24). The result for higher powers holds so long as the tracer concentration is non-negative, which is always the case for material tracers. 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.

37.3.4 Connecting tracer variance to the diffusion operator

The diffusion operator is a linear self-adjoint operator. Consequently, it has an associated negative semidefinite functional (e.g., [Courant and Hilbert, 1953, 1962](#)). For example, the Laplacian operator $\nabla^2 C$ is identified with the functional derivative $\nabla^2 C = \delta\mathcal{F}/\delta C$, where

$$\mathcal{F} \equiv -(1/2) \int |\nabla C|^2 \rho d^3x \quad (37.26)$$

is the associated functional. In this subsection we prove this result for a general diffusion tensor K^{mn} acting on an arbitrary passive tracer concentration, C . As detailed by [Griffies et al. \(1998\)](#) and Chapter 16 of [Griffies \(2004\)](#), the connection between the diffusion operator and the functional derivative of the diffusion dissipation provides a useful framework for deriving numerical discretizations of the diffusion operator.

For this subsection it is useful to write

$$d^3x = dV \quad (37.27)$$

for the volume element. The reason will become apparent at equation (37.34).

Derivative of the diffusion dissipation functional

We introduce the *diffusion dissipation functional*

$$\mathcal{F} = \int \mathcal{L} d^3x \quad (37.28)$$

where the integrand is the negative semi-definite quadratic form

$$2\mathcal{L} = \mathbf{J} \cdot \nabla C = -\rho \partial_m C \mathbb{K}_{mn} \partial_n C \leq 0. \quad (37.29)$$

Our goal is to relate the diffusion operator, given by the convergence of the diffusion flux, $\mathcal{R} = -\nabla \cdot \mathbf{J}$, to the functional derivative of \mathcal{F} . To compute the functional derivative requires us to insert variations³ to the tracer field δC into the dissipation functional

$$\delta\mathcal{F} = \int \left[\delta C \frac{\delta\mathcal{L}}{\delta C} + \delta(\partial_m C) \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right] d^3x. \quad (37.30)$$

³Functional *variations* are small perturbations to the functional form.

Integration by parts on the second term leads to

$$\delta\mathcal{F} = \int \left[\delta C \frac{\delta\mathcal{L}}{\delta C} + \partial_m \left(\delta C \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) - \delta C \partial_m \left(\frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) \right] d^3x. \quad (37.31)$$

The middle term is a total derivative that integrates to a boundary contribution and the associated *natural boundary condition*

$$\hat{\mathbf{n}} \cdot \frac{\delta\mathcal{L}}{\delta \nabla C} = \hat{\mathbf{n}} \cdot \mathbf{J} = \text{boundary flux}, \quad (37.32)$$

with $\hat{\mathbf{n}}$ the boundary outward normal. 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

$$\delta\mathcal{F} = \int \delta C \left[\frac{\delta\mathcal{L}}{\delta C} - \partial_m \left(\frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) \right] d^3x. \quad (37.33)$$

Consequently, the functional derivative is given by

$$(d^3y)^{-1} \frac{\delta\mathcal{F}}{\delta C(\mathbf{y})} = \frac{\delta\mathcal{L}}{\delta C} - \partial_m \left[\frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right], \quad (37.34)$$

where d^3y is the volume element at the field point \mathbf{y} . To reach the last step required the identity

$$\frac{\delta C(\mathbf{x})}{\delta C(\mathbf{y})} = d^3y \delta(\mathbf{x} - \mathbf{y}), \quad (37.35)$$

where $\delta(\mathbf{x} - \mathbf{y})$ is the Dirac delta function satisfying

$$\int \delta(\mathbf{x} - \mathbf{y}) d^3y = 1, \quad (37.36)$$

so long as the integration range includes the singular point $\mathbf{x} = \mathbf{y}$. Note that the delta function has dimensions of inverse volume, which necessitates the appearance of the volume factor, d^3y , on the right hand side of equation (37.35).

Connection to the diffusion operator

Reintroducing the specific form of the diffusion integrand $2\mathcal{L} = -\rho \partial_m C \mathbb{K}_{mn} \partial_n C$ leads to

$$(d^3y)^{-1} \frac{\delta\mathcal{F}}{\delta C(\mathbf{y})} = -\partial_m \left[\frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right] = \partial_m (\rho \mathbb{K}_{mn} \partial_n C). \quad (37.37)$$

The second equality identifies the diffusion operator, thus revealing the connection between the dissipation functional, the diffusion fluxes, and the diffusion operator

$$(d^3y)^{-1} \frac{\delta\mathcal{F}}{\delta C(\mathbf{y})} = -\nabla \cdot \mathbf{J}(\mathbf{y}) = \mathcal{R}. \quad (37.38)$$

37.4 Advection physics

A *perfect* or *ideal* fluid is comprised of material fluid elements whose matter content and thermodynamic properties remain fixed. From the discussion of molecular diffusion in Section 37.2, 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 among fluid elements. This matter 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 any mixing.

37.4.1 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 with the fluid. 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*

$$\frac{DC}{Dt} = \frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C = 0. \quad (37.39)$$

The first equality relates the material time derivative to the Eulerian time derivative plus advective transport (see Section 16.4), with \mathbf{v} the velocity of a fluid element. We can convert the *material* form of the advection equation (37.39) into its flux-form by combining with the mass continuity equation (17.9)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (37.40)$$

which yields

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v}) = 0. \quad (37.41)$$

Again, the material form of the advection equation is the trivial statement that tracer concentration remains constant on a fluid element in the absence of sources or mixing. Hence, a general solution to the advection equation is given by

$$C(\mathbf{x}, t) = C[\mathbf{X}(0)], \quad (37.42)$$

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, we know the tracer concentration for all space and time. For those cases where trajectories are unknown, it is useful to make use of the Eulerian form of the advection equation in order to deduce the evolution of tracer.

37.4.2 Eulerian time tendencies from advection

At a point in the fluid, the advection equation (37.39) leads to the Eulerian time tendency for tracer concentration

$$\frac{\partial C}{\partial t} = -\mathbf{v} \cdot \nabla C. \quad (37.43)$$

Geometrically, the tendency arises from the projection of the fluid velocity onto the normal to concentration iso-surfaces. The concentration remains fixed at points where the velocity is parallel to concentration iso-surfaces. From the flux-form advection equation (37.41), the density-weighted tracer concentration (the tracer mass per volume) has an Eulerian time tendency given by the convergence of the advective flux

$$\frac{\partial(\rho C)}{\partial t} = -\nabla \cdot (\rho C \mathbf{v}). \quad (37.44)$$

The tendency vanishes at a point if there is no convergence of tracer mass towards the point.

37.5 Advection maths

We now explore various mathematical properties of the advection equation. For that purpose, recall the mass continuity equation (37.40) and flux-form tracer advection equation (37.44)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (37.45a)$$

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v}) = 0. \quad (37.45b)$$

These equations are manifestly compatible in that the tracer equation (37.45b) reduces to the continuity equation (37.45a) if the tracer concentration is spatially uniform (see Section 18.1.5 for more discussion of compatibility).

37.5.1 Material constancy of C^γ

A trivial consequence of the material constancy of tracer concentration is that C^γ is also materially constant, for γ an arbitrary number. We show this property mathematically by noting that the chain rule holds for a material time derivative, so that

$$\frac{DC^\gamma}{Dt} = n C^{\gamma-1} \frac{DC}{Dt} = 0. \quad (37.46)$$

Likewise, making use of the Eulerian form yields

$$\frac{\partial C^\gamma}{\partial t} + \mathbf{v} \cdot \nabla C^\gamma = \gamma C^{\gamma-1} \left[\frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C \right] = 0. \quad (37.47)$$

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, which contrasts to the case of diffusion considered in Section 37.3.3.

37.5.2 Mass transport

The mass density time tendency

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\mathbf{v} \rho) \quad (37.48)$$

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

$$\rho \mathbf{v} \rightarrow \rho \mathbf{v}^\dagger = \rho \mathbf{v} + \nabla \wedge (\rho \boldsymbol{\Psi}). \quad (37.49)$$

As in Section 19.5.2, the arbitrariness manifest in equation (37.49) is known as a *gauge symmetry*. The additional mass flux, $\nabla \wedge (\rho \boldsymbol{\Psi})$, leads to no accumulation of mass at a point since it has zero divergence. In the Boussinesq case with ρ set to a constant ρ_0 , the divergent-free velocity $\nabla \wedge \boldsymbol{\Psi}$ leads to zero accumulation of volume at a point.

The non-divergent mass flux

$$\rho \mathbf{v}^* \equiv \nabla \wedge (\rho \boldsymbol{\Psi}) \quad (37.50)$$

often arises when we decompose the mass flux into a mean and a non-divergent fluctuation. In that context, we make use of the following terminology:

$$\mathbf{v} = \text{Eulerian mean velocity} \quad (37.51\text{a})$$

$$\rho \mathbf{v} = \text{Eulerian mean mass flux} \quad (37.51\text{b})$$

$$\mathbf{v}^* = \text{eddy-induced velocity} \quad (37.51\text{c})$$

$$\rho \mathbf{v}^* = \nabla \wedge (\rho \Psi) = \text{eddy-induced mass flux} \quad (37.51\text{d})$$

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^* = \text{residual mean velocity} \quad (37.51\text{e})$$

$$\rho \mathbf{v}^\dagger = \rho (\mathbf{v} + \mathbf{v}^*) = \text{residual mean mass flux.} \quad (37.51\text{f})$$

The name “residual mean” is motivated since the sum $\mathbf{v} + \mathbf{v}^*$ 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.

37.5.3 Advection tracer fluxes and skew tracer fluxes

Following from the previous discussion, we consider the tracer advection equation with the advective tracer transport determined by the residual mean velocity

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v}^\dagger) = 0. \quad (37.52)$$

Given the form (37.50) for the eddy mass flux $\rho \mathbf{v}^*$, we can write the advective tracer flux as

$$\rho C \mathbf{v}^\dagger = C (\rho \mathbf{v} + \rho \mathbf{v}^*) \quad (37.53\text{a})$$

$$= C \rho \mathbf{v} + C \nabla \wedge (\rho \Psi) \quad (37.53\text{b})$$

$$= C \rho \mathbf{v} + \nabla \wedge (C \rho \Psi) - \nabla C \wedge \rho \Psi. \quad (37.53\text{c})$$

It is the divergence of the tracer flux that determines the time tendency, in which case the total curl plays no role

$$\nabla \cdot (\rho C \mathbf{v}^\dagger) = \nabla \cdot (\rho C \mathbf{v} + \rho C \mathbf{v}^*) \quad (37.54\text{a})$$

$$= \nabla \cdot (\rho C \mathbf{v} - \nabla C \wedge \rho \Psi). \quad (37.54\text{b})$$

That is, the divergence of the advective mass flux equals to the divergence of the skew tracer flux

$$\underbrace{\nabla \cdot (\rho C \mathbf{v}^*)}_{\text{advective flux divergence}} = \underbrace{\nabla \cdot (-\nabla C \wedge \rho \Psi)}_{\text{skew flux divergence}} \quad (37.55)$$

since the advective flux and skew flux differ by a rotational flux

$$\mathbf{J}^{\text{adv}} = \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}} \quad (37.56)$$

where

$$\mathbf{J}^{\text{adv}} = C \rho \mathbf{v}^* \quad \mathbf{J}^{\text{skew}} = -\nabla C \wedge \rho \Psi \quad \mathbf{J}^{\text{rot}} = \nabla \wedge (\rho C \Psi). \quad (37.57)$$

Notably, the skew tracer flux is neither upgradient nor downgradient. Rather, it is oriented parallel to iso-surfaces of tracer concentration

$$\nabla C \cdot \mathbf{J}^{\text{skew}} = \nabla C \cdot (-\nabla C \wedge \rho \Psi) = 0. \quad (37.58)$$

This orientation serves as motivation for the name “skew.” Figure 37.3 provides a schematic of the skew tracer fluxes.

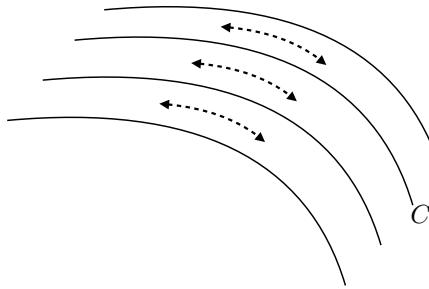


Figure 37.3: Skew fluxes (dashed lines with arrows) for a tracer C are oriented parallel to lines of constant tracer concentration (tracer isolines are the solid lines).

37.5.4 Skew diffusion

Introducing tensor labels brings the skew tracer flux into the form

$$J_m^{\text{skew}} = -(\nabla C \wedge \rho \Psi)_m \quad (37.59\text{a})$$

$$= -\epsilon_{mnp} \frac{\partial C}{\partial x^n} \rho \Psi_p \quad (37.59\text{b})$$

$$= -\rho \mathbb{A}_{mn} \frac{\partial C}{\partial x^n}, \quad (37.59\text{c})$$

where we defined the anti-symmetric *skew diffusion* tensor

$$\mathbb{A}_{mn} = \epsilon_{mnp} \Psi_p = \begin{bmatrix} 0 & \Psi_3 & -\Psi_2 \\ -\Psi_3 & 0 & \Psi_1 \\ \Psi_2 & -\Psi_1 & 0 \end{bmatrix}. \quad (37.60)$$

We thus conclude that advection by a non-divergent mass flux is equivalent to skew-diffusion through the action of an anti-symmetric tensor.

Although leading to the same stirring operator, skew and advective fluxes possess rather complementary properties as listed here.

- **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. That is, the fluxes in effect swap the placement of the derivative operator. Correspondingly, the advective flux vanishes only if the velocity vanishes, whereas the skew flux vanishes when the tracer gradient vanishes (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.
- **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), whereas a skew flux is not interpreted as the passage of matter across an area per time. This distinction is particularly important when interpreting boundary conditions discussed in Section 37.6.2.

In Section 37.6 we pursue the above points to further detail the dual relation between advection and skewson.

37.5.5 Further study

Ideas 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 fluids with rotation and/or magnetic fields. [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.

37.6 Advection and skewson

We introduced skew diffusion in Section 37.5.4 and will again encounter it in Chapter 38. Following the terminology of Section 9.2 of [Griffies \(2004\)](#), we refer to *skewson* as any process that leads to tracer stirring via skew fluxes, with skew diffusion a particular example. There are occasions where it is conceptually more convenient to use advective fluxes, such as when considering the stirring of tracers by the flow field explicitly resolved by a numerical simulation. In contrast, skew fluxes are often more convenient for certain subgrid scale eddy parameterizations, such as the one discussed in Section 39.1. We here consider facets of advection and skewson for those interested in diving deeper into the details.

37.6.1 Choosing a gauge

Consider an arbitrary divergent-free mass transport

$$\nabla \cdot (\rho \mathbf{v}^*) = 0, \quad (37.61)$$

where the divergent-free constraint is satisfied by introducing a vector streamfunction

$$\rho \mathbf{v}^* = \nabla \wedge (\rho \Psi). \quad (37.62)$$

The streamfunction is arbitrary up to a gauge transformation

$$\rho \Psi' = \rho \Psi + \nabla(\rho \Lambda), \quad (37.63)$$

where Λ 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, $\mathbf{J}^{\text{skew}} = -\nabla C \wedge \rho \Psi$, changes. Nonetheless, the divergence of the skew flux is invariant, as we see by noting that

$$\nabla C \wedge [\rho \Psi + \nabla(\rho \Lambda)] = \nabla C \wedge (\rho \Psi) + \nabla \wedge [C \nabla(\rho \Lambda)]. \quad (37.64)$$

and since $\nabla \cdot \nabla \wedge [C \nabla(\rho \Lambda)] = 0$, the flux divergence, $\nabla \cdot \mathbf{J}^{\text{skew}}$, remains unchanged.

Coulomb gauge

One has a lot of freedom in specifying the gauge function. The simplest case 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. A gauge commonly used in electrodynamics (e.g., [Jackson \(1975\)](#)) is the *Coulomb gauge*, in which

$$\nabla \cdot \Psi = 0 \quad \text{Coulomb gauge.} \quad (37.65)$$

Making use of the curl identity (4.36c) leads to the Poisson equation for the vector potential

$$\nabla^2(\rho\Psi) = -\nabla \wedge (\rho \mathbf{v}^*). \quad (37.66)$$

This equation has a free-space Green's function given by the Coulomb-Ampere expression

$$\rho(\mathbf{x}, t)\Psi(\mathbf{x}, t) = \nabla \wedge \int \frac{\rho(\mathbf{x}', t)\mathbf{v}^*(\mathbf{x}', t)}{4\pi|\mathbf{x} - \mathbf{x}'|} dV', \quad (37.67)$$

where dV' is the volume element for integration over the test points, \mathbf{x}' . Although common in electrodynamics, we know of no geophysical fluid application making use of the Coulomb gauge.

Vertical gauge

As introduced in Section 19.5.3, one gauge used for eddy parameterizations (Section 39.1) 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 mass transport field. A common choice is the *vertical gauge* in which

$$\Psi_3 = 0 \quad \text{vertical gauge.} \quad (37.68)$$

Let us see how we can generally make this gauge choice. Consider a vector streamfunction Φ that has all three components nonzero. Now consider the alternative streamfunction $\rho\Psi = \rho\Phi + \nabla(\rho\Lambda)$, with $\nabla(\rho\Lambda) = -\hat{\mathbf{z}}\rho\Phi_3$. This choice in turn means that the third component of Ψ is zero.

To further specify the vertical gauge we invert the relations

$$\rho u^* = -\partial_z(\rho\Psi_2) \quad \rho v^* = \partial_z(\rho\Psi_1) \quad \rho w^* = \partial_x(\rho\Psi_2) - \partial_y(\rho\Psi_1) \quad (37.69)$$

to render the vector streamfunction

$$\rho\Psi = \hat{\mathbf{z}} \wedge \int_{-H}^z \rho \mathbf{u}^* dz' = \hat{\mathbf{z}} \wedge \underline{\mathbf{U}}^{(*\rho)} \quad (37.70)$$

where

$$\underline{\mathbf{U}}^{(*\rho)}(z) = \int_{-H}^z \rho \mathbf{u}^* dz' \quad (37.71)$$

is the horizontal mass transport associated with \mathbf{u}^* passing between the bottom and a depth $z \geq -H$. The anti-symmetric stirring tensor for the vertical gauge is given by

$$\rho \mathbb{A}_{mn} = \begin{pmatrix} 0 & 0 & \underline{U}^{(*\rho)} \\ 0 & 0 & \underline{V}^{(*\rho)} \\ -\underline{U}^{(*\rho)} & -\underline{V}^{(*\rho)} & 0 \end{pmatrix}, \quad (37.72)$$

and the corresponding skew, rotational, and advective fluxes are

$$\mathbf{J}^{\text{skew}} = -\underline{\mathbf{U}}^{(*\rho)} \partial_z C + \hat{\mathbf{z}} \underline{\mathbf{U}}^{(*\rho)} \cdot \nabla_z C \quad (37.73a)$$

$$\mathbf{J}^{\text{rot}} = \partial_z(C \underline{\mathbf{U}}^{(*\rho)}) - \hat{\mathbf{z}} \nabla_z \cdot (C \underline{\mathbf{U}}^{(*\rho)}) \quad (37.73b)$$

$$\mathbf{J}^{\text{adv}} = C(\partial_z \underline{\mathbf{U}}^{(*\rho)}) - \hat{\mathbf{z}} C \nabla_z \cdot \underline{\mathbf{U}}^{(*\rho)}. \quad (37.73c)$$

Note that the identity $\mathbf{J}^{\text{adv}} = \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}$ is maintained by 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.

37.6.2 Boundary conditions

We assume that all domain boundaries are material in regards to the velocity \mathbf{v}^* . Furthermore, even for moving boundaries, we assume that the suite of kinematic boundary conditions is based on the barycentric velocity \mathbf{v} (see Section 17.4), so that \mathbf{v}^* satisfies the no-normal flow condition even on moving boundaries

$$\hat{\mathbf{n}} \cdot \mathbf{v}^* = 0. \quad (37.74)$$

Hence, the advective flux for tracers also has a no-normal boundary condition on all boundaries

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{adv}} = \hat{\mathbf{n}} \cdot \mathbf{v}^* \rho C = 0. \quad (37.75)$$

The corresponding boundary condition for the skew flux is found by inserting the relation (37.56) into the advective flux boundary condition (37.75) to render

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{adv}} = \hat{\mathbf{n}} \cdot [\mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}] = 0. \quad (37.76)$$

Hence, the skew flux generally has a non-zero normal component at the solid boundaries as determined by the rotational flux

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{skew}} = -\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{rot}}. \quad (37.77)$$

Even so, we may have occasions in which $\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{skew}} = 0$, which is ensured so long as

$$(-\nabla C \wedge \rho \Psi) \cdot \hat{\mathbf{n}} = -(\rho \Psi \wedge \hat{\mathbf{n}}) \cdot \nabla C = 0. \quad (37.78)$$

A sufficient condition is to have $\Psi \wedge \hat{\mathbf{n}} = 0$, 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 39.1 as prescribed by the [Gent et al. \(1995\)](#) eddy parameterization scheme.

37.7 Active tracers and dia-surface flow

An *active tracer* is a tracer that impacts the dynamics, with potential temperature the canonical example. Potential temperature directly impacts the buoyancy (e.g., we commonly assume buoyancy is linearly proportional to potential temperature), which in turn affects pressure and the velocity. Hence, the advection-diffusion equation for active tracers is nonlinear since the velocity field is dependent on active tracers

$$\rho \frac{D\theta}{Dt} = \rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \theta = -\nabla \cdot \mathbf{J}(\theta). \quad (37.79)$$

37.7.1 Adiabatic flow

In an adiabatic fluid, potential temperature is materially invariant

$$\frac{D\theta}{Dt} = \frac{\partial \theta}{\partial t} + \mathbf{v} \cdot \nabla \theta = 0 \quad \text{adiabatic.} \quad (37.80)$$

Furthermore, following the kinematics of Section 17.4.2, the adiabatic flow field does not penetrate surfaces of constant potential temperature (θ -isosurfaces are impermeable) since

$$\mathbf{v} \cdot \nabla \theta = -\frac{\partial \theta}{\partial t}. \quad (37.81)$$

In this case we say that advection reversibly stirs the potential temperature field.

37.7.2 Diabatic processes generating dia-surface transport

Potential temperature is no longer materially invariant in the presence of diabatic processes such as mixing. Correspondingly, the velocity picks up a component, w^{dia} , that crosses the moving potential temperature surface thus making it permeable. We follow the kinematics from Section 21.3 to render the expression (21.24) for w^{dia} , which we here write as

$$w^{\text{dia}} \equiv \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\theta)}) = \frac{1}{|\nabla\theta|} \frac{D\theta}{Dt} \quad (37.82)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla\theta}{|\nabla\theta|} \quad \text{and} \quad \mathbf{v}^{(\theta)} \cdot \nabla\theta = -\frac{\partial\theta}{\partial t}. \quad (37.83)$$

Rearrangement of equation (37.82) renders the kinematic identity

$$\frac{D\theta}{Dt} = \frac{\partial\theta}{\partial t} + \mathbf{v} \cdot \nabla\theta = w^{\text{dia}} |\nabla\theta|. \quad (37.84)$$

We offer examples to illustrate the physics underlying this identity.

Diffusion with no fluid motion

Diffusion is the canonical example of a diabatic process (Section 37.2)

$$\rho \frac{D\theta}{Dt} = \nabla \cdot (\kappa \rho \nabla\theta), \quad (37.85)$$

with $\kappa > 0$ the scalar kinematic diffusivity. Diffusion in turn drives a diabatic transport velocity

$$\rho w^{\text{dia}} = \frac{\nabla \cdot (\kappa \rho \nabla\theta)}{|\nabla\theta|}. \quad (37.86)$$

Consider a horizontally homogeneous potential temperature field. If buoyancy is alone determined by potential temperature then there is no fluid motion since buoyancy surfaces are flat. Yet in the presence of vertical diffusion and vertical stratification there is a diabatic transport velocity

$$\rho w^{\text{dia}} = \frac{\partial_z(\kappa \rho \partial_z\theta)}{|\partial_z\theta|}. \quad (37.87)$$

In the absence of fluid flow, this transport velocity is determined solely by movement of the potential temperature surfaces. Correspondingly, potential temperature evolution is determined only by vertical diffusion (there is no advection since $\mathbf{v} = 0$)

$$\frac{\partial\theta}{\partial t} = \rho^{-1} \frac{\partial}{\partial z} \left[\kappa \rho \frac{\partial\theta}{\partial z} \right] = w^{\text{dia}} \left| \frac{\partial\theta}{\partial z} \right|. \quad (37.88)$$

Steady state advective-diabatic balance

A steady state in the presence of diabatic processes can only be realized via a balance between advective transport and diabatic transport

$$\mathbf{v} \cdot \nabla\theta = w^{\text{dia}} |\nabla\theta| \quad \text{steady state.} \quad (37.89)$$

The advective transport via fluid motion crossing θ surfaces (left hand side) is exactly balanced by diabatic transport through diffusion or other diabatic processes (right hand side), with this balance required to maintain static θ -surfaces ($\partial_t\theta = 0$). The particular case of a diffusion in a vertically stratified horizontally homogeneous potential temperature field results in the vertical advective-diffusive balance

$$w \partial_z\theta = w^{\text{dia}} |\partial_z\theta| = \rho^{-1} \partial_z(\kappa \rho \partial_z\theta). \quad (37.90)$$

37.8 Exercises

EXERCISE 37.1: 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 from scaling the diffusion equation, which reveals that it takes the form

$$\tau_{\text{diffusion}} = \frac{\Delta^2}{\kappa} \quad (37.91)$$

where Δ is the length scale and κ 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 37.2: ANALYTICAL SOLUTION TO ONE-DIMENSIONAL DIFFUSION EQUATION

Consider a one-dimensional diffusion equation

$$\frac{\partial C}{\partial t} = \kappa \frac{\partial^2 C}{\partial z^2}, \quad (37.92)$$

where C is a tracer concentration (e.g., temperature or salinity), κ 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

$$\frac{\partial C}{\partial t} = \kappa \frac{\partial^2 C}{\partial z^2}, \quad (37.93a)$$

$$\text{no-flux boundary condition from part (b)} \quad (37.93b)$$

$$C(z, t = 0) = C_0 \cos(Kz), \quad (37.93c)$$

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 37.3: 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 fluid in which case the density factors are all constant and so can be dropped

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial z} \left[\kappa \frac{\partial C}{\partial z} \right] \quad (37.94a)$$

$$= \frac{\partial \kappa}{\partial z} \frac{\partial C}{\partial z} + \kappa \frac{\partial^2 C}{\partial z^2}, \quad (37.94b)$$

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

$$\frac{\partial C^*}{\partial t} = \kappa \frac{\partial^2 C^*}{\partial z^2}. \quad (37.95)$$

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)?

(b) THREE-DIMENSIONAL DIFFUSION

Generalize the above one dimensional result to three dimensions, whereby the diffusivity κ becomes a symmetric positive-definite diffusion *tensor*, in which case

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x^m} \left[\kappa^{mn} \frac{\partial C}{\partial x^n} \right]. \quad (37.96)$$

Now consider an extrema in the tracer field, which is defined by

$$\frac{\partial C^*}{\partial x^n} = 0 \quad \forall n = 1, 2, 3. \quad (37.97)$$

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 37.4: ONE-DIMENSIONAL ADVECTION

Consider the advection equation in one space dimension without boundaries

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = 0 \quad (37.98a)$$

$$C(x, z, t = 0) = C_0 \cos(k x) \quad (37.98b)$$

$$u(z, t) = \alpha z \cos(\omega t). \quad (37.98c)$$

The specified zonal velocity is non-divergent, oscillatory in time, and vertically sheared

$$\frac{\partial u}{\partial z} = \alpha \cos(\omega t), \quad (37.99)$$

with ω the radial frequency of the temporal oscillations. What is the tracer concentration at times $t > 0$? Hint: make use of the exact solution given by equation (37.42).

EXERCISE 37.5: SKEW FLUX FOR OCEAN MESOSCALE EDDIES

Consider a middle-latitude mesoscale ocean eddy respecting geostrophic balance (see Section 29.3) on an f -plane (constant Coriolis parameter) and incompressibility. In this case, the horizontal eddy-induced velocity at the ocean surface is non-divergent

$$\mathbf{u}^* = \nabla \wedge \hat{\mathbf{z}} \psi. \quad (37.100)$$

In this equation, the geostrophic streamfunction is given by

$$\psi = -\hat{\mathbf{z}} \frac{g \eta}{f}, \quad (37.101)$$

with f the Coriolis parameter, g the gravitational acceleration, and η 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, ρ_0 .

- (a) Determine the skew diffusion tensor (37.60).
- (b) Determine the skew tracer flux (37.73a).

EXERCISE 37.6: 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.

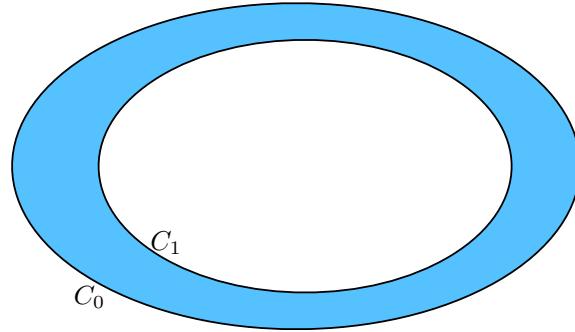


Figure 37.4: Illustrating the area contained between two closed tracer contours, $C_0 \leq C(x, y, t) \leq C_1$. Exercise 37.6 develops some mathematical expressions for integration within this area, with the resulting expressions of use for the analyses of tracer transport.

- (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 37.4. Derive the following expression for the area enclosed by the two contours

$$\mathcal{A} = \int_{C_0}^{C_1} dC \oint \frac{dl}{|\nabla C|}. \quad (37.102)$$

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

$$\mathcal{A}(C) = \int_{C_0}^C dC' \oint \frac{dl}{|\nabla C'|} \quad (37.103)$$

we have the identity

$$\frac{\partial \mathcal{A}(C)}{\partial C} = \oint \frac{dl}{|\nabla C|}. \quad (37.104)$$

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

$$\frac{\partial}{\partial C} \left[\int \Phi(\mathbf{x}) d\mathcal{A} \right] = \frac{\partial}{\partial C} \left[\int_{C_0}^C dC' \oint \frac{\Phi dl}{|\nabla C'|} \right] \quad (37.105a)$$

$$= \oint \frac{\Phi dl}{|\nabla C|}, \quad (37.105b)$$

with Φ an arbitrary function. This is a remarkable identity with many useful applications such as those in [Marshall et al. \(2006\)](#).

EXERCISE 37.7: EVOLUTION OF TRACER CENTER OF MASS IN A STATIC DOMAIN

The exercise introduces us to how the tracer center of mass evolves within a Boussinesq fluid. We define the tracer center of mass as

$$\langle \mathbf{x} \rangle^C = \frac{\int \mathbf{x} C dV}{\int C dV}, \quad (37.106)$$

with C the tracer concentration, \mathbf{x} the coordinate of a point in the fluid, and integration 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

$$\mathcal{V} = \int dV \quad C = \int C dV. \quad (37.107)$$

Furthermore, when computing the time derivative acting on the integral, make use of the kinematic results from Section 18.3.3, in which for any integrand φ

$$\frac{d}{dt} \int \varphi dV = \int \frac{\partial \varphi}{\partial t} dV, \quad (37.108)$$

which follows since the region boundaries are assumed to be static.

- (a) Consider a tracer concentration whose tendency at a point in space is affected only by advection

$$\frac{\partial C}{\partial t} = -\nabla \cdot (\mathbf{v} C), \quad (37.109)$$

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

$$\frac{d\langle \mathbf{x} \rangle^C}{dt} = \langle \mathbf{v} \rangle^C, \quad (37.110)$$

where the tracer center of mass velocity is given by

$$\langle \mathbf{v} \rangle^C = \frac{\int \mathbf{v} C dV}{\int C dV} = \frac{1}{C} \int \mathbf{v} C dV. \quad (37.111)$$

- (b) Consider a tracer concentration whose tendency at a point in space affected only by diffusion

$$\frac{\partial C}{\partial t} = \nabla \cdot (K \nabla C), \quad (37.112)$$

where $K = K(\mathbf{x}) > 0$ is a diffusivity (dimensions of squared length per time) that is assumed to vanish at the domain boundaries. Show that the tracer center of mass drifts up the diffusivity gradient

$$\frac{d\langle \mathbf{x} \rangle^C}{dt} = \langle \nabla K \rangle^C. \quad (37.113)$$

- (c) Consider an initial tracer concentration that is a function only of latitude,

$$C(x, y, z, t = 0) = C_0(y), \quad (37.114)$$

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. Introduce an eddy stirring that breaks the zonal symmetry. Qualitatively discuss the process whereby this turbulent diffusion causes the tracer center of mass to drift towards the turbulent region.

EXERCISE 37.8: 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

$$\frac{DC}{Dt} = \dot{C} \neq 0. \quad (37.115)$$

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

$$\langle \mathbf{x} \rangle^C = \frac{\int_{\mathcal{R}(\mathbf{v})} \mathbf{x} C \rho dV}{\int_{\mathcal{R}(\mathbf{v})} C \rho dV}. \quad (37.116)$$

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 Section 18.3.5 since the region has a constant mass. Consequently, we can set

$$\frac{d}{dt} \int_{\mathcal{R}(\mathbf{v})} \psi \rho dV = \int_{\mathcal{R}(\mathbf{v})} \frac{D\psi}{Dt} \rho dV. \quad (37.117)$$

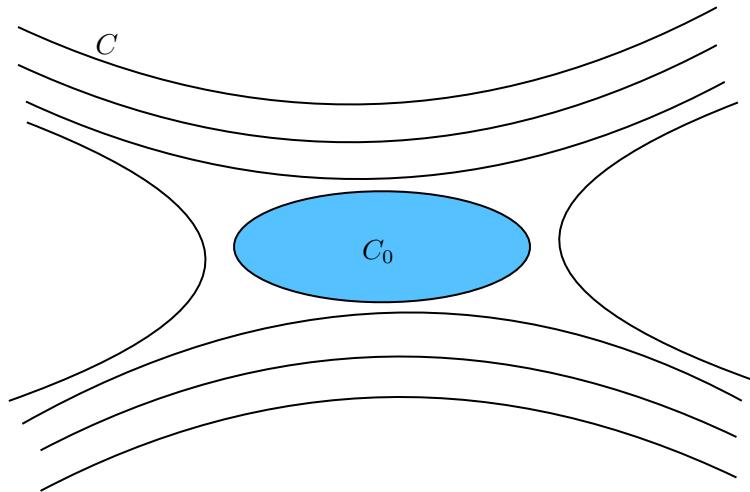


Figure 37.5: In a steady state flow, the tracer concentration within a constant C contour is constant. Diffusion has thus acted to remove all variations in tracer concentration within the region. In this figure, the concentration within the closed region has constant value $C = C_0$, whereas the region with open tracer contours remains non-homogeneous.

EXERCISE 37.9: DIFFUSIVE HOMOGENIZATION OF SCALARS INSIDE CLOSED CONTOURS

The advection-diffusion equation for a tracer concentration is given by

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho \mathbf{v} C) = -\nabla \cdot \mathbf{J} \quad (37.118)$$

where

$$\mathbf{J} = -\rho \mathbb{K} \cdot \nabla C \quad (37.119)$$

is a downgradient diffusive flux with \mathbb{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

$$\nabla \cdot (\rho \mathbf{v} C) = -\nabla \cdot \mathbf{J}. \quad (37.120)$$

Consider a two-dimensional steady state flow and consider a region enclosed by a constant tracer contour. Prove that the tracer concentration is homogeneous (i.e., a spatially constant) within the contour of constant C , as shown in Figure 37.5. Hence, 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.

Hint: make use of a *reductio ad absurdum* argument.

38

Tracer kinematics

Geophysical fluid flows exhibit 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 fluctuation relative to the mean. We perform an eddy-mean decomposition when interest concerns the mean field and impacts on the mean by the fluctuating instantaneous flow, with such impacts often termed *rectified* effects. The mean field can be defined in many fashions with subjective choices based on particulars of the flow and the analysis goals. The definition for the mean in turn affects what we refer to as the fluctuation. Quite generally, 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 originally motivated by the analysis of scalar transport induced by small amplitude wave-like eddying features, but is also of use for turbulent processes and their parameterizations (e.g., Chapter 39).

We consider two 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, 2014](#)). For our purposes it is sufficient to use only a small piece of the GLM framework to help unpack the kinematics of eddy tracer fluxes. The second kinematic method makes use isopycnal vertical coordinates. We connect an isopycnal description to the GLM by applying the GLM just in the vertical direction. In this sense the isopycnal approach is quasi-Lagrangian since it fixes the horizontal position (Eulerian) yet allows the vertical to follow an adiabatic fluid parcel (Lagrangian). The isopycnal approach is frequently used to help understand how ocean mesoscale eddies affect stratification and tracer transport in stably stratified flows. Our presentation follows the methods developed by [McDougall and McIntosh \(2001\)](#) and summarized in Chapter 9 of [Griffies \(2004\)](#).

READER'S GUIDE TO THIS CHAPTER

Material in this chapter relies on an understanding of the tracer equation as derived in Section 18.1 and the maths and physics of the advection-diffusion equation explored in Chapter 37. We focus most discussion on incompressible flows discussed in Chapter 19 and applicable to the Boussinesq fluid commonly assumed for the ocean (Chapter 28). Generalizations to compressible fluids are straightforward, with examples provided by [Griffies and Greatbatch \(2012\)](#). The kinematics of isopycnal fluid layers in a perfect fluid (Sections 38.4 and 38.6) are posed using the isopycnal vertical coordinates detailed in Chapter 11 and further pursued in Chapter 32.

38.1	Reynold's decomposition	570
38.2	Basic kinematics of the GLM	571
38.2.1	Motivation	572
38.2.2	Length scales and the small parameter	573
38.2.3	Decomposing the particle trajectory	573
38.2.4	GLM and the Stokes mean	574
38.2.5	An example linear wave	575
38.2.6	GLM with a materially constant scalar	577
38.2.7	Further study	578
38.3	Kinematics of eddy tracer fluxes	578
38.3.1	Particle displacements and eddy tracer fluxes	578
38.3.2	Symmetric and skew-symmetric tracer fluxes	579
38.3.3	Massaging the mean field tracer equation	582
38.3.4	Connection to Stokes drift	582
38.3.5	A linear rotating periodic wave example	583
38.3.6	Further study	584
38.4	Kinematics of volume transport in isopycnal layers	584
38.4.1	Isopycnal mean	585
38.4.2	Modified mean	585
38.4.3	Transformed residual mean (TRM)	587
38.4.4	Volume conservation and the thickness equation	588
38.4.5	Ensemble mean kinematics in isopycnal coordinates	590
38.4.6	Ensemble mean kinematics in geopotential coordinates	591
38.4.7	Approximate ensemble mean kinematics in geopotential coordinates	591
38.4.8	Further study	593
38.5	Thickness transport for a linear longitudinal wave	593
38.5.1	An undulating fluid layer	593
38.5.2	Stokes drift	594
38.5.3	Linearized thickness perturbations	595
38.5.4	Correlation between thickness and velocity	595
38.5.5	Comments	596
38.6	Mean tracer equation	596
38.6.1	Thickness weighted means	596
38.6.2	Isopycnal mean thickness weighted tracer equation	596
38.6.3	Subgrid scale tracer transport tensor	597
38.6.4	Mean tracer transport beneath a density surface	597
38.6.5	Summary of the tracer parameterization problem	598
38.6.6	Comments	598

38.1 Reynold's decomposition

At any point in space and time, we can decompose a field into a mean, $\bar{\Phi}(\mathbf{x}, t)$, and a departure from the mean, $\Phi'(\mathbf{x}, t)$

$$\Phi(\mathbf{x}, t) = \bar{\Phi}(\mathbf{x}, t) + \Phi'(\mathbf{x}, t). \quad (38.1)$$

The departure from the mean is generally termed the “eddy” or the “fluctuation”. The following offers a non-exhaustive list of mean operators.

- TIME MEAN: If the mean operator is based on a long time mean, then the mean fields are assumed to be time independent: $\bar{\Phi}(\mathbf{x}, t) = \bar{\Phi}(\mathbf{x})$. This is a common operator when interest is focused on the long term mean fluid properties.

- PHASE AVERAGE: Rather than a time mean, we may choose to average over the phase (or period) of a wave. This choice is particularly relevant when the fluctuating field involves quasi-linear waves.
- ZONAL MEAN: If the mean operator is based on an average along a particular coordinate direction (e.g., zonal average), then the mean tracer concentration is independent of the “averaged out” direction.
- COARSE-GRAINING: If the mean operator is based on an average over a spatial and temporal region, such as the mesoscale, then such coarse-graining averages out smaller scales.
- ENSEMBLE MEAN: Rather than a space or time mean operation, we may consider the mean computed over an ensemble of many flow realizations. For many purposes this is the most analytically convenient operator.

If a mean operator satisfies the following properties then it is said to provide a “Reynold’s decomposition”

$$\overline{\Phi'} = 0 \tag{38.2a}$$

$$\overline{\overline{\Phi}} = \overline{\Phi} \tag{38.2b}$$

$$\overline{\gamma \overline{\Phi}} = \gamma \overline{\Phi} \quad \text{for } \gamma \text{ a constant.} \tag{38.2c}$$

Equation (38.2a) says that the mean of an eddy fluctuation vanishes. The equality (38.2b) says that the mean of a mean field returns the mean field. The final equality, (38.2c), says that a constant 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.

38.2 Basic kinematics of the GLM

We here consider basic elements of generalized Lagrangian mean (GLM) theory. GLM is distinct from both the Eulerian mean and the Lagrangian mean. Rather, 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*

$$\text{Lagrangian mean} = \text{Eulerian mean} + \text{Stokes mean.} \tag{38.3}$$

This name is motivated from the *Stokes drift* introduced in Section 36.3, which we again encounter in Section 38.2.4. Note that the literature typically 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. We instead propose that a mean operator is subjectively chosen based on its suitability to a particular scientific question. Furthermore, no mean operator is suitable for all questions.

38.2.1 Motivation

Consider a materially constant scalar field

$$\frac{D\Phi}{Dt} = \frac{\partial\Phi}{\partial t} + \mathbf{v} \cdot \nabla\Phi = 0. \quad (38.4)$$

The scalar Φ is constant following fluid particles whose trajectories are integral curves of the fluid velocity \mathbf{v} . The question arises how to develop a mean operator that averages over fluctuations in the trajectories while preserving the material constancy nature of the instantaneous equation $D\Phi/Dt = 0$. This aspiration is not trivial.

Eulerian mean

An Eulerian mean operator considered in Section 38.1 leads to the mean field equation

$$\frac{\partial\bar{\Phi}}{\partial t} + \bar{\mathbf{v}} \cdot \nabla\bar{\Phi} = -\overline{\mathbf{v}' \cdot \nabla\Phi'}. \quad (38.5)$$

Whereas Φ is materially constant when following the instantaneous flow field \mathbf{v} , the Eulerian mean $\bar{\Phi}$ is not materially constant when following $\bar{\mathbf{v}}$ due to the source term $-\overline{\mathbf{v}' \cdot \nabla\Phi'}$ provided by the eddy correlation. Furthermore, when given information only about the mean fields, then we must develop a closure for the unresolved correlation. Such closures are the topic of extensive research typical of eddy-mean decompositions. Nonetheless, we ask whether there are methods that offer insights into mean field behaviour even without making a closure assumption. GLM is one such method.

Lagrangian mean

An alternative approach is to remain in the Lagrangian frame, where material constancy of Φ takes on the linear form

$$\frac{\partial\Phi(\mathbf{a}, t)}{\partial t} = 0. \quad (38.6)$$

Consider a mean operator computed as an average over a region of material space coordinate \mathbf{a} . For example, if \mathbf{a} is the initial fluid particle position, then an average coordinate, $\bar{\mathbf{a}}$, and corresponding averaged field, $\bar{\Phi}$, render a coarse-graining over the initial positions. Since each member of the Lagrangian average satisfies the linear equation (38.6), so too does the Lagrangian mean

$$\frac{\partial\bar{\Phi}(\bar{\mathbf{a}}, t)}{\partial t} = 0. \quad (38.7)$$

Although this equation retains the simplicity of the unaveraged version, it still requires information about trajectories. Trajectories are computed based on the flow map (i.e., the velocity field), with trajectories an impractical means for describing chaotic or turbulent fluids. GLM offers an alternative that aims to meld elements of the Eulerian (e.g., computability) to the Lagrangian (e.g., material constancy).

Generalized Lagrangian mean

The GLM approach produces a GLM field that remains constant following the GLM velocity

$$\frac{\partial\bar{\Phi}^{(L)}}{\partial t} + \bar{\mathbf{v}}^{(L)} \cdot \nabla\bar{\Phi}^{(L)} = 0. \quad (38.8)$$

Hence, GLM maintains the desirable properties of the Lagrangian mean. However, it does so using Eulerian methods which can prove to be more practical for many cases. Notably, even if the Eulerian velocity is non-divergent, as for a Boussinesq fluid, the GLM velocity is generally divergent. Although we will not prove the GLM result (38.8), we will motivate the GLM average from the analysis of small amplitude eddying motions.

38.2.2 Length scales and the small parameter

There are two length scales associated with an eddy or wave fluctuation. One characterizes the size of the eddy whose length scale we write as λ . If the eddy is a monochromatic wave, then λ is its wave length. The other length scale characterizes the size of particle displacements, $|\xi|$. In the following, we assume the particle displacements are small relative to λ

$$|\xi| \ll \lambda \quad \text{small amplitude waves.} \quad (38.9)$$

We thus introduce the small non-dimensional ratio of length scales for the following analysis

$$\alpha = \frac{|\xi|}{\lambda} \ll 1. \quad (38.10)$$

38.2.3 Decomposing the particle trajectory

Recall the discussion of fluid particle trajectories given in Chapter 16. In this description, the trajectory of a particle is determined by integrating the relation between the particle trajectory and the particle velocity

$$\left[\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \right]_{\mathbf{a}} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t)] \implies \mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, 0) + \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t')] dt', \quad (38.11)$$

so that the trajectory measures the position of a particle relative to a chosen origin. 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. In turn, we introduce an Eulerian field, $\xi(\mathbf{x}, t)$, that measures the position of a fluid particle relative to its mean position. Correspondingly, the Eulerian mean of the disturbance field vanishes

$$\overline{\xi(\mathbf{x}, t)} = 0. \quad (38.12)$$

Note that the Eulerian mean operator can be any of the operators (or others) satisfying the Reynold's decomposition property discussed in Section 38.1

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 linear waves such as shown in Figure 38.1, for which the particle displacement amplitude is much smaller than the wavelength of the disturbance

$$\alpha = \frac{|\xi|}{\lambda} \ll 1. \quad (38.13)$$

In this case the disturbance field is constructed by time integration of the eddy velocity field

$$\left[\frac{\partial \xi(\mathbf{x}, t)}{\partial t} \right]_{\mathbf{x}} = \mathbf{v}'(\mathbf{x}, t) \implies \xi(\mathbf{x}, t) = \int^t \mathbf{v}'(\mathbf{x}, t') dt'. \quad (38.14)$$

With this specification for the disturbance field, we see that if the eddy velocity is non-divergent then so is the disturbance particle position field

$$\nabla \cdot \mathbf{v}' = 0 \implies \nabla \cdot \boldsymbol{\xi} = 0. \quad (38.15)$$

The definition (38.14) for the disturbance field, $\boldsymbol{\xi}(\mathbf{x}, t)$, is directly analogous to the particle trajectory position, $\mathbf{X}(\mathbf{a}, t)$, given by equation (38.11). However, there are important distinctions. Namely, the disturbance, $\boldsymbol{\xi}(\mathbf{x}, t)$, is an Eulerian field that measures the position of a fluid particle relative to its mean position, 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 field that is attached to each fluid particle and measures the position of that particle relative to a unique origin.

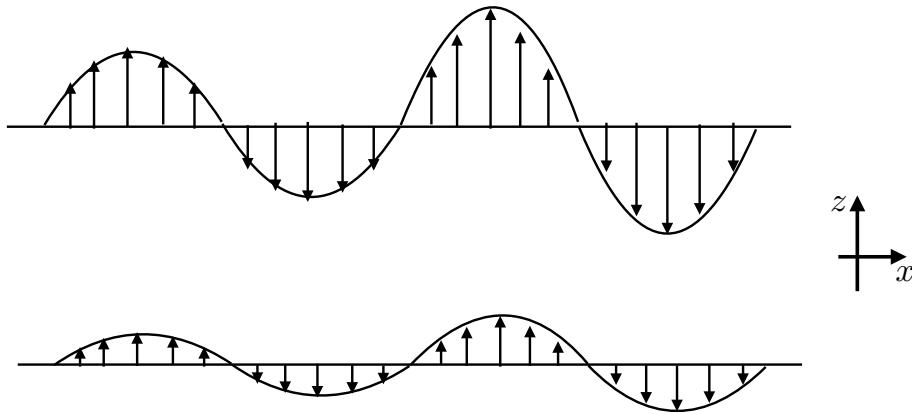


Figure 38.1: Illustrating the displacement of fluid particles at two selective vertical positions due to a linear transverse wave disturbance. The particle positions, $\mathbf{x}^{(\xi)} = \mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t)$, have a disturbance field of the form $\boldsymbol{\xi}(\mathbf{x}, t) = \hat{z} \xi_0(x, z) \sin(kx - \omega t)$, with $\xi_0(x, z)$ a spatially dependent wave amplitude, $\lambda = 2\pi/k$ the wavelength, $\mathbf{k} = \hat{x} k$ the wavevector, $\omega = ck$ the angular frequency, and c the wavespeed. Small amplitude waves satisfy $|\xi_0| \ll \lambda$. Note that this wave does not produce a Stokes drift since particle displacements are perpendicular to the wavevector: $\boldsymbol{\xi} \cdot \mathbf{k} = 0$ (see Section 38.2.5), whereas Stokes drift requires particle motion to have a nonzero component in the wave direction (see Figures 36.4 and 36.5). Even so, it does generally produce a Stokes mean for an arbitrary field Φ (Section 38.2.4).

38.2.4 GLM and the Stokes mean

The mean of a fluid property, Φ , is generally a function of how the property is sampled to compute the mean. For example, the mean of Φ sampled on a fluctuating fluid particle differs from the mean sampled at the particle's mean position. Mathematically, this distinction means that

$$\underbrace{\overline{\Phi(\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t))}}_{\text{GLM}} \neq \underbrace{\overline{\Phi(\mathbf{x}, t)}}_{\text{Eulerian}}, \quad (38.16)$$

where it is common to make use of the shorthand¹

$$\mathbf{x}^{(\xi)}(\mathbf{x}, t) \equiv \mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t) \quad (38.17)$$

for the instantaneous position of the fluid particle. The average

$$\overline{\Phi}^{(L)}(\mathbf{x}, t) \equiv \overline{\Phi(\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t), t)} = \overline{\Phi(\mathbf{x}^{(\xi)})} \quad (38.18)$$

¹We place superscripts ξ, S, L, E inside parentheses to distinguish from tensor labels.

defines the generalized Lagrangian mean. As defined, the GLM is computed by evaluating the property Φ at the position of a fluid particle, $\mathbf{x}^{(\xi)}(\mathbf{x}, t) = \mathbf{x} + \underline{\boldsymbol{\xi}}(\mathbf{x}, t)$, where \mathbf{x} is both an arbitrary Eulerian field point and the mean position of a fluid particle, $\mathbf{x}^{(\xi)} = \mathbf{x}$. In contrast, the Eulerian mean is determined by evaluating Φ at the fixed Eulerian point in space

$$\overline{\Phi}^{(E)}(\mathbf{x}, t) \equiv \overline{\Phi}(\mathbf{x}, t). \quad (38.19)$$

Following our discussion at the start of Section 38.2, we define the difference between the GLM and Eulerian mean as the Stokes mean

$$\overline{\Phi}^{(S)}(\mathbf{x}, t) = \overline{\Phi}^{(L)}(\mathbf{x}, t) - \overline{\Phi}^{(E)}(\mathbf{x}, t). \quad (38.20)$$

The Stokes mean arises from inhomogeneities in Φ , which in turn lead to differences in its mean depending on how that field is sampled, whether 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

$$\Phi(\mathbf{x} + \boldsymbol{\xi}, t) = \Phi(\mathbf{x}, t) + \boldsymbol{\xi} \cdot \nabla \Phi(\mathbf{x}, t) + \frac{1}{2} \xi_m \xi_n \partial_m \partial_n \Phi(\mathbf{x}, t) + \mathcal{O}(\alpha^3). \quad (38.21)$$

The non-dimensional ratio $\alpha = |\boldsymbol{\xi}|/\lambda \ll 1$ was introduced in equation (38.13), which measures the ratio of the amplitude for particle displacements to the wavelength, λ , of fluctuations in the field Φ . Taking the mean of equation (38.21) then leads to an expression for the Stokes mean

$$\overline{\Phi}^{(S)}(\mathbf{x}, t) = \overline{\Phi}^{(L)}(\mathbf{x}, t) - \overline{\Phi}^{(E)}(\mathbf{x}, t) \quad (38.22a)$$

$$= \overline{\boldsymbol{\xi} \cdot \nabla \Phi} + \frac{1}{2} \overline{\xi_m \xi_n \partial_m \partial_n \Phi} + \mathcal{O}(\alpha^3). \quad (38.22b)$$

$$= \overline{\boldsymbol{\xi} \cdot \nabla \Phi'} + \frac{1}{2} \overline{\xi_m \xi_n} \partial_m \partial_n \overline{\Phi}^{(E)} + \mathcal{O}(\alpha^3), \quad (38.22c)$$

where we introduced the Eulerian fluctuation

$$\Phi'(\mathbf{x}, t) = \Phi(\mathbf{x}, t) - \overline{\Phi}^{(E)}(\mathbf{x}, t) \quad (38.23)$$

and all terms on the right hand side of equation (38.22c) are evaluated at (\mathbf{x}, t) . The Stokes drift (Section 36.3) associated with the GLM arises from setting Φ equal to one of the velocity components

$$\overline{v}_p^{(S)} = \overline{\boldsymbol{\xi} \cdot \nabla v'_p} + \frac{1}{2} \overline{\xi_m \xi_n} \partial_m \partial_n \overline{v}_p^{(E)} + \mathcal{O}(\alpha^3). \quad (38.24)$$

38.2.5 An example linear wave

We exemplify the previous discussion by considering the small amplitude linear wave

$$\boldsymbol{\xi} = -\frac{\mathbf{U}(\mathbf{x})}{\omega} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (38.25a)$$

$$\mathbf{v}' = \partial_t \boldsymbol{\xi} = \mathbf{U}(\mathbf{x}) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (38.25b)$$

$$\nabla v'_p = \nabla U_p \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \mathbf{k} U_p \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (38.25c)$$

$$\nabla \cdot \mathbf{v}' = (\nabla \cdot \mathbf{U}) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \mathbf{k} \cdot \mathbf{U} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (38.25d)$$

where \mathbf{U} is the velocity amplitude that is generally a function of space, \mathbf{k} is the wavevector, and

$$T = 2\pi/\omega \quad (38.26)$$

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

$$\bar{\boldsymbol{\xi}}(\mathbf{x}) = \frac{1}{T} \int_0^T \boldsymbol{\xi}(\mathbf{x}, t') dt' = 0 \quad (38.27a)$$

$$\bar{\mathbf{v}}'(\mathbf{x}) = \frac{1}{T} \int_0^T \mathbf{v}'(\mathbf{x}, t') dt' = 0. \quad (38.27b)$$

To maintain a non-divergent eddy velocity at arbitrary times requires

$$\nabla \cdot \mathbf{v}' = 0 \implies \nabla \cdot \mathbf{U} = \mathbf{U} \cdot \mathbf{k} = 0. \quad (38.28)$$

The second condition means that the wave is transverse, so that particle displacements are orthogonal to the wavevector (e.g., Figure 38.1)

Stokes drift

Specializing to the velocity field (38.25b), substituting into the Stokes drift expression (38.24), and making use of an average over a wave period yields

$$\bar{\boldsymbol{\xi}} \cdot \nabla \bar{v}_p' = \frac{U_p \mathbf{U} \cdot \mathbf{k}}{2\omega} \quad (38.29a)$$

$$\bar{v}_p^{(E)} = 0. \quad (38.29b)$$

The second equality holds since the velocity at a point arises just from the wave field, which has a zero Eulerian mean. Hence, to $\mathcal{O}(\alpha^2)$, the Stokes drift velocity associated with the GLM is given by

$$\bar{\mathbf{v}}^{(S)} = \frac{\mathbf{U} (\mathbf{U} \cdot \mathbf{k})}{2\omega} + \mathcal{O}(\alpha^2). \quad (38.30)$$

The Stokes drift vanishes at this order of accuracy for transverse waves in which $\mathbf{U} \cdot \mathbf{k} = 0$.

As a check on our formalism we consider a one-dimensional longitudinal wave, in which the Stokes drift is given by

$$\bar{v}^{(S)} = \frac{U^2}{2c} + \mathcal{O}(\alpha^2). \quad (38.31)$$

This result agrees with that derived using Lagrangian trajectories in Section 36.3 (see Exercise 36.1). Use of 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

$$\bar{\Phi}^{(S)}(\mathbf{x}, t) = -\omega^{-1} \mathbf{U} \cdot \overline{\nabla \Phi' \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)} + \mathcal{O}(\alpha^2) \quad (38.32a)$$

$$= -\omega^{-1} \overline{\nabla \cdot (\mathbf{U} \Phi')} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) + \mathcal{O}(\alpha^2), \quad (38.32b)$$

where the second equality made use of the non-divergent nature of the wave field (38.28). To second order in wave amplitude, the Stokes mean is determined by the projection of the gradient of the Eulerian fluctuation, $\nabla\Phi'$, onto the wave amplitude, \mathbf{U} . For example, consider a transverse wave such as that shown in Figure 38.1. Even though the Stokes drift vanishes to order $\mathcal{O}(\alpha^2)$, the Stokes mean, $\bar{\Phi}^{(S)}(\mathbf{x}, t)$, can be nonzero so long as there is a nonzero vertical gradient in the Eulerian fluctuation.

38.2.6 GLM with a materially constant scalar

Consider a materially constant scalar field, such as a tracer concentration in the absence of mixing and sources

$$\frac{DC}{Dt} = 0. \quad (38.33)$$

How the GLM for C is related to the instantaneous C

The GLM for C equals to the value of C on a fluid particle

$$\bar{C}^{(L)}(\mathbf{x}, t) = C(\mathbf{x} + \boldsymbol{\xi}, t). \quad (38.34)$$

This is a very important identity that packs in a lot of information. In words, it says that when evaluated at the mean fluid particle position, \mathbf{x} , the GLM field $\bar{C}^{(L)}(\mathbf{x}, t)$ equals to the concentration, C , evaluated on a fluid particle at $\mathbf{x} + \boldsymbol{\xi}$. One means to understand this identity is to assume the GLM is an ensemble mean following fluid particles. Since C is constant on fluid particles, each ensemble member has the same value for C , in which case the GLM for C clearly equals the value of C for each ensemble member. We make particular use of the identity (38.34) when considering isopycnal kinematics in Sections 38.4 and 38.6.

Relating the particle disturbance field to Eulerian properties of C

There is a frequently used consequence of the identity (38.34) involving the disturbance field, the Eulerian fluctuation

$$C'(\mathbf{x}, t) = C(\mathbf{x}, t) - \bar{C}(\mathbf{x}, t) \quad (38.35)$$

and the Eulerian mean

$$C^{(E)}(\mathbf{x}, t) = \bar{C}(\mathbf{x}, t). \quad (38.36)$$

To derive it, recall the Taylor series expansion (38.21) truncated here to first order accuracy

$$C(\mathbf{x} + \boldsymbol{\xi}, t) = C(\mathbf{x}, t) + \boldsymbol{\xi} \cdot \nabla \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (38.37)$$

Taking the Eulerian mean of both sides renders

$$\bar{C}(\mathbf{x} + \boldsymbol{\xi}, t) = \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (38.38)$$

This identity means that the GLM equals to the Eulerian mean to order $\mathcal{O}(\alpha^2)$, which is a result consistent with the Stokes mean being an order $\mathcal{O}(\alpha^2)$ quantity as seen by equation (38.22c). From equation (38.34) we know that $C(\mathbf{x} + \boldsymbol{\xi}, t) = \bar{C}(\mathbf{x} + \boldsymbol{\xi}, t)$, so that we can subtract equations (38.37) and (38.38) to find

$$C'(\mathbf{x}, t) = -\boldsymbol{\xi} \cdot \nabla \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (38.39)$$

Hence, to first order accuracy, the Eulerian fluctuation equals to minus the disturbance field projected onto the gradient of the mean field; i.e., the Eulerian fluctuation in the tracer is first order in the disturbance. We make use of this result when discussing the kinematics of eddy tracer fluxes in Section 38.3. Furthermore, for the isopycnal kinematics in Sections 38.4 and 38.6, we focus on vertical particle displacements, $\xi = \xi \hat{z}$, in which case the Eulerian fluctuation is given by

$$C'(z, t) = -\xi \partial_z \bar{C}(z, t) + \mathcal{O}(\alpha^2). \quad (38.40)$$

38.2.7 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 [Bühler \(2014\)](#).

38.3 Kinematics of eddy tracer fluxes

Consider the Eulerian eddy-mean decomposition for a materially constant tracer in an incompressible fluid. The advection equation for this tracer is given by

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{v} C) = 0, \quad (38.41)$$

and its Eulerian mean is

$$\frac{\partial \bar{C}}{\partial t} + \nabla \cdot (\bar{\mathbf{v}} \bar{C}) = -\nabla \cdot \bar{\mathbf{v}}' C'. \quad (38.42)$$

The eddy advective flux, $\mathbf{v}' C'$, is the product of the eddy velocity and eddy tracer concentration. Its mean provides the correlation or mean eddy flux, $\bar{\mathbf{v}}' \bar{C}'$. The convergence of this mean eddy flux provides a source to the advection equation for the Eulerian mean tracer concentration.

In this section we make use of the particle disturbance field of Section 38.2 to unpack the kinematics of eddy tracer fluxes induced by small amplitude waves. Although not offering new dynamical information, the particle disturbance field is a very useful means to frame the kinematics of tracer eddy fluxes.

38.3.1 Particle displacements and eddy tracer fluxes

Following Section 38.2, we here introduce a particle disturbance vector corresponding to small amplitude eddy fluctuations

$$\partial_t \boldsymbol{\xi}(\mathbf{x}, t) = \mathbf{v}'(\mathbf{x}, t) + \mathcal{O}(\alpha^2) \quad (38.43a)$$

$$\bar{\boldsymbol{\xi}} = 0. \quad (38.43b)$$

Correspondingly, each spatial point, \mathbf{x} , is the mean position of a fluid particle whose instantaneous position is $\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t)$. Following the results from Section 38.2.6, to leading order we can write the Eulerian fluctuation in terms of the particle displacement (equation (38.39))

$$C'(\mathbf{x}, t) = -\boldsymbol{\xi} \cdot \nabla \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2) \quad (38.44)$$

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 eddy motions lead to a nonzero tracer fluctuation with the eddy tracer flux taking on the form

$$\mathbf{v}' C' = -\partial_t \boldsymbol{\xi} (\boldsymbol{\xi} \cdot \nabla) \bar{C} + \mathcal{O}(\alpha^2). \quad (38.45)$$

We unpack this expression for the purpose of characterizing kinematic properties of the eddy tracer flux

38.3.2 Symmetric and skew-symmetric tracer fluxes

From equation (38.45), the m'th component of the eddy tracer flux is given by

$$\bar{v}'_m C' = -[(\partial_t \xi_m) \xi_n] \partial_n \bar{C}. \quad (38.46)$$

We here decompose this flux in order to characterize its kinematic properties.

Decomposing the tracer flux

Let us decompose the second order tensor $(\partial_t \xi_m) \xi_n$ into its symmetric and anti-symmetric components²

$$2(\partial_t \xi_m) \xi_n = [(\partial_t \xi_m) \xi_n + (\partial_t \xi_n) \xi_m] + [(\partial_t \xi_m) \xi_n - (\partial_t \xi_n) \xi_m] \quad (38.47a)$$

$$= \partial_t(\xi_m \xi_n) + [(\partial_t \xi_m) \xi_n - (\partial_t \xi_n) \xi_m]. \quad (38.47b)$$

Introducing the symmetric and anti-symmetric correlation tensors

$$2\mathbb{K}_{mn} \equiv \overline{\partial_t(\xi_m \xi_n)} \quad (38.48a)$$

$$2\mathbb{A}_{mn} \equiv \overline{(\partial_t \xi_m) \xi_n} - \overline{(\partial_t \xi_n) \xi_m} \quad (38.48b)$$

allows us to write the mean eddy tracer flux

$$\overline{\bar{v}'_m C'} = -(\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} \quad (38.49)$$

and the mean field tracer equation (38.42)

$$\frac{\partial \bar{C}}{\partial t} + \nabla \cdot (\bar{\mathbf{v}} \bar{C}) = \nabla \cdot [(\mathbb{K} + \mathbb{A}) \cdot \nabla \bar{C}]. \quad (38.50)$$

The right hand side of this equation equals to the convergence of the symmetric and skew-symmetric tracer fluxes³

$$\nabla \cdot [(\mathbb{K} + \mathbb{A}) \cdot \nabla \bar{C}] = -\nabla \cdot (\mathbf{F}^{(\text{sym})} + \mathbf{F}^{(\text{skew})}), \quad (38.51)$$

where

$$\mathbf{F}^{(\text{sym})} = -\mathbb{K} \cdot \nabla \bar{C} \quad (38.52a)$$

$$\mathbf{F}^{(\text{skew})} = -\mathbb{A} \cdot \nabla \bar{C} \quad (38.52b)$$

$$\overline{\bar{v}' C'} = \mathbf{F}^{(\text{sym})} + \mathbf{F}^{(\text{skew})}. \quad (38.52c)$$

²See Section 20.2.4 for a similar decomposition of the velocity gradient tensor.

³We place parentheses around “skew” and “sym” to distinguish the name for these vectors from what may otherwise appear to be tensor labels.

The symmetric flux

In terms of particle displacements, the symmetric flux (38.52a) is given by

$$F_m^{(\text{sym})} = -\mathbb{K}_{mn} \partial_n \bar{C} = -\frac{1}{2} \overline{\partial_t(\xi_m \xi_n)} \partial_n \bar{C}. \quad (38.53)$$

The symmetric tensor \mathbb{K} vanishes when the average is over the period of a periodic wave, in which the particle displacements undergo reversible periodic excursions (see Section 38.3.5). For waves that decay in amplitude over the averaging period, particle displacements decrease in magnitude thus leading to an upgradient symmetric flux. 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 dispersive tracer mixing. Dispersive mixing is well parameterized by diffusion, and we have more to say about diffusive parameterizations of lateral dispersion in Section 39.3.

The skew, advective, and rotational fluxes

Following our discussion in Section 37.6, we write the skew flux as

$$F_m^{(\text{skew})} = -\mathbb{A}_{mn} \partial_n \bar{C} = -\epsilon_{mnp} \Psi_p \partial_n \bar{C} = -(\nabla \bar{C} \wedge \Psi)_m, \quad (38.54)$$

where we introduced the vector streamfunction (dimensions squared length per time)⁴

$$\Psi = \frac{1}{2} \overline{\partial_t \xi \wedge \xi} = \frac{1}{2} \overline{\mathbf{v}' \wedge \xi}. \quad (38.55)$$

The vector streamfunction is 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, in which case the wave field is said to be *polarized*.

The skew flux can be written

$$\mathbf{F}^{(\text{skew})} = -\nabla \bar{C} \wedge \Psi \quad (38.56a)$$

$$= (\nabla \wedge \Psi) \bar{C} - \nabla \wedge (\bar{C} \Psi) \quad (38.56b)$$

$$= \mathcal{U}^{(\Psi)} \bar{C} - \nabla \wedge (\bar{C} \Psi) \quad (38.56c)$$

$$= \mathbf{F}^{(\text{adv})} - \mathbf{F}^{(\text{rot})}, \quad (38.56d)$$

so that the skew flux equals to an advective flux minus a rotational flux. We here introduced the non-divergent velocity

$$\mathcal{U}^{(\Psi)} = \nabla \wedge \Psi \quad (38.57)$$

and the non-divergent rotational flux

$$\mathbf{F}^{(\text{rot})} = \nabla \wedge (\bar{C} \Psi). \quad (38.58)$$

The divergence of the skew flux equals to the divergence of the advective flux

$$\nabla \cdot \mathbf{F}^{(\text{skew})} = \nabla \cdot (\mathbf{F}^{(\text{adv})} - \mathbf{F}^{(\text{rot})}) = \nabla \cdot \mathbf{F}^{(\text{adv})}, \quad (38.59)$$

so that the rotational flux has no impact on evolution of the mean tracer field.

⁴ Middleton and Loder (1989) and Garrett (2006) introduce a skew-diffusivity, \mathbf{D} , which is opposite in sign to the vector streamfunction: $\Psi = -\mathbf{D}$.

What does a point measurement estimate?

From equation (38.52c), we see that a point measurement of the correlation $\overline{\mathbf{v}' C'}$ provides an estimate of the diffusive and skew diffusive tracer fluxes

$$\overline{\mathbf{v}' C'} = \mathbf{F}^{(\text{sym})} + \mathbf{F}^{(\text{skew})} = -(\mathbb{K} + \mathbb{A}) \cdot \nabla \overline{C}. \quad (38.60)$$

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, $-\nabla \overline{C} \wedge \Psi$. One might instead presume that the point measurement offers a direct estimate of the advective flux, $\overline{C} \mathcal{U}^{(\Psi)}$, rather than the skew flux. But that presumption is wrong. Instead, since the skew flux equals to a rotational flux plus the advective flux, we have

$$\overline{\mathbf{v}' C'} = -\mathbb{K} \cdot \nabla \overline{C} - \nabla \overline{C} \wedge \Psi \quad (38.61a)$$

$$= -\mathbb{K} \cdot \nabla \overline{C} - \nabla \wedge (\overline{C} \Psi) + \overline{C} \nabla \wedge \Psi \quad (38.61b)$$

$$= -\mathbb{K} \cdot \nabla \overline{C} - \nabla \wedge (\overline{C} \Psi) + \overline{C} \mathcal{U}^{(\Psi)}. \quad (38.61c)$$

The rotational flux is generally nontrivial for polarized waves and so cannot be ignored. As detailed by [Fox-Kemper et al. \(2003\)](#), there is no general method for removing the rotational flux. We therefore find it more convenient to work directly with the skew flux than the advective flux.

Area integrated tracer flux

We now offer an interpretation for the rotational contribution by considering the mean of the tracer flux integrated over a static area \mathcal{S}

$$\mathcal{T} = \overline{\int_{\mathcal{S}} \mathbf{v} C \cdot \hat{\mathbf{n}} dS} = \int_{\mathcal{S}} \overline{\mathbf{v} C} \cdot \hat{\mathbf{n}} dS = \int_{\mathcal{S}} [\overline{\mathbf{v}} \overline{C} + \overline{\mathbf{v}' C'}] \cdot \hat{\mathbf{n}} dS. \quad (38.62)$$

In terms of particle displacements, the eddy correlation, $\overline{\mathbf{v}' C'}$, equals to the sum of the symmetric flux and the skew flux as in equation (38.60). Introducing the diffusive, advective, and rotational flux then renders

$$\mathcal{T} = \int_{\mathcal{S}} [\overline{\mathbf{v}} \overline{C} + \mathcal{U}^{(\Psi)} \overline{C} - \nabla \wedge (\overline{C} \Psi) - \mathbb{K} \cdot \nabla \overline{C}] \cdot \hat{\mathbf{n}} dS. \quad (38.63)$$

Use of Stokes' Theorem transforms the rotational term to a line integral around the boundary of the area

$$\mathcal{T} = \int_{\mathcal{S}} [\overline{\mathbf{v}} \overline{C} + \mathcal{U}^{(\Psi)} \overline{C} - \mathbb{K} \cdot \nabla \overline{C}] \cdot \hat{\mathbf{n}} dS - \oint_{\partial \mathcal{S}} \overline{C} \Psi \cdot d\mathbf{l}. \quad (38.64)$$

Following Section 2b of [Middleton and Loder \(1989\)](#), we interpret the boundary term as a Stokes contribution associated with the correlation of particle motion and perturbation velocity along the boundary

$$\oint_{\partial \mathcal{S}} \overline{C} \Psi \cdot d\mathbf{l} = (1/2) \oint_{\partial \mathcal{S}} \overline{C} (\overline{\mathbf{v}' \wedge \xi}) \cdot d\mathbf{l}. \quad (38.65)$$

We further this interpretation when considering the transport beneath a fluctuating isopycnal surface in Section 38.6.4.

38.3.3 Massaging the mean field tracer equation

We here write the mean tracer equation (38.50) in various forms that can be found throughout the literature. For this purpose, write the right hand side of equation (38.50) in the form

$$\nabla \cdot [(\mathbb{K} + \mathbb{A}) \cdot \nabla \bar{C}] = \partial_m [(\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C}] \quad (38.66a)$$

$$= \partial_m (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} + (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_m \partial_n \bar{C} \quad (38.66b)$$

$$= \partial_m (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} + \mathbb{K}_{mn} \partial_m \partial_n \bar{C}. \quad (38.66c)$$

The final equality follows from the identity

$$\mathbb{A}_{mn} \partial_m \partial_n \bar{C} = 0, \quad (38.67)$$

which results from the contraction of the anti-symmetric, \mathbb{A}_{mn} , to the symmetric operator $\partial_m \partial_n$. The second term, $\mathbb{K}_{mn} \partial_m \partial_n \bar{C}$, is a diffusion operator if symmetric tensor \mathbb{K} is also positive-definite. The first term in equation (38.66c) can be interpreted as an advection operator through the action of a non-divergent plus a divergent advection velocity

$$\partial_m (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} = [\mathcal{U}^{(K)} + \mathcal{U}^{(\Psi)}] \cdot \nabla \bar{C}, \quad (38.68)$$

where⁵

$$\mathcal{U}_n^{(K)} \equiv -\partial_m \mathbb{K}_{mn} \implies \nabla \cdot \mathcal{U}^{(K)} = -\partial_n \partial_m \mathbb{K}_{mn} \neq 0 \quad (38.69a)$$

$$\mathcal{U}_n^{(\Psi)} \equiv -\partial_m \mathbb{A}_{mn} \implies \nabla \cdot \mathcal{U}^{(\Psi)} = \partial_n \partial_m \mathbb{A}_{mn} = 0. \quad (38.69b)$$

Bringing the above results together allows us to write the mean field tracer equation (38.50) in the following equivalent forms

$$\frac{\partial \bar{C}}{\partial t} + [\bar{\mathbf{v}} + \mathcal{U}^{(\Psi)} + \mathcal{U}^{(K)}] \cdot \nabla \bar{C} = \mathbb{K}_{mn} \partial_m \partial_n \bar{C} \quad \text{advective form} \quad (38.70a)$$

$$\frac{\partial \bar{C}}{\partial t} + \nabla \cdot ([\bar{\mathbf{v}} + \mathcal{U}^{(\Psi)}] \bar{C}) = \nabla \cdot (\mathbb{K} \cdot \nabla \bar{C}) \quad \text{flux form,} \quad (38.70b)$$

where we made use of the identities

$$\nabla \cdot \bar{\mathbf{v}} = 0 \quad \nabla \cdot \mathcal{U}^{(\Psi)} = 0 \quad \nabla \cdot \mathcal{U}^{(K)} \neq 0. \quad (38.71)$$

38.3.4 Connection to Stokes drift

From equation (38.24) we have the leading order expression for the Stokes drift

$$\bar{v}_p^{(S)} = \overline{\xi_n \partial_n \partial_t \xi_p} + \mathcal{O}(\alpha^2). \quad (38.72)$$

As noted in equation (38.15), with $\partial_t \boldsymbol{\xi} = \mathbf{v}'$ and with $\nabla \cdot \mathbf{v}' = 0$, the corresponding particle displacements are non-divergent, $\nabla \cdot \mathbf{v}' = 0 \Rightarrow \nabla \cdot \boldsymbol{\xi} = 0$. Consequently, to second order accuracy, the Stokes drift velocity can be written

$$\bar{v}_p^{(S)} = \overline{\xi_n \partial_n \partial_t \xi_p} \quad (38.73a)$$

$$= \partial_n [\overline{(\partial_t \xi_p) \xi_n}] \quad (38.73b)$$

$$= \partial_n (\mathbb{K}_{pn} + \mathbb{A}_{pn}) \quad (38.73c)$$

$$= \partial_n (\mathbb{K}_{np} - \mathbb{A}_{np}) \quad (38.73d)$$

$$= -\mathcal{U}_p^{(K)} + \mathcal{U}_p^{(\Psi)}. \quad (38.73e)$$

⁵Note that [Middleton and Loder \(1989\)](#) define $\mathcal{U}_n^{(K)} \equiv +\partial_m \mathbb{K}_{mn}$, which is the opposite sign to that used here in equation (38.69a), whereas they define $\mathcal{U}_n^{(\Psi)} = -\partial_m \mathbb{A}_{mn}$ as in equation (38.69b).

For the case of periodic waves, the Stokes drift velocity equals to the non-divergent skew velocity

$$\bar{\mathbf{v}}^{(S)} = \mathcal{U}^{(\Psi)} \Rightarrow \nabla \cdot \bar{\mathbf{v}}^{(S)} = 0 \quad \text{periodic waves.} \quad (38.74)$$

More generally, for non-periodic waves the divergent velocity is non-zero so that the Stokes velocity is also divergent

$$\bar{\mathbf{v}}^{(S)} = \mathcal{U}^{(\Psi)} - \mathcal{U}^{(K)} \Rightarrow \nabla \cdot \bar{\mathbf{v}}^{(S)} = -\nabla \cdot \mathcal{U}^{(K)} \neq 0 \quad \text{non-periodic waves.} \quad (38.75)$$

38.3.5 A linear rotating periodic wave example

We illustrate some of the previous analysis by considering Consider a displacement vector comprised of a linear periodic wave in two-dimensions

$$\xi(\mathbf{x}, t) = \Gamma [\hat{\mathbf{x}} \cos(kx - \omega t) + \hat{\mathbf{y}} \sin(kx - \omega t)] \quad (38.76a)$$

$$\partial_t \xi(\mathbf{x}, t) = \omega \Gamma [\hat{\mathbf{x}} \sin(kx - \omega t) - \hat{\mathbf{y}} \cos(kx - \omega t)], \quad (38.76b)$$

where Γ the time-independent wave amplitude, $T = 2\pi/\omega$ is the wave period, and $\lambda = 2\pi/k$ is the wavelength. The fluid particles exhibit counter-clockwise circular motion in the horizontal plane with squared radius

$$\xi \cdot \xi = \Gamma^2. \quad (38.77)$$

We are motivated to let the mean operator be a phase average

$$\bar{\phi} = \frac{1}{T} \int_0^T \phi(t) dt, \quad (38.78)$$

which is the traditional operator used when examining the impacts of linear waves on mean fields. For spatially constant wave amplitude, we will see the the mean tracer concentration, \bar{C} , remains unchanged by these waves. The absence of a rectified change to \bar{C} reflects the linear periodic nature of the wave field.

Symmetric mixing tensor

The symmetric mixing tensor

$$\mathbb{K}_{mn} = \frac{\Gamma^2}{2T} \int_0^T dt \frac{\partial}{\partial t} \begin{bmatrix} \cos^2(kx - \omega t) & \cos(kx - \omega t) \sin(kx - \omega t) \\ \cos(kx - \omega t) \sin(kx - \omega t) & \sin^2(kx - \omega t) \end{bmatrix} \quad (38.79)$$

vanishes identically since the wave field is periodic so that the particle motion has an amplitude whose growing phase is exactly matched by its decaying phase.

Skew-symmetric stirring tensor

In contrast, the skew-symmetric tensor has non-zero components

$$\mathbb{A}_{12} = -\mathbb{A}_{21} = \frac{\Gamma^2 \omega}{2} [\sin^2(kx - \omega t) + \cos^2(kx - \omega t)] = \frac{\Gamma^2 \omega}{2}, \quad (38.80)$$

which reflects the counter-clockwise polarization. The corresponding vector streamfunction is vertical

$$\Psi = \frac{\Gamma^2 \omega}{2} \hat{z}, \quad (38.81)$$

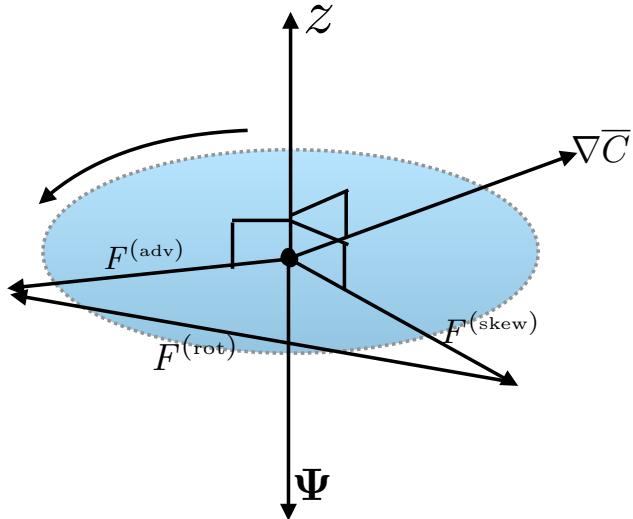


Figure 38.2: Sketch of the various tracer fluxes associated with the polarized displacement vector (38.76a). The particles are moving on the horizontal plane in a circle with time-independent radius Γ . The vector streamfunction (38.81) points in the negative \hat{z} direction. The mean concentration gradient, $\nabla\bar{C}$, generally points outside of the horizontal plane. However, it is only the horizontal components that contribute since the displacement vector is in the horizontal plane, thus resulting in horizontal skew, advective, and rotational fluxes.

and the skew flux is horizontal

$$\mathbf{F}^{(\text{skew})} = \frac{\Gamma^2 \omega}{2} (\hat{z} \wedge \nabla\bar{C}). \quad (38.82)$$

Finally, the advective velocity is given by

$$\nabla \wedge \Psi = \omega \Gamma \nabla\Gamma \wedge \hat{z}, \quad (38.83)$$

with the advective velocity vanishing when the wave amplitude, Γ , is a constant. In this case, the advective tracer flux is zero, although the skew flux is non-zero. With a constant Γ , the skew flux has a zero divergence (it is a purely rotational flux when Γ is constant). Hence, for a constant wave amplitude, neither the skew flux nor the advective flux affect the evolution of \bar{C} . Figure 38.2 offers a schematic of the skew, advective, and rotational fluxes induced by a linear rotating particle wave in the horizontal plane.

38.3.6 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. Additional treatments can be found in the review article of [Moffatt \(1983\)](#), who considers rotating fluids and magnetic fluids.

38.4 Kinematics of volume transport in isopycnal layers

In this section we consider the reversible stirring of fluid parcels by turbulent flow in a perfect (i.e., no mixing or sources) stratified Boussinesq fluid. As the fluid parcels are stirred, they preserve their volume while changing their shape and stretching into finer scale filaments. Stirring by ocean mesoscale/baroclinic eddies offers the canonical example of such stirring. Eventually, small-scale

processes, such as those summarized in Section 40.2.4 mix properties irreversibly. We are here focused just on the stirring part of this scenario.

Over space and time scales larger than the mesoscale, the stirring by ocean mesoscale eddies can be considered chaotic, which in turn motivates a stochastic perspective in which an ensemble of eddies is considered. The goal is to describe the ensemble mean properties of the perfect fluid, with a focus in this section on the kinematics of parcel rearrangement. Hence, eddy correlations in the present section appear between the thickness of a fluid layer and the velocity. We introduce trace matter in Section 38.6, at which point we also consider eddy correlations between velocity and tracer as in Section 38.3.

The material in this section is rather detailed. However, its mastery comes readily by keeping in mind the more general (and somewhat simpler) presentation of GLM offered in Section 38.2. We are motivated to provide full details in this section since the kinematics of isopycnal ensembles appears throughout the study of wave-mean flow interactions in adiabatic geophysical fluid mechanics, such as in the study of ocean mesoscale eddies.

38.4.1 Isopycnal mean

Each fluid parcel in a stably stratified perfect Boussinesq fluid preserves its potential density. We are interested in following the vertical motion of potential density layer interfaces as waves and turbulent processes transport layer thickness from one region to another. In contrast, we are not concerned with following the lateral position of a fluid parcel within a layer. Here we introduce the isopycnal mean, which is based on describing ensembles of perfect fluid parcels using isopycnal coordinates. In Section 38.4.3, we relate this isopycnal approach to the vertical/isopycnal GLM.

Defining the isopycnal ensemble

An overbar with a potential density label, $\overline{(\)}^{(\sigma)}$, denotes a mean over an ensemble of fluid parcels, each having the same potential density, σ , the same horizontal position, (x, y) , and the same time, t . Isopycnals undulate in space and time, which means that each ensemble member has a vertical position that is generally distinct from the ensemble mean depth, z . Furthermore, when the context is clear, it is useful to drop the dependence on (x, y, t) to highlight the dependence on potential density and/or the vertical position.

Isopycnal ensemble mean

The isopycnal ensemble mean makes use of potential density as a vertical coordinate (Chapters 11 and 32), with the mean field denoted by

$$\overline{\Phi}^{(\sigma)}(x, y, \sigma, t) \equiv \text{ensemble mean using isopycnal vertical coordinates.} \quad (38.84)$$

This average is straightforward to compute when using isopycnal coordinates, thus producing an isopycnal mean that is a function of the potential density, σ .

38.4.2 Modified mean

As a complement to the isopycnal approach in Section 38.4.1, we here introduce the vertical/isopycnal GLM, also known as the modified mean.

Figure 38.3: Schematic of the ensemble mean depth $\bar{z}^\rho(x, y, \rho, t)$ of a particular potential density surface ρ . In general, different members of an isentropic ensemble live at different depths. Therefore, when considering ensemble members with the same potential density, the ensemble mean depth is the average over the different members. For the case of a two-member ensemble, as shown here, $2\bar{z}^\rho(\rho) = z(1, \rho) + z(2, \rho)$, where the depth $z(1, \rho)$ is generally different from $z(2, \rho)$.

Vertical/isopycnal GLM

The discussion in Section 38.2 considered a three-vector particle displacement vector $\xi(x, t)$. In contrast, we are here interested just in the vertical displacement of an isopycnal layer interface

$$\xi(x, y, \sigma, t) = \hat{z} \xi(x, y, \sigma, t). \quad (38.85)$$

The displacement field $\xi(x, y, \sigma, t)$ measures the vertical position of a potential density interface, σ , relative to its ensemble mean depth. For any particular ensemble member with potential density σ , we write its vertical position as (dropping x, y, t dependence for brevity)

$$z(\sigma) = \bar{z}^{(\sigma)} + \xi(\sigma), \quad (38.86)$$

where

$$\bar{z}^{(\sigma)} = \overline{z(\sigma)}^{(\sigma)} \quad (38.87)$$

is the isopycnal ensemble mean depth, and the displacement field has a zero ensemble mean

$$\overline{\xi(\sigma)}^{(\sigma)} = 0. \quad (38.88)$$

Given the above definitions for the vertical position, we define the *vertical/isopycnal GLM* for an arbitrary function

$$\tilde{\Phi}(x, y, \bar{z}^{(\sigma)}, t) \equiv \overline{\Phi(x, y, \bar{z}^{(\sigma)} + \xi(x, y, \sigma, t), t)}^{(\sigma)}. \quad (38.89)$$

As defined, the vertical/isopycnal GLM, $\tilde{\Phi}$, is a function of the ensemble mean vertical position, $\bar{z}^{(\sigma)}$ (left hand side), and is determined by an ensemble mean of Φ sampled at the depth of each ensemble member, $\bar{z}^{(\sigma)} + \xi(\sigma)$. [McDougall and McIntosh \(2001\)](#) refer to the vertical/isopycnal GLM (38.89) as the *modified mean*.

Relating the modified mean to the isopycnal mean

Following the general result (38.34), we know that the modified mean potential density, $\tilde{\sigma}(x, y, \bar{z}^{(\sigma)}, t)$, equals to the potential density of each ensemble member so that

$$\tilde{\sigma}(x, y, \bar{z}^{(\sigma)}, t) = \sigma(x, y, \bar{z}^{(\sigma)} + \xi(x, y, \sigma, t), t). \quad (38.90)$$

This relation means that the modified mean potential density is the functional inverse of the isopycnal ensemble mean vertical position. Consequently, the isopycnal ensemble mean of a function, $\overline{\Phi}^{(\sigma)}$ (equation (38.84)), when evaluated at the modified mean potential density, $\tilde{\sigma}$, equals to the modified mean $\tilde{\Phi}$ when evaluated at the vertical position of the mean density

$$\overline{\Phi}^{(\sigma)}(x, y, \tilde{\sigma}, t) = \tilde{\Phi}(x, y, \bar{z}^{(\sigma)}, t). \quad (38.91)$$

This is a very important identity that will be used in the following.

38.4.3 Transformed residual mean (TRM)

When working with isopycnal layers, it is very useful to use thickness weighting to account for the net amount of material within a layer, or to measure the net transport in the layer. We make use of thickness weighted fields, $h \Phi$, and the corresponding thickness weighted isopycnal ensemble mean

$$\hat{\Phi}(\sigma) = \frac{\overline{h \Phi}^{(\sigma)}}{\overline{h}^{(\sigma)}}. \quad (38.92)$$

The identity (38.90) then renders

$$\overline{\Phi}^{\#}(x, y, \bar{z}^{(\sigma)}, t) \equiv \hat{\Phi}(x, y, \tilde{\sigma}, t), \quad (38.93)$$

where $\overline{\Phi}^{\#}$ is the *transformed residual mean* (TRM) evaluated at the isopycnal ensemble mean vertical position. This is yet another important identity that will be used in the following.

Depth integrated TRM transport

A particularly key TRM field is the TRM horizontal velocity

$$\hat{\mathbf{u}}(x, y, \tilde{\sigma}, t) = \overline{\mathbf{u}}^{\#}(x, y, \bar{z}^{(\sigma)}, t). \quad (38.94)$$

Following the discussion of the vertical gauge in Section 37.6.1 (see in particular equation (37.71)), we are led to define the depth integrated TRM transport

$$\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)}) = \int_{-H}^{\bar{z}^{(\sigma)}} \overline{\mathbf{u}}^{\#}(z) dz = \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \hat{\mathbf{u}}(\gamma) \overline{h}^{(\gamma)} d\gamma, \quad (38.95)$$

with the second equality following from a change of coordinates from geopotential to isopycnal. We can go further with this expression by writing

$$\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)}) = \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \hat{\mathbf{u}}(\gamma) \overline{h}^{(\gamma)} d\gamma \quad \text{from equation (38.95)} \quad (38.96a)$$

$$= \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \overline{\mathbf{u} h}^{(\gamma)} d\gamma \quad \text{from equation (38.92)} \quad (38.96b)$$

$$= \int_{\sigma(-H)}^{\sigma(\bar{z}^{(\sigma)} + \xi)} \overline{\mathbf{u} h}^{(\gamma)} d\gamma \quad \text{from equation (38.90).} \quad (38.96c)$$

The final equality makes it clear that the TRM transport, $\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)})$, is the ensemble mean volume transport for fluid denser than $\sigma(\bar{z}^{(\sigma)} + \xi) = \tilde{\sigma}(\bar{z}^{(\sigma)})$. This transport can also be written using geopotential coordinates

$$\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)}) = \overline{\int_{-H}^{\bar{z}^{(\sigma)} + \xi} \mathbf{u} dz}. \quad (38.97)$$

The transport from each ensemble member is determined by integrating from the bottom to the depth, $\bar{z}^{(\sigma)} + \xi$, and then the TRM transport is determined by computing the ensemble mean for this transport.

Quasi-Stokes transport

The TRM transport (38.97) can be decomposed into an Eulerian mean plus the correlation of a fluctuation

$$\bar{U}^\#(\bar{z}^{(\sigma)}) \equiv \bar{U}(\bar{z}^{(\sigma)}) + \bar{U}^{\text{qs}}(\bar{z}^{(\sigma)}). \quad (38.98)$$

The first term,

$$\bar{U}(\bar{z}^{(\sigma)}) = \int_{-H}^{\bar{z}^{(\sigma)}} \mathbf{u} dz \quad (38.99)$$

is the ensemble mean transport between the bottom and the ensemble mean depth, $\bar{z}^{(\sigma)}$. We interpret this transport as an Eulerian mean since the depth ranges are fixed. In contrast, the *quasi-Stokes* transport

$$\bar{U}^{\text{qs}}(\bar{z}^{(\sigma)}) \equiv \overline{\int_{\bar{z}^{(\sigma)}}^{\bar{z}^{(\sigma)} + \xi} \mathbf{u} dz} \quad (38.100)$$

measures the ensemble mean transport between the mean vertical position of an isopycnal, $\bar{z}^{(\sigma)}$, and that of each ensemble member, $\bar{z}^{(\sigma)} + \xi(\sigma)$. We refer to transport as “quasi-Stokes” given that is the difference between an isopycnal (i.e., quasi-Lagrangian) mean and an Eulerian mean (see Section 38.2)

$$\bar{U}^{\text{qs}} = \bar{U}^\# - \bar{U}. \quad (38.101)$$

As for the traditional Stokes drift discussed in Sections 36.3, and 38.2.5, which arises from a correlation between larger velocity when a wave crest is present, so too does the quasi-Stokes transport arise from a correlation between a larger velocity and a larger undulation in isopycnal thickness.

Three-component TRM velocity

Following from the vertical gauge expression (37.70), we introduce the TRM vector streamfunction

$$\bar{\Psi}^\# = \bar{U}^\# \wedge \hat{z}, \quad (38.102)$$

and the corresponding three-dimensional non-divergent TRM velocity

$$\bar{v}^\# = \nabla \wedge \bar{\Psi}^\#. \quad (38.103)$$

The vertical component,

$$\bar{w}^\# = \hat{z} \cdot (\nabla \wedge \bar{\Psi}^\#), \quad (38.104)$$

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, $\bar{U}^\#$, so the two descriptions in effect make use of the same number of degrees of freedom.

38.4.4 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 parcel is written

$$\delta V = \delta x \delta y \delta z = \delta x \delta y \delta \sigma h, \quad (38.105)$$

where we introduced the specific thickness

$$h = \frac{\partial z}{\partial \sigma}. \quad (38.106)$$

As discussed in Section 11.9.1, specific thickness is the Jacobian of transformation between geopotential coordinates, (x, y, z, t) , and isopycnal coordinates, (x, y, σ, 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 (Section 25.3.5)

$$N^2 = -\frac{g}{\rho_0} \frac{\partial \sigma}{\partial z} = -\frac{g}{\rho_0 h} \quad (38.107)$$

Geometrically, the product $|h \delta \sigma|$ represents the vertical distance, or *thickness*, between the two infinitesimally close density classes σ and $\sigma + \delta \sigma$. 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 45.1.3)

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) = 0, \quad (38.108)$$

with \mathbf{u} the horizontal velocity field, the time derivative is computed with σ held fixed, and

$$\nabla_\sigma = \nabla_z + \mathbf{S} \frac{\partial}{\partial z} \quad (38.109)$$

is the horizontal derivative operator with σ held fixed and

$$\mathbf{S} = \nabla_\sigma z \quad (38.110)$$

is the horizontal slope of the potential density surface.

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, $D\sigma/Dt = 0$. When written as skewson rather than advection, the natural gauge is the vertical gauge introduced in Section 37.6.1 (equation (37.68)), since this gauge only requires the same horizontal velocity field \mathbf{u} used with the isopycnal coordinate description. This gauge has an associated skew flux of potential density $\mathbf{F}^{(\text{skew})} = -\nabla \sigma \wedge \Psi$, which leads to the evolution

$$\frac{\partial \sigma}{\partial t} = \nabla \cdot (\nabla \sigma \wedge \Psi), \quad (38.111)$$

where all derivatives are here taken with fixed Eulerian (geopotential) coordinates, (x, y, z) , and the divergence operator is three-dimensional.

Figure 38.4: Schematic of the ensemble averaged potential density as measured by an observer at a fixed point (x, y, z, t) in space-time. In general, different members of the ensemble have potential density surfaces that live at different depths. That is, a fixed Eulerian space-time observer measures an ensemble mean potential density as the average over different potential density surfaces. For the case of a two-member ensemble as shown here, $2\bar{\rho}^z(z) = \rho(1, z) + \rho(2, z)$, where $\rho(1, z)$ is generally different from $\rho(2, z)$.

38.4.5 Ensemble mean kinematics in isopycnal coordinates

Consider an ensemble of stably stratified (so that the layer specific thickness h is single-signed and nonvanishing) perfect Boussinesq fluid parcels with the same infinitesimal volume, $\delta V = \delta x \delta y \delta z = \delta x \delta y h \delta \sigma$, and same potential density, σ . 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, 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, σ, t) , the thickness equation (38.108) is satisfied by each ensemble member

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) = 0. \quad (38.112)$$

The ensemble mean computed over these fluid parcels with potential density σ satisfies

$$\partial_t \bar{h}^{(\sigma)} + \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \bar{\mathbf{u}}^{(\sigma)} + \bar{h}' \bar{\mathbf{u}}'^{(\sigma)}) = 0, \quad (38.113)$$

where primed variables represent deviations from the isopycnal mean. The mean specific thickness $\bar{h}^{(\sigma)}$ of parcels with potential density σ therefore satisfies the conservation equation

$$\partial_t \bar{h}^{(\sigma)} + \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \hat{\mathbf{u}}) = 0. \quad (38.114)$$

In this equation we introduced the thickness weighted isopycnal ensemble mean horizontal velocity

$$\hat{\mathbf{u}} = \frac{\bar{h} \bar{\mathbf{u}}^{(\sigma)}}{\bar{h}^{(\sigma)}} = \bar{\mathbf{u}}^{(\sigma)} + \frac{\bar{h}' \bar{\mathbf{u}}'^{(\sigma)}}{\bar{h}^{(\sigma)}} \equiv \bar{\mathbf{u}}^{(\sigma)} + \mathbf{u}^{\text{bolus}}, \quad (38.115)$$

along with the isopycnal ensemble mean horizontal velocity, $\bar{\mathbf{u}}^{(\sigma)}$, 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

$$\bar{h}^{(\sigma)} \mathbf{u}^{\text{bolus}} = \bar{h}^{(\sigma)} (\hat{\mathbf{u}} - \bar{\mathbf{u}}^{(\sigma)}) = \bar{h}' \bar{\mathbf{u}}'^{(\sigma)} \quad (38.116)$$

arises from the along-isopycnal correlations between specific thickness and horizontal velocity.

Quite conveniently, the mean conservation equation (38.114) takes the *same* mathematical form as the conservation equation (38.112) satisfied by each ensemble member. The key difference is that the isopycnal ensemble mean thickness $\bar{h}^{(\sigma)}$ 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 (38.114) is afforded by use of the Lagrangian vertical coordinate σ .

38.4.6 Ensemble mean kinematics in geopotential coordinates

We 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, $\bar{z}^{(\sigma)}$. Consequently, mean fields defined at the fixed vertical position correspond to either modified mean fields when not thickness weighted (equation (38.89)), or TRM fields when thickness weighted (equation (38.93)).

Evolution of modified mean density

Following the skew-symmetric formulation from Section 37.6, at the ensemble mean depth $z = \bar{z}^\rho$, the streamfunction $\bar{\Psi}^\#$ defines an effective skew flux of the modified mean potential density given by

$$\bar{\mathbf{F}}^\# = -\nabla \tilde{\sigma} \wedge \bar{\Psi}^\#. \quad (38.117)$$

Using the identity $\bar{\Psi}^\# = \bar{\mathbf{U}}^\# \wedge \hat{\mathbf{z}}$, we can write this expression in one of the forms

$$\bar{\mathbf{F}}^\# = -\bar{\mathbf{U}}^\# \partial_z \tilde{\sigma} + \hat{\mathbf{z}} \bar{\mathbf{U}}^\# \cdot \nabla_z \tilde{\sigma} \quad (38.118a)$$

$$= -(\bar{\mathbf{U}}^\# + \hat{\mathbf{z}} \mathbf{S} \cdot \bar{\mathbf{U}}^\#) \partial_z \tilde{\sigma}, \quad (38.118b)$$

where

$$\mathbf{S} = -\frac{\nabla_z \tilde{\sigma}}{\partial_z \tilde{\sigma}} \quad (38.119)$$

is the slope of the modified mean density field and $\nabla_z = (\partial_x, \partial_y, 0)$ is the horizontal gradient operator taken with constant depth $z = \bar{z}^{(\sigma)}$. The convergence of the effective skew flux leads to a stirring of the modified mean density $\tilde{\sigma}$ at the mean depth $z = \bar{z}^{(\sigma)}$,

$$\frac{\partial \tilde{\sigma}}{\partial t} = \nabla \cdot (\nabla \tilde{\sigma} \wedge \bar{\Psi}^\#). \quad (38.120)$$

This equation represents an 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 (38.111) satisfied at depth z by a single member of the ensemble.

38.4.7 Approximate ensemble mean kinematics in geopotential coordinates

Equation (38.120) 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 Lagrangian information used to realize this exact description must be approximated.

Estimating the quasi-Stokes transport

The approximation requires us to estimate the quasi-Stokes transport $\bar{\mathbf{U}}^{\text{qs}}$ defined by equation (38.100). We addressed a similar estimation in Section 38.2.4 when discussing the Stokes mean. Here, we expand the TRM transport in a Taylor series about the geopotential $z = \bar{z}^{(\sigma)}$

$$\bar{\mathbf{U}}^\#(z) = \overline{\int_{-H}^{z+\xi} \mathbf{u}(s) ds} \quad (38.121a)$$

$$= \bar{\mathbf{U}}(z) + \overline{\mathbf{u} \xi^{(z)}} + \frac{1}{2} \overline{\partial_z \mathbf{u} \xi \xi^{(z)}} + \mathcal{O}(\alpha^3), \quad (38.121b)$$

where neglected terms are third order in deviation quantities. Note that all ensemble means are taken at fixed vertical position, which accords with taking a Taylor series about the ensemble mean depth $z = \bar{z}^{(\sigma)}$.

The ensemble means in equation (38.121b) are interpreted as follows. The first term is the Eulerian mean horizontal transport passing beneath the ensemble mean depth, $z = \bar{z}^{(\sigma)}$. The second term, $\bar{\mathbf{u}}\xi$ is the horizontal velocity evaluated at the ensemble mean depth and multiplied by the deviation, ξ , of the potential density surface from its mean depth, all averaged at fixed depth. An Eulerian split of the horizontal velocity \mathbf{u} into its Eulerian mean $\bar{\mathbf{u}}^{(z)}$ and deviation \mathbf{u}' leads to the correlation

$$\bar{\mathbf{u}}\xi^{(z)} = \bar{\mathbf{u}'}\xi^{(z)}. \quad (38.122)$$

For the second order term, similar considerations lead to

$$\partial_z \bar{\mathbf{u}}\xi\xi^{(z)} \approx \partial_z \bar{\mathbf{u}}^z \xi\xi^{(z)}, \quad (38.123)$$

where neglected terms are third order and higher. Combining these relations leads to the second order accurate expression

$$\bar{\mathbf{U}}^\# \approx \bar{\mathbf{U}} + \bar{\mathbf{u}'}\xi^{(z)} + \frac{1}{2}\xi\xi^z \partial_z \bar{\mathbf{u}}^{(z)}. \quad (38.124)$$

The disturbance field

Following the discussion in Section 38.2.6, we here determine the disturbance field, ξ , in terms of fields at constant depth. For this purpose, use the identity (38.90) to give

$$\tilde{\sigma}(z) = \sigma(z + \xi) \quad (38.125a)$$

$$= \sigma(z) + \partial_z \sigma(z) \xi + \frac{1}{2} \partial_{zz} \sigma(z) \xi^2 + \mathcal{O}(\alpha^3). \quad (38.125b)$$

Subtracting the Eulerian mean of equation (38.125b) from the unaveraged equation (38.125b), and noting that $\tilde{\sigma}$ is already a mean field, leads to the second order accurate expression for the deviation

$$\xi = -\sigma'(z)/\partial_z \bar{\sigma}^{(z)} + \mathcal{O}(\alpha^2), \quad (38.126)$$

where

$$\sigma(z) = \bar{\sigma}^{(z)} + \sigma'(z). \quad (38.127)$$

To within the same order, the deviation can be written

$$\xi = -\sigma'(z)/\partial_z \tilde{\sigma}(z) + \mathcal{O}(\alpha^2). \quad (38.128)$$

Approximate quasi-Stokes transport

Substituting the deviation (38.128) into the approximate expression (38.121b) for the TRM transport yields an approximate expression for the Stokes transport

$$\mathbf{U}^{\text{qs}} = -\frac{\bar{\mathbf{u}'}\sigma'^{(z)}}{\partial_z \tilde{\sigma}} + \frac{\bar{\phi}^{(z)} \partial_z \bar{\mathbf{u}}^{(z)}}{(\partial_z \tilde{\sigma})^2} + \mathcal{O}(\alpha^3), \quad (38.129)$$

where

$$\bar{\phi}^{(z)} = \frac{1}{2} \overline{\sigma' \sigma'}^{(z)} \quad (38.130)$$

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 (38.129). We have more to say regarding this parameterization in Section 39.1.

Substituting the deviation (38.128) into the approximate expression (38.125b) yields, to within terms of third order, the relation

$$\tilde{\sigma} = \bar{\sigma}^{(z)} - \partial_z \left[\frac{\bar{\phi}^{(z)}}{\partial_z \bar{\sigma}^{(z)}} \right] + \mathcal{O}(\alpha^3). \quad (38.131)$$

As for the Stokes transport, the modified mean density and Eulerian mean density, when evaluated at the same depth, differ by terms that are second order in eddy amplitude.

38.4.8 Further study

This section is largely based on approaches used by [DeSzoeke 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 \(2004b\)](#), [Wolfe \(2014\)](#).

38.5 Thickness transport for a linear longitudinal wave

We here study the Stokes drift within a layer of constant density fluid, showing that this drift leads to a net transport of volume per horizontal area; i.e., the layer thickness. To simplify the maths, we focus on transport from linear longitudinal waves. This discussion provides a specific example of the general considerations of thickness transport by eddies presented in Section 38.4.5. Indeed, much of our intuition for the general case is based on this relatively simple worked example.⁶

38.5.1 An undulating fluid layer

Figure 38.5 shows a layer of constant density fluid within a stably stratified fluid. The total volume of fluid within this layer is assumed to remain constant, which means the layer does not mix with surrounding fluid layers; i.e., it is an immiscible fluid layer. In its unperturbed state with flat layer interfaces, the meridional velocity in the fluid layer is zero and the thickness is a constant, h_o . When perturbed, the thickness is written

$$h(y, t) = h_o + h'(y, t). \quad (38.132)$$

The layer thickness changes in time according to the convergence of the advective transport of thickness

$$\frac{\partial h}{\partial t} = -\nabla \cdot (h \mathbf{u}), \quad (38.133)$$

where the convergence is computed within the layer. We offer a derivation of this volume conservation equation in Section 42.1.3. Even without working through that derivation, the truth of this equation follows from Figure 38.5, whereby 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

$$\frac{\partial h}{\partial t} = -\frac{\partial (h v)}{\partial y}. \quad (38.134)$$

⁶This worked example is based Section 2 of [Lee et al. \(1997\)](#).

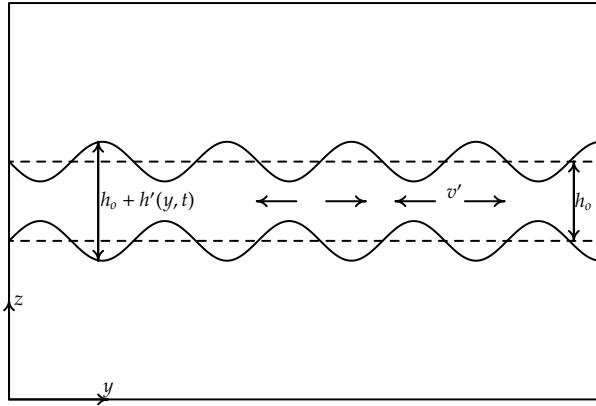


Figure 38.5: Shown here is a single layer of constant density fluid, with resting thickness $h = h_o$ and instantaneous thickness $h = h_o + h'(y, t)$. Associated with the undulations in thickness are fluctuations in the meridional velocity $v' = v_o \sin(k y - \omega t)$, depicted here by the alternating vectors within the layer. Vertical-meridional axes are shown in the lower left corner for orientation. We are not concerned with boundaries in the meridional direction.

38.5.2 Stokes drift

Consider a linear wave perturbation in the meridional velocity that propagates in the meridional direction

$$v'(y, t) = v_o \sin(k y - \omega t), \quad (38.135)$$

where k is a constant wave number and ω is a constant frequency. This longitudinal wave is depicted in Figure 38.5. We now follow the general formalism developed in Section 36.3 or equivalently in Section 38.2.4 to determine 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

$$\frac{dY}{dt} = v_o \sin(k Y - \omega t), \quad (38.136)$$

where $Y = Y(Y_o, t)$ is the meridional trajectory with initial position Y_o . Following equation (36.43) 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)

$$\frac{dY}{dt} - v(y, t) = v_o^2 k \cos(ky - \omega t) \int_0^t \sin(ky - \omega t') dt' \quad (38.137a)$$

$$= \frac{v_o^2 k}{\omega} (\cos^2(ky - \omega t) - \cos(ky - \omega t) \cos(ky)). \quad (38.137b)$$

Time averaging over a single wave period,

$$T = 2\pi/\omega \quad (38.138)$$

leads to the Stokes drift as per the general expression in equation (36.45)

$$V_{\text{Stokes}} = \frac{v_o^2 k}{2\omega}. \quad (38.139)$$

Introducing the phase speed for the wave $c = \omega/k$ allows us to write the Stokes drift as

$$V_{\text{Stokes}} = \frac{v_o^2}{2c}. \quad (38.140)$$

The Stokes drift becomes small when the phase speed is large, since the fluid particles have only a short time to feel each wave. In this case, there is only a small difference between the Eulerian and Lagrangian velocities. The converse holds for slow phase speeds, where Eulerian and Lagrangian velocities differ more. Note that if we were to consider a more careful asymptotic expansion, then the case of relatively slow phase speeds would require us to keep more terms in the expansion.

38.5.3 Linearized thickness perturbations

The velocity and thickness are written in terms of their rest state plus a perturbation due to the wave

$$h = h_o + h' \quad (38.141a)$$

$$v = v', \quad (38.141b)$$

where the velocity vanishes when the wave is absent. The thickness equation (38.134) thus takes the form

$$\frac{\partial h'}{\partial t} + h_o \frac{\partial v'}{\partial y} + v' \frac{\partial h'}{\partial y} = 0. \quad (38.142)$$

Linearizing this equation, and using the wave perturbation (38.135), leads to

$$\frac{\partial h'}{\partial t} + h_o v_o k \cos(ky - \omega t) = 0, \quad (38.143)$$

thus yielding the thickness perturbation

$$h' = h_o \frac{v'}{c}. \quad (38.144)$$

Hence, to leading order, the thickness perturbation is directly proportional to and in phase with the velocity perturbation.

38.5.4 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

$$\overline{h' v'} = \frac{1}{T} \int_0^T h' v' dt \quad (38.145a)$$

$$= \frac{h_o}{c T} \int_0^T v' v' dt \quad (38.145b)$$

$$= \frac{v_o^2 h_o}{c T} \int_0^T \sin^2(ky - \omega t) dt \quad (38.145c)$$

$$= \frac{v_o^2 h_o}{2 c} \quad (38.145d)$$

$$= h_o V_{\text{Stokes}}, \quad (38.145e)$$

where we introduced the Stokes drift (38.140) to reach the final equality. A nonzero correlation $\overline{h' v'}$ means that the thickness has a nonzero tendency when averaged over a wave period.

38.5.5 Comments

The nonzero correlation in equation (38.145e) 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 that for more general waves and eddies moving through fluid layers, such as considered in the ensemble mean isopycnal layer transport discussed in Section 38.4.5. For the general case, a nonzero bolus velocity (38.116), 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, thus prompting certain authors to make the equality in general.

38.6 Mean tracer equation

We now include a tracer field to the ideal Boussinesq parcel 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.

38.6.1 Thickness weighted means

Equation (38.115) introduced a specific thickness weighted mean operator, which will prove to be quite useful when considering the mean tracer equation. In general, for any field Φ associated with a potential density layer σ , we define the decomposition into thickness weighted mean and deviation

$$\Phi(\sigma) = \hat{\Phi}(\sigma) + \Phi''(\sigma) \quad (38.146a)$$

$$= \frac{\overline{h \Phi}^{(\sigma)}}{\overline{h}^{(\sigma)}} + \Phi''. \quad (38.146b)$$

It follows by definition that

$$\overline{h \Phi''}^{(\sigma)} = 0. \quad (38.147)$$

38.6.2 Isopycnal mean thickness weighted tracer equation

When attaching a tracer to fluid parcels, each member of the ensemble satisfies the isopycnal tracer equation

$$\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla_{\sigma} C = 0. \quad (38.148)$$

Combining the tracer and thickness equations leads to the thickness weighted tracer equation

$$\frac{\partial(hC)}{\partial t} + \nabla_{\sigma} \cdot (h \mathbf{u} C) = 0. \quad (38.149)$$

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 to give

$$\partial_t[h(\hat{C} + C'')] + \nabla_{\sigma} \cdot [h(\hat{\mathbf{u}} + \mathbf{u}'')](\hat{C} + C'') = 0. \quad (38.150)$$

Taking an ensemble average over fluid elements with the same potential density, and using equation (38.147), yield the mean thickness weighted tracer equation

$$\partial_t (\bar{h}^{(\sigma)} \hat{C}) + \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \hat{C} \hat{\mathbf{u}}) = -\nabla_\sigma \cdot (\bar{h} C'' \bar{\mathbf{u}}''). \quad (38.151)$$

Now introduce the correlation

$$\bar{h} C'' \bar{\mathbf{u}}'' = \bar{h}^{(\sigma)} \widehat{C'' \mathbf{u}''} \quad (38.152)$$

(see equation (38.146b)), and recall that the mean thickness $\bar{h}^{(\sigma)}$ satisfies the mean thickness equation (38.114). These two points lead to the evolution equation for the mean thickness weighted tracer concentration

$$(\partial_t + \hat{\mathbf{u}} \cdot \nabla_\sigma) \hat{C} = -\frac{1}{\bar{h}^{(\sigma)}} \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \widehat{C'' \mathbf{u}''}). \quad (38.153)$$

38.6.3 Subgrid scale tracer transport tensor

The correlation between tracer and velocity found on the right-hand side of the mean thickness weighted tracer equation (38.153) is typically written in terms of a subgrid scale tracer transport tensor

$$\widehat{C'' \mathbf{u}''} = -\mathbb{J} \cdot \nabla_\sigma \hat{C}. \quad (38.154)$$

This definition leads to the evolution equation

$$(\partial_t + \hat{\mathbf{u}} \cdot \nabla_\sigma) \hat{C} = \frac{1}{\bar{h}^{(\sigma)}} \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \mathbb{J} \cdot \nabla_\sigma \hat{C}). \quad (38.155)$$

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 11.15. However, in addition to symmetric diffusion processes, this operator includes skewed fluxes that lead to skew diffusion as discussed in Section 38.3.2. Whereas the diffusive aspect is commonly parameterized as dianeutral diffusion and neutral diffusion (Section 39.1), there is no parameterization for the skewed correlations for use in ocean models. We comment further on this situation in Section 39.3.8.

38.6.4 Mean tracer transport beneath a density surface

It is useful to further elucidate the relevance of mean thickness weighted fields. For this purpose, consider the mean horizontal tracer transport occurring beneath a particular potential density surface $\sigma = \tilde{\sigma}$,

$$\overline{C}^\#(\bar{z}^{(\sigma)}) = \overline{\int_{-H}^{\bar{z}^{(\sigma)}+\xi} C \mathbf{u} dz}. \quad (38.156)$$

Setting tracer concentration to unity recovers the expression (38.97) for the TRM transport. Changing coordinates and making use of the tracer correlation tensor renders

$$\overline{C}^\#(\bar{z}^{(\sigma)}) = \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \overline{C \mathbf{u} h^{(\sigma)}} d\sigma \quad (38.157a)$$

$$= \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \bar{h}^{(\sigma)} d\sigma (\hat{C} \hat{\mathbf{u}} + \widehat{C'' \mathbf{u}''}) \quad (38.157b)$$

$$= \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \bar{h}^{(\sigma)} d\sigma (\hat{C} \hat{\mathbf{u}} - \mathbb{J} \cdot \nabla_\sigma \hat{C}) \quad (38.157c)$$

$$= \int_{-H}^{\bar{z}^{(\sigma)}} dz (\hat{C} \hat{\mathbf{u}} - \mathbb{J} \cdot \nabla_\sigma \hat{C}). \quad (38.157d)$$

Hence, the mean thickness weighted fields naturally appear when considering such physically interesting quantities as the mean horizontal transport of a tracer beneath the modified mean potential density surface.

38.6.5 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 again is related to the thickness weighted horizontal velocity via

$$\widehat{\mathbf{u}}((\sigma)) = \frac{\overline{h} \mathbf{u}^{(\sigma)}}{\overline{h}^{(\sigma)}} = \overline{\mathbf{u}}^{(\sigma)} + \frac{\overline{h'} \mathbf{u}'^{(\sigma)}}{\overline{h}^{(\sigma)}} = \overline{\mathbf{u}}^{(\sigma)} + \mathbf{u}^{\text{bolus}}. \quad (38.158)$$

In addition to the bolus velocity, it is necessary to parameterize the subgrid scale tracer transport tensor

$$\widehat{C'' \mathbf{u}''} = -\mathbb{J} \cdot \nabla_{\sigma} \widehat{C}, \quad (38.159)$$

which generally has symmetric (diffusive) and antisymmetric (stirring) components (Section 38.3).

For a geopotential coordinate description, equation (38.93) is used to relate thickness weighted mean fields, defined as a function of σ , and TRM fields, defined as a function of the mean vertical position of σ , to write for the tracer field

$$\widehat{C}(x, y, \tilde{\sigma}, t) = \overline{C}^{\#}(x, y, \overline{z}^{(\sigma)}, t). \quad (38.160)$$

Equation (38.160), and the developed formalism, leads to the mean field tracer equation in geopotential coordinates

$$\partial_t \overline{C}^{\#} = \nabla \cdot (\nabla \overline{C}^{\#} \wedge \overline{\Psi}^{\#}) + R(\overline{C}^{\#}), \quad (38.161)$$

where $R(\overline{C}^{\#})$ is the geopotential coordinate form of the mixing/stirring operator on the right-hand side of equation (38.155). Details for the transformation of the mixing/stirring operator from isopycnal to geopotential coordinates are provided in Section 11.15.

38.6.6 Comments

Much in this section follows from [Smith \(1999\)](#), [McDougall and McIntosh \(2001\)](#), and [Young \(2012\)](#), each of which focused on the hydrostatic primitive equations assuming a vertically stable buoyancy stratification. The paper by [Young \(2012\)](#) is a milestone in the literature as he succeeded in formulating the ensemble mean primitive equations in a form where only the thickness weighted (residual mean) velocity appears. Prior attempts failed due to their insufficient mathematical framework. 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.

Even so, realistic ocean general circulation models are not formulated as “residual mean” models. The key reason is that outside of the stably stratified interior, as in boundary layers, thickness weighted averaging is inappropriate. Instead, we need Eulerian averaged fields when formulating boundary layer closures (e.g., [Large et al., 1994](#)). [Young \(2012\)](#) thus provides a compelling method to decompose the flow into eddies and mean within the stably stratified interior. However, it is not sufficient to capture the full suite of flow regimes represented or parameterized by realistic ocean circulation models.

39

Subgrid scale tracer transport

We are here concerned with the general properties of parameterization of processes affecting tracer distributions. These parameterizations aim to summarize physical processes too small to observe and/or simulate and how they impact on the larger scales. Such parameterizations in the tracer equation generally take the form of subgrid scale advection and diffusion. This *subgrid scale parameterization problem* is broader and deeper than available from a single chapter. Instead, we here aim to synthesize a range of physical and mathematical topics associated with ocean tracers, offering a platform for further study of a vast and growing literature.

READER'S GUIDE FOR THIS CHAPTER

Tracers evolve according to the advection-diffusion equation discussed in Chapter 37 and further unpacked in the tracer kinematics Chapter 38.

- add section on anisotropic neutral diffusion
- add section on anisotropic GM

39.1	Parameterizing eddy induced tracer transport	600
39.1.1	Framework based on tracer variance cascade	600
39.1.2	Density changes from molecular diffusion	601
39.1.3	Mixing from small (or fine) scale processes	601
39.1.4	Mesoscale eddy-induced stirring and mixing	601
39.1.5	Synthesis	604
39.1.6	Lateral versus diapycnal diffusion	605
39.2	Quasi-Stokes induced tracer stirring	605
39.2.1	Gent and McWilliams skewson	606
39.2.2	Local adiabatic dissipation of potential energy	606
39.2.3	Meridional overturning streamfunction	608
39.2.4	Connection to form stress	609
39.2.5	Isopycnal thickness diffusion and GM	610
39.3	Neutral diffusion	612
39.3.1	Redi neutral diffusion	612
39.3.2	Small slope neutral diffusion	613
39.3.3	Neutral tangent plane neutral diffusion	613
39.3.4	Neutrality condition	614
39.3.5	Symmetry condition	615
39.3.6	GM skewson plus small slope neutral diffusion	615
39.3.7	Small slope neutral diffusion in generalized vertical coordinates	615
39.3.8	Comments	616
39.4	Cabbeling and thermobaricity	617
39.4.1	Basic manipulations	617
39.4.2	A tidy form	618
39.4.3	Cabbeling	619
39.4.4	Thermobaricity	619
39.4.5	Comments	620

39.1 Parameterizing eddy induced tracer transport

In this section, we present a theoretical framework commonly used for parameterizing tracer transport. In turn, we build on the discussion from Section 40.2 to further detail how *in situ* density evolves in the presence of eddy parameterizations.

39.1.1 Framework based on tracer variance cascade

In the presence of turbulent processes, tracer variance directly cascades to the small scales. This 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 8.5 of [Vallis \(2006\)](#)). At this scale, tracer gradients are sufficiently large in magnitude that molecular diffusion can readily act to dissipate tracer variance through irreversible diffusive mixing. Hence, tracer transport at scales larger than the Batchelor scale is dominated by 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, Δ , is generally much larger than the Batchelor scale.

39.1.2 Density changes from molecular diffusion

Ignoring cross-diffusion processes (see [IOC et al. \(2010\)](#), Section 2.5 of [Olbers et al. \(2012\)](#), and [Graham and McDougall \(2013\)](#) for discussion), the molecular diffusion of Θ and S lead to the material evolution equations

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot [\rho \kappa_\Theta \nabla \Theta] \quad (39.1a)$$

$$\rho \frac{DS}{Dt} = \nabla \cdot [\rho \kappa_S \nabla S], \quad (39.1b)$$

where $\kappa_\Theta > 0$ and $\kappa_S > 0$ are the molecular kinematic diffusivities for Θ and S , respectively. Following equation (40.28), we see that these molecular tracer fluxes lead to the material evolution of *in situ* density

$$\frac{D\rho}{Dt} - \frac{\omega}{c_s^2} = -\alpha \nabla \cdot (\rho \kappa_\Theta \nabla \Theta) + \beta \nabla \cdot (\rho \kappa_S \nabla S) \quad (39.2a)$$

$$= -\nabla \cdot [\rho (\kappa_\Theta \alpha \nabla \Theta - \kappa_S \beta \nabla S)] + \rho (\kappa_\Theta \nabla \Theta \cdot \nabla \alpha - \kappa_S \nabla S \cdot \nabla \beta). \quad (39.2b)$$

Density evolves from molecular tracer diffusion through the convergence of a buoyancy flux as well as through processes associated with the nonlinear equation of state that give rise to spatial dependence for α and β .

39.1.3 Mixing from small (or fine) scale processes

For a model grid scale, Δ , larger than the scale where gravity waves break and dissipate kinetic energy (i.e., tens to hundreds of metres), diffusion is commonly used to parameterize 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, breaking leewaves, etc.

Small scale mixing generally takes place in an isotropic manner. Its parameterization thus appears just as for isotropic molecular diffusion given by equation (39.1b), yet with a far larger eddy diffusivity $\kappa \gg \kappa_\Theta, \kappa_S$ that is a function of the flow. This eddy tracer diffusion dissipates tracer variance at the grid scale, and in turn it mixes *in situ* density according to

$$\frac{D\rho}{Dt} - \frac{\omega}{c_s^2} = -\alpha \nabla \cdot (\rho \kappa \nabla \Theta) + \beta \nabla \cdot (\rho \kappa \nabla S) \quad (39.3a)$$

$$= -\nabla \cdot [\rho \kappa (\alpha \nabla \Theta - \beta \nabla S)] + \rho \kappa (\nabla \Theta \cdot \nabla \alpha - \nabla S \cdot \nabla \beta). \quad (39.3b)$$

Note that since vertical stratification is generally much larger than horizontal stratification, the isotropic diffusion operator is commonly approximated by a vertical or diapycnal diffusion operator (see [McDougall et al. \(2014\)](#) for further discussion).

39.1.4 Mesoscale eddy-induced stirring and mixing

Stirring from turbulent scales smaller than the grid scale is commonly parameterized by an eddy-induced stirring velocity, \mathbf{v}^* . For mesoscale eddies, such parameterized stirring generally follows a variant of [Gent et al. \(1995\)](#). In addition, mixing is promoted by the direct cascade from stirring. This mixing is parameterized by a diffusion operator distinct from that used for the small scale mixing discussed in Section 39.1.3. The general form of the diffusion operator is inferred in this section.

We mathematically frame our discussion by introducing a second order subgrid scale transport tensor, \mathbb{M} , meant to parameterize both subgrid scale eddy stirring and eddy mixing. With this tensor, the evolution of salinity and Conservative Temperature takes the form

$$\rho \frac{DS}{Dt} = \nabla \cdot (\rho \mathbb{M} \cdot \nabla S) \quad (39.4a)$$

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot (\rho \mathbb{M} \cdot \nabla \Theta). \quad (39.4b)$$

Note that we use the same transport tensor for both S and Θ . 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.

As discussed in Chapter 37, it is useful to decompose the second order transport tensor into the sum of its symmetric and anti-symmetric components

$$\mathbb{M} = \mathbb{K} + \mathbb{A}. \quad (39.5)$$

The symmetric tensor, \mathbb{K} , gives rise to downgradient diffusion whereas the anti-symmetric tensor, \mathbb{A} , gives rise to skew-diffusion or eddy-induced advection.

Mesoscale eddy-induced stirring

The anti-symmetric tensor, \mathbb{A} , contributes to the parameterized transport according to

$$\nabla \cdot (\rho \mathbb{A} \cdot \nabla S) = \partial_m (\rho A^{mn} \partial_n S) \quad (39.6a)$$

$$= \partial_m (\rho A^{mn}) \partial_n S + \rho A^{mn} \partial_m \partial_n S \quad (39.6b)$$

$$= -\rho v^{*n} \partial_n S, \quad (39.6c)$$

where we made use of the Einstein index notation with repeated indices summed over their range $m, n = 1, 2, 3$, and where A^{mn} are the components to the anti-symmetric transport tensor \mathbb{A} . Additionally, we noted that

$$\rho A^{mn} \partial_m \partial_n S = 0 \quad (39.7)$$

since A^{mn} is anti-symmetric whereas $\partial_m \partial_n S$ is symmetric. Finally, we introduced a density-weighted eddy-induced velocity

$$\rho v^{*n} = -\partial_m (\rho A^{mn}). \quad (39.8)$$

Importantly, $\rho \mathbf{v}^*$ has a zero divergence, again due to anti-symmetry of A^{mn}

$$\nabla \cdot (\rho \mathbf{v}^*) = \partial_n (\rho v^{*n}) = -\partial_n \partial_m (\rho A^{mn}) = 0. \quad (39.9)$$

A zero-divergence for $\rho \mathbf{v}^*$ means that it contributes no mass sources or sinks to the fluid.¹

Transport from the anti-symmetric tensor thus adds a means to stir tracers due to unresolved eddy processes. The mathematical form of the stirring can be either through skew-diffusion or through advection (see Section 37.5). Choosing to make use of the advection form allows us to

¹For a Boussinesq fluid, the density factor is replaced by the constant reference density, ρ_0 , so that $\nabla \cdot \mathbf{v}^* = 0$ in the Boussinesq fluid. See section 7 of [Griffies and Greatbatch \(2012\)](#) for more details of the Boussinesq and non-Boussinesq forms for the parameterized eddy-induced transport.

combine the contribution from the anti-symmetric transport tensor with the resolved advection operator, thus resulting in a residual mean material transport equation

$$\rho \frac{D^\dagger S}{Dt} = \nabla \cdot (\rho \mathbb{K} \cdot \nabla S) \quad (39.10a)$$

$$\rho \frac{D^\dagger \Theta}{Dt} = \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta), \quad (39.10b)$$

where the residual mean material time derivative is given by

$$\frac{D^\dagger}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^\dagger \cdot \nabla \quad (39.11)$$

and the residual mean velocity is

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*. \quad (39.12)$$

Making use of the residual mean velocity then leads to the material evolution of the *in situ* density

$$\frac{D\rho}{Dt} - \frac{1}{c_s^2} \frac{Dp}{Dt} = -\rho \mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S) - \alpha \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbb{K} \cdot \nabla S), \quad (39.13)$$

which can be written in terms of the residual mean material time derivative

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = -\alpha \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbb{K} \cdot \nabla S). \quad (39.14)$$

Mesoscale eddy-induced diffusion

Transport from the symmetric tensor, \mathbb{K} , corresponds to diffusion so long as the tensor is positive semi-definite. The diffusion operator in the residual mean evolution equation (39.14) can be written

$$-\alpha \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbb{K} \cdot \nabla S) = \nabla \cdot [\rho \mathbb{K} \cdot (-\alpha \nabla \Theta + \beta \nabla S)] + \rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S, \quad (39.15)$$

so that the *in situ* density evolves according to

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = -\underbrace{\nabla \cdot [\rho \mathbb{K} \cdot (\alpha \nabla \Theta - \beta \nabla S)]}_{\text{conservative processes}} + \underbrace{\rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S}_{\text{sources from nonlinear EOS processes}}. \quad (39.16)$$

We now discuss the physical processes associated with the right hand side terms.

- NONLINEAR EQUATION OF STATE: A nonlinear equation of state is characterized by spatially dependent thermal expansion and haline contraction coefficients. Mixing of Θ 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 summary of these processes in Section 39.4.
- CONSERVATIVE PROCESSES AND NEUTRAL DIFFUSION: A linear equation of state is independent of pressure, so that the evolution equation (39.16) takes the form

$$\frac{D^\dagger \rho}{Dt} = -\nabla \cdot [\rho \mathbb{K} \cdot (\alpha \nabla \Theta - \beta \nabla S)]. \quad (39.17)$$

Under the residual mean transport, density remains materially constant in the absence of any diffusion. Additionally, it remains constant if the diffusive fluxes of Θ and S are density-compensated so that

$$\mathbb{K} \cdot (\alpha \nabla \Theta - \beta \nabla S) = 0. \quad (39.18)$$

As detailed in Section 39.3, various forms of this *neutral diffusion* satisfy this *neutrality condition* even for general equations of state that have pressure dependence.

We thus propose that the mesoscale eddy stirring-induced mixing of Θ and S be parameterized as density-compensated diffusion, otherwise known as neutral diffusion. Neutral diffusion leaves *in situ* density unchanged in the absence of nonlinear equation of state processes so that

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = \underbrace{\rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S}_{\text{sources from nonlinear EOS processes}}. \quad (39.19)$$

What is the evidence for a diffusion operator oriented according to neutral directions? To answer this question, consider a diffusion tensor that does *not* maintain the constraint (39.18). In this case, additional diffusive mixing appears, adding to that already parameterized from small scale mixing processes such as breaking gravity waves. 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 diffusive fluxes is proportional to the squared tangent of the angle between the proposed new direction and the neutral tangent plane. Estimates for interior ocean mixing constrain the magnitude of the tangent to be less than 10^{-4} . This number is very small, indeed it is zero within error bars of field measurements. Measurements thus support the use of a neutral diffusion operator oriented so to respect the constraint (39.18). In Section 39.3 we dive into the details of neutral diffusion.

39.1.5 Synthesis

In summary, the proposed evolution equation for *in situ* density in the presence of subgrid scale processes takes the form

$$\frac{D\rho}{Dt} - \underbrace{\frac{1}{c_s^2} \frac{Dp}{Dt}}_{\text{compressibility}} = -\underbrace{\mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S)}_{\text{eddy-induced stirring}} - \underbrace{\alpha \nabla \cdot (\rho \kappa \nabla \Theta) + \beta \nabla \cdot (\rho \kappa \nabla S)}_{\text{small scale mixing and nonlinear EOS}} + \underbrace{\rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S}_{\text{nonlinear EOS processes from eddy mixing}}. \quad (39.20)$$

In summary, we have the following physical processes contributing to the evolution of *in situ* density.

- **SMALL SCALE MIXING:** To parameterize mixing induced by the suite of subgrid small scale processes (e.g., breaking gravity waves, lee waves, turbulent boundary layer processes, double diffusion) we introduce an isotropic diffusion operator (39.3b) with an eddy-diffusivity, κ , that is a function of the flow. This 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 horizontal, it is common to approximate the isotropic diffusion operator with a vertical diffusion operator.
- **EDDY-INDUCED STIRRING:** For subgrid scale stirring, such as from mesoscale (and submesoscale) eddies, we introduce a parameterized eddy-induced advection operator. When combined with the resolved advection, we are led to a residual mean material time derivative, D^\dagger/Dt .
- **EDDY-INDUCED MIXING:** Subgrid scale eddy-induced stirring leads to a direct cascade of Θ and S variance to the small scales. Mixing arising from this cascade is parameterized by neutral diffusion, whereby the diffusive fluxes of Θ and S are density compensated according to the constraint (39.18).

- NONLINEAR EOS PROCESSES: Mixing of Θ and S in the presence of a nonlinear equation of state means that *in situ* density evolves due to cabbeling and thermobaricity (Section 39.4). The dominant contributions to these processes arise from eddy-stirring induced mixing (i.e., neutral diffusion) ([McDougall, 1987b](#)), though small scale mixing also has a contribution as seen by writing

$$-\alpha \nabla \cdot (\rho \kappa \nabla \Theta) + \beta \nabla \cdot (\rho \kappa \nabla S) = -\nabla \cdot [\rho \kappa (\alpha \nabla \Theta - \beta \nabla S)] + \rho \kappa (\nabla \alpha \cdot \nabla \Theta - \nabla \beta \cdot \nabla S). \quad (39.21)$$

39.1.6 Lateral versus diapycnal diffusion

What is more important for setting tracer distributions: lateral or diapycnal diffusion? Although the lateral eddy diffusivity is many orders larger than the eddy diapycnal diffusivity, the gradients on which they act are very different. So to help answer the question, consider a scaling in which we consider a constant coefficient lateral diffusivity and a constant coefficient isotropic diffusivity. Furthermore, to simplify the analysis assume Cartesian orientation of the diffusion operators and assume the isotropic diffusion is dominated by vertical diffusion

$$\text{horizontal diffusion} = \kappa_h \nabla_z^2 C \quad \text{vertical diffusion} = \kappa_v \partial_{zz} C. \quad (39.22)$$

Now introduce a vertical scale H and horizontal scale L over which the tracer concentration changes by δC . Doing so leads to the scaled diffusion operators

$$\text{horizontal diffusion} \sim (\kappa_h/L^2) \delta C \quad \text{vertical diffusion} \sim (\kappa_v/H^2) \delta C. \quad (39.23)$$

These operators have the same scale when

$$\kappa_v = (H/L)^2 \kappa_h. \quad (39.24)$$

Choosing $L = 10^5$ m and $H = 10^1$ m leads to

$$\kappa_v = 10^{-8} \kappa_h. \quad (39.25)$$

Furthermore, if $\kappa_h = 10^3$ m² s⁻¹, then the two operators provide a similar contribution to tracer evolution if $\kappa_v = 10^{-5}$ m² s⁻¹. 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](#)).² This scaling is crude since the length scales are dependent on details of the flow regime as are the eddy diffusivities. Even so, the scaling indicates that even a relatively small turbulent diapycnal diffusivity arising from the background gravity wave spectrum can contribute to tracer distributions a similar amount as from lateral diffusion.

39.2 Quasi-Stokes induced tracer stirring

As mentioned in Section 38.6.3, there are two processes that contribute to eddy-induced stirring. One involves the correlations between eddy fluctuations in the velocity and tracer fields. In Section 38.3, 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. There is currently no method

²Molecular diffusivities are roughly $\approx 10^{-6}$ m² s⁻¹ for temperature and $\approx 10^{-7}$ m² s⁻¹ for salinity and other material tracers. (SMG: check these values).

available 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 38.4, 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 stirring in turn affects an eddy-induced tracer stirring within isopycnal layers. Transient mesoscale eddies are the canonical dynamical process leading to this form of stirring. 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. Elements of this scheme are detailed in this section.

In this section we refer to ρ as the potential density, with the assumption of a linear equation of state so that buoyancy equals to potential density (Section 25.2.6). This assumption is for convenience only, and can be readily generalized to an arbitrary equation of state.

39.2.1 Gent and McWilliams skewson

[Gent et al. \(1995\)](#) parameterize the divergent part of the quasi-Stokes transport by setting

$$\mathbf{U}^{\text{qs}} = -\kappa \mathbf{S}. \quad (39.26)$$

In this expression, \mathbf{S} is the slope of the potential density surfaces (equation (38.119)), and $\kappa > 0$ is a kinematic diffusivity (dimensions of velocity times a length). The corresponding three-dimensional non-divergent eddy-induced velocity is given by

$$\mathbf{v}^* = -\partial_z(\kappa \mathbf{S}) + \hat{\mathbf{z}} \nabla_z \cdot (\kappa \mathbf{S}), \quad (39.27)$$

and the antisymmetric stirring tensor is

$$\mathbb{A}_{mn} = \begin{bmatrix} 0 & 0 & -\kappa S_x \\ 0 & 0 & -\kappa S_y \\ \kappa S_x & \kappa S_y & 0 \end{bmatrix}. \quad (39.28)$$

The parameterized skew flux of potential density, ρ , due to the quasi-Stokes transport is given by

$$\mathbf{F}^{\text{skew}} = -\mathbf{U}^{\text{qs}} \partial_z \rho + \hat{\mathbf{z}} (\mathbf{U}^{\text{qs}} \cdot \nabla_z) \rho \quad (39.29a)$$

$$= -\kappa \nabla_z \rho + \hat{\mathbf{z}} S^2 \kappa \partial_z \rho \quad (39.29b)$$

$$= -\kappa \nabla_z \rho - \hat{\mathbf{z}} (\kappa \rho_0/g) (S N)^2. \quad (39.29c)$$

This parameterization yields horizontal downgradient diffusion of potential density, combined with a vertical upgradient diffusion. So long as the stratification is stable ($N^2 > 0$), the vertical component to the skew flux is vertically downward. Additionally, [Gent et al. \(1995\)](#) prescribe a diffusivity that vanishes on all boundaries, including the ocean surface. [McIntosh and McDougall \(1996\)](#) and [McDougall and McIntosh \(2001\)](#) present more discussion of vertical boundary conditions, which can be understood by considering the exact form of the quasi-Stokes transport defined by equation (38.100).

39.2.2 Local adiabatic dissipation of potential energy

We here consider the effects from the [Gent et al. \(1995\)](#) scheme on the potential energy and available potential energy (APE). We express the behavior using both skew fluxes and advective fluxes. Note that since we are assuming the parameterization is adiabatic, the change in potential energy is identical to the change in available potential energy (APE).

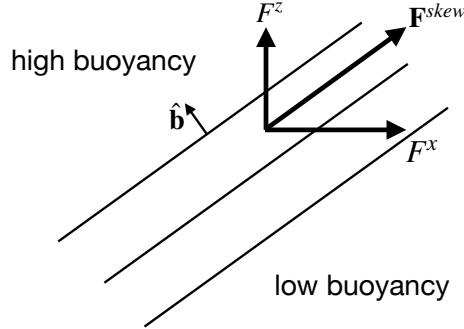


Figure 39.1: Orientation of the skew flux for buoyancy as proposed by Gent et al. (1995) and described by Griffies (1998), where buoyancy is $b = -(g/\rho_0)(\rho - \rho_0)$ so that $\nabla b = -(g/\rho_0)\nabla\rho$. The horizontal flux is downgradient (moving high buoyancy water to low buoyancy) whereas the vertical flux component is upgradient. The net effect is a flux that is oriented parallel to constant buoyancy lines (i.e., skewed relative to the buoyancy gradient).

Skew flux approach

Let us approach the parameterization problem from the perspective of satisfying two general properties: (I) the subgrid scale operator adiabatically stirs while maintaining the same amount of fluid within isopyncal layers, (II) the subgrid operator locally dissipates potential energy through an adiabatic rearrangement of the density surfaces, with the 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?

Adiabatic stirring of potential density can be realized via the convergence of a skew flux oriented parallel to potential density surfaces

$$\mathbf{F} = -\nabla\rho \wedge \Psi, \quad (39.30)$$

where Ψ remains to be found. To see what the local dissipation of APE imposes, consider the gravitational potential energy of the adiabatic Boussinesq system

$$P = g \int \rho z \, dV, \quad (39.31)$$

where we assume the *in situ* density equals to the potential density as per a linear equation of state (Section 25.2.6). Assuming all boundaries are material and static allows us to focus on the time tendency of potential energy associated with the unknown flux

$$\frac{dP}{dt} = g \int z \frac{\partial \rho}{\partial t} \, dV \quad (39.32a)$$

$$= -g \int (z \nabla \cdot \mathbf{F}) \, dV \quad (39.32b)$$

$$= -g \int (z \partial_z F^{(z)}) \, dV \quad (39.32c)$$

$$= g \int F^{(z)} \, dV, \quad (39.32d)$$

where $F^{(z)}$ is the vertical flux component. We drop boundary effects by assuming the subgrid scale flux vanishes on all boundaries. To provide a *local* APE sink requires

$$F^{(z)} \leq 0, \quad (39.33)$$

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 $\partial_z \rho < 0$, the following form provides a local APE sink

$$F^{(z)} = \kappa S^2 \frac{\partial \rho}{\partial z} = -(\kappa \rho_0/g) (S N)^2 \leq 0, \quad (39.34)$$

where $\kappa > 0$ is a kinematic diffusivity setting the strength of the flux and S^2 is the squared isopycnal slope. The corresponding horizontal flux is given by a downgradient diffusive flux

$$\mathbf{F}^{(h)} = -\kappa \nabla_z \rho. \quad (39.35)$$

We have thus recovered the skew flux (39.29c) 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 APE sink, thus deriving a more general subgrid scale operator.

Advection 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 (39.32d) and make use of the vertical component of the advective flux rather than the skew flux

$$\frac{dP}{dt} = g \int F^{(z)} dV \quad (39.36a)$$

$$= g \int \rho w^* dV \quad (39.36b)$$

$$= g \int \rho \nabla_z \cdot (\kappa \mathbf{S}) dV \quad (39.36c)$$

$$= g \int \nabla_z \cdot (\rho \kappa \mathbf{S}) dV - g \int \nabla_z \rho \cdot \kappa \mathbf{S} dV \quad (39.36d)$$

$$= -\rho_0 \int \kappa (S N)^2 dV, \quad (39.36e)$$

which is the same result as for the skew flux.

39.2.3 Meridional overturning streamfunction

It is often of interest to compute the net transport of volume across a portion of the ocean. In particular, meridional-depth or meridional-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 transport of fluid across a basin and passing beneath a particular depth in the form (the minus sign is conventional)

$$\mathcal{T}(y, z, t) = - \int dx \int_{-H}^z (v + v^*) dz \quad (39.37a)$$

$$= - \int dx \int_{-H}^z v dz + \int dx \int_{-H}^z \partial_z(\kappa S_y) dz \quad (39.37b)$$

$$= - \int dx \int_{-H}^z v dz + \int \kappa S_y dx \quad (39.37c)$$

$$\equiv \mathcal{T}^{\text{eulerian}}(y, z, t) + \mathcal{T}^{\text{qS}}(y, z, t). \quad (39.37d)$$

For the penultimate step we set the parameterized quasi-Stokes transport to zero at the ocean bottom. We thus see that the parameterized quasi-Stokes transport adds a contribution that scales linearly with basin size, isopyncal slope, and diffusivity,

$$\mathcal{T}^{\text{qs}} \sim L S \kappa. \quad (39.38)$$

As an example, let $\kappa = 10^3 \text{ m}^2 \text{s}^{-1}$, $S = 10^{-3}$, and $L = 10^7 \text{ m}$, which yields $\mathcal{T} \approx 10 \times 10^6 \text{ m}^3 \text{s}^{-1} = 10 \text{ Sv}$. Such transport can represent a nontrivial addition to that from the resolved scale velocity field.

39.2.4 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 anticipate our general discussion of form stress in Section 26.8 and more detailed discussions in Sections 29.5 and 43.2. In those discussions, we identify form stress as the horizontal pressure force acting on a sloped surface, with our present concern with surfaces of constant isopycnals as discussed in Sections 29.5 and 43.2.

[Young \(2012\)](#) provides a general means for making the connection between [Gent et al. \(1995\)](#) and form stress for a continuously stratified fluid. 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). For this purpose, assume the fluid is in Boussinesq planetary geostrophic balance (detailed in Section 29.4) whereby the horizontal momentum satisfies

$$\rho_o f (\hat{\mathbf{z}} \wedge \mathbf{u}) = -\nabla_z p + \partial_z \boldsymbol{\tau}. \quad (39.39)$$

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 surface where it arises from turbulent air-sea interactions; i.e., wind stress. Furthermore, it can be large at the bottom through turbulent bottom stresses.

To make the connection between [Gent et al. \(1995\)](#) and the vertical transfer of horizontal form stress, add $\rho_o f (\hat{\mathbf{z}} \wedge \mathbf{u}^*)$ to both sides of equation (39.39) to obtain

$$\rho_o f (\hat{\mathbf{z}} \wedge \mathbf{u}^\dagger) = -\nabla_z p + \partial_z \boldsymbol{\tau} + \rho_o f (\hat{\mathbf{z}} \wedge \mathbf{u}^*), \quad (39.40)$$

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, vertical 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 (Section 29.4.5), whereby

$$f \frac{\partial \mathbf{u}}{\partial z} = -(g/\rho_0) \hat{\mathbf{z}} \wedge \nabla \rho = -\hat{\mathbf{z}} \wedge N^2 \mathbf{S}, \quad (39.41)$$

with

$$\mathbf{S} = -\frac{\nabla_z \rho}{\partial_z \rho} = \frac{g}{\rho_0} \frac{\nabla_z \rho}{N^2} \quad (39.42)$$

the isopycnal slope. We can thus write the Coriolis acceleration from the eddy-induced velocity in the form

$$f(\hat{\mathbf{z}} \wedge \mathbf{u}^*) = -f[\hat{\mathbf{z}} \wedge \partial_z(\kappa \mathbf{S})] \quad (39.43a)$$

$$= -\partial_z[\hat{\mathbf{z}} \wedge (f \kappa \mathbf{S})] \quad (39.43b)$$

$$= \frac{\partial}{\partial z} \left[\frac{\kappa f^2}{N^2} \frac{\partial \mathbf{u}}{\partial z} \right] \quad (39.43c)$$

$$= \frac{\partial}{\partial z} \left[\nu_e \frac{\partial \mathbf{u}}{\partial z} \right], \quad (39.43d)$$

where the final equality introduced an eddy-induced vertical viscosity

$$\nu_e \equiv \kappa(f^2/N^2). \quad (39.44)$$

Making use of this result in the planetary geostrophic equation (39.40) thus leads to

$$\rho_0 f(\hat{\mathbf{z}} \wedge \mathbf{u}^\dagger) = -\nabla_z p + \partial_z(\boldsymbol{\tau} + \boldsymbol{\tau}_e), \quad (39.45)$$

where

$$\boldsymbol{\tau}_e = \rho_0 \nu_e \frac{\partial \mathbf{u}}{\partial z} \quad (39.46)$$

defines a horizontal mesoscale eddy stress arising from the thermal wind shears. Equation (39.45) 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 determined by a viscosity $\nu_e = \kappa(f/N)^2$. 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 discussed in Section 29.5.

39.2.5 Isopycnal thickness diffusion and GM

Recall the ensemble mean thickness equation (38.114) derived in Section 38.4.5

$$\partial_t h + \nabla_\rho \cdot (h \hat{\mathbf{u}}) = 0, \quad (39.47)$$

where

$$\hat{\mathbf{u}} = \mathbf{u} + \mathbf{u}^{\text{bolus}} \quad (39.48)$$

is the thickness weighted transport velocity affecting evolution of the ensemble mean thickness h . Note that for brevity we here drop the nomenclature $(\)^{(\rho)}$ used in Section 38.4.5.

Isopycnal correlations of horizontal velocity and layer thickness define the bolus velocity via

$$h \mathbf{u}^{\text{bolus}} = \overline{h' \mathbf{u}'} \quad (39.49)$$

Now consider a downgradient diffusive closure for this correlation

$$h \mathbf{u}^{\text{bolus}} = \overline{h' \mathbf{u}'^{(\rho)}} \quad (39.50\text{a})$$

$$= -\mathbf{K} \cdot \nabla_\rho h \quad (39.50\text{b})$$

with \mathbf{K} a symmetric and positive-definite 2×2 diffusion tensor. The mean thickness equation thus takes the form of an advection-diffusion equation in isopycnal coordinates

$$\partial_t h + \nabla_\rho \cdot (h \mathbf{u}) = \nabla_\rho \cdot (\mathbf{K} \cdot \nabla_\rho h). \quad (39.51)$$

To make a connection between the thickness diffusion closure (39.50b) and the [Gent et al. \(1995\)](#) closure discussed in Section 39.2.1, note that the specific thickness is the inverse of the vertical derivative of the potential density

$$h = (\partial_z \rho)^{-1}. \quad (39.52)$$

Correspondingly, using the relation between derivative operators, $\nabla_\rho = \nabla_z + \mathbf{S} \partial_z$, gives

$$h^{-1} \nabla_\rho h = -h \nabla_\rho (1/h) \quad (39.53\text{a})$$

$$= -(\partial_z \rho)^{-1} (\nabla_z + \mathbf{S} \partial_z) \partial_z \rho \quad (39.53\text{b})$$

$$= -\frac{\partial_z (\nabla_z \rho)}{\partial_z \rho} + \frac{\partial_{zz} \rho \nabla_z \rho}{(\partial_z \rho)^2} \quad (39.53\text{c})$$

$$= \partial_z \mathbf{S}. \quad (39.53\text{d})$$

Consequently, the bolus velocity takes the form

$$\mathbf{u}^{\text{bolus}} = -h^{-1} \mathbf{K} \cdot \nabla_\rho h = -\mathbf{K} \cdot \partial_z \mathbf{S}. \quad (39.54)$$

The special case of depth independent diffusivity

For the special case where \mathbf{K} is independent of depth and proportional to the 2×2 identity matrix, then

$$\mathbf{u}^{\text{bolus}} = -\partial_z (\kappa \mathbf{S}) = \mathbf{u}^*, \quad (39.55)$$

where the horizontal component of the [Gent et al. \(1995\)](#) velocity \mathbf{u}^* was identified from equation (39.27). 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. \(2000a\)](#), 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.

39.3 Neutral diffusion

Neutral diffusion parameterizes the mixing induced by mesoscale eddy stirring acting preferentially along neutral directions. By construction, the neutral diffusive flux of a tracer is oriented along a neutral direction or a *neutral tangent plane*. As detailed in Section 25.3, neutral directions are directions in a stratified fluid that allow for mixing of Θ and S without modifying the locally defined buoyancy. The neutral diffusive tracer flux for an arbitrary tracer, C , is perpendicular to the dianeutral unit vector

$$\mathbf{J} \cdot \hat{\gamma} = 0 \quad (39.56)$$

where (equation (25.42))

$$\hat{\gamma} = \frac{\rho_\theta \nabla \theta + \rho_S \nabla S}{|\rho_\theta \nabla \theta + \rho_S \nabla S|} = \frac{-\alpha \nabla \Theta + \beta \nabla S}{|-\alpha \nabla \Theta + \beta \nabla S|}. \quad (39.57)$$

39.3.1 Redi neutral diffusion

One diffusive flux satisfying the property (39.56) is given by

$$\mathbf{J}^{\text{redi}} = -\rho A [\nabla C - \hat{\gamma} (\hat{\gamma} \cdot \nabla C)], \quad (39.58)$$

where $A > 0$ is the neutral diffusivity (dimensions of squared length per time). We confirm that \mathbf{J}^{redi} is oriented down the tracer gradient by noting that

$$\mathbf{J}^{\text{redi}} \cdot \nabla C = -\rho A [|\nabla C|^2 - (\hat{\gamma} \cdot \nabla C)^2] \leq 0. \quad (39.59)$$

The flux \mathbf{J}^{redi} is precisely that resulting from the neutral diffusion tensor of [Redi \(1982\)](#) (see also Section 14.1.6 of [Griffies \(2004\)](#)), as can be seen by writing the flux as the product

$$J^m = -\rho K^{mn} \partial_n C, \quad (39.60)$$

where the Redi diffusion tensor is given by

$$K^{mn} = \frac{A}{1 + S_x^2 + S_y^2} \begin{bmatrix} 1 + S_y^2 & -S_x S_y & S_x \\ -S_x S_y & 1 + S_x^2 & S_y \\ S_x & S_y & S_x^2 + S_y^2 \end{bmatrix} \quad (39.61)$$

with the corresponding neutral diffusion operator given by the three-dimensional flux convergence

$$\mathcal{R}^{\text{redi}} = -\nabla \cdot \mathbf{J}^{\text{redi}} = \partial_m (\rho K^{mn} \partial_n C). \quad (39.62)$$

In the Redi tensor (39.61) we introduced the components of the horizontal vector, $\mathbf{S} = (S_x, S_y, 0)$, with

$$\mathbf{S} = - \begin{bmatrix} -\alpha \nabla_z \Theta + \beta \nabla_z S \\ -\alpha \partial_z \Theta + \beta \partial_z S \end{bmatrix} = \frac{g (-\alpha \nabla_z \Theta + \beta \nabla_z S)}{N^2} \quad (39.63)$$

the slope of the neutral tangent plane relative to the (x, y) horizontal plane, with

$$N^2 = -\frac{g}{\rho_0} \frac{\partial \gamma}{\partial z} = -g [-\alpha \partial_z \Theta + \beta \partial_z S] \quad (39.64)$$

the squared buoyancy frequency (Section 25.3). Notably, it is useful to introduce the slope vector only when the fluid is stratified in the vertical so that the slope magnitude, $|\mathbf{S}|$, is finite.

39.3.2 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 most of the ocean interior even in frontal regions. With this approximation, the small slope neutral diffusion tensor takes the form

$$(K^{\text{small}})^{mn} = A \begin{bmatrix} 1 & 0 & S_x \\ 0 & 1 & S_y \\ S_x & S_y & S_x^2 + S_y^2 \end{bmatrix}. \quad (39.65)$$

The corresponding small slope neutral diffusive flux is

$$\mathbf{J}^{\text{small}} = -\rho A [\nabla_\gamma + \hat{\mathbf{z}} (\mathbf{S} \cdot \nabla_\gamma)] C \quad (39.66)$$

where

$$\nabla_\gamma = \nabla_z + \mathbf{S} \partial_z \quad (39.67)$$

is the horizontal derivative operator computed on the neutral tangent plane (see equation (11.64)). To show that $\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = 0$, we make use of the identity

$$\hat{\gamma} = \frac{\mathbf{S} - \hat{\mathbf{z}}}{(1 + \mathbf{S} \cdot \mathbf{S})^{1/2}}, \quad (39.68)$$

so that

$$\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = \frac{\mathbf{J}^{\text{small}} \cdot \mathbf{S} - \mathbf{J}^{\text{small}} \cdot \mathbf{S}}{(1 + \mathbf{S} \cdot \mathbf{S})^{1/2}} = 0. \quad (39.69)$$

Furthermore, we confirm that $\mathbf{J}^{\text{small}}$ is oriented down the tracer gradient by noting that

$$\mathbf{J}^{\text{small}} \cdot \nabla C = -\rho A [\nabla_\gamma C \cdot \nabla_z C + (\mathbf{S} \cdot \nabla_\gamma C) \partial_z C] \quad (39.70a)$$

$$= -\rho A [|\nabla_z C|^2 + 2(\mathbf{S} \cdot \nabla_z C) \partial_z C + |\mathbf{S} \partial_z C|^2] \quad (39.70b)$$

$$= -\rho A |\nabla_z C + \mathbf{S} \partial_z C|^2 \quad (39.70c)$$

$$= -\rho A |\nabla_\gamma C|^2 \quad (39.70d)$$

$$\leq 0. \quad (39.70e)$$

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 (39.66) 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

$$\mathcal{R}^{\text{small}} = -\nabla \cdot \mathbf{J}^{\text{small}} = \nabla_z \cdot (\rho A \nabla_\gamma C) + \partial_z (\rho A \mathbf{S} \cdot \nabla_\gamma C). \quad (39.71)$$

39.3.3 Neutral tangent plane neutral diffusion

A third method to compute neutral diffusion is motivated by the form of isopycnal diffusion in isopycnal layered models. 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 11.

Following the derivations given in Section 11.15, the neutral diffusive flux in the neutral tangent frame is given by the horizontal flux

$$\mathbf{J}^{\text{ntp}} = -\rho A \nabla_\gamma C. \quad (39.72)$$

This flux is oriented down the tracer gradient as oriented along neutral directions

$$\mathbf{J}^{\text{ntp}} \cdot \nabla_{\gamma} C = -\rho A |\nabla_{\gamma} C|^2, \quad (39.73)$$

which is the same as equation (39.70d) for the small slope fluxes. However, as a purely horizontal flux, \mathbf{J}^{ntp} is not oriented along neutral directions

$$\mathbf{J}^{\text{ntp}} \cdot \hat{\gamma} \neq 0. \quad (39.74)$$

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}^{ntp} along the neutral tangent plane as per equation (11.84)

$$\mathcal{R}^{\text{ntp}} = -\frac{1}{h^{\gamma}} [\nabla_{\gamma} \cdot (h^{\gamma} \mathbf{J}^{\text{ntp}})] = \frac{1}{h^{\gamma}} [\nabla_{\gamma} \cdot (h^{\gamma} \rho A \nabla_{\gamma} C)], \quad (39.75)$$

where

$$h^{\gamma} = \frac{\partial z}{\partial \gamma} d\gamma = - \left[\frac{g}{\rho_0 N^2} \right] d\gamma \quad (39.76)$$

measures the thickness of a layer defined by two neutral tangent planes (see equation (11.81)).

As detailed in Section 11.15, \mathcal{R}^{ntp} is identical to the small slope neutral diffusion operator (39.71)

$$\mathcal{R}^{\text{ntp}} = \mathcal{R}^{\text{small}}. \quad (39.77)$$

In principle, it is a matter of convenience which form f 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}^{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. \(2012\)](#)). 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 37.3). 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.

39.3.4 Neutrality condition

Given the expression (39.57) for the dianeutral unit vector, $\hat{\gamma}$, it is straightforward to show that the neutral diffusive flux for Conservative Temperature balances that for salinity

$$\alpha \mathbf{J}(\Theta) = \beta \mathbf{J}(S). \quad (39.78)$$

We refer to this balance as the *neutrality condition*. It reflects the vanishing of the neutral diffusive flux when acting on locally referenced potential density. It is maintained by the diffusive flux (39.58) of [Redi \(1982\)](#), the small slope flux (39.66) of [Gent and McWilliams \(1990\)](#), and the neutral tangent frame neutral diffusive flux (39.72). However, it is not maintained by the small slope fluxes from [Cox \(1987\)](#). Indeed, [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 \(1987\)](#).

39.3.5 Symmetry condition

Since the neutral diffusion tensor is symmetric (as are all diffusion tensors; see Section 37.3), we have

$$\mathbf{J}(\Theta) \cdot \nabla S = -A \rho K^{mn} \partial_n \Theta \partial_m S \quad (39.79a)$$

$$= -A \rho K^{nm} \partial_n S \partial_m \Theta \quad (39.79b)$$

$$= -A \rho K^{mn} \partial_n S \partial_m \Theta \quad (39.79c)$$

$$= \mathbf{J}(S) \cdot \nabla \Theta. \quad (39.79d)$$

This symmetry condition will be useful in our discussion of cabbeling and thermobaricity in Section 39.4.

39.3.6 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 39.2.1) and small slope neutral diffusion (Section 39.3.2). The combined tracer flux takes the form

$$\mathbf{F} = -A \nabla_z C + (\kappa - A) \mathbf{S} \partial_z C - \hat{\mathbf{z}} [(A + \kappa) \mathbf{S} \cdot \nabla_z C + A S^2 \partial_z C], \quad (39.80)$$

which can be written in terms of a subgrid scale transport tensor (*Griffies*, 1998)

$$\begin{bmatrix} F^{(x)} \\ F^{(y)} \\ F^{(z)} \end{bmatrix} = \begin{bmatrix} A & 0 & (A - \kappa) S_x \\ 0 & A & (A - \kappa) S_y \\ (A + \kappa) S_x & (A + \kappa) S_y & A S^2 \end{bmatrix} \begin{bmatrix} \partial_x C \\ \partial_y C \\ \partial_z C \end{bmatrix}. \quad (39.81)$$

In the 1990s and throughout much of the 2000s, it was common to assume that $A = \kappa$, in which case the combined subgrid scale flux simplifies to

$$\mathbf{F} = -\kappa \nabla_z C - \hat{\mathbf{z}} \kappa (2 \mathbf{S} \cdot \nabla_z C + S^2 C_{,z}). \quad (39.82)$$

Notably, the 2×2 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 compelling and 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), and *Ferrari et al.* (2010), the boundary conditions for neutral diffusion and GM skewusion 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 and the other parameterizes downgradient diffusion along neutral directions.

39.3.7 Small slope neutral diffusion in 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 the generalized vertical coordinates (GVCs) detailed in Chapters 11 and 21. This formulation is relevant for the now common use of generalized vertical coordinates for ocean modeling, with an example algorithm discussed in Section 33.4.

We start by recalling the expression (11.83) for a general diffusion operator written in terms of the generalized vertical coordinate, $\sigma = \sigma(x, y, z, t)$

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \delta_\sigma(z_\sigma \nabla_\sigma \cdot \mathbf{J}) \right], \quad (39.83)$$

where $\delta_\sigma \equiv d\sigma/d\sigma$ is the dimensionless derivative operator, and the thickness of a σ -layer is

$$h^\sigma = dz = z_\sigma d\sigma = \frac{\partial z}{\partial \sigma} d\sigma. \quad (39.84)$$

Now assume the flux, \mathbf{J} , is given by equation (39.66) for small slope neutral diffusion. Transforming to GVCs leads to the horizontal flux component

$$\mathbf{J}^{\text{small, h}} = -\rho A \nabla_\gamma C \quad (39.85a)$$

$$= -\rho A [\nabla_z + (\nabla_\gamma z) \partial_z] C \quad (39.85b)$$

$$= -\rho A [\nabla_\sigma + (-\nabla_\sigma z + \nabla_\gamma z) \partial_z] C \quad (39.85c)$$

$$= -\rho A [\nabla_\sigma + (-\mathbf{S}^{(\sigma/z)} + \mathbf{S}^{(\gamma/z)}) \partial_z] C \quad (39.85d)$$

$$= -\rho A (\nabla_\sigma + \mathbf{S}^{(\gamma/\sigma)} \partial_z) C, \quad (39.85e)$$

where the neutral slopes as shown in Figure 39.2 satisfy the identity

$$\mathbf{S}^{(\sigma/z)} = \mathbf{S}^{(\gamma/z)} - \mathbf{S}^{(\gamma/\sigma)}. \quad (39.86)$$

Furthermore, we made use of the identity (11.64) relating the partial derivative operators

$$\nabla_\gamma = \nabla_z + (\nabla_\gamma z) \partial_z \quad \nabla_z = \nabla_\sigma - (\nabla_\sigma z) \partial_z. \quad (39.87)$$

The horizontal flux (39.85e) has the same form as when written using geopotential coordinates, only now with the derivative operator ∇_σ and the slope $\mathbf{S}^{(\gamma/\sigma)}$. Correspondingly, the vertical flux component

$$J^{\text{small, z}} = \mathbf{J}^{\text{small, h}} \cdot \mathbf{S}^{(\gamma/z)} \quad (39.88)$$

takes the form

$$z_\sigma \nabla_\sigma \cdot \mathbf{J}^{\text{small}} = -\mathbf{S}^{(\sigma/z)} \cdot \mathbf{J}^{\text{small, h}} + J^{\text{small, z}} = \mathbf{J}^{\text{small, h}} \cdot \mathbf{S}^{(\gamma/\sigma)}, \quad (39.89)$$

which in turn yields the diffusion operator (39.83)

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^{\text{small, h}}) + \delta_\sigma(\mathbf{J}^{\text{small, h}} \cdot \mathbf{S}^{(\gamma/\sigma)}) \right]. \quad (39.90)$$

In the special case when σ is parallel to the neutral direction so that $\mathbf{S}^{(\gamma/\sigma)} = 0$, the diffusion operator (39.90) reduces to the neutral tangent plane version given by equation (39.75).

39.3.8 Comments

As noted in Section 38.6.3, 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. To parameterize the skew fluxes arising from tracer-velocity correlations requires one to study the polarization of the eddies giving rise to the skew flux, as per the discussion in Section 38.3.2 and Middleton and Loder (1989).

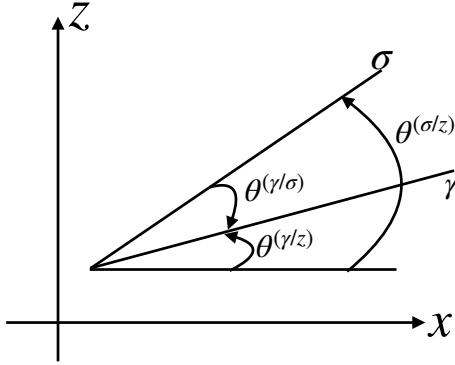


Figure 39.2: Slopes of neutral directions (denoted by γ isolines) relative to both the horizontal plane, $\tan \theta^{(\gamma/z)}$, and relative to a σ -isoline, $\tan \theta^{(\gamma/\sigma)}$. We assume positive angles as measure counter-clockwise relative to the horizontal and relative to the σ -isoline, 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 σ and γ isolines.

39.4 Cabbeling and thermobaricity

We now return to the density equation (40.28)

$$\frac{D \ln \rho}{Dt} = \nabla \cdot [\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S)] - (\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S) + \frac{\omega}{\rho c^2}. \quad (39.91)$$

We focus here on temperature and salinity fluxes just from neutral diffusion. The neutrality condition (39.78) takes the following form in terms of specific volume

$$\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S) = 0. \quad (39.92)$$

Consequently, neutral diffusion affects density evolution only through the source term

$$\left[\frac{D \ln \rho}{Dt} \right]_{\text{ntral diff}} = -\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta - \mathbf{J}(S) \cdot \nabla \nu_S. \quad (39.93)$$

39.4.1 Basic manipulations

As a first step, eliminate the salt flux by using the neutrality condition (39.92) so that

$$\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S = \mathbf{J}(\Theta) \cdot [\nu_S \nabla \nu_\Theta - \nu_\Theta \nabla \nu_S] / \nu_S. \quad (39.94)$$

Next, expand the gradients of the specific volume to write

$$\nabla \nu_\Theta = \nu_{\Theta\Theta} \nabla \Theta + \nu_{\Theta S} \nabla S + \nu_{\Theta p} \nabla p \quad (39.95a)$$

$$\nabla \nu_S = \nu_{S S} \nabla S + \nu_{\Theta S} \nabla \Theta + \nu_{S p} \nabla p, \quad (39.95b)$$

so that

$$\begin{aligned} \nu_S \nabla \nu_\Theta - \nu_\Theta \nabla \nu_S &= \nabla \Theta (\nu_S \nu_{\Theta\Theta} - \nu_\Theta \nu_{\Theta S}) \\ &\quad + \nabla S (\nu_S \nu_{\Theta S} - \nu_\Theta \nu_{S S}) + \nabla p (\nu_S \nu_{\Theta p} - \nu_\Theta \nu_{S p}). \end{aligned} \quad (39.96)$$

We again make use of the neutrality condition (39.92), as well as the symmetry condition (39.79d) to write

$$\mathbf{J}(\Theta) \cdot \nabla S (\nu_S \nu_{\Theta S} - \nu_\Theta \nu_{SS}) = -\mathbf{J}(\Theta) \cdot \nabla \Theta \left(\nu_\Theta \nu_{\Theta S} - \nu_{SS} \frac{(\nu_\Theta)^2}{\nu_S} \right). \quad (39.97)$$

Bringing these results together leads to

$$\begin{aligned} \mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S &= \mathbf{J}(\Theta) \cdot \nabla p \left[\nu_{\Theta p} - \nu_{pS} \frac{\nu_\Theta}{\nu_S} \right] \\ &\quad + \mathbf{J}(\Theta) \cdot \nabla \Theta \left[\nu_{\Theta \Theta} - 2 \nu_{\Theta S} \frac{\nu_\Theta}{\nu_S} + \nu_{SS} \left(\frac{\nu_\Theta}{\nu_S} \right)^2 \right], \end{aligned} \quad (39.98)$$

which can be written in terms of density partial derivatives as

$$\begin{aligned} \mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S &= -\rho^{-2} \mathbf{J}(\Theta) \cdot \nabla p \left[\rho_{\Theta p} - \rho_{pS} \frac{\rho_\Theta}{\rho_S} \right] \\ &\quad - \rho^{-2} \mathbf{J}(\Theta) \cdot \nabla \Theta \left[\rho_{\Theta \Theta} - 2 \rho_{\Theta S} \frac{\rho_\Theta}{\rho_S} + \rho_{SS} \left(\frac{\rho_\Theta}{\rho_S} \right)^2 \right]. \end{aligned} \quad (39.99)$$

39.4.2 A tidy form

We next write the bracket terms in forms consistent with those introduced by [McDougall \(1987b\)](#) in his classic paper discussing cabbeling and thermobaricity. For that purpose, introduce the *thermobaricity* parameter (dimensions of inverse temperature times inverse pressure) whose form is given by

$$\mathcal{T} = \beta \partial_p \left[\frac{\alpha}{\beta} \right] \quad (39.100a)$$

$$= \frac{\partial \alpha}{\partial p} - \frac{\alpha}{\beta} \frac{\partial \beta}{\partial p} \quad (39.100b)$$

$$= \rho \nu_S \partial_p \left[\frac{\nu_\Theta}{\nu_S} \right] \quad (39.100c)$$

$$= -\rho^{-1} \rho_S \partial_p \left[\frac{\rho_\Theta}{\rho_S} \right] \quad (39.100d)$$

$$= -\rho^{-1} \left[\rho_{\Theta p} - \rho_{pS} \left[\frac{\rho_\Theta}{\rho_S} \right] \right], \quad (39.100e)$$

and the *cabbeling* parameter (dimensions of squared inverse temperature)

$$\mathcal{C} = \frac{\partial \alpha}{\partial \Theta} + 2 \frac{\alpha}{\beta} \frac{\partial \alpha}{\partial S} - \left(\frac{\alpha}{\beta} \right)^2 \frac{\partial \beta}{\partial S} \quad (39.101a)$$

$$= -\rho^{-1} \left[\rho_{\Theta \Theta} - 2 \rho_{\Theta S} \left[\frac{\rho_\Theta}{\rho_S} \right] + \rho_{SS} \left[\frac{\rho_\Theta}{\rho_S} \right]^2 \right] \quad (39.101b)$$

$$= \rho \left[\nu_{\Theta \Theta} - 2 \nu_{\Theta S} \left[\frac{\nu_\Theta}{\nu_S} \right] + \nu_{SS} \left[\frac{\nu_\Theta}{\nu_S} \right]^2 \right] \quad (39.101c)$$

to render the very compact result

$$\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S = \rho^{-1} \mathbf{J}(\Theta) \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta) \quad (39.102)$$

which in turn yields the material evolution of *in situ* density due to neutral diffusion

$$\left[\frac{D\rho}{Dt} \right]_{\text{ntral diff}} = -\mathbf{J}(\Theta) \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta). \quad (39.103)$$

39.4.3 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 Θ 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 actually greater than the unmixed elements. This densification upon mixing is a physical process known as *cabbeling* ([McDougall, 1987b](#)).

The sign definite nature of cabbeling (i.e., cabbeling always results in denser fluid elements after mixing) is a direct result of the geometry of the locally referenced potential density surface when viewed in Conservative Temperature and salinity space. This property in turn manifests with the following inequality for the cabbeling parameter

$$\mathcal{C} = \frac{\partial \alpha}{\partial \Theta} + 2 \frac{\alpha}{\beta} \frac{\partial \alpha}{\partial S} - \left[\frac{\alpha}{\beta} \right]^2 \frac{\partial \beta}{\partial S} \geq 0. \quad (39.104)$$

Given the downgradient nature of the neutral diffusive fluxes, we have

$$\text{Cabbeling} \equiv -\mathcal{C} \mathbf{J}(\Theta) \cdot \nabla \Theta \geq 0, \quad (39.105)$$

thus providing a mathematical expression for the cabbeling source (with dimensions of density per time). That is, cabbeling results in a positive material evolution of density; i.e., density increases due to cabbeling. An increase in the density within a column of fluid results in the reduction of the sea level due to compression of the column.

39.4.4 Thermobaricity

The thermobaricity parameter

$$\mathcal{T} = \beta \frac{\partial}{\partial p} \left[\frac{\alpha}{\beta} \right] \quad (39.106)$$

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

$$\begin{aligned} \mathcal{T} &= \frac{\partial \alpha}{\partial p} - \frac{\alpha}{\beta} \frac{\partial \beta}{\partial p} \\ &= -\frac{\rho_{\Theta p}}{\rho} + \frac{\rho_{\Theta}}{\rho_S} \frac{\rho_{pS}}{\rho} \end{aligned} \quad (39.107)$$

Thermobaricity is the common name for the sum, since pressure variations in the thermal expansion coefficient dominate those of the haline contraction coefficient. The thermal expansion coefficient generally increases as pressure increases, thus making the thermobaric parameter positive.

Since neutral gradient of temperature 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)

$$\text{Thermobaricity} \equiv -\mathcal{T} \mathbf{J}(\Theta) \cdot \nabla p \quad (39.108)$$

appearing in equation (39.102). 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.

39.4.5 Comments

[Griffies and Greatbatch \(2012\)](#) discuss the impacts on global mean sea level from thermobaricity and cabling as diagnosed from an ocean model. Given that cabling 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 cabling and thermobaricity from observational based measurements, with [Groeskamp et al. \(2019\)](#) also offering a more robust numerical method for performing that diagnostic calculation.

Although cabling and thermobaricity lead to watermass transformation and associated transport of water across neutral directions, they are distinct from other mixing processes such as breaking gravity waves (Section 39.1). Namely, cabling and thermobaricity arise from the strong stirring by mesoscale eddies along neutral directions, which in turn leads to neutral diffusion acting on Conservative Temperature and salinity. Consequently, cabling and thermobaricity are independent of the amount of mechanical energy dissipation by breaking gravity waves ([McDougall et al., 2003](#)). Mesoscale eddies impart a downscale cascade of tracer variance that is ultimately halted by irreversible molecular mixing, or microscale processes active before reaching the molecular level. This mixing is the ultimate cause for cabling and thermobaricity. However, the strength of the cabling and thermobaricity processes are functions of mesoscale stirring rather than irreversible mixing.

40

Ocean buoyancy

Conservative temperature, Θ , is the preferred means to measure the transport of heat in the ocean, and salinity, S , measures the concentration of dissolved salt matter. These two scalar fields are referred to as *active* tracers as they both impact buoyancy and in turn affect pressure and ocean currents. In this chapter we establish the evolution equations for Θ and S and in turn determine how their affects on buoyancy.

READER'S GUIDE FOR THIS CHAPTER

Development in Section 40.1 presumes an understanding of how the tracer equation is derived from the kinematics of mass conserving fluid elements detailed in Section 18.1. Basic notions of thermodynamics (in particular Section 23.7) motivate the use of Conservative temperature (or potential temperature) as a measure of ocean heat transfer.

40.1	Salt and freshwater budgets	622
40.1.1	Mass budgets	622
40.1.2	Kinematic boundary conditons	623
40.1.3	Further reading	625
40.2	Material evolution of <i>in situ</i> density	625
40.2.1	Material changes to pressure	625
40.2.2	Material changes to Θ and S	626
40.2.3	Summary of density changes	626
40.2.4	A synopsis of ocean mixing processes	626
40.3	Boundary fluxes of buoyancy	627
40.3.1	Outlining the boundary fluxes of heat and salt	627
40.3.2	Transport of scalars across the ocean surface	628
40.3.3	Scalar budgets for a surface ocean model grid cell	628
40.3.4	Salt fluxes from sea ice melt and formation	629
40.3.5	Salt and heat fluxes associated with water transport	629
40.3.6	Non-penetrative surface heat fluxes	630
40.3.7	The case of frazil	631
40.3.8	Penetrative shortwave radiation	631
40.3.9	Buoyancy budget for a surface ocean model grid cell	632
40.3.10	Comments	632

40.1 Salt and freshwater budgets

Seawater is comprised of freshwater with 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 *salinity*

$$S = \frac{\text{mass of salt}}{\text{mass of seawater}} = \frac{\text{mass of salt}}{\text{mass of freshwater} + \text{mass of salt}} \quad (40.1)$$

to specify the amount of salt within an element of seawater. The complement to salinity is the freshwater concentration or mass fraction for an element of seawater

$$F = \frac{\text{mass of freshwater}}{\text{mass of seawater}} = \frac{\text{mass of freshwater}}{\text{mass of freshwater} + \text{mass of salt}} = 1 - S. \quad (40.2)$$

Other trace matter occurs at very low concentrations so as to make seawater a two-component fluid consisting of freshwater plus dissolved salt. We here derive the mass budget for salt and freshwater as well as the associated kinematic boundary conditions.

40.1.1 Mass budgets

Following our discussion of the tracer equation in Section 18.1, the mass budget equations for an element of seawater take the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad \text{seawater} \quad (40.3)$$

$$\frac{\partial(\rho S)}{\partial t} + \nabla \cdot (\rho \mathbf{v} S + \mathbf{J}^{(S)}) = 0 \quad \text{salt} \quad (40.4)$$

$$\frac{\partial(\rho F)}{\partial t} + \nabla \cdot (\rho \mathbf{v} F + \mathbf{J}^{(F)}) = 0 \quad \text{freshwater}, \quad (40.5)$$

where ρ is the seawater mass density. Equation (40.3) is the mass budget for seawater and equation (40.4) is the mass budget for salt. The freshwater budget (40.5) is derived by subtracting the salt budget (40.4) from the seawater mass budget (40.3). Hence, only two of the three mass budget equations (40.3)-(40.5) are independent.

We make use of the barycentric velocity in the above conservation laws, where the barycentric velocity is given by

$$\mathbf{v} = S \mathbf{v}^{(S)} + F \mathbf{v}^{(F)}. \quad (40.6)$$

The velocities $\mathbf{v}^{(S)}$ and $\mathbf{v}^{(F)}$ are, respectively, the molecular center of mass velocities for salt and freshwater contained in a fluid element, in which case

$$\rho \left[\frac{\partial S}{\partial t} + \mathbf{v}^{(S)} \cdot \nabla S \right] = \frac{\partial(\rho S)}{\partial t} + \nabla \cdot (\rho \mathbf{v}^{(S)} S) = 0 \quad (40.7a)$$

$$\rho \left[\frac{\partial F}{\partial t} + \mathbf{v}^{(F)} \cdot \nabla F \right] = \frac{\partial(\rho F)}{\partial t} + \nabla \cdot (\rho \mathbf{v}^{(F)} F) = 0. \quad (40.7b)$$

Furthermore, the fluxes $\mathbf{J}^{(S)}$ and $\mathbf{J}^{(F)}$ arise from the difference between the salt and freshwater velocities from the barycentric velocity

$$\mathbf{J}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v}) \quad \mathbf{J}^{(F)} = \rho F (\mathbf{v}^{(F)} - \mathbf{v}). \quad (40.8)$$

These fluxes are generally parameterized by downgradient diffusive fluxes

$$\mathbf{J}^{(S)} = -\rho \kappa_S \nabla S \quad \mathbf{J}^{(F)} = -\rho \kappa_S \nabla F, \quad (40.9)$$

where $\kappa_S > 0$ is the kinematic diffusivity for salt in seawater ([Gill, 1982](#)). Note that we use the same diffusivity for salt and freshwater, as the diffusion of one is balanced by the other. Furthermore, as discussed in Section 39.1, the effective diffusivity is enhanced beyond the molecular value in the presence of subgrid scale eddy effects.

The advective flux of seawater is comprised of a salt flux plus a freshwater flux

$$\rho \mathbf{v} = \rho S \mathbf{v}^{(S)} + \rho F \mathbf{v}^{(F)}. \quad (40.10)$$

Conversely, the salt flux and freshwater flux can be represented as a diffusive flux plus an advective flux where advection is via the barycentric velocity

$$\rho S \mathbf{v}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v}) + \rho S \mathbf{v} = \mathbf{J}^{(S)} + \rho S \mathbf{v} \quad (40.11a)$$

$$\rho F \mathbf{v}^{(F)} = \rho F (\mathbf{v}^{(F)} - \mathbf{v}) + \rho F \mathbf{v} = \mathbf{J}^{(F)} + \rho F \mathbf{v}. \quad (40.11b)$$

The center of mass velocities offer a conceptual framework of use for a formulation of kinematic boundary conditions in Section 40.1.2. Even so, they offer no new information beyond the parameterized fluxes $\mathbf{J}^{(S)}$ and $\mathbf{J}^{(F)}$: knowledge of one is sufficient for determining the other.

40.1.2 Kinematic boundary conditions

Recall the boundary condition (17.61) derived in Section 17.4.3 for matter crossing the ocean surface

$$\rho (\mathbf{v} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_m, \quad (40.12)$$

where \mathcal{Q}_m is the net mass flux of freshwater plus salt crossing the surface boundary, and $\mathbf{v}^{(\eta)}$ is the velocity of a point on the free surface.

Boundary conditions in terms of salt and freshwater velocities

In deriving the boundary condition (40.12), we made use of the barycentric velocity \mathbf{v} for an element of seawater. A directly analogous procedure can be applied to the salt and freshwater crossing the surface to render

$$\rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_S \quad (40.13a)$$

$$\rho (\mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_F, \quad (40.13b)$$

where

$$\mathcal{Q}_m = \mathcal{Q}_S + \mathcal{Q}_F \quad (40.14)$$

relates the mass fluxes of salt and freshwater to the total mass flux crossing the boundary (mass per time per surface normalized area). Adding the boundary conditions (40.13a) and (40.13b), and using the relation $S + F = 1$, recovers the boundary condition (40.12) written in terms of the barycentric velocity.

In many regions, the ocean surface is impermeable to salt, in which case the ocean surface acts as a material surface in terms of the salt velocity

$$\rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{zero surface salt flux.} \quad (40.15)$$

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.

Diffusive flux boundary condition for salt

We find it useful to make use of relation (40.11a) to eliminate the salt velocity $\mathbf{v}^{(S)}$ in favor of the diffusive flux $\mathbf{J}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v})$, in which case the kinematic boundary condition (40.13a) takes the form

$$\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = S \mathcal{Q}_F. \quad (40.16)$$

This kinematic boundary condition relates the surface freshwater flux crossing the ocean surface (right hand side) to the normal component of the subgrid scale salt flux (left hand side). Notably, this result holds whether or not there is a surface salt flux.

Now assume $\mathbf{J}^{(S)}$ is in the form of a diffusive flux

$$\mathbf{J}^{(S)} = -\rho \kappa_S \nabla S, \quad (40.17)$$

in which case the boundary condition (40.16) takes the form

$$\rho \kappa_S \nabla S \cdot \hat{\mathbf{n}} = -S \mathcal{Q}_F. \quad (40.18)$$

Diffusive mixing of salinity within the ocean thus mediates the incorporation of boundary freshwater fluxes into the ocean. Since it is the mass of a fluid element that is constant, any transfer of freshwater into that element must be compensated by a removal of salt, and vice versa. Through the act of salt diffusion in one direction, freshwater diffuses in the opposite. For example, when adding freshwater to the ocean, $P - E > 0$, it enters the ocean (moves downward) so long as salt diffuses upward toward the surface. Correspondingly, in the absence of diffusive mixing, boundary freshwater is not incorporated into the ambient ocean fluid. Rather, it remains an unmixed lens sitting on top of the seawater.

40.1.3 Further reading

Nurser and Griffies (2019) offer a more detailed discussion of this material.

40.2 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). As discussed in Section 25.2.2, it is common to write the seawater equation of state for density as a function of potential temperature, salinity, and pressure. A somewhat more accurate approach makes use of the Conservative Temperature rather than the potential temperature, where the Conservative Temperature, Θ , is the potential enthalpy divided by a constant heat capacity (McDougall, 2003; IOC et al., 2010). We thus make use of the empirical relation for the seawater density in the functional form

$$\rho = \rho(\Theta, S, p). \quad (40.19)$$

We formulate the material evolution of density as weighted by the specific volume

$$\nu = \rho^{-1}, \quad (40.20)$$

so that we consider

$$\frac{D \ln \rho}{Dt} = \frac{\partial \ln \rho}{\partial \Theta} \frac{D\Theta}{Dt} + \frac{\partial \ln \rho}{\partial S} \frac{DS}{Dt} + \frac{1}{\rho} \frac{\partial \rho}{\partial p} \frac{Dp}{Dt} \quad (40.21a)$$

$$= -\alpha \frac{D\Theta}{Dt} + \beta \frac{DS}{Dt} + \frac{\omega}{\rho c^2}. \quad (40.21b)$$

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 coordinates

$$\alpha = - \left[\frac{\partial \ln \rho}{\partial \Theta} \right]_{p,S} \quad \beta = \left[\frac{\partial \ln \rho}{\partial S} \right]_{p,\Theta} \quad c_s^2 = \left[\frac{\partial p}{\partial \rho} \right]_{S,\Theta} \quad \omega = \frac{Dp}{Dt}. \quad (40.22)$$

40.2.1 Material changes to pressure

To garner some exposure to the physics of ω as it appears in equation (40.21), we consider the special 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 21.3.5)

$$\begin{aligned} w^{(p)} &= \frac{\partial z}{\partial p} \frac{Dp}{Dt} \\ &= -(\rho g)^{-1} \omega. \end{aligned} \quad (40.23)$$

The transport measured by $w^{(p)}$ is the pressure-coordinate analog of the vertical velocity component $w = Dz/Dt$ in a geopotential coordinate representation of the vertical. That is, fluid moving into regions of increasing hydrostatic pressure ($\omega > 0$) 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.

40.2.2 Material changes to Θ and S

We now focus on the salinity and temperature contributions to the evolution of *in situ* density. To do so, assume that the material evolution of Θ and S are given by the convergence of a subgrid scale flux

$$\rho \frac{D\Theta}{Dt} = -\nabla \cdot \mathbf{J}(\Theta) \quad (40.24a)$$

$$\rho \frac{DS}{Dt} = -\nabla \cdot \mathbf{J}(S). \quad (40.24b)$$

This form for material changes in temperature and salinity then lead to

$$-\alpha \frac{D\Theta}{Dt} + \beta \frac{DS}{Dt} = \nu_\Theta \nabla \cdot \mathbf{J}(\Theta) + \nu_S \nabla \cdot \mathbf{J}(S) \quad (40.25a)$$

$$= \nabla \cdot [\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S)] - [\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S] \quad (40.25b)$$

where again $\nu = \rho^{-1}$ is the specific volume and its partial derivatives are

$$\nu_\Theta = \frac{\partial \nu}{\partial \Theta} = \frac{\alpha}{\rho} \quad \text{and} \quad \nu_S = \frac{\partial \nu}{\partial S} = -\frac{\beta}{\rho}. \quad (40.26)$$

40.2.3 Summary of density changes

Bringing the above results together leads to the density equation

$$\frac{D \ln \rho}{Dt} - \frac{\omega}{\rho c_s^2} = \nabla \cdot [\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S)] - [\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S], \quad (40.27)$$

which has the equivalent form

$$\frac{D\rho}{Dt} - \frac{\omega}{c_s^2} = \nabla \cdot [\alpha \mathbf{J}(\Theta) - \beta \mathbf{J}(S)] - [\mathbf{J}(\Theta) \cdot \nabla \alpha - \mathbf{J}(S) \cdot \nabla \beta]. \quad (40.28)$$

We brought the adiabatic source term from motion across pressure surfaces (Section 40.2.1) onto the left hand side, as this term appears in the absence of mixing whereas terms on the right hand side require mixing. 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., temperature Θ reducing and S increasing). These fluxes arise from a variety of mixing processes, some of which are surveyed in Section 40.2.4. The second term on the right hand side of equations (40.27) and (40.28) relates to properties of the locally referenced potential density surface. We study this source term arising from neutral diffusion in Section 39.4, where we encounter cabbeling and thermobaricity. Further effects arise from unresolved eddy-induced stirring, with that process contributing to the material time derivative operator to render a residual mean velocity (Section 39.1.4).

40.2.4 A synopsis of ocean mixing processes

Irreversible mixing in the ocean takes place at the millimeter scale through the process of molecular (Brownian) motion acting to dissipate property gradients. This mixing is generally represented by downgradient molecular diffusion ([Einstein, 1905](#)). 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 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 molecular diffusion, and following the pioneering work of [Taylor \(1921\)](#), it is common to parameterize mixing induced by 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 is the notable counter-example to this equivalence [Schmitt \(1994\)](#).

Mixing induced by eddies of length scale $\mathcal{O}(\text{centimeters-metres})$ is associated with, among other processes, gravitational instability, shear instability and breaking internal gravity waves ([MacKinnon et al., 2013](#)), as well as a suite of boundary layer processes ([Large et al., 1994](#)). 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 quiescent ocean interior ([Polzin et al., 1997](#); [Whalen et al., 2012](#); [Waterhouse et al., 2014](#)).

Mesoscale eddies, with size $\mathcal{O}(10 - 100) \text{ km}$, preferentially stir tracers along neutral directions ([McDougall, 1987a,b](#); [McDougall et al., 2014](#)). The mesoscale eddy stirring in turn induces a mixing that is parametrized by downgradient diffusion along neutral directions (Section 39.3). 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](#)). 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](#)).

Although the isotropic diffusivity is much smaller than the mesoscale diffusivity, the isotropic diffusivity multiplies the generally larger tracer gradients crossing neutral directions and thus supports a critical form of watermass transformation and an induced ocean circulation ([Munk, 1966](#); [Munk and Wunsch, 1998](#)).

40.3 Boundary fluxes of buoyancy

As introduced in Chapter 25, buoyancy measures the gravitational acceleration of a fluid element relative to that of the fluid environment surrounding the element. A reduction in density for the fluid element is associated with an increase in buoyancy; that is, the fluid element becomes more *buoyant*. 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 way, buoyancy changes are directly related to processes that impact locally referenced potential density through changes in the temperature and salinity of a fluid element.

In this section we derive the equation describing the changes in ocean buoyancy due to heat, salt, and water fluxes crossing the ocean boundaries. 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 heat across the ocean surface boundary. Furthermore, we are only concerned with the upper ocean boundary, so that we ignore geothermal fluxes crossing the ocean bottom boundary.

40.3.1 Outlining the boundary fluxes of heat and salt

Broadly, the boundary fluxes are associated with the following physical processes.

- Turbulent processes transfer heat 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.
- Salt is transferred between the liquid ocean and sea ice when sea ice melts and forms. This transfer is proportional to the water mass flux and the difference in salinity between the liquid ocean and sea ice. More generally, we simply consider this process to be associated with a salt flux between sea ice and ocean.
- Advective processes transfer heat and salt across the ocean surface through the transfer of water mass across the interface.

40.3.2 Transport of scalars across the ocean surface

To develop a quantitative understanding of how buoyancy is impacted by surface boundary fluxes, we develop the evolution equations for temperature, salinity, and mass in an arbitrary top model grid cell, and focus on evolution arising from surface boundary fluxes. For this purpose, recall the derivation of scalar budgets for an arbitrary grid cell depicted in Figure 18.2, and specialize to the case of a grid cell with upper boundary the air-sea interface. From equation (18.33), the transport of an arbitrary scalar, ψ , crossing the upper ocean interface is written

$$Q_\psi \, dA = -[\rho \psi (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} \, dS, \quad (40.29)$$

where \mathbf{J} is a subgrid scale flux, dS is the area along the surface interface, and dA is the horizontal projection of that area. Equation (17.71c) renders the surface kinematic boundary condition

$$\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} \, dS = -Q_m \, dA \quad (40.30)$$

for the mass per time of material crossing the surface through such processes as precipitation, evaporation, and river runoff. We introduce the analogous expression for the subgrid scalar transport

$$\mathbf{J} \cdot \hat{\mathbf{n}} \, dS = -Q_{sgs} \, dA, \quad (40.31)$$

where Q_{sgs} encapsulates the turbulent fluxes of scalar, such as those for heat and salt listed above. We are thus left with scalar flux across the surface ocean in the form

$$Q_\psi = Q_m \psi + Q_{sgs}. \quad (40.32)$$

By construction, a positive contribution to Q_ψ increases the amount of ψ within the ocean domain.

40.3.3 Scalar budgets for a surface ocean model grid cell

We now make use of the surface flux expression (40.32) within the finite volume budget equation (18.33) to develop budget equations for temperature, salinity, and mass for a grid cell. Again, we focus exclusively on contributions from surface boundary fluxes, with lateral transport important but not our present concern. The budget equations are thus given by

$$\frac{\partial (\rho dz \Theta)}{\partial t} = Q_m \Theta_m + Q_\Theta^{\text{non-pen}} + [Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z)] \quad (40.33)$$

$$\frac{\partial (\rho dz S)}{\partial t} = Q_m S_m + Q_S \quad (40.34)$$

$$\frac{\partial (\rho dz)}{\partial t} = Q_m. \quad (40.35)$$

We now detail the terms appearing in these equations.

- ρdz is the mass per horizontal area of seawater in the grid cell. For a volume conserving Boussinesq fluid discussed in Chapter 28, the *in situ* density, ρ , is set to the constant reference density ρ_0
- Θ_m is the temperature of water crossing the ocean surface, and $C_p^o Q_m \Theta_m$ is the associated enthalpy flux (W m^{-2}). We further discuss this flux in Section 40.3.5.
- S_m is the salinity of water crossing the ocean surface, and $Q_m S_m$ is the associated mass flux of salt. S_m is typically taken to be zero, as for precipitation and evaporation. However, rivers can contain a nonzero salt concentration, as can sea ice melt. So we keep S_m for the following formulation. We further discuss this salt flux in Section 40.3.5.
- C_p^o is the seawater heat capacity at constant pressure ($\text{J kg}^{-1} \text{ }^\circ\text{C}^{-1}$). [IOC et al. \(2010\)](#) provides the most precise value appropriate for an ocean with heat measured through conservative temperature.
- Q_S is the flux of salt ($\text{kg m}^{-2} \text{ sec}^{-1}$) that crosses the ocean surface, with $Q_S > 0$ when salt enters the ocean. This flux arises in the turbulent transfer of salt when sea ice forms and melts. We further discuss this salt flux in Section 40.3.4.
- $C_p^o Q_\Theta^{\text{non-pen}}$ is the non-penetrative surface heat flux associated with turbulent processes (latent and sensible) and radiative longwave cooling (W m^{-2}) localized to the upper ocean interface. $Q_\Theta^{\text{non-pen}} > 0$ for heat entering the ocean surface (i.e., ocean warming). We further discuss this heat flux in Section 40.3.6.
- $C_p^o Q_\Theta^{\text{pen}}(z = \eta)$ is the radiative shortwave heat flux (W m^{-2}) entering the ocean through its surface at $z = \eta$, with $Q_\Theta^{\text{pen}}(\eta) > 0$ warming the ocean surface. Likewise, $C_p^o Q_\Theta^{\text{pen}}(z = -\Delta z)$ is the radiative shortwave heat flux leaving the top cell through its bottom interface. It is the difference $[Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z)]$ that contributes to the net warming of the grid cell from shortwave radiation. We further discuss the shortwave heat flux in Section 40.3.8.

40.3.4 Salt fluxes from sea ice melt and formation

The turbulent mass flux of salt Q_S ($\text{kg m}^{-2} \text{ sec}^{-1}$) is positive for salt entering the ocean. There is transport of salt across the ocean surface when sea ice forms and melts, due to the nonzero salt content in sea ice. Otherwise, the surface salt flux is generally zero for the large scale ocean. For ocean models, however, the salt flux can be nonzero when formulating the surface boundary in terms of virtual salt fluxes rather than real water fluxes ([Huang, 1993](#); [Griffies et al., 2001](#)). It can also be non-zero when using an ocean-ice model that is not coupled to an atmosphere or land model, in which case salt restoring is required to maintain stability of the overturning circulation (see Section 3 of [Griffies et al. \(2009\)](#)).

40.3.5 Salt and heat fluxes associated with water transport

In most cases, salinity in the water fluxed across the ocean surface is zero, so that $S_m = 0$. However, there are some cases where rivers have a nonzero salinity so that $S_m \neq 0$ and the product $Q_m S_m$ leads to an advective transport of salt across the ocean surface. Also, in some contexts the advective flux of salt due to sea ice melt and formation is encapsulated in the $Q_m S_m$ term rather than the turbulent salt flux Q_S .

Since water transported across the ocean has a nonzero heat content, this transport in turn affects the ocean heat content. One can either prescribe the temperature of this water, Θ_m , or the product $Q_m \Theta_m$. Consider the case where the product is specified for river water entering the ocean, which is the case with certain river models such as in [Dunne et al. \(2012\)](#). In this case, the heat flux with respect to $0^\circ C$ (in units of W m^{-2}) of liquid river runoff $\mathcal{H}^{\text{liquid runoff}}$ is given to the ocean from the river model, so that

$$Q_m \Theta_m = \frac{\mathcal{H}^{\text{liquid runoff}}}{C_p^{\text{liquid runoff}}}, \quad (40.36)$$

with $C_p^{\text{liquid runoff}}$ the heat capacity of the water coming in from the river runoff. Likewise, if the heat associated with frozen runoff (e.g., calving land ice) is provided by the land model, then we have

$$Q_m \Theta_m = \frac{\mathcal{H}^{\text{solid runoff}}}{C_p^{\text{solid runoff}}}, \quad (40.37)$$

with $C_p^{\text{solid runoff}}$ the heat capacity of the solid runoff. These two heat capacities are typically provided by the component model (i.e., the river model) used to compute the runoff fields. Similar considerations hold for transfer of water between sea ice models and the ocean.

40.3.6 Non-penetrative surface heat fluxes

The heat flux $C_p^o Q_\Theta^{\text{non-pen}}$ (W m^{-2}) is positive for heat entering the ocean. This flux is comprised of the following contributions (e.g., see page 34 of [Gill, 1982](#))

$$C_p^o Q_\Theta^{\text{non-pen}} = Q_{\text{long}} + Q_{\text{latent}} + Q_{\text{sens}}. \quad (40.38)$$

These fluxes are termed non-penetrative since they are deposited or withdrawn from a small region near the ocean surface interface.

Longwave radiation

Longwave radiation leaves the ocean in the form of the $\sigma_{\text{SB}} T^4$ Stefan-Boltzmann Law, with T the ocean skin temperature and

$$\sigma_{\text{SB}} = 5.6734 \times 10^{-8} \text{ W m}^{-2} \text{ }^\circ\text{K}^{-4} \quad (40.39)$$

the Stefan-Boltzmann constant. $Q_{\text{long}} < 0$ since the longwave heat flux removes heat from the ocean surface and sends it back to the atmosphere.

Latent heat fluxes

Q_{latent} arises from phase changes whereby liquid seawater either evaporates, or it acts to melt frozen precipitation. In either case, $Q_{\text{latent}} < 0$ since the liquid ocean loses heat to energize the phase changes.

When seawater evaporates, the latent heat lost by the ocean is determined by the latent heat of vaporization for fresh water

$$H^{\text{vapor}} = 2.5 \times 10^6 \text{ J kg}^{-1}, \quad (40.40)$$

so that

$$Q_{\text{evap}} = H^{\text{vapor}} Q_m^{\text{evap}} \quad (40.41)$$

where $Q_m^{\text{evap}} < 0$ is the mass flux ($\text{kg m}^{-2} \text{ sec}^{-1}$) of fresh water leaving the ocean due to evaporation. A similar expression holds when seawater melts frozen precipitation (e.g., snow), in which case

$$H^{\text{fusion}} = 3.34 \times 10^5 \text{ J kg}^{-1}, \quad (40.42)$$

so that

$$Q_{\text{melt}} = -H^{\text{fusion}} Q_m^{\text{frozen precip}}, \quad (40.43)$$

where $Q_m^{\text{frozen precip}} > 0$ is the mass flux ($\text{kg m}^{-2} \text{ sec}^{-1}$) of frozen precipitation falling onto the ocean surface. Again, both Q_{evap} and Q_{melt} are negative since latent heating extracts heat from the ocean.

Sensible heat fluxes

Q_{sens} is the sensible heat transfer proportional to the difference between the ocean temperature and that of the atmosphere, sea ice, or land ice. Sensible heating generally acts to cool the ocean ($Q_{\text{sens}} < 0$), particularly near western boundary currents such as the Gulf Stream, Kuroshio, and Agulhas.

40.3.7 The case of frazil

As the temperature of seawater cools to the freezing point, sea ice is formed, initially through the production of frazil ice. Frazil can generally form at various levels in the upper ocean, though many ocean models assume frazil production occurs just in the top grid cell. Operationally in an ocean model, liquid water can be supercooled at any particular time step through surface fluxes and transport. An adjustment process is used to heat the liquid water back to the freezing point, with this positive heat flux $Q_{\text{frazil}} > 0$ extracted from the ice model as frazil sea ice is formed.

40.3.8 Penetrative shortwave radiation

Momentum and buoyancy are transferred across the upper ocean surface boundary, with ocean processes such as advection and mixing then transporting the boundary momentum and buoyancy laterally as well as into the ocean interior. In contrast, penetrative shortwave radiation is absorbed into the ocean absent ocean transport processes, with such absorption a function of ocean optical properties. Much of the shortwave radiation is absorbed in the upper ocean turbulent boundary layer, though a fraction leaks through to the interior. In general, such non-turbulent and non-advection transport of buoyancy via penetrative radiation represents a fundamentally novel aspect of ocean boundary layer physics relative to the atmosphere. Namely, for the atmosphere, radiative absorption is far less relevant than in the upper ocean, since the atmosphere is largely transparent to radiation. It is therefore useful to distinguish penetrative shortwave radiation from other buoyancy fluxes when formulating how boundary fluxes impact the ocean.

The penetrative shortwave radiative heat flux $C_p^o Q_\Theta^{\text{pen}} > 0$ arises from the net shortwave radiation entering through the ocean surface and absorbed by seawater. As noted above, a fraction of this radiation generally penetrates beneath the surface ocean grid cell, with the fraction depending on the optical properties of seawater and thickness of the grid cell. Hence, we subtract a heat flux $C_p^o Q_\Theta^{\text{pen}}(z = -\Delta z)$, which represents the radiative shortwave heat flux passing through the bottom of the surface ocean cell at $z = -\Delta z$. It is the difference

$$\text{net shortwave heating of surface grid cell} = C_p^o [Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z)] \quad (40.44)$$

that stays in the surface grid cell.

40.3.9 Buoyancy budget for a surface ocean model grid cell

We now bring the previous fluxes together to form the budget for buoyancy in a surface grid cell due to the impacts of surface fluxes. The resulting expression is then used to derive an expression for the buoyancy forcing that acts on the ocean surface boundary layer. Buoyancy has a time tendency given by

$$-\left[\frac{\rho_o}{g}\right] \frac{\partial b}{\partial t} = \rho_\Theta \frac{\partial \Theta}{\partial t} + \rho_S \frac{\partial S}{\partial t}, \quad (40.45)$$

where we introduced the shorthand

$$\rho_\Theta = \left[\frac{\partial \rho}{\partial \Theta}\right]_{S,p} \quad \rho_S = \left[\frac{\partial \rho}{\partial S}\right]_{\Theta,p} \quad (40.46)$$

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 (40.33) by ρ_Θ and add to the surface salinity equation (40.34) multiplied by ρ_S

$$\rho_\Theta \left[\frac{\partial(\rho dz \Theta)}{\partial t} \right] + \rho_S \left[\frac{\partial(\rho dz S)}{\partial t} \right] = Q_m (\rho_\Theta \Theta_m + \rho_S S_m) + \rho_\Theta [Q_\Theta^{\text{non-pen}} + \delta_k Q_\Theta^{\text{pen}}] + \rho_S Q_S, \quad (40.47)$$

where we introduced the shorthand

$$\delta_k Q_\Theta^{\text{pen}} = Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z). \quad (40.48)$$

We now use the mass budget (40.35) and introduce the buoyancy tendency according to equation (40.45) to realize an expression for the time tendency of the surface ocean buoyancy

$$(\rho_o/g) \rho dz \frac{\partial b}{\partial t} = Q_m [\rho_\Theta (\Theta - \Theta_m) + \rho_S (S - S_m)] + \rho_\Theta [Q_\Theta^{\text{non-pen}} + \delta_k Q_\Theta^{\text{pen}}] - \rho_S Q_S. \quad (40.49)$$

Introducing the thermal expansion and saline contraction coefficients

$$\alpha = -\frac{1}{\rho} \left[\frac{\partial \rho}{\partial \Theta} \right]_{S,p} \quad \beta = \frac{1}{\rho} \left[\frac{\partial \rho}{\partial S} \right]_{\Theta,p} \quad (40.50)$$

to render

$$dz \frac{\partial b}{\partial t} = \frac{g}{\rho_o} (Q_m [-\alpha (\Theta - \Theta_m) + \beta (S - S_m)] + \alpha (\delta_k Q_\Theta^{\text{pen}} + Q_\Theta^{\text{non-pen}}) - \beta Q_S). \quad (40.51)$$

40.3.10 Comments

The buoyancy flux expression (40.51) is of use for boundary layer parameterizations, such as the KPP scheme of [Large et al. \(1994\)](#) and [Roekel et al. \(2018\)](#). It is furthermore used when studying water mass transformations as reviewed by [Groeskamp et al. \(2019\)](#) and summarized in Chatper 41.

41

Water mass phase space

A *water mass* is a region of seawater characterized by a selection of scalar properties that distinguish it from other regions. Water masses are typically formed through boundary processes, such as the extremely large buoyancy fluxes at the high latitudes that form the Antarctic Bottom Water (AABW) and North Atlantic Deep Water (NADW). As these waters enter the ocean interior they are advected over basin scales while they are also eroded or *transformed* by irreversible mixing processes. Water masses offer a useful conceptual means to partition the ocean into distinct classes whose origin, movement, and modification can be measured, modeled, and studied. Furthermore, scalar properties are generally simpler to measure than vector properties such as velocity, with a water mass perspective offering the means to infer general properties of the ocean circulation without directly measuring the velocity field.

Water mass analysis is a mathematical formalism aimed at determining budgets for seawater mass and tracer mass within layers defined by scalar properties such as buoyancy, Conservative Temperature, salinity, or biogeochemical tracers. We speak of a *water mass phase space*, which is three dimensional just like the geographic/depth space (x, y, z) and yet with some or all of the coordinates determined by scalar properties. A water mass phase space views the ocean circulation through the lens afforded by layers rather than fixed regions (Eulerian “boxes”) or moving fluid regions (Lagrangian fluid “particles”). Hence, a water mass view complements that from Eulerian and Lagrangian kinematics. Notably, there is no assumption of a one-to-one mapping between geographic/depth space and water mass phase space. A point in water mass phase space can thus be comprised of many points in geographic/depth space. This property is both liberating and frustrating. It is liberating since it provides a framework to infer general properties of ocean circulation even without measuring velocity. It is frustrating since those properties often come without full geographic/depth information, though partial spatial information is common.

Determining ocean circulation in a water mass phase space is a problem of determining how the layers are modified or *transformed* by irreversible boundary and interior ocean processes. A water mass perspective thus focuses on irreversible changes that lead to the *formation* or destruction of water mass classes. This focus on irreversible change means there is precisely zero motion in water mass phase space in the absence of the irreversible. So if one is interested in reversible processes (e.g., linear waves) then the water mass phase space is not appropriate. Indeed, given the generally non-local geographic aspect of water mass phase space, one cannot study local forces, stresses, and accelerations within this space. Instead, dynamics in water mass phase space arises solely from irreversible processes affecting the scalar fields that define the layers, rather than the dynamics from Newton’s Laws affecting motion of fluid elements.

With its focus on the irreversible, a water mass phase space perspective is natural for quantifying and understanding steady state and transient properties of the ocean climate state, including ocean stratification and the movement and transformation of scalar properties. For these reasons, a water mass phase space viewpoint is very useful for addressing questions of primary interest in the Anthropocene.

READER'S GUIDE FOR THIS CHAPTER

A study of circulation in water mass phase space requires the vector calculus encountered in Chapter 4, elements of the generalized vertical coordinates from Chapters 11 and 21, and features of the tracer transport and mixing discussed in Chapter 39. Most of these pre-requisites are reviewed here to keep the discussion reasonably self-contained.

41.1	Transformation and formation	635
41.1.1	A three-layer thought experiment	635
41.1.2	Characterizing how processes lead to transformations	636
41.2	Seawater mass budget within layers	637
41.2.1	Dia-surface transport and water mass transformation	638
41.2.2	The volume of a λ -layer	638
41.2.3	Practical expression for the water mass transformation	640
41.2.4	Transport crossing open boundaries and ocean surface	641
41.2.5	The area between two λ -interfaces intersecting a surface	642
41.2.6	The layer mass budget	642
41.2.7	Circulation in λ -space	643
41.2.8	Further study	644
41.3	Tracer mass budget within layers	645
41.3.1	General form of the mass budget	645
41.3.2	Tracer processes	646
41.3.3	Transport across an interior layer interface	647
41.3.4	Transport across interior and surface boundaries	647
41.3.5	The layer tracer budget	647
41.4	$\Theta - S$ water mass analysis	648
41.5	Tracer transport pathways	648
41.6	Green's functions and ventilation time scales	648

41.1 Transformation and formation

The most common scalar for partitioning the ocean is buoyancy, or its approximate form given by potential density (Chapter 25). In this section we introduce the notions of *transformation* and *formation* when partitioning the ocean according to density (γ) classes that locally measure buoyancy. The ideas extend to any scalar field used to bin the ocean fluid.

Water mass *transformation* measures the mass per time of water that moves across the boundaries of a layer. By convention, the transformation is positive if water moves into a region of larger density (more generally into a region of larger value for the scalar field) and negative if it enters a lighter density layer. Water mass *formation* refers to the difference in transformation across a layer, so that it measures the change in mass of the layer. That is, formation is the convergence of transformation. Both transformation and formation have dimensions of mass per time (or volume per time when considering Boussinesq fluids), and is typically measured in Sverdrup units.

41.1.1 A three-layer thought experiment

To illustrate the concepts of transformation and formation, consider an ocean that is binned into three density layers bounded by four density interfaces:

$$\text{light density layer} = [\gamma - \delta\gamma/2, \gamma + \delta\gamma/2] \quad (41.1a)$$

$$\text{middle density layer} = [\gamma + \delta\gamma/2, \gamma + 3\delta\gamma/2] \quad (41.1b)$$

$$\text{heavy density layer} = [\gamma + 3\delta\gamma/2, \gamma + 5\delta\gamma/2]. \quad (41.1c)$$

Figure 41.1 depicts a sample mass distribution for this ocean; i.e., the mass census for water binned into these three density layers. Now introduce a process that results in water leaving the middle density layer and entering both the light layer and the heavy layer. Let $G(\sigma)$ measure the mass

per time that water crosses the density interfaces; i.e., $G(\sigma)$ is the transformation. This particular thought experiment has the following transformations across the various layer interfaces

$$G(\sigma) = \begin{cases} 0 & \sigma = \gamma - \delta\gamma/2 \\ < 0 & \sigma = \gamma + \delta\gamma/2 \\ > 0 & \sigma = \gamma + 3\delta\gamma/2 \\ 0 & \sigma = \gamma + 5\delta\gamma/2 \end{cases} \quad \begin{array}{l} \text{closed boundary} \\ \text{mass moves to light density from middle density} \\ \text{mass moves from middle density to heavy density} \\ \text{closed boundary.} \end{array} \quad (41.2)$$

The difference in the transformation across the interface boundaries of a particular layer determines the formation/destruction of water into 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:

$$\text{light-formation} = -[G(\gamma + \delta\gamma/2) - G(\gamma - \delta\gamma/2)] > 0 \quad (41.3a)$$

$$\text{middle-formation} = -[G(\gamma + 3\delta\gamma/2) - G(\gamma + \delta\gamma/2)] < 0 \quad (41.3b)$$

$$\text{heavy-formation} = -[G(\gamma + 5\delta\gamma/2) - G(\gamma + 3\delta\gamma/2)] > 0. \quad (41.3c)$$

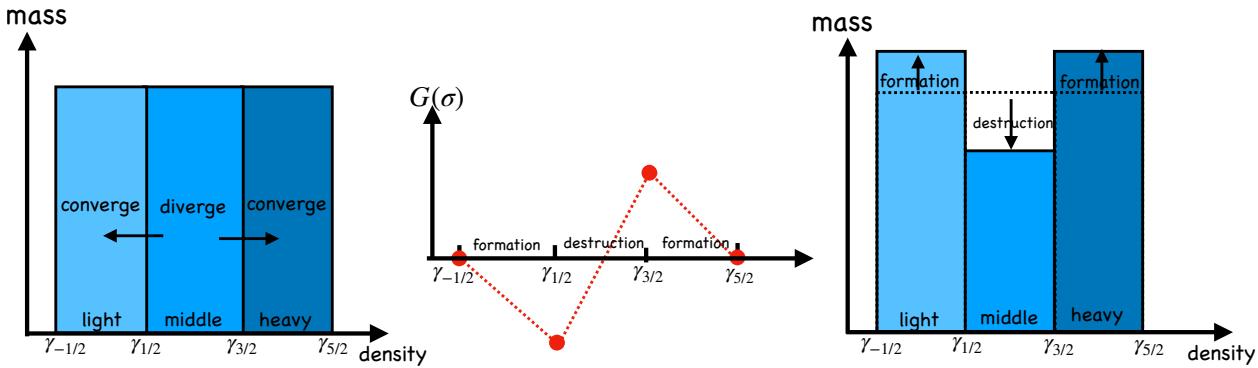


Figure 41.1: A sample mass distribution of the ocean binned into density layers (light, middle, heavy) bounded by the four density interfaces: $\gamma_{n/2} = \gamma + n \delta\gamma/2$ for $n = -1, 1, 3, 5$. The left panel depicts an ocean state with equal mass in each layer. A process then causes water 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 of mass) whereas the light and heavy layers have experienced a positive formation. The middle panel depicts the transformation, $G(\sigma)$, 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 $G(\sigma)$) but it does not modify the basic ideas illustrated in this thought experiment.

41.1.2 Characterizing how processes lead to transformations

The key focus of water mass analysis is the movement of water between layers, with this water movement modifying the distribution of mass within the layers. There are a number of processes that affect the movement of mass across layer boundaries. One involves the passage of mass across domain boundaries such as the surface ocean. In particular, rain and evaporation alter the mass of the ocean, with these mass changes appearing in the mass budget for those density layers that outcrop at the ocean surface.

Mixing provides a second means to move water across layer boundaries, with no scalar field (of non-uniform distribution) materially invariant in the presence of mixing. As discussed in Chapter

18, mixing causes trace matter to move between fluid elements even as mixing does not alter the net mass of fluid elements (recall our discussion of barycentric velocity in Section 18.1.2). Hence, seawater fluid elements retain a fixed mass and yet this mass is redistributed among layers since the layer boundaries move. A particularly striking example occurs in the upper ocean boundary layer where surface forcing leads to the seasonal migration of density outcrops. The associated lateral movement of density layers is commonly faster than the movement of fluid elements. Hence, through the seasonal cycle a density bin can inflate or deflate by moving the layer boundaries so that the layer entrains or detrains mass.

41.2 Seawater mass budget within layers

In this section we put some mathematical flesh onto the conceptual bones from Section 41.1 by developing a formalism to quantify water mass budgets within a layer such as in Figure 41.2. For generality, we formulate water mass transformation analysis in terms of an arbitrary smooth scalar field, $\lambda = \lambda(\mathbf{x}, t)$, with seawater distributed into discrete bins or classes according to the value of λ . We thus seek a formalism to quantify the evolution of seawater distributions within the λ -classes.

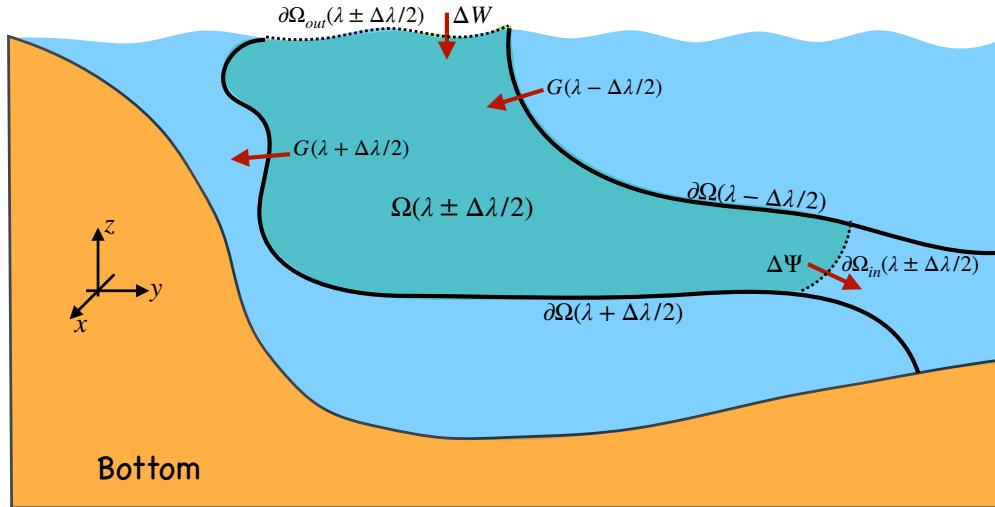


Figure 41.2: A layer of seawater with scalar property λ within the range $\lambda \in [\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$ and defined over a geographical domain $\Omega(\lambda \pm \Delta\lambda/2)$. The water mass transformations crossing the layer interfaces $\partial\Omega(\lambda \pm \Delta\lambda/2)$ are written $G(\lambda \pm \Delta\lambda/2)$, with $G > 0$ for water moving to regions of larger λ . The seawater mass crossing the layer through the geographical bounds $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$ is written $\Delta\Psi(\lambda \pm \Delta\lambda/2)$, with $\Delta\Psi(\lambda \pm \Delta\lambda/2) > 0$ 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. The mass crossing the sea surface, $\partial\Omega_{out}(\lambda \pm \Delta\lambda/2)$, through rain, evaporation, melt, or rivers 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 unstable 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 $\partial\Omega(\lambda \pm \Delta\lambda/2) = \partial\Omega_{in}(\lambda \pm \Delta\lambda/2) + \partial\Omega_{out}(\lambda \pm \Delta\lambda/2) + \partial\Omega(\lambda + \Delta\lambda/2) + \partial\Omega(\lambda - \Delta\lambda/2)$. This figure is based on Figure 3 of Groeskamp et al. (2019) and is oriented according to the Southern Ocean with Antarctica to the left.

In Section 41.1 we assumed λ is a density field, $\lambda = \gamma$, whereas here it is a generic scalar field. Furthermore, in contrast to the case of a generalized vertical coordinate (Chapters 11 and 21), we make no assumption regarding the stratification of λ . Rather, λ -isosurfaces are here free to overturn (e.g., see Figure 41.2) and to disconnect. This freedom is motivated by the behavior of most oceanographic scalar property fields, even buoyancy within boundary layers. As noted in the chapter introduction, this added freedom comes at a cost. Namely, we no longer have a one-to-one

relation between a point in water mass phase space and a point in geographic/depth space. Yet we choose to pay that price given the novel insights into ocean circulation when viewed within water mass phase space.

41.2.1 Dia-surface transport and water mass transformation

A fundamental element in the analysis of water mass transformation is the dia-surface velocity from Section 21.3.6

$$w^{\text{dia}} = \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\lambda)}) = \frac{\dot{\lambda}}{|\nabla\lambda|} \quad \text{with} \quad \hat{\mathbf{n}} = \frac{\nabla\lambda}{|\nabla\lambda|} \quad \text{and} \quad \dot{\lambda} = \frac{D\lambda}{Dt}. \quad (41.4)$$

The transport w^{dia} locally measures the amount of seawater that penetrates a λ -interface in the direction of increasing λ , with the λ -interface moving with a non-zero velocity, $\mathbf{v}^{(\lambda)}$, defined according to

$$\frac{\partial\lambda}{\partial t} + \mathbf{v}^{(\lambda)} \cdot \nabla\lambda = 0. \quad (41.5)$$

For a particular λ -interface labelled by $\lambda(\mathbf{x}, t) = \lambda^*$ we define the transformation, $G(\lambda^*)$, as the area integral of ρw^{dia} over the λ^* -interface

$$G(\lambda^*) = \int_{\partial\Omega(\lambda^*)} \rho w^{\text{dia}} dS = \int_{\partial\Omega(\lambda^*)} \frac{\rho \dot{\lambda}}{|\nabla\lambda|} dS \implies \text{dimensions} = M T^{-1}. \quad (41.6)$$

We write $\partial\Omega(\lambda^*)$ for the geographical area occupied by the interface (see Figure 41.2).

Recall our conceptual description of water mass transformation in Section 41.1.2. Based on the definition (41.1.2), we see that water mass transformation across a λ -interface occurs when there is a material change, $\dot{\lambda} \neq 0$. Such changes arise from mixing and boundary processes thus causing irreversible changes to the matter content within λ -layers by driving seawater to cross λ -interfaces. *Formation* occurs when more water is transformed into a layer than that leaving the layer, thus leading to a modified distribution of mass within the λ -layers.

Equation (41.6) for the water mass transformation is written as a surface integral of the dia-surface transport across the λ interface. In practice, the surface integral is not easy to compute diagnostically since the λ -surface can be highly corrugated and disconnected, and discrete estimates of $|\nabla\lambda|^{-1}$ are prone to nontrivial errors and noise. A practical evaluation of the water mass transformation thus requires an alternative method. In the next few sections we derive mathematical results that allow us to develop the practical means for computing the transformation.

41.2.2 The volume of a λ -layer

Consider how we measure the volume within an infinitesimal λ -layer bounded by two interfaces $\lambda \in [\lambda - \delta\lambda/2, \lambda + \delta\lambda/2]$, with Figure 41.3 providing an example. The volume within a tiny cylinder extending from one interface to the other is given by

$$\delta V = \delta h \delta S \quad (41.7)$$

where dS is the area element and δh is the layer thickness. The differential λ increment separating the two interfaces is given by

$$\delta\lambda = \nabla\lambda \cdot \delta\mathbf{x} = |\nabla\lambda| \hat{\mathbf{n}} \cdot \delta\mathbf{x} = |\nabla\lambda| \delta h, \quad (41.8)$$

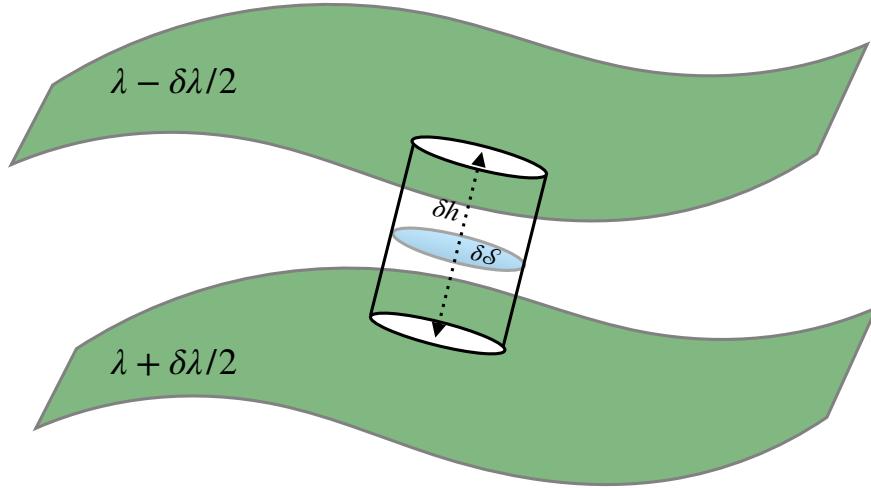


Figure 41.3: This schematic shows an infinitesimally thin λ -layer bounded by two interfaces $\lambda(\mathbf{x}, t) \in [\lambda - \delta\lambda/2, \lambda + \delta\lambda/2]$. We here assume the λ interfaces are monotonically layered in one of the three directions (e.g., stably stratified density in the vertical). However, this assumption is not generally needed for water mass analysis. The cylinder region extends between the two iso-surfaces and it has thickness $\delta h = \delta\lambda/|\nabla\lambda|$ and cross-sectional area δS . The cylinder is oriented according to the normal direction $\hat{n} = |\nabla\lambda|^{-1} \nabla\lambda$, where we assume that $|\nabla\lambda| \neq 0$.

so that the layer thickness is

$$\delta h = \frac{\delta\lambda}{|\nabla\lambda|}. \quad (41.9)$$

Hence, for a given increment, $\delta\lambda > 0$, the layer thickness is inversely proportional to the modulus of the λ -gradient. The thickness is oriented according to the normal direction, $\hat{n} = \nabla\lambda/|\nabla\lambda|$, so that it always measures the distance between the interfaces in the direction of the normal.

Given the expression (41.9) for the layer thickness, the volume for the cylinder between two infinitesimally close λ -interfaces is

$$\delta V = \delta h \delta S = \frac{\delta\lambda \delta S}{|\nabla\lambda|}. \quad (41.10)$$

It follows that the volume of a finite layer bounded by two interfaces is given by the integral over the layer

$$\Delta V(\lambda) \equiv \int_{\Omega(\lambda \pm \Delta\lambda/2)} dV = \int_{\lambda - \Delta\lambda/2}^{\lambda + \Delta\lambda/2} d\lambda \int_{\partial\Omega(\lambda')} \frac{dS}{|\nabla\lambda'|}. \quad (41.11)$$

We also find it useful to express the volume of seawater with values of λ greater than a reference value; i.e., for $\lambda_0 \leq \lambda$ where λ_0 is an arbitrary value. In Figure 41.4 we illustrate this case where λ_0 is less than any value of λ realized in the full domain. The volume of water in this region is written

$$V(\lambda) \equiv \int_{\Omega(\lambda_0 \leq \lambda)} dV = \int_{\lambda_0}^{\lambda} d\lambda \int_{\partial\Omega(\lambda')} \frac{dS}{|\nabla\lambda'|}, \quad (41.12)$$

where the integral excludes regions that are not in the ocean. The volume within a λ -layer $[\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$ is given by the difference

$$\Delta V(\lambda) = V(\lambda + \Delta\lambda/2) - V(\lambda - \Delta\lambda/2). \quad (41.13)$$

Notably, the arbitrary limit λ_0 drops out from the difference. Taking the limit $\Delta\lambda \rightarrow 0$ renders the derivative

$$\frac{\partial V(\lambda)}{\partial \lambda} = \lim_{\Delta\lambda \rightarrow 0} \frac{V(\lambda + \Delta\lambda/2) - V(\lambda - \Delta\lambda/2)}{\Delta\lambda} = \int_{\partial\Omega(\lambda)} \frac{d\mathcal{S}}{|\nabla\lambda|}, \quad (41.14)$$

which provides a measure for the volume per unit λ within a layer. This is an important result that will be used in the following.

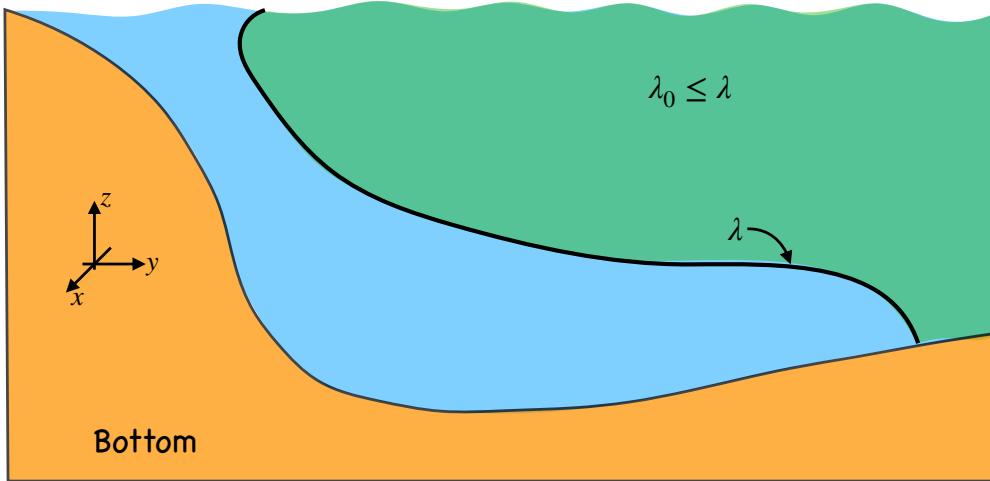


Figure 41.4: Depicting the volume of fluid with $\lambda_0 \leq \lambda$, where λ_0 is an arbitrary reference value. Here we assume λ_0 is less than any value of λ realized in the ocean. For example, if $\lambda = \gamma$, and if λ_0 is less than any density realized in the ocean, then the region with $\lambda_0 \leq \lambda$ encompasses all of the fluid lighter than λ .

41.2.3 Practical expression for the water mass transformation

We now extend the identities given by equation (41.14) to include the volume integral of an arbitrary function, $\Phi(\mathbf{x}, t)$

$$g_\Phi(\lambda) = \int_{\Omega(\lambda_0 \leq \lambda)} \Phi dV = \int_{\lambda_0}^{\lambda} d\lambda' \int_{\partial\Omega(\lambda')} \frac{\Phi d\mathcal{S}}{|\nabla\lambda'|}. \quad (41.15)$$

Taking the λ -derivative of the volume integral and using the fundamental theorem of calculus picks out the surface integral

$$\frac{\partial g_\Phi(\lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda} \left[\int_{\lambda_0}^{\lambda} d\lambda' \int_{\partial\Omega(\lambda')} \frac{\Phi d\mathcal{S}}{|\nabla\lambda'|} \right] = \int_{\partial\Omega(\lambda)} \frac{\Phi d\mathcal{S}}{|\nabla\lambda|}. \quad (41.16)$$

The identity (41.16) is the key to a practical discrete approximation of the surface integral

$$\int_{\partial\Omega(\lambda)} \frac{\Phi dS}{|\nabla\lambda|} = \frac{\partial\mathcal{I}_\Phi(\lambda)}{\partial\lambda} \quad (41.17a)$$

$$= \lim_{\Delta\lambda \rightarrow 0} \frac{\mathcal{I}_\Phi(\lambda + \Delta\lambda/2) - \mathcal{I}_\Phi(\lambda - \Delta\lambda/2)}{\Delta\lambda} \quad (41.17b)$$

$$= \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \left[\int_{\lambda_0}^{\lambda+\Delta\lambda/2} d\lambda' \int_{\partial\Omega(\lambda')} \frac{\Phi dS}{|\nabla\lambda'|} - \int_{\lambda_0}^{\lambda-\Delta\lambda/2} d\lambda' \int_{\partial\Omega(\lambda')} \frac{\Phi dS}{|\nabla\lambda'|} \right] \quad (41.17c)$$

$$= \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \left[\int_{\lambda-\Delta\lambda/2}^{\lambda+\Delta\lambda/2} d\lambda' \int_{\partial\Omega(\lambda')} \frac{\Phi dS}{|\nabla\lambda'|} \right] \quad (41.17d)$$

$$= \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \int_{\Omega(\lambda \pm \Delta\lambda/2)} \Phi dV. \quad (41.17e)$$

The final expression is a volume integral of Φ computed within a thin layer surrounding the λ -interface.

Setting $\Phi = \rho \dot{\lambda}$ in equation (41.17e) renders the transformation across the λ -interface

$$G(\lambda) = \int_{\partial\Omega(\lambda)} \rho u^{\text{dia}} dS = \int_{\partial\Omega(\lambda)} \frac{\rho \dot{\lambda} dS}{|\nabla\lambda|} = \frac{\partial}{\partial\lambda} \int_{\Omega(\lambda_0 \leq \lambda)} \rho \dot{\lambda} dV = \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho \dot{\lambda} dV. \quad (41.18)$$

The final expression lends itself to a straightforward binning algorithm that forms the basis for the practice of water mass transformation analysis (see Section 7.5 of [Groeskamp et al. \(2019\)](#)). In this manner, calculation of water mass transformation in a model or from observational measurements requires the various contributions to the material time change, $\dot{\lambda}$, a weighting of those time changes according to the mass of the fluid element, ρdV , and binning of $\rho \dot{\lambda} dV$ according to λ -classes. Notably, many scalars materially evolve according to $\rho \dot{\lambda} = -\nabla \cdot \mathbf{J}$, so that the volume integral in equation (41.18) reduces, through Gauss' divergence theorem, to surface integrals of $\hat{\mathbf{n}} \cdot \mathbf{J}$ along the layer boundaries. However, in practice it is more convenient to accumulate $\rho \dot{\lambda} dV$ within the thin layer according to the discretized volume integral, rather than to perform the area integral of the flux around the domain boundary.

41.2.4 Transport crossing open boundaries and ocean surface

According to Figure 41.2, we need two more transport terms to develop a layer mass budget. The mass transport leaving the layer through an interior open boundary is given by

$$\Delta\Psi(\lambda \pm \Delta\lambda/2) = \int_{\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS, \quad (41.19)$$

where $\mathbf{v}^{(b)}$ is the velocity for a point on the boundary and $\hat{\mathbf{n}}$ is the outward normal along the boundary. As an explicit example, set the boundary to be at a particular latitude so that $\mathbf{v}^{(b)} = 0$ and $\hat{\mathbf{n}} = \hat{\mathbf{y}}$, in which case

$$\Delta\Psi(\lambda \pm \Delta\lambda/2) = \int_{\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)} \rho v dx dz, \quad (41.20)$$

where $\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)$ specifies the depth and longitude range over the constant latitude boundary.

In a similar manner we write the mass transport crossing the ocean free surface

$$\Delta W(\lambda \pm \Delta\lambda/2) = - \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} Q_m dA, \quad (41.21)$$

where we made use of the surface kinematic boundary condition (equation (17.71c)) to write

$$\rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} \equiv -Q_m dA, \quad (41.22)$$

where $Q_m dA$ is the mass transport of water crossing the free surface ($Q_m > 0$ for water entering the ocean) and $dA = dx dy$ is the horizontal area element.

41.2.5 The area between two λ -interfaces intersecting a surface

The transports $\Delta W(\lambda \pm \Delta\lambda/2)$ and $\Delta\Psi(\lambda \pm \Delta\lambda/2)$ are computed over the area of a surface bounded by where the interfaces $\lambda \pm \Delta\lambda/2$ intersect that surface. We can express these area integrals in a manner analogous to the volume integrals considered previously. For that purpose consider the area between two λ values along a surface, such as the ocean surface $\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)$ or the open interior boundary $\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)$. Figure 41.5 illustrates the case for the ocean surface. From this figure we write the integral of an arbitrary function Φ over the region of a surface with $\lambda_0 \leq \lambda$

$$\Sigma_\Phi(\lambda_0 \leq \lambda) = \int_{\partial\Omega(\lambda_0 \leq \lambda)} \Phi d\mathcal{S}. \quad (41.23)$$

By construction the area integral of an arbitrary function within the layer $[\lambda - \Delta\lambda/2, \lambda + \Delta\lambda/2]$ is given by the difference

$$\Delta\Sigma_\Phi(\lambda \pm \Delta\lambda/2) = \Sigma_\Phi(\lambda_0 \leq \lambda + \Delta\lambda/2) - \Sigma_\Phi(\lambda_0 \leq \lambda - \Delta\lambda/2) = \int_{\partial\Omega(\lambda \pm \Delta\lambda/2)} \Phi d\mathcal{S}, \quad (41.24)$$

with the reference value dropping out. We make use of this result in Section 41.2.6 when developing the layer mass budget.

41.2.6 The layer mass budget

Bringing the above pieces together leads to the mass budget for a layer

$$\frac{d\Delta M}{dt} + \Delta\Psi = \Delta W - \Delta G, \quad (41.25)$$

where

$$\Delta M(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho dV \quad (41.26)$$

is the layer mass, and for brevity we dropped $\lambda \pm \Delta\lambda/2$ arguments on all terms. It is common to define the layer mass *formation* as the mass accumulation within the layer

$$\Delta F \equiv \frac{d\Delta M}{dt} + \Delta\Psi = \Delta W - \Delta G. \quad (41.27)$$

The first equality indicates that water mass formation occurs in the layer if there is a trend in the mass within the layer and/or net mass crossing the interior open boundary. The second equality indicates that water mass formation occurs if there is mass crossing the ocean surface and/or mass converging through transformation across interior layer interfaces.

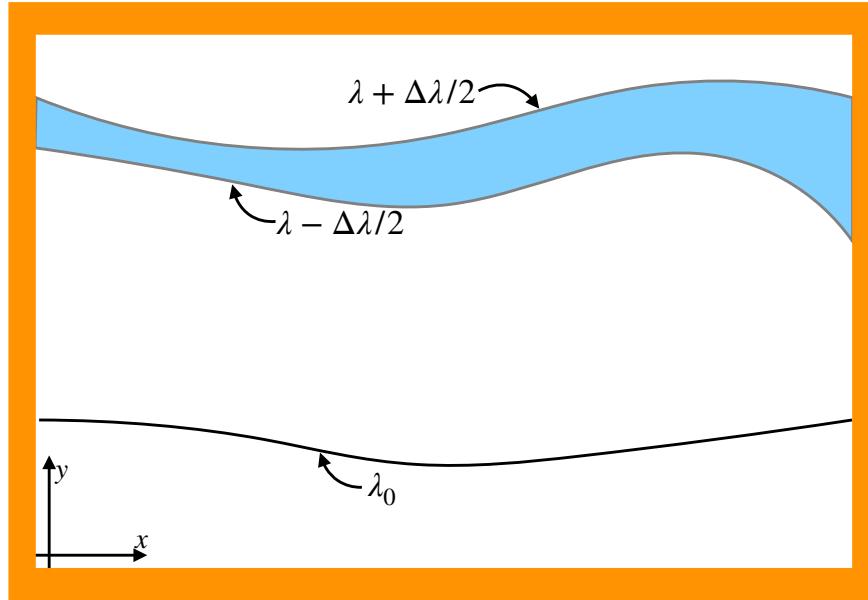


Figure 41.5: A closed ocean basin and the area between two λ -interfaces along the ocean surface where the λ -interfaces outcrop. The reference value λ_0 is arbitrary.

We arrive at a differential equation for the mass budget (41.27) by dividing through by the layer increment, $\Delta\lambda$, and taking the limit as this increment tends to zero

$$\frac{\partial F}{\partial \lambda} = \frac{\partial}{\partial t} \left[\frac{\partial M}{\partial \lambda} \right] + \frac{\partial \Psi}{\partial \lambda} = \frac{\partial W}{\partial \lambda} - \frac{\partial G}{\partial \lambda}. \quad (41.28)$$

Interchanging the λ and time derivatives and integrating from a reference value λ_0 to λ leads to

$$\int_{\lambda_0}^{\lambda} \frac{\partial \Psi}{\partial \lambda} d\lambda = \int_{\lambda_0}^{\lambda} \frac{\partial}{\partial \lambda} \left[-\frac{\partial M}{\partial t} + W - G \right] d\lambda \implies \Psi = -\frac{\partial M}{\partial t} + W - G, \quad (41.29)$$

where we dropped the contribution from the reference value at the lower limit since it is assumed to sit outside of the ocean domain. The water mass equation (41.29) is a continuous version of the discrete equation (41.27).

41.2.7 Circulation in λ -space

The layer mass budget (41.27) and its continuous expression (41.29) provide the basis for many general inferences about circulation and transformation. As an illustration, consider the continuous expression (41.29) and integrate from the reference value up to λ

$$\int_{\lambda_0}^{\lambda} \Psi d\lambda' = \int_{\lambda_0}^{\lambda} \left[-\frac{\partial M}{\partial t} + W - G \right] d\lambda'. \quad (41.30)$$

The left hand side is an expression for the overturning circulation in λ -space. We thus see that a nonzero overturning circulation is driven by mass through the ocean surface, convergence of mass transformed across the λ -interfaces, and/or time changes to the mass within the domain. Correspondingly, in the absence of surface boundary mass fluxes, a steady circulation is driven just by water mass transformation

$$\int_{\lambda_0}^{\lambda} \Psi d\lambda' = - \int_{\lambda_0}^{\lambda} G d\lambda'. \quad (41.31)$$

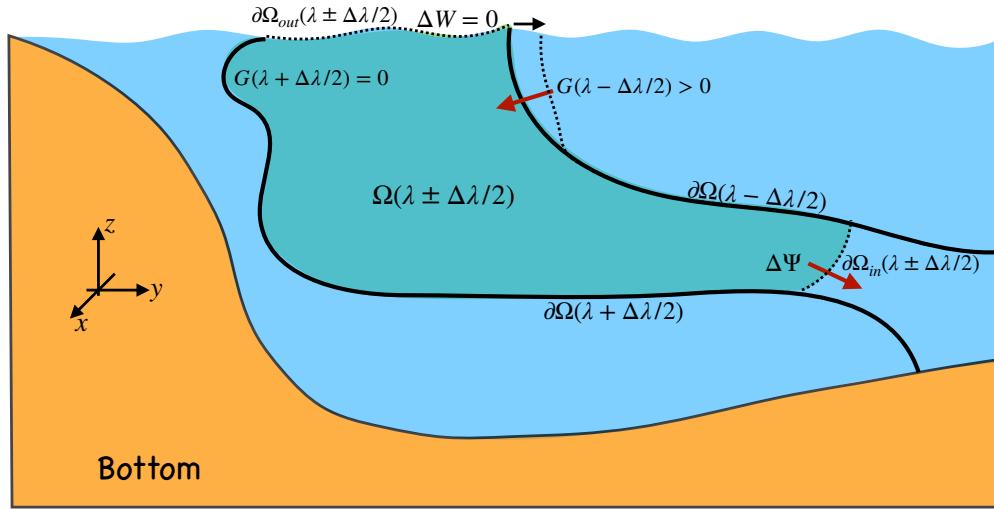


Figure 41.6: An example of a surface transformation driven circulation oriented according to the Southern Ocean with Antarctica to the left. Here we depict a layer that is exposed to air-sea interactions that cause the $\lambda - \Delta\lambda/2$ -interface to move meridionally (e.g., buoyancy fluxes in the case of $\lambda = \gamma$). Movement of the $\lambda - \Delta\lambda/2$ -interface causes seawater 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). In turn, the boundary $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$ moves in the opposite direction 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 (e.g., by having less mass move across the other layer interface at $\partial\Omega_{\lambda+\Delta\lambda/2}$), then mass accumulates within the layer. There is a steady state (i.e., layer mass is constant) only if the same amount of mass that converges into the layer via surface water mass transformation leaves the layer through circulation at the boundary $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$.

We depict an example in Figure 41.6 where the surface outcrop of the layer is exposed to air-sea interactions that lead to a meridional drift of the $\lambda - \Delta\lambda/2$ -interface. This movement laterally entrains mass into the layer. If there is a net convergence of mass into the layer, then the layer mass increases. A steady state mass budget for the layer is reached if the amount of mass entrained through surface transformation is reflected in the same mass leaving through the circulation at the open boundary $\partial\Omega_{in}(\lambda \pm \Delta\lambda/2)$. This and other statements related to the water mass budget are rather routine mathematically. Yet since the mass budget is formulated over layers, the mass budget offers the means to make very general statements about the circulation without a direct measurement of the flow.

41.2.8 Further study

[Walsh \(1977\)](#) and [Walsh \(1982\)](#) developed a layer formalism for studying water mass transformation and the associated circulation. [Marshall et al. \(1999\)](#) introduced the formalism surrounding equation (41.16), thus providing a mathematical framework for the practice of water mass analysis. A somewhat pedantic point of note is that [Marshall et al. \(1999\)](#) refer to equation (41.16) as a *generalized Leibniz rule*. Instead, it is a direct result of the fundamental theorem of calculus.

Much of the formalism in this section follows that reviewed by [Groeskamp et al. \(2019\)](#). This paper offers specific examples of water mass transformation analysis as well as citations to numerous papers.

41.3 Tracer mass budget within layers

Consider a tracer concentration, C , localized in a region within a layer of buoyancy (commonly approximated by potential density) as depicted in Figure 41.7. The upper panels to this figure illustrate the tracer patch in geographic/depth space along with isolines of buoyancy, γ , whereas the lower panels show the tracer distribution (histogram) binned within the buoyancy classes. If the tracer is mixed within a layer, such as via the neutral diffusion process of Section 39.3, then the tracer patch is spread laterally within the layer and yet the distribution (lower panel) is unchanged. In contrast, if the tracer is mixed across layer interfaces 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 geographical space. That is, if the tracer moves at a velocity distinct from the buoyancy field, then its distribution within buoyancy classes will change.

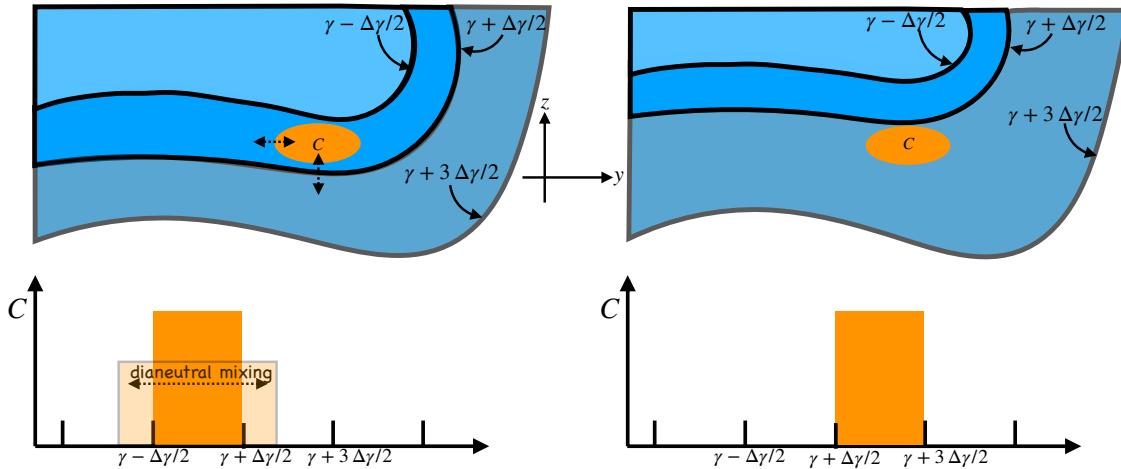


Figure 41.7: This figure shows a tracer patch within the buoyancy layer bounded by the interface values $[\gamma - \Delta\gamma/2, \gamma + \Delta\gamma/2]$ (left panel) and $[\gamma + \Delta\gamma/2, \gamma + 3\Delta\gamma/2]$ (right panel). The upper panels show the tracer and buoyancy in geographic/depth space whereas the bottom panels show the tracer distribution (histogram) binned according to buoyancy. 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 39.3). 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 $[\gamma + \Delta\gamma/2, \gamma + 3\Delta\gamma/2]$.

41.3.1 General form of the mass budget

As depicted in Figure 41.7 for buoyancy 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, λ , used to classify water masses. We can quantify these two processes by writing the expression for the time change of tracer content within a λ -layer. This expression is arrived at by applying the Leibniz-Reynolds transport theorem from Section 18.3.4 to a λ -layer

$$\frac{d}{dt} \Delta M_C(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho \dot{C} dV - \oint_{\partial\Omega(\lambda \pm \Delta\lambda/2)} \rho C (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} dS, \quad (41.32)$$

where

$$\Delta M_C(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho C dV \quad (41.33)$$

is the mass of tracer within the layer. The volume integral on the right hand side of equation (41.25) arises from material time changes to the tracer within the layer, whereas the surface integral arises from dia-surface transport across the layer boundary.

41.3.2 Tracer processes

We often parameterize the material time changes for a tracer according to the convergence of a flux

$$\rho \dot{C} = \rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}. \quad (41.34)$$

Many biogeochemical tracers have additional source terms as well, but these are ignored here for brevity. Gauss' 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 diagnostically to bin the volume integrated material time changes within the λ -classes. In contrast, the surface and bottom boundary contributions are fed into the budget via Neumann boundary conditions applied to the flux \mathbf{J}

$$\mathbf{J} \cdot \hat{\mathbf{n}} dS = \text{boundary tracer transport.} \quad (41.35)$$

Note that when there is an advective/skew diffusive component to the subgrid scale flux (Chapters 38 and 39), then it adds to the resolved advective component to render a residual mean material time operator

$$\rho \frac{D^\dagger C}{Dt} = -\nabla \cdot \mathbf{J}_{\text{non-advect}}, \quad (41.36)$$

where

$$\frac{D^\dagger}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v} + \mathbf{v}^*) \cdot \nabla, \quad (41.37)$$

with \mathbf{v}^* an eddy-induced velocity (see Section 39.1). For the purposes of water mass transformation analysis, we write

$$\dot{C} = \frac{D^\dagger C}{Dt}, \quad (41.38)$$

thus incorporating the eddy-induced stirring into the kinematic expression for the material time derivative.

There are many interior and boundary processes that contribute to \dot{C} within a layer. As a general expression for these contributions to the layer budget we write

$$\Delta E_C(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho \dot{C} dV, \quad (41.39)$$

which is sometimes usefully decomposed into interior and boundary processes

$$\Delta E_C^{\text{interior}}(\lambda \pm \Delta\lambda/2) + \Delta E_C^{\text{boundary}}(\lambda \pm \Delta\lambda/2) = \int_{\Omega(\lambda \pm \Delta\lambda/2)} \rho (\dot{C}^{\text{interior}} + \dot{C}^{\text{boundary}}) dV, \quad (41.40)$$

where diffusion is the canonical mixing process contributing to $\dot{C}^{\text{interior}}$.

41.3.3 Transport across an interior layer interface

The surface integral in the budget (41.32) involves transport across the layer interfaces, with this transport requiring motion of the interface relative to a fluid particle. The same formalism introduced in Section 41.2 can be used to compute this transport. That is, we can generalize the transformation equation (41.18) to write

$$G_C(\lambda) = \int_{\partial\Omega(\lambda)} \rho C (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \frac{\partial}{\partial \lambda} \int_{\Omega(\lambda_0 \leq \lambda)} \rho \dot{\lambda} C dV. \quad (41.41)$$

As a sanity check, we note that for the special case where the tracer concentration is a constant along the layer interface, then $G_C(\lambda) = C G(\lambda)$.

41.3.4 Transport across interior and surface boundaries

Following the surface kinematic boundary condition detailed in Section 41.2.6 for the outcrop region we write the budget contribution from mass fluxes crossing the boundary in the following form, which is based on the surface kinematic boundary condition given by equation (18.51),

$$\Delta W_C = \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} \rho C (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\partial\Omega_{\text{out}}(\lambda \pm \Delta\lambda/2)} Q_m C_m dA, \quad (41.42)$$

where C_m is the tracer concentration within the mass transported across the boundary.¹ As a sanity check, note that in the special case of a constant tracer concentration at the sea surface, then $\Delta W_C = C \Delta W$, where ΔW is the water mass transported across the ocean free surface as given by equation (41.21).

For the interior open boundary the contribution is written in the generic manner

$$\Delta\Psi_C = \int_{\partial\Omega_{\text{in}}(\lambda \pm \Delta\lambda/2)} C \rho (\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (41.43)$$

Again, in the special case where the tracer concentration is a constant along the interior boundary, then $\Delta\Psi_C = C \Delta\Psi$, where $\Delta\Psi$ is the seawater mass transport given by equation (41.19).

41.3.5 The layer tracer budget

The layer tracer mass budget is written

$$\frac{d\Delta M_C}{dt} + \Delta\Psi_C = \Delta E_C + \Delta W_C - [G_C(\lambda + \Delta\lambda/2) - G_C(\lambda - \Delta\lambda/2)], \quad (41.44)$$

which is directly analogous to the seawater layer mass budget (41.25), with the added term ΔE_C arising from material tracer changes. As for the seawater mass budget discussed in Section 41.2.6, the layer tracer budget (41.44) provides the framework for rather general inferences about tracer transport within λ -classes.

¹Note that equation 26 in [Groeskamp et al. \(2019\)](#) incorrectly writes the integrand in equation (41.42) as $Q_m (C_m - C)$.

41.4 $\Theta - S$ water mass analysis

41.5 Tracer transport pathways

- Lagrangian methods as per [*van Sebille et al. \(2018\)*](#).
- Stochastic tracer methods and Fokker-Planck as summarized in [*van Sebille et al. \(2018\)*](#).

41.6 Green's functions and ventilation time scales

- Summarize formalism from Haine et al review in progress.

Part VII

Shallow water mechanics

Shallow water models consist of constant density fluid layers whose interfaces are material (i.e., no matter is transferred between the immiscible layers). Thermodynamic processes are absent from the system, thus allowing us to focus purely on the dynamics of perfect fluid layers. Momentum is transferred between layers through pressure forces that act on sloping layer interfaces. Furthermore, dynamical motion occurs in columns, with horizontal velocity independent of vertical position within a layer.

The shallow water model provides us with a suite of versatile theoretical models of use to deduce fluid dynamical impacts from both rotation and stratification. It is among the most popular fluid models for theorists. Consequently, the shallow water model features heavily in many areas of geophysical fluid mechanics as well as in applications to the ocean and atmosphere.

42

Shallow water models

We here formulate the kinematic and dynamic equations for a suite of shallow water models by developing equations for a single shallow water layer; multiple shallow water layers (stacked shallow water); and reduced gravity models (models with one layer that is dynamically inactive).

READER'S GUIDE TO THIS CHAPTER

This chapter is the mathematical and physical basis for subsequent discussions that involve the shallow water model. It relies on elements of fluid kinematics and dynamics described in earlier chapters.

42.1	A single shallow water layer	652
42.1.1	Pressure within the fluid layer	652
42.1.2	Momentum equation	653
42.1.3	Thickness equation	654
42.1.4	Eulerian flux-form momentum equation	656
42.1.5	Kinematic boundary conditions	656
42.1.6	Stretching and vertical velocity	658
42.1.7	Further study	659
42.2	Emphasizing the hydrostatic approximation	659
42.2.1	Hydrostatic balance with respect to background density	660
42.2.2	Density is a uniform constant	660
42.2.3	Fluid is incompressible	660
42.2.4	Non-hydrostatic and hydrostatic forces in a homogeneous layer	660
42.2.5	Comments	661
42.3	Shallow water fluid in a rotating tank	661
42.3.1	Equations of motion	661
42.3.2	Free surface shape in solid-body rotation	661
42.3.3	Further study	662
42.4	Reduced gravity model for the upper ocean	662
42.4.1	Momentum and thickness equations for the active layer	662
42.4.2	Relating undulations of the top and bottom layer interfaces	663
42.4.3	Momentum equation with reduced gravity	663
42.4.4	Further study	663
42.5	Stacked shallow water equations	664
42.5.1	Model formulation	664
42.5.2	Further study	666

42.1 A single shallow water layer

Consider a homogeneous layer of fluid in a uniform effective gravitational field (gravity plus planetary centrifugal), contained on its side boundaries by vertical 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 (volume in a uniform density layer) means that thicker fluid regions must come at the cost of thinner fluid regions. Furthermore, layer thickness gradients create pressure differences (thicker water has larger hydrostatic bottom pressure than thinner water), 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 depth dependence to the horizontal motion.

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 that ensures that the accumulation of fluid in one region is balanced by depletion of fluid in another. Pressure gradients act to homogenize the layer thickness. However, rotation allows for layer thickness to be non-constant even in a steady state.

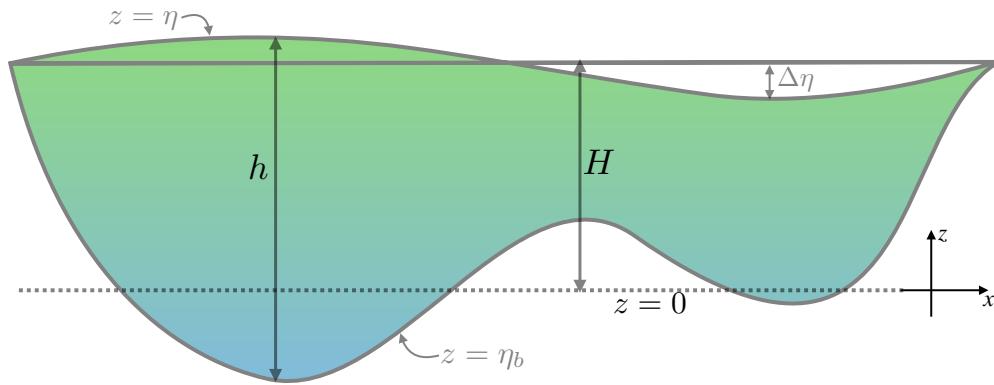


Figure 42.1: A single layer of shallow water fluid with h the thickness of the water column, H the averaged fluid depth (if the area integral of η_b is zero), $z = \eta$ the free surface height measured with respect to $z = 0$, and $z = \eta_b$ the height of the bottom boundary measured with respect to $z = 0$. Note that $z = 0$ is usefully chosen so that the area average of η_b is zero. Furthermore, with $\Delta\eta$ the deviation of the free surface relative to the averaged depth, H , volume conservation means that the area integral of $\Delta\eta$ vanishes. In summary, we have $\eta(x, y, t) = \eta_b(x, y) + h(x, y, t) = H + \Delta\eta(x, y, t)$ as well as $\int \eta_b \, dx \, dy = \int \Delta\eta \, dx \, dy = 0$. Additionally, we are concerned with fluctuations that leave the free surface monotonic; i.e., we do not consider overturns or breaking shallow water waves.

42.1.1 Pressure within the fluid layer

Figure 42.1 shows a single shallow water layer with a generally non-flat bottom and an undulating free surface height. We assume the fluid to be in hydrostatic balance, so that the vertical momentum equation reduces to

$$\frac{\partial p}{\partial z} = -\rho g. \quad (42.1)$$

Recall from Section 27.3 that the hydrostatic balance is consistent with lateral length scales much larger than vertical (small vertical to horizontal aspect ratio). Hence, a shallow water fluid is a relevant idealization if we are considering large horizontal scales relative to the vertical. This configuration is common for large-scale geophysical fluids.

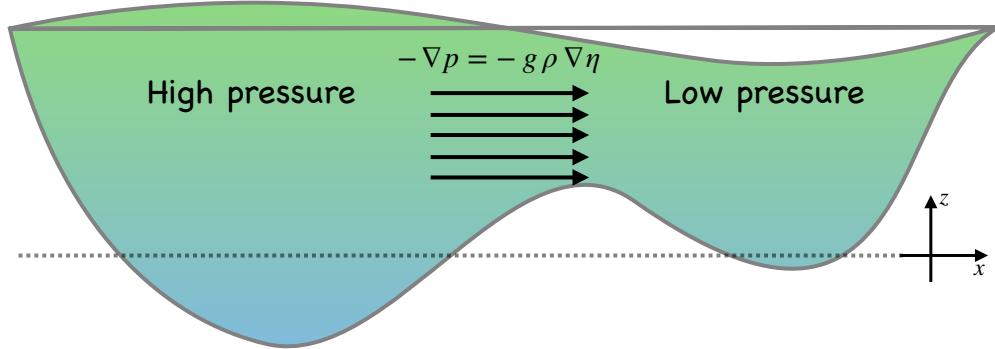


Figure 42.2: The horizontal acceleration from pressure within a single shallow water fluid layer is determined by the surface height, $-\nabla p = -g \nabla \eta$. The acceleration is uniform throughout the layer and points from sea level highs towards sea level lows.

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 layer

$$p(x, y, z, t) = p_o(x, y, t) + g \rho \int_z^\eta dz \quad (42.2a)$$

$$= p_o(x, y, t) + g \rho [\eta(x, y, t) - z], \quad (42.2b)$$

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

$$\nabla_z p = \nabla_z p_o + g \rho \nabla_z \eta. \quad (42.3)$$

Since p_o and η are independent of z , there is no need to expose the z subscript on the gradient operator on the right hand side. We thus drop the subscript when no ambiguity results. We generally ignore the applied surface pressure, p_o , since the fluid above the layer is assumed to have zero inertia. However, that assumption can be relaxed in order to study the effects of atmospheric pressure on a single layer ocean, for example. With $p_o = 0$, horizontal pressure forces within the fluid layer are determined solely by undulations in the free surface

$$\nabla_z p = g \rho \nabla \eta \quad \text{if } p_o = 0. \quad (42.4)$$

As depicted in Figure 42.2, acceleration from the pressure force is uniform throughout the layer and it points from sea level highs to sea level lows.

42.1.2 Momentum equation

If there is no friction anywhere in the fluid, then the horizontal momentum is effected only by the Coriolis and pressure forces. In this case, the horizontal momentum equation takes the form

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta, \quad (42.5)$$

where

$$\mathbf{v} = (\mathbf{u}, w) \quad (42.6)$$

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

The Coriolis parameter, f , is independent of depth, as is the horizontal pressure force. Consequently, if the horizontal velocity is initially independent of depth, it will remain so for all time. The material time derivative thus only has contributions from the local time derivative and from horizontal advection

$$\frac{D\mathbf{u}}{Dt} = \left[\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \right] \mathbf{u} \quad (42.7)$$

so that the shallow water momentum equation (42.5) takes on the form

$$\left[\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + f \hat{\mathbf{z}} \wedge \right] \mathbf{u} = -g \nabla \eta. \quad (42.8)$$

42.1.3 Thickness equation

The mass of a shallow water layer is constant in the absence of mixing, sources, or sinks. 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 Exercise 42.2. Note that since the fluid density is constant, mass conservation is the same as volume conservation. Hence, the terms “mass conservation” and “volume conservation” are commonly used interchangeably.

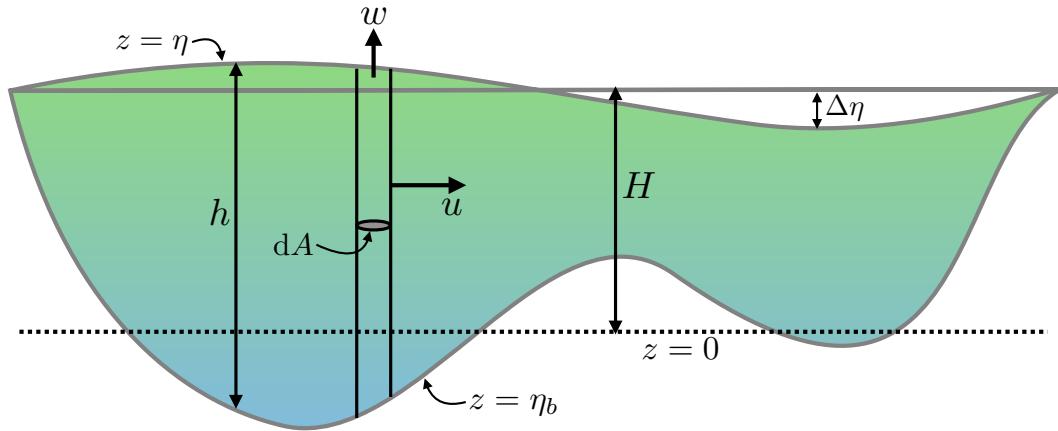


Figure 42.3: Mass budget for a column of shallow water fluid with cross-sectional area dA , constant density ρ , and thickness h . In the absence of boundary mass fluxes, the column mass is affected only by horizontal transport.

Consider an infinitesimal vertical column of shallow water fluid that is fixed in space. Let the horizontal cross-sectional area be written as dA and the thickness be

$$h = \eta - \eta_b \quad (42.9)$$

(see Figure 42.3). The total mass of fluid in this column is given by

$$M = \int dA \int_{\eta_b}^{\eta} \rho dz = \rho \int (\eta - \eta_b) dA = \rho \int h dA. \quad (42.10)$$

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

$$\frac{dM}{dt} = \rho \int \frac{\partial h}{\partial t} dA, \quad (42.11)$$

where

$$\frac{\partial h}{\partial t} = \frac{\partial (\eta - \eta_b)}{\partial t} = \frac{\partial \eta}{\partial t}, \quad (42.12)$$

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 vertical column boundaries (again, no mass crosses the top or bottom interfaces). The mass flux penetrating the vertical boundary is given by

$$\text{mass per time entering column} = -\rho \oint \mathbf{u} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (42.13)$$

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

$$\text{mass per time entering column} = -\rho \oint dl \int \mathbf{u} \cdot \hat{\mathbf{n}} dz, \quad (42.14)$$

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

$$-\rho \oint dl \int \mathbf{u} \cdot \hat{\mathbf{n}} dz = -\rho \oint h \mathbf{u} \cdot \hat{\mathbf{n}} dl \quad (42.15a)$$

$$= -\rho \int \nabla \cdot (h \mathbf{u}) dA, \quad (42.15b)$$

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 (42.11), and noting that the horizontal cross-sectional area is arbitrary, yields an equation for the layer thickness

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = 0. \quad (42.16)$$

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 (42.16) in the material form

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (42.17)$$

Hence, thickness of a material fluid column increases in regions where the horizontal flow converges.

Special case with a rectangular column

To help solidify our understanding of the step (42.15b) in the above derivation, consider the special case of a rectangular column, for which the mass per time of fluid entering the column is given by

$$\text{mass per time entering column} = -\rho \int [(u h)_{\text{east}} - (u h)_{\text{west}}] dy - \rho \int [(v h)_{\text{north}} - (v h)_{\text{south}}] dx. \quad (42.18)$$

Taking the limit as the column becomes infinitesimal leads to

$$\text{mass per time entering column} = -\rho \int \left[\frac{\partial(u h)}{\partial x} + \frac{\partial(v h)}{\partial y} \right] dx dy = -\rho \int \nabla \cdot (h \mathbf{u}) dA, \quad (42.19)$$

thus recovering the result (42.15b).

42.1.4 Eulerian flux-form momentum equation

The momentum and thickness equations are given in their Lagrangian form as

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (42.20)$$

To write the Eulerian flux-form momentum equation, we write the thickness weighted material acceleration as

$$h \frac{D\mathbf{u}}{Dt} = h \frac{D\mathbf{u}}{Dt} + \mathbf{u} \left[\frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} \right] = \frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot [\mathbf{u} (h \mathbf{u})], \quad (42.21)$$

so that the momentum equation takes the form

$$\frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot [\mathbf{u} (h \mathbf{u})] + \mathbf{f} \wedge (h \mathbf{u}) = -g h \nabla \eta. \quad (42.22)$$

Writing $\eta = \eta_b + h$ allows us to more clearly expose the contributions from bottom topography

$$\frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot [\mathbf{u} (h \mathbf{u})] + \mathbf{f} \wedge (h \mathbf{u}) = -g [(1/2) \nabla h^2 + h \nabla \eta_b]. \quad (42.23)$$

Expanding this vector equation into its two horizontal components renders

$$\frac{\partial(h u)}{\partial t} + \partial_x(h u^2 + g h^2/2) + \partial_y(h u v) - v h f = -g h \partial_x \eta_b \quad (42.24a)$$

$$\frac{\partial(h v)}{\partial t} + \partial_x(h u v) + \partial_y(h v^2 + g h^2/2) + u h f = -g h \partial_y \eta_b. \quad (42.24b)$$

We make use of this form when discussing force balances in a shallow water layer in Section 43.2.6.

42.1.5 Kinematic boundary conditions

Kinematic boundary conditions arise from geometric constraints placed on the fluid system. We consider here the kinematic boundary conditions at the ocean surface and bottom in the case where there is no flow through either interface. Recall from our discussion of fluid kinematics in Part III, we use the term *material surface* for any continuous surface or interface that is impenetrable to mass flow. The kinematics of such material surfaces is found throughout geophysical fluid mechanics. In Section 17.4 we derived the kinematic boundary conditions for a fluid, and we here apply those ideas to the shallow water system.

Bottom kinematic boundary condition

The ocean bottom is located at a vertical position

$$z = \eta_b(x, y). \quad (42.25)$$

This location can equivalently be specified mathematically by the surface

$$s(x, y, z) = \eta_b(x, y) - z = 0. \quad (42.26)$$

The outward normal at this surface is thus given by

$$\hat{\mathbf{n}} = \frac{\nabla s}{|\nabla s|} = \frac{\nabla \eta_b - \hat{\mathbf{z}}}{\sqrt{1 + \nabla \eta_b \cdot \nabla \eta_b}}. \quad (42.27)$$

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

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta_b. \quad (42.28)$$

That is, fluid can move tangentially to the bottom, but not normal to the bottom. Making use of the bottom outward normal (42.27) leads to

$$w = \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \eta_b. \quad (42.29)$$

For a flat bottom, with $\nabla \eta_b = 0$, the no-normal flow condition means that $w = 0$ at the bottom interface. But more generally, sloping bottoms lead to a nonzero vertical velocity component whose value depends on the projection of the horizontal velocity onto the bottom slope.

The kinematic result (42.29) is written in an Eulerian sense, with the velocity constrained to satisfy this relation at each point along the bottom interface. It has a complementary material interpretation based on acknowledging that the bottom interface is a material surface. A parcel 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. We can ensure this constraint by setting

$$\frac{Ds}{Dt} = \frac{D(z - \eta_b)}{Dt} = 0 \quad \text{at } z = \eta_b. \quad (42.30)$$

Rearrangement of this result leads to the Eulerian constraint (42.29). Equivalently, we can write this boundary condition in the form

$$w = \frac{D\eta_b}{Dt} \quad \text{at } z = \eta_b. \quad (42.31)$$

Surface kinematic boundary condition

We again assume the surface boundary is a material surface. Consequently, the surface kinematic boundary condition follows analogously to the bottom. However, there is a fundamentally new feature in that the fluid free surface is a time dependent moving boundary. The free surface is located at a vertical position $z = \eta$. Equivalently, the free surface can be specified by a surface of constant s , where

$$s(x, y, z, t) = z - \eta(x, y, t) = 0. \quad (42.32)$$

The outward normal to the free surface is thus given by

$$\hat{\mathbf{n}} = \frac{\nabla s}{|\nabla s|} = \frac{\hat{\mathbf{z}} - \nabla \eta}{\sqrt{1 + \nabla \eta \cdot \nabla \eta}}. \quad (42.33)$$

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

$$(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta(x, y, t), \quad (42.34)$$

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

$$\frac{\partial s}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla s = 0. \quad (42.35)$$

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

$$\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = -\frac{\partial_t s}{|\nabla s|} = \frac{\partial_t \eta}{\sqrt{1 + \nabla \eta \cdot \nabla \eta}}. \quad (42.36)$$

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 (42.36) in the no-normal flow constraint (42.34) then leads to the surface kinematic boundary condition

$$w - \mathbf{u} \cdot \nabla \eta = \frac{\partial \eta}{\partial t} \quad \text{at } z = \eta. \quad (42.37)$$

As for the bottom kinematic boundary condition written as (42.30), we can interpret this result materially, in which case

$$\frac{Ds}{Dt} = \frac{D(z - \eta)}{Dt} = 0 \quad \text{at } z = \eta. \quad (42.38)$$

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

$$w = \frac{D\eta}{Dt} \quad \text{at } z = \eta. \quad (42.39)$$

42.1.6 Stretching and vertical velocity

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

$$\nabla \cdot \mathbf{v} = 0 \Rightarrow \frac{\partial w}{\partial z} = -\nabla \cdot \mathbf{u}. \quad (42.40)$$

This result also follows since material parcels in the constant density shallow water layer maintain a constant volume (see Section 19.1). Furthermore, since the horizontal velocity has no depth dependence, we can vertically integrate the continuity equation from the bottom to an arbitrary depth within the layer to render

$$w(z) = w(\eta_b) - (z - \eta_b) \nabla \cdot \mathbf{u}, \quad (42.41)$$

so that the vertical velocity is a linear function of depth. Applying this equation at the ocean surface yields

$$w(\eta) = w(\eta_b) - (\eta - \eta_b) \nabla \cdot \mathbf{u}. \quad (42.42)$$

Eliminating the horizontal convergence between equations (42.41) and (42.42) leads to

$$w(z) - w(\eta_b) = \left(\frac{z - \eta_b}{\eta - \eta_b} \right) [w(\eta) - w(\eta_b)]. \quad (42.43)$$

Making use of the surface kinematic boundary condition (42.39) and bottom kinematic boundary condition (42.31) renders the material form

$$\frac{1}{z - \eta_b} \left(\frac{D(z - \eta_b)}{Dt} \right) = \frac{1}{\eta - \eta_b} \left(\frac{D(\eta - \eta_b)}{Dt} \right). \quad (42.44)$$

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

$$\frac{D}{Dt} \left(\frac{z - \eta_b}{h} \right) = 0. \quad (42.45)$$

Again, $h = \eta - \eta_b$ is the layer thickness and $z - \eta_b$ is the height of a fluid parcel from the bottom interface (see Figure 42.1). Consequently, equation (42.45) means that the ratio of the parcel height above the bottom to the layer thickness remains constant as the parcel moves through the shallow water fluid. That is, a column of shallow water fluid stretches or squeezes uniformly within a shallow water fluid. Shallow water dynamics thus comprises the dynamics of moving coherent fluid columns within a layer. This constrained behaviour results from the linear dependence with depth of the vertical velocity, which itself is a result of the depth independence of the horizontal velocity.

42.1.7 Further study

The shallow water model is ubiquitous in the geophysical fluid literature. An early application of the single shallow water layer was Laplace's studies of tides on a sphere (see exercise 42.4). Our presentation largely follows Chapter 3 of [Vallis \(2017\)](#) with another lucid discussion given by [Salmon \(1998\)](#).

42.2 Emphasizing the hydrostatic approximation

We revisit the above formulation to emphasize the hydrostatic approximation as the fundamental assumption leading to the shallow water equations. For this purpose, start from the compressible equations of motion for a perfect fluid in a rotating reference frame

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \quad (42.46a)$$

$$\rho \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \rho \mathbf{v} = -\hat{\mathbf{z}} g \rho - \nabla p. \quad (42.46b)$$

In a manner analogous to our discussion of the Boussinesq approximation in Section 28.1, decompose the density and pressure into a depth dependent term and a deviation

$$\rho(x, y, z, t) = \rho_r(z) + \rho'(x, y, z, t) \quad (42.47a)$$

$$p(x, y, z, t) = p_r(z) + p'(x, y, z, t) \quad (42.47b)$$

where the reference pressure is in hydrostatic balance with the reference density

$$\frac{dp_r}{dz} = -\rho_r g. \quad (42.48)$$

Inserting this decomposition into the equation of motion leads to

$$\rho \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \rho \mathbf{v} = -\hat{\mathbf{z}} g \rho' - \nabla p'. \quad (42.49)$$

42.2.1 Hydrostatic balance with respect to background density

Now introduce the following rather strict approximation: *The full fluid is in hydrostatic balance with respect to the background density, so that $\rho' = 0$.* As shown below, this assumption means that the density is a uniform constant and that the fluid is in turn incompressible.

42.2.2 Density is a uniform constant

With $\rho = \rho_r(z)$, the mass continuity equation implies

$$w \frac{d\rho_r}{dz} = 0. \quad (42.50)$$

For a nonzero vertical velocity, this constraint means that the density is itself a constant in space and time

$$\rho = \text{constant}. \quad (42.51)$$

The perturbation hydrostatic pressure is thus given by

$$p'(x, y, z, t) = g \rho (\eta - z), \quad (42.52)$$

where $\eta = \eta(x, y, t)$ is the free surface height. The horizontal momentum equation thus takes the form

$$\frac{D\mathbf{u}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} = -g \nabla \eta. \quad (42.53)$$

An initial horizontal velocity that is depth independent will thus remain such.

42.2.3 Fluid is incompressible

With a uniform density, the mass continuity implies that the fluid is incompressible

$$\nabla \cdot \mathbf{v} = 0. \quad (42.54)$$

With \mathbf{u} depth independent, we can depth integrate $\nabla \cdot \mathbf{v} = 0$ as well as the kinematic boundary conditions in Section 42.1.5 to render the thickness equation

$$\frac{1}{h} \frac{Dh}{Dt} = -\nabla \cdot \mathbf{u}. \quad (42.55)$$

42.2.4 Non-hydrostatic and hydrostatic forces in a homogeneous layer

In Section 36.1, we studied surface gravity waves in a homogeneous (constant density) fluid layer. Such waves have an amplitude that exponentially decays with depth so that the horizontal and vertical fluid motion associated with the waves have a non-zero vertical shear. Now the hydrostatic pressure within a homogeneous layer has a depth-independent horizontal gradient

$$\partial_z (\nabla_z p_{\text{hydrostatic}}) = -g \nabla_z \rho = 0, \quad (42.56)$$

so the hydrostatic pressure gradients can only drive a horizontal flow that is depth independent. Hence, a vertically sheared horizontal flow within a homogeneous fluid must arise from a non-hydrostatic pressure force. As discussed in Section 36.1.3, surface gravity waves indeed involve non-hydrostatic pressure forces that drive the vertical dependence to the wave amplitude.

In contrast, as shown in Section 27.3.3, a non-homogeneous density field, in particular density with a horizontal gradient, leads to a vertically dependent hydrostatic pressure gradient. This hydrostatic pressure force can in turn drive vertically sheared flow, with thermal wind the canonical example in rotating fluids (Section 29.3.4).

42.2.5 Comments

The basic assumption leading to the shallow water model is the hydrostatic approximation for the full fluid state relative to the background density ρ_r . Mass continuity then implies ρ is a uniform constant, further implying the fluid is incompressible. The shallow water momentum and thickness equations then follow.

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 (Section 27.3). The term “water” refers to the incompressible nature of the fluid, which is a more relevant approximation for water than for the atmosphere (see Section 28.1).

42.3 Shallow water fluid in a rotating tank

We introduced the geopotential in Section 14.1.2, which are surfaces where the effective gravitational force (sum of central gravity plus planetary centrifugal) is constant. Correspondingly, we introduced geopotential vertical coordinates in Section 14.2.3. We here revisit that discussion in relation to a shallow water fluid undergoing constant rotation about the vertical ($f = 2\Omega\hat{z}$). In this case, we see how the centrifugal acceleration (due to rotation of the tank) leads to a parabolic shape for the surface of a rotating tank fluid in solid-body rotation.

42.3.1 Equations of motion

The equation of motion for a fluid in a rotating tank is given by (see also Exercise 27.5)

$$\frac{D\mathbf{u}}{Dt} + f\hat{z} \wedge \mathbf{u} = -\nabla(p/\rho + g_e z - \Omega^2 r^2/2). \quad (42.57)$$

Note the use of both a gravitational potential and centrifugal potential. As for planetary motion we know these two potentials can be combined into a geopotential. However, we here keep them separate since we are interested in details of the motion as seen from the laboratory. We make use of polar coordinates as defined by

$$x = r \cos \theta \quad (42.58)$$

$$y = r \sin \theta, \quad (42.59)$$

with r the radial distance from the rotation axis and θ the angle made in a counter-clockwise direction from the positive x -axis.

42.3.2 Free surface shape in solid-body rotation

Consider a fluid at rest in the rotating frame, thus undergoing solid-body rotation. Static equilibrium in the rotating frame is realized when the forcing on the right hand side vanishes so that

$$p/\rho + g_e z - \Omega^2 r^2/2 = p_0/\rho, \quad (42.60)$$

where p_0 is a constant pressure to be specified below. At the free surface where $z = \eta = h$ (recall flat bottom), the pressure equals to that applied to the free surface by the overlying media, $p = p_a$ (e.g., atmospheric pressure). The equilibrium layer thickness is thus given by

$$h = \frac{(\Omega r)^2}{2g_e} + \left(\frac{p_0 - p_a}{\rho g_e} \right). \quad (42.61)$$

For simplicity, assume the applied pressure is spatially constant. Hence, we specify p_0 according to the thickness at $r = 0$, so that

$$h(r) = \frac{(\Omega r)^2}{2 g_e} + h(0). \quad (42.62)$$

This column thickness is therefore parabolic in shape, with increasing thickness moving away from the rotation axis.

42.3.3 Further study

The discussion in this section parallels that in Exercise 27.5. We consider the angular momentum for this system in Section 43.5. Refer to Section 6.6.4 of *Marshall and Plumb (2008)* for more discussion of rotating tank laboratory experiments.

42.4 Reduced gravity model for the upper ocean

The reduced gravity model describes an active layer of uniform density, ρ_1 , above an stagnant layer of density ρ_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 zero motion. In this way, we introduce the *level of no motion*, below which (baroclinic) currents vanish.

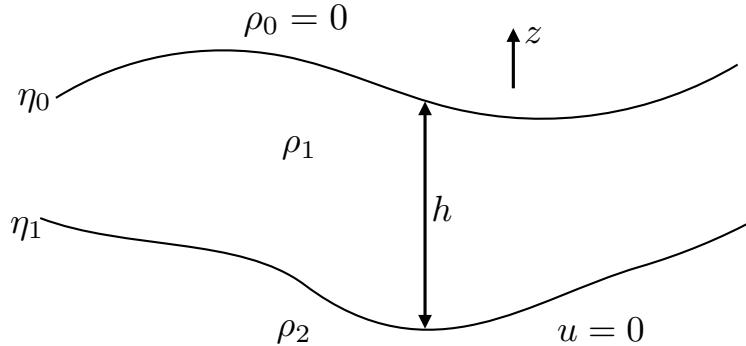


Figure 42.4: Reduced gravity model of shallow water fluid. The lower layer with density ρ_2 is dynamically inactive and thus has a zero velocity. The upper layer is dynamically active with thickness h and density ρ_1 . The dynamically active layer is bounded above by a zero density layer. The lower inactive layer is assumed to be infinitely deep so that its continuity equation can be ignored; i.e., even though there is zero currents within this layer it can still undulate.

42.4.1 Momentum and thickness equations for the active layer

We develop the momentum equations for the reduced gravity model by assuming hydrostatic balance, in which pressure at a depth, z , in the upper layer is computed as (see Figure 42.4)

$$p_1(x, y, z, t) = p_0(x, y, t) + g \rho_1 (\eta_0 - z). \quad (42.63)$$

Since the fluid above the upper layer has zero density, we set

$$p_0 = 0. \quad (42.64)$$

We immediately see that the horizontal momentum equation for the upper layer is written

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta_0. \quad (42.65)$$

The equations for the upper layer are completed by mass (volume) conservation that yields the thickness equation

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (42.66)$$

42.4.2 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

$$p_2(x, y, z, t) = g \rho_1 (\eta_0 - \eta_1) + g \rho_2 (\eta_1 - z). \quad (42.67)$$

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

$$\nabla_z p_2 = g \rho_1 \nabla(\eta_0 - \eta_1) + g \rho_2 \nabla \eta_1 = 0. \quad (42.68)$$

This relation provides a constraint that links undulations of the top and bottom interfaces of the dynamically active layer

$$\eta_0 = -\eta_1 \left[\frac{\rho_2 - \rho_1}{\rho_1} \right] + \text{constant}. \quad (42.69)$$

The density ratio on the right hand side is positive but typically much smaller than unity. Hence, the relation (42.69) means that undulations of the free surface, η_0 , are of opposite sign and of much smaller amplitude than undulations in the lower interface, η_1 . This behaviour is typical of undulations of the thermocline region of the ocean and the free surface (see Figure 42.5).

42.4.3 Momentum equation with reduced gravity

Relation (42.69) can be used to write the momentum equation (42.65) in the form

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = +g'_1 \nabla \eta_1, \quad (42.70)$$

where

$$g'_1 = g \left[\frac{\rho_2 - \rho_1}{\rho_1} \right] \ll g \quad (42.71)$$

defines the *reduced gravity*. It is more typical to make use of the momentum equation in the form (42.70), than the original form (42.65). The reason is that historically, ocean hydrography measurements¹ have allowed for an estimate of the pycnocline slope, $\nabla \eta_1$, whereas it was not until satellite altimetry measurements (post-1993) that we could estimate the sea level slope, $\nabla \eta_0$.

42.4.4 Further study

The material in this section is a summary of that in Section 3.2 of [Vallis \(2017\)](#). [Tomczak and Godfrey \(1994\)](#) make use of the 1.5-layer reduced gravity model for interpreting aspects of the observed ocean. Additional use is made by [Griffies et al. \(2014\)](#) for interpreting patterns of sea level in the ocean. Figure 42.5 is based on Figure 3.3 from [Tomczak and Godfrey \(1994\)](#) as well as Figure 37 from [Griffies et al. \(2014\)](#).

¹In oceanography, hydrography refers to measurements of temperature, salinity, and pressure; see [Talley et al. \(2011\)](#).

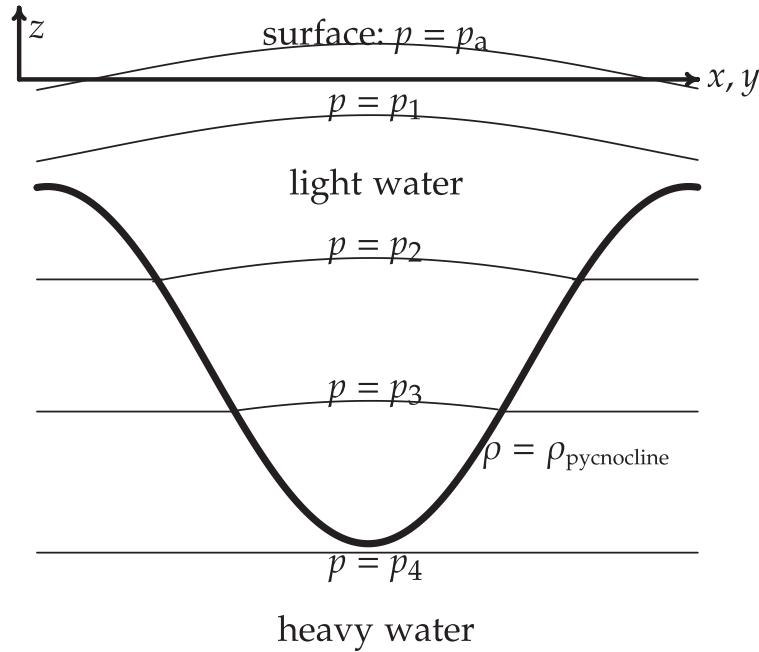


Figure 42.5: A vertical slice through a reduced gravity, or 1.5 layer, ocean in hydrostatic balance. Shown here is a plug of dynamically active light water, as may occur in a warm core eddy to the subtropical gyres, sitting on top of heavy water of zero motion. The free surface corresponds to η_0 in Figure 42.4, whereas the pycnocline (heavy black line) corresponds to the lower interface η_1 of Figure 42.4. 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 m. Note that the essential hydrostatic features of this example are contained in Figure 27.1 when considering the pressure difference between two columns of fluid with unequal density.

42.5 Stacked shallow water equations

We here consider two dynamically active shallow water layers as shown in Figure 42.6. This model offers the canonical tool for theoretical studies of baroclinic behaviour. The equations for more than two layers follow by induction from the two-layer case.

42.5.1 Model formulation

Each shallow water layer satisfies its own independent thickness equation, representing the conservation of mass for each layer

$$\frac{\partial h_1}{\partial t} + \nabla \cdot (h_1 \mathbf{u}_1) = 0 \quad (42.72)$$

$$\frac{\partial h_2}{\partial t} + \nabla \cdot (h_2 \mathbf{u}_2) = 0. \quad (42.73)$$

We emphasize that there is no coupling between these equations, as each layer separately must satisfy volume conservation.

We next need the pressure in each layer to formulate the pressure forces for driving flow. As before, make use of the hydrostatic balance and integrate down from the surface (assuming zero

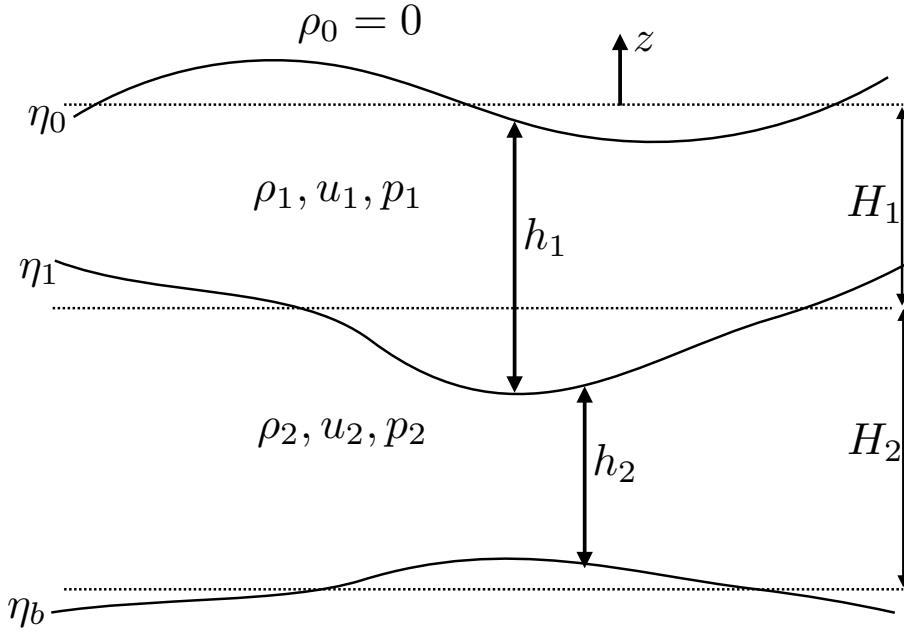


Figure 42.6: Two dynamically active layers of stacked shallow water fluid. The notation corresponds to that for the reduced gravity model of Figure 42.4, with two dynamically active layers shown in the present figure.

mass layer above), which results in the pressure fields

$$p_1 = \rho_1 g (\eta_0 - z) \quad (42.74)$$

$$p_2 = \rho_1 g (\eta_0 - \eta_1) + \rho_2 g (\eta_1 - z). \quad (42.75)$$

It is convenient to write pressure in layer-two using the reduced gravity, which leads to

$$p_2 = \rho_1 g (\eta_0 - \eta_1) + \rho_2 g (\eta_1 - z) \quad (42.76a)$$

$$= g \eta_1 (\rho_2 - \rho_1) + g \rho_1 \eta_0 - g \rho_2 z \quad (42.76b)$$

$$= \rho_1 \left[g \eta_0 + g \eta_1 \frac{\rho_2 - \rho_1}{\rho_1} \right] - g \rho_2 z \quad (42.76c)$$

$$= \rho_1 (g \eta_0 + g'_1 \eta_1) - g \rho_2 z, \quad (42.76d)$$

where we introduced the reduced gravity between layers one and two

$$g'_1 = g \left[\frac{\rho_2 - \rho_1}{\rho_1} \right]. \quad (42.77)$$

The terms \$g \rho_i z\$ appearing in the layer pressures \$p_i\$ have zero horizontal gradient. Hence, they play no dynamical role in determining the horizontal velocity.

The horizontal momentum equations for the two layers take the general form

$$\rho_1 \left[\frac{D_1 \mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 \right] = -\nabla p_1 \quad (42.78a)$$

$$\rho_2 \left[\frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 \right] = -\nabla p_2, \quad (42.78b)$$

where we introduced the layer material time derivatives

$$\frac{D_n}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}_n \cdot \nabla. \quad (42.79)$$

Making use of expressions (42.74) and (42.76d) for layer pressures leads to

$$\frac{D_1 \mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -g \nabla \eta_0 \quad (42.80)$$

$$\frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\frac{\rho_1}{\rho_2} (g \nabla \eta_0 + g'_1 \nabla \eta_1). \quad (42.81)$$

Finally, it is convenient to express pressure 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 (42.72) and (42.73). We thus write

$$\eta_0 = \eta_b + h_1 + h_2 \quad (42.82a)$$

$$\eta_1 = \eta_b + h_2, \quad (42.82b)$$

so that

$$p_1 = \rho_1 g (\eta_b + h_1 + h_2) \quad (42.83a)$$

$$p_2 = \rho_1 [g (\eta_b + h_1 + h_2) + g'_1 (\eta_b + h_2)], \quad (42.83b)$$

thus resulting in the momentum equations

$$\frac{D_1 \mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -g \nabla (\eta_b + h_1 + h_2) \quad (42.84a)$$

$$\frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\frac{\rho_1}{\rho_2} \nabla [g (\eta_b + h_1 + h_2) + g'_1 (\eta_b + h_2)]. \quad (42.84b)$$

Notice how layer thickness from one layer is coupled to the other layer through the 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. Finally, the Boussinesq approximation sets the density ratio ρ_1/ρ_2 to unity so that

$$\frac{D_1 \mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -g \nabla (\eta_b + h_1 + h_2) \quad (42.85a)$$

$$\frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\nabla [g (\eta_b + h_1 + h_2) + g'_1 (\eta_b + h_2)]. \quad (42.85b)$$

Notice how the difference in layer velocities, $\mathbf{u}_1 - \mathbf{u}_2$, is affected by a pressure gradient arising just from bottom topography and the interior layer thickness, h_2

$$\frac{D_1 \mathbf{u}_1}{Dt} - \frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge (\mathbf{u}_1 - \mathbf{u}_2) = g'_1 \nabla (\eta_b + h_2). \quad (42.86)$$

This vertical “shear” does not directly feel undulations of the free surface. Rather, it feels these surface undulations only indirectly via nonlinear terms appearing in the advection terms on the left hand side. We further discuss this result in Section 43.1.2 by introducing thermal wind and the Margules’ Relation.

42.5.2 Further study

We generalize the stacked shallow water equations in Chapter 45 by developing the equations for a Boussinesq isopycnal ocean model.

42.6 Exercises

EXERCISE 42.1: RELATIONS FOR VERTICAL VELOCITY (EXERCISE (3.2) OF *Vallis (2006)*)
 Show that the vertical velocity within a shallow water system is given by

$$w = \left(\frac{z - \eta_b}{h} \right) \frac{Dh}{Dt} + \frac{D\eta_b}{Dt}. \quad (42.87)$$

Interpret the result, showing that it gives sensible answers at the top and bottom of the fluid layer.

EXERCISE 42.2: SHALLOW WATER LAYER WITH SURFACE VOLUME SOURCES

In Section 42.1 we assumed a zero volume crossing the boundary of the shallow water fluid. Consequently, both the surface and bottom boundaries are material surfaces. For this exercise we introduce a surface volume source as occurs across the ocean surface through evaporation, precipitation, and river runoff. This surface volume transfer in turn means the surface boundary is no longer a material surface. For this problem, let \mathcal{V} be the volume per time per horizontal area of fluid entering across the surface of the shallow water layer (\mathcal{V} has dimensions of length per time and $\mathcal{V} > 0$ means volume enters the shallow water layer). Assume the water in \mathcal{V} has the same density and same velocity as the shallow water layer. Hence, there is no modification to the layer stratification (i.e., it remains a homogeneous layer).

- (a) Equations (42.38) and (42.39) offer equivalent expressions for the surface kinematic boundary conditions in the absence of a volume transfer across the surface. How are these expressions modified in the presence of $\mathcal{V} \neq 0$?
- (b) What is the layer thickness equation in the presence of $\mathcal{V} \neq 0$?
- (c) Equation (42.45) shows that in the absence of volume sources, a column of shallow water fluid stretches or squeezes uniformly. How is this relation modified in the presence of $\mathcal{V} \neq 0$?
- (d) Equation (47.25) shows that the potential vorticity $Q = (\zeta + f)/h$ is materially conserved for an inviscid shallow water fluid layer in the absence of volume sources, $DQ/Dt = 0$. How is this material conservation equation modified in the presence of $\mathcal{V} \neq 0$? Answering this question requires knowledge of the shallow water potential vorticity derivation given in Section 47.3.

EXERCISE 42.3: NON-ROTATING HYDRAULIC CONTROL

Consider the steady flow in a non-rotating shallow water layer where the flow is purely one-dimensional in the zonal direction.

- (a) Show that the steady flow satisfies the balance

$$\partial_x h [1 - Fr^2] = \partial_x \eta_b \quad (42.88)$$

where the Froude number is given by

$$Fr = \frac{u}{\sqrt{gh}}. \quad (42.89)$$

The Froude number is the ratio of the speed for a fluid particle to the speed of a shallow water gravity wave.

- (b) Discuss the case in which $\partial_x \eta_b = 0$ yet $\partial_x h \neq 0$. This case is known as *hydraulic control*.

EXERCISE 42.4: SHALLOW WATER EQUATIONS WITH TIDES

In Chapter 34 we derive the equations for a primitive equation ocean in the presence of astronomical forcing that leads to tides. Specialize the general results from that chapter to derive the thickness and momentum equations for a single layer of shallow water fluid in the presence of astronomical tidal forcing. As in Section 34.4, assume the perturbation geopotential is depth independent.

EXERCISE 42.5: REDUCED GRAVITY MODEL FOR THE ATMOSPHERE

Derive the shallow water equations for a single moving layer of fluid of density ρ_2 above a rigid floor, and where above the moving fluid is a stationary fluid of density ρ_1 , with $\rho_1 < \rho_2$. Show that as $\rho_1/\rho_2 \rightarrow 0$ the single layer shallow water equations emerge. Make use of notation from the two-layer system shown in Figure 42.6.

43

Shallow water dynamics

We focus in this chapter on developing dynamical features for the shallow water system, including geostrophy, thermal wind (in the form of Margules' relation), interfacial contact pressure forces (i.e., form stress), mechanical energy, and available potential energy.

READER'S GUIDE TO THIS CHAPTER

This chapter builds from the shallow water formulations in Chapter 42, and we make use of the dynamical results in many of the subsequent chapters.

43.1	Geostrophic balance and thermal wind	670
43.1.1	Geostrophy for a single shallow water layer	670
43.1.2	Geostrophy and thermal wind for two shallow water layers	670
43.1.3	Comments	671
43.2	Contact pressure forces and form stress	672
43.2.1	Contact pressure force along vertical sides	672
43.2.2	Contact pressure force along the top and bottom interfaces	674
43.2.3	Form stress	675
43.2.4	Net integrated contact pressure force on the column	675
43.2.5	Contact pressure force on a single layer	675
43.2.6	Forces on a single layer in a zonally re-entrant channel with topography	676
43.2.7	Forces on multiple layers in a zonally re-entrant channel with topography	678
43.3	Mechanical energy budget for a shallow water layer	678
43.3.1	Gravitational potential energy	679
43.3.2	Kinetic energy	679
43.3.3	Mechanical energy	680
43.4	Available potential energy of a shallow water layer	681
43.5	Angular momentum for fluid in a rotating cylindrical tank	681
43.5.1	Angular momentum for a column of shallow water fluid	682
43.5.2	Material time evolution of the angular momentum	682
43.5.3	Materially invariant angular momentum	683
43.5.4	Comments	684
43.6	Exercises	684

43.1 Geostrophic balance and thermal wind

As described in Chapter 29, geostrophic balance arises from dropping the material time derivative in the inviscid horizontal momentum equation. 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.

43.1.1 Geostrophy for a single shallow water layer

The geostrophic balance for a single shallow water fluid layer takes the form

$$\mathbf{f} \wedge \mathbf{u}_g = -g \nabla \eta, \quad (43.1)$$

or in component form

$$u_g = -\frac{g}{f} \frac{\partial \eta}{\partial y} \quad \text{and} \quad v_g = \frac{g}{f} \frac{\partial \eta}{\partial x}. \quad (43.2)$$

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 43.1 shows a schematic of the geostrophic balance for a single shallow water layer.

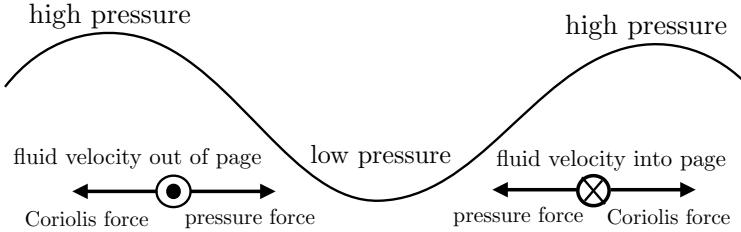


Figure 43.1: Side view of geostrophic balance for a single shallow water layer, here shown with two high pressure centers surrounding a low pressure center. The Coriolis force balances the pressure gradient force. In the northern hemisphere, where $f > 0$, geostrophic flow is counter-clockwise around a low pressure center and clockwise around a high pressure center.

43.1.2 Geostrophy and thermal wind for two shallow water layers

Now consider two shallow water layers as in Figure 42.6. Recall the layer pressure equations (42.74) and (42.75), which leads to the pressure difference

$$p_1 - p_2 = g \eta_1 (\rho_1 - \rho_2) + g z (\rho_2 - \rho_1) \quad (43.3a)$$

$$= -g'_1 \rho_1 \eta_1 + g'_1 \rho_1 z, \quad (43.3b)$$

$$= g'_1 \rho_1 (-\eta_1 + z), \quad (43.3c)$$

where the reduced gravity is given by

$$g'_1 = g \left[\frac{\rho_2 - \rho_1}{\rho_1} \right] \ll g. \quad (43.4)$$

The density difference $\rho_2 - \rho_1$ is generally much smaller than either density, so that $g'_1 \ll g$. For a Boussinesq shallow water system, the momentum equations are given by

$$\frac{D_1 \mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -\rho_1^{-1} \nabla p_1 \quad (43.5a)$$

$$\frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\rho_1^{-1} \nabla p_2, \quad (43.5b)$$

where we used the top layer density as the reference density for the Boussinesq fluid, and where we introduced the material time derivatives for each layer

$$\frac{D_n}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}_n \cdot \nabla. \quad (43.6)$$

Making use of the pressure difference (43.3c) renders

$$\frac{D_1 \mathbf{u}_1}{Dt} - \frac{D_2 \mathbf{u}_2}{Dt} + \mathbf{f} \wedge \Delta \mathbf{u} = -\rho_1^{-1} \nabla(p_1 - p_2) \quad (43.7a)$$

$$= g'_1 \nabla \eta_1, \quad (43.7b)$$

where

$$\Delta \mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2 \quad (43.8)$$

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

$$\mathbf{f} \wedge \Delta \mathbf{u}_g = g'_1 \nabla \eta_1, \quad (43.9)$$

which is equivalent to

$$f \Delta \mathbf{u}_g = -\hat{z} \wedge g'_1 \nabla \eta_1, \quad (43.10)$$

or in component form

$$\Delta u_g = +\frac{g'_1}{f} \frac{\partial \eta_1}{\partial y} \quad \Delta v_g = -\frac{g'_1}{f} \frac{\partial \eta_1}{\partial x}. \quad (43.11)$$

These equations represent 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. We illustrate this relation in Figure 43.2. The Margules relation is a discrete (two-layer) version of the thermal wind relation discussed in Section 29.4.5.

43.1.3 Comments

An alternative definition of reduced gravity in equation (43.4) uses the average density,

$$\bar{\rho} = (\rho_1 + \rho_2)/2 \quad (43.12)$$

for the denominator, in which case the modified reduced gravity is

$$\bar{g}'_1 = g \left[\frac{\rho_2 - \rho_1}{\bar{\rho}} \right]. \quad (43.13)$$

Correspondingly, the pressure gradient on the right hand side of the momentum equations (43.5a) and (43.5b) have $1/\rho_1$ replaced by $1/\bar{\rho}$. Even so, the equation for the vertical shear evolution, (43.7b), remains unchanged.

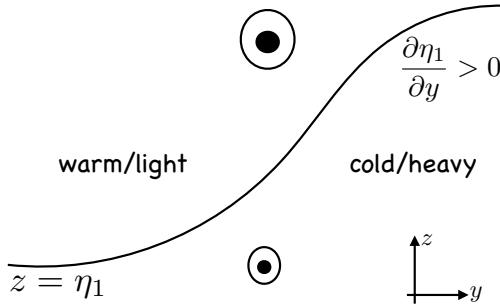


Figure 43.2: Illustrating Margule’s relation for the northern hemisphere ($f > 0$). Here we show the interface between a cold/heavy layer to the right and a warm/light layer to the left. The slope of the interface is positive, $\partial\eta_1/\partial y > 0$, thus leading to an increase in the eastward 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.

43.2 Contact pressure forces and form stress

Thus far in our discussion of shallow water models, we have considered the pressure force in the mathematical form of a pressure gradient acting within a fluid volume. In this way, the pressure force appears as a body force just like gravity and Coriolis. Alternatively, we may consider pressure to be a contact force per area acting at the interface between arbitrary fluid regions. If the contact pressure force integrates to a nonzero value over the region boundaries, then pressure accelerates the region.

As already discussed in Sections 26.1.3 and 32.2, the connection between the two forms for the pressure force arise through an application of Gauss’s divergence theorem to scalar fields (see Section 4.7.2)

$$\mathbf{F}_{\mathcal{R}}^{\text{press}} = - \int_{\mathcal{R}} \nabla p \, dV = - \oint_{\partial\mathcal{R}} p \hat{\mathbf{n}} \, dS. \quad (43.14)$$

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.

We discussed this dual representation of the pressure force per area in Section 22.1. Here, we pursue the contact force perspective as a means to understand the *form stress* acting between layers of shallow water fluid. There is also a form stress acting between a fluid layer and the solid earth, as well as between a fluid layer and the overlying atmosphere (when that atmosphere has a non-zero mass).

43.2.1 Contact pressure force along vertical sides

The pressure at a vertical position within a shallow water layer is given by (see Figure 43.3)

$$p(z) = \rho_i g (\eta_{i-1} - z) + p_{i-1}. \quad (43.15)$$

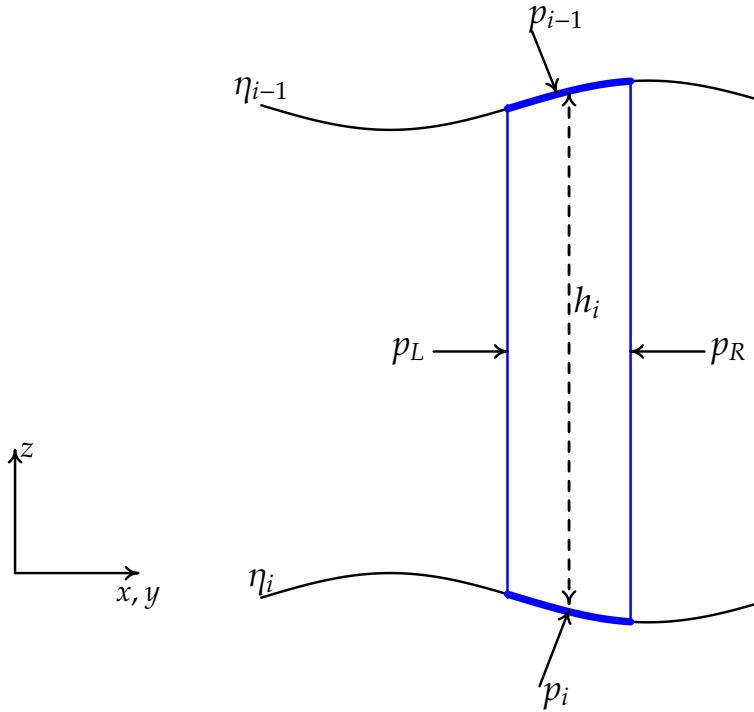


Figure 43.3: A schematic of the contact pressure force per area acting on the boundaries of a vertical column region within a shallow water layer of density ρ_i . The horizontal cross-sectional area of the column is depth independent. The interface at the lower boundary is at the vertical position $z = \eta_i$, and the upper interface is at $z = \eta_{i-1}$. The layer thickness is the difference between the interface positions, $h_i = \eta_{i-1} - \eta_i$. The boundaries of the blue region feel a contact pressure acting inward. The left boundary experiences a pressure p_L ; the right boundary has p_R ; the upper interface has a pressure p_{i-1} acting between the layer $i - 1$ and layer i , and the lower interface has a pressure p_i acting between the layer $i + 1$ and layer i . Note that pressures are continuous across each interface, according to Newton's Third Law. The area integral of the pressure force per area around the region leads to a net pressure force acting on the region.

Integrating this pressure over the layer thickness yields

$$\int_{\eta_i}^{\eta_{i-1}} p(z) dz = \rho_i g [\eta_{i-1} (\eta_{i-1} - \eta_i) - (1/2) (\eta_{i-1}^2 - \eta_i^2)] + p_{i-1} h_i \quad (43.16a)$$

$$= \rho_i g h_i^2 / 2 + h_i p_{i-1}. \quad (43.16b)$$

For simplicity assume the column to be rectangular. The zonal pressure force, in the limit that the column becomes thin, takes the form

$$dy \int_{\eta_i}^{\eta_{i-1}} (p_L - p_R) dz = -dx dy \left((1/2) \rho_i g \frac{\partial h_i^2}{\partial x} + \frac{\partial (h_i p_{i-1})}{\partial x} \right) \quad (43.17a)$$

$$= -dx dy \frac{\partial}{\partial x} ((1/2) \rho_i g h_i^2 + h_i p_{i-1}), \quad (43.17b)$$

where $dx dy$ is the horizontal cross-sectional area of the column. A similar result holds for the meridional direction, thus rendering the net contact pressure force acting on the vertical sides

$$\mathbf{F}_{\text{sides}}^{\text{press}} = -dx dy \nabla ((1/2) \rho_i g h_i^2 + h_i p_{i-1}). \quad (43.18)$$

That is, the contact force on the sides reduces, as the column becomes thin, to a gradient force. This exercise reveals no more than the integral theorem (43.14). However, it is useful as a means to see just how the integral theorem manifests within the shallow water system.

43.2.2 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. We term the horizontal pressure acting on the sloped interface the *form stress*. In addition to acting between two fluid layers with sloped interfaces, form stress also acts between a fluid and the solid-earth bottom, as well as the fluid and the atmosphere.

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

$$\hat{\mathbf{n}}_{i-1} = \frac{\nabla(z - \eta_{i-1})}{|\nabla(z - \eta_{i-1})|} = \frac{\hat{\mathbf{z}} - \nabla\eta_{i-1}}{\sqrt{1 + (\nabla\eta_{i-1})^2}}. \quad (43.19)$$

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

$$\hat{\mathbf{n}}_{i-1} = \frac{\hat{\mathbf{z}} - \hat{\mathbf{x}} \partial_x \eta_{i-1}}{\sqrt{1 + (\partial_x \eta_{i-1})^2}} \quad (43.20a)$$

$$= \frac{\hat{\mathbf{z}} - \hat{\mathbf{x}} \tan \phi_{i-1}}{\sqrt{1 + \tan^2 \phi_{i-1}}} \quad (43.20b)$$

$$= (\hat{\mathbf{z}} - \hat{\mathbf{x}} \tan \phi_{i-1}) \cos \phi_{i-1}, \quad (43.20c)$$

where we defined the interface slope as

$$\frac{\partial \eta_{i-1}}{\partial x} = \tan \phi_{i-1}, \quad (43.21)$$

with ϕ_{i-1} the angle between the horizontal plane and the interface. Trigonometry leads to an expression for the area of the top of the column¹

$$d\mathcal{S}_{i-1} = \frac{dx dy}{\cos \phi_{i-1}}. \quad (43.22)$$

Hence, the product of the area and the outward normal is given by

$$\hat{\mathbf{n}}_{i-1} d\mathcal{S}_{i-1} = dx dy (\hat{\mathbf{z}} - \hat{\mathbf{x}} \partial_x \eta_{i-1}). \quad (43.23)$$

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

$$\hat{\mathbf{n}}_{i-1} d\mathcal{S}_{i-1} = dx dy (\hat{\mathbf{z}} - \nabla \eta_{i-1}). \quad (43.24)$$

The contact pressure force at the top of the column is therefore given by

$$\mathbf{F}_{\text{top}}^{\text{press}} = -dx dy (\hat{\mathbf{z}} - \nabla \eta_{i-1}) p_{i-1}. \quad (43.25)$$

Analogous considerations lead to the contact pressure force at the bottom of the column

$$\mathbf{F}_{\text{bot}}^{\text{press}} = dx dy (\hat{\mathbf{z}} - \nabla \eta_i) p_i. \quad (43.26)$$

¹Equation (43.22) was also found in Section 17.4.3 when developing the kinematic boundary condition for a material interface.

43.2.3 Form stress

As noted earlier, form stress is the horizontal projection of the contact pressure acting on the sloped top or bottom interface of the fluid column. The corresponding force is the area multiplied by the form stress so that

$$\mathbf{F}_{\text{top}}^{\text{form stress}} = dx dy (p_{i-1} \nabla \eta_{i-1}) \quad (43.27)$$

$$\mathbf{F}_{\text{bott}}^{\text{form stress}} = -dx dy (p_i \nabla \eta_i). \quad (43.28)$$

These forces are associated with the vertical exchange of horizontal momentum. This momentum exchange occurs without any exchange of matter. Rather, it is an inviscid exchange that occurs only through the mechanical imbalance of forces on the interfaces.

For a specific case, consider a top interface that slopes upward in the \hat{x} direction (e.g., see Figure 43.3). Form stress acts on this interface to accelerate the column in the $+ \hat{x}$ direction. For the bottom interface, a negatively sloped bottom interface is accelerated in the $+ \hat{x}$ direction (e.g., see Figure 43.3). Form stress at the ocean surface arises from the weight of the atmosphere above. Likewise, form stress at the ocean bottom arises from an exchange of momentum between the fluid and the solid-earth.

43.2.4 Net integrated contact pressure force on the column

Summing the contact pressure forces (43.18), (43.25), and (43.26) leads to the net pressure force

$$\mathbf{F}_{\text{net}}^{\text{press}} = -dx dy [\nabla ((1/2) \rho_i g h_i^2 + h_i p_{i-1}) + (\hat{z} - \nabla \eta_{i-1}) p_{i-1} - (\hat{z} - \nabla \eta_i) p_i] \quad (43.29a)$$

$$= -M_i g \hat{z} - dx dy [\nabla ((1/2) \rho_i g h_i^2 + h_i p_{i-1}) - p_{i-1} \nabla \eta_{i-1} + p_i \nabla \eta_i]. \quad (43.29b)$$

To reach this result, we used the hydrostatic relation for the vertical force

$$p_{i-1} - p_i = -\rho_i g h_i, \quad (43.30)$$

and introduced the column mass

$$M_i = \rho_i h_i dx dy. \quad (43.31)$$

The vertical component of the net contact pressure force is therefore the weight of the column, which is expected since the fluid is assumed to be in hydrostatic balance. The horizontal contact pressure force arises from a total horizontal gradient plus the form stress at the surface and bottom interfaces. In the ocean, the gradient term is removed when integrating horizontally over the full domain, assuming 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.

43.2.5 Contact pressure force on a single layer

As a check on our calculation of the contact pressure force (43.29b), consider a single shallow water layer under a massless atmosphere. In this case, the contact pressure force per mass is given by

$$\frac{\mathbf{F}_{\text{net}}^{\text{press}}}{M} = -g \hat{z} - \frac{1}{\rho h} [\nabla ((1/2) \rho g h^2) + p_b \nabla \eta_b] \quad (43.32a)$$

$$= -g \hat{z} - \frac{1}{\rho h} [\rho g h \nabla h + \rho g h \nabla \eta_b] \quad (43.32b)$$

$$= -g \hat{z} - g \nabla \eta. \quad (43.32c)$$

The final equality made use of the identity

$$\eta = h + \eta_b \quad (43.33)$$

as shown in Figure 42.1. As expected, the horizontal component of this force equals to the pressure gradient body force per mass detailed in Section 42.1.1.

43.2.6 Forces on a single layer in a zonally re-entrant channel with topography

To illustrate pressure form stress, consider flow in a single shallow water layer in a zonally re-entrant channel with solid vertical boundaries at the north and south and nonzero topography in the interior. Applying a zonal surface stress, τ^x (dimensions of force per area), inserts momentum to the fluid. What are the domain integrated zonal and meridional force balances at steady state? This question was examined in the context of the axial angular momentum budget in Section 26.10.5. We find it useful to here revisit this question in the context of the shallow water system.

Volume transport for steady flow

Before considering the steady state force balance, let us establish a constraint based on volume conservation for the single layer of fluid by considering the thickness equation

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = 0. \quad (43.34)$$

At steady state the thickness weighted flow is horizontally non-divergent, $\nabla \cdot (h \mathbf{u}) = 0$. Zonally integrating the steady flow and using zonal periodicity renders $\int \partial_y(h v) dx = 0$. Hence, at steady state there can be no meridional accumulation of meridional transport when integrated over a latitude circle. Furthermore, since the meridional flow vanishes at both the southern and northern vertical walls, the only way to ensure zero convergence of zonally integrated meridional transport is for

$$\int h v dx = 0 \quad (43.35)$$

along each latitude circle. That is, the zonal integral of the meridional transport (dimensions of length cube per time) vanishes in the steady state. This constraint reflects the inability of the steady flow to build up or deplete the fluid along a latitude circle.

Area integrated meridional force balance

To study the meridional force balance we make use of the Eulerian flux-form momentum equation (42.24b)

$$\frac{\partial(h v)}{\partial t} + \partial_x(h u v) + \partial_y(h v^2 + g h^2/2) + u h f = -g h \partial_y \eta_b. \quad (43.36)$$

Integrating horizontally over the total area of the channel removes the zonal transport term, $\partial_x(h u v)$, due to periodicity. Furthermore, the no-normal flow boundary condition eliminates the contribution from $\partial_y(h v^2)$. We are thus left with the area integrated meridional momentum balance

$$\frac{d}{dt} \int h v dA = - \int [g h \partial_y \eta + u h f] dA, \quad (43.37)$$

where we set $\eta = \eta_b + h$ according to Figure 42.1. In a steady state, the Coriolis force arising from zonal flow balances the meridional gradient in the surface height

$$\int h(u f + g \partial_y \eta) dA = 0. \quad (43.38)$$

Steady volume integrated zonal geostrophic transport is supported by the northern and southern walls that allow for a nonzero meridional gradient in surface height.

Area integrated zonal force balance

Now consider the zonal momentum equation written in its flux-form (see equation (42.24a))

$$\frac{\partial(h u)}{\partial t} + \partial_x(h u^2 + g h^2/2) + \partial_y(h u v) - v h f = -g h \partial_x \eta_b + \tau^x / \rho. \quad (43.39)$$

Integrating horizontally over the total area of the channel removes all nonlinear transport terms from the left hand side: zonal periodicity eliminates the contribution from $\partial_x(h u^2 + g h^2/2)$ and the no-normal flow boundary condition eliminates the contribution from $\partial_y(h u v)$. Furthermore, volume conservation in the form of equation (43.35) eliminates the Coriolis force. We are thus left with the area integrated zonal momentum balance

$$\frac{d}{dt} \int \rho h u dA = \int [-g h \rho \partial_x \eta_b + \tau^x] dA. \quad (43.40)$$

We identify the pressure at the bottom of the layer due to the constant density fluid within the layer

$$p_{\text{bot}} = g h \rho, \quad (43.41)$$

so that the area integrated zonal momentum equation is

$$\frac{d}{dt} \int h \rho u dA = \int [-p_{\text{bot}} \partial_x \eta_b + \tau^x] dA. \quad (43.42)$$

A steady state is realized when there is a balance between the area integrated contact forces from zonal surface stress and bottom pressure form stress

$$\int p_{\text{bot}} \partial_x \eta_b dA = \int \tau^x dA. \quad (43.43)$$

Correlation between surface height and topography slope

Let us pursue the balance (43.43) for the particular case of an eastward zonal wind stress, $\tau^x > 0$, that comes into balance with a westward bottom form stress. A westward bottom form stress is established by 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 thickness of the column. Hence, to establish the anomalous bottom pressures there must be an anomalously thick fluid column upstream of topographic bumps and thin fluid column downstream. This situation is illustrated in Figure 43.4.

We can more explicitly reveal the correlation between surface height and bottom topography by writing $p_{\text{bot}} = \rho g h$ and using $\eta = h + \eta_b$ so that

$$p_{\text{bot}} \partial_x \eta_b = \rho g (\eta - \eta_b) \partial_x \eta_b = \rho g \eta \partial_x \eta_b - (g/2) \partial_x \eta_b^2. \quad (43.44)$$

The balance (43.43) thus becomes

$$\rho g \int \eta \partial_x \eta_b dA = \int \tau^x dA. \quad (43.45)$$

Hence, if $\int \tau^x dA > 0$, then the surface height is positively correlated with the bottom slope, $\int \eta \partial_x \eta_b dA > 0$, so that the surface height is high where topography slopes are positive and low where topography slopes are negative.

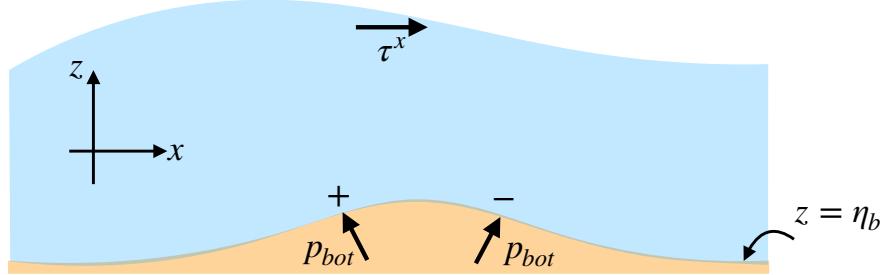


Figure 43.4: Side view of the wind stress and bottom pressure acting on a steady state layer of shallow water fluid flowing over a mountain. The eastward surface stress acts in the $+\hat{x}$ direction. As the shallow water layer acts on the solid-earth through its bottom pressure, the solid earth in turn reacts back (Newton's 3rd law). This force from the earth acting on the fluid constitutes the form stress on the bottom interface of the fluid column. In regions where the topographic slope is positive, $\partial\eta_b/\partial x > 0$, form stress provides an acceleration in the $-\hat{x}$ direction, and conversely where the slope is negative. The greater the slope, the greater the zonal acceleration from the form stress. When zonally integrating the steady state flow around the channel, the Coriolis force vanishes due to volume conservation, thus leading to a steady state domain integrated balance between bottom form stress and wind stress. For the bottom 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. Compare this figure to the analogous schematic in Figure 26.10.

43.2.7 Forces on multiple layers in a zonally re-entrant channel with topography

The above balances are modified for a stack of shallow water layers, especially when allowing for the vertical transfer of volume between the layers to admit a meridional-depth overturning circulation. In this case, there can be net meridional motion along a latitude circle to thus add the Coriolis force to the steady force balance, including an Ekman transport (balance between Coriolis and surface stress as in Section 31.1) for the layer feeling the zonal surface stress. Realization of a steady state balance requires the surface stress imparted to the upper layer to be vertically transmitted through sloped layer interfaces to the bottom topography. Section 21.7 of [Vallis \(2017\)](#) provides a discussion of flow in the Antarctic Circumpolar Current, in which interfacial pressure form stress developed from baroclinic eddies provides a mechanism for transferring horizontal momentum from the surface stress to the solid-earth bottom. We here illustrate the basic features of this interfacial form stress, and its connection to thickness fluxes.

43.3 Mechanical energy budget for a shallow water layer

We here derive the mechanical energy budget for a single shallow water layer sitting on top of a generally non-flat bottom.

43.3.1 Gravitational potential energy

The gravitational potential energy per horizontal area of a shallow water fluid of constant density is given by

$$\mathcal{P} = g \rho \int_{\eta_b}^{\eta} z \, dz \quad (43.46a)$$

$$= \frac{g \rho}{2} (\eta^2 - \eta_b^2) \quad (43.46b)$$

$$= g \rho h (\eta - h/2). \quad (43.46c)$$

Note how the gravitational potential energy vanishes when $\eta^2 = \eta_b^2$. For the case $\eta = \eta_b$, there is no fluid since the free surface sits on top of the bottom, 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. Since potential energy is computed with respect to a reference state $z = 0$, potential energy vanishes for the case where the same mass of fluid sits beneath $z = 0$ as above. Furthermore, note that in the flat bottom case, $\eta_b = 0$ so that $h = \eta - \eta_b = \eta$, in which case the potential energy (43.46c) reduces to

$$\mathcal{P}_{\text{flat}} = g \rho h^2 / 2. \quad (43.47)$$

The material time tendency of the potential energy written in the form (43.46b) is

$$\frac{D\mathcal{P}}{Dt} = g \rho \left[\eta \frac{D\eta}{Dt} - \eta_b \frac{D\eta_b}{Dt} \right] \quad (43.48a)$$

$$= g \rho (\eta w_\eta - \eta_b w_b), \quad (43.48b)$$

where we used equations for the vertical velocity component from Section 42.1.6. This equation states that the potential energy increases if the thickness of the layer increases through vertical motion along the top and bottom interfaces. Making further use of these equations renders

$$\frac{D\mathcal{P}}{Dt} = g \rho (\eta w_\eta - \eta_b w_b) \quad (43.49a)$$

$$= g \rho [\eta (w_b - h \nabla \cdot \mathbf{u}) - \eta_b w_b] \quad (43.49b)$$

$$= g \rho [w_b (\eta - \eta_b) - \eta h \nabla \cdot \mathbf{u}] \quad (43.49c)$$

$$= g \rho h (w_b - \eta \nabla \cdot \mathbf{u}), \quad (43.49d)$$

where we used the definition $h = \eta - \eta_b$. We make use of this form of the potential energy change for developing the mechanical energy budget in Section 43.48b.

43.3.2 Kinetic energy

As for the flat bottom case, the kinetic energy per horizontal area is

$$\mathcal{K} = \frac{1}{2} \int_{\eta_b}^{\eta} \rho \mathbf{u}^2 \, dz = \frac{1}{2} \rho h \mathbf{u}^2, \quad (43.50)$$

which has a material time derivative given by

$$\frac{D\mathcal{K}}{Dt} = -g \rho h \mathbf{u} \cdot \nabla \eta - \frac{\rho h \mathbf{u}^2}{2} \nabla \cdot \mathbf{u} \quad (43.51a)$$

$$= -g \rho h \mathbf{u} \cdot \nabla \eta - \mathcal{K} \nabla \cdot \mathbf{u}. \quad (43.51b)$$

$$= -g \rho h \mathbf{u} \cdot \nabla \eta + \frac{\mathcal{K}}{h} \frac{Dh}{Dt}. \quad (43.51c)$$

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 increases if the flow has a component that is directed down the pressure gradient. The second term increases kinetic energy for cases where the thickness increases via the convergence of horizontal velocity.

43.3.3 Mechanical energy

The material time derivative of the mechanical energy per horizontal area is given by

$$\frac{D(\mathcal{K} + \mathcal{P})}{Dt} = -g h \rho \mathbf{u} \cdot \nabla \eta - \mathcal{K} \nabla \cdot \mathbf{u} + g h \rho (w_b - \eta \nabla \cdot \mathbf{u}). \quad (43.52)$$

Note there is no internal energy for the shallow water fluid, so the mechanical energy equals to the total energy.

Expanding the material time derivative in equation (43.52) into its Eulerian components leads to

$$\frac{\partial}{\partial t} (\mathcal{K} + \mathcal{P}) + \mathbf{u} \cdot \nabla (\mathcal{K} + \mathcal{P}) = -g h \rho \mathbf{u} \cdot \nabla \eta - \mathcal{K} \nabla \cdot \mathbf{u} + g h \rho (w_b - \eta \nabla \cdot \mathbf{u}), \quad (43.53)$$

with rearrangement rendering

$$\frac{\partial}{\partial t} (\mathcal{K} + \mathcal{P}) + \nabla \cdot [\mathbf{u} (\mathcal{K} + \mathcal{P})] = \mathcal{P} \nabla \cdot \mathbf{u} - g h \rho \mathbf{u} \cdot \nabla \eta + g h \rho (w_b - \eta \nabla \cdot \mathbf{u}) \quad (43.54a)$$

$$= (\mathcal{P} - g h \rho \eta) \nabla \cdot \mathbf{u} + g h \rho (w_b - \mathbf{u} \cdot \nabla \eta) \quad (43.54b)$$

$$= -(g \rho h^2 / 2) \nabla \cdot \mathbf{u} + g h \rho (w_b - \mathbf{u} \cdot \nabla \eta) \quad (43.54c)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2) + g \rho h \mathbf{u} \cdot \nabla h + g h \rho (w_b - \mathbf{u} \cdot \nabla \eta) \quad (43.54d)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2) + g \rho h (\mathbf{u} \cdot \nabla (h - \eta) + w_b) \quad (43.54e)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2) + g \rho h (-\mathbf{u} \cdot \nabla \eta_b + w_b) \quad (43.54f)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2), \quad (43.54g)$$

where we used the identity

$$w_b = \frac{D\eta_b}{Dt} = \mathbf{u} \cdot \nabla \eta_b, \quad (43.55)$$

which follows since $\partial \eta_b / \partial t = 0$. We are thus left with the conservation law of the form

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathbf{F} = 0, \quad (43.56)$$

which has the specific expression

$$\frac{\partial}{\partial t} (\mathcal{K} + \mathcal{P}) + \nabla \cdot [\mathbf{u} (\mathcal{K} + \mathcal{P} + g \rho h^2 / 2)] = 0, \quad (43.57)$$

where the total mechanical energy is

$$\mathcal{E} = \mathcal{K} + \mathcal{P} = \frac{1}{2} \rho h \mathbf{u}^2 + \frac{1}{2} \rho g (\eta^2 - \eta_b^2), \quad (43.58)$$

and the flux of mechanical energy is

$$\mathbf{F} = \mathbf{u} (\mathcal{K} + \mathcal{P} + g \rho h^2 / 2) \quad (43.59a)$$

$$= \mathbf{u} (\mathcal{K} + \rho g \eta h). \quad (43.59b)$$

Recall in our discussion of total energy for a continuously stratified fluid in Section 24.2.2, where we identified the the mechanical injection work term that appears in the energy flux in addition to the advection of energy. Here, that term is given by $g \rho h^2 / 2$ for the shallow water fluid.

43.4 Available potential energy of a shallow water layer

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 that amount of the gravitational potential energy that can be converted to kinetic energy through a reversible rearrangement of the fluid.

The gravitational potential energy for a single shallow water layer is given by

$$P = g \rho \int dA \int_0^\eta z dz = \frac{g \rho}{2} \int \eta^2 dA, \quad (43.60)$$

where $\int dA$ is the horizontal integral over the full domain of the fluid and we choose to measure the potential energy relative to $z = 0$. The background or reference potential energy is realized by relaxing the sea surface interface to a uniform value $z = H$, so that

$$P_{\text{ref}} = \frac{g \rho}{2} \int H^2 dA. \quad (43.61)$$

The available potential energy is the difference

$$E_{\text{APE}} = P - P_{\text{ref}} \quad (43.62a)$$

$$= \frac{g \rho}{2} \int (\eta^2 - H^2) dA \quad (43.62b)$$

$$= \frac{g \rho}{2} \int (\eta^2 - \bar{\eta}^2) dA \quad (43.62c)$$

$$= \frac{g \rho}{2} \int (\eta')^2 dA, \quad (43.62d)$$

where

$$\bar{\eta} = \frac{\int \eta dA}{A} \quad (43.63)$$

is the area averaged surface height, and

$$\eta' = \eta - \bar{\eta} \quad (43.64)$$

is the anomalous sea surface. Furthermore, we set

$$\bar{\eta} = H \quad (43.65)$$

since the total volume of the shallow water layer is constant. Equation (43.62d) shows that the APE is non-negative for the shallow water layer. That is, any slope to the shallow water layer represents a store of positive APE.

43.5 Angular momentum for fluid in a rotating cylindrical tank

We here study angular momentum for a layer of inviscid shallow water fluid in a rotating cylindrical tank. This system was first discussed in Section 42.3, where we developed the horizontal equation of motion

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla (p/\rho + g_e z - \Omega^2 r^2/2), \quad (43.66)$$

where $r^2 = x^2 + y^2$ is the radial distance from the rotational axis,

$$\Omega = f/2 \quad (43.67)$$

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 Chapter 10.3) in the following, in which case

$$x = r \cos \theta \quad (43.68a)$$

$$y = r \sin \theta, \quad (43.68b)$$

with the polar angle θ measured counter-clockwise from the positive x -axis.

43.5.1 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 (see Sections 15.7 and 22.4)

$$L^z = \delta M [\mathbf{x} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] \cdot \hat{\mathbf{z}}, \quad (43.69)$$

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 solid-body rotation velocity is

$$\mathbf{U}_{\text{solid}} = (f/2) \hat{\mathbf{z}} \wedge \mathbf{x} = r \Omega \hat{\boldsymbol{\theta}}, \quad (43.70)$$

where $\hat{\mathbf{z}} \wedge \hat{\mathbf{r}} = \hat{\boldsymbol{\theta}}$ is the azimuthal unit vector pointing counter-clockwise around the origin.

We can further massage the expression for the angular momentum by writing

$$\mathbf{x} \wedge \mathbf{u} = (x v - y u) \hat{\mathbf{z}} = r^2 \dot{\theta} \hat{\mathbf{z}}, \quad (43.71)$$

where $\dot{\theta} = D\theta/Dt$ is the angular velocity. Likewise, we have

$$\mathbf{x} \wedge \mathbf{U}_{\text{solid}} = r^2 \Omega \hat{\mathbf{z}}, \quad (43.72)$$

so that the angular momentum can be written

$$L^z = \delta M [\mathbf{x} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] \cdot \hat{\mathbf{z}} = \delta M r^2 (\dot{\theta} + \Omega). \quad (43.73)$$

43.5.2 Material time evolution of the angular momentum

The material time evolution for the angular momentum is given by

$$\frac{DL^z}{Dt} = \delta M [\mathbf{u} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] \cdot \hat{\mathbf{z}} + \delta M \left[\mathbf{x} \wedge \left(\frac{D\mathbf{u}}{Dt} + \frac{D\mathbf{U}_{\text{solid}}}{Dt} \right) \right] \cdot \hat{\mathbf{z}}. \quad (43.74)$$

Using the solid-body rotation velocity given by equation (43.70), and with a constant rotation rate, yields

$$\mathbf{u} \wedge \mathbf{U}_{\text{solid}} + \mathbf{x} \wedge \frac{D\mathbf{U}_{\text{solid}}}{Dt} = \mathbf{u} \wedge (\boldsymbol{\Omega} \wedge \mathbf{x}) + \mathbf{x} \wedge (\boldsymbol{\Omega} \wedge \mathbf{u}) \quad (43.75a)$$

$$= (\mathbf{x} \cdot \mathbf{u}) f \hat{\mathbf{z}}. \quad (43.75b)$$

Making use of the material evolution of the horizontal velocity given by equation (43.66) renders

$$\left[\boldsymbol{x} \wedge \frac{D\boldsymbol{u}}{Dt} \right] \cdot \hat{\boldsymbol{z}} = (\boldsymbol{x} \wedge [-f \hat{\boldsymbol{z}} \wedge \boldsymbol{u} - \nabla(p/\rho + g_e z - \Omega^2 r^2/2)]) \cdot \hat{\boldsymbol{z}} \quad (43.76a)$$

$$= -f(\boldsymbol{x} \cdot \boldsymbol{u}) - (\boldsymbol{x} \wedge g \nabla \eta) \cdot \hat{\boldsymbol{z}}. \quad (43.76b)$$

The centrifugal term dropped out since

$$\boldsymbol{x} \wedge \nabla r^2 = 2\boldsymbol{x} \wedge r \hat{\boldsymbol{r}} = 2\boldsymbol{x} \wedge \boldsymbol{x} = 0. \quad (43.77)$$

The gravitational term dropped out since

$$(\boldsymbol{x} \wedge \nabla z) \cdot \hat{\boldsymbol{z}} = (\boldsymbol{x} \wedge \hat{\boldsymbol{z}}) \cdot \hat{\boldsymbol{z}} = 0, \quad (43.78)$$

as does the vertical component to the pressure gradient. We are thus left with

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g(\boldsymbol{x} \wedge \nabla \eta) \cdot \hat{\boldsymbol{z}}. \quad (43.79)$$

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.

The evolution of angular momentum is the same regardless the frame of reference. Hence, there can be no dependence on the Coriolis parameter, which indeed is the case for equation (43.79). That is, the angular momentum is a frame invariant property, so that its evolution is the same whether measured in an inertial or a non-inertial reference frame.

We can bring the expression (43.79) into a more transparent form by switching to polar coordinates

$$\boldsymbol{x} \wedge \nabla \eta = r \hat{\boldsymbol{r}} \wedge \left[\hat{\boldsymbol{r}} \frac{\partial \eta}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial \eta}{\partial \theta} \right] = \frac{\partial \eta}{\partial \theta} \hat{\boldsymbol{z}}, \quad (43.80)$$

so that

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g \frac{\partial \eta}{\partial \theta}. \quad (43.81)$$

This result is directly analogous to the angular momentum evolution for a fluid moving around a sphere as derived in Section 22.4. 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.

43.5.3 Materially invariant angular momentum

The angular momentum for a fluid column is materially invariant (i.e., a constant on a material fluid parcel) if

$$\frac{DL^z}{Dt} = 0 \iff \frac{\partial \eta}{\partial \theta} = 0. \quad (43.82)$$

For a flat bottom, equation (42.62) says that the free surface takes on a radial parabolic shape when the fluid is in solid-body rotation. In this case, $\nabla \eta$ is in the radial direction, in which case $\boldsymbol{x} \wedge \nabla \eta = 0$. Consequently, when the fluid is in solid-body rotation, the angular momentum for each fluid column remains materially constant.

43.5.4 Comments

The material evolution equation (43.79) 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 (see Section 14.1.2). 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 Plumbe \(2008\)](#) for more discussion of rotating tank experiments.

43.6 Exercises

EXERCISE 43.1: POTENTIAL TEMPERATURE SLOPES IN ATMOSPHERE AND OCEAN

Use the two-layer thermal wind relations from Section 43.1.2, also known as Margules' relation, to estimate the slope of the potential temperature surfaces in the atmosphere and ocean. This question is based on exercise 3.2 of [Vallis \(2006\)](#).

- (a) 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 Plumbe \(2008\)](#) for physical scales.
- (b) 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 Plumbe \(2008\)](#)).

EXERCISE 43.2: CIRCULAR STEADY GEOSTROPHIC FLOW

Consider a single layer of shallow water fluid in steady geostrophic balance on a f -plane so that

$$f \hat{z} \wedge \mathbf{u}_g = -g \nabla \eta, \quad (43.83)$$

where $f > 0$ (northern hemisphere). Assume the free surface has a circular Gaussian shape

$$\eta = \eta_0 e^{-r^2/(2\sigma^2)} \quad (43.84)$$

where $r^2 = x^2 + y^2$ is the squared radius and σ is the standard deviation of the Gaussian.

- (a) Determine the horizontal geostrophic velocity components corresponding to this free surface undulation.
- (b) Determine the streamlines for the flow. Hint: recall the discussion in Section 16.7.2.

EXERCISE 43.3: 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

$$\mathbf{P} = \int \rho \mathbf{u} dV = \int \rho h \mathbf{u} dA. \quad (43.85)$$

Show that

$$\frac{dP}{dt} = 0 \quad (43.86)$$

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 43.4: 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 43.5: APE FOR A SINGLE SHALLOW WATER LAYER WITH BOTTOM TOPOGRAPHY

Generalize the APE discussion in Section 43.4 to allow for a nonzero bottom topography, $z = \eta_b(x, y)$. Show that the APE is non-negative, just as for the flat bottom case. Assume the domain is simply connected.

EXERCISE 43.6: APE FOR TWO SHALLOW WATER LAYERS

Compute the APE for two shallow water layers as in Figure 42.6 with nontrivial bottom topography, $z = \eta_b(x, y)$. Show that the APE is non-negative. Assume the domain is simply connected.

EXERCISE 43.7: 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

$$\frac{E_{\text{KE}}}{E_{\text{APE}}} \sim \left[\frac{L_d^{\text{ext}}}{L} \right]^2. \quad (43.87)$$

In this equation, L is the horizontal length scale for the fluctuation of the free surface η (i.e., $\nabla \eta \sim \eta'/L$), and

$$L_d^{\text{ext}} = \frac{\sqrt{g H}}{f} \quad (43.88)$$

is the external deformation radius. The scaling (43.87) means that for scales larger than the external deformation radius, the available potential energy is larger than the kinetic energy.

EXERCISE 43.8: RATIO OF KE TO APE FOR 1.5 LAYER f -PLANE GEOSTROPHY

Consider a reduced gravity system (Section 42.4) in geostrophic balance on an f -plane. Show that the ratio of kinetic energy to available potential energy scales like

$$\frac{E_{\text{KE}}}{E_{\text{APE}}} \sim \left[\frac{L_d^{\text{int}}}{L} \right]^2. \quad (43.89)$$

In this equation, L is the horizontal length scale for the fluctuation of the internal interface η_1 (i.e., $\nabla \eta_1 \sim \eta'_1/L$), and

$$L_d^{\text{int}} = \frac{\sqrt{g'_1 h}}{f} \quad (43.90)$$

is the internal deformation radius with $g'_1 = g(\rho_2 - \rho_1)/\rho_1$ the reduced gravity and $h = A^{-1} \int (\eta_0 - \eta_1) dA$ the area averaged layer thickness (see Figure 42.4). The scaling (43.89) means that for scales

larger than the internal deformation radius, the available potential energy is larger than the kinetic energy.

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 42.5.
- The scaling (43.89) is identical to that found for the quasi-geostrophic system in Section 56.5.4. However, to solve this exercise it is not sufficient to merely reproduce the scaling discussed in Section 56.5.4. Instead, make use here of the expressions for APE and KE in the shallow water system.

44

Gravity waves and geostrophic adjustment

Waves are oscillatory fluctuations that result from a restoring force. We here consider linear waves in a single layer of shallow water fluid with a flat bottom on an f -plane. Waves in this system arise from the gravitational restoring force.¹ To develop the mathematical equations for these gravity waves, we linearize the equations of motion and then develop constraints that must be satisfied for the existence of nontrivial solutions. We identify basic properties of the gravity waves, and then consider an initial value problem to illustrate the adjustment of the fluid from an unbalanced state to a geostrophically balanced state. Material in this chapter is largely a summary of that given in Section 3.9 of [Vallis \(2017\)](#), where more details are available.

44.1	The linearized shallow water system	687
44.1.1	Linearizing the shallow water equations	688
44.1.2	Relative vorticity of linear shallow water fluctuations	688
44.1.3	Potential vorticity of linear shallow water fluctuations	689
44.2	Non-rotating shallow water gravity waves	689
44.2.1	Dispersionless waves	690
44.2.2	Vanishing relative vorticity	690
44.3	Inertia-gravity (Poincaré) waves	691
44.3.1	Non-dimensionalization	691
44.3.2	Dispersion relation	691
44.3.3	Shortwave limit	693
44.3.4	Longwave limit	693
44.4	Shallow water Kelvin waves	693
44.5	Geostrophic adjustment	695
44.5.1	Posing the initial value problem	695
44.5.2	Adjustment in the absence of rotation	695
44.5.3	Adjustment with rotation	696
44.5.4	Comments	697
44.6	Exercises	698

44.1 The linearized shallow water system

We here develop the linear shallow water equations, including the velocity, surface height, and potential vorticity equations.

¹Rossby waves also arise when considering a non-constant Coriolis parameter.

44.1.1 Linearizing the shallow water equations

Recall the shallow water equations of motion are given by the momentum and continuity equations, written here in their Eulerian form

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (44.1a)$$

$$\frac{\partial h}{\partial t} + h \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla h = 0. \quad (44.1b)$$

Since the bottom is assumed flat, the surface height equals to the thickness (see Figure 42.1)

$$\eta = h \quad \text{flat bottom.} \quad (44.2)$$

Now consider small fluctuations of the thickness and velocity relative to a state of rest

$$\eta(x, y, t) = H + \eta'(x, y, t) \quad (44.3a)$$

$$\mathbf{u}(x, y, t) = 0 + \mathbf{u}'(x, y, t). \quad (44.3b)$$

Substitution into the thickness equation (44.1b) leads to

$$\frac{\partial \eta'}{\partial t} + (H + \eta') \nabla \cdot \mathbf{u}' + \mathbf{u}' \cdot \nabla \eta' = 0. \quad (44.4)$$

The products $\eta' \nabla \cdot \mathbf{u}'$ and $\mathbf{u}' \cdot \nabla \eta'$ are second order in fluctuating quantities. Dropping these terms leads to the linearized surface height (or thickness) equation

$$\frac{\partial \eta'}{\partial t} + H \nabla \cdot \mathbf{u}' = 0 \quad \text{linearized surface height equation.} \quad (44.5)$$

Similarly, the linearized momentum equation takes the form (dropping the nonlinear advection term $(\mathbf{u}' \cdot \nabla) \mathbf{u}'$)

$$\frac{\partial \mathbf{u}'}{\partial t} + f \hat{\mathbf{z}} \wedge \mathbf{u}' = -g \nabla \eta' \quad \text{linearized momentum equation.} \quad (44.6)$$

44.1.2 Relative vorticity of linear shallow water fluctuations

We here consider how the vertical component of relative vorticity

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}) \quad (44.7)$$

evolves for the small amplitude shallow water fluctuations determined by equations (44.5) and (44.6). Taking the curl of the linearized momentum equation (44.6) leads to

$$\frac{\partial \zeta'}{\partial t} = -\nabla \cdot (f \mathbf{u}'). \quad (44.8)$$

On the f -plane and with the non-divergence condition $\nabla \cdot \mathbf{u}' + \partial_z w' = 0$, we have

$$\frac{\partial \zeta'}{\partial t} = f \frac{\partial w'}{\partial z}. \quad (44.9)$$

We can use the linearized thickness equation (44.5) to render an equivalent result

$$\frac{\partial \zeta'}{\partial t} = \frac{f}{H} \frac{\partial \eta'}{\partial t}. \quad (44.10)$$

In either case, relative vorticity of the linearized shallow water system is modified by vertical stretching in the presence of planetary rotation ($f \neq 0$). In the absence of planetary rotation, the relative vorticity remains static at each point in space. Consequently, if the relative vorticity for a non-rotating system starts with a zero value, it will remain so throughout the linearized evolution.

44.1.3 Potential vorticity of linear shallow water fluctuations

Rearranging the linearized vorticity equation (44.10) leads to the local (i.e., no advection) conservation law

$$\frac{\partial q}{\partial t} = 0, \quad (44.11)$$

where the linearized shallow water potential vorticity is given by²

$$q = \zeta' - \frac{f\eta'}{H}. \quad (44.12)$$

As mentioned in Section 44.1.2, relative vorticity is locally constant in the non-rotating case. For the rotating case, local conservation of potential vorticity provides a critical constraint on the resulting steady state after the linear fluctuations (i.e., waves) pass (see Section 44.5).

44.2 Non-rotating shallow water gravity waves

For the non-rotating case ($f = 0$), the linear velocity and thickness equations are

$$\frac{\partial \eta'}{\partial t} = -H \nabla \cdot \mathbf{u}' \quad (44.13a)$$

$$\frac{\partial \mathbf{u}'}{\partial t} = -g \nabla \eta'. \quad (44.13b)$$

Equation (44.13a) reveals that horizontal convergence drives temporal changes in the free surface height. Hence, a nonzero horizontal flow convergence is required for gravity waves to exist.

Take the time derivative of the thickness equation and the divergence of the momentum equation

$$\frac{\partial^2 \eta'}{\partial t^2} = -H \frac{\partial(\nabla \cdot \mathbf{u}')}{\partial t} \quad (44.14a)$$

$$\frac{\partial(\nabla \cdot \mathbf{u}')}{\partial t} = -g \nabla^2 \eta'. \quad (44.14b)$$

Time changes in the horizontal divergence are thus driven by curvature in the free surface. Substitution then reveals that the perturbation surface height satisfies the linear wave equation³

$$\frac{\partial^2 \eta'}{\partial t^2} - g H \nabla^2 \eta' = 0. \quad (44.15)$$

The complement substitution leads to the slightly more general wave equation for the perturbation horizontal velocity

$$\frac{\partial^2 \mathbf{u}'}{\partial t^2} - g H \nabla(\nabla \cdot \mathbf{u}') = 0. \quad (44.16)$$

²The linearized potential vorticity (44.12) can be obtained by taking the limit of $|f| \gg |\zeta|$ and $H \gg \Delta\eta$ in the shallow water potential vorticity (see Section 47.3) $Q = (f + \zeta)/h$.

³In Section 5.6, we discuss the linear wave equation and show that it is the canonical hyperbolic partial differential equation.

44.2.1 Dispersionless waves

Assume a linear wave solution of the form

$$\eta' = \gamma \cos(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (44.17)$$

where γ is a constant amplitude,

$$\mathbf{k} = \hat{\mathbf{x}} k + \hat{\mathbf{y}} l \quad (44.18)$$

is the vector wave-number, ω is the radial frequency with $2\pi/\omega$ the wave period. This form for the surface height fluctuation leads, through the linearized momentum equation (44.13b), to the velocity fluctuation⁴

$$\mathbf{u}' = \mathbf{k} \frac{g \eta'}{\omega}, \quad (44.19)$$

which can be readily shown to satisfy the linearized velocity equation (44.16). Substitution of the surface height fluctuation (44.17) into the wave equation (44.15) leads to the relation between frequency and wave-number

$$\omega = \pm c |\mathbf{k}|, \quad (44.20)$$

where

$$c = \sqrt{g H} \quad (44.21)$$

is the shallow water gravity wave speed. The *dispersion relation* (44.20) indicates that each wave-number corresponds to a single frequency. Hence, there is no mixing, or dispersion, between waves of different wavenumber or frequency. Non-rotating shallow water gravity waves are a realization of *dispersionless* waves.

44.2.2 Vanishing relative vorticity

Equation (44.14b) indicates that the linear fluctuations have a horizontal divergence driven by curvature in the surface height. In contrast, because the linearized velocity fluctuation is driven by the gradient of the surface height (see equation (44.13b)), the associated vorticity has a zero time tendency

$$\frac{\partial(\nabla \wedge \mathbf{u}')}{\partial t} = 0. \quad (44.22)$$

This result follows from the discussion of vorticity for the shallow water waves in Section 44.1.2. Hence, if the initial flow configuration has zero vorticity, the linear gravity waves retain zero vorticity as they propagate. The velocity fluctuation (44.19) indeed has zero vorticity since

$$\hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}') = \frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y} \quad (44.23a)$$

$$= \frac{g}{\omega} (l \partial_x - k \partial_y) \eta' \quad (44.23b)$$

$$= -\frac{g \gamma}{\omega} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) (l k - k l) \quad (44.23c)$$

$$= 0. \quad (44.23d)$$

⁴Formally, we have \mathbf{u}' specified only up to an arbitrary function of space. We set that function to zero without loss of generality.

44.3 Inertia-gravity (Poincaré) waves

Now let the Coriolis parameter, f , be a nonzero constant, so that the linearized thickness equation (44.5) and momentum equation (44.6) take the form

$$\frac{\partial \eta'}{\partial t} = -H \nabla \cdot \mathbf{u}' \quad (44.24a)$$

$$\frac{\partial u'}{\partial t} - f v' = -g \frac{\partial \eta'}{\partial x} \quad (44.24b)$$

$$\frac{\partial v'}{\partial t} + f u' = -g \frac{\partial \eta'}{\partial y}. \quad (44.24c)$$

The resulting linear fluctuations are known as inertia-gravity or Poincaré waves. The name “inertia-gravity” is due to the presence of both the Coriolis frequency f and gravity g . Recall our discussion of inertial oscillations in Section 15.3, which describe free particle motion in a rotating reference frame. Inertia-gravity waves arise from the combination of inertial oscillations and non-rotating gravity wave oscillations from Section 44.2. That is, both f and g play a role as restoring forces for the waves.

The free surface equation (44.24a) remains the same as for the non-rotating case in equation (44.13a). Like the non-rotating case, convergence in the horizontal flow drives surface height tendencies. That is, horizontal flow convergence is required to support gravity waves in both the rotating and non-rotating systems.

44.3.1 Non-dimensionalization

It is convenient to non-dimensionalize the linear equations (44.24a)-(44.24c) by writing

$$\mathbf{x} = L \hat{\mathbf{x}}, \quad \mathbf{u}' = U \hat{\mathbf{u}}, \quad t = \frac{L \hat{t}}{U}, \quad f = \hat{f} T^{-1}, \quad \eta' = H \hat{\eta}, \quad (44.25)$$

where L is a length scale, T is a time scale, U is a velocity scale, and H is the resting layer thickness. All variables with hats are non-dimensional and not to be confused with unit vectors. Substitution into equations (44.24a)-(44.24c) leads to the non-dimensional system

$$\frac{\partial \hat{\eta}}{\partial \hat{t}} + \frac{\partial \hat{u}}{\partial \hat{x}} + \frac{\partial \hat{v}}{\partial \hat{y}} = 0 \quad (44.26a)$$

$$\frac{\partial \hat{u}}{\partial \hat{t}} - \hat{f} \hat{v} = -\hat{c}^2 \frac{\partial \hat{\eta}}{\partial \hat{x}} \quad (44.26b)$$

$$\frac{\partial \hat{v}}{\partial \hat{t}} + \hat{f} \hat{u} = -\hat{c}^2 \frac{\partial \hat{\eta}}{\partial \hat{y}}, \quad (44.26c)$$

where

$$\hat{c} = \frac{\sqrt{g H}}{U} \quad (44.27)$$

is the non-dimensional gravity wave speed. It is also the ratio of a wave speed to a velocity scale, which is an inverse Froude number.

44.3.2 Dispersion relation

To obtain a dispersion relation we let

$$(\hat{u}, \hat{v}, \hat{\eta}) = (\tilde{u}, \tilde{v}, \tilde{\eta}) e^{i(\hat{k} \cdot \hat{\mathbf{x}} - \hat{\omega} \hat{t})}, \quad (44.28)$$

where the real part of the right hand side is assumed, and where \hat{k} is the non-dimensional wave number and $\hat{\omega}$ is the non-dimensional frequency. We are motivated to seek the linear wave solution (44.28) given the horizontal symmetry of the linearized system (44.26a)-(44.26c). Substitution into equations (44.26a)-(44.26c) leads to the dispersion relation

$$\begin{pmatrix} -i\hat{\omega} & -\hat{f} & i\hat{c}^2\hat{k} \\ \hat{f} & -i\hat{\omega} & i\hat{c}^2\hat{l} \\ i\hat{k} & i\hat{l} & -i\hat{\omega} \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\eta} \end{pmatrix} = 0. \quad (44.29)$$

This is a homogeneous system of linear equations. There is a non-trivial solution only when the determinant of the matrix vanishes, which in turn leads to the dispersion relation

$$\hat{\omega} \left[\hat{\omega}^2 - \hat{f}^2 - \hat{c}^2 (\hat{k}^2 + \hat{l}^2) \right] = 0. \quad (44.30)$$

The $\hat{\omega} = 0$ solution corresponds to time-independent geostrophic motion. Reintroducing dimensions, the second solution satisfies the dispersion relation

$$\omega^2 = f^2 + c^2 (k^2 + l^2). \quad (44.31)$$

Figure 44.1 illustrates this relation. We discuss the shortwave and longwave limits in the following.

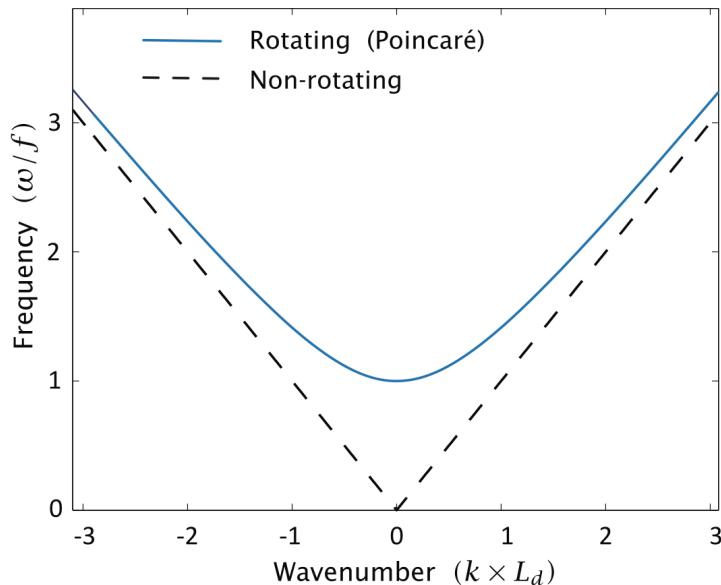


Figure 44.1: This is Figure 3.8 of [Vallis \(2017\)](#), illustrating the dispersion relation (44.31) for Poincaré waves. Frequency is scaled by the Coriolis frequency, f , and wavenumber by the inverse deformation radius \sqrt{gH}/f . For small wave number (large wave length relative to the deformation radius) the Poincaré wave frequency approaches the inertia frequency, f . We expect this result since waves large relative to the deformation radius feel the Coriolis acceleration. At the opposite extreme of high wave numbers (small wave length relative to the deformation radius), the Poincaré wave frequency approaches the non-rotating gravity wave frequency. Waves small relative to the deformation radius do not feel the Coriolis acceleration and thus converge to non-rotating gravity waves.

44.3.3 Shortwave limit

The short wave limit is in the regime where

$$k^2 + l^2 \gg \frac{f^2}{g H} = \frac{1}{L_d^2}, \quad (44.32)$$

where we introduced the shallow water deformation radius⁵

$$L_d = \frac{c}{f} = \frac{\sqrt{g H}}{f}. \quad (44.33)$$

So the shortwave limit occurs when the wavelength is much shorter than the deformation radius. For a wave moving in the \hat{x} direction, and the wavelength $\lambda = 2\pi/k$, the shortwave limit occurs when

$$\lambda \ll 2\pi L_d. \quad (44.34)$$

Note that to remain consistent with the shallow water limit with a small vertical to horizontal aspect ratio (i.e., hydrostatic layer), the wavelength must be longer than the layer thickness, H . Finally, for the shortwave limit, the dispersion relation (44.31) reduces to the non-rotating dispersion relation (44.20)

$$\omega \approx \pm c |\mathbf{k}|. \quad (44.35)$$

We see that waves much smaller than the deformation radius are too small to feel the effects of the Coriolis acceleration and thus reduce to linear non-rotating gravity waves.

44.3.4 Longwave limit

The opposite limit occurs when

$$k^2 + l^2 \ll \frac{1}{L_d^2}, \quad (44.36)$$

so the waves are much longer than the deformation radius. In this limit the dispersion relation is

$$\omega = \pm f, \quad (44.37)$$

which are known as inertial oscillations. As discussed in Section 15.3, inertial oscillations are unaffected by pressure forces otherwise arising from surface height undulations. Instead, they are determined only by the Coriolis frequency.

44.4 Shallow water Kelvin waves

The Kelvin wave is an inertia-gravity wave that arises from the presence of a boundary⁶ and rotation. Orient the f plane with a boundary at $y = 0$. The meridional velocity component must vanish at $y = 0$ to satisfy the no-normal flow condition. We are thus motivated to seek nontrivial

⁵We motivate the name “deformation radius” in Section 44.5.3.

⁶Kelvin waves also occur at the equator, which acts as a boundary due to the change in sign of f . We only consider Kelvin waves that arise from a vertical side wall.

solutions with $v' = 0$ everywhere. In this case the linearized equations of motion are given by

$$\frac{\partial \eta'}{\partial t} = -H \frac{\partial u'}{\partial x} \quad (44.38a)$$

$$\frac{\partial u'}{\partial t} = -g \frac{\partial \eta'}{\partial x} \quad (44.38b)$$

$$f u' = -g \frac{\partial \eta'}{\partial y}. \quad (44.38c)$$

Equations (44.38a) and (44.38c) lead to the one-dimensional wave equation for the zonal velocity fluctuation

$$\frac{\partial^2 u'}{\partial t^2} - c^2 \frac{\partial^2 u'}{\partial x^2} = 0, \quad (44.39)$$

where $c^2 = g H$ is the one-dimensional shallow water gravity wave speed. Solutions are propagating signals, which can be written in the form⁷

$$u'(x, y, t) = F_1(x + ct, y) + F_2(x - ct, y), \quad (44.40)$$

with corresponding surface height displacement

$$\eta'(x, y, t) = \sqrt{H/g} [-F_1(x + ct, y) + F_2(x - ct, y)]. \quad (44.41)$$

We now substitute this form of the solution into equation (44.38c) to determine the y -dependence

$$\frac{\partial F_1}{\partial y} = \frac{f F_1}{c} \quad (44.42a)$$

$$\frac{\partial F_2}{\partial y} = -\frac{f F_2}{c} \quad (44.42b)$$

with solutions

$$F_1 = F(x + ct) e^{y/L_d} \quad (44.43a)$$

$$F_2 = G(x - ct) e^{-y/L_d} \quad (44.43b)$$

where $L_d = c/f$ is the shallow water deformation radius (equation (44.33)). To ensure boundedness in the region $y > 0$ where the fluid is assumed to exist, we drop the F_1 solution, thus leaving

$$u' = e^{-y/L_d} G(x - ct) \quad (44.44a)$$

$$v' = 0 \quad (44.44b)$$

$$\eta' = (H/g)^{1/2} e^{-y/L_d} G(x - ct). \quad (44.44c)$$

These wave signals are propagating in the positive \hat{x} direction, in which case the boundary $y = 0$ is on the right. This orientation holds for any boundary orientation in the northern hemisphere, whereby Kelvin waves propagate with the solid boundary on the right when looking in the direction of wave movement. For the southern hemisphere Kelvin waves propagate with the boundary to the left of the wave motion. Hence, Kelvin waves propagate in a cyclonic direction.

⁷See Section 5.6 for more on general solutions to the linear wave equation.

44.5 Geostrophic adjustment

The geostrophic balance presented in Sections 29.3 and 43.1 is very well maintained by the observed large-scale atmosphere and ocean. Hence, geostrophy (and the associated thermal wind) is a powerful diagnostic. In this section, we examine how a flow state that is initially not in geostrophic balance evolves towards geostrophy. We thus study the dynamical processes associated with the *geostrophic adjustment* problem.

A single shallow water layer on a flat f -plane is sufficient to introduce the main physical ideas. Furthermore, we focus on linear perturbations so that the governing equations are those derived in Section 44.1. Consequently, the adjustment consists of linear inertia-gravity waves that maintain a locally static potential vorticity (Section 44.1.3). For brevity in notation, we here drop all primes on the linear fluctuating terms.

44.5.1 Posing the initial value problem

We solve for the $t > 0$ evolution of surface height and velocity by making use of the linearized equations

$$\frac{\partial \mathbf{u}}{\partial t} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (44.45)$$

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0 \quad (44.46)$$

$$\zeta - \frac{f\eta}{H} = q(x, y) \quad (44.47)$$

where we dropped the primes for brevity, and where $q(x, y)$ is the static potential vorticity determined by the initial conditions (Section 44.1.3). To illustrate the geostrophic adjustment in an analytically tractable manner, consider the following step initial conditions for the surface height

$$\eta(x, t = 0) = \begin{cases} +\eta_0 & x < 0 \\ -\eta_0 & x > 0, \end{cases} \quad (44.48)$$

which can be written

$$\eta(x, t = 0) = -\eta_0 \operatorname{sgn}(x). \quad (44.49)$$

The velocity is assumed to be zero initially

$$\mathbf{u}(x, y, t = 0) = 0. \quad (44.50)$$

Correspondingly, the initial relative vorticity vanishes so that the linearized potential vorticity is initialized as

$$q(x, y) = \frac{f\eta_0}{H} \operatorname{sgn}(x). \quad (44.51)$$

Since $\partial q / \partial t = 0$, this value of the potential vorticity is maintained at each point throughout the adjustment process. The velocity and surface height adjustment is thus constrained to keep potential vorticity static. This rather basic point is key to determining their evolution.

44.5.2 Adjustment in the absence of rotation

In the absence of rotation ($f = 0$), relative vorticity is constant at each grid point. With a zero initial velocity, relative vorticity remains zero throughout the adjustment. The adjustment is thus

quite simple. Namely, it consists of linear gravity waves propagating away from the initial step, converting the potential energy of the step into kinetic energy of waves that propagate to infinity. As the linear gravity waves are non-dispersive, they carry the initial pulse out to infinity without distortion in the form

$$\eta(x, t) = -\frac{\eta_0}{2} [\operatorname{sgn}(x + ct) + \operatorname{sgn}(x - ct)], \quad (44.52)$$

where $c = \sqrt{gH}$ is the speed for non-rotating gravity waves. The meridional velocity remains zero, whereas the zonal velocity

$$\frac{\partial u}{\partial t} = -g \frac{\partial \eta}{\partial x} \quad (44.53)$$

is given by

$$u(x, t) = \frac{g\eta_0}{2c} [\operatorname{sgn}(x + ct) - \operatorname{sgn}(x - ct)]. \quad (44.54)$$

After the transient waves have passed, the steady solution is a flat surface height with zero velocity. This steady solution is familiar from the case of a rock dropped into a still pond. After dropping the rock into the pond, the surface gravity waves radiate outward from the rock and are eventually damped upon reaching the shore. In equilibrium, the pond returns to a state of rest with a flat surface height.

44.5.3 Adjustment with rotation

With rotation, the transient solution consists of inertia-gravity waves that transmit information about the initial surface height perturbation out to infinity. After the transient waves have passed, the steady solution is either the trivial solution with flat surface height (as for the non-rotating case), or a nontrivial solution that is in geostrophic balance

$$f\hat{z} \wedge \mathbf{u} = -g \nabla \eta \quad (44.55)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (44.56)$$

$$q = \frac{f\eta_0}{H} \operatorname{sgn}(x). \quad (44.57)$$

Conservation of potential vorticity constrains the solution so that the steady state surface height is indeed sloped according to a geostrophically balanced state. That is, an equilibrium state of no-motion is not allowed by potential vorticity conservation.

Computing the equilibrium state

As the flow is geostrophic on an f -plane, we make use of the geostrophic streamfunction

$$\psi = \frac{g\eta}{f}. \quad (44.58)$$

The equilibrium state is written in terms of the streamfunction according to

$$u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x} \quad (\nabla^2 - L_d^{-2}) \psi = q(x, y), \quad (44.59)$$

where we introduced the *deformation radius*

$$L_d = \frac{\sqrt{gH}}{f}. \quad (44.60)$$

We motivate the name “deformation radius” in the following.

The initial condition (44.49) has no y -dependence. Furthermore, there is nothing in the adjustment process that will break this meridional symmetry. Hence, the equilibrium state is a function only of x , in which case the streamfunction satisfies the ordinary differential equation

$$\frac{d^2\psi}{dx^2} - L_d^{-2} \psi = \frac{f\eta_0}{H} \operatorname{sgn}(x). \quad (44.61)$$

We solve this equation separately for $x > 0$ and $x < 0$ and then match the function and its first derivative at $x = 0$. Furthermore, we constrain the streamfunction to vanish at $\pm\infty$. The $x > 0$ streamfunction satisfies

$$\frac{d^2\psi}{dx^2} - L_d^{-2} \psi = \frac{f\eta_0}{H}. \quad (44.62)$$

The particular solution is

$$\psi_p = -L_d^2 \frac{f\eta_0}{H} = -\frac{g\eta_0}{f} \quad (44.63)$$

and the homogeneous solution is

$$\psi_h = \frac{g\eta_0}{f} e^{-x/L_d} \quad (44.64)$$

so that

$$\psi = -\frac{g\eta_0}{f} \left[1 - e^{-x/L_d} \right]. \quad (44.65)$$

The $x < 0$ solution is found similarly, so that

$$\psi = \frac{g\eta_0}{f} \begin{cases} -\left(1 - e^{-x/L_d}\right) & x > 0 \\ \left(1 - e^{x/L_d}\right) & x < 0, \end{cases} \quad (44.66)$$

which means that the equilibrium surface height is

$$\eta = \eta_0 \begin{cases} -\left(1 - e^{-x/L_d}\right) & x > 0 \\ \left(1 - e^{x/L_d}\right) & x < 0, \end{cases} \quad (44.67)$$

Note that the streamfunction vanishes at $x = 0$ and has a first derivative of $-\eta_0 \sqrt{gH}/H$. Since the streamfunction only has a zonal dependence, the equilibrium velocity is purely meridional

$$u = 0 \quad v = -\frac{g\eta_0}{f L_d} e^{-|x|/L_d}. \quad (44.68)$$

The equilibrium velocity thus consists of a jet that is perpendicular to the surface height front.

44.5.4 Comments

As illustrated in Figure 44.2, the equilibrium profiles for the surface height and velocity both have an exponential decay, with decay length scale given by the deformation radius. It is this length scale over which the solution is affected or “deformed” by rotation, thus motivating the name “deformation radius”.

The key feature of the rotating case is that some of the potential energy contained within the initial perturbed surface height remains in the equilibrium geostrophic flow. The conservation of potential vorticity constrains the flow so that all of the initial potential energy cannot be converted to kinetic energy. Rather, the adjustment occurs only within a deformation radius distance from the initial perturbation.

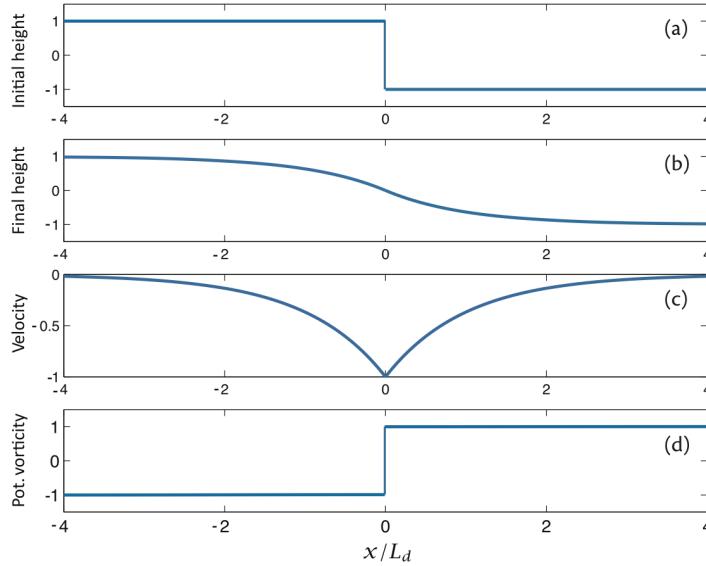


Figure 44.2: This is Figure 3.10 of [Vallis \(2017\)](#), illustrating solutions to the linear geostrophic adjustment of a rotating shallow water layer to a surface height perturbation. The top panel shows the initial surface height (44.49), and the second panel the equilibrium surface height (44.67). The third panel shows the equilibrium meridional velocity (44.68) comprised of a jet centered at $x = 0$. The final panel shows the static potential vorticity (44.57).

44.6 Exercises

EXERCISE 44.1: DEFORMATION RADIUS

The deformation radius appears in many contexts within rotating fluid dynamics. Here, we compute this length scale for selective geophysical flow regimes.

- (a) Compute the shallow water deformation radius for an ocean continental shelf of depth 500 m.
- (b) Compute the shallow water deformation radius for the deep ocean with depth 5000 m.
- (c) The deformation radius defined in this chapter is sometimes called the *external deformation radius* as it makes use of the full depth of the fluid and the gravitational acceleration. In contrast, the deformation radius defined in terms of internal layer thickness and reduced gravity g' leads to the internal deformation radius. The internal deformation radius, $L_d^{\text{int}} = \sqrt{g' h} / f$ is the appropriate rotational length scale for density layers in the interior of the ocean or isentropic layers in the interior of the atmosphere. Compute the deformation radius for a density layer of thickness $h = 200$ m and reduced gravity of $g' = g/1000$.

EXERCISE 44.2: GEOSTROPHIC ADJUSTMENT

This exercise revisits many of the points raised in this chapter with the aim to ensure a thorough understanding of geostrophic adjustment. We consider again the linearized shallow water equations on an f -plane with flat bottom topography

$$\frac{\partial \mathbf{u}}{\partial t} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (44.69a)$$

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0, \quad (44.69b)$$

where

$$h(x, y, t) = H + \eta(x, y, t) \quad (44.70)$$

is the layer thickness, η is the free surface undulation relative to the resting layer, H is the resting layer thickness, f is the constant Coriolis parameter, and g is the constant gravitational acceleration. We dropped primes to reduce notational clutter.

- (a) Derive the linearized shallow water potential vorticity conservation equation directly using the linearized shallow water equations.
- (b) For the special case of a geostrophic flow, show how to express the horizontal velocity $\mathbf{u} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}}$ and the thickness h in terms of a scalar streamfunction, ψ .
- (c) Consider an arbitrary initial condition for u, v, h and *assume* the system evolves into a state of geostrophic balance at infinite time. Using the results from parts (a) and (b), write down the streamfunction for the geostrophically adjusted final state.
- (d) Specialize the result in part (c) to the case of an initial condition of zero motion and initial height $\eta(x, y, t = 0) = A$ inside a circle centered at $(x, y) = (0, 0)$ and $\eta = 0$ outside the circle. Write an equation for the geostrophically adjusted final state in plane polar coordinates coordinates.

45

Isopycnal models

For stably stratified fluids, buoyancy is a particularly useful generalized vertical coordinate. Most notably, physical processes away from turbulent boundary layers are oriented according to these surfaces, and horizontal buoyancy gradients give rise to thermal wind shears in a geostrophically balanced flow. For this reason buoyancy (or entropy) plays a key role in theoretical and numerical models of ocean and atmospheric circulation. In this chapter we derive the hydrostatic Boussinesq equations using buoyancy as the vertical coordinate. The resulting equation set forms the basis for *isopycnal* models of the ocean or isentropic models of the atmosphere. We pay particular attention to the needs of integrating the equations over discrete layers, as required to develop discrete numerical models. In the adiabatic limit, the isopycnal equations reduce to the stacked shallow water equations.

READER'S GUIDE TO THIS CHAPTER

We assume a working knowledge of the shallow water system as described in Chapters 42 and 43; the mathematics of generalized vertical coordinates (GVCs) detailed in Chapter 11; and the corresponding kinematics in Chapter 21 and dynamics in Chapter 32. We also make particular use of the layer integrated notions introduced for mass continuity and the tracer equations in Sections 21.9 and 21.10. Cartesian horizontal coordinates are sufficient for this chapter.

45.1	Isopycnal equations	702
45.1.1	Montgomery potential and the pressure force	702
45.1.2	Material time derivative	704
45.1.3	Layer integrated thickness equation	704
45.1.4	Ocean equations	705
45.1.5	Vector-invariant horizontal momentum equation	705
45.1.6	Connection to the shallow water equations	706
45.2	Transfer across layer boundaries	706
45.2.1	Diapycnal transfer	706
45.2.2	Momentum transfer	707
45.2.3	Allowing for layers to vanish and reappear	707

45.1 Isopycnal equations

In this section we derive the hydrostatic Boussinesq equations using buoyancy as the vertical coordinate, which we refer to as *isopycnal* vertical coordinates in accordance with common usage in the ocean literature. Rather than specializing the general expressions provided in Section 32.1, we find it pedagogical to start from the equations written using the geopotential vertical coordinate (see Section 28.1.6)

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z \psi + \mathbf{F} \quad \text{horizontal momentum} \quad (45.1a)$$

$$\frac{\partial \psi}{\partial z} = b \quad \text{hydrostatic} \quad (45.1b)$$

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad \text{continuity} \quad (45.1c)$$

$$\frac{Db}{Dt} = \dot{b} \quad \text{thermodynamics} \quad (45.1d)$$

$$\frac{DC}{Dt} = \dot{C} \quad \text{tracers,} \quad (45.1e)$$

where $\mathbf{v} = (\mathbf{u}, w)$ is the velocity field, \mathbf{u} is its horizontal component, ψ is the dynamic pressure, b is the buoyancy, and C is an arbitrary tracer concentration. A discrete realization of the isopycnal layer-integrated form of these equations is depicted in Figure 45.1, with the remainder of this section detailing the formulation using isopycnal coordinates for the vertical.

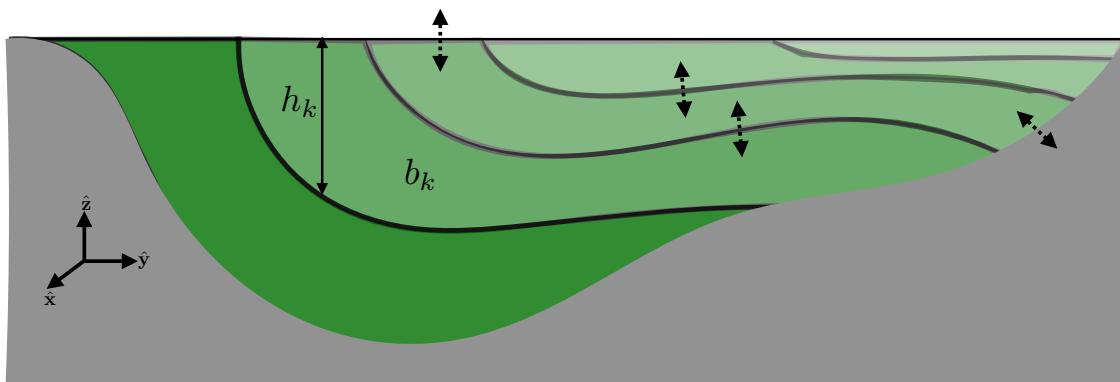


Figure 45.1: Schematic of an isopycnal model, formulated as stacked shallow water layers (green layers) that generally allow for the transfer of matter and energy across the layer interfaces as well as across the ocean surface and ocean bottom (as depicted by the double-headed dashed arrows). The dark gray region is land.

45.1.1 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

Following the discussion in Section 32.1.2, the horizontal pressure gradient transforms as

$$\nabla_z \psi = \nabla_b \psi - \frac{\partial \psi}{\partial z} \nabla_b z \quad (45.2a)$$

$$= \nabla_b \psi - b \nabla_b z \quad (45.2b)$$

$$= \nabla_b (\psi - b z) \quad (45.2c)$$

$$= \nabla_b M, \quad (45.2d)$$

where

$$M = \psi - b z \quad (45.3)$$

defines the Montgomery potential. As the contribution to the horizontal pressure force, the Montgomery potential is the geostrophic streamfunction in buoyancy coordinates (see Section 45.1.4).

The horizontal pressure gradient force for numerical models

It is notable that the horizontal pressure gradient force is determined by the horizontal isopycnal gradient of a single term, the Montgomery potential. Furthermore, as shown below, the Montgomery potential satisfies the buoyancy coordinate form of the hydrostatic balance. Hence, numerical isopycnal models do not suffer from problems with computing the horizontal pressure gradient that can occur with other generalized vertical coordinate models, such as terrain-following models (see Figure 32.1).

Equation (45.2e) is the key step in the formulation, whereby we made use of $\nabla_b b = 0$. This step is available only under certain cases that utilize an idealized equation of state for seawater. In more realistic cases, the buoyancy determining the hydrostatic pressure (i.e., the *mass buoyancy*) is defined locally whereas the generalized vertical coordinate must be defined globally. As a result, there are two terms contributing to the pressure gradient in a manner similar to terrain-following models. [Sun et al. \(1999\)](#) and [Hallberg \(2005\)](#) discuss this issue in the context of numerical ocean modeling. For present purposes we ignore this detail and continue to assume a simplified equation of state so that $\nabla_b b = 0$.

Hydrostatic balance

Supporting our use of the Montgomery potential as a pressure field, the hydrostatic balance takes the form

$$\frac{\partial M}{\partial b} = \frac{\partial \psi}{\partial b} - b \frac{\partial z}{\partial b} - z \quad (45.4a)$$

$$= \frac{\partial \psi}{\partial z} \frac{\partial z}{\partial b} - b \frac{\partial z}{\partial b} - z \quad (45.4b)$$

$$= -z, \quad (45.4c)$$

where we made use of the hydrostatic balance $\partial \psi / \partial z = b$ (equation (45.1b)). This result means that M is the buoyancy coordinate version of pressure.

45.1.2 Material time derivative

As seen in Section 21.4, there are two equivalent forms for the material time derivative

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad \text{geopotential form} \quad (45.5a)$$

$$= \left[\frac{\partial}{\partial t} \right]_b + \mathbf{u} \cdot \nabla_b + w^{(b)} \frac{\partial}{\partial z} \quad \text{isopycnal form,} \quad (45.5b)$$

where

$$w^{(b)} = \frac{\partial z}{\partial b} \frac{Db}{Dt} \quad (45.6)$$

is the diapycnal velocity component that measures the rate of flow crossing buoyancy surfaces (Section 21.3). Besides differences in the spatial operators, it is important to note that the time derivative operators in equations (45.5a) and (45.5b) 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

$$(u, v) = \frac{D(x, y)}{Dt}. \quad (45.7)$$

45.1.3 Layer integrated thickness equation

The continuity equation, $\nabla_z \cdot \mathbf{u} + \partial_z w = 0$, is an expression of volume conservation. We already derived the GVC version of this equation in Section 21.9.4, and thus quote the layer thickness result here

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) + \Delta_b (w^{(b)}) = 0, \quad (45.8)$$

where the isopycnal layer thickness (dimensions of length) is given by

$$h = \int_{b-\delta b/2}^{b+\delta b/2} \frac{\partial z}{\partial b} db \approx \frac{\partial z}{\partial b} \delta b = N^{-2} \delta b, \quad (45.9)$$

where the approximation holds when assuming the stratification is constant across the extent of a layer. We here introduced the squared buoyancy frequency

$$N^2 = \frac{\partial b}{\partial z} \quad (45.10)$$

and note that its inverse is the specific thickness

$$\frac{\partial z}{\partial b} = \text{specific thickness.} \quad (45.11)$$

Furthermore,

$$\Delta_b = \delta b \frac{\partial}{\partial b} \quad (45.12)$$

is the non-dimensional differential b -operator.

When $w^{(b)} \neq 0$, the three terms in the full thickness equation (45.8) are coupled, with discussion given in Section 21.5 as part of our discussion of the vertical velocity versus the dia-surface velocity. For idealized studies we often consider the adiabatic case in which $\dot{b} = 0$. In this case layer thickness

is altered only through horizontal rearrangement of volume within a layer according to the adiabatic thickness equation

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) = 0. \quad (45.13)$$

As further discussed in Section 45.1.6, the adiabatic limit brings the discrete isopycnal model into accord with the stacked shallow water system.

45.1.4 Ocean equations

For the tracer equation we make use of the general development in Section 21.10, here specialized to the case of a Boussinesq model using isopycnal coordinates. Bringing the pieces together leads to the isopycnal version of the hydrostatic Boussinesq equations

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \mathbf{u} + (w^{(b)} \partial_z) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b M + \mathbf{F}^h \quad (45.14a)$$

$$\frac{\partial M}{\partial b} = -z \quad (45.14b)$$

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) + \Delta_b w^{(b)} = 0 \quad (45.14c)$$

$$\left[\frac{\partial (h C)}{\partial t} \right]_b + \nabla_b \cdot (h C \mathbf{u} + h \mathbf{J}^h) + \Delta_b (C w^{(b)} + J^{(b)}) = 0, \quad (45.14d)$$

where the tracer equation includes possible subgrid scale flux contributions as well as advective transport. Notice how the advective transport is two-dimensional in the adiabatic case with $\dot{b} = 0$, 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 (45.14a) gives

$$f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b M \implies f u = - \left[\frac{\partial M}{\partial y} \right]_b \quad \text{and} \quad f v = \left[\frac{\partial M}{\partial x} \right]_b \quad \text{geostrophy.} \quad (45.15)$$

Hence, the Montgomery potential is the streamfunction for geostrophic flow as represented using buoyancy coordinates.

45.1.5 Vector-invariant horizontal momentum equation

It is common for isopycnal models to make use of the vector-invariant form of the momentum equation derived in Section 32.1.4. Introducing the isopycnal version of the relative vorticity (see Section 52.2.1)

$$\hat{\mathbf{z}} \tilde{\zeta} \equiv \nabla_b \wedge \mathbf{u} \quad (45.16)$$

renders the vector-invariant horizontal momentum equation

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (w^{(b)} \partial_z) \mathbf{u} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b \mathcal{B} + \mathbf{F}^h, \quad (45.17)$$

where

$$\mathcal{B} = M + \mathbf{u} \cdot \mathbf{u} / 2 = \psi - b z + \mathbf{u} \cdot \mathbf{u} / 2 \quad (45.18)$$

is the Bernoulli potential for a hydrostatic Boussinesq fluid (see Section 24.2.3). Note that we can further introduce the isopycnal potential vorticity (Section 52.2.2)

$$\tilde{Q} = \frac{f + \tilde{\zeta}}{h} \quad (45.19)$$

to bring the momentum equation to the form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (w^{(b)} \partial_z) \mathbf{u} + \tilde{Q} h \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b \mathcal{B} + \mathbf{F}^h. \quad (45.20)$$

This form is commonly used as the starting point for certain theoretical analyses, particularly when considering the adiabatic limit in which $w^{(b)} = 0$.

45.1.6 Connection to the shallow water equations

We can make use of the material time derivative operator (45.5b) to write the material form of the adiabatic and inviscid equations (45.14a)-(45.14c)

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b M \quad (45.21a)$$

$$\frac{\partial M}{\partial b} = -z \quad (45.21b)$$

$$\frac{Dh}{Dt} + h \nabla_b \cdot \mathbf{u} = 0. \quad (45.21c)$$

These isopycnal momentum and thickness equations are isomorphic to those for a single layer of adiabatic shallow water fluid (see Section 42.1). This isomorphism allows us to derive the vorticity and potential vorticity equations in Section 52.2, making use of the shallow water manipulations from Section 47.3. When doing so, note that for the isopycnal case, lateral gradient operations are computed along surfaces of constant buoyancy, thus making use of the isopycnal gradient operator, ∇_b , rather than the horizontal gradient operator, ∇_z , used in geopotential coordinates.

45.2 Transfer across layer boundaries

We here briefly summarize the treatment of boundaries between isopycnal layers as well as the intersection of layers with the surface and bottom boundaries.

45.2.1 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 isopycnal layers in the presence of irreversible processes such as mixing.

45.2.2 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 Section 43.2 and more generally in Section 26.8. Furthermore, there are in general 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.

45.2.3 Allowing for layers to vanish and reappear

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 45.1). The seasonal cycle of warming and cooling is a canonical example of layer outcropping at the surface ocean. A formulational expedient to handle vanishing layers is to assume that all layers exist everywhere horizontally across the ocean domain, but to allow for zero layer thickness where a layer has zero volume. 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.

Part VIII

Vorticity and potential vorticity

Vorticity is a kinematic fluid property that locally measures the spin of a fluid element. Its kinematic and dynamic properties are fundamental to understanding and predicting fluid flow. 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 planet, also renders a nonzero vorticity for geophysical fluids even when at rest in the rotating earth reference frame. Potential vorticity (PV) is a strategically chosen component of vorticity that eliminates the baroclinicity source, with PV of great use for understanding and predicting geophysical fluid flows.

We start this part of the book by introducing vorticity and circulation in Chapter 46, making use of Stokes' Theorem to trivially show that the area integral of vorticity over a finite region yields the circulation around the region's boundary. We follow up in Chapter 47 with an introduction to the mechanics of vorticity and potential vorticity within a shallow water fluid. Chapter 48 then dives into the fundamentals of vorticity and circulation. It is here that we encounter Kelvin's Circulation Theorem, which identifies the materially conserved nature of circulation around an arbitrary loop in a perfect barotropic flow. Chapter 50 presents the foundations of potential vorticity and Chapter 52 dives into the details by considering a suite of ocean examples.

This part of the book makes extensive use of the Cartesian tensor algebra and vector calculus detailed in Chapters 3 and 4. Further specialized insights can be garnered through use of generalized vertical coordinates from Chapters 21 and 32.

46

Vorticity and circulation

Vorticity is defined pointwise whereas circulation is defined over a region. Helmholtz was an early proponent of vorticity whereas Kelvin was a proponent of circulation. These two properties of fluid motion are connected through Stokes' theorem. We here introduce the kinematic notions of vorticity and circulation, leaving dynamical discussions for later chapters.

READER'S GUIDE FOR THIS CHAPTER

This chapter makes use of vector calculus identities for Cartesian coordinates as detailed in Chapter 4. The ideas introduced here are fundamental to the remaining chapters in this part of the book.

46.1	Vorticity	712
46.1.1	Vorticity generates the rotation of line elements	713
46.1.2	Rotating reference frame	713
46.2	Irrational flows	713
46.3	Circulation of the velocity field	714
46.4	The free vortex	715
46.4.1	Motion of a fluid particle	715
46.4.2	Velocity circulation	716
46.5	Translation and solid-body rotation	716
46.5.1	Absolute vorticity equals planetary vorticity plus relative vorticity . .	717
46.5.2	Solid-body rotation on a plane	717
46.5.3	Circulation for solid-body rotation	718
46.5.4	Comments	718
46.6	Exercises	718

46.1 Vorticity

Vorticity is the curl of the velocity field

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v}. \quad (46.1)$$

Vorticity measures the rotation or spin of fluid motion and it does so without reference to an origin. Angular momentum also provides a measure of the rotational motion but it does so with respect to a subjectively chosen origin. Though related (see Chapter 49), vorticity and angular momentum are generally not the same in the presence of strain.

Figure 46.1 provides an example zonal flow with a meridional strain (shear). The vertical component to the vorticity is negative for this flow. Hence, an imaginary test “paddle wheel” placed anywhere within this flow will spin clockwise about its axis. We next provide some rigor to this “test paddle wheel” characterization of vorticity.

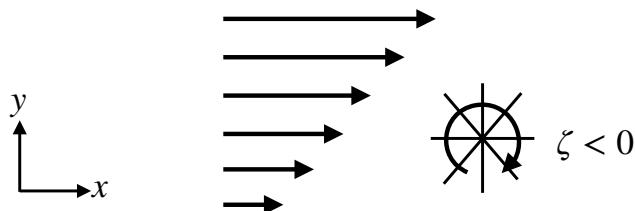


Figure 46.1: An example zonal flow with a meridional shear, $\mathbf{v} = u(y)\hat{x}$ and a corresponding vertical vorticity component that is negative: $\zeta = \hat{z} \cdot (\nabla \wedge \mathbf{v}) = \partial v / \partial x - \partial u / \partial y = -\partial u / \partial y < 0$. The clockwise arrow denotes the negative vorticity flow, with the sign of the vorticity determined through the right hand rule. 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.

46.1.1 Vorticity generates the rotation of line elements

In Section 20.2 we considered the kinematic evolution of a line element, $\delta\mathbf{x}$, separating two fluid particles, with that evolution provided by equation (20.20)

$$\frac{D(\delta x_m)}{Dt} = \delta x_n \frac{\partial v_m}{\partial x_n} \implies \frac{D(\delta \mathbf{x})}{Dt} = (\delta \mathbf{x} \cdot \nabla) \mathbf{v}. \quad (46.2)$$

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 rate of strain tensor, $2 \mathbb{S}_{mn} = \partial_n v_m + \partial_m v_n$, whose action generates changes in the distance between the fluid particles (Section 20.2.6). The anti-symmetric portion is the rotation tensor, $2 \mathbb{A}_{mn} = \partial_n v_m - \partial_m v_n$, which is related to vorticity via equation (20.32)

$$\mathbb{A}_{mn} = -\epsilon_{mnp} \omega_p / 2 \iff \mathbb{A} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}, \quad (46.3)$$

so that

$$2 \mathbb{A}_{mn} \delta x_n = -\epsilon_{mnp} \omega_p \delta x_n \implies 2 \mathbb{A} \cdot \delta \mathbf{x} = \boldsymbol{\omega} \wedge \delta \mathbf{x}. \quad (46.4)$$

From our discussion of rotation in Section 13.5, this equation means that vorticity generates a rigid rotation of a material line element (Section 20.2.7). This result accords with Figure 46.1, whereby vorticity leads to the spin of a test paddle wheel.

46.1.2 Rotating reference frame

View the flow field from a reference frame that rotates with a constant angular velocity $\boldsymbol{\Gamma}$. Following equation (13.36b), we know that the velocity observed in the non-rotating reference frame, \mathbf{v} , is related to the rotating reference frame velocity, \mathbf{v}_{rot} , via

$$\mathbf{v} = \mathbf{v}_{\text{rot}} + \boldsymbol{\Gamma} \wedge \mathbf{x}. \quad (46.5)$$

The vorticity in the two reference frames is thus related by

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} = \nabla \wedge \mathbf{v}_{\text{rot}} + \nabla \wedge (\boldsymbol{\Gamma} \wedge \mathbf{x}) = \boldsymbol{\omega}_{\text{rot}} + 2 \boldsymbol{\Gamma}. \quad (46.6)$$

If there is a point in the fluid whereby the rotating reference frame's angular velocity equals to one-half the local vorticity, $\boldsymbol{\Gamma} = \boldsymbol{\omega}/2$, then the rotating reference frame's vorticity vanishes at that point

$$\boldsymbol{\Gamma} = \boldsymbol{\omega}/2 \implies \boldsymbol{\omega}_{\text{rot}} = 0. \quad (46.7)$$

Hence, we may interpret $\boldsymbol{\omega}/2$ as the angular velocity of the fluid circulating around any point in the fluid. Furthermore, if the vorticity, $\boldsymbol{\omega}$, is spatially constant, then we can move to a rotating reference frame in which the flow is everywhere irrotational.

46.2 Irrotational flows

Most geophysical flows have nonzero vorticity. However, there are some geophysically relevant flows with vanishing vorticity, such as linear gravity waves in the absence of planetary rotation (Section 44.2.2). Irrotational fluid flow is characterized by a zero vorticity

$$\boldsymbol{\omega} = 0 = \text{irrotational flow.} \quad (46.8)$$

Since the curl of a gradient vanishes, irrotational flow has a velocity field equal to the gradient of a velocity potential

$$\nabla \wedge \mathbf{v} = 0 \Rightarrow \mathbf{v} = \nabla \Psi. \quad (46.9)$$

Irrotational flow is therefore sometimes called *potential flow*. Figure 46.2 illustrates a two-dimensional flow field generated by taking the gradient of a scalar potential so that the flow has zero vorticity. The vertical component of the vorticity vanishes at each point since $\partial v / \partial x = \partial u / \partial y$.

If the flow is incompressible, as in a Boussinesq fluid (Section 28.1), then the velocity potential is a harmonic function

$$\nabla \cdot \mathbf{v} = 0 \Rightarrow \nabla^2 \Psi = 0. \quad (46.10)$$

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 flow for incompressible fluids.

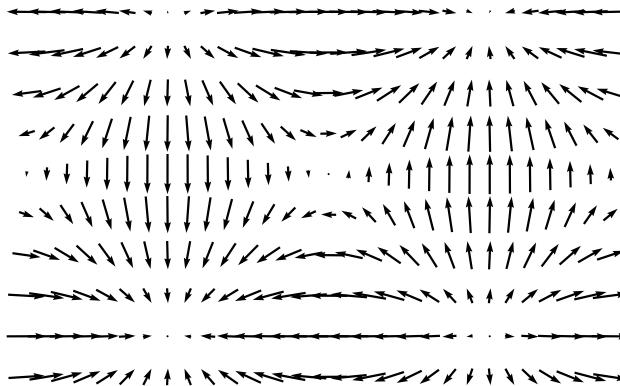


Figure 46.2: An example horizontal flow based on a potential, $\Psi = \sin(x/5) \sin(y/5)$. The flow has zero vorticity, $\omega \cdot \hat{z} = \zeta = \partial v / \partial x - \partial u / \partial y = 0$, since the flow is based on a scalar potential: $\omega = \nabla \wedge \mathbf{v} = \nabla \wedge \nabla \Psi = 0$. This example illustrates how irrotational flow may have nontrivial structure even though a test paddle wheel will not spin since there is zero vorticity.

46.3 Circulation of the velocity field

The velocity circulation, or more briefly the *circulation*, is defined as the oriented closed loop integral of velocity as projected onto the path of the loop

$$\mathcal{C} \equiv \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r}, \quad (46.11)$$

with Figure 46.3 offering a schematic. The line element, $d\mathbf{r}$, is oriented in the counter-clockwise direction around the circuit ∂S . More precisely, let $\mathbf{r}(\varphi)$ 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 (see Section 4.4). The difference between two very close positions along the circuit defines the increment

$$d\mathbf{r} = \mathbf{r}(\varphi + \delta\varphi) - \mathbf{r}(\varphi). \quad (46.12)$$

By construction, $d\mathbf{r}$ is tangent to the circuit so that $\mathbf{v} \cdot d\mathbf{r}$ picks out the component of the velocity that is tangent to the path.

Stokes' Theorem (Section 4.6) renders the very important identity

$$\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int_S (\nabla \wedge \mathbf{v}) \cdot \hat{\mathbf{n}} dS = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (46.13)$$

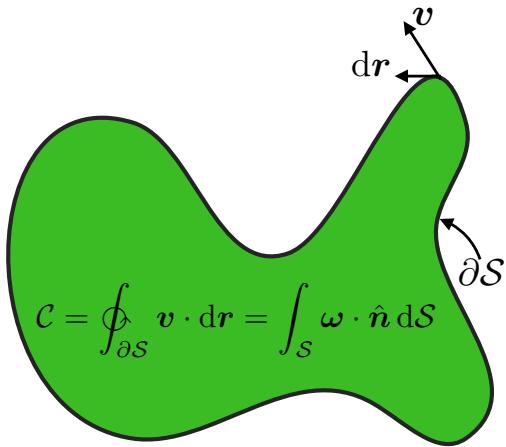


Figure 46.3: The velocity circulation around the boundary of a region, $\partial\mathcal{S}$, is determined by the line integral of the velocity projected into the direction of the line integral, $\mathbf{v} \cdot d\mathbf{r}$. Stokes' Theorem shows that the velocity circulation computed as a line integral is identical to the normal projection of the vorticity integrated over the area of the region, $C = \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} = \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS$.

where $\hat{\mathbf{n}}$ is the outward 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 “flux of vorticity” penetrating the area. Stokes’ theorem provides the means to connect the vorticity theories promoted by Helmholtz to the circulation theories of Kelvin.

46.4 The free vortex

Consider a two-dimensional rotating fluid in the $x - y$ plane with angular velocity given by

$$\boldsymbol{\Omega} = \frac{\mathbf{x} \wedge \mathbf{v}}{r^2} = \frac{K \hat{\mathbf{z}}}{r^2}. \quad (46.14)$$

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. Some refer to this flow as the “vr-vortex” whereas we refer to it as a *free vortex*. As shown here, the fluid flow associated with the free vortex has zero vorticity and velocity circulation for all points except the origin, yet the same points also have a constant angular momentum relative to the origin, with the angular momentum associated with nonzero strain within the fluid (see Exercise 46.3).

46.4.1 Motion of a fluid particle

A fluid particle moves in a circular orbit when in the free vortex flow field. Hence, the particle velocity is perpendicular to its position vector, $\mathbf{x} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}}$, with respect to the origin

$$\mathbf{v} \cdot \mathbf{x} = 0. \quad (46.15)$$

The velocity for this pure rotational flow is given by (see Section 13.5)

$$\mathbf{v} = \boldsymbol{\Omega} \wedge \mathbf{x} = \frac{K(-y \hat{\mathbf{x}} + x \hat{\mathbf{y}})}{r^2} = \frac{K \hat{\theta}}{r}, \quad (46.16)$$

where $\hat{\theta}$ is the polar angle unit vector oriented in the counter-clockwise direction (see Section 10.3). Each component of the velocity falls off as $1/r$ when moving away from the origin. Away from the origin, the vorticity vector vanishes

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} = 0, \quad (46.17)$$

whereas it is singular at the origin.

Although vorticity is zero everywhere except at the origin, the angular momentum is nonzero, as expected since the fluid is rotating around the vortex center. The angular momentum for this system arises just from the strain in the fluid (see Exercise 46.3), 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

$$\mathbf{r} \wedge \mathbf{v} = r \hat{\mathbf{r}} \wedge (K/r) \hat{\theta} = K \hat{\mathbf{z}}. \quad (46.18)$$

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.

We illustrate the free vortex velocity field (46.16) in Figure 46.4 along with two test paddle wheels. The paddle wheels remain stationary when placed away from the origin, in the region where the vorticity vanishes. As the paddle wheel centers move counter-clockwise with the flow, the marked paddle wheel blades remain oriented at the same angle. That is, the paddle wheels orbit around the vortex center but they do not spin. The free vortex thus illustrates a fluid flow with non-zero angular momentum yet with zero vorticity.

46.4.2 Velocity circulation

The velocity circulation vanishes for any circuit bounded away from the origin, and it does so trivially since vorticity vanishes away from the origin. However, the circulation is nonzero for any circuit enclosing the origin

$$\mathcal{C} = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int_0^{2\pi} \mathbf{v} \cdot \hat{\theta} r d\theta = 2\pi K. \quad (46.19)$$

To reach this result, we set the line element to

$$d\mathbf{r} = \hat{\theta} r d\theta \quad (46.20)$$

and inserted the velocity (46.16) represented in cylindrical polar coordinates, $\mathbf{v} \cdot \hat{\theta} = K/r$. Hence, the singular point vortex at $r = 0$ induces a nonzero circulation for all circuits that enclose the vortex.

46.5 Translation and solid-body rotation

Rigid or solid-body fluid motion occurs when all fluid particles are rigidly locked into their relative positions, as if in a rigid solid body. There are two kinds of rigid body motion: translation and rotation. The velocity field for this motion is given by

$$\mathbf{v} = \mathbf{U} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (46.21)$$

where \mathbf{x} is the position vector relative to the origin, \mathbf{U} is a translation velocity, and $\boldsymbol{\Omega}$ is an angular velocity. For rigid body motion, both \mathbf{U} and $\boldsymbol{\Omega}$ are spatially uniform, but can in general be time

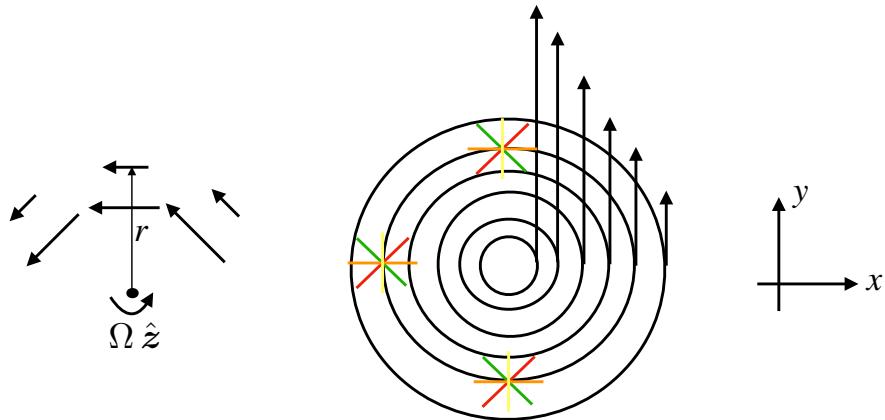


Figure 46.4: Irrotational counter-clockwise planar flow in the presence of a free vortex with velocity $\mathbf{v} = (K/r)\hat{\theta}$. The tangential velocity decays as $1/r$ from the origin and the vorticity, $\nabla \wedge \mathbf{v}$, vanishes for all points except the origin. Also, test paddle wheels do not spin when removed from the center since there is zero vorticity. Notice also that 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 .

dependent. The rate of strain tensor vanishes for uniform translation or solid-body motion (see Exercise 46.1)

$$S_{mn} = \frac{1}{2}(\partial_m v_n + \partial_n v_m) = 0. \quad (46.22)$$

However, the vorticity is nonzero (see Exercise 46.1)

$$\boldsymbol{\omega} = \nabla \wedge (\boldsymbol{\Omega} \wedge \mathbf{x}) = 2\boldsymbol{\Omega}. \quad (46.23)$$

We encountered this vorticity in Section 46.1.2 when connecting vorticity and angular velocity. Note that the factor of two in this equation is geometric; it needs no physical explanation.

46.5.1 Absolute vorticity equals planetary vorticity plus relative vorticity

For planetary fluid dynamics, rotation of the planet 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

$$\boldsymbol{\omega}_{\text{absolute}} = \boldsymbol{\omega}_{\text{planet}} + \boldsymbol{\omega}. \quad (46.24)$$

In this equation,

$$\boldsymbol{\omega}_{\text{planet}} = 2\boldsymbol{\Omega}_{\text{planet}} \quad (46.25)$$

is the planetary vorticity associated with solid-body motion of a fluid particle stationary with respect to the planet, and

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (46.26)$$

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.

46.5.2 Solid-body rotation on a plane

Consider the circular solid-body rotation on a plane shown in Figure 46.5, in which the velocity is purely tangential and linearly proportional to the distance from the center

$$\mathbf{v} = \boldsymbol{\Omega} \wedge \mathbf{x} = |\boldsymbol{\Omega}|(-y\hat{\mathbf{x}} + x\hat{\mathbf{y}}) = |\boldsymbol{\Omega}|r\hat{\theta}. \quad (46.27)$$

Assuming the center of mass to be at the circle center, the angular momentum for the flow is the same as that for a solid-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 46.5.

46.5.3 Circulation for solid-body rotation

For solid-body rotation, the velocity circulation around a circular path of radius R is given by

$$\mathcal{C} = \oint \mathbf{v} \cdot d\mathbf{r} = \oint (\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{r} = R^2 |\boldsymbol{\Omega}| \oint d\theta = 2\pi R^2 |\boldsymbol{\Omega}| = 2A |\boldsymbol{\Omega}|, \quad (46.28)$$

where $A = \pi R^2$ is the area of the circle. Hence, the velocity circulation per area for solid-body rotating fluid flow is twice the angular rotation rate, which is the magnitude of the vorticity

$$\mathcal{C}/A = |\boldsymbol{\omega}| = 2 |\boldsymbol{\Omega}|. \quad (46.29)$$

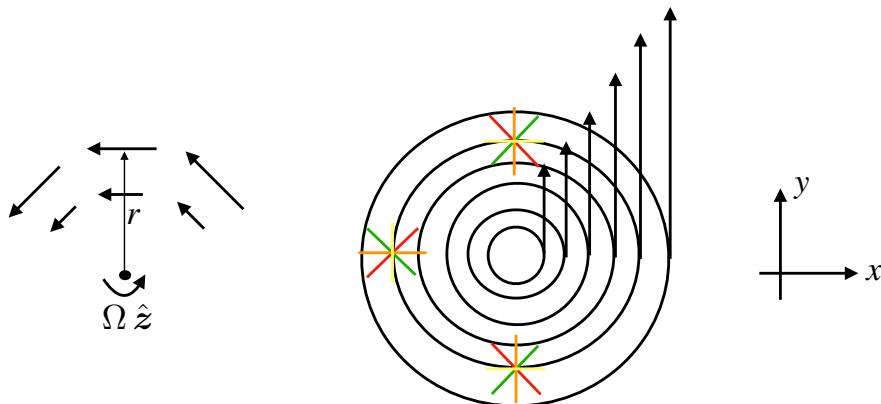


Figure 46.5: Rigid body fluid motion, whereby the fluid velocity is purely tangential and linearly proportional to the radial distance from the vortex center, $\mathbf{v} = |\boldsymbol{\Omega}| r \hat{\theta}$; fluid particles maintain a fixed relative position; and vorticity is constant and points perpendicular to the page, $\boldsymbol{\omega} = 2\boldsymbol{\Omega} = 2|\boldsymbol{\Omega}|\hat{z}$. Test paddle wheels rigidly move around the center, and they exhibit a spin about their axis thus manifesting the nonzero vorticity.

46.5.4 Comments

As seen in Section 46.4, 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 solid-body rotating fluid flow has a nonzero vorticity, nonzero angular momentum, yet a zero strain. Chapter 49 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 nicely illustrated in [this short video](#)

46.6 Exercises

EXERCISE 46.1: VORTICITY FOR SOLID-BODY ROTATION

Show that a fluid in solid-body rotation with angular velocity

$$\mathbf{v}_{\text{solid-body}} = \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (46.30)$$

has a vorticity given by

$$\nabla \wedge \mathbf{v}_{\text{solid-body}} = 2\boldsymbol{\Omega}. \quad (46.31)$$

EXERCISE 46.2: VELOCITY POTENTIAL FOR THE FREE VORTEX

What is the velocity potential (46.9) for the free vortex whose velocity field is given by (46.16)? Hint: The problem is two-dimensional and rotationally symmetric, so it is convenient to make use of polar coordinates $x = r \cos \theta$ and $y = r \sin \theta$ as in Appendix 10.3.

EXERCISE 46.3: STRAIN TENSOR FOR THE FREE VORTEX

Determine all components to the strain tensor

$$S_{pq} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{1}{2} \left[\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right] \\ \frac{1}{2} \left[\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right] & \frac{\partial v}{\partial y} \end{bmatrix} \quad (46.32)$$

for the free vortex as specified by the velocity field (46.16). Present the answer in the form of a 2×2 matrix.

EXERCISE 46.4: VANISHING VISCOUS FRICTION FOR SOLID-BODY MOTION

As seen in Section 37.2.5, viscous effects from molecular viscosity in an incompressible fluid appear in the momentum equation as a Laplacian weighted by a constant molecular viscosity

$$\text{viscous force per mass} = \mu \nabla^2 \mathbf{v}, \quad (46.33)$$

where μ is the molecular kinematic viscosity, assumed here to be a constant. Show that the viscous operator vanishes for a fluid in solid-body rotation. That is, solid-body motion engenders no frictional dissipation. This result reflects the lack of frictional interaction in a fluid absent shears and strains.

47

Shallow water vorticity and potential vorticity

We introduced the basic kinematics of vorticity in Chapter 46. Here, we further that discussion by considering vorticity within the shallow water system and then extending it to potential vorticity. The shallow water fluid offers a useful conceptual model to introduce the dynamical equations for vorticity and potential vorticity while encountering a relatively modest level of mathematical details. In this chapter we derive the evolution equation for vorticity by taking the curl of the momentum 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 very important mechanical constraint on the fluid flow.

READER'S GUIDE FOR THIS CHAPTER

We here make use of vector calculus identities for Cartesian coordinates as detailed in Chapter 4 as well as the shallow water mechanics from Chapters 42 and 43, building from the vorticity kinematics introduced in Chapter 46. The concepts and methods developed in this chapter are fundamental to the remaining chapters in this part of the book.

47.1	Vorticity dynamics	722
47.2	PV for a rotating deformable cylinder	723
47.2.1	Mass conservation	724
47.2.2	Angular momentum conservation	724
47.2.3	Potential vorticity conservation	725
47.2.4	Connecting angular momentum and vorticity	725
47.2.5	Comments and further reading	725
47.3	Potential vorticity for a shallow water layer	726
47.3.1	Mass conservation + the vorticity equation	726
47.3.2	Motivating the name	727
47.3.3	Mass conservation + Kelvin's circulation theorem	727
47.3.4	Material conservation of an arbitrary function of PV	729
47.4	Example implications of PV material invariance	729
47.4.1	Absolute vorticity invariance when h is fixed	729
47.4.2	Rigid lid and planetary geostrophic PV	730
47.4.3	Spin up of converging flow	731
47.4.4	Further study	732
47.5	Exercises	732

47.1 Vorticity dynamics

When working with a shallow water fluid it is useful to introduce vorticity for the full velocity field as well as that just for the horizontal flow

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (47.1a)$$

$$\boldsymbol{\omega}^* = \nabla \wedge \mathbf{u} = \left[\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right] \hat{z} = \zeta \hat{z}, \quad (47.1b)$$

where

$$\zeta = \hat{z} \cdot \boldsymbol{\omega}^* = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad (47.2)$$

is the vertical component to the relative vorticity.

The vector identity (see Section 4.3.4)

$$(\mathbf{u} \cdot \nabla) \mathbf{u} = (1/2) \nabla(\mathbf{u} \cdot \mathbf{u}) - \mathbf{u} \wedge (\nabla \wedge \mathbf{u}) \quad (47.3)$$

brings the inviscid shallow water momentum equation (42.5)

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{z} \wedge \mathbf{u} = -g \nabla \eta \quad (47.4)$$

into its “vector-invariant” form

$$\frac{\partial \mathbf{u}}{\partial t} + \boldsymbol{\omega}_a^* \wedge \mathbf{u} = -\nabla (g \eta + \mathbf{u}^2/2). \quad (47.5)$$

We here introduced the absolute vorticity

$$\boldsymbol{\omega}_a^* = (\zeta + f) \hat{z}, \quad (47.6)$$

which is the sum of the relative vorticity of the horizontal flow, $\omega^* = \zeta \hat{z}$, plus the solid-body vorticity, $f \hat{z}$, due to motion of the rotating reference frame (recall Section 46.5.1). We next make use of the vector identity from Section 4.3.4

$$\nabla \wedge (\omega_a^* \wedge \mathbf{u}) = \omega_a^* (\nabla \cdot \mathbf{u}) - \mathbf{u} (\omega_a^* \cdot \nabla) + (\mathbf{u} \cdot \nabla) \omega_a^* - (\omega_a^* \cdot \nabla) \mathbf{u} \quad (47.7a)$$

$$= \omega_a^* (\nabla \cdot \mathbf{u}) + (\mathbf{u} \cdot \nabla) \omega_a^*. \quad (47.7b)$$

This result required setting

$$\nabla \cdot \omega_a^* = \nabla \cdot \omega^* + \nabla \cdot (f \hat{z}) = 0, \quad (47.8)$$

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

$$(\omega_a^* \cdot \nabla) \mathbf{u} = |\omega_a^*| \partial_z \mathbf{u} = 0, \quad (47.9)$$

which follows since the horizontal velocity in a shallow water fluid is depth independent (see Section 42.1). Applying the operator $\hat{z} \cdot (\nabla \wedge)$ onto the vector-invariant momentum equation (47.5) yields a prognostic equation for the absolute vorticity

$$\frac{\partial \zeta_a}{\partial t} + \nabla \cdot (\mathbf{u} \zeta_a) = 0, \quad (47.10)$$

where

$$\zeta_a = \hat{z} \cdot (\omega^* + \hat{z} f) = \hat{z} \cdot (\omega + \hat{z} f) = \zeta + f \quad (47.11)$$

is the vertical component of the absolute vorticity.

The Eulerian flux-form evolution equation (47.10) means that the vertical component to the absolute vorticity at a point in the inviscid shallow water fluid changes according to the horizontal convergence of vorticity advected to that point

$$\frac{\partial \zeta_a}{\partial t} = -\nabla \cdot (\mathbf{u} \zeta_a). \quad (47.12)$$

Alternatively, we can write the vorticity equation (47.10) in the material form

$$\frac{D \zeta_a}{Dt} = -\zeta_a \nabla \cdot \mathbf{u}, \quad (47.13)$$

where the material time derivative is determined by the horizontal flow

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla. \quad (47.14)$$

The material evolution equation (47.13) means that the absolute vorticity of a shallow water fluid particle, moving with the horizontal flow, changes according to the horizontal convergence of the fluid flow as multiplied by the absolute vorticity.

47.2 PV for a rotating deformable cylinder

To conceptually introduce potential vorticity, consider a deformable cylinder of constant mass M , constant density ρ , variable radius R , and variable height H , and assume the cylinder exhibits solid-body rotation about its central axis. This example may appear contrived for a fluid. However, it offers a useful conceptual picture for a rotating material region of a constant density fluid, in which time derivatives in the following are interpreted as material derivatives. It also lends useful intuition for the motion of spinning fluids constrained by mass and angular momentum conservation.

47.2.1 Mass conservation

Mass conservation is a kinematic property of the cylinder. With a constant density, mass conservation means that the volume of the cylinder is fixed. Hence, mass conservation constrains the relative changes to the radius and height of the cylinder. A materially constant cylinder mass thus renders

$$M = \pi R^2 H \rho \quad (47.15)$$

implies

$$\frac{2}{R} \frac{DR}{Dt} = -\frac{1}{H} \frac{DH}{Dt}. \quad (47.16)$$

That is, mass conservation means that as twice the relative radius increases the relative height decreases.

47.2.2 Angular momentum conservation

A second constraint on cylinder rotation arises from angular momentum conservation. Since the cylinder is rotating as a solid-body, angular momentum is straightforward to compute. For simplicity, 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

$$\boldsymbol{\Omega} = \Omega \hat{z}. \quad (47.17)$$

With this axis orientation, the solid-body 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 by (e.g., [Marion and Thornton, 1988](#))

$$I_{mn} = \delta_{mn} \frac{MR^2}{2}. \quad (47.18)$$

The moment of inertia is a measure of the rotational inertia of the cylinder, and is seen to be directly related to the cylinder mass (assumed fixed here) and radius (can change). Note that the moment of inertia is not a function of the cylinder height. The angular momentum for the cylinder is thus given by

$$\mathbf{L} = \frac{MR^2}{2} \Omega \hat{z}. \quad (47.19)$$

The familiar “ice skater” example occurs when the cylinder radius changes (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 Ω increases (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

$$\frac{2}{R} \frac{DR}{Dt} = -\frac{1}{\Omega} \frac{D\Omega}{Dt}. \quad (47.20)$$

This tradeoff between spin rate and radius holds in general and we encounter it again in Figure 47.7 for a layer of shallow water fluid. Namely, reducing the moment of inertia for a constant mass body by bringing its mass distribution towards the central axis leads (converging mass), through angular momentum conservation, to an increase in rotation speed. The opposite occurs when mass diverges from a region, thus reducing the rotation speed.

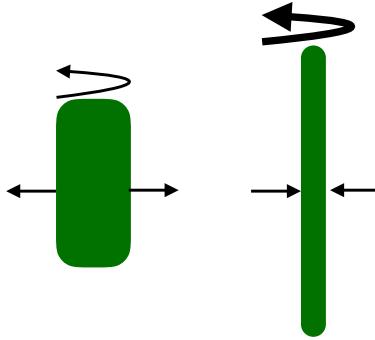


Figure 47.1: Illustrating the conservation of angular momentum for a constant mass rotating cylinder undergoing solid-body rotation around its central axis. The moment of inertia (relative to the central axis) for the left configuration is larger since more of its mass is distributed away from the central axis than in the right configuration. Assuming the two configurations have identical angular momentum means that the right configuration rotates more rapidly since its moment of inertia is smaller. This example exemplifies the familiar ice-skater experience, whereby the skater's spin increases when bringing arms (mass) inward towards the central axis of the body (depicted by the inward arrows on the right panel), and the skater slows when extending the arms outward (depicted by the outward arrows on the left panel).

47.2.3 Potential vorticity conservation

Combining angular momentum conservation (47.20) with mass conservation (47.16) leads to the conservation law

$$\frac{D}{Dt} \frac{\Omega}{H} = 0. \quad (47.21)$$

Equation (47.21) is a statement of potential vorticity conservation for the material fluid column, with potential vorticity given by

$$Q \equiv \frac{\Omega}{H}. \quad (47.22)$$

For example, if the column thickens then the rotational velocity increases to maintain $Q = \Omega/H$ constant.

47.2.4 Connecting angular momentum and vorticity

When allowing the fluid to exhibit motion that is more general than a solid-body, then the angular rotation rate appearing in the potential vorticity (47.22) is generalized to the absolute vorticity. We encounter this generalization in Section 47.3. Furthermore, as shown in Section 46.5, the vorticity for solid body motion equals to twice the rotation rate, 2Ω . Hence, the numerator for the potential vorticity of the solid-body rotating cylinder equals to one-half the vorticity. In Chapter 49, we connect angular momentum and vorticity (and strain) for arbitrary fluid motion.

47.2.5 Comments and further reading

The discussion in this section is motivated by a similar presentation given by [Salmon \(1998\)](#). The solid-body rotating cylinder succinctly identifies the two mechanical properties contributing to the potential vorticity conservation law (47.21): a kinematic property (mass conservation) and a dynamical property (angular momentum conservation). Notably, for the solid-body rotating cylinder, the implications of PV conservation are well gleaned from the separate mass and angular momentum conservation principles. Hence, PV conservation lends little novel insight for the cylinder.

However, PV conservation is of fundamental use for studies of general fluid motions, particularly stratified and rotating fluids.

47.3 Potential vorticity for a shallow water layer

We now consider the potential vorticity (PV) for a single layer of shallow water fluid. The result is directly analogous to that derived for the rotating cylinder in Section 47.2. However, the derivation here makes use of fluid mechanical notions rather than rigid body dynamics. 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 (we discuss Kelvin's Theorem in Chapter 48). Figure 47.2 summarizes key elements leading to PV conservation for a shallow water fluid layer. Namely, as shown in this section, shallow water PV conservation arises from combining the kinematic constraint of mass conservation (material conservation of $h A$) with either the vorticity equation or Kelvin's circulation theorem for a small aspect ratio fluid.

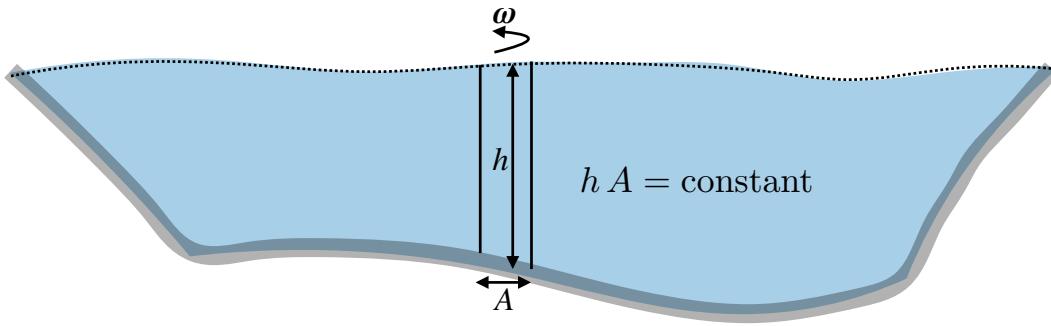


Figure 47.2: Illustrating the conservation of PV for a layer of shallow water fluid. PV conservation results from merging mass conservation (material conservation of the column volume, $h A$), to either the vorticity equation or Kelvin's circulation theorem for a small aspect ratio fluid (material conservation of ζA).

47.3.1 Mass conservation + the vorticity equation

To derive the potential vorticity equation, we here make use of the vorticity equation (47.13) and combine it with mass conservation.

Shallow water vorticity and vortex stretching

Mass conservation in the form of the material thickness equation (42.17) leads to the following expression for the divergence of the horizontal velocity

$$\nabla \cdot \mathbf{u} = -\frac{1}{h} \frac{Dh}{Dt}. \quad (47.23)$$

Making use of this result in the vorticity equation (47.13) allows us to eliminate the horizontal divergence, thus yielding

$$\frac{D\zeta_a}{Dt} = \frac{\zeta_a}{h} \frac{Dh}{Dt}. \quad (47.24)$$

This equation says that material changes in shallow water absolute vorticity arise only from material changes in the layer thickness. Hence, shallow water absolute vorticity changes if the fluid column

stretches or compresses. We see in Section 48.5.2 that vorticity in more general fluids is affected by vortex stretching as well as vortex tilting and torques from baroclinicity. The shallow water fluid is thus a very special case whereby absolute vorticity changes materially only through vortex stretching. This behavior is a result of the depth independence of the horizontal velocity and the associated column-like nature of shallow water fluid motion.

Material invariance of PV

Equation (47.24) can be written as an expression of the material invariance of the shallow water potential vorticity

$$\frac{DQ}{Dt} = 0, \quad (47.25)$$

where

$$Q = \frac{\zeta_a}{h} = \frac{\zeta + f}{h} \quad (47.26)$$

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 (47.25) says that this ratio remains constant for the shallow water layer.

47.3.2 Motivating the name

We can understand the “potential” in the name by noting that potential vorticity measures the potential 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 (see Figure 47.3). 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” as described in Section 23.7.

47.3.3 Mass conservation + Kelvin’s circulation theorem

Although we have yet to discuss Kelvin’s Theorem (Section 48.2), we here invoke it to illustrate another way to derive shallow water PV conservation. When applied to an infinitesimal circuit in an inviscid and constant density fluid, Kelvin’s theorem says that

$$\frac{D(\omega_a \cdot \hat{n} \delta S)}{Dt} = 0, \quad (47.27)$$

where ω_a is the absolute vorticity

$$\omega_a = \omega + f \hat{z}, \quad (47.28)$$

$\hat{n} \delta S$ is the infinitesimal surface area enclosed by the closed circuit, with \hat{n} the outward normal to the surface. For the shallow water system, we decompose absolute vorticity into

$$\omega_a = \hat{z} (\zeta + f) + \omega_h, \quad (47.29)$$

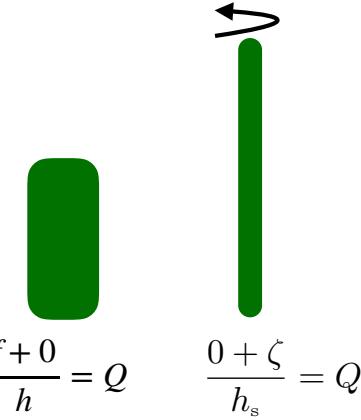


Figure 47.3: Left panel: an arbitrary shallow water column with zero relative vorticity and potential vorticity $Q = f/h$. Right panel: the same fluid column moved to the equator (where $f = 0$) and stretched to have the standard thickness, $h_s > h$. The relative vorticity of the column at the equator is given by $\zeta = f(h_s/h)$, with $f > 0$ assumed for this figure (northern hemisphere). 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 as part of the name.

where ζ is the vertical component to the relative vorticity (equation (47.2)), and

$$\omega_h = \hat{x} \frac{\partial w}{\partial y} - \hat{y} \frac{\partial w}{\partial x} \quad (47.30)$$

is the horizontal component to the relative vorticity, making use of the depth independence of the horizontal velocity components for the shallow water fluid ($\partial u/\partial z = \partial v/\partial z = 0$). Inserting the absolute vorticity (47.29) into Kelvin’s theorem (47.29) leads to

$$\frac{D}{Dt} [(\zeta + f) \delta A + \omega_h \cdot \hat{n} \delta S] = 0, \quad (47.31)$$

where the horizontal area element, δA , is the projection of the surface area element onto the vertical direction

$$\delta A = \hat{z} \cdot \hat{n} dS. \quad (47.32)$$

Shallow water fluid dynamics arises from considering a constant density fluid layer under the small aspect ratio limit ($H/L \ll 1$, with H the vertical length scale and L the horizontal length scale). Under this limit, the second term in equation (47.31) is much smaller than the first. Ignoring this term then leads to

$$\frac{D}{Dt} \left[\left(\frac{\zeta + f}{h} \right) h \delta A \right] = 0, \quad (47.33)$$

where h is the layer thickness and $h \delta A$ is the volume of an infinitesimal fluid column. The volume of a column of shallow water fluid is materially constant

$$\frac{D(h \delta A)}{Dt} = 0, \quad (47.34)$$

so that equation (47.33) yields the material invariance of shallow water PV

$$\frac{D}{Dt} \left[\frac{\zeta + f}{h} \right] = \frac{DQ}{Dt} = 0, \quad (47.35)$$

where $Q = (\zeta + f)/h$ is the same shallow water potential vorticity derived above in Section 47.3.1. The essence of this derivation is depicted in Figure 47.4.

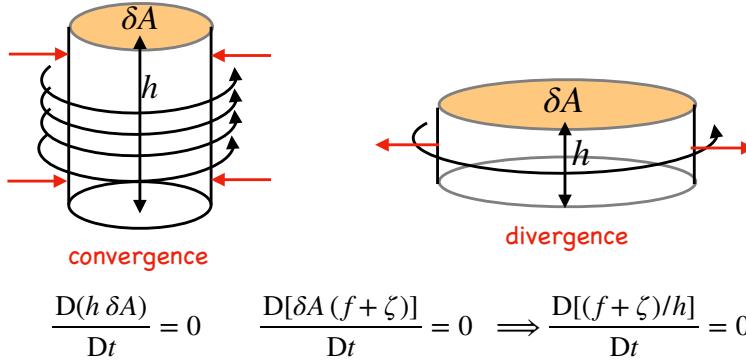


Figure 47.4: Material invariance for PV in a shallow water layer results from combining material invariance of the volume of the fluid column to the invariance of the area weighted absolute vorticity. As the cross-sectional area of the column decreases, as in a converging flow, the column spins faster by increasing its relative vorticity, just like the spinning cylinder in Figure 47.1. Conversely, as the cross-sectional area increases, as in a diverging flow, then the column spins slower by reducing its relative vorticity. This process of column stretching and squashing is a fundamental means for changing the vorticity while maintaining constant potential vorticity.

47.3.4 Material conservation of an arbitrary function of PV

The material conservation law for PV, equation (47.25), means that any function, $F(Q)$ is also materially constant. We see this property through the chain rule

$$\frac{DF}{Dt} = F'(Q) \left[\frac{DQ}{Dt} \right] = 0. \quad (47.36)$$

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.

47.4 Example implications of PV material invariance

The material invariance of shallow water PV provides a nontrivial constraint on the fluid motion, saying that f, h, ζ cannot change independently of the other. Rather, the combination $Q = (f + \zeta)/h$ must remain materially unchanged. There are a huge 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 PV-thinking.

47.4.1 Absolute vorticity invariance when h is fixed

To illustrate the power of material invariance of PV to constrain the flow, consider a column of shallow water fluid that maintains a fixed thickness. In this case the material invariance of PV reduces to the material invariance of absolute vorticity

$$\frac{D(\zeta + f)}{Dt} = 0 \quad \text{if } h = \text{constant}. \quad (47.37)$$

This situation holds for the two-dimensional barotropic flows described in Chapter 53. To make use of this constraint for a column of shallow water fluid, we make the observation that a column of fluid curving to the left (facing downstream) is associated with a positive relative vorticity for the column, $\zeta > 0$. In contrast, the oppositely curved column has negative relative vorticity, $\zeta < 0$.

If the motion is initially zonal but then turns, the meridional motion on the sphere is associated with a change in planetary vorticity (f changes). To maintain constant absolute vorticity, $\zeta + f$, when a column moves meridionally requires the relative vorticity induced by the curved motion to counteract the change in planetary vorticity. We now show that the constraint of fixed absolute vorticity means that eastward flow (westerly winds) cannot turn meridionally while maintaining fixed absolute vorticity, whereas westward flow (easterly winds) can.

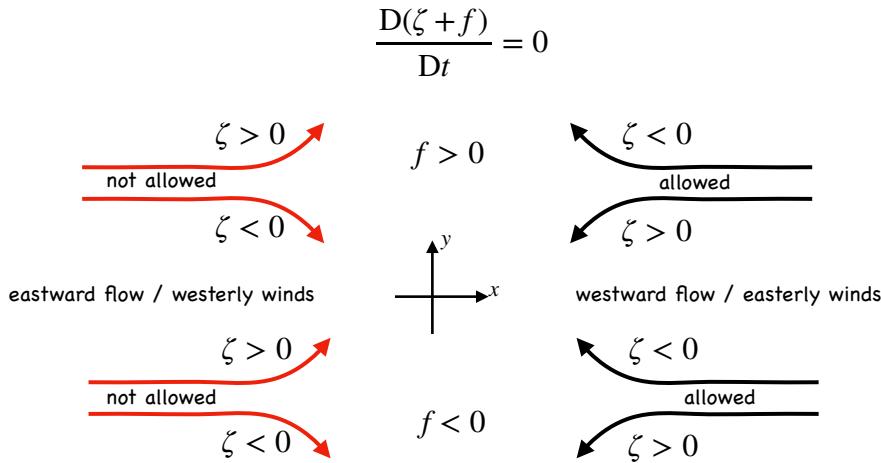


Figure 47.5: Illustrating the constraints on shallow water flow imposed by material invariance of absolute vorticity: $\zeta + f = \text{constant}$, which holds if the layer thickness is constant. In each case the entering flow has zero relative vorticity, $\zeta = 0$, so that absolute vorticity must remain constant at the initial Coriolis parameter, $\zeta_a = f$. The red eastward flow (westerly winds) is disallowed by material invariance of absolute vorticity. In these cases 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 show motion for both the northern and southern hemispheres. Taken after Figure 4.8 of Holton (1992).

Consider westward flow in the northern hemisphere ($f > 0$). If the flow turns to the north (to the right facing downstream) then this flow has an associated $\zeta < 0$ and an increase in the planetary vorticity (f increases). Likewise, a westward flowing fluid column that turns equatorward (to the left) has a positive 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 hold in the southern hemisphere. These scenarios are depicted in Figure 47.5.

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 $\zeta > 0$ 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 induces $\zeta < 0$ 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.

47.4.2 Rigid lid and planetary geostrophic PV

As introduced in Section 29.4, planetary geostrophy (PG) is used to study the large-scale laminar ocean circulation where relative vorticity is ignored. Furthermore, as shown in Section 54.3, the

inviscid and adiabatic PG system materially preserves the PG potential vorticity, $Q = f/h$, so that

$$\frac{D(f/h)}{Dt} = 0. \quad (47.38)$$

Consequently, fluid particles respecting the inviscid PG equations follow contours of constant f/h . If we furthermore assume the ocean surface is a flat rigid lid (not a bad assumption for large-scale flow), then shallow water columns follow trajectories of constant f/H , where $z = -H(x, y)$ is the vertical position of the bottom topography. Figure 47.6 illustrates sample f/H contours for a topographic bump in the northern hemisphere. Fluid columns are steered towards the south as they encounter the bump. If the bump is too high, then there is no way for the flow to go over that part of the bump while maintaining f/H constant. Flow maintaining constant f/H can only go around large bumps, in which case we say the bump is topographically blocked.

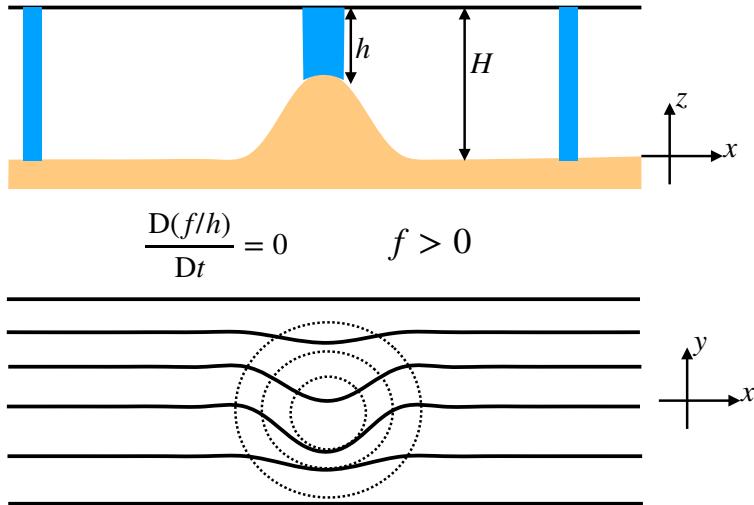


Figure 47.6: When a shallow water fluid has a rigid lid and is governed by the inviscid and adiabatic planetary geostrophic equations, fluid columns maintain fixed PG potential vorticity, $D(f/h)/Dt = 0$. Fluid column trajectories thus follow contours of constant f/H , where $z = -H(x, y)$ is the bottom topography. We illustrate these contours for a topographic bump in the northern hemisphere. As the column moves onto the bump it must move south to keep f/H fixed at its initial value far from the bump. (SMG: PRODUCE A BETTER LOWER FIGURE USING PYTHON.)

47.4.3 Spin up of converging flow

Consider the flow shown in Figure 47.7, whereby mass in the shallow water layer converges into a region. Just as described in the PV derivation Figure 47.4, 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 47.2, 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 PV. Finally, note that the opposite occurs in a region of diverging fluid, whereby the PV invariance implies that the relative vorticity has a negative tendency ($\partial\zeta/\partial t < 0$) (see also Figure 47.4).

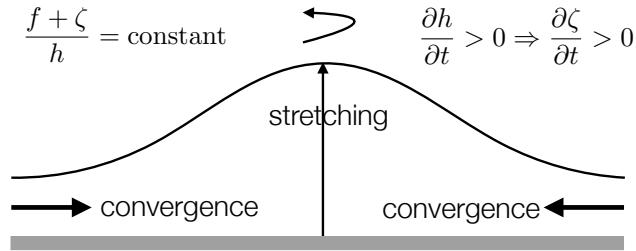


Figure 47.7: Illustrating the implications of material PV invariance for a shallow water fluid on an f -plane. If mass converges into a region, thus stretching the fluid column, then PV invariance implies the relative vorticity increases, $\partial\zeta/\partial t > 0$. This result is directly analogous to the rotating cylinder example considered in Figure 47.1. Namely, converging a region of constant mass reduces its moment of inertia so that angular momentum conservation leads to an increase in positive spin.

47.4.4 Further study

Section 4.3 of Holton (1992) discusses the case of flow over topography where the full shallow water PV 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) deflects 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 classic video from Prof. Dave Fultz of the University of Chicago](#). 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.

47.5 Exercises

EXERCISE 47.1: 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

$$\mathcal{V} = \int dA \int dz = \int h dA = \int (H + \Delta\eta - \eta_b) dA = H \mathcal{A}, \quad (47.39)$$

where \mathcal{A} 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 η_b is the undulation of the bottom topography (see Figure 42.1). Additionally, recall that $z = 0$ is set according to

$$\int \eta_b dA = 0, \quad (47.40)$$

and volume conservation ensures that

$$\int \Delta\eta dA = 0. \quad (47.41)$$

- (a) Determine the volume average of the vorticity $\hat{z} \cdot \omega_{\text{solid}}$ arising from the solid-body rotation

$$\langle \hat{z} \cdot \omega_{\text{solid}} \rangle = V^{-1} \int \hat{z} \cdot \omega_{\text{solid}} dV. \quad (47.42)$$

- (b) Determine the area average of the relative vorticity,

$$\bar{\zeta} = \mathcal{A}^{-1} \int \hat{z} \cdot \omega dA, \quad (47.43)$$

in terms of the circulation around the boundary of the domain.

- (c) Determine the volume average of the relative vorticity

$$\langle \zeta \rangle = V^{-1} \int \hat{z} \cdot \omega dV. \quad (47.44)$$

Write the expression in terms of the area average vorticity, $\bar{\zeta}$, the resting layer thickness, H , and the deviation of the surface height from resting, $\Delta\eta$.

EXERCISE 47.2: APPLICATIONS OF SHALLOW WATER PV CONSERVATION

In an adiabatic shallow water fluid in a rotating reference frame, show that the potential vorticity conservation law is

$$\frac{D}{Dt} \left[\frac{\zeta + f}{\eta - \eta_b} \right] = 0, \quad (47.45)$$

where η is the height of the free surface and η_b is the height of the bottom topography (see Figure 42.1). For both of the following questions, assume constant volume for the fluid column. Also, assume the column rotates about its axis as a solid-body.

- (a) A cylindrical column of air at 30° 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°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°N , starting from the relative vorticity it obtained from the plateau. What is its new relative vorticity?

EXERCISE 47.3: APPLICATION OF SHALLOW WATER PV CONSERVATION

An air column at 60°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°N , what is its absolute vorticity and relative vorticity as it passes the mountaintop? Hint: Use PV conservation for a shallow-water fluid, and assume the top of the column remains at 10 km.

EXERCISE 47.4: SHALLOW WATER PV WITH FRICTION

Consider the rotating shallow water equations with friction

$$\frac{Du}{Dt} + f \hat{z} \wedge \mathbf{u} = -g \nabla \eta + \mathbf{F} \quad \text{and} \quad \frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} = 0, \quad (47.46)$$

where \mathbf{F} is a friction operator. Let this fluid be in a simply-connected domain, \mathcal{D} , with a static lateral boundary $\partial\mathcal{D}$ and no-normal flow boundary conditions.

- (a) What is the material evolution equation for PV?
- (b) Show that the time evolution of globally integrated shallow water PV is determined only by contributions from friction along the lateral boundary.
- (c) Likewise, show that the time tendency for the circulation around the lateral boundary is effected only by the “circulation” of friction around the boundary (i.e., the oriented line integral of friction around the boundary).

EXERCISE 47.5: SOME FLOW PROPERTIES OF THE STEADY STATE SHALLOW WATER

Consider a single layer of shallow water fluid in steady state (i.e., all Eulerian time derivatives vanish).

- (a) Show that there exists a streamfunction for the steady state thickness weighted horizontal flow

$$h \mathbf{u} = \nabla \wedge (\hat{\mathbf{z}} \Psi). \quad (47.47)$$

- (b) What are the physical dimensions of Ψ ?
- (c) Show that the shallow water potential vorticity is a constant along the steady state streamlines of the thickness weighted flow

$$Q = Q(\Psi). \quad (47.48)$$

- (d) Show that the Bernoulli function,

$$B = g \eta + \mathbf{u} \cdot \mathbf{u}/2 \quad (47.49)$$

is also a constant along the same streamlines; i.e.,

$$B = B(\Psi). \quad (47.50)$$

- (e) What is the functional relation between the Bernoulli function and the potential vorticity?

EXERCISE 47.6: ZONALLY SYMMETRIC SHALLOW WATER FRONT

Consider a single layer of shallow water fluid on a β -plane with a flat bottom. Assume all fields possess zonal symmetry as in the zonal front shown in Figure 47.8.

- (a) Write the potential vorticity, Q , assuming the fluid is in geostrophic balance. Write in terms of meridional derivatives of the layer thickness.
- (b) From the shallow water equations, explicitly show that the PV is materially constant (i.e., it is a Lagrangian invariant). To do so, work through the usual shallow water PV conservation derivation yet make use of the zonally symmetric equations of motion. Show all relevant steps.
- (c) 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 ?
- (d) 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.

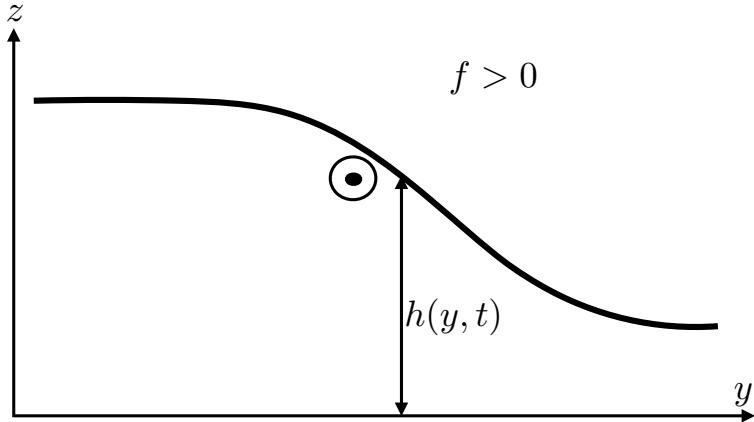


Figure 47.8: Schematic of a zonally symmetric front in a shallow water layer in the northern hemisphere ($f > 0$). The thickness decreases to the north. If the flow is in geostrophic balance, then the northward pressure gradient is in geostrophic balance with a southward Coriolis acceleration arising from an eastward (out of the page) geostrophic current (see also Figure 43.1).

EXERCISE 47.7: 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 Section 29.3 or the 2d barotropic equation in Section 53.1). 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

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla(g\eta + \alpha\Gamma) \quad (47.51a)$$

$$\frac{Dh}{Dt} = -h\nabla \cdot \mathbf{u}. \quad (47.51b)$$

The parameter $\alpha > 0$ is a constant and the field Γ is given by the Laplacian of the horizontal flow divergence

$$\Gamma = \nabla^2 \mathcal{D}, \quad (47.52)$$

where

$$\mathcal{D} = \nabla \cdot \mathbf{u}. \quad (47.53)$$

The divergence has physical dimensions of inverse time (T^{-1}), so that its Laplacian, Γ , has dimensions of $L^{-2} T^{-1}$, and the coefficient α 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

$$\tilde{\eta} = \eta + \frac{\alpha \Gamma}{g}. \quad (47.54)$$

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 (47.13) remains unchanged in the presence of divergence-damping.
- (b) Show that the potential vorticity equation (47.25) remains unchanged in the presence of divergence-damping.
- (c) Show that the horizontal divergence evolves according to

$$\frac{\partial \mathcal{D}}{\partial t} = \left[\frac{\partial \mathcal{D}}{\partial t} \right]_{\alpha=0} - \alpha \nabla^2 \Gamma. \quad (47.55)$$

- (d) Show that the evolution of gravitational potential energy per horizontal area

$$\mathcal{P} = g \rho \int_{\eta_b}^{\eta} z \, dz \quad (47.56)$$

remains unchanged from that determined in Section 43.3.1.

- (e) Show that the kinetic energy per horizontal area evolves according to

$$\frac{\partial \mathcal{K}}{\partial t} + \nabla \cdot (\mathbf{u} \mathcal{K}) = -h \rho g \nabla \tilde{\eta}, \quad (47.57)$$

where

$$\mathcal{K} = \frac{1}{2} \int_{\eta_b}^{\eta} \rho \mathbf{u}^2 \, dz = \rho h \mathbf{u}^2 / 2, \quad (47.58)$$

is the horizontal kinetic energy per area (Section 43.3.2).

- (f) Determine the evolution equation for global integrated kinetic energy

$$\frac{\partial}{\partial t} \left[\int \mathcal{K} \, dA \right] = \frac{\partial}{\partial t} \left[\int \int_{\eta_b}^{\eta} (\rho \mathbf{u}^2 / 2) \, dz \, dA \right]. \quad (47.59)$$

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 43.5. Show that the material evolution of angular momentum relative to the vertical rotational axis is given by

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g \frac{\partial \eta}{\partial \phi} + \mathcal{T}. \quad (47.60)$$

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

$$\frac{\partial}{\partial t} \int L^z = \alpha \rho \int \Gamma \frac{\partial \eta}{\partial \phi} dA. \quad (47.61)$$

where we assume the bottom topography is flat so that $h = \eta$.

- (i) The linearized thickness equation (see Section 44.3) for a flat bottom is given by

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0, \quad (47.62)$$

where H is the thickness of the resting fluid layer. Show that the time change for the global integrated angular momentum is given by

$$\frac{\partial}{\partial t} \int L^z = -\frac{\alpha \rho}{H} \int \left[\frac{\partial}{\partial t} \nabla^2 \eta \right] \frac{\partial \eta}{\partial \phi} dA. \quad (47.63)$$

48

Vorticity mechanics

This chapter develops the basic kinematics and dynamics of vorticity, building from the shallow water discussion of Chapter 47 to here consider fully three-dimensional flows. Doing so provides us with a mathematical and physical framework to understand how physical processes affect vorticity in geophysical fluids.

READER'S GUIDE FOR THIS CHAPTER

This chapter requires notions from earlier chapters in this part of the book as well as the fluid kinematics from Part III and fluid dynamics from Part IV. As for the shallow water vorticity discussed in Chapter 47, we here make use of vector calculus identities for Cartesian coordinates as detailed in Chapter 4. The concepts and methods developed in this chapter are fundamental to the notions of vorticity, much of which is encountered in the remainder of this part of the book as well as for the balanced models considered in Part IX.

48.1	Kinematics of vortex lines and vortex tubes	740
48.1.1	Vortex lines and vortex tubes	740
48.1.2	Kinematic properties	741
48.1.3	Helmholz's theorems	742
48.1.4	Further study	743
48.2	Kelvin's Circulation Theorem	743
48.2.1	Formulation	744
48.2.2	Barotropic flow	745
48.2.3	Pressure gradient force for barotropic fluids	745
48.3	Mechanics of baroclinicity	745
48.4	Vorticity dynamics	746
48.4.1	Vector-invariant velocity equation	746
48.4.2	Basic form of the vorticity equation	747
48.4.3	Massaged form of the vorticity equation	747
48.5	Vorticity filaments and material line elements	748
48.5.1	Frozen-in nature of vorticity	748
48.5.2	Stretching and tilting of vortex tubes	748
48.6	Circulation and vorticity for rotating fluids	751
48.6.1	Material evolution of absolute circulation	752
48.6.2	The beta effect	753
48.6.3	A two-dimensional fluid example	754
48.7	Exercises	755

48.1 Kinematics of vortex lines and vortex tubes

We here develop the basics of vorticity kinematics, with this discussion closely following from the kinematics of material line elements discussed in Section 20.2.3.

48.1.1 Vortex lines and vortex tubes

A *vortex line* is a line drawn through the fluid that is tangent, at each instance in time, to the vorticity at each spatial point. A vortex line is mathematically parameterized just like any other line, whereby we write the spatial coordinates along the line as a function of a suitable parameter φ (e.g., the arc-length)

$$\mathbf{x}(\varphi) = x(\varphi) \hat{\mathbf{x}} + y(\varphi) \hat{\mathbf{y}} + z(\varphi) \hat{\mathbf{z}}. \quad (48.1)$$

The three coordinates of the line are constrained so that the line is tangent to vorticity at each point, which means

$$\frac{dx/d\varphi}{\omega_x} = \frac{dy/d\varphi}{\omega_y} = \frac{dz/d\varphi}{\omega_z}. \quad (48.2)$$

These equations are directly analogous to those satisfied by velocity streamlines (Section 16.7.2)

$$\frac{dx/d\varsigma}{u} = \frac{dy/d\varsigma}{v} = \frac{dz/d\varsigma}{w}, \quad (48.3)$$

where ς is the parameter along the streamline. Notably, the velocity is not constant along a velocity streamline, nor is vorticity constant along a vortex line. In a steady state, streamlines map the trajectory of a fluid particle (see Section 16.7). However, a vortex line does not offer an interpretation in terms of trajectories.

A *vortex tube* is a bundle of vortex lines that pass through a simple closed curve, with Figure 48.1 illustrating a sample tube. By definition, the sides of the vortex tube are parallel to the vorticity field, since the sides are constructed from vortex lines. We defined a similar notion, the streamtube, for a non-divergent velocity field in Figure 16.5.

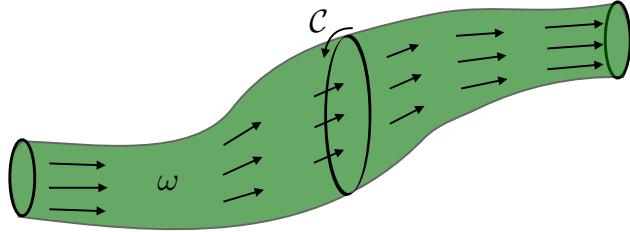


Figure 48.1: A vortex line is a line in the fluid that is everywhere tangent to the vorticity vector. A vortex tube is the accumulation of vortex lines passing through a closed loop. A vortex tube is sometimes referred to as a vortex filament. We here depict a vortex tube within the fluid and illustrate the circulation around the tube $C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS$. Since vorticity has zero divergence, the circulation is the same for any loop around the vortex tube (Helmholtz's first theorem from Section 48.1.3). 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 where the tube has a large area.

48.1.2 Kinematic properties

Vorticity has zero divergence

$$\nabla \cdot \boldsymbol{\omega} = \nabla \cdot (\nabla \wedge \mathbf{v}) = 0, \quad (48.4)$$

which follows since the divergence of a curl vanishes. Integrating the non-divergence relation over an arbitrary closed volume within the fluid leads to

$$\int_{\mathcal{R}} \nabla \cdot \boldsymbol{\omega} dV = \oint_{\partial \mathcal{R}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0, \quad (48.5)$$

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 closed region. Consequently, there are no sources or sinks of vorticity within the fluid. In turn, there is no accumulation of vorticity within any arbitrary closed region.

Now specialize the surface integral in equation (48.5) to a volume along a chosen vortex tube such as in Figure 48.1.¹ The two ends of the tube generally have different cross-sectional areas. The integral over the sides of the vortex tube vanishes, since the vorticity is parallel to the tube sides. Hence, the surface integral only picks up contributions from the two ends of the tube

$$\int_A \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS_A + \int_B \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS_B = 0. \quad (48.6)$$

Note that the outward normals point in the opposite direction, so that we see that the flux of vorticity is independent of position along the tube. Correspondingly, Stoke's Theorem transfers

¹In Exercise 19.8, we developed a similar set of results for a streamtube in an incompressible fluid.

the vorticity constraint to a constraint on the circulation around the circumference of the tube, so that

$$\oint_A \mathbf{v} \cdot d\mathbf{r} + \oint_B \mathbf{v} \cdot d\mathbf{r} = 0. \quad (48.7)$$

We thus see that the circulation around the vortex tube is the same no matter where it is computed. The circulation constraints (48.6) and (48.7) are kinematic, holding for any vorticity field. We now consider some consequences of this constraint.

48.1.3 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.

Helmholz's first theorem

Since the cross-sectional slices used to derive the circulation constraint (48.7) are arbitrary, the constraint holds throughout the full extent of the vortex tube. Hence, as noted following equation (48.7), the circulation is the same for any position along the vortex tube; i.e., the strength of a vortex tube is constant along its length (see Figure 48.1). This result is known as Helmholtz's first theorem.

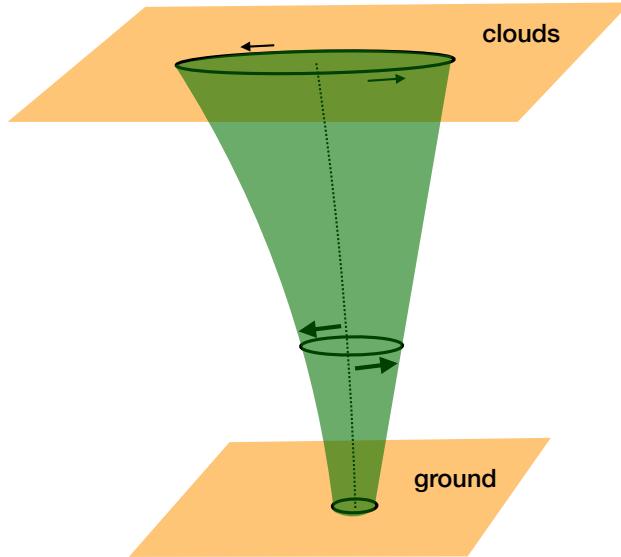


Figure 48.2: A vortex tube idealization of a tornado. Since the circulation around the tube is uniform (Helmholz's first theorem), the tangential velocity of a fluid particle has a larger magnitude in regions where the vortex area is smaller, such as near the ground. As the tornado reaches into the clouds, 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 (48.6) to note that changes in the vortex tube cross-sectional area are 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 48.2, 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.

Helmholz's second theorem

The vorticity constraint (48.6) 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 48.3), or intersect a boundary (as for a tornado in Figure 48.2, where the ground and clouds form the boundary).

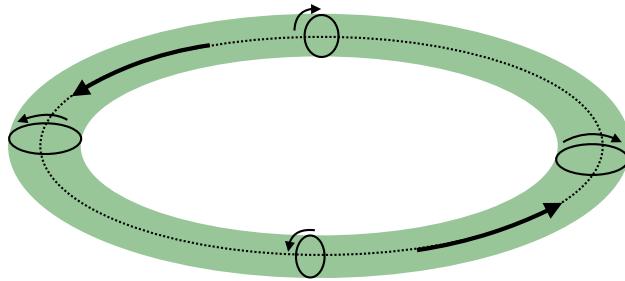


Figure 48.3: A vortex ring (torus) is a vortex tube that closes on itself. We here depict a vortex ring with vorticity pointing counter-clockwise around the ring. The tangential velocity is oriented as shown so that the vorticity points according to the right hand rule. That is, orient the fingers on the right hand according to the tangential velocity. The thumb of the right hand then points in the direction of the vorticity vector.

Helmholz's third theorem

Helmholz's third theorem states that an unforced inviscid barotropic fluid that has zero vorticity will remain irrotational forever. This theorem is a special case of Kelvin's Circulation Theorem, which is discussed in Section 48.2.

48.1.4 Further study

A particularly insightful and pedagogical discussion of these ideas 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.
- A rotating tank experiment shown near the 15 minute mark of [this video from Prof. Dave Fultz of the University of Chicago](#) shows how vorticity is affected by vortex stretching.
- [This video](#) offers a classic tutorial on vorticity in non-rotating fluids from Prof. A. Shapiro of MIT.

48.2 Kelvin's Circulation Theorem

Kelvin's Circulation Theorem is concerned with the evolution of circulation around a closed material loop, or equivalently (through Stokes' Theorem) with the change in vorticity penetrating the

enclosed area

$$\frac{dC}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} \mathbf{v} \cdot d\mathbf{r} = \frac{d}{dt} \int_{S(\mathbf{v})} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (48.8)$$

where $S(\mathbf{v})$ designates a surface whose points all move with the fluid flow. We here consider non-rotating flow, with the non-rotating flow directly applicable to the absolute circulation (relative circulation plus circulation due to planetary rotation) found for the rotating case in Section 48.6.

48.2.1 Formulation

We are here concerned with circulation for material loops. Hence, the time derivative acting on the circulation integral in equation (48.8) moves inside the integral as a material time operator

$$\frac{dC}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} \mathbf{v} \cdot d\mathbf{r} = \oint_{\partial S(\mathbf{v})} \frac{D(\mathbf{v} \cdot d\mathbf{r})}{Dt}. \quad (48.9)$$

The material evolution of \mathbf{v} is determined by Newton's Law of motion, which for a non-rotating flow is given by (see Section 22.1.4)

$$\frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F}. \quad (48.10)$$

In this equation, p is the pressure, ρ is the mass density, Φ is the geopotential and/or the potential for any conservative force, and \mathbf{F} is a force arising from non-conservative viscous stresses.

The material time derivative of the differential line element moving around the circuit equals to the differential of the velocity on the circuit

$$\frac{D(d\mathbf{r})}{Dt} = d\mathbf{v}. \quad (48.11)$$

This result follows since all points along the circuit are defined by fluid particles. Consequently, evolution of circulation following a material loop becomes

$$\frac{dC}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} \mathbf{v} \cdot d\mathbf{r} \quad (48.12a)$$

$$= \oint_{\partial S(\mathbf{v})} \left[\left(-\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F} \right) \cdot d\mathbf{r} + \frac{1}{2} d\mathbf{v}^2 \right] \quad (48.12b)$$

$$= \oint_{\partial S(\mathbf{v})} \left[-\frac{1}{\rho} \nabla p + \mathbf{F} \right] \cdot d\mathbf{r} \quad (48.12c)$$

$$= \int_{S(\mathbf{v})} \left[-\nabla \wedge \left(\frac{1}{\rho} \nabla p \right) + \nabla \wedge \mathbf{F} \right] \cdot \hat{\mathbf{n}} dS \quad (48.12d)$$

$$= \int_{S(\mathbf{v})} (\mathbf{B} + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad (48.12e)$$

The third equality follows since $\nabla \Phi \cdot d\mathbf{r} = d\Phi$ has zero integral around a closed circuit, as does $d\mathbf{v}^2$. The fourth equality made use of Stokes' theorem. The final equality introduced the solenoidal vector

$$\mathbf{B} = \frac{\nabla \rho \wedge \nabla p}{\rho^2}, \quad (48.13)$$

which we also refer to as the *baroclinicity vector*. The baroclinicity vector has physical dimensions of inverse squared time, T^{-2} . Equation (48.12e) says that the circulation around a material loop is affected by two processes: baroclinicity and the curl of any non-conservative (e.g., friction) forces.

48.2.2 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

$$p = p(\rho) \Rightarrow \text{barotropic flow.} \quad (48.14)$$

Kelvin's theorem then follows, which states that for inviscid barotropic flow the circulation around any closed material circuit remains constant

$$\frac{d\mathcal{C}}{dt} = \frac{d}{dt} \oint_{\partial\mathcal{S}(\mathbf{v})} \mathbf{v} \cdot d\mathbf{r} = \frac{d}{dt} \int_{\mathcal{S}(\mathbf{v})} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} = 0 \quad \Leftarrow \text{inviscid barotropic flow.} \quad (48.15)$$

That is, the circulation around any vortex tube in a perfect barotropic fluid moves in a manner that keeps the circulation materially constant. This remarkable result greatly constrains the motion of a barotropic perfect fluid.

48.2.3 Pressure gradient force for barotropic fluids

As noted above, a barotropic fluid has pressure that is a function just of density, and conversely the density is a function just of the pressure. Hence, the curl of the pressure gradient force vanishes

$$\nabla \wedge (\rho^{-1} \nabla p) = 0 \quad (48.16)$$

and there is a corresponding scalar potential for the pressure gradient force

$$\nabla \Phi_p = \rho^{-1} \nabla p. \quad (48.17)$$

Integration renders (see page 79 of [Müller \(2006\)](#))

$$\Phi_p = \int_{p_0}^p \frac{dp'}{\rho(p')}, \quad (48.18)$$

where p_0 is an arbitrary reference pressure. We make use of the scalar potential Φ_p in Exercise 48.5.

48.3 Mechanics of baroclinicity

Baroclinicity is present in all realistic geophysical flows, thus affecting the material evolution of circulation, and correspondingly the evolution of vorticity as seen in Section 48.4. Flow with a nonzero baroclinicity vector is generally referred to as *baroclinic flow*. We here offer a mechanical interpretation of baroclinicity as a torque that modifies vorticity/circulation.

First note that the baroclinicity vector can be written as

$$\mathbf{B} = \frac{\nabla\rho \wedge \nabla p}{\rho^2} = -\nabla\rho^{-1} \wedge \nabla p. \quad (48.19)$$

A solenoid is a tube perpendicular to both $\nabla\rho$ and ∇p . Solenoids vanish for barotropic flows, whereby $p = p(\rho)$ (see equation (48.14)). For baroclinic flow, solenoids introduce a torque at each point which in turn affects vorticity.

To further understand the mechanical interpretation of solenoids in terms of a torque, consider the cross product

$$\rho \mathbf{B} = (-\rho^{-1} \nabla p) \wedge \nabla\rho. \quad (48.20)$$

The first term on the right hand side is the pressure gradient force that acts down the pressure gradient. Now consider a tiny fluid element such as shown in Figure 48.4. 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. Since the pressure force does not pass through the center of mass, it imparts a torque to the fluid element. This torque then modifies the vorticity and hence the circulation around the boundary of the element. Only when the pressure force is aligned with the density gradient (barotropic flow), or if the density is spatially uniform (e.g., constant density shallow water layer) does the pressure force pass through the center of mass, thus creating no torque and not affecting vorticity or circulation.

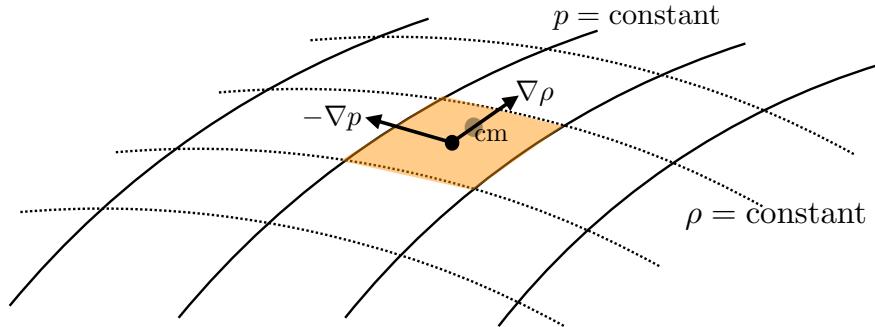


Figure 48.4: A mechanical interpretation of the baroclinicity vector. We consider a tiny fluid element bounded by surfaces of constant pressure and density. By construction, the pressure force acts at the geometric center of the element, whereas the center of mass for the element is off-center due to the density gradient across the element. The pressure 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 $(-\rho^{-1} \nabla p) \wedge \nabla \rho$), so that this baroclinicity spins-up a clockwise circulation around the element, or equivalently a clockwise vorticity. This figure is based on Figure 14.9 of [Thorne and Blandford \(2017\)](#).

48.4 Vorticity dynamics

We now move from the circulation around a macroscopic circuit to the vorticity at a point. In particular, we seek information for how vorticity changes in time. What physical processes lead to these changes? As for Kelvin's theorem, we make use of Newton's law of motion, written here in the form for a rotating fluid (see Section 22.1.4)

$$\rho \left[\frac{D}{Dt} + 2\boldsymbol{\Omega} \wedge \right] \mathbf{v} = -\nabla p - \rho \nabla \Phi + \rho \mathbf{F}, \quad (48.21)$$

where $\boldsymbol{\Omega}$ is the angular velocity of the rotating reference frame.

48.4.1 Vector-invariant velocity equation

As for the shallow water fluid in Section 47.1, we find it useful to convert the advective-form momentum equation to vector-invariant form. For this purpose, make use of the vector identity (see Section 4.3.4)

$$\boldsymbol{\omega} \wedge \mathbf{v} = -(1/2) \nabla(\mathbf{v} \cdot \mathbf{v}) + (\mathbf{v} \cdot \nabla) \mathbf{v} \quad (48.22)$$

to eliminate velocity self-advection in favor of vorticity

$$\frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega}_a \wedge \mathbf{v} = -\frac{1}{\rho} \nabla p - \nabla \left[\frac{1}{2} \mathbf{v}^2 + \Phi \right] + \mathbf{F}. \quad (48.23)$$

We here introduced the absolute vorticity

$$\boldsymbol{\omega}_a = \boldsymbol{\omega} + 2 \boldsymbol{\Omega}, \quad (48.24)$$

which is the sum of the relative vorticity and the planetary vorticity (see Section 46.5.1).

48.4.2 Basic form of the vorticity equation

Taking the curl of the vector-invariant momentum equation (48.23) removes the mechanical energy per mass, $\mathbf{v}^2/2 + \Phi$, thus leaving

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v}) = \frac{1}{\rho^2} (\nabla \rho \wedge \nabla p) + \nabla \wedge \mathbf{F}. \quad (48.25)$$

For geophysical fluids we generally assume that $\boldsymbol{\Omega}$ has zero time tendency, so that

$$\frac{\partial \boldsymbol{\omega}_a}{\partial t} = \frac{\partial (\boldsymbol{\omega} + 2 \boldsymbol{\Omega})}{\partial t} = \frac{\partial \boldsymbol{\omega}}{\partial t}, \quad (48.26)$$

in which case equation (48.25) can be written

$$\frac{\partial \boldsymbol{\omega}_a}{\partial t} + \nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v}) = \mathbf{B} + \nabla \wedge \mathbf{F}, \quad (48.27)$$

where \mathbf{B} is the baroclinicity vector introduced by equation (48.13).

48.4.3 Massaged form of the vorticity equation

Physical interpretation of the term $\nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v})$ appearing in the prognostic equation (48.27) can be made more transparent by using yet another vector identity

$$\nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v}) = (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \boldsymbol{\omega}_a \nabla \cdot \mathbf{v} - \mathbf{v} \nabla \cdot \boldsymbol{\omega}_a \quad (48.28a)$$

$$= (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} - \frac{\boldsymbol{\omega}_a}{\rho} \frac{D\rho}{Dt}. \quad (48.28b)$$

The second equality required the continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}, \quad (48.29)$$

and the non-divergent nature of the absolute vorticity

$$\nabla \cdot \boldsymbol{\omega}_a = \nabla \cdot (\nabla \wedge \mathbf{v} + 2 \boldsymbol{\Omega}) = 0. \quad (48.30)$$

Equation (48.27) thus takes the form

$$\frac{D\boldsymbol{\omega}_a}{Dt} - \frac{\boldsymbol{\omega}_a}{\rho} \frac{D\rho}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \frac{1}{\rho^2} (\nabla \rho \wedge \nabla p) + \nabla \wedge \mathbf{F}, \quad (48.31)$$

which can be written

$$\rho \frac{D(\omega_a/\rho)}{Dt} = (\omega_a \cdot \nabla) v + B + \nabla \wedge F. \quad (48.32)$$

Equation (48.32) is the desired form of the vorticity evolution equation. Each term on the right hand side represents a distinct physical process that impacts material evolution of vorticity. The first term, $(\omega_a \cdot \nabla) v$, will be explored in Section 48.5 in the simplified context of a barotropic fluid. The second term arises from baroclinicity as introduced in equation (48.13) and given a mechanical interpretation in Section 48.3. The third term arises from the curl of the friction vector, thus contributing especially in boundary layer regions where friction is large.

48.5 Vorticity filaments and material line elements

To help develop an understanding for the term $(\omega_a \cdot \nabla) v$ appearing in the vorticity equation (48.32), consider the special case of an incompressible, inviscid, non-rotating barotropic fluid, in which case the vorticity equation (48.32) reduces to

$$\frac{D\omega}{Dt} = (\omega \cdot \nabla) v. \quad (48.33)$$

Evolution of vorticity due to the source term $(\omega \cdot \nabla) v$ is kinematically identical to evolution of a material line element detailed in Section 20.2. We pursue this connection in the following.

48.5.1 Frozen-in nature of vorticity

Recall from Section 48.1.1 that a vortex line is a line drawn through the fluid that is everywhere parallel to the vorticity. Such a line connects material fluid elements, so that a vortex line constitutes a material line. At time $t = 0$, let the vorticity on an infinitesimal vortex line be related to the initial material line element

$$\delta x(0) = \Gamma \omega(x, 0), \quad (48.34)$$

where Γ is scalar with dimensions LT that is determined by the initial vorticity and initial line element. Importantly, this relation follows by construction. We are free to draw an infinitesimal vortex line and call it a material line. The key point is that the vorticity equation (48.33) has precisely the same mathematical form as the material line element equation (20.19)

$$\frac{D(\delta x)}{Dt} = (\delta x \cdot \nabla) v. \quad (48.35)$$

Consequently, the relation (48.34) holds for all time with Γ a constant. That is, the line element and vorticity evolve according to the same material equation, so the line element and vorticity forever maintain the relation $\delta x = \Gamma \omega$. In this way, the vorticity is attached to the particular material line element, motivating one to say that vorticity is a “frozen-in” property. This property is illustrated in Figure 48.5. Again, this property holds only for the case of an incompressible, inviscid, barotropic fluid. Nonetheless, it offers great insight into the more general situation occurring in real fluids.

48.5.2 Stretching and tilting of vortex tubes

Vorticity responds when vortex lines or tubes are stretched or bent. To help understand the response, consider again the barotropic incompressible vorticity equation (48.33) and focus just the vertical vorticity component

$$\frac{D\omega^z}{Dt} = \omega^x \frac{\partial w}{\partial x} + \omega^y \left(\frac{\partial w}{\partial y} \right) + \omega^z \left(\frac{\partial w}{\partial z} \right). \quad (48.36)$$

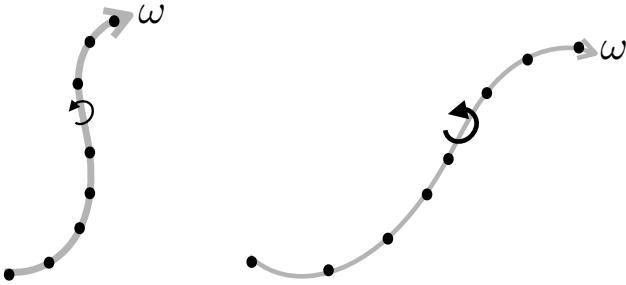


Figure 48.5: For an incompressible perfect barotropic fluid, vortex lines (also known as vortex filaments) are also material lines. This property means that for an arbitrary vortex line drawn in the fluid, the fluid particles that are initially on the vortex line will remain on the line as it moves through the fluid. 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 48.5.2. The material property of vortex lines is known as their *frozen-in nature*. The frozen-in nature of vortex lines strictly holds only for perfect incompressible barotropic fluid, yet it offers useful insight into vorticity dynamics for more general fluids.

Note that since $\nabla \cdot \boldsymbol{\omega} = 0$, we can write this equation as

$$\frac{D\omega^z}{Dt} = \nabla \cdot (\boldsymbol{w} \boldsymbol{\omega}), \quad (48.37)$$

though in the following we find it more useful to focus on the form given by equation (48.36). The following discussion closely emulates that given for a material line element in Section 20.2.5.

Stretching

Consider the vortex tube to be initially aligned with the (vertical) z -axis, so that $\omega^x = \omega^y = 0$, in which case there is only a single term impacting vertical vorticity²

$$\frac{D\omega^z}{Dt} = \omega^z \frac{\partial w}{\partial z}. \quad (48.38)$$

Since the fluid is incompressible, the volume of an infinitesimal portion of the vortex tube is materially constant

$$\frac{D(\delta V)}{Dt} = 0, \quad (48.39)$$

which means that the vertical extent, δz , and cross-sectional area, δA , are constrained

$$\frac{1}{\delta z} \frac{D(\delta z)}{Dt} + \frac{1}{\delta A} \frac{D(\delta A)}{Dt} = 0. \quad (48.40)$$

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 (48.35)) allows us to write

$$\frac{1}{\delta z} \frac{D(\delta z)}{Dt} = \frac{\partial w}{\partial z}, \quad (48.41)$$

²Be mindful to distinguish the symbols for the vertical component of vorticity, ω^z , and the vertical component of velocity, w .

so that the vorticity equation (48.38) becomes

$$\frac{D\omega^z}{Dt} = \omega^z \frac{\partial w}{\partial z} \quad (48.42a)$$

$$= \omega^z \left[\frac{1}{\delta z} \frac{D(\delta z)}{Dt} \right] \quad (48.42b)$$

$$= -\omega^z \left[\frac{1}{\delta A} \frac{D(\delta A)}{Dt} \right]. \quad (48.42c)$$

Rearrangement leads to

$$\frac{D(\omega^z \delta A)}{Dt} = 0, \quad (48.43)$$

which is an expression of Kelvin's circulation theorem (equation (48.15)) for a cross-section of the vortex tube.

The above manipulations suggest the following interpretation for the *stretching* term $\omega^z (\partial w / \partial z)$ appearing in the vertical vorticity equation (48.36) and illustrated in Figure 48.6. 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 48.1.3). 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 also accords with our understanding of angular momentum conservation as discussed for the rotating cylinder in Section 47.2.2 and depicted by Figure 47.1.

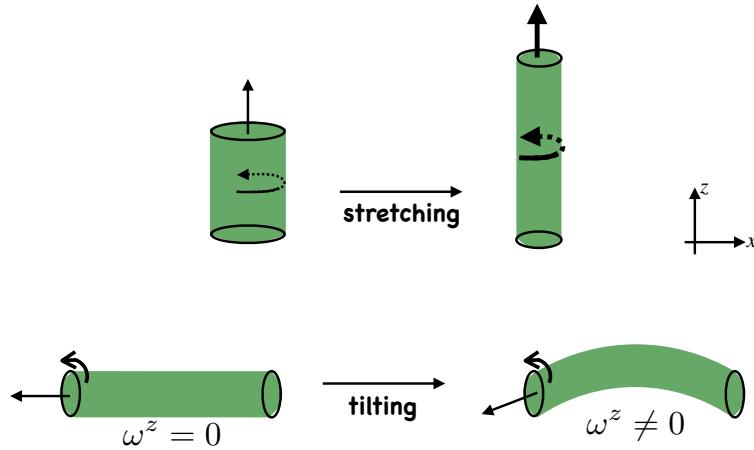


Figure 48.6: Illustrating how stretching and tilting of a vortex tube impacts on the vorticity. Top panels: As the cross-sectional area of the vortex tube shrinks, and the vertical extent of the tube stretches, the magnitude of the vorticity along the axis of the tube increases. This result accords with our understanding of angular momentum conservation as discussed for the rotating cylinder in Section 47.2.2 and depicted by Figure 47.1, as well as with Helmholtz's first theorem in Section 48.1.3 and Figure 48.2. 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 (i.e., $\partial w / \partial x \neq 0$) 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 48.6 so that $\omega^z = 0$. Focus on just one of the two horizontal directions, so that equation (48.36) for the vertical

vorticity becomes

$$\frac{D\omega^z}{Dt} = \omega^x \frac{\partial w}{\partial x}. \quad (48.44)$$

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 (48.35)) is given by

$$\frac{D(\delta\mathbf{x})}{Dt} = \delta x \frac{\partial \mathbf{v}}{\partial x}. \quad (48.45)$$

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.

48.6 Circulation and vorticity for rotating fluids

The previous sections focused on circulation and vorticity for non-rotating fluids. The discussion here for rotating fluids is a straightforward extension. Even so, the implications are quite profound for the motion of geophysical fluids.

We start by recalling the expression from Section 13.9.1 for the inertial or absolute velocity (i.e., velocity measured in an inertial frame)

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (48.46)$$

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

$$\mathcal{C}_a = \oint_{\partial\mathcal{S}} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x}) \cdot d\mathbf{r} = \mathcal{C} + \mathcal{C}_{\text{planet}}, \quad (48.47)$$

where the circulation measured in the rotating frame is

$$\mathcal{C} = \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} \quad (48.48)$$

and the circulation associated with the rotating planet is

$$\mathcal{C}_{\text{planet}} = \oint_{\partial\mathcal{S}} (\boldsymbol{\Omega} \wedge \mathbf{x}) \cdot d\mathbf{r}. \quad (48.49)$$

Again, $d\mathbf{r}$ is the differential line element moving around the circuit as mentioned in Section 46.3.

A fluid particle at rest in the rotating frame will still have a nonzero absolute circulation as given by the planetary circulation. Making use of Stokes' theorem leads to the equivalent forms for the circulations

$$\mathcal{C} = \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} = \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad \text{relative circulation} \quad (48.50a)$$

$$\mathcal{C}_{\text{planet}} = \oint_{\partial\mathcal{S}} (\boldsymbol{\Omega} \wedge \mathbf{x}) \cdot d\mathbf{r} = \int_{\mathcal{S}} \boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{n}} d\mathcal{S} \quad \text{planetary circulation} \quad (48.50b)$$

$$\mathcal{C}_a = \oint_{\partial\mathcal{S}} \mathbf{v}_a \cdot d\mathbf{r} = \int_{\mathcal{S}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad \text{absolute circulation,} \quad (48.50c)$$

where

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad \text{relative vorticity} \quad (48.51\text{a})$$

$$\boldsymbol{\omega}_{\text{planet}} = \nabla \wedge (\boldsymbol{\Omega} \wedge \mathbf{x}) = 2\boldsymbol{\Omega} \quad \text{planetary vorticity} \quad (48.51\text{b})$$

$$\boldsymbol{\omega}_{\text{a}} = \boldsymbol{\omega} + \boldsymbol{\omega}_{\text{planet}} \quad \text{absolute vorticity.} \quad (48.51\text{c})$$

Thus far we have merely substituted in the expression (48.46) 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 will see how the relative and planetary circulations interact.

48.6.1 Material evolution of absolute circulation

We now consider how the absolute circulation evolves for a material circuit that moves with the fluid

$$\frac{dC_{\text{a}}}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{r}, \quad (48.52)$$

where

$$\mathbf{r} = \mathbf{x} \quad (48.53)$$

is now the position of a fluid particle on the circuit, and $d\mathbf{r}$ is a differential line element around a circuit of material fluid particles. We measure fluid motion in the rotating frame so that the material time derivative contains advection by the velocity \mathbf{v} rather than the absolute velocity \mathbf{v}_{a} . Following the discussion for non-rotating Kelvin's Circulation Theorem in Section 48.2 leads to

$$\frac{dC_{\text{a}}}{dt} = \frac{d}{dt} \oint_{\partial S(\mathbf{v})} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{r} \quad (48.54\text{a})$$

$$= \oint_{\partial S(\mathbf{v})} \left[\frac{D\mathbf{v}}{Dt} + \boldsymbol{\Omega} \wedge \frac{D\mathbf{r}}{Dt} \right] \cdot d\mathbf{r} + \oint_{\partial S(\mathbf{v})} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{v} \quad (48.54\text{b})$$

$$= \oint_{\partial S(\mathbf{v})} \left[\frac{D\mathbf{v}}{Dt} + \boldsymbol{\Omega} \wedge \mathbf{v} \right] \cdot d\mathbf{r} + \oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{v} \quad (48.54\text{c})$$

$$= \oint_{\partial S(\mathbf{v})} \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] \cdot d\mathbf{r}. \quad (48.54\text{d})$$

To reach this result we set

$$\mathbf{v} = \frac{D\mathbf{r}}{Dt}, \quad (48.55)$$

for the velocity of a fluid particle on the circuit. We also used the identity

$$\oint_{\partial S(\mathbf{v})} \mathbf{v} \cdot d\mathbf{v} = \frac{1}{2} \oint_{\partial S(\mathbf{v})} d\mathbf{v}^2 = 0 \quad (48.56)$$

as well as

$$\oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{v} = \oint_{\partial S(\mathbf{v})} d[(\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot \mathbf{v}] - \oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \wedge d\mathbf{r}) \cdot \mathbf{v} = \oint_{\partial S(\mathbf{v})} (\boldsymbol{\Omega} \wedge \mathbf{v}) \cdot d\mathbf{r}. \quad (48.57)$$

Now insert the momentum equation (48.21) into equation (48.54d) to yield

$$\frac{dC_a}{dt} = \oint_{\partial S(v)} \left[\frac{Dv}{Dt} + 2\Omega \wedge v \right] \cdot dr. \quad (48.58a)$$

$$= \oint_{\partial S(v)} \left[-\frac{1}{\rho} \nabla p - \nabla \Phi + F \right] \cdot dr. \quad (48.58b)$$

$$= \oint_{\partial S(v)} \left[-\frac{dp}{\rho} + F \cdot dr \right]. \quad (48.58c)$$

Making use of Stokes' Theorem leads to the evolution of absolute circulation around a material loop

$$\frac{dC_a}{dt} = \oint_{\partial S(v)} \left[-\frac{dp}{\rho} + F \cdot dr \right] = \int_{S(v)} (\mathbf{B} + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS, \quad (48.59)$$

where $\mathbf{B} = \rho^{-2} \nabla \rho \wedge \nabla p$ is the baroclinicity vector from equation (48.19).

The circulation theorem (48.59) is the same as obtained for the non-rotating Kelvin's Circulation Theorem discussed in Section 48.2 (see equation (48.12e)). This equivalence proves that circulation is an objective (frame invariant) property of the fluid, in which its evolution is unchanged when moving to a non-inertial rotating frame.

48.6.2 The beta effect

As given by equation (48.47), 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)

$$C_a = C + C_{\text{planet}} = C + 2 \int_S \Omega \cdot \hat{\mathbf{n}} dS \iff \frac{dC_a}{dt} = \frac{dC}{dt} + \frac{dC_{\text{planet}}}{dt}. \quad (48.60)$$

We can determine the processes that affect the absolute circulation around a material loop by using the circulation equation (48.59)

$$\frac{dC}{dt} = -\frac{dC_{\text{planet}}}{dt} + \frac{dC_a}{dt} \quad (48.61a)$$

$$= -2 \frac{d}{dt} \left[\int_{S(v)} \Omega \cdot \hat{\mathbf{n}} dS \right] + \int_{S(v)} (\mathbf{B} + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad (48.61b)$$

We generally assume that the planetary rotation is a constant in time and points through the north pole of the sphere³ $\Omega = \Omega \hat{\mathbf{Z}}$, so that

$$\int_{S(v)} \Omega \cdot \hat{\mathbf{n}} dS = \Omega \int_{S(v)} \hat{\mathbf{Z}} \cdot \hat{\mathbf{n}} dS = \Omega A_{\perp}. \quad (48.62)$$

The area A_{\perp} is the projection of the spherical area enclosed by the circuit onto the horizontal equatorial plane, with Figure 48.7 illustrating the geometry. This result has profound impact on

³We follow the notational conventions of Figure 13.4 with one exception. Here, 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.

large scale geophysical fluid motion, whereby relative circulation around a material circuit in the rotating frame changes according to

$$\frac{dC}{dt} = \underbrace{-2\Omega \frac{dA_{\perp}}{dt}}_{\text{beta effect}} + \underbrace{\int_S (\mathbf{B} + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS(\mathbf{v})}_{\text{solenoids plus friction}}. \quad (48.63)$$

[Holton \(1992\)](#) calls equation (48.63) the *Bjerknes Circulation Theorem* (see his equation (4.5)). The second term, comprised of baroclinicity and friction, also appears in the non-rotating case. It has already been discussed in Sections 48.2 and 48.3.

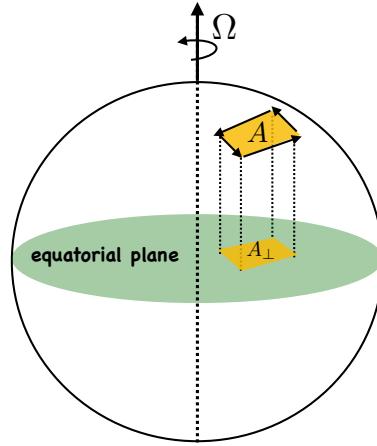


Figure 48.7: Geometry of the beta effect. According to the Bjerknes circulation theorem (48.63), the circulation for a material loop on the surface of a rotating sphere is affected by baroclinicity and friction, as for a non-rotating sphere, as well as latitudinal motion of the loop. The latitudinal motion alters the area of the loop as projected onto the equatorial plane. When multiplied by the magnitude of the planetary vorticity, 2Ω , 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*. The beta effect requires both rotation (2Ω) and curvature of the sphere ($\partial_y f = \beta$); it is therefore absent on the *f*-plane.

The first term in the circulation theorem (48.63) 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. Note that longitudinal motion has no impact on A_{\perp} , so that longitudinal motion does not alter the relative circulation. The material change in A_{\perp} , when multiplied by the magnitude of the planetary vorticity, modifies the relative circulation around the material circuit. We refer to *planetary induction* as the process whereby relative circulation changes due to latitudinal motion of a material circuit on a rotating sphere. Or more commonly, planetary induction is referred to as the *beta effect*, given its connection to the latitudinal gradient of the Coriolis parameter, $\beta = \partial_y f$. In certain theories of large-scale laminar planetary flows, the solenoidal/baroclinicity and friction terms are sub-dominant. In those cases the material evolution of relative circulation is dominated by the beta effect. Planetary geostrophic flow is just such an example, as studied in Section 29.4.

48.6.3 A two-dimensional fluid example

To garner more insight into the beta effect, consider a perfect incompressible and two-dimensional flow (zero vertical velocity) on a rotating sphere. In this case there is only a vertical component to

vorticity and both baroclinicity and friction have no role in vorticity dynamics. Hence, vorticity is affected only via the beta effect.

In the rotating frame, circulation around an infinitesimal closed material loop is

$$\mathcal{C} = A \zeta, \quad (48.64)$$

where ζ is the relative vorticity and A is the area enclosed by the loop. Because the fluid is incompressible, the loop area A remains constant even as the loop becomes contorted (see Section 19.6). The material evolution of circulation is therefore given by

$$\frac{D\mathcal{C}}{Dt} = A \frac{D\zeta}{Dt} = -2\Omega \frac{DA_{\perp}}{Dt}, \quad (48.65)$$

where the second equality follows from Bjerknes' circulation theorem (48.63). Let the material circuit be at a latitude ϕ so that the projection of the loop area onto the equatorial plane is (see Figure 48.7)

$$A_{\perp} = A \sin \phi. \quad (48.66)$$

Hence, material evolution of the circulation is

$$\frac{D\mathcal{C}}{Dt} = A \frac{D\zeta}{Dt} \quad (48.67a)$$

$$= -2\Omega \frac{DA_{\perp}}{Dt} \quad (48.67b)$$

$$= -2A\Omega \frac{D \sin \phi}{Dt} \quad (48.67c)$$

$$= -2A\Omega \cos \phi \frac{D\phi}{Dt} \quad (48.67d)$$

$$= -A \left[\frac{2\Omega \cos \phi}{R} \right] \left[R \frac{D\phi}{Dt} \right] \quad (48.67e)$$

$$= -A\beta v, \quad (48.67f)$$

where we introduced the meridional velocity component

$$v = R \frac{D\phi}{Dt} \quad (48.68)$$

and the meridional derivative of the planetary vorticity

$$\beta = \frac{df}{dy} = \frac{1}{R} \frac{d}{d\phi} (2\Omega \sin \phi) = \frac{2\Omega \cos \phi}{R}. \quad (48.69)$$

The result (48.67f) shows how meridional motion on a rotating sphere induces relative circulation. It furthermore motivates the name *beta effect* for planetary induction of relative vorticity.

48.7 Exercises

EXERCISE 48.1: FRICTION IN THE VORTICITY EQUATION

Add a viscous term of the form

$$\mathbf{F} = \mu \nabla^2 \mathbf{v}, \quad (48.70)$$

with μ a constant molecular viscosity. How is the vorticity equation modified?

EXERCISE 48.2: GENERATION OF VORTICITY BY BAROCLINICITY

Consider an initially resting 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

$$\rho(x) = \rho_0(1 - \gamma|x|) \quad (48.71)$$

where ρ_0 and γ 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 48.4 will help with this exercise.

- (a) Compute the density gradient $\nabla\rho$ and draw a schematic.
- (b) Compute the pressure gradient ∇p . Draw a schematic at $x = 0$.
- (c) Compute the baroclinicity/solenoidal vector $\mathbf{B} = \rho^{-2}(\nabla\rho \wedge \nabla p)$. Draw a schematic.
- (d) Describe the vorticity induced by the baroclinicity vector.

EXERCISE 48.3: CIRCULATION WITH ISLANDS

Our discussion of Stokes' Theorem has been thus far restricted to a simply connected domain, in which

$$\oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} = \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (48.72)$$

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 48.8 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 in an AquaPlanet or the global atmosphere.

Derive the following expression for the circulation in multiply-connected regions

$$\oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r} = \sum_{n=1}^N \left(\oint_{\partial\mathcal{S}_n} \mathbf{v} \cdot d\mathbf{r} \right) + \int_{\mathcal{S}} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (48.73)$$

where N is the number of islands and \mathcal{S}_n is the contour surrounding each island. 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 region. Removing the islands allows the island contours to be shrunk to zero size, in which case we recover the simply connected result (48.72).

EXERCISE 48.4: EVOLUTION OF CIRCULATION AROUND ISLANDS

The momentum equation for a homogeneous layer of inviscid shallow water on a tangent plane is given by

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta. \quad (48.74)$$

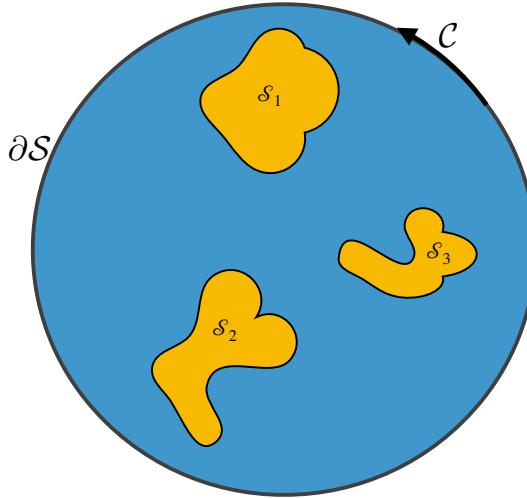


Figure 48.8: A region of the ocean consisting of three islands, S_1 , S_2 , and S_3 , each with boundaries ∂S_n and with the closed contour, ∂S , drawn around the three islands. The contour ∂S cannot be shrunk to a point without crossing over the islands, thus indicating that the domain is multiply connected. Exercise 48.3 is concerned with deriving an expression for the circulation of the depth-integrated flow as defined along the closed contour, ∂S .

In this equation, $\mathbf{u} = (u, v)$ is the horizontal velocity, f is the Coriolis parameter, g is the effective gravitational acceleration, and η is the deviation of the free surface from its horizontal resting position. All spatial derivatives are horizontal, so that

$$\mathbf{u} \cdot \nabla = u \partial_x + v \partial_y. \quad (48.75)$$

Use of a vector identity allows us to write

$$\frac{\partial \mathbf{u}}{\partial t} + (f + \zeta) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla (\mathbf{u}^2/2 + g \eta), \quad (48.76)$$

where

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}) \quad (48.77)$$

is the vorticity of the shallow water fluid.

Consider an island, such as one shown in Figure 48.8. Each island is static and impenetrable to fluid flow, which means that

$$\mathbf{u} \cdot \hat{\mathbf{n}} = 0 \quad (48.78)$$

where $\hat{\mathbf{n}}$ is the outward normal on an island boundary. This no-normal flow constraint means that the velocity just next to an island is parallel to the island⁴

$$\mathbf{u} \wedge d\mathbf{r} = 0. \quad (48.79)$$

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

$$\frac{d}{dt} \oint_I \mathbf{u} \cdot d\mathbf{r} = 0. \quad (48.80)$$

⁴This 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 condition.

Recall that Kelvin's circulation theorem is formulated for a material contour in an inviscid fluid. This exercise shows that the circulation theorem also holds for a material contour enclosing a static solid boundary.

EXERCISE 48.5: HELICITY FOR AN IDEAL BAROTROPIC FLUID

Consider a closed material volume, \mathcal{R} , of isentropic non-rotating homogenous barotropic fluid. 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,

$$\hat{\mathbf{n}} \cdot \boldsymbol{\omega} = 0. \quad (48.81)$$

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

$$\mathbb{H} = \int_{\mathcal{R}(v)} \mathbf{v} \cdot \boldsymbol{\omega} dV, \quad (48.82)$$

where the volume $\mathcal{R}(v)$ is material. In Cartesian coordinates, the helicity density takes the form

$$\mathbf{v} \cdot \boldsymbol{\omega} = u(\partial_y w - \partial_z v) + v(\partial_z u - \partial_x w) + w(\partial_x v - \partial_y u). \quad (48.83)$$

Although the helicity density vanishes for some common examples, such as for a fluid in solid-body rotation, it need not vanish in general.

- (a) Show that helicity is materially constant following the material volume

$$\frac{d\mathbb{H}}{dt} = 0. \quad (48.84)$$

- (b) Discuss why helicity is not defined for a shallow water fluid.

Make use of the following hints.

- Make use of Φ_p that satisfies equation (48.17).
- 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, only the vertical component of vorticity is nontrivial.

49

Angular momentum, vorticity, and strain

As noted in Section 46.5.4, fluid flow in the presence of a free vortex (Section 46.4) 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, solid-body fluid flow (Section 46.5) has a nonzero vorticity, nonzero angular momentum, yet a zero strain. In this chapter we detail the formal connection between vorticity, strain, and angular momentum in a fluid flow.

READER'S GUIDE TO THIS CHAPTER

We assume an understanding of vorticity as given in Chapter 48, and make use of Cartesian tensors as presented in Chapter 3. No subsequent chapter makes use of the results here. Rather, this brief chapter serves only to satisfy the curiosity of interested readers.

49.1	A resume of point particle mechanics	760
49.1.1	Angular velocity and moment of inertia	760
49.1.2	Relating angular velocity to velocity	761
49.2	Linear momentum for material fluid regions	761
49.3	Angular momentum for material fluid regions	762
49.3.1	Taylor expanding the velocity	763
49.3.2	Strain and vorticity	764
49.3.3	Relating angular momentum to strain and vorticity	765
49.3.4	Comments	766

49.1 A resume of point particle mechanics

Much of this section follows from our earlier treatment of particle mechanics in Part II. We revisit salient points to emphasize elements of interest for the present chapter.

The linear momentum of a point particle is written¹

$$\mathbf{P} = M \mathbf{V}, \quad (49.1)$$

where M is the particle's mass, which is a measure of the particle's inertia. The velocity, \mathbf{V} , is the time change of the particle position,

$$\mathbf{V} = \frac{d\mathbf{X}}{dt}. \quad (49.2)$$

The corresponding angular momentum is given by

$$\mathbf{L} = \mathbf{X} \wedge \mathbf{P} = M (\mathbf{X} \wedge \mathbf{V}). \quad (49.3)$$

The angular momentum is a function of the origin of the chosen coordinate system. The utility and relevance of angular momentum stems from its conservation for systems exhibiting rotational symmetry about special points or special directions. For example, motion on a smooth sphere exhibits rotational symmetry with respect to the center of the sphere. Consequently, all components of angular momentum for a particle are constant in the absence of externally applied torques. Likewise, for motion on a smooth rotating sphere, we showed in Section 15.7 that the component of angular momentum about the rotation axis is a constant of the motion.

49.1.1 Angular velocity and moment of inertia

Whereas linear momentum has physical dimensions of

$$[\mathbf{P}] \equiv \text{mass} \times \text{length} \times \text{time}^{-1}, \quad (49.4)$$

angular momentum has dimensions of

$$[\mathbf{L}] \equiv \text{mass} \times \text{length}^2 \times \text{time}^{-1}. \quad (49.5)$$

We can pursue the analog by introducing the angular velocity

$$\boldsymbol{\Omega} = \frac{\mathbf{X} \wedge \mathbf{V}}{|\mathbf{X}|^2}. \quad (49.6)$$

¹We use capital letters to accord with our usage of Lagrangian fluid particle trajectories.

The angular velocity has physical dimensions of inverse time, and it is defined with respect to the chosen coordinate origin. Furthermore, by construction the angular velocity vector is orthogonal to both the velocity and to the position

$$\boldsymbol{\Omega} \cdot \mathbf{X} = \boldsymbol{\Omega} \cdot \mathbf{V} = 0. \quad (49.7)$$

The angular velocity is not defined at the origin since $|\mathbf{X}| = 0$.

Inserting the definition of the angular velocity (49.6) into the angular momentum (49.3) renders

$$\mathbf{L} = M (\mathbf{X} \wedge \mathbf{V}) \quad (49.8a)$$

$$= M |\mathbf{X}|^2 \boldsymbol{\Omega} \quad (49.8b)$$

$$\equiv I \boldsymbol{\Omega}. \quad (49.8c)$$

In the final equality we introduced the moment of inertia for a point particle

$$I = M |\mathbf{X}|^2. \quad (49.9)$$

The moment of inertia measures the inertia appropriate for determining angular momentum relative to a chosen coordinate origin. The moment of inertia scalar, I , generalizes to the moment of inertia tensor, I_{mn} , when considering angular momentum for extended matter, such as a rigid body or a material fluid region (Section 49.3.3).

49.1.2 Relating angular velocity to velocity

The cross product of the position vector with the angular velocity (49.6) is given by

$$\boldsymbol{\Omega} \wedge \mathbf{X} = \mathbf{V} - \mathbf{X} \frac{\mathbf{V} \cdot \mathbf{X}}{|\mathbf{X}|^2}. \quad (49.10)$$

For the special case of velocity \mathbf{V} orthogonal to the position vector, \mathbf{X} , we have

$$\mathbf{V} = \boldsymbol{\Omega} \wedge \mathbf{X} \quad \text{when } \mathbf{V} \cdot \mathbf{X} = 0. \quad (49.11)$$

In particular, for circular motion the velocity is orthogonal to the position.

49.2 Linear momentum for material fluid regions

We now consider the velocity and linear momentum of a connected material fluid region denoted by $\mathcal{R}(\mathbf{v})$, with each point of the region moving with the local fluid velocity $\mathbf{v}(\mathbf{x}, t)$. Let an arbitrary fluid parcel 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

$$\mathbf{V}(\mathbf{a}, t) = \frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t}, \quad (49.12)$$

where the time derivative is computed holding the material label fixed. Since the parcel is within a finite material region, we find it useful to decompose the motion of the parcel into the motion of the region's center of mass plus the motion of the parcel relative to the center of mass

$$\mathbf{V}(\mathbf{a}, t) = \frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \quad (49.13a)$$

$$= \frac{\partial(\bar{\mathbf{X}} + \mathbf{X}')}{\partial t} \quad (49.13b)$$

$$= \bar{\mathbf{V}}(t) + \mathbf{V}'(\mathbf{a}, t). \quad (49.13c)$$

In this equation, we introduced the velocity \mathbf{V}' defined relative to the center of mass of the region. Furthermore, the center of mass velocity is given by

$$\bar{\mathbf{V}} = \frac{d\bar{\mathbf{X}}}{dt} \quad (49.14a)$$

$$= \frac{d}{dt} \left[\frac{\int_{\mathcal{R}(\mathbf{v})} \mathbf{x} \rho dV}{\int_{\mathcal{R}(\mathbf{v})} \rho dV} \right] \quad (49.14b)$$

$$= \frac{1}{M} \int_{\mathcal{R}(\mathbf{v})} \frac{D\mathbf{x}}{Dt} \rho dV \quad (49.14c)$$

$$= \frac{1}{M} \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV. \quad (49.14d)$$

The identity (49.14c) follows since the material region maintains a constant mass,

$$M = \int_{\mathcal{R}(\mathbf{v})} \rho dV \implies \frac{dM}{dt} = 0, \quad (49.15)$$

allowing the denominator to come outside the derivative. Additionally, each of the fluid parcels in the region maintains constant mass. As per Reynold's transport theorem (Section 17.3.4), the time derivative moves across the integral to act materially on the position vector. The final equality, (49.14d), follows since the material time derivative of a parcel trajectory when evaluated at a point, \mathbf{x} , equals to the velocity field at that point

$$\mathbf{v}(\mathbf{x}, t) = \frac{D\mathbf{x}}{Dt}. \quad (49.16)$$

It follows that the linear momentum for the material fluid region is given by

$$\mathbf{P} = \int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV = M \bar{\mathbf{V}}. \quad (49.17)$$

We conclude that the total linear momentum of an extended body equals to that of a point particle of mass $M = \int_{\mathcal{R}(\mathbf{v})} \rho dV$ moving with the center of mass velocity, $\bar{\mathbf{V}}$.

49.3 Angular momentum for material fluid regions

We here consider angular momentum for a material fluid region, which is determined by the integral over that region of the angular momentum for each fluid parcel

$$\mathbf{L} = \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x} \wedge \mathbf{v}) \rho dV. \quad (49.18)$$

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

$$\mathbf{L} = \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x} \wedge \mathbf{v}) \rho dV \quad (49.19a)$$

$$= \int_{\mathcal{R}(\mathbf{v})} [(\bar{\mathbf{x}} + \mathbf{x}') \wedge \mathbf{v}] \rho dV \quad (49.19b)$$

$$= \bar{\mathbf{X}} \wedge \left[\int_{\mathcal{R}(\mathbf{v})} \mathbf{v} \rho dV \right] + \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \wedge \mathbf{v}) \rho dV \quad (49.19c)$$

$$= (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \wedge \mathbf{v}) \rho dV. \quad (49.19d)$$

The final equality introduced the linear momentum, (49.17), for the fluid region. The first term in equation (49.19d) is the angular momentum of the region with respect to the position of the center of mass. The second term is associated with deviations of parcel positions relative to the center of mass.

We now focus on how the deviation term, $\int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \wedge \mathbf{v}) \rho dV$, contributes to the angular momentum (49.18). As we will see, this analysis exposes how angular momentum of the extended material fluid region is affected by vorticity and strain in the fluid flow. To facilitate some of the manipulations, we make use of basic Cartesian tensor analysis from Chapter 3, including the summation convention whereby repeated indices are summed over their range. Additionally, we introduce components to the totally anti-symmetric Levi-Civita tensor, ϵ_{mnp} so that the vector cross product is written (see Section 3.4)

$$(\mathbf{A} \wedge \mathbf{B})_m = \epsilon_{mnp} A_n B_p. \quad (49.20)$$

49.3.1 Taylor expanding the velocity

We now 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 term²

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}(\bar{\mathbf{x}} + \mathbf{x}') \approx \mathbf{v}(\bar{\mathbf{x}}) + (\mathbf{x}' \cdot \nabla) \mathbf{v}|_{\mathbf{x}=\bar{\mathbf{x}}}. \quad (49.21)$$

We are thus left with

$$\mathbf{L} = (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \wedge \mathbf{v}) \rho dV \quad (49.22a)$$

$$= (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \wedge \mathbf{v}(\bar{\mathbf{x}})] \rho dV + \int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \wedge (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})] \rho dV. \quad (49.22b)$$

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 (49.22b) vanishes

$$\int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \wedge \mathbf{v}(\bar{\mathbf{x}})] \rho dV = \left[\int_{\mathcal{R}(\mathbf{v})} \mathbf{x}' \rho dV \right] \wedge \mathbf{v}(\bar{\mathbf{x}}) = 0, \quad (49.23)$$

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

$$\mathbf{L} = (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \wedge (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})] \rho dV. \quad (49.24)$$

²The velocity field evaluated at the center of mass position, $\mathbf{v}(\bar{\mathbf{x}})$, is not equal to the center of mass velocity: $\mathbf{v}(\bar{\mathbf{x}}) \neq \bar{\mathbf{v}}$.

The m' th component of the second term can be written

$$\int_{\mathcal{R}(\mathbf{v})} [\mathbf{x}' \wedge (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})]_m \rho dV = \epsilon_{mnp} \int_{\mathcal{R}(\mathbf{v})} x'_n [(\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})]_p \rho dV \quad (49.25a)$$

$$= \epsilon_{mnp} \int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \partial_q v(\bar{\mathbf{x}})_p \rho dV \quad (49.25b)$$

$$= \epsilon_{mnp} \left[\int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \partial_q v(\bar{\mathbf{x}})_p. \quad (49.25c)$$

We removed the velocity derivatives

$$\partial_q v(\bar{\mathbf{x}})_p = \left[\frac{\partial v_p}{\partial x_q} \right]_{\mathbf{x}=\bar{\mathbf{x}}} \quad (49.26)$$

from the integral, since they are evaluated at the center of mass point and so do not participate in the integration.

49.3.2 Strain and vorticity

Following from the discussion in Section 20.2.4, we know that the velocity derivatives $\partial_q v_p$ appearing in equation (49.25c) form the components to a second order tensor. To expose the kinematics of this tensor, decompose it into its symmetric and anti-symmetric components

$$\partial_q v_p = \frac{1}{2}(\partial_q v_p + \partial_p v_q) + \frac{1}{2}(\partial_q v_p - \partial_p v_q) \quad (49.27a)$$

$$\equiv S_{qp} + A_{qp}. \quad (49.27b)$$

The symmetric tensor

$$S_{qp} = \frac{1}{2}(\partial_q v_p + \partial_p v_q) \quad (49.28)$$

is associated with deformations in the fluid arising from strains; it is therefore called the *deformation* or *rate of strain* tensor. The anti-symmetric tensor can be written as

$$2A_{qp} = \partial_q v_p - \partial_p v_q \quad (49.29a)$$

$$= (\delta_{qm} \delta_{pn} - \delta_{qn} \delta_{pm}) \partial_m v_n \quad (49.29b)$$

$$= \epsilon_{sqp} \epsilon_{smn} \partial_m v_n \quad (49.29c)$$

$$= \epsilon_{sqp} \omega_s, \quad (49.29d)$$

where $\omega_s = \epsilon_{smn} \partial_m v_n$ are components to the vorticity pseudo-vector

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v}. \quad (49.30)$$

49.3.3 Relating angular momentum to strain and vorticity

Making use of the strain and vorticity brings the angular momentum for a connected material fluid into the form

$$L_m = (\bar{\mathbf{X}} \wedge \mathbf{P})_m + \epsilon_{mnp} \left[\int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] S_{qp} + \epsilon_{mnp} \left[\int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] A_{qp} \quad (49.31a)$$

$$= (\bar{\mathbf{X}} \wedge \mathbf{P})_m + \epsilon_{mnp} \left[\int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] S_{qp} + \frac{1}{2} \epsilon_{mnp} \epsilon_{sqp} \left[\int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] \omega_s \quad (49.31b)$$

$$= \underbrace{(\bar{\mathbf{X}} \wedge \mathbf{P})_m}_{\text{center of mass}} + \underbrace{\epsilon_{mnp} \left[\int_{\mathcal{R}(\mathbf{v})} x'_n x'_q \rho dV \right] S_{qp}}_{\text{strain contribution}} + \underbrace{\frac{1}{2} \left[\int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \cdot \mathbf{x}' \delta_{ms} - x'_m x'_s) \rho dV \right] \omega_s}_{\text{vorticity contribution}}. \quad (49.31c)$$

As each point in the fluid can be considered the center of mass for an arbitrary material region, the decomposition (49.31c) is general.

- **CENTER OF MASS ANGULAR MOMENTUM:** The first term on the right hand side of equation (49.31c) arises from the angular momentum of the material region as measured with respect to the center of mass position. It has the form of that for a point particle (see equation (49.8a)). This term vanishes if the origin of the coordinate system is taken at the center of mass.
- **STRAINS:** The second contribution is proportional to fluid deformations acting to dilate or strain the fluid region (see Section 20.2.4). At each point of the fluid, deformations are measured by the deformation tensor S_{qp} . A rigid body moves by uniform translations and/or solid-body rotations, with the deformation tensor vanishing for rigid body motions (see Section 46.5). It is for this reason that the deformation tensor is so-named, as this tensor measures motions that are deviations or deformations relative to the motion of a rigid body. The contribution from these deformations is weighted by an integral of deviations of parcel position from the center of mass position. A closed form expression for this integral is available only for special shapes.
- **VORTICITY:** The third contributor to angular momentum in equation (49.31c) contains the vorticity as weighted by the moment of inertia tensor

$$I_{ms} \equiv \int_{\mathcal{R}(\mathbf{v})} (\mathbf{x}' \cdot \mathbf{x}' \delta_{ms} - x'_m x'_s) \rho dV. \quad (49.32)$$

Since the material region is evolving and is not rigid, the moment of inertia tensor is time dependent. The contribution

$$L_m^{\text{vorticity}} \equiv \frac{1}{2} I_{ms} \omega_s \quad (49.33)$$

has the same form as angular momentum for a rigid body, with one-half the vorticity playing the role of angular velocity (see equation (49.8c) for the point particle expression). Fluid vorticity hence contributes to angular momentum for a material region via its product with the moment of inertia tensor.

49.3.4 Comments

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 quite distinct, especially when the fluid has non-zero strains. It is only for the special case of a solid-body motion that the strain contribution to angular momentum vanishes.

The discussion in this chapter supplements that from [Chatwin \(1973\)](#), as well as online notes “The Vorticity Equation and Conservation of Angular Momentum” from A.J. DeCaria.

50

Potential vorticity mechanics

The chapter details the foundational properties of potential vorticity (PV) and its evolution. The PV we encounter is sometimes referred to as the *Ertel* PV, which is the most fundamental form of PV arising in geophysical fluids ([Ertel, 1942](#)). The barotropic fluid forms a pedagogically useful starting point for the discussion. However, realistic geophysical flows are baroclinic, which are fluids where “PV thinking” is most powerful. The general method (“trick”) exploited for the construction of PV is to choose a scalar field to strategically orient the vorticity. If the scalar is a material invariant, and it annihilates the baroclinicity vector, then PV is a material invariant. Notably, as so constructed, PV is a function of the chosen scalar.

For a barotropic fluid, the choice of scalar field is rather arbitrary, with preference given to one that is materially invariant. However, for a baroclinic fluid we are much more constrained, since the scalar must orient vorticity in a direction to annihilate the torque from baroclinicity and, ideally, be itself materially invariant. We can only annihilate baroclinicity under certain restricted cases. Nonetheless, even when PV is not materially invariant, it remains a very important fluid property that partially constrains the fluid motion and allows for insights into the interpretation of that motion.

READER’S GUIDE FOR THIS CHAPTER

This chapter requires a firm understanding of vorticity from Chapter 48 as well as good skills with vector calculus identities for Cartesian coordinates as detailed in Chapter 4. The concepts and methods developed in this chapter are fundamental to the notions of potential vorticity, much of which is encountered in the remainder of this part of the book as well as when studying balanced models in Part IX.

50.1	Material invariance of PV in perfect fluids	768
50.1.1	Perfect barotropic fluid	768
50.1.2	Region between two scalar isosurfaces	768
50.1.3	Material invariance	769
50.1.4	Perfect baroclinic fluid	770
50.1.5	Some general remarks	771
50.2	PV and the seawater equation of state	772
50.2.1	Baroclinicity vector	772
50.2.2	PV based on potential density	772
50.2.3	An example EOS admitting a materially invariant PV	773
50.2.4	Further reading	774
50.3	PV evolution with non-conservative processes	774
50.4	Eulerian flux-form PV budget	775
50.4.1	Deriving the Eulerian flux-form PV budget	776
50.4.2	PV-substance and the PV flux	776
50.5	Exercises	777

50.1 Material invariance of PV in perfect fluids

In this section we derive the material invariance of potential vorticity (PV) for a perfect fluid (i.e., fluid without mixing). We make use of Kelvin's circulation theorem for an infinitesimal loop, in which case the primary object of interest is a particular component of the absolute vorticity. The discussion starts with a barotropic fluid, in which baroclinicity vanishes (Sections 48.2 and 48.3), and then we generalize to a baroclinic fluid.

50.1.1 Perfect barotropic fluid

Consider a perfect barotropic fluid. As for the shallow water discussion in Section 47.3.3, we apply Kelvin's circulation theorem (Section 48.2.2) to an infinitesimal material circuit within the fluid

$$\frac{D}{Dt}(\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \delta\mathcal{S}) = 0, \quad (50.1)$$

with $\delta\mathcal{S}$ the area of the circuit. The conservation of PV is built from specializing this result. For that purpose, introduce a materially invariant scalar field

$$\frac{DC}{Dt} = 0. \quad (50.2)$$

50.1.2 Region between two scalar isosurfaces

We make use of isosurfaces of C to orient the material circuit and hence to orient the vorticity. In particular, referring to Figure 50.1, let the circuit bound a small cylinder whose two ends sit on two isosurfaces, $C - \delta C/2$ and $C + \delta C/2$. The tube volume is given by

$$\delta V = \delta\mathcal{S} \delta h, \quad (50.3)$$

where δh is the distance between the isosurfaces. The unit normal direction orienting the area $\delta\mathcal{S}$ is given by

$$\hat{\mathbf{n}} = \frac{\nabla C}{|\nabla C|}. \quad (50.4)$$

The separation between the two isosurfaces is related to the increment δC through

$$\delta C = \nabla C \cdot \delta \mathbf{x} = |\nabla C| \hat{\mathbf{n}} \cdot \delta \mathbf{x} = |\nabla C| \delta h. \quad (50.5)$$

Note that we can write this result in the equivalent manner

$$\delta C = |\nabla C| \delta h = (\hat{\mathbf{n}} \cdot \nabla C) \delta h = \frac{\delta C}{\delta n} \delta h. \quad (50.6)$$

Hence, the distance (or thickness) between the two isosurfaces is

$$\delta h = \frac{\delta C}{|\nabla C|}. \quad (50.7)$$

This equation has a straightforward geometric interpretation indicated in Figure 50.1. Namely, the geometric separation between the two isosurfaces is reduced in regions of strong scalar gradients and increased in regions of weak gradients.

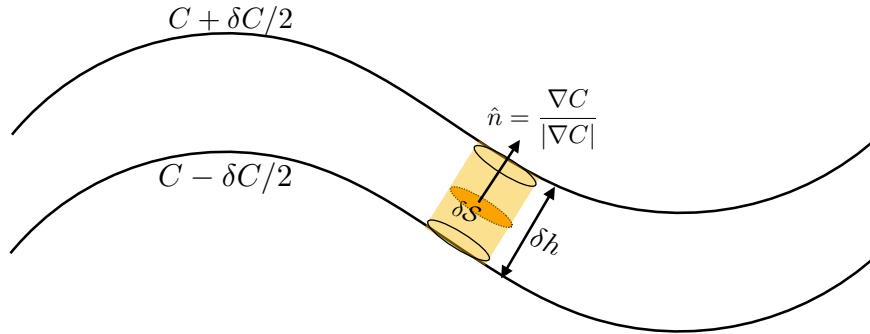


Figure 50.1: Illustrating the geometry of a cylindrical region of fluid between two iso-surfaces of a scalar field C . The volume of the region is $\delta V = \delta S \delta h$, with δh the thickness and δS the area. By convention, the normal vector, $\hat{\mathbf{n}}$, points towards larger values of C . Notably, if C is a material invariant so that $DC/Dt = 0$, then so is its infinitesimal increment, $D(\delta C)/Dt = 0$. As per equation (50.7), the geometric thickness between the isosurfaces is related to the scalar field increment by $\delta h = \delta C/|\nabla C|$, so that the stronger the gradient in the scalar field the smaller the layer thickness. For a baroclinic fluid material invariance of PV holds only if we can find a scalar field such that $\hat{\mathbf{n}} \cdot \mathbf{B} = 0$, with $\mathbf{B} = (\nabla \rho \wedge \nabla p)/\rho^2$ the baroclinicity.

50.1.3 Material invariance

We now have the necessary pieces in place to write

$$\omega_a \cdot \hat{\mathbf{n}} \delta S = \frac{\omega_a \cdot \nabla C}{|\nabla C|} \delta S \quad (50.8a)$$

$$= \frac{\omega_a \cdot \nabla C}{|\nabla C|} \frac{\delta V}{\delta h} \quad (50.8b)$$

$$= (\omega_a \cdot \nabla C) \frac{\delta V}{\delta C} \quad (50.8c)$$

$$= \frac{\omega_a \cdot \nabla C}{\rho} \frac{\rho \delta V}{\delta C}. \quad (50.8d)$$

Mass is materially invariant so that

$$\frac{D(\rho \delta V)}{Dt} = 0. \quad (50.9)$$

Likewise, by assumption C is materially invariant so that the increment between two C isosurfaces is materially invariant

$$\frac{D(\delta C)}{Dt} = 0. \quad (50.10)$$

Bringing these elements into Kelvin's circulation theorem (50.1) leads us to conclude that the potential vorticity, Q , is also materially invariant

$$Q = \frac{\boldsymbol{\omega}_a \cdot \nabla C}{\rho} = \frac{\nabla \cdot (\boldsymbol{\omega}_a C)}{\rho} \implies \frac{DQ}{Dt} = 0. \quad (50.11)$$

This expression for the PV is the most general form and it is often referred to as the *Ertel PV* ([Ertel, 1942](#)). The first expression shows the numerator as the projection of the absolute vorticity into the direction normal to tracer isosurfaces. Conversely, it is a measure of the C 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. We make use of this second form when discussing PV budgets in Section 50.4 and layer integrated PV Section 51.3.

50.1.4 Perfect baroclinic fluid

Consider the case of a perfect baroclinic fluid, in which Kelvin's circulation theorem for an infinitesimal circuit takes the form

$$\frac{D}{Dt}(\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \delta S) = \mathbf{B} \cdot \hat{\mathbf{n}} \delta S, \quad (50.12)$$

where the source on the right hand side arises from the baroclinicity vector discussed in Sections 48.2 and 48.3.

$$\mathbf{B} = \frac{\nabla \rho \wedge \nabla p}{\rho^2}. \quad (50.13)$$

Now assume there exists a materially invariant scalar, θ , that annihilates the baroclinicity vector as in Figure 50.2, so that

$$\mathbf{B} \cdot \hat{\mathbf{n}} = \frac{\mathbf{B} \cdot \nabla \theta}{|\nabla \theta|} = 0. \quad (50.14)$$

In that case, the barotropic derivation detailed earlier in this section follows for the baroclinic case so that PV remains materially invariant

$$\frac{DQ}{Dt} = 0 \quad \text{where} \quad Q = \frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{\rho}. \quad (50.15)$$

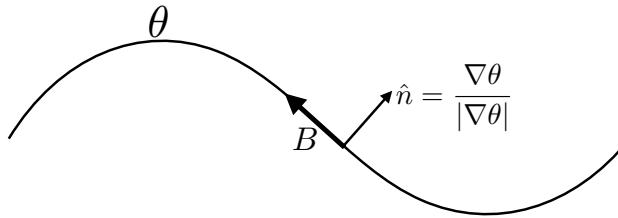


Figure 50.2: Material invariance of PV is ensured for perfect fluids that allow for a materially invariant scalar to annihilate baroclinicity, meaning that $\hat{\mathbf{n}} \cdot \mathbf{B} = 0$ where $\hat{\mathbf{n}} = \nabla \theta / |\nabla \theta|$. In this figure we depict the potential temperature field with a baroclinicity vector aligned with isolines of θ .

Existence of a materially invariant PV for perfect baroclinic fluids depends on the existence of a materially invariant scalar that annihilates the baroclinicity vector. The most common choice for

this scalar in geophysical fluid applications is the buoyancy, which is typically measured by entropy or potential temperature in the atmosphere and potential density in the ocean. We have more to say on the chosen scalar field in the remainder of this chapter as well as in Section 51.2.

50.1.5 Some general remarks

Perfect fluid PV material invariance \leftrightarrow Kelvin's circulation theorem

Kelvin's circulation theorem from Section 48.2.2 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 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 scalar field. For the perfect barotropic fluid, we merely require that both mass and scalar to be materially invariant to realize a potential vorticity that is also materially invariant. Material invariance of PV for a baroclinic fluid requires a scalar field that is both materially invariant *and* that annihilates the baroclinicity vector. We have more to say in regards to the availability of such scalar fields when considering entropic PV in the remainder of this chapter, whereby potential temperature is the chosen scalar field.

There are numerous forms for PV

The expression (50.11) is, on first glance, quite distinct from the shallow water expression $Q = (\zeta + f)/h$ explored in Chapter 47 (see equation (47.26)). However, as shown in Section 52.2, they are closely related for the special case of entropic PV in an incompressible fluid when formulated using isopycnal/isentropic coordinates. Even so, there are a variety of other forms for PV, with the forms (and physical dimensions) depending on the assumptions made in regards to the fluid flow and its thermodynamic properties. We encounter some further forms of PV in the remainder of this chapter, as well as in the oceanic PV discussions of Chapter 52 and in our study of balanced models in Part IX. The review paper by Müller (1995) offers a lucid presentation of PV and its many forms from physical oceanography.

Motivating the adjective "potential"

Write the PV in the form

$$Q = \frac{\omega_a \cdot \hat{n}}{\rho} |\nabla\theta| \quad \text{with} \quad \hat{n} = \frac{\nabla\theta}{|\nabla\theta|}. \quad (50.16)$$

As noted in Section 2.1 of Müller (1995), in cases where ρ is roughly a constant, and when Q is materially invariant, the component of the absolute vorticity in the direction parallel to $\nabla\theta$ increases when the fluid parcel moves into a region where $|\nabla\theta|$ decreases. There is hence a "release" of absolute vorticity aligned with $\nabla\theta$ in regions where θ isosurfaces are spread, and a withdrawal of absolute vorticity where θ isosurfaces are tightly packed. We thus conceive of PV as the "potential" for releasing absolute vorticity oriented in the direction parallel to $\nabla\theta$.

PV as a dynamical tracer

We refer to PV as a *dynamical tracer* since it depends directly on the velocity field (as well as the thermodynamic tracer, θ). In contrast, material tracers (e.g., salinity) and thermodynamic tracers

(e.g., θ), are properties of the fluid and they do not directly depend on velocity, even though they are affected by velocity through advection.

PV is the grand unifier for geophysical fluid mechanics

Entropic PV provides a connection between vorticity (mechanics) and stratification (thermodynamics). By connecting these two basic facets of geophysical fluid flows, the study of PV 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 PV is sometimes considered the grand unifying concept in geophysical fluid mechanics.

50.2 PV and the seawater equation of state

As seen in Section 50.1, material invariance of PV for a perfect fluid requires a materially invariant scalar field to annihilate the baroclinicity vector. In the ocean there is generally no such scalar when there is also a realistic seawater equation of state. Hence, the ocean cannot be considered a perfect fluid even in the absence of friction and diabatic processes. Nonetheless, as discussed in this section, there are important approximate cases that allow for material PV invariance.

50.2.1 Baroclinicity vector

Recall the baroclinicity vector given by (Sections 48.2 and 48.3)

$$\mathbf{B} = \frac{\nabla\rho \wedge \nabla p}{\rho^2}. \quad (50.17)$$

If we take the *in situ* density as the scalar field to define PV, then $\mathbf{B} \cdot \nabla\rho = 0$. However, *in situ* density is not a conserved scalar in the ocean due to pressure effects. Namely, the material time derivative of $\rho = \rho(S, \theta, p)$ is

$$\frac{D\rho}{Dt} = \frac{\partial\rho}{\partial S} \frac{DS}{Dt} + \frac{\partial\rho}{\partial\theta} \frac{D\theta}{Dt} + \frac{\partial\rho}{\partial p} \frac{Dp}{Dt}. \quad (50.18)$$

Even when salinity and potential temperature are materially constant, $DS/Dt = 0$ and $D\theta/Dt = 0$, the *in situ* density has a nonzero material time evolution due to mechanical effects leading to pressure changes, $Dp/Dt \neq 0$. Hence, *in situ* density is not an appropriate scalar for use in developing a materially invariant PV.

50.2.2 PV based on potential density

Potential density is commonly used in oceanography (see Section 25.2.5), with potential density the *in situ* density referenced to a chosen pressure.¹ We write potential density as

$$\sigma(S, \theta) = \rho(S, \theta, p = p_R). \quad (50.19)$$

Potential density provides a global measure of buoyancy, with accuracy of the measure compromised in the presence of a nonlinear equation of state (Section 25.5).

¹Oceanographers often choose the reference pressure as the standard atmospheric surface pressure. However, that is not required for the following formalism to hold; any reference pressure is suitable.

The material time derivative of potential density is

$$\frac{D\sigma}{Dt} = \frac{\partial\sigma}{\partial S} \frac{DS}{Dt} + \frac{\partial\sigma}{\partial\theta} \frac{D\theta}{Dt}, \quad (50.20)$$

which vanishes in the absence of irreversible material changes to salinity and potential temperature. When using potential density as the scalar field for PV, the baroclinicity vector projects onto the diapycnal direction according to

$$\rho^2 \mathbf{B} \cdot \nabla\sigma = (\nabla\rho \wedge \nabla p) \cdot \nabla\sigma \quad (50.21a)$$

$$= (\nabla\sigma \wedge \nabla\rho) \cdot \nabla p \quad (50.21b)$$

$$= [(\sigma_S \nabla S + \sigma_\theta \nabla\theta) \wedge (\rho_S \nabla S + \rho_\theta \nabla\theta + \rho_p \nabla p)] \cdot \nabla p \quad (50.21c)$$

$$= [(\sigma_S \nabla S + \sigma_\theta \nabla\theta) \wedge (\rho_S \nabla S + \rho_\theta \nabla\theta)] \cdot \nabla p \quad (50.21d)$$

$$= [\sigma_S \nabla S \wedge \rho_\theta \nabla\theta + \sigma_\theta \nabla\theta \wedge \rho_S \nabla S] \cdot \nabla p \quad (50.21e)$$

$$= (\sigma_S \rho_\theta - \sigma_\theta \rho_S) (\nabla S \wedge \nabla\theta) \cdot \nabla p, \quad (50.21f)$$

where we used the shorthand notation for partial derivatives

$$\rho_S = \frac{\partial\rho}{\partial S} \quad \sigma_S = \frac{\partial\sigma}{\partial S} \quad (50.22a)$$

$$\rho_\theta = \frac{\partial\rho}{\partial\theta} \quad \sigma_\theta = \frac{\partial\sigma}{\partial\theta}. \quad (50.22b)$$

Note that the triple product, $(\nabla S \wedge \nabla\theta) \cdot \nabla p$, also appears in the discussion of neutral helicity in Section 25.6 (see equation (25.57)). Equation (50.21f) allows us to identify cases where the baroclinicity vector is annihilated, $\mathbf{B} \cdot \nabla\sigma = 0$, thus yielding a materially invariant PV in the absence of irreversible processes.

- **UNIFORM SALINITY OR UNIFORM POTENTIAL TEMPERATURE:** If salinity or potential temperature are spatially uniform, then $\mathbf{B} \cdot \nabla\sigma = 0$.
- **ADDITIVE PRESSURE DEPENDENCE TO THE *in situ* DENSITY:** There is an exact PV conservation principle if the thermodynamic pre-factor $\sigma_S \rho_\theta - \sigma_\theta \rho_S$ vanishes. However, the ocean has a pressure dependent equation of state and this pressure dependence generally means that $\mathbf{B} \cdot \nabla\sigma \neq 0$. Nonetheless, the baroclinicity vector is annihilated if the *in situ* density has a pressure dependence that is additive, so that we can write

$$\rho(S, \theta, p) = \sigma(S, \theta) + F(p) - F(p_R) \Rightarrow \sigma_S \rho_\theta - \sigma_\theta \rho_S = 0, \quad (50.23)$$

which then leads to a materially invariant PV. Notably, we did not assume the equation of state to be linear; only that it has the special functional form in equation (50.23). For some cases, we may assume F to be a constant, in which case there is no pressure dependence; i.e., the *in situ* density is the same as potential density.

50.2.3 An example EOS admitting a materially invariant PV

An explicit realization of the equation of state (50.23) 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 potential temperature

$$\rho(S, \theta, p) \approx \sigma(S, \theta) + (p - p_R) \underbrace{\left[\frac{\partial \rho}{\partial p} \right]_{S=S_R, \theta=\theta_R, p=p_R}}_{F(p) - F(p_R)} + H.O.T. \quad (50.24)$$

where *H.O.T.* symbolizes higher order terms. This approach ignores the salinity and potential 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

$$c_s^{-2} = \frac{\partial \rho}{\partial p} \approx \text{constant}. \quad (50.25)$$

In this case, the equation of state takes the form

$$\rho(S, \theta, p) \approx \sigma(S, \theta) + \frac{p - p_R}{c_s^2}, \quad (50.26)$$

where again $\sigma(S, \theta) = \rho(S, \theta, p_R)$ is the potential density referenced to $p = p_R$.

50.2.4 Further reading

[Straub \(1999\)](#) presents a discussion of ocean potential vorticity with a focus on the source of potential vorticity arising from a nonzero thermobaricity parameter, $\mathcal{T} = \partial_p(\alpha/\beta)$ (see Section 39.4.4).

50.3 PV evolution with non-conservative processes

Thus far we have considered perfect fluids, with the use of Kelvin's circulation theorem a suitable framework to derive the material invariance of PV. In this section we move beyond the perfect fluid by considering a real fluid that contains non-conservative processes. PV is no longer materially invariant when exposed to non-conservative processes. However, as detailed in Chapter 51, there remain some rather remarkable global conservation properties for PV budgets within isentropic layers.

To develop the PV budget in the presence of non-conservative processes, we find it more convenient to pursue an algebraic approach that starts from the vorticity equation (48.32)

$$\rho \frac{D(\boldsymbol{\omega}_a / \rho)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} + \mathbf{B} + \nabla \wedge \mathbf{F}, \quad (50.27)$$

where \mathbf{F} is the frictional acceleration vector and \mathbf{B} the baroclinicity vector. Furthermore, we introduce a non-conservative scalar field

$$\frac{DC}{Dt} = \dot{C}, \quad (50.28)$$

with \dot{C} arising from diffusion or other irreversible processes.

As part of the manipulations we make use of the identity

$$(\boldsymbol{\omega}_a \cdot \nabla) \frac{DC}{Dt} = \boldsymbol{\omega}_a \cdot \frac{D(\nabla C)}{Dt} + [(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}] \cdot \nabla C, \quad (50.29)$$

which is readily proven by expanding terms and assuming Cartesian coordinates. Rearrangement, and use of the scalar equation (50.28), leads to

$$\boldsymbol{\omega}_a \cdot \frac{D(\nabla C)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{C} - [(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}] \cdot \nabla C. \quad (50.30)$$

Now project the vorticity equation (50.27) onto the direction normal to the C isosurfaces

$$\rho \nabla C \cdot \frac{D(\boldsymbol{\omega}_a/\rho)}{Dt} = \nabla C \cdot [(\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v}] + \nabla C \cdot (\mathbf{B} + \nabla \wedge \mathbf{F}). \quad (50.31)$$

The sum of equations (50.30) and (50.31) leads to

$$\rho \frac{D(\nabla C \cdot \boldsymbol{\omega}_a/\rho)}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{C} + \nabla C \cdot (\mathbf{B} + \nabla \wedge \mathbf{F}). \quad (50.32)$$

This equation is general, and thus applies to any scalar field.

To simplify the source terms on the right hand side of equation (50.32), follow the discussion from Section 50.1.4 by assuming we can find a special scalar field that annihilates the baroclinicity vector.² We generally denote this scalar field as θ , which we use in this chapter as a general designation for potential temperature, specific entropy, or potential density. We thus have

$$\nabla \theta \cdot \mathbf{B} = 0, \quad (50.33)$$

which in turn leads to the potential vorticity equation in the presence of irreversible processes

$$\rho \frac{DQ}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla) \dot{\theta} + \nabla \theta \cdot (\nabla \wedge \mathbf{F}), \quad (50.34)$$

where the entropic-PV is again given by

$$Q = \frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{\rho}. \quad (50.35)$$

As defined, the material evolution of PV is affected by diabatic processes (heating and cooling) as well as friction. Absent these irreversible processes leaves us with the same PV conservation statement (50.15) derived for the perfect fluid using Kelvin's circulation theorem. In their presence, the PV of a fluid particle can be either generated or destroyed depending on details of the irreversible process. Such source/sink regions of potential vorticity are often localized to regions of mixing as well as to boundaries where strong mechanical and/or buoyancy processes are active. The study of how PV is materially modified by irreversible processes forms an important area of research in PV dynamics.

50.4 Eulerian flux-form PV budget

The material invariance of PV is an example of a material or Lagrangian conservation property of perfect fluids, with the material conservation statement $\rho DQ/Dt = 0$ having its Eulerian flux-form expression

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot (\rho \mathbf{v} Q) = 0 \quad \text{perfect fluid.} \quad (50.36)$$

²In Section 51.2 we study what happens when no such scalar exists.

Following the formalism established for material tracers in Section 18.3, the flux form local conservation law (50.36) leads then to conservation properties over finite regions, which we refer to as *global conservation* laws. In this section we examine the Eulerian flux-form budget in the presence of non-conservative processes. As we will see, the Eulerian evolution of PV will continue to be determined by the convergence of a flux, thus allowing for natural extensions to global conservation laws.

50.4.1 Deriving the Eulerian flux-form PV budget

To transform the material evolution equation (50.34) into a flux-form Eulerian equation we make use of the following identities

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad \text{relating material and Eulerian time changes} \quad (50.37a)$$

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v} \quad \text{mass conservation} \quad (50.37b)$$

$$(\boldsymbol{\omega}_a \cdot \nabla) \dot{\theta} = \nabla \cdot (\boldsymbol{\omega}_a \dot{\theta}) \quad \nabla \cdot \boldsymbol{\omega}_a = 0 \quad (50.37c)$$

$$\nabla \theta \cdot (\nabla \wedge \mathbf{F}) = \nabla \cdot (\mathbf{F} \wedge \nabla \theta) \quad \text{divergence of curl vanishes.} \quad (50.37d)$$

The identity (50.37d) follows from

$$\nabla \theta \cdot (\nabla \wedge \mathbf{F}) = \nabla \cdot (\theta \nabla \wedge \mathbf{F}) = \nabla \cdot [\nabla \wedge (\theta \mathbf{F}) - \nabla \theta \wedge \mathbf{F}] = \nabla \cdot (\mathbf{F} \wedge \nabla \theta), \quad (50.38)$$

where we set the divergence of a curl to zero to reach the first and third equalities. These identities then lead to the material evolution equation

$$\rho \frac{DQ}{Dt} = \nabla \cdot [\boldsymbol{\omega}_a \dot{\theta} + \mathbf{F} \wedge \nabla \theta]. \quad (50.39)$$

Now converting the material time derivative into its Eulerian expression, and making use of mass conservation, renders the flux-form Eulerian PV budget equation

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot [\rho Q \mathbf{v} - \boldsymbol{\omega}_a \dot{\theta} - \mathbf{F} \wedge \nabla \theta] = 0. \quad (50.40)$$

50.4.2 PV-substance and the PV flux

The budget equation (50.40) says that the density-weighted potential vorticity

$$\rho Q = \boldsymbol{\omega}_a \cdot \nabla \theta \quad (50.41)$$

has a local time tendency determined by the convergence of the PV flux vector

$$\frac{\partial(\rho Q)}{\partial t} = -\nabla \cdot \mathbf{J}_Q \quad \text{with} \quad \mathbf{J}_Q = \rho Q \mathbf{v} - \boldsymbol{\omega}_a \dot{\theta} - \mathbf{F} \wedge \nabla \theta. \quad (50.42)$$

The budget (50.40) follows a form similar to material tracers detailed in Chapter 18, though with some specific terms in the flux vector, \mathbf{J}_Q . The correspondence suggests that one consider equation (50.40) as the local budget for *PV-substance*, with Q the concentration of PV-substance and \mathbf{J}_Q its flux. This interpretation is pursued further in Chapter 51 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 (50.42) arises from the advection of PV-substance; the second contribution arises from diabatic processes; and the third from the curl of the friction

vector. Note that there is a nonzero friction contribution only when the frictional acceleration has a component that is not parallel to $\nabla\theta$. In this manner we can think of the friction vector as contributing to a “torque” that rotates the θ isosurfaces as it contributes to the evolution of PV-substance (see Figure 50.3).

The time tendency for PV-substance, $\partial_t(\rho Q)$, is unchanged by adding the curl of a vector to \mathbf{J}_Q . This ambiguity manifests a *gauge freedom* afforded the PV-substance flux vector. We offer a general discussion of gauge freedom in Section 19.5.2 and provide more specifics for PV in Chapter 51.

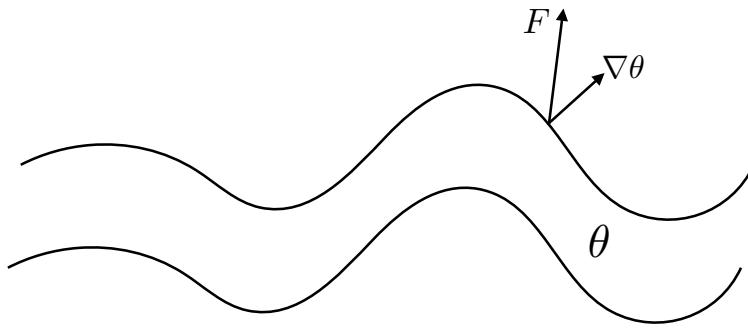


Figure 50.3: The contribution from friction to the PV flux is given by $\mathbf{J}_{\text{friction}} = \nabla\theta \wedge \mathbf{F}$, which is nonzero only when the frictional acceleration is not fully aligned with $\nabla\theta$. Friction creates PV by acting to rotate θ isosurfaces, so if friction is aligned with $\nabla\theta$, or when there is no spatial structure to $\nabla\theta \wedge \mathbf{F}$ (i.e., zero divergence), then friction does not contribute to PV evolution.

50.5 Exercises

EXERCISE 50.1: PV FOR A PERFECT BOUSSINESQ FLUID

Consider a non-hydrostatic perfect Boussinesq fluid on a rotating β -plane. Let density be linearly proportional to potential temperature,

$$\rho = \rho_0 (1 - \alpha \theta), \quad (50.43)$$

with the thermal expansion coefficient, $\alpha > 0$, assumed constant, as is the reference density ρ_0 . With this equation of state, the buoyancy takes the form

$$b = -g \left[\frac{\rho - \rho_0}{\rho_0} \right] = \alpha g \theta. \quad (50.44)$$

The governing equations for a perfect fluid version of this system are given by

$$\frac{D\mathbf{v}}{Dt} + f(\hat{\mathbf{z}} \wedge \mathbf{v}) = -\nabla\phi + b\hat{\mathbf{z}} \quad (50.45)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (50.46)$$

$$\frac{D\theta}{Dt} = 0, \quad (50.47)$$

where

$$\phi = \frac{\delta p}{\rho_0} \quad (50.48)$$

is the perturbation pressure.

- (a) Derive the equation for the material time evolution of potential vorticity in this fluid system.
 (b) Show that the vertical portion of Q^{bouss} can be written

$$Q_{\text{vert}}^{\text{bouss}} = (\zeta + f) N^2 \quad (50.49)$$

where ζ is the vertical component to the relative vorticity and N^2 is the squared buoyancy frequency (Section 25.3.4). Hint: this is a rather trivial question.

- (c) If flow maintains geostrophic and thermal wind balance, show that the horizontal portion of Q^{bouss} can be written

$$Q_{\text{horz}}^{\text{bouss}} = \boldsymbol{\omega} \cdot \nabla_z b \approx -f^{-1} |\nabla_z b|^2. \quad (50.50)$$

Hint: recall that for geostrophic flow, the vertical velocity is much smaller than horizontal.

EXERCISE 50.2: PV FOR NON-HYDROSTATIC BOUSSINESQ FLOW

Reconsider Exercise 50.1 in the presence of irreversible friction and buoyancy sources so that the governing equations are

$$\frac{D\mathbf{v}}{Dt} + f(\hat{\mathbf{z}} \wedge \mathbf{v}) = -\nabla\phi + b\hat{\mathbf{z}} + \mathbf{F} \quad (50.51)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (50.52)$$

$$\frac{D\theta}{Dt} = \dot{\theta} \quad (50.53)$$

where $\dot{\theta}$ is a diabatic heating source/sink (units of degrees per second), and \mathbf{F} is a friction operator (units of acceleration).

- (a) Derive the equation for the material time evolution of potential vorticity in this fluid system, including the irreversible contributions from friction and heating.
 (b) Derive an equation for the potential vorticity time tendency (i.e., Eulerian time derivative), written in the form

$$\frac{\partial Q}{\partial t} = -\nabla \cdot \mathbf{J}_Q. \quad (50.54)$$

What is the PV 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

$$\dot{b} = -\mu(b - b^*), \quad (50.55)$$

where μ is a constant “Newtonian” damping coefficient (units 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

$$\mathbf{F} = -\lambda \mathbf{u}, \quad (50.56)$$

with λ a constant Rayleigh damping parameter 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.

51

Impermeability theorem and PV budgets

The chapter examines a particularly striking property of the flux of potential vorticity substance. Namely, there are forms for this flux (differing by gauge choices) that have identically zero flux of PV-substance crossing an isentrope. This *impermeability theorem* holds even when there are mass and entropy fluxes crossing the isentrope. Isentropes are thus semi-permeable membranes, open to the transport of matter and heat but closed to the transport of PV-substance. This kinematic result offers further understanding for why, as developed in Section 51.3, ρQ remains constant within an isentrope unless the isentrope intersects a boundary.

READER'S GUIDE FOR THIS CHAPTER

This chapter builds from the potential vorticity mechanics introduced in Chapter 50.

51.1	Variations on the impermeability theorem	782
51.1.1	Impermeability for the Haynes-McIntyre PV flux	782
51.1.2	A kinematic derivation of impermeability	783
51.1.3	Impermeability for all components of absolute vorticity	785
51.1.4	Comments	785
51.2	Impermeability theorem for a realistic seawater EOS	785
51.2.1	Ocean PV in terms of potential density	786
51.2.2	A modified PV-substance flux	786
51.2.3	Integral constraints for steady state	788
51.3	Integrated PV	788
51.3.1	Region bounded by a buoyancy surface	789
51.3.2	Region bounded by two buoyancy surfaces	790
51.3.3	Region bounded by land and a buoyancy surface	790
51.3.4	An ocean layer that outcrops	791
51.4	Layer integrated PV budget	791
51.4.1	Layer integrated budget	792
51.4.2	Impermeability across interior layer interfaces	793
51.4.3	PV flux at the land-sea boundary	793
51.4.4	PV flux at the air-sea boundary	794
51.4.5	Thought experiments	795
51.4.6	Is there a preferred form of the PV-substance flux?	795

51.1 Variations on the impermeability theorem

In this section we derive the impermeability theorem by making use of a variety of forms for the flux vector of potential vorticity substance. The forms differ by a gauge function, so that the divergence of all the flux forms are the same.

51.1.1 Impermeability for the Haynes-McIntyre PV flux

Recall the Eulerian flux-form evolution equation for PV-substance (50.40)

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot \mathbf{J}_Q = 0 \quad \text{with} \quad \mathbf{J}_Q = \rho \mathbf{v} Q - \dot{\theta} \boldsymbol{\omega}_a + \nabla \theta \wedge \mathbf{F}, \quad (51.1)$$

with the PV-substance flux vector, \mathbf{J}_Q , here given in the form examined in [Haynes and McIntyre \(1987\)](#). Now decompose the velocity into two components, one oriented parallel to isentropes and one oriented perpendicular

$$\mathbf{v}_{\parallel} = \mathbf{v} - \hat{\mathbf{n}} (\mathbf{v} \cdot \hat{\mathbf{n}}) \quad \text{and} \quad \mathbf{v}_{\perp} = -\frac{\hat{\mathbf{n}} \partial \theta / \partial t}{|\nabla \theta|} \quad \Rightarrow \mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} + \frac{\hat{\mathbf{n}} \dot{\theta}}{|\nabla \theta|} \quad (51.2)$$

where $\hat{\mathbf{n}} = \nabla \theta / |\nabla \theta|$ is the normal vector on an isentrope. The velocity \mathbf{v}_{\perp} , by definition, satisfies

$$\frac{\partial \theta}{\partial t} + \mathbf{v}_{\perp} \cdot \nabla \theta = 0. \quad (51.3)$$

Hence, according to the kinematics detailed in Section 17.4.2,

$$\mathbf{v}_{\perp} \cdot \hat{\mathbf{n}} = \mathbf{v}_{\theta} \cdot \hat{\mathbf{n}}, \quad (51.4)$$

where \mathbf{v}_θ is the velocity of a point fixed on the θ isosurface. We make use of this key identity below.

With the velocity decomposition (51.2), the PV-substance flux vector takes the form

$$\mathbf{J}_Q = \rho \mathbf{v} Q - \dot{\theta} \boldsymbol{\omega}_a + \nabla \theta \wedge \mathbf{F} \quad (51.5a)$$

$$= \left[\mathbf{v}_{\parallel} + \mathbf{v}_{\perp} + \frac{\dot{\theta} \nabla \theta}{|\nabla \theta|^2} \right] \rho Q - \dot{\theta} \boldsymbol{\omega}_a + \nabla \theta \wedge \mathbf{F} \quad (51.5b)$$

$$= (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) \rho Q - \dot{\theta} [\boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}}] + \nabla \theta \wedge \mathbf{F} \quad (51.5c)$$

$$= (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) \rho Q - \dot{\theta} (\boldsymbol{\omega}_a)_{\parallel} + \nabla \theta \wedge \mathbf{F} \quad (51.5d)$$

$$= \mathbf{v}_{\perp} \rho Q + \left[\rho Q \mathbf{v}_{\parallel} - \dot{\theta} (\boldsymbol{\omega}_a)_{\parallel} \right] + \nabla \theta \wedge \mathbf{F} \quad (51.5e)$$

$$\equiv \mathbf{J}_{\perp} + \mathbf{J}_{\parallel}, \quad (51.5f)$$

where

$$(\boldsymbol{\omega}_a)_{\parallel} = \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \boldsymbol{\omega}_a - \left[\frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{|\nabla \theta|^2} \right] \nabla \theta = \boldsymbol{\omega}_a - \frac{\rho Q}{|\nabla \theta|} \hat{\mathbf{n}}. \quad (51.6)$$

The above results motivate us to we write the PV-substance budget equation (51.1) in the alternative form

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot (\mathbf{v}_Q \rho Q) = 0, \quad (51.7)$$

where

$$\mathbf{v}_Q = \frac{\mathbf{J}_Q}{\rho Q} = \mathbf{v}_{\perp} + \mathbf{v}_{\parallel} - \frac{\dot{\theta} (\boldsymbol{\omega}_a)_{\parallel} + \mathbf{F} \wedge \nabla \theta}{\rho Q} \quad (51.8)$$

is an effective velocity that advects the PV-substance through the fluid. A direct calculation shows that \mathbf{v}_Q satisfies the following key property

$$\mathbf{v}_Q \cdot \hat{\mathbf{n}} = \mathbf{v}_{\perp} \cdot \hat{\mathbf{n}} = \mathbf{v}_\theta \cdot \hat{\mathbf{n}}, \quad (51.9)$$

where the final equality made use of the identity (51.4). As a result, the velocity \mathbf{v}_Q , which advects PV-substance, has a normal component that is identical to that of the velocity of a point fixed on the isentrope

$$\frac{\partial \theta}{\partial t} + \mathbf{v}_Q \cdot \nabla \theta = 0. \quad (51.10)$$

We depict this result in Figure 51.1, whereby the PV-substance flux never crosses the isentrope, even as the isentrope moves and even in the presence of irreversible processes that allow for matter and heat to cross the isentrope. This result holds since the θ isosurface moves in a way to precisely track the PV-substance flux. In the presence of non-conservative processes, isentropes are permeable to matter and heat but impermeable to PV-substance. This is a rather remarkable kinematic result that has important implications for budgets of PV-substance, as shown by the discussion in the remainder of this chapter as well as in Chapter 52.

51.1.2 A kinematic derivation of impermeability

We here offer a second derivation of the impermeability theorem that emphasizes its kinematic origins. That is, we make no use of the vorticity equation or the potential temperature equation. Instead, we merely make use of the definition of potential vorticity and, critically, the non-divergence of absolute vorticity.

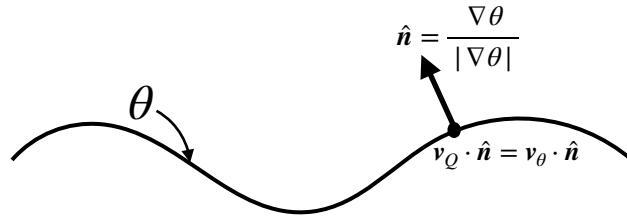


Figure 51.1: The flux, \mathbf{J}_Q , of PV-substance does not penetrate a surface of constant potential temperature (isentrope). This kinematic result follows since the effective velocity of PV-substance, $\mathbf{v}_Q = \mathbf{J}_Q^{\text{HM}}/(\rho Q)$, has the same normal component as a point fixed on a potential temperature iso-surface, $\mathbf{v}_Q \cdot \hat{\mathbf{n}} = \mathbf{v}_\theta \cdot \hat{\mathbf{n}}$. Consequently, the potential temperature 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 potential temperature surfaces are impermeable to the flux of PV-substance.

The key identity we need is the following

$$\rho Q = \boldsymbol{\omega}_a \cdot \nabla \theta = \nabla \cdot (\boldsymbol{\omega}_a \theta), \quad (51.11)$$

so that ρQ is a pure divergence. Taking the Eulerian time derivative then leads to

$$\frac{\partial(\rho Q)}{\partial t} = -\nabla \cdot \mathbf{J}_Q^{\text{kin}}, \quad (51.12)$$

where

$$\mathbf{J}_Q^{\text{kin}} = -\frac{\partial(\boldsymbol{\omega}_a \theta)}{\partial t} \quad (51.13)$$

is the kinematic form of the PV-substance flux. By construction, this flux vanishes in the steady state

$$\mathbf{J}_Q^{\text{kin}} = 0 \quad \text{in steady state}, \quad (51.14)$$

which contrasts to the Haynes-McIntyre form given by equation (51.1). We return to this point in Section 51.2 where we introduce a strategically chosen gauge term (a total curl) to render a nonzero flux in the steady state.

Introducing the velocity seen from an inertial reference frame (also called the absolute velocity; Section 13.9.1)

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x} \quad (51.15)$$

leads to

$$-\frac{\partial \boldsymbol{\omega}_a}{\partial t} \theta = -\frac{\partial(\nabla \wedge \mathbf{v}_a)}{\partial t} \theta = -\left[\nabla \wedge \frac{\partial \mathbf{v}_a}{\partial t} \right] \theta = -\nabla \wedge \left[\frac{\partial \mathbf{v}_a}{\partial t} \theta \right] + \nabla \theta \wedge \frac{\partial \mathbf{v}_a}{\partial t}. \quad (51.16)$$

Dropping the total curl yields the modified kinematic form for the PV-substance flux

$$\tilde{\mathbf{J}}_Q^{\text{kin}} = -\frac{\partial \mathbf{v}_a}{\partial t} \wedge \nabla \theta - \boldsymbol{\omega}_a \frac{\partial \theta}{\partial t}. \quad (51.17)$$

It follows that

$$\tilde{\mathbf{v}}_Q^{\text{kin}} \cdot \hat{\mathbf{n}} = \frac{\tilde{\mathbf{J}}_Q^{\text{kin}} \cdot \hat{\mathbf{n}}}{\rho Q} = -\frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{\rho Q} \frac{\partial \theta}{\partial t} \frac{1}{|\nabla \theta|} = -\frac{\partial \theta}{\partial t} \frac{1}{|\nabla \theta|} = \mathbf{v}_\theta \cdot \hat{\mathbf{n}}, \quad (51.18)$$

which then leads to the impermeability theorem for $\tilde{\mathbf{J}}_Q^{\text{kin}}$. We again emphasize that there has been no use of the dynamical equations for vorticity or potential temperature. Instead, this expression of impermeability only used the definition of potential vorticity along with the central property $\nabla \cdot \boldsymbol{\omega}_a = 0$.

51.1.3 Impermeability for all components of absolute vorticity

The above kinematic result suggests that the impermeability theorem is both remarkable and trivial. We support that dual characterization by here showing that all components of absolute vorticity satisfy an impermeability theorem with respect to the direction along which the components are defined. For that purpose, consider an arbitrary smooth scalar field λ and use it to project out that component of absolute vorticity normal to λ isosurfaces

$$\rho Q_\lambda \equiv \boldsymbol{\omega}_a \cdot \nabla \lambda = \nabla \cdot (\boldsymbol{\omega}_a \lambda). \quad (51.19)$$

Following the derivation given in Section 51.1.2 leads to

$$\frac{\partial(\rho Q_\lambda)}{\partial t} = \nabla \cdot \left[\frac{\partial \mathbf{v}_a}{\partial t} \wedge \nabla \lambda + \boldsymbol{\omega}_a \frac{\partial \lambda}{\partial t} \right] \equiv -\nabla \cdot \mathbf{J}_{Q_\lambda}. \quad (51.20)$$

It follows that

$$\frac{\mathbf{J}_{Q_\lambda} \cdot \hat{\mathbf{n}}_\lambda}{\rho Q_\lambda} = -\frac{\boldsymbol{\omega}_a \cdot \nabla \lambda}{\rho Q_\lambda} \frac{\partial \lambda}{\partial t} \frac{1}{|\nabla \lambda|} = -\frac{\partial \lambda}{\partial t} \frac{1}{|\nabla \lambda|} = \mathbf{v}_\lambda \cdot \hat{\mathbf{n}}_\lambda, \quad (51.21)$$

where

$$\hat{\mathbf{n}}_\lambda = \frac{\nabla \lambda}{|\nabla \lambda|} \quad \text{and} \quad \frac{\partial \lambda}{\partial t} + \mathbf{v}_\lambda \cdot \nabla \lambda = 0. \quad (51.22)$$

We are thus led to an impermeability theorem for the λ -surface. 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, with Section 5 of [Haynes and McIntyre \(1987\)](#) offering further mathematical insights into the inevitability of the result based on the structure of the vorticity equation. Even so, it does not reduce the importance of the entropic PV impermeability theorem for studies of stratified flows, since entropic PV is a special component of the absolute vorticity that has direct connection to dynamics and thermodynamics.

51.1.4 Comments

The impermeability theorem was introduced by [Haynes and McIntyre \(1987\)](#), with their paper met by some confusion that prompted [Haynes and McIntyre \(1990\)](#). The impermeability theorem emphasizes the importance of boundary processes when studying how potential vorticity substance changes within a domain bounded by isentropes. For example, the emphasis on boundary processes is a focus of oceanographic studies of submesoscale instabilities such as [Thomas et al. \(2008\)](#).

Besides exposing the purely kinematic aspects of the impermeability theorem, the presentation in this section reveals that there are multiple PV flux vectors that satisfy the impermeability theorem, with the vectors differing by a gauge transformation. Which 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 an example of such considerations in Section 51.2.

51.2 Impermeability theorem for a realistic seawater EOS

As seen from Section 51.1.3, 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 50.1.4). As shown in

Section 50.2, there is no such materially invariant scalar 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 PV budget.

51.2.1 Ocean PV in terms of potential density

The distinction between impermeability and material invariance was emphasized by [Marshall et al. \(2001\)](#), and we here summarize features from that paper.¹ For that purpose, introduce an ocean potential vorticity field according to

$$Q^{\text{ocn}} = \frac{\nabla b \cdot \omega_a}{\rho}, \quad (51.23)$$

where the globally defined buoyancy field, b , is approximated by a chosen potential density. As shown in Section 50.2, a globally defined buoyancy does not annihilate the baroclinicity vector for a realistic seawater equation of state

$$\mathbf{B} \cdot \nabla b = [-\nabla(1/\rho) \wedge \nabla p] \cdot \nabla b \neq 0. \quad (51.24)$$

Consequently, $DQ^{\text{ocn}}/Dt \neq 0$ even in the absence of irreversible processes. Nonetheless, the Eulerian budget for PV-substance satisfies

$$\frac{\partial(\rho Q^{\text{ocn}})}{\partial t} = -\nabla \cdot \tilde{\mathbf{J}}_Q^{\text{ocn}}, \quad (51.25)$$

and $\tilde{\mathbf{J}}_Q^{\text{ocn}}$ satisfies the impermeability theorem for b -surfaces. Both properties greatly facilitate the study of budgets for PV-substance even within an ocean with a realistic equation of state. Proof of this assertion follows from the discussion in Section 51.1.2, where we know that the kinematic flux

$$\tilde{\mathbf{J}}_Q^{\text{ocn}} = -\frac{\partial \mathbf{v}_a}{\partial t} \wedge \nabla b - \omega_a \frac{\partial b}{\partial t} = -\frac{\partial \mathbf{v}}{\partial t} \wedge \nabla b - \omega_a \frac{\partial b}{\partial t} \quad (51.26)$$

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 (51.26) follows since

$$\frac{\partial \mathbf{v}_a}{\partial t} = \frac{\partial(\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x})}{\partial t} = \frac{\partial \mathbf{v}}{\partial t}, \quad (51.27)$$

given that the Eulerian time derivative is computed at a fixed position, \mathbf{x} , and the planetary rotation is assumed constant.

51.2.2 A modified PV-substance flux

The kinematic flux (51.26) vanishes in the steady state. We here motivate a gauge transformed flux that leads to the same flux divergence yet that renders a more physically interesting steady state flux. For this purpose we make use of the vector-invariant velocity equation (equation (48.23))

$$\frac{\partial \mathbf{v}}{\partial t} + \omega_a \wedge \mathbf{v} = -\rho^{-1} \nabla p - \nabla m + \mathbf{F}, \quad (51.28)$$

¹ [Marshall et al. \(2001\)](#) builds from the generalized Bernoulli theorem of [Schär \(1993\)](#) and [Bretherton and Schär \(1993\)](#). We prefer the derivation given by [Marshall et al. \(2001\)](#) since they do not make use of the questionable thermodynamic relation (14) used by [Schär \(1993\)](#). We also consider these topics for a Boussinesq hydrostatic fluid in Section 52.1.5.

where

$$m = \frac{1}{2} \mathbf{v}^2 + \Phi \quad (51.29)$$

is the mechanical energy per mass of a fluid element. Bringing the pressure term inside of the gradient operator leads to

$$\frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega}_a \wedge \mathbf{v} = p \nabla(1/\rho) - \nabla(m + p/\rho) + \mathbf{F}. \quad (51.30)$$

[Marshall et al. \(2001\)](#) refer to $m + p/\rho$ as “a Bernoulli function”. We prefer to reserve that name for equation (24.19), in which the Bernoulli function in a compressible (non-Boussinesq) fluid, $\mathcal{B} = m + p/\rho + \mathcal{I}$, also includes the internal energy per mass, \mathcal{I} .

The vector-invariant velocity equation (51.30) leads to the cross-product

$$\frac{\partial \mathbf{v}}{\partial t} \wedge \nabla b = -(\boldsymbol{\omega}_a \wedge \mathbf{v}) \wedge \nabla b + [p \nabla(1/\rho) - \nabla(m + p/\rho) + \mathbf{F}] \wedge \nabla b \quad (51.31a)$$

$$= -(\nabla b \cdot \boldsymbol{\omega}_a) \mathbf{v} + (\nabla b \cdot \mathbf{v}) \boldsymbol{\omega}_a + [p \nabla(1/\rho) - \nabla(m + p/\rho) + \mathbf{F}] \wedge \nabla b \quad (51.31b)$$

$$= -\mathbf{v} \rho Q^{ocn} + (\dot{b} - \partial_t b) \boldsymbol{\omega}_a + [p \nabla(1/\rho) - \nabla(m + p/\rho) + \mathbf{F}] \wedge \nabla b. \quad (51.31c)$$

Use of this result leads to the flux (51.26)

$$\tilde{\mathbf{J}}_Q^{ocn} = -\frac{\partial \mathbf{v}}{\partial t} \wedge \nabla b - \boldsymbol{\omega}_a \frac{\partial b}{\partial t} \quad (51.32a)$$

$$= \mathbf{v} \rho Q^{ocn} - \dot{b} \boldsymbol{\omega}_a - \mathbf{F} \wedge \nabla b + [\nabla(m + p/\rho) - p \nabla(1/\rho)] \wedge \nabla b \quad (51.32b)$$

$$= \mathbf{J}_Q + [\nabla(m + p/\rho) - p \nabla(1/\rho)] \wedge \nabla b, \quad (51.32c)$$

where \mathbf{J}_Q is the Haynes-McIntyre form of the PV-substance flux given by equation (51.1). The term

$$\nabla(m + p/\rho) \wedge \nabla b = \nabla \wedge [(m + p/\rho) \nabla b] \quad (51.33)$$

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

$$\mathbf{J}_Q^{\text{marshall}} = \tilde{\mathbf{J}}_Q^{ocn} - \nabla(m + p/\rho) \wedge \nabla b = -\left[\frac{\partial \mathbf{v}}{\partial t} + \nabla(m + p/\rho)\right] \wedge \nabla b - \boldsymbol{\omega}_a \frac{\partial b}{\partial t}. \quad (51.34)$$

Diagnostically desirable features of $\mathbf{J}_Q^{\text{marshall}}$ include the following

- $\nabla b \cdot \mathbf{J}_Q^{\text{marshall}} / (\rho Q) = \partial_t b$, thus satisfying the impermeability theorem.
- The flux $\mathbf{J}_Q^{\text{marshall}}$ has no reference to irreversible processes. Consequently, it can in some cases be simpler to diagnose than the Haynes-McIntyre flux \mathbf{J}_Q .
- In a steady state, the flux is given by

$$\mathbf{J}_Q^{\text{marshall}} = \nabla b \wedge \nabla(m + p/\rho) = \nabla \wedge [b \nabla(m + p/\rho)]. \quad \text{steady state.} \quad (51.35)$$

Consequently, $m + p/\rho$ 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.

51.2.3 Integral constraints for steady state

The steady state PV-substance flux in the form (51.35) can be used to develop some integral constraints on the steady flow. For this purpose consider the steady form of $\mathbf{J}_Q^{\text{marshall}}$ and integrate over an arbitrary simply connected area making use of Stokes' Theorem

$$\int_S \nabla \wedge [b \nabla B] \cdot \hat{\mathbf{n}} dS = \oint_{\partial S} b \nabla B \cdot d\mathbf{r} = \oint_{\partial S} b dB = - \oint_{\partial S} B db, \quad (51.36a)$$

where we wrote

$$B = m + p/\rho, \quad (51.37)$$

which is the Bernoulli potential for a Boussinesq fluid. The first equality set

$$\nabla B \cdot d\mathbf{r} = dB, \quad (51.38)$$

and the final equality made use of

$$b dB = d(b B) - B db \quad (51.39)$$

and noted that

$$\oint_{\partial S} d(B b) = 0. \quad (51.40)$$

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

$$\int_S \mathbf{J}_Q^{\text{marshall}} \cdot \hat{\mathbf{n}} dS = 0 \quad \text{area enclosed by contour with } m + p/\rho \text{ constant or } b \text{ constant.} \quad (51.41)$$

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 fluid.

51.3 Integrated PV

In this section we derive some properties of integrated potential vorticity, with these properties relying solely on the definition of PV. We write potential vorticity using a global buoyancy field, b , as in our discussion of ocean potential vorticity in Section 51.2

$$Q = \frac{\boldsymbol{\omega}_a \cdot \nabla b}{\rho} = \frac{\nabla \cdot (\boldsymbol{\omega}_a b)}{\rho}. \quad (51.42)$$

Notably, the properties exhibited in this section hold even when there is no materially invariant potential vorticity. They rely solely on the non-divergent nature of the absolute vorticity.

In this section 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 ρQ the amount of PV-substance per volume.² With this interpretation, the amount of PV substance within an arbitrary finite region is determined by the volume integral of ρQ

$$g = \int_R Q \rho dV = \int_R \nabla \cdot (\boldsymbol{\omega}_a b) dV = \oint_{\partial R} \boldsymbol{\omega}_a b \cdot \hat{\mathbf{n}} dS, \quad (51.43)$$

where the final equality used Gauss's divergence theorem. Hence, the volume integrated PV substance in a region is determined solely by values on the region boundary. This property is strikingly distinct from material tracers. We next explore some implications of this result.

²Recall our discussion of extensive and intensive fluid properties in Section 18.3.1.

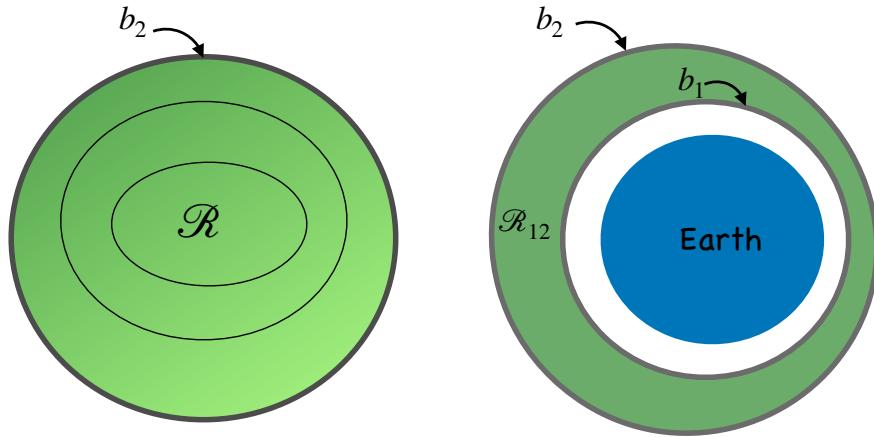


Figure 51.2: Integrating PV over regions bounded by potential temperature surfaces that do not intersect the ground. The left panel considers a single buoyancy surface, $b = b_2$, bounding the bubble-like fluid region \mathcal{R} . Notably, the region inside the bubble generally has nontrivial buoyancy distribution, as depicted here by exhibiting a couple of contours inside the b_2 -bounded domain. The only assumption is that it is wholly contained inside the $b = b_2$ contour. The right panel considers 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 PV in both \mathcal{R} and \mathcal{R}_{12} . Hence, if there is any nontrivial distribution of PV in either domain, there must be as much integrated positive values as there are negative.

51.3.1 Region bounded by a buoyancy surface

Consider a volume of fluid bounded by a single buoyancy surface as shown in the bubble-like region in the left panel of Figure 51.2. 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 (51.43) so that

$$\mathcal{I} = \oint_{\partial\mathcal{R}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \, dS = b_2 \oint_{\partial\mathcal{R}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \, dS. \quad (51.44)$$

We can now use the divergence theorem to return to the volume integral, only now with b outside of the integral

$$\mathcal{I} = b_2 \int_{\mathcal{R}} \nabla \cdot \boldsymbol{\omega}_a \, dV = 0, \quad (51.45)$$

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 $\mathcal{I} = 0$ (see Section 4.7.4).

The remarkable identity (51.45) follows from kinematics alone since no dynamical information was used in its derivation. It says that there is zero integrated PV-substance contained within any region bounded solely by a single buoyancy surface. Hence, within the domain there is just as much positive PV-substance as there is negative PV-substance. So if PV changes locally within the domain, then somewhere else it must experience an oppositely signed change so to leave a net integrated value of zero. The result holds whether there are reversible or irreversible processes acting on the buoyancy surface. Our only assumption is that the domain is fully enclosed by a b -surface.

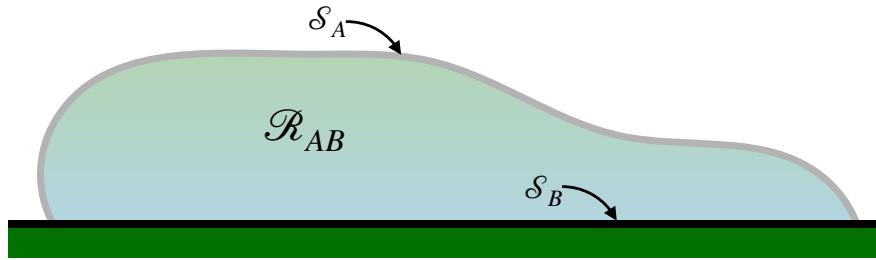


Figure 51.3: A fluid region, \mathcal{R}_{AB} , bounded by two surfaces, \mathcal{S}_A and \mathcal{S}_B . The upper surface \mathcal{S}_A is defined by a buoyancy isosurface, $b = b_A$, with this surface intersecting the ground. The lower surface, \mathcal{S}_B , is along the ground and has a buoyancy that is a function of space and time, $b_B(\mathbf{x}, t)$.

51.3.2 Region bounded by two buoyancy surfaces

The identity (51.45) has a corollary, in which we consider a region bounded by two b -surfaces such as \mathcal{R}_{12} shown in the right panel of Figure 51.2. 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

$$\mathcal{I} = \int_{\mathcal{R}_{12}} \nabla \cdot (\boldsymbol{\omega}_a b) dV = \int_{\mathcal{R}_2} \nabla \cdot (\boldsymbol{\omega}_a b) dV - \int_{\mathcal{R}_1} \nabla \cdot (\boldsymbol{\omega}_a b) dV, \quad (51.46)$$

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 single isentrope domain \mathcal{R} in Figure 51.2. Invoking the single isentrope result we see that both integrals separately vanish. We are thus led to a vanishing integral for the layer

$$\mathcal{I} = \int_{\mathcal{R}_{12}} \nabla \cdot (\boldsymbol{\omega}_a b) dV = 0. \quad (51.47)$$

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 single isentrope result.

51.3.3 Region bounded by land and a buoyancy surface

Now consider a fluid domain bounded by a buoyancy surface that intersects the ground (atmospheric example) or ocean surface (ocean example), as shown in Figure 51.2. The integrated PV is given

by

$$\mathcal{I} = \int_{\mathcal{R}_{AB}} \nabla \cdot (\boldsymbol{\omega}_a b) dV \quad (51.48a)$$

$$= \int_{\mathcal{S}_A} \boldsymbol{\omega}_a b \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_B} \boldsymbol{\omega}_a b \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (51.48b)$$

$$= \theta_A \int_{\mathcal{S}_A} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_B} b_B \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (51.48c)$$

$$= \theta_A \left[\int_{\mathcal{S}_A} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_B} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \right] + \int_{\mathcal{S}_B} (b_B - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (51.48d)$$

$$= b_A \int_{\mathcal{R}_{AB}} \nabla \cdot \boldsymbol{\omega}_a dV + \int_{\mathcal{S}_B} (b_B - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (51.48e)$$

$$= \int_{\mathcal{S}_B} (b_B - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (51.48f)$$

where we made use $\nabla \cdot \boldsymbol{\omega}_a = 0$ to reach the final equality. Consequently, the PV substance in a region enclosed by buoyancy surfaces can change only when the buoyancy surfaces intersect a boundary that has a non-constant buoyancy. As both the ground and the ocean surface have buoyancy gradients, they contribute to changes in the PV substance within the region they bound.

51.3.4 An ocean layer that outcrops

Figure 51.4 depicts a buoyancy layer in the ocean that outcrops at the ocean surface at both of its ends. Following the derivation in Section 51.3.3 we have the following integrated PV contents

$$\mathcal{I}_A \equiv \int_{\mathcal{S}_A} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3} (b - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (51.49a)$$

$$\mathcal{I}_B \equiv \int_{\mathcal{S}_B} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_2} b \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{S}_2} (b - b_B) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (51.49b)$$

Taking the difference in the integrals leads to the integrated PV within the layer \mathcal{R}_{AB}

$$\mathcal{I}_{AB} \equiv \int_{\mathcal{R}_{AB}} \nabla \cdot (\boldsymbol{\omega}_a b) dV \quad (51.50a)$$

$$= \mathcal{I}_A - \mathcal{I}_B \quad (51.50b)$$

$$= \int_{\mathcal{S}_1} (b - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + (b_B - b_A) \int_{\mathcal{S}_2} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\mathcal{S}_3} (b - b_A) \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (51.50c)$$

51.4 Layer integrated PV budget

In Section 51.3 we developed expressions for the PV-substance integrated over a selection of volumes determined by buoyancy surfaces. That discussion illustrated how the volume integrate potential vorticity 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 illustrate implications of the impermeability theorem of Section 51.1 by considering a buoyancy (isopycnal) layer within the ocean that intersects the bottom on one side and

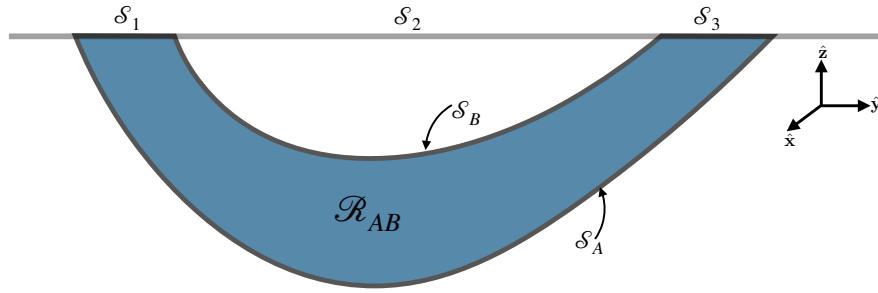


Figure 51.4: This figure depicts a buoyancy layer in the ocean that outcrops at both ends of the layer. The bounds of the regions are given by the following surfaces. Surfaces S_A and S_B are defined by buoyancy isosurfaces with $b_A < b_B$. The sea surface is decomposed into three regions, S_1 , S_2 , and S_3 according to the outcrop locations of S_A and S_B .

the atmosphere on the other (Figure 51.5). This discussion aims to expose the physical processes affecting changes to the PV-substance within the layer by unpacking the boundary fluxes.

An isopycnal layer generally moves as it expands and contracts due to both reversible and irreversible processes (waves, currents, entrainment, detrainment). The impermeability theorem means that the total potential vorticity substance for the layer changes only through exchanges at the bottom (boundary between the solid earth and ocean) and air-sea boundaries. Removing interior interfaces from the layer PV budget simplifies the budget analysis, as already revealed in Section 51.3. As per the discussion of Section 51.2, the material in this section applies even when there is no materially invariant potential vorticity. All we require is an Eulerian flux-form budget.

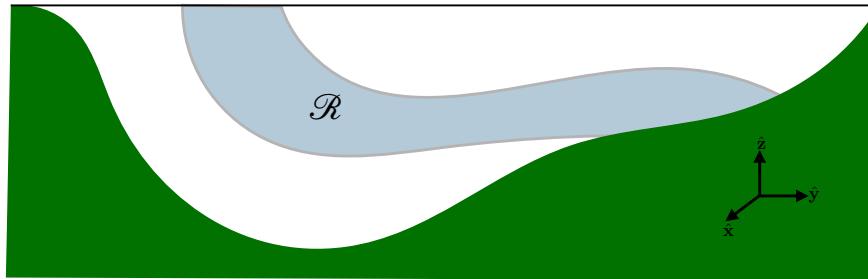


Figure 51.5: An isopycnal layer of seawater denoted by \mathcal{R} , with the layer intersecting bottom topography on one side and the atmosphere on the other.

51.4.1 Layer integrated budget

In addition to waves, currents, entrainment, and detrainment 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 general formulation framework, we derive the isopycnal layer PV budget making use of the Leibniz-Reynolds transport theorem derived in Section 18.3.4. Just as for the layer integrated tracer budget considered in Section 18.4, applying Leibniz-Reynolds to the layer integrated PV budget renders

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho Q dV \right] = \int_{\mathcal{R}} \left[\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot (\rho Q \dot{\mathbf{x}}) \right] dV, \quad (51.51)$$

where \mathcal{R} is the domain defined by the layer and

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} \quad (51.52)$$

is the velocity for a point on the layer boundary. Making use of the PV equation, $\partial(\rho Q)/\partial t = -\nabla \cdot \mathbf{J}_Q$, and the divergence theorem renders

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho Q dV \right] = \int_{\mathcal{R}} \nabla \cdot (-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) dV = \int_{\partial\mathcal{R}} (-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (51.53)$$

This result holds around the full domain boundary. Now we decompose that boundary into portions defined by isopycnal layer interfaces and those along the air-sea and land-sea boundaries.

51.4.2 Impermeability across interior layer interfaces

Rather than invoking the impermeability theorem derived in Section 51.1, we rederive it within the present context to further our confidence in its validity. We thus consider the following for interior layer interfaces

$$[-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}] \cdot \hat{\mathbf{n}} = [\rho Q (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a - \nabla b \wedge \mathbf{F}] \cdot \hat{\mathbf{n}} \quad (51.54a)$$

$$= [(\boldsymbol{\omega}_a \cdot \nabla b) (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a] \cdot \hat{\mathbf{n}} \quad (51.54b)$$

$$= [(\boldsymbol{\omega}_a \cdot \nabla b) (\dot{\mathbf{x}} - \mathbf{v}) + (\partial b / \partial t + \mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a] \cdot \hat{\mathbf{n}} \quad (51.54c)$$

$$= [(\boldsymbol{\omega}_a \cdot \nabla b) \dot{\mathbf{x}} + (\partial b / \partial t) \boldsymbol{\omega}_a] \cdot \hat{\mathbf{n}}, \quad (51.54d)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla b}{|\nabla b|} \quad (51.55)$$

is the outward normal direction pointing to regions of higher buoyancy. Now recall that the velocity of a point fixed on an isopycnal interface has a normal component that satisfies equation (51.9) (here applied to isopycnals)

$$\dot{\mathbf{x}} \cdot \hat{\mathbf{n}} = -\frac{\partial b / \partial t}{|\nabla b|}. \quad (51.56)$$

This result then leads to the impermeability statement for isopycnal interfaces in the fluid interior

$$(-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = 0. \quad (51.57)$$

We thus conclude that changes to the layer integrated PV occur only via transfer across the land-sea boundary and the air-sea boundary

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho Q dV \right] = \int_{\text{land-sea}} [-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}] \cdot \hat{\mathbf{n}} d\mathcal{S} + \int_{\text{air-sea}} [-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}] \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (51.58)$$

51.4.3 PV flux at the land-sea boundary

We here evaluate the PV flux from equation (51.5a) at the land-sea boundary

$$-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}} = \rho Q (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}. \quad (51.59)$$

At a solid 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 $\dot{\mathbf{x}} \cdot \hat{\mathbf{n}} = 0$. The bottom boundary condition is thus affected just by irreversible processes

$$(-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = (\dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{n}}. \quad (51.60)$$

In many parts of the ocean, geothermal heating is negligible so that there is no buoyancy input at the bottom, thus leaving just the contribution from friction

$$(-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} \approx (\nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} = (\hat{\mathbf{n}} \wedge \nabla b) \cdot \mathbf{F}. \quad (51.61)$$

Furthermore, in the absence of geothermal heating the buoyancy satisfies a no-flux boundary condition, which can be ensured by having the buoyancy field maintaining

$$\hat{\mathbf{n}} \cdot \nabla b = 0. \quad (51.62)$$

Buoyancy isolines thus intersect the bottom parallel to the bottom outward normal, as shown in Figure 51.6. Assuming buoyancy increases upward along the sloping bottom, then $(\hat{\mathbf{n}} \wedge \nabla b) \cdot \mathbf{F}$ projects out that component of the friction vector pointing parallel to the bottom. Within the bottom boundary layer, quadratic bottom drag is a common parameterization of the deceleration associated with turbulent frictional processes

$$\mathbf{F} = -C_d |\mathbf{u}| \mathbf{u}, \quad (51.63)$$

where C_d is a non-dimensional drag coefficient and \mathbf{u} is the horizontal velocity. In this case the boundary condition for PV takes the form

$$(\nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} = C_d |\mathbf{u}| (\nabla b \wedge \hat{\mathbf{n}}) \cdot \mathbf{u}. \quad (51.64)$$

This boundary flux provides a positive source for PV in cases where the bottom boundary layer flow is clockwise around abyssal hills and counter-clockwise around abyssal bowls, and conversely for oppositely oriented flow. In contrast, bottom flows that are parallel to ∇b (i.e., flows that are orthogonal to buoyancy isosurfaces at the bottom boundary) provide a zero source for PV since $\nabla b \wedge \mathbf{F} = 0$. This result is expected from the discussion in Section 50.4 and Figure 50.3, where we note that friction changes PV by rotating buoyancy surfaces, with that rotation realized only when friction is not aligned with ∇b .

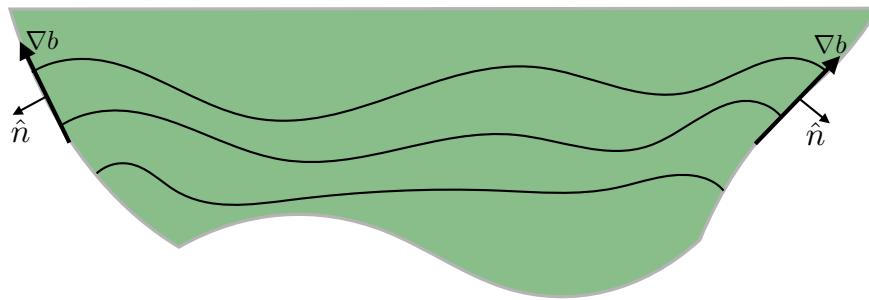


Figure 51.6: Buoyancy isosurfaces that intersect the bottom. As discussed in Section 18.4 and depicted in Figure 18.4, in the absence of geothermal heating buoyancy isosurfaces satisfy the no-normal flux bottom boundary condition, $\hat{\mathbf{n}} \cdot \nabla b = 0$, which requires the buoyancy surfaces to be orthogonal to the bottom. This structure for the buoyancy surfaces affects how friction impacts on the layer-integrated PV budget.

51.4.4 PV flux at the air-sea boundary

We make use of the kinematic boundary condition derived in Section 17.4.3 for the permeable air-sea boundary, where the boundary condition (17.61) leads to

$$\rho \hat{\mathbf{n}} \cdot (\dot{\mathbf{x}} - \mathbf{v}) = \mathcal{Q}_m \quad \text{air-sea boundary} \quad (51.65)$$

with \mathcal{Q}_m the mass per time per surface area crossing the boundary. We are thus led to the air-sea boundary condition

$$(-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = Q \mathcal{Q}_m + (\dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{n}}. \quad (51.66)$$

Besides the irreversible processes, PV 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 PV across the boundary via the boundary mass transport. Note that we might choose to approximate the sea surface as nearly flat, so that

$$\hat{\mathbf{n}} \approx \hat{\mathbf{z}}, \quad (51.67)$$

in which case

$$(-\mathbf{J}_Q + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} \approx Q \mathcal{Q}_m + \dot{b}(\zeta + f) + (\nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{z}}. \quad (51.68)$$

The buoyancy source term appears multiplied by the vertical component to the absolute vorticity, so that modifications to PV appear in proportion to the sign of both \dot{b} and $\zeta + f$. We think of this term as acting to stretch/compress the fluid column so to alter vorticity and hence PV. The friction source arises from the mis-alignment of the horizontal friction vector and the horizontal buoyancy gradient, thus acting like a torque to spin the fluid analogous to the action from baroclinicity described in Section 48.3.

51.4.5 Thought experiments

The surface PV flux (51.68) provides an explicit expression for how surface boundary fluxes affect the PV budget within a buoyancy layer outcropping at the ocean surface. It contains a wealth of physics with many scenarios for exploring via thought experiments. For example, consider a fluid region initially with zero baroclinicity and zero flow so that the PV is given by $f N^2$, with N^2 the squared buoyancy frequency. The surface PV flux (51.68) creates PV via the mass flux term and through heating/cooling. If this term alone affected the PV, and it did so uniformly in space, then it would alter PV 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 hence generate a nonzero friction contribution to the PV flux.

Consider an initially homogenous box of seawater with zero PV. In this case it is only the buoyancy term, $\dot{b} f$, that contributes to initial changes in PV. Northern hemisphere ($f > 0$) surface cooling ($\dot{b} < 0$) adds negative PV to the box. Cooling also initiates gravitational instability that mixes the water and in turn spreads the negative PV 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 PV into the fluid interior. Notably, the discussion in Section 57.5 indicates that a region with $f Q < 0$ is locally unstable to symmetric instability, with the generated symmetric instability that locally brings the flow towards a zero PV state. However, the constraints from impermeability mean that the net PV-substance remains unchanged within a buoyancy layer, even in the presence of energetic mixing.

51.4.6 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 (51.1))

$$\mathbf{J}_Q = \rho \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}. \quad (51.69)$$

We could have instead chosen to work with the Marshall form (equation (51.34))

$$\mathbf{J}_Q^{\text{marshall}} = - \left[\frac{\partial \mathbf{v}}{\partial t} + \nabla(m + p/\rho) \right] \wedge \nabla b - \boldsymbol{\omega}_{\text{a}} \frac{\partial b}{\partial t}, \quad (51.70)$$

or the modified kinematic form (equation (51.17))

$$\tilde{\mathbf{J}}_Q^{\text{kin}} = - \frac{\partial \mathbf{v}_{\text{a}}}{\partial t} \wedge \nabla \theta - \boldsymbol{\omega}_{\text{a}} \frac{\partial \theta}{\partial t}. \quad (51.71)$$

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.

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

$$\mathbf{J}_Q = \rho \mathbf{v} Q - (\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_{\text{a}} + \nabla b \wedge \mathbf{F} \quad (51.72)$$

$$\mathbf{J}_Q^{\text{marshall}} = - \nabla(m + p/\rho) \wedge \nabla b \quad (51.73)$$

$$\tilde{\mathbf{J}}_Q^{\text{kin}} = 0. \quad (51.74)$$

The physical picture for $\tilde{\mathbf{J}}_Q^{\text{kin}}$ 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 $\mathbf{J}_Q^{\text{marshall}}$ 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 $\mathbf{J}_Q^{\text{marshall}}$ for steady budgets, and we explore facets of that approach in Section 52.1 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 PV. Such forcing exposes the flow to a variety of local instabilities (symmetric, centrifugal, gravitational; see Chapter 57). [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 PV flux and to furthermore ensure it is properly formulated within numerical models (e.g., [Hallberg and Rhines \(1996\)](#)).

52

Hydrostatic and Boussinesq ocean potential vorticity

This chapter formulates the potential vorticity budget for the hydrostatic and Boussinesq ocean equations. The aim is to expose details for the practitioner interested in the analysis of ocean potential vorticity and its budgets. Relatively few new concepts are introduced beyond those already discussed in Chapter 50 and earlier.

READER'S GUIDE FOR THIS CHAPTER

The material here is concerned with the details PV budgets with a focus on the ocean. It requires a working understanding of potential vorticity from Chapter 50 and the budget of PV-substance in Chapter 51. We also encounter mechanics as formulated using the generalized vertical coordinates from Chapters 11, 21, and 32. Material from this chapter is not used by subsequent chapters.

52.1	Perfect fluid PV using Cartesian coordinates	798
52.1.1	Vorticity equation	799
52.1.2	Potential vorticity	801
52.1.3	Potential vorticity flux vector	801
52.1.4	Kinematic derivation of the PV-substance flux	802
52.1.5	Potential vorticity flux vector	803
52.2	Perfect fluid PV using isopycnal coordinates	805
52.2.1	Derivation of the vorticity equation	805
52.2.2	Derivation of the potential vorticity equation	805
52.2.3	A note about specific thickness	806
52.2.4	Coordinate transforming vorticity and potential vorticity	806
52.3	Isopycnal coordinate PV with irreversible processes	807
52.3.1	Derivation method I	807
52.3.2	Derivation method II	809

52.1 Perfect fluid PV using Cartesian coordinates

In this section we develop the vorticity and potential vorticity substance budgets for a hydrostatic Boussinesq fluid in the presence of diabatic and frictional forcing. This system is of particular importance for models of the ocean circulation with the governing equations given by

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{v} = -\nabla_z \psi + \mathbf{F} \quad (52.1a)$$

$$\frac{\partial \psi}{\partial z} = b \quad (52.1b)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (52.1c)$$

$$\frac{Db}{Dt} = \dot{b}. \quad (52.1d)$$

In this equation, the non-divergent velocity field is written

$$\mathbf{v} = (\mathbf{u}, w) = \mathbf{u} + w \hat{\mathbf{z}}. \quad (52.2)$$

The perturbation pressure is

$$\psi = \frac{\delta p}{\rho_0} = \frac{p - p_0}{\rho_0}, \quad (52.3)$$

with $p_0 = p_0(z)$ in hydrostatic balance with the constant reference density

$$\frac{dp_0}{dz} = -g \rho_0 \quad (52.4)$$

and p the full hydrostatic pressure satisfying the hydrostatic balance

$$\frac{\partial p}{\partial z} = -g \rho. \quad (52.5)$$

The buoyancy is given by

$$b = -g \left[\frac{\rho - \rho_0}{\rho_0} \right], \quad (52.6)$$

with negative buoyancy for fluid more dense than the reference density, ρ_0 , and vice versa. We assume the Coriolis parameter of the form $\mathbf{f} = f \hat{\mathbf{z}}$, in which case

$$\mathbf{f} \wedge \mathbf{v} = \mathbf{f} \wedge \mathbf{u}. \quad (52.7)$$

Finally, the horizontal friction acceleration vector is given by

$$\mathbf{F} = (F^x, F^y, 0) \quad (52.8)$$

and the gradient operator is

$$\nabla = \nabla_z + \hat{\mathbf{z}} \partial_z. \quad (52.9)$$

52.1.1 Vorticity equation

To derive the vorticity equation, it is useful to combine the horizontal momentum equation with the hydrostatic balance, in which case

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{v} = -\nabla \psi + b \hat{\mathbf{z}} + \mathbf{F}. \quad (52.10)$$

As for the non-hydrostatic case (Section 48.4.1), we rewrite the self-advection operator, $(\mathbf{v} \cdot \nabla) \mathbf{u}$, before taking the curl. We will in turn introduce the hydrostatic relative vorticity, defined as the curl of the horizontal velocity

$$\boldsymbol{\omega}^{\text{hy}} = \nabla \wedge \mathbf{u} = -\hat{\mathbf{x}} \partial_z v + \hat{\mathbf{y}} \partial_z u + \hat{\mathbf{z}} \zeta, \quad (52.11)$$

where

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad (52.12)$$

is the vertical component to the relative vorticity. It is then straightforward to show that

$$\boldsymbol{\omega}^{\text{hy}} \wedge \mathbf{v} = \hat{\mathbf{x}} (w \partial_z u - v \partial_x v + v \partial_y u) + \hat{\mathbf{y}} (w \partial_z v - u \partial_y u + u \partial_x v) - \hat{\mathbf{z}} \partial_z (u^2 + v^2)/2 \quad (52.13a)$$

$$= w \partial_z \mathbf{u} + \zeta (\hat{\mathbf{x}} v - \hat{\mathbf{u}} u) - \hat{\mathbf{z}} \partial_z (u^2 + v^2)/2, \quad (52.13b)$$

in which case

$$\nabla(\mathbf{u}^2/2) + \boldsymbol{\omega}^{\text{hy}} \wedge \mathbf{v} = \nabla(u^2 + v^2)/2 - \hat{\mathbf{z}} \partial_z (u^2 + v^2)/2 + w \partial_z \mathbf{u} + \zeta (u \hat{\mathbf{y}} - v \hat{\mathbf{x}}) \quad (52.14a)$$

$$= (u \partial_x + v \partial_y + w \partial_z) \mathbf{u} \quad (52.14b)$$

$$= (\mathbf{v} \cdot \nabla) \mathbf{u}. \quad (52.14c)$$

The material time derivative of the horizontal velocity can thus be written as

$$\frac{D\mathbf{u}}{Dt} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{u} = \frac{\partial \mathbf{u}}{\partial t} + \boldsymbol{\omega}^{\text{hy}} \wedge \mathbf{v} + \nabla(\mathbf{u}^2/2), \quad (52.15)$$

which then leads to the vector-invariant horizontal velocity equation

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{f} + \boldsymbol{\omega}^{\text{hy}}) \wedge \mathbf{v} = -\nabla(\psi + \mathbf{u}^2/2) + \hat{\mathbf{z}} b + \mathbf{F}, \quad (52.16)$$

which can be written in the equivalent form

$$\left[\frac{\partial}{\partial t} + w \frac{\partial}{\partial z} + (f + \zeta) \hat{\mathbf{z}} \wedge \right] \mathbf{u} = -\nabla(\psi + \mathbf{u}^2/2) + \hat{\mathbf{z}} b + \mathbf{F}. \quad (52.17)$$

Now take the curl of equation (52.16), and make use of the identity

$$\nabla \wedge (\boldsymbol{\omega}_a^{hy} \wedge \mathbf{v}) = (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla) \mathbf{v}, \quad (52.18)$$

where we introduced the absolute vorticity for the hydrostatic fluid

$$\boldsymbol{\omega}_a^{hy} = \mathbf{f} + \boldsymbol{\omega}^{hy} \quad (52.19)$$

thus resulting in the vorticity equation

$$\frac{\partial \boldsymbol{\omega}_a^{hy}}{\partial t} + (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_a^{hy} = (\boldsymbol{\omega}_a^{hy} \cdot \nabla) \mathbf{v} + \nabla \wedge \hat{\mathbf{z}} b + \nabla \wedge \mathbf{F}. \quad (52.20)$$

Since the Coriolis parameter is time independent, we can add it to the time derivative to yield

$$\frac{D\boldsymbol{\omega}_a^{hy}}{Dt} = \underbrace{(\boldsymbol{\omega}_a^{hy} \cdot \nabla) \mathbf{v}}_{\text{stretching + tilting}} + \underbrace{\nabla \wedge \hat{\mathbf{z}} b}_{\text{baroclinicity}} + \underbrace{\nabla \wedge \mathbf{F}}_{\text{friction curl}}. \quad (52.21)$$

Equation (52.21) is the vorticity equation for a hydrostatic Boussinesq fluid. We can compare this equation to the vorticity equation for a non-hydrostatic and non-Boussinesq fluid (equation (48.32)). Both equations have a vorticity source due to stretching and tilting. However, the baroclinicity vector for the Boussinesq fluid is given by

$$\mathbf{B}_{\text{hydro-bouss}} = \nabla \wedge \hat{\mathbf{z}} b = \nabla b \wedge \hat{\mathbf{z}}, \quad (52.22)$$

which is simpler than

$$\mathbf{B} = (\nabla \rho \wedge \nabla p) / \rho^2 \quad (52.23)$$

(equation (48.19)) for the non-hydrostatic and non-Boussinesq fluid. We see in Exercise 50.2 that the same baroclinicity vector (52.22) appears for the non-hydrostatic Boussinesq fluid. So in general, we can diagnose the presence of baroclinicity for the Boussinesq fluid merely by noting whether there is a slope to the buoyancy surfaces relative to the horizontal, such as in Figure 52.1. That is, a sloping buoyancy surface acts as a vorticity source for the Boussinesq fluid by providing a torque to fluid elements (see Section 48.3).

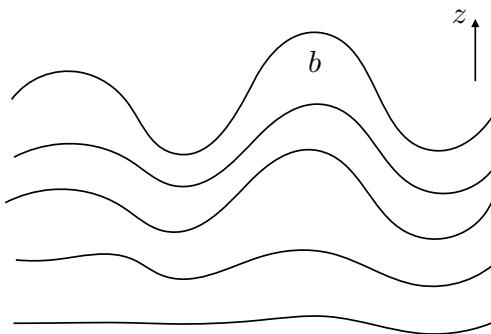


Figure 52.1: Baroclinicity in a Boussinesq fluid is manifest by nonzero horizontal gradients in the buoyancy field. Here we depict a region of relatively strong baroclinicity above a region of weaker baroclinicity. A sloping buoyancy surface is therefore synonymous with a nontrivial baroclinic structure.

52.1.2 Potential vorticity

Potential vorticity evolves in the absence of baroclinicity, which is eliminated from the vorticity equation by projecting the absolute vorticity onto the direction normal to buoyancy surfaces

$$\nabla b \cdot \frac{D\omega_a^{hy}}{Dt} = \nabla b \cdot [(\omega_a^{hy} \cdot \nabla) \mathbf{v}] + \nabla b \cdot (\nabla \wedge \mathbf{F}), \quad (52.24)$$

where we used

$$\nabla b \cdot (\nabla \wedge \hat{\mathbf{z}} b) = 0. \quad (52.25)$$

We next make use of the identity

$$\frac{D(\partial b / \partial x^i)}{Dt} = \frac{\partial}{\partial x^i} \left[\frac{Db}{Dt} \right] - \nabla b \cdot \frac{\partial \mathbf{v}}{\partial x^i} \quad (52.26a)$$

$$= \frac{\partial \dot{b}}{\partial x^i} - \nabla b \cdot \frac{\partial \mathbf{v}}{\partial x^i}, \quad (52.26b)$$

so that

$$\omega_a^{hy} \cdot \left[\frac{D \nabla b}{Dt} \right] = \omega_a^{hy} \cdot \nabla \dot{b} - \nabla b \cdot [(\omega_a^{hy} \cdot \nabla) \mathbf{v}]. \quad (52.27)$$

Making use of this result in equation (52.24) renders

$$\nabla b \cdot \frac{D\omega_a^{hy}}{Dt} + \omega_a^{hy} \cdot \frac{D\nabla b}{Dt} = \omega_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \wedge \mathbf{F}), \quad (52.28)$$

which leads to

$$\frac{DQ}{Dt} = \omega_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \wedge \mathbf{F}) \quad (52.29)$$

where

$$Q = \omega_a^{hy} \cdot \nabla b = \omega_a^{hy} \cdot \nabla b + f \frac{\partial b}{\partial z} \quad (52.30)$$

is the potential vorticity for a rotating hydrostatic Boussinesq fluid. 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 (52.11). In this way, PV takes on the form

$$Q = \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + (\zeta + f) \frac{\partial b}{\partial z} = \hat{\mathbf{z}} \cdot \left[\frac{\partial \mathbf{u}}{\partial z} \wedge \nabla b \right] + (\zeta + f) \frac{\partial b}{\partial z}. \quad (52.31)$$

52.1.3 Potential vorticity flux vector

The material form of the PV equation (52.29) is converted into its Eulerian flux-form via

$$\frac{\partial Q}{\partial t} + \nabla \cdot (\mathbf{v} Q) = \omega_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \wedge \mathbf{F}) \quad (52.32a)$$

$$= \nabla \cdot [\dot{b} \omega_a^{hy} + b (\nabla \wedge \mathbf{F})] \quad (52.32b)$$

$$= \nabla \cdot [\dot{b} \omega_a^{hy} + \nabla \wedge (b \mathbf{F}) - \nabla b \wedge \mathbf{F}] \quad (52.32c)$$

$$= \nabla \cdot [\dot{b} \omega_a^{hy} - \nabla b \wedge \mathbf{F}], \quad (52.32d)$$

where we used

$$\nabla \cdot \mathbf{v} = 0 \quad (52.33a)$$

$$\nabla \cdot \boldsymbol{\omega}_a^{hy} = 0 \quad (52.33b)$$

$$\nabla \cdot (\nabla \wedge \mathbf{F}) = 0 \quad (52.33c)$$

$$\nabla \cdot [\nabla \wedge (b \mathbf{F})] = 0. \quad (52.33d)$$

The conservation equation (52.32d) allows us to identify a potential vorticity flux vector

$$\mathbf{J}_Q = \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A}, \quad (52.34)$$

so that the PV equation takes the form

$$\frac{\partial Q}{\partial t} + \nabla \cdot \mathbf{J}_Q = 0. \quad (52.35)$$

The potential vorticity flux (52.34) is comprised of an advective term

$$\mathbf{J}_{\text{advective}} = \mathbf{v} Q, \quad (52.36)$$

and non-advective terms arising from diabatic and frictional forcing,

$$\mathbf{J}_{\text{non-advective}} = -\dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F}, \quad (52.37)$$

as well as a gauge term,

$$\mathbf{J}_{\text{gauge}} = \nabla \wedge \mathbf{A}. \quad (52.38)$$

52.1.4 Kinematic derivation of the PV-substance flux

Following the discussion in Section 51.1.2, we consider a purely kinematic derivation of the PV flux vector. For that purpose write the hydrostatic Boussinesq PV in the form

$$Q = \boldsymbol{\omega}_a^{hy} \cdot \nabla b = \nabla \cdot (b \boldsymbol{\omega}_a^{hy}), \quad (52.39)$$

which follows since $\nabla \cdot \boldsymbol{\omega}_a^{hy} = 0$. Taking the Eulerian time derivative then leads to

$$\frac{\partial Q}{\partial t} = \nabla \cdot \left[\frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + b \frac{\partial \boldsymbol{\omega}_a^{hy}}{\partial t} \right] \quad (52.40a)$$

$$= \nabla \cdot \left[\frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \frac{\partial \mathbf{u}}{\partial t} \wedge \nabla b \right], \quad (52.40b)$$

which made use of the identities

$$\frac{\partial \boldsymbol{\omega}_a^{hy}}{\partial t} = \frac{\partial \boldsymbol{\omega}}{\partial t} \quad \text{and} \quad \nabla \cdot [\nabla \wedge (b \mathbf{u})] = 0. \quad (52.41)$$

The PV thus evolves according to

$$\frac{\partial Q}{\partial t} = -\nabla \cdot \tilde{\mathbf{J}}_Q \quad (52.42)$$

where the PV flux is given by

$$\tilde{\mathbf{J}}_Q = -\frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} - \frac{\partial \mathbf{u}}{\partial t} \wedge \nabla b. \quad (52.43)$$

This form of the PV flux manifests the impermeability property since

$$\hat{\mathbf{n}} \cdot (\tilde{\mathbf{J}}_Q/Q) = -\frac{1}{|\nabla b|} \frac{\partial b}{\partial t} = \hat{\mathbf{n}} \cdot \mathbf{v}_{b\perp}, \quad (52.44)$$

where $\hat{\mathbf{n}} = \nabla b |\nabla b|^{-1}$ is the unit normal for a buoyancy surface, and with $\mathbf{v}_{b\perp}$ the velocity of a point on the isopycnal that satisfies

$$\frac{\partial b}{\partial t} + \mathbf{v}_{b\perp} \cdot \nabla b = 0. \quad (52.45)$$

52.1.5 Potential vorticity flux vector

We introduced the Bernoulli potential and Bernoulli theorem in Section 24.2.3. [Schär \(1993\)](#) provided a generalization of Bernoulli's Theorem, with [Marshall \(2000\)](#), [Marshall et al. \(2001\)](#), and [Polton and Marshall \(2007\)](#) applying this theorem to oceanic contexts. In particular, [Marshall \(2000\)](#) and [Polton and Marshall \(2007\)](#) work with a Boussinesq and hydrostatic ocean. We here consider the details of their analysis, with some of the material in this section building from the more general discussion in Section 51.2.1.

Momentum equation

We start by exposing the Boussinesq form of the Bernoulli potential from the vector-invariant velocity equation. For this purpose, return to the horizontal momentum equation (52.16), and expand the expressions for the perturbation pressure and the buoyancy

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{f} + \boldsymbol{\omega}^{\text{hy}}) \wedge \mathbf{v} = -\nabla(\psi + \mathbf{u}^2/2) + \hat{\mathbf{z}} b + \mathbf{F} \quad (52.46a)$$

$$= -\nabla(\mathbf{u}^2/2) - \frac{1}{\rho_0} \nabla(p - p_0) - \hat{\mathbf{z}} \left[\frac{g(\rho - \rho_0)}{\rho_0} \right] + \mathbf{F} \quad (52.46b)$$

$$= -\nabla(\mathbf{u}^2/2) - \frac{1}{\rho_0} \nabla p - \hat{\mathbf{z}} \frac{g\rho}{\rho_0} + \mathbf{F} \quad (52.46c)$$

$$= -\nabla \left[\frac{\mathbf{u}^2}{2} + \frac{p}{\rho_0} \right] - \hat{\mathbf{z}} \left[\frac{g\rho - g\rho_o + g\rho_0}{\rho_0} \right] + \mathbf{F} \quad (52.46d)$$

$$= -\nabla \left[\frac{\mathbf{u}^2}{2} + \frac{p}{\rho_0} + g z \right] - \hat{\mathbf{z}} \left[\frac{g(\rho - \rho_o)}{\rho_0} \right] + \mathbf{F} \quad (52.46e)$$

$$= -\nabla \mathcal{B} + \hat{\mathbf{z}} b + \mathbf{F}, \quad (52.46f)$$

where we introduced the Bernoulli potential for a hydrostatic and Boussinesq fluid

$$\mathcal{B} = \frac{\mathbf{u}^2}{2} + \frac{p}{\rho_0} + g z. \quad (52.47)$$

Note that for a non-Boussinesq fluid the internal energy appears in the Bernoulli function (Section 24.2.2).

Potential vorticity flux

The flux-form potential vorticity conservation statement remains as given by equation (52.32d), and the PV-substance flux is given by equation (52.34). However, we can make use of the gauge

invariance of the PV-substance flux to write it in a manner conducive to analyzing steady state conditions. For this purpose, operate with $\nabla b \wedge$ on the velocity equation (52.46f) to have

$$\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} + \nabla b \wedge (\boldsymbol{\omega}_a^{hy} \wedge \mathbf{v}) = -\nabla b \wedge \nabla \mathcal{B} + \nabla b \wedge \hat{\mathbf{z}} b + \nabla b \wedge \mathbf{F}. \quad (52.48)$$

Make use of the identity

$$\nabla b \wedge (\boldsymbol{\omega}_a^{hy} \wedge \mathbf{v}) = (\nabla b \cdot \mathbf{v}) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \mathbf{v} \quad (52.49)$$

in equation (52.48) to render

$$(\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \mathbf{v} = -\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} - \nabla b \wedge \nabla \mathcal{B} + \nabla b \wedge \hat{\mathbf{z}} b + \nabla b \wedge \mathbf{F}. \quad (52.50)$$

Now write the flux given by equation (52.34) in the form

$$\mathbf{J}_Q = \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (52.51a)$$

$$= \mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) - \left[\frac{\partial b}{\partial t} + \mathbf{v} \cdot \nabla b \right] \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (52.51b)$$

$$= [\mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) - (\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a^{hy}] - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (52.51c)$$

$$= \left[\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} + \nabla b \wedge \nabla \mathcal{B} - \nabla b \wedge \hat{\mathbf{z}} b - \nabla b \wedge \mathbf{F} \right] - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (52.51d)$$

$$= \nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \nabla \mathcal{B} - \nabla b \wedge \hat{\mathbf{z}} b + \nabla \wedge \mathbf{A}. \quad (52.51e)$$

$$= \nabla b \wedge \left[\frac{\partial \mathbf{u}}{\partial t} + \nabla \mathcal{B} \right] - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla \wedge (\mathbf{A} - \hat{\mathbf{z}} b^2/2). \quad (52.51f)$$

Choosing the gauge function according to

$$\mathbf{A} = \hat{\mathbf{z}} (b^2/2) \quad (52.52)$$

renders the flux vector

$$\mathbf{J}_Q = \nabla b \wedge \left(\frac{\partial \mathbf{u}}{\partial t} + \nabla \mathcal{B} \right) - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} \quad (52.53a)$$

$$= \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F}. \quad (52.53b)$$

The second equality reintroduced the first form of the PV flux given by equation (52.34); made use of the gauge choice (52.52); and furthermore dropped the non-divergent term

$$\nabla b \wedge b \hat{\mathbf{z}} = (1/2) \nabla b^2 \wedge \hat{\mathbf{z}} \quad \text{with} \quad \nabla \cdot (\nabla b \wedge b \hat{\mathbf{z}}) = 0. \quad (52.54)$$

In the steady state, the PV flux (52.53a) reduces to

$$\mathbf{J}_Q^{ss} = \nabla b \wedge \nabla \mathcal{B}. \quad (52.55)$$

Hence, the steady state potential vorticity flux is aligned with the intersection of surfaces of constant buoyancy and Bernoulli potential

$$\nabla b \cdot \mathbf{J}_Q^{ss} = 0 \quad \text{and} \quad \nabla \mathcal{B} \cdot \mathbf{J}_Q^{ss} = 0. \quad (52.56)$$

This result is the Boussinesq/hydrostatic form of the more general result derived in Section 51.2.3.

52.2 Perfect fluid PV using isopycnal coordinates

In Section 52.1, we show how the absolute vorticity in a Boussinesq hydrostatic fluid with a simplified seawater equation of state (Section 51.2), when projected into the direction normal to constant buoyancy surface, $\omega_a \cdot \nabla b$, is not affected by baroclinicity; i.e., that 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 feature of buoyancy surfaces. Hence, derivation of the PV equation in using isopycnal coordinates never requires us to explicitly eliminate the baroclinicity vector. Instead, we use the relevant scalar field, buoyancy, to partition the vertical direction in formulating the equations of motion and the vorticity equation. In so doing, isopycnal coordinates remove baroclinicity *a priori*, thus allowing the potential vorticity derivation to proceed as for the shallow water system in Section 47.3. The central assumption needed for this approach is that surfaces of constant buoyancy must remain stably stratified in the vertical.

Throughout this section we make use of the horizontal derivatives on constant buoyancy surfaces (derived in Section 11.12), written here in the form

$$\nabla_b = \hat{\mathbf{x}} \left[\frac{\partial}{\partial x} \right]_b + \hat{\mathbf{y}} \left[\frac{\partial}{\partial y} \right]_b. \quad (52.57)$$

There is a corresponding isopycnal relative vorticity that appears in the following

$$\tilde{\zeta} = \left[\frac{\partial v}{\partial x} \right]_b - \left[\frac{\partial u}{\partial y} \right]_b = \hat{\mathbf{z}} \cdot (\nabla_b \wedge \mathbf{u}). \quad (52.58)$$

52.2.1 Derivation of the vorticity equation

Acting with the vertical projection of the curl, $\hat{\mathbf{z}} \cdot (\nabla_b \wedge)$, onto the adiabatic form of the momentum equation (45.14a) leads to the isopycnal vorticity equation

$$\left[\frac{\partial \tilde{\zeta}_a}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \tilde{\zeta}_a = -\tilde{\zeta}_a \nabla_b \cdot \mathbf{u} \quad (52.59)$$

where

$$\tilde{\zeta}_a = f + \hat{\mathbf{z}} \cdot (\nabla_b \wedge \mathbf{u}) = f + \tilde{\zeta} \quad (52.60)$$

is the absolute vorticity, written as the planetary vorticity plus the isopycnal relative vorticity. The left hand side of equation (52.59) is the material time derivative of absolute vorticity (see equation (45.5b)), so that we can write

$$\frac{D\tilde{\zeta}_a}{Dt} = -\tilde{\zeta}_a \nabla_b \cdot \mathbf{u}. \quad (52.61)$$

52.2.2 Derivation of the potential vorticity equation

We now make use of the thickness equation derived in Section 45.1.3, here realized in its material form to eliminate the convergence $\nabla_b \cdot \mathbf{u}$ on the right hand side of equation (52.61), thus leading to

$$\frac{D\tilde{\zeta}_a}{Dt} - \frac{\tilde{\zeta}_a}{h} \frac{Dh}{Dt} = 0. \quad (52.62)$$

Introducing the isopycnal potential vorticity

$$\tilde{Q} = \frac{\tilde{\zeta}_a}{h} = \frac{f + \tilde{\zeta}}{h} \quad (52.63)$$

leads to

$$\frac{D\tilde{Q}}{Dt} = 0. \quad (52.64)$$

Expanding the material time derivative into its components according to equation (45.5b), and making use of the adiabatic form of the thickness equation leads to the Eulerian flux-form equation

$$\left[\frac{\partial(h\tilde{Q})}{\partial t} \right]_b + \nabla_b \cdot (h\tilde{Q}\mathbf{u}) = 0. \quad (52.65)$$

52.2.3 A note about specific thickness

Our use of the thickness (see equation (45.9)),

$$h = (\partial z / \partial b) db, \quad (52.66)$$

for defining the potential vorticity (52.63) accords with the shallow water case as detailed in Chapter 47. A slightly modified form makes use of the specific thickness

$$\hbar = \frac{\partial z}{\partial z}, \quad (52.67)$$

which, for an adiabatic fluid, satisfies the same equation as the thickness

$$\frac{D\hbar}{Dt} = -\hbar \nabla_b \cdot \mathbf{u}. \quad (52.68)$$

Hence, PV defined according to the specific thickness

$$Q = \frac{f + \tilde{\zeta}}{\hbar} \quad (52.69)$$

is also materially constant and thus satisfies the Eulerian flux-form equation

$$\left[\frac{\partial(hQ)}{\partial t} \right]_b + \nabla_b \cdot (hQ\mathbf{u}) = 0 \quad (52.70)$$

or equivalently using the specific thickness

$$\left[\frac{\partial(\hbar Q)}{\partial t} \right]_b + \nabla_b \cdot (\hbar Q\mathbf{u}) = 0. \quad (52.71)$$

52.2.4 Coordinate transforming vorticity and potential vorticity

As just shown, PV for a hydrostatic Boussinesq fluid can be written in the relatively simple form of a shallow water PV when choosing isopycnal coordinates. Here is a direct transformation from

Cartesian to isopyncal coordinates that also reveals this form

$$(\boldsymbol{\omega}^{\text{hy}} + f \hat{z}) \cdot \nabla b = -\frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + f \right) \frac{\partial b}{\partial z} \quad (52.72\text{a})$$

$$= \frac{\partial b}{\partial z} \left[f + \left(\frac{\partial v}{\partial x} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} \right) - \left(\frac{\partial u}{\partial y} - \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} \right) \right] \quad (52.72\text{b})$$

$$= \frac{\partial b}{\partial z} \left[f + \left(\frac{\partial v}{\partial x} \right)_b - \left(\frac{\partial u}{\partial y} \right)_b \right] \quad (52.72\text{c})$$

$$= \frac{f + (\partial v / \partial x)_b - (\partial u / \partial y)_b}{\partial z / \partial b} \quad (52.72\text{d})$$

$$= \frac{f + \tilde{\zeta}}{h} \quad (52.72\text{e})$$

$$= Q. \quad (52.72\text{f})$$

52.3 Isopycnal coordinate PV with irreversible processes

In Section 52.2, we considered the PV equation for an adiabatic, inviscid, hydrostatic, Boussinesq fluid using isopycnal vertical coordinates. We here extend to the case of friction in the momentum equation and diabatic heating in the buoyancy equation. We consider two ways to derive the governing equations. One is to convert the non-hydrostatic PV equation in Exercise 50.2 to isopycnal coordinates, after making the hydrostatic approximation. The second is to start from the equations of motion in isopycnal coordinates and derive the vorticity equation and then the PV equation.

52.3.1 Derivation method I

The equations of motion with diabatic heating and friction, written using isopycnal (or buoyancy) vertical coordinates, take the form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \mathbf{u} + \dot{b} \frac{\partial \mathbf{u}}{\partial b} + \mathbf{f} \wedge \mathbf{u} = -\nabla_b M + \mathbf{F} \quad (52.73\text{a})$$

$$\frac{\partial M}{\partial b} = -z \quad (52.73\text{b})$$

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) = -\frac{\partial(h \dot{b})}{\partial b} \quad (52.73\text{c})$$

$$\frac{D b}{D t} = \dot{b}. \quad (52.73\text{d})$$

As seen by the thickness equation (52.73c), the diabatic term, \dot{b} , affects transport across surfaces of constant buoyancy. Its specification depends on knowledge of heating sources/sinks in the fluid. We are not concerned with the details of this term, only that it is nonzero. Furthermore, we make use of the material time derivative operator (45.5b) to write the material form of the equations

$$\frac{D \mathbf{u}}{D t} + \mathbf{f} \wedge \mathbf{u} = -\nabla_b M + \mathbf{F} \quad (52.74\text{a})$$

$$\frac{\partial M}{\partial b} = -z \quad (52.74\text{b})$$

$$\frac{D h}{D t} + h \nabla_b \cdot \mathbf{u} = -h \frac{\partial \dot{b}}{\partial b}. \quad (52.74\text{c})$$

We start taking the curl, $\nabla_b \wedge$, of the momentum equation (52.73a), thus leading to the isopycnal vorticity equation

$$\left[\frac{\partial \tilde{\zeta}_a}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \tilde{\zeta}_a + \dot{b} \left[\frac{\partial \tilde{\zeta}_a}{\partial b} \right] = -\tilde{\zeta}_a \nabla_b \cdot \mathbf{u} + \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial b} \wedge \nabla_b \dot{b} + \nabla_b \wedge \mathbf{F} \right]. \quad (52.75)$$

The left hand side of equation (52.75) is the material time derivative of absolute vorticity (see equation (45.5b)), so that

$$\frac{D\tilde{\zeta}_a}{Dt} = -\tilde{\zeta}_a \nabla_b \cdot \mathbf{u} + \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial b} \wedge \nabla_b \dot{b} + \nabla_b \wedge \mathbf{F} \right]. \quad (52.76)$$

Now make use of the thickness equation in the material form (52.74c) to eliminate the convergence $\nabla_b \cdot \mathbf{u}$ on the right hand side, thus leading to

$$\frac{D\tilde{\zeta}_a}{Dt} - \frac{\tilde{\zeta}_a}{h} \left[\frac{Dh}{Dt} - h \frac{\partial \dot{b}}{\partial b} \right] = \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial b} \wedge \nabla_b \dot{b} + \nabla_b \wedge \mathbf{F} \right]. \quad (52.77)$$

Introducing the isopycnal potential vorticity

$$Q = \frac{\tilde{\zeta}_a}{h} = \frac{\tilde{\zeta} + f}{h} \quad (52.78)$$

leads to

$$h \frac{DQ}{Dt} = \zeta_a \frac{\partial \dot{b}}{\partial b} + \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial b} \wedge \nabla_b \dot{b} + \nabla_b \wedge \mathbf{F} \right]. \quad (52.79)$$

The diabatic terms can be written

$$\zeta_a \frac{\partial \dot{b}}{\partial b} + \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial b} \wedge \nabla_b \dot{b} \right] = \zeta_a \frac{\partial \dot{b}}{\partial b} + \dot{b} \frac{\partial \tilde{\zeta}}{\partial b} - \hat{z} \cdot \left[\nabla_b \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \quad (52.80a)$$

$$= \zeta_a \frac{\partial \dot{b}}{\partial b} + \dot{b} \frac{\partial \tilde{\zeta}_a}{\partial b} - \hat{z} \cdot \left[\nabla_b \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \quad (52.80b)$$

$$= \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} - \hat{z} \cdot \left[\nabla_b \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \quad (52.80c)$$

$$= \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} + \nabla_b \cdot \left[\hat{z} \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right], \quad (52.80d)$$

where we noted that the Coriolis parameter is independent of the buoyancy. The PV equation thus takes the material form

$$h \left[\frac{DQ}{Dt} \right] = \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} + \nabla_b \cdot \left[\hat{z} \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} - \hat{z} \wedge \mathbf{F} \right]. \quad (52.81)$$

Expanding the material time derivative into its components (45.5b), and making use of the thickness equation (52.73c), leads to the Eulerian flux-form equation

$$\left[\frac{\partial(h Q)}{\partial t} \right]_b + \nabla_b \cdot (h Q \mathbf{u}) + \frac{\partial(h Q \dot{b})}{\partial b} = \frac{\partial(\tilde{\zeta}_a \dot{b})}{\partial b} + \nabla_b \cdot \left[\hat{z} \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} - \hat{z} \wedge \mathbf{F} \right]. \quad (52.82)$$

Since $h Q = \tilde{\zeta}_a$, the ∂_b terms cancel, thus leaving

$$\left[\frac{\partial(h Q)}{\partial t} \right]_b = -\nabla_b \cdot \left[h Q \mathbf{u} - \hat{z} \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} + \hat{z} \wedge \mathbf{F} \right]. \quad (52.83)$$

52.3.2 Derivation method II

The alternative method to derive the PV equation in isopycnal coordinates is to start from the hydrostatic Boussinesq PV equation in geopotential vertical coordinates, and directly transform to isopycnal coordinates. Equation (??) provides an expression for material evolution of PV for a non-hydrostatic Boussinesq fluid. We can specialize this result to the hydrostatic case by reducing the non-hydrostatic vorticity to the hydrostatic form

$$\frac{DQ}{Dt} = \nabla \cdot \left[(f \hat{\mathbf{z}} + \boldsymbol{\omega}_{hy}) \dot{b} + b \nabla \wedge \mathbf{F} \right], \quad (52.84)$$

where the hydrostatic relative vorticity is given by

$$\boldsymbol{\omega}_{hy} = -\hat{\mathbf{x}} \frac{\partial v}{\partial z} + \hat{\mathbf{y}} \frac{\partial u}{\partial z} + \hat{\mathbf{z}} \left[\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right]. \quad (52.85)$$

The simplest term to transform to isopycnal coordinates is the curl of the horizontal friction vector

$$\hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{F}) = \hat{\mathbf{z}} \cdot \nabla \wedge (F^x, F^y, 0), \quad (52.86)$$

which takes the form

$$\hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{F}) = \kappa^{-1} \hat{\mathbf{z}} \cdot (\nabla_b \wedge \mathbf{F}) = -\kappa^{-1} \nabla_b \cdot (\hat{\mathbf{z}} \wedge \mathbf{F}). \quad (52.87)$$

The diabatic term requires some more work. Since the vorticity has zero divergence, the diabatic term can be written as

$$\nabla \cdot \left[(f \hat{\mathbf{z}} + \boldsymbol{\omega}_{hy}) \dot{b} \right] = (f \hat{\mathbf{z}} + \boldsymbol{\omega}_{hy}) \cdot \nabla \dot{b} \quad (52.88a)$$

$$= f \frac{\partial \dot{b}}{\partial z} - \frac{\partial v}{\partial z} \frac{\partial \dot{b}}{\partial x} + \frac{\partial u}{\partial z} \frac{\partial \dot{b}}{\partial y} + \left[\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right] \frac{\partial \dot{b}}{\partial z}. \quad (52.88b)$$

We now introduce horizontal derivatives on isopycnal surfaces according to the following relation (see Section 11.12.2)

$$\nabla_z = \nabla_b + \nabla_b b \left[\frac{\partial z}{\partial b} \right] \frac{\partial}{\partial z} \quad (52.89)$$

Doing so leads to

$$\begin{aligned}
 \nabla \cdot [(f \hat{z} + \boldsymbol{\omega}_{hy}) \dot{b}] &= f \frac{\partial \dot{b}}{\partial z} - \frac{\partial v}{\partial z} \frac{\partial \dot{b}}{\partial x} + \frac{\partial u}{\partial z} \frac{\partial \dot{b}}{\partial y} + \left[\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right] \frac{\partial \dot{b}}{\partial z} \\
 &= f \frac{\partial \dot{b}}{\partial z} - \frac{\partial v}{\partial z} \left(\left[\frac{\partial \dot{b}}{\partial x} \right]_b + \left[\frac{\partial b}{\partial x} \right]_z \frac{\partial z}{\partial b} \frac{\partial \dot{b}}{\partial z} \right) + \frac{\partial u}{\partial z} \left(\left[\frac{\partial \dot{b}}{\partial y} \right]_b + \left[\frac{\partial b}{\partial y} \right]_z \frac{\partial z}{\partial b} \frac{\partial \dot{b}}{\partial z} \right) \\
 &\quad + \frac{\partial \dot{b}}{\partial z} \left(\left[\frac{\partial v}{\partial x} \right]_b + \left[\frac{\partial b}{\partial x} \right]_z \frac{\partial z}{\partial b} \frac{\partial v}{\partial z} \right) - \frac{\partial \dot{b}}{\partial z} \left(\left[\frac{\partial u}{\partial y} \right]_b + \left[\frac{\partial b}{\partial y} \right]_z \frac{\partial z}{\partial b} \frac{\partial u}{\partial z} \right) \\
 &= f \frac{\partial \dot{b}}{\partial z} - \frac{\partial v}{\partial z} \left[\frac{\partial \dot{b}}{\partial x} \right]_b + \frac{\partial u}{\partial z} \left[\frac{\partial \dot{b}}{\partial y} \right]_b + \tilde{\zeta} \left[\frac{\partial \dot{b}}{\partial z} \right] \\
 &= \frac{\partial b}{\partial z} \left(\tilde{\zeta}_a \left[\frac{\partial \dot{b}}{\partial b} \right] - \frac{\partial v}{\partial b} \left[\frac{\partial \dot{b}}{\partial x} \right]_b + \frac{\partial u}{\partial b} \left[\frac{\partial \dot{b}}{\partial y} \right]_b \right) \\
 &= h^{-1} \left(\tilde{\zeta}_a \left[\frac{\partial \dot{b}}{\partial b} \right] + \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial b} \wedge \nabla_b \dot{b} \right] \right) \\
 &= h^{-1} \left(\tilde{\zeta}_a \left[\frac{\partial \dot{b}}{\partial b} \right] + \dot{b} \left[\frac{\partial \tilde{\zeta}_a}{\partial b} \right] + \nabla_b \cdot \left[\hat{z} \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \right) \\
 &= h^{-1} \left(\frac{\partial(\dot{b} \tilde{\zeta}_a)}{\partial b} + \nabla_b \cdot \left[\hat{z} \wedge \dot{b} \frac{\partial \mathbf{u}}{\partial b} \right] \right). \tag{52.90}
 \end{aligned}$$

To reach the penultimate step we noted that $\partial f / \partial b = 0$ so that we could form the derivative of the absolute vorticity. Bringing the pieces together leads to the PV equation (52.81) derived starting from the isopycnal version of the equations of motion.

Part IX

Balanced models

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 Chapter 29, the geostrophic balance is diagnostic so that it offers no means to compute the time evolution of the fluid. To obtain a prognostic equation requires going to next order in Rossby number within the asymptotic expansion. 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.

Perhaps the simplest balance model of geophysical relevance is the two-dimensional barotropic model. We develop the theory for this model in Chapter 53. The nuts and bolts of the next two chapters involve methods of scaling analysis and asymptotic analysis via perturbation series. In Chapter 54, we use these tools to derive equations for planetary geostrophy (PG) and quasi-geostrophy (QG) within the shallow water system, and then PG and QG for continuously stratified flows in Chapters 55 and 56 . PG and QG are useful theoretical models lending insights into different aspects of ocean and atmospheric fluid mechanics. In particular, PG is commonly used to study features of the large-scale laminar ocean circulation, and QG is ubiquitous in studies of both oceanic and atmospheric flows at or near the deformation radius including geostrophic turbulence. PG and QG represent two examples of *balance models*, in which knowledge of potential vorticity is sufficient to determine the stratification, pressure, and velocity.

53

Two-dimensional barotropic flows

A single layer of shallow water fluid is among the simplest models available for the study of rotating flows. However, the presence of divergence adds analytical complexity that may not be necessary physically nor desirable analytically. Additionally, we may choose to focus on the low frequency planetary (Rossby wave) modes rather than the divergent and higher frequency gravity wave mode (Section 44.3). For these reasons, we introduce in this chapter a fluid dynamical model in which the two-dimensional circulation has zero divergence. With no vertical stratification and a vanishing horizontal divergence, the vertical velocity is formally zero but in fact it is never even needed. Gravity waves are also absent, as the gravity wave speed has in effect been set to infinite.

The resulting set of equations forms the *two-dimensional barotropic model*, which is a well used theoretical model of geophysical flows. It furthermore offers a suitable starting point for the study of balanced flow. Vorticity is the primary dynamical field for the two-dimensional barotropic model, with its knowledge sufficient to fully determine the velocity.

- Rossby waves and wave-mean interactions
- Analytic vortex solutions

53.1	Governing equations	814
53.2	Vorticity equation	814
53.2.1	Deriving the vorticity equation	815
53.2.2	Potential vorticity	815
53.2.3	Poisson equation for the streamfunction	815
53.3	Pressure	816
53.3.1	Lid pressure	816
53.3.2	Diagnostic relation for the pressure	816
53.4	Further study	816
53.5	Exercises	817

53.1 Governing equations

The governing equations for 2d barotropic flow are the shallow water equations with the horizontal circulation assumed to be non-divergent

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z \phi \quad (53.1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (53.2)$$

where the pressure is normalized according to

$$\phi = p/\rho_0 \quad (53.3)$$

with ρ_0 a constant reference density, and where material evolution occurs with the two-dimensional non-divergent flow

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla. \quad (53.4)$$

With the horizontal flow $\mathbf{u} = (u, v, 0)$ non-divergent, we can introduce a streamfunction

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi \Rightarrow u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x}. \quad (53.5)$$

The advection operator can thus be written

$$\mathbf{u} \cdot \nabla \zeta = u \partial_x \zeta + v \partial_y \zeta \quad (53.6a)$$

$$= -\partial_y \psi \partial_x \zeta + \partial_x \psi \partial_y \zeta \quad (53.6b)$$

$$= \hat{\mathbf{z}} \cdot (\nabla \psi \wedge \nabla \zeta) \quad (53.6c)$$

$$\equiv J(\psi, \zeta), \quad (53.6d)$$

where J is the Jacobian operator

$$J(A, B) = \hat{\mathbf{z}} \cdot (\nabla A \wedge \nabla B) = \frac{\partial A}{\partial x} \frac{\partial B}{\partial y} - \frac{\partial A}{\partial y} \frac{\partial B}{\partial x}. \quad (53.7)$$

53.2 Vorticity equation

The vertical component of the relative vorticity is given by

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}) = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla^2 \psi. \quad (53.8)$$

There are many applications of 2d barotropic flow for a sphere. We are concerned here with the case of a tangent plane configuration in which the Coriolis parameter is given by the β -plane approximation

$$f = f_0 + \beta (y - y_0). \quad (53.9)$$

53.2.1 Deriving the vorticity equation

To form the vorticity equation, take the zonal derivative of the meridional momentum equation and meridional derivative of the zonal momentum equation (see equation (53.1)) to arrive at

$$\frac{\partial}{\partial t} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) + \frac{\partial}{\partial x} [\nabla \cdot (\mathbf{u} v)] - \frac{\partial}{\partial y} [\nabla \cdot (\mathbf{u} u)] + f \nabla \cdot \mathbf{u} + \beta v = 0. \quad (53.10)$$

Notice how the pressure gradient force dropped out, which is a key reason to study the vorticity equation. Straightforward manipulations, and use of the non-divergence condition $\nabla \cdot \mathbf{u} = 0$, then leads to the identity

$$\frac{\partial}{\partial x} [\nabla \cdot (\mathbf{u} v)] - \frac{\partial}{\partial y} [\nabla \cdot (\mathbf{u} u)] = \mathbf{u} \cdot \nabla \zeta, \quad (53.11)$$

so that the 2d vorticity equation takes the form

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta + \beta v = 0, \quad (53.12)$$

which takes on the material evolution form

$$\frac{D\zeta}{Dt} = -\beta v. \quad (53.13)$$

53.2.2 Potential vorticity

For 2d non-divergent flow, the potential vorticity $q = \zeta + f$ is the same as absolute vorticity. Since f is time independent, we can write the vorticity equation in the alternative form

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot (\mathbf{u} q). \quad (53.14)$$

Similarly, we have the potential vorticity equation

$$\frac{\partial q}{\partial t} = -\nabla \cdot (\mathbf{u} q), \quad (53.15)$$

or the material conservation form

$$\frac{Dq}{Dt} = 0. \quad (53.16)$$

53.2.3 Poisson equation for the streamfunction

Given initial and boundary conditions, the barotropic vorticity equation (53.14) allows us to determine the evolution of vorticity. We can in turn invert the Poisson equation (see Section 5.4 for discussion of the Poisson equation)

$$\nabla^2 \psi = \zeta \quad (53.17)$$

to determine the streamfunction and then the velocity field $\mathbf{u} = \hat{z} \wedge \nabla \psi$. Hence, time integration of the barotropic potential vorticity equation is sufficient to fully specify time evolution of the 2d barotropic flow field. We do not need to explicitly determine pressure to determine the flow.

53.3 Pressure

The 2d barotropic system is so highly constrained as to be almost be non-fluid like. Notably, pressure forces are present yet we do not need their expression to determine the flow evolution. Furthermore, in the absence of two-dimensional divergence, the free surface is flat, as if there was a flat rigid lid placed on the surface. Indeed, this *rigid lid* approximation is commonly employed for studies of large-scale ocean circulation. But in a homogeneous fluid, how can we generate pressure variations in the absence of free surface gradients?

53.3.1 Lid pressure

Pressure gradients do exist in the barotropic fluid even though the surface is flat, with these pressure gradients generated through gradients in the *lid pressure*. To keep the free surface flat requires a lid pressure, with gradients in the lid pressure driving flow. Furthermore, as the flow is homogeneous, pressure applied at the surface is transmitted throughout the body of the fluid to drive the flow.

53.3.2 Diagnostic relation for the pressure

Even though we do not need the pressure to time step velocity, it is of interest to determine pressure to better understand the dynamics. We derive an equation for the pressure by using the two-dimensional non-divergence property of the horizontal flow and then developing a diagnostic relation for the pressure. We can eliminate the time derivative from equation (53.1) 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

$$-\nabla^2\phi = \partial_x[\nabla \cdot (\mathbf{u} u)] + \partial_y[\nabla \cdot (\mathbf{u} v)] - f\zeta + \beta u. \quad (53.18)$$

Given sufficient boundary conditions, this elliptic partial differential equation (Section 5.4) can be inverted to find the pressure field. In Exercise 28.2, we encountered a similar elliptic problem for the pressure in a three-dimensional incompressible fluid.

Numerically inverting an elliptic operator 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 fall into shallow minima, making it difficult to find the true solution. This algorithmic complexity is one reason numerical barotropic models are less commonly used for realistic numerical experimentation than the more general shallow water models.

53.4 Further study

In exercise 53.2 we develop some integral properties of the 2d barotropic system. Chapter 3 of [McWilliams \(2006\)](#) provides further analysis of this system, offering an exploration of analytical solutions associated with vortices. Some work with the 2d barotropic system was motivated by studies of coherent vortex structures, such as those found by the simulations documented in [McWilliams \(1984\)](#).

[Bryan \(1969\)](#) provided the first working numerical algorithm to determine the ocean general circulation model. Bryan's method made use of the rigid lid approximation so that the depth integrated circulation is assumed to be non-divergent. 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 \(2004\)](#)).

53.5 Exercises

EXERCISE 53.1: 2D BAROTROPIC SYSTEM AND GRAVITY WAVES

Are there gravity waves for the 2d barotropic system described in Section 53.1. Why? Hint: recall the discussion of gravity waves for the shallow water system in Section 44.3. A one-sentence answer is sufficient.

EXERCISE 53.2: INTEGRAL PROPERTIES OF THE INVISCID 2D BAROTROPIC MODEL

In this exercise, we establish some global conservation properties for inviscid two-dimensional non-divergent flow on a β -plane. We assume that the geometry is a flat plane defined over a region \mathcal{A} . The region can either be infinite, in which case all fields decay to zero at infinity, or a finite domain surrounded by static material boundaries. 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

$$\mathcal{K} = \frac{1}{2} \int_{\mathcal{A}} \mathbf{u} \cdot \mathbf{u} \, dA, \quad (53.19)$$

where the horizontal integral extends over the full fluid domain \mathcal{A} .

- (b) Show that the domain integrated vorticity (equal also to the circulation) is constant in time

$$\mathcal{C} = \int_{\mathcal{A}} \zeta \, dA. \quad (53.20)$$

- (c) Show that the domain integrated enstrophy is constant in time for f -plane motion ($\beta = 0$)

$$\mathcal{Z}^{(\zeta)} = \int_{\mathcal{A}} \zeta^2 \, dA. \quad (53.21)$$

- (d) Show that the domain integrated potential enstrophy is constant in time even with $\beta \neq 0$

$$\mathcal{Z}^{(q)} = \int_{\mathcal{A}} q^2 \, dA. \quad (53.22)$$

EXERCISE 53.3: CIRCULATION IN A 2D BAROTROPIC FLOW

Consider a two-dimensional barotropic flow on a β -plane in the presence of a biharmonic friction operator, where the governing vorticity equation is

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta + \beta y) = -\nu \nabla^4 \zeta, \quad (53.23)$$

with $\nu > 0$ a constant biharmonic viscosity with dimensions of $L^4 T^{-1}$. Show that the circulation around a fixed area in the fluid evolves according to

$$\frac{d\mathcal{C}}{dt} = - \oint \left[\psi \frac{\partial q}{\partial s} + \nu \frac{\partial (\nabla^2 \zeta)}{\partial n} \right] ds, \quad (53.24)$$

where s is the arc-length along the boundary of the area and n is a coordinate normal to the boundary.

EXERCISE 53.4: DYNAMICS OF VORTICITY GRADIENTS

For many purposes it is of interest to develop equations describing the evolution of scalar gradients. We developed a general expression in Exercise 16.1. Here, we develop a similar equation for the gradient of relative vorticity in a two-dimensional barotropic flow. For this purpose, consider the inviscid barotropic vorticity equation on an f -plane

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) = 0. \quad (53.25)$$

- (a) Show that the material evolution of the vorticity gradient is given by

$$\frac{D(\nabla \zeta)}{Dt} = -J(\nabla \psi, \zeta). \quad (53.26)$$

- (b) Show that the material evolution of the squared vorticity gradient is given by

$$\frac{D|\nabla \zeta|^2}{Dt} = 2 J(\zeta, \nabla \psi) \cdot \nabla \zeta. \quad (53.27)$$

EXERCISE 53.5: ANGULAR MOMENTUM

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 43.5.

As in Section 43.5.1, the relative angular momentum for a region of fluid is given by

$$\mathbf{L} = \int dA \int (\mathbf{x} \wedge \mathbf{v}) \rho dz, \quad (53.28)$$

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 solid body. For a barotropic fluid of constant density and constant thickness, and correspondingly a zero vertical velocity, the relative angular momentum reduces to

$$\mathbf{L} = \rho H \int_{\mathcal{A}} (\mathbf{x} \wedge \mathbf{u}) dA, \quad (53.29)$$

with \mathbf{u} the horizontal velocity and \mathcal{A} the horizontal region. For barotropic motion on a tangent plane we are interested in the vertical component of the relative angular momentum

$$L^z = \rho H \int_{\mathcal{A}} \hat{\mathbf{z}} \cdot (\mathbf{x} \wedge \mathbf{u}) dA. \quad (53.30)$$

Show for a simply connected and bounded region, L^z can be written

$$L^z = 2 \rho H \int_{\mathcal{A}} (\psi_b - \psi) dA \quad (53.31)$$

where ψ is the streamfunction satisfying $\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi$, and ψ_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 Section 19.4.2 that the streamfunction equals to a spatial constant when evaluated along the domain boundary.

54

Shallow water planetary and quasi-geostrophy

READER'S GUIDE FOR THIS CHAPTER

We here derive the mechanical equations for planetary geostrophy (PG) and quasi-geostrophy (QG) within the shallow water fluid system. The derivation requires elements from asymptotic analysis, with salient details introduced here. This chapter is largely technical in nature, with certain details put to use in the continuously stratified systems discussed in Chapters 55 and 56.

54.1	Scaling analysis and the Buckingham-II theorem	820
54.2	Shallow water equations	821
54.2.1	Dimensional scales	821
54.2.2	Physical dimensions	822
54.2.3	Number of non-dimensional parameters	822
54.2.4	Choosing the non-dimensional parameters	822
54.2.5	Assumed values for the non-dimensional parameters	824
54.2.6	Deformation radius and the free surface undulation scale	825
54.2.7	Non-dimensional shallow water equations	825
54.3	Shallow water planetary geostrophy	827
54.4	Shallow water quasi-geostrophy	827
54.4.1	Quasi-geostrophic scaling	828
54.4.2	Outlining the asymptotic method	828
54.4.3	Zeroth order asymptotic equations	829
54.4.4	First order asymptotic equations	830
54.4.5	Dimensional equations and quasi-geostrophic PV	831
54.5	Exercises	832

54.1 Scaling analysis and the Buckingham-II theorem

Scaling analysis is ubiquitous in physics, with the Buckingham-II theorem providing a useful framework for scaling. This theorem states that the number of dimensionless parameters in a physical system is a function of the number of dimensional parameters or scales K (e.g., scales for the velocity, rotation rate, pressure force, friction force, gravitational acceleration) and the number of physical dimensions R (e.g., time, length, mass). Precisely, Buckingham-II states that the number of dimensionless parameters is

$$N_{\text{dimensionless}} = K - R. \quad (54.1)$$

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, 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. We focus here on the regime of large-scale atmospheric and oceanic flow where the 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.

54.2 Shallow water equations

A single-layer of inviscid shallow water fluid is governed by the momentum and thickness equations (Chapter 42)

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad (54.2a)$$

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = 0, \quad (54.2b)$$

where (see Figure 42.1)

$$\eta = \eta_b + h = H + \Delta\eta. \quad (54.3)$$

54.2.1 Dimensional scales

We identify nine dimensional parameters for the shallow water system.

- LENGTH SCALES

- ★ H = depth scale of the fluid; for the shallow water system, H is the area average fluid thickness (see Figure 42.1).
- ★ L = horizontal/lateral length scale of motions under consideration. We assume both horizontal directions to have the same length scale. This assumption is not necessarily valid on a rotating planet, where zonal (east-west) length scales can be longer than meridional (north-south) scales. Nonetheless, this choice does not preclude the dynamical emergence of anisotropic length scales. Indeed, by not *a priori* introducing anisotropic length scales, we ensure that the emergence of anisotropy naturally arises from the dynamics.
- ★ R = 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 (β -plane).
- ★ \mathcal{H} = scale for free surface height undulations, $\Delta\eta$.
- ★ \mathcal{B} = scale for undulations of the bottom topography, η_b .

- VELOCITY SCALES

- ★ U = velocity scale for fluid particle motion via advection; i.e., the speed for horizontal currents or winds.
- ★ c = wave speed scale. For the shallow water model, the wave speed scale is given by the shallow water gravity wave

$$c = \sqrt{g H}. \quad (54.4)$$

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.

- BODY FORCES: There are two body forces acting on the fluid; one from gravity and one from Coriolis.

- ★ g = gravitational acceleration.
- ★ f = Coriolis frequency.

If we were interested in other forces, such as electromagnetic or frictional forces, then we would have other dimensional parameters. But for our purposes, nine is all we are interested in for a single layer of inviscid shallow water fluid.

54.2.2 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 so that mass is described by area times height

$$M = \int \rho dV [\equiv] L^2 H \rho. \quad (54.5)$$

Relatedly, we are unconcerned with the fluid density since it is uniform and does not explicitly appear in any of the governing equations.

54.2.3 Number of non-dimensional parameters

The Buckingham-II theorem then says we have

$$N_{\text{dimensionless}} = 9 - 2 = 7 \quad (54.6)$$

non-dimensional parameters. What if we incorrectly count the physical dimensions or the dimensional scales/parameters? Fortunately, the process of determining the non-dimensional parameters is largely self-correcting. Namely, in the process of non-dimensionalizing the shallow water equations, the seven non-dimensional parameters will arise as part of the analysis. Hence, making use of Buckingham-II is useful but it is not essential. If one left out a physical dimension or a physical parameter, then it would appear somewhere in the subsequent analysis, often not until near the end where something mathematically or physically inconsistent appears. One must always be cognizant of the need to self-correct when moving onward with the analysis.

54.2.4 Choosing the non-dimensional parameters

There is no unique choice for the non-dimensional parameters. Our choice is guided by experience and interest.

1. VERTICAL TO HORIZONTAL ASPECT RATIO: The ratio of the vertical scale to the horizontal scale defines the aspect ratio

$$\delta_{\text{vertical/horizontal}} = \frac{\text{vertical depth scale}}{\text{horizontal length scale}} = \frac{H}{L}. \quad (54.7)$$

2. RATIO OF HORIZONTAL SCALE TO PLANETARY SCALE: The ratio of the lateral length scale to the planet radius is

$$\delta_{\text{horizontal/planet}} = \frac{\text{lateral length scale}}{\text{planetary length scale}} = \frac{L}{R}. \quad (54.8)$$

3. RATIO OF FREE SURFACE UNDULATION TO DEPTH: The ratio of the free surface undulation scale to the vertical depth scale is

$$\delta_{\text{free surface/depth}} = \frac{\text{free surface undulation scale}}{\text{vertical depth scale}} = \frac{\mathcal{H}}{H}. \quad (54.9)$$

4. RATIO OF BOTTOM TOPOGRAPHY UNDULATION TO FREE SURFACE UNDULATION: The ratio of the bottom topography undulation scale to the free surface undulation scale is

$$\delta_{\text{bottom/free surface}} = \frac{\text{bottom topography undulation scale}}{\text{free surface undulation scale}} = \frac{\mathcal{B}}{\mathcal{H}}. \quad (54.10)$$

5. FROUDE NUMBER: The Froude number is the ratio of the fluid particle speed to the wave speed. For the shallow water system, this ratio is

$$Fr = \frac{U}{c} = \frac{U}{\sqrt{g H}}. \quad (54.11)$$

Note that the Froude number is not directly utilized in the following, though its introduction is useful in other contexts.

6. ROSSBY NUMBER: The Rossby number is the ratio of the fluid particle acceleration scale to the Coriolis acceleration

$$Ro = \frac{\text{parcel acceleration}}{\text{Coriolis acceleration}}. \quad (54.12)$$

The particle acceleration scale is determined by the local time tendency and advection

$$\text{particle acceleration} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \quad (54.13a)$$

$$\sim \frac{U}{T} + \frac{U^2}{L}. \quad (54.13b)$$

We assume that the time scale is determined by advection, so that

$$T \sim \frac{L}{U} \Rightarrow \frac{U^2}{L} = \frac{U}{T}, \quad (54.14)$$

in which case the Rossby number is given by

$$Ro = \frac{1}{f T} = \frac{U}{f L}. \quad (54.15)$$

Another interpretation for the Rossby number is the ratio of the relative vorticity to the planetary vorticity

$$Ro = \frac{\text{relative vorticity}}{\text{planetary vorticity}} \quad (54.16)$$

With the relative vorticity scaling as U/L and the planetary vorticity scaling as f , we recover the expression (54.15) for the Rossby number.

7. GEOSTROPHIC NUMBER: We define the geostrophic number as the ratio of the Coriolis acceleration to the pressure gradient acceleration¹

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}}. \quad (54.17)$$

The Coriolis acceleration scales as

$$\text{Coriolis acceleration} \sim f U \quad (54.18)$$

whereas the pressure gradient acceleration, $-g \nabla \eta$, scales as

$$\text{pressure gradient acceleration} \sim \frac{g \mathcal{H}}{L}, \quad (54.19)$$

so that

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}} = \frac{f U}{(g/L) \mathcal{H}}. \quad (54.20)$$

¹The geostrophic number is generally not introduced in the literature, since it will later be assumed equal to unity. We find it pedagogical to introduce it in order to enumerate the seven non-dimensional parameters available for the shallow water system.

54.2.5 Assumed values for the non-dimensional parameters

We now enumerate the assumed values for the non-dimensional parameters.

1. SMALL VERTICAL TO HORIZONTAL ASPECT RATIO: The aspect ratio is generally small for large-scale atmospheric and oceanic fluid systems

$$\delta_{\text{vertical/horizontal}} \ll 1. \quad (54.21)$$

This assumption was made when formulating the shallow water system, which is based on hydrostatic balance (see Section 42.1). We thus retain this assumption as we further scale the shallow water system.

2. SMALL OR ORDER ONE RATIO OF HORIZONTAL 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

$$\delta_{\text{horizontal/planet}} \ll 1 \quad \text{quasi-geostrophy} \quad (54.22a)$$

$$\delta_{\text{horizontal/planet}} \sim 1 \quad \text{planetary geostrophy.} \quad (54.22b)$$

3. RATIO OF FREE SURFACE UNDULATION TO DEPTH: This ratio will be implied by other scaling assumptions along with the dynamical equations.
4. RATIO OF BOTTOM TOPOGRAPHY UNDULATION TO FREE SURFACE UNDULATION: We assume that gradients in the bottom topography are small relative to gradients in the surface height, which can be assured if the scales for the bottom topography are much smaller than scales for the surface height

$$\delta_{\text{bottom/free surface}} \ll 1. \quad (54.23)$$

5. FROUDE NUMBER: The Froude number is implied by sizes assumed for the other non-dimensional numbers.
6. SMALL ROSSBY NUMBER: The Rossby number is assumed small

$$Ro = \frac{U}{f L} \ll 1, \quad (54.24)$$

which means that the Coriolis acceleration is a leading order term in the horizontal momentum equation (54.2a).

7. UNIT GEOSTROPHIC NUMBER: The geostrophic number is assumed to be order unity

$$Ge \sim 1. \quad (54.25)$$

This assumption means that the Coriolis acceleration and pressure gradient acceleration scale together

$$f U \sim (g/L) \mathcal{H}. \quad (54.26)$$

Making use of the momentum equation (54.2a), we see that this scaling is consistent only so long as the Rossby number is small, $Ro \ll 1$. Furthermore, this scaling constrains the scale of the free surface undulation, \mathcal{H} , as we discuss in Section 54.2.6.

54.2.6 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 geostrophy number. For this purpose, start from the geostrophic scaling of Coriolis and pressure gradient accelerations, (54.26), to express the free surface undulation scale according to

$$\Delta\eta \sim \mathcal{H} = \frac{f U L}{g} = Ro \frac{f^2 L^2}{g} = Ro H \frac{f^2 L^2}{g H} = Ro H \left[\frac{L}{L_d} \right]^2. \quad (54.27)$$

In the final equality we introduced the deformation radius

$$L_d = \frac{\sqrt{g H}}{f}. \quad (54.28)$$

We encountered the deformation radius when discussing geostrophic adjustment in Section 44.5. Furthermore, note that the deformation radius is the scale whereby the relative vorticity and the surface height (vortex stretching) make equal contributions to the potential vorticity (see page 92 of *Pedlosky (1987)*). The deformation radius decreases toward the poles, so that rotational effects are felt by smaller scales in the high latitudes than in the tropics. We can use L_d to rewrite the Froude number as the ratio of the advection speed to the rotational speed

$$Fr = \frac{U}{\sqrt{g H}} = \frac{U}{f L_d} = Ro \frac{L}{L_d}. \quad (54.29)$$

Furthermore, the squared ratio of the deformation radius to the lateral length scale is termed the Burger number

$$F^{-1} = Bu = \left[\frac{L_d}{L} \right]^2. \quad (54.30)$$

Use of the Burger number allows us to write the Froude number as

$$Fr = \frac{Ro}{\sqrt{Bu}} = Ro \sqrt{F} \quad (54.31)$$

and the free surface height undulation scale as

$$\mathcal{H} = H Ro \left[\frac{L}{L_d} \right]^2 = H \frac{Ro}{Bu} = H Ro F = H \frac{Fr^2}{Ro}. \quad (54.32)$$

Hence, the ratio of the free surface undulations to the depth scale is given by

$$\delta_{\text{free surface/depth}} = \frac{\mathcal{H}}{H} = Ro \left[\frac{L}{L_d} \right]^2 = \frac{Ro}{Bu} = Ro F = \frac{Fr^2}{Ro}. \quad (54.33)$$

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.

54.2.7 Non-dimensional shallow water equations

To non-dimensionalize the shallow water equations, introduce non-dimensional variables, denoted by a hat, according to

$$t = T \hat{t} \quad (x, y) = L (\hat{x}, \hat{y}) \quad \partial_t = T^{-1} \partial_{\hat{t}} \quad \nabla = L^{-1} \hat{\nabla} \quad (54.34a)$$

$$(u, v) = U (\hat{u}, \hat{v}) \quad f = f_o \hat{f} \quad \Delta\eta = \mathcal{H} \hat{\eta} \quad \eta = H + \Delta\eta = H + \mathcal{H} \hat{\eta}. \quad (54.34b)$$

where f_o is a typical scale for the Coriolis parameter. Importantly, we assume that the non-dimensional variables (the hat-variables) are order unity. Also note that the non-dimensional hatted variables should not be confused with the unit vector notation used elsewhere in this book.

Non-dimensional momentum equation

Introducing the above variables into the shallow water momentum equation (54.2a) renders

$$\frac{U}{T} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + \frac{U^2}{L} (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{\mathbf{u}} + f_o U (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = - \frac{g \mathcal{H}}{L} \hat{\nabla} \hat{\eta}. \quad (54.35)$$

As before, we assume the time scale is given by the advection time

$$T = \frac{L}{U} = \frac{1}{R_o f_0}, \quad (54.36)$$

so that dividing by $f_0 U$ leads to

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{\mathbf{u}} \right] + (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = - \left[\frac{g H}{f_o L U} \frac{Ro}{Bu} \right] \hat{\nabla} \hat{\eta}, \quad (54.37)$$

where we set $\mathcal{H} = H (Ro/Bu)$ according to equation (54.33). We reduce the factor on the right hand side according to

$$\frac{g H}{f_o L U} \frac{Ro}{Bu} = \frac{g H}{f_o L U} \frac{U}{f_o L} \frac{L^2}{L_d^2} = \frac{g H}{f_o L U} \frac{U}{f_o L} \frac{L^2 f_o^2}{g H} = 1. \quad (54.38)$$

Hence, the non-dimensional inviscid shallow water momentum equation takes on the rather elegant form

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{\mathbf{u}} \right] + \hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = - \hat{\nabla} \hat{\eta}. \quad (54.39)$$

Introducing the non-dimensional material time derivative

$$\frac{D}{D\hat{t}} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}} \cdot \hat{\nabla} \quad (54.40)$$

brings the momentum equation to

$$Ro \frac{D\hat{\mathbf{u}}}{D\hat{t}} + \hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = - \hat{\nabla} \hat{\eta}. \quad (54.41)$$

We see that the momentum 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.

Non-dimensional continuity (thickness) equation

The continuity equation (54.2b) can be written as

$$\frac{\partial \Delta \eta}{\partial t} + (H + \Delta \eta - \eta_b) \nabla \cdot \mathbf{u} + (\mathbf{u} \cdot \nabla) (H + \Delta \eta - \eta_b) = 0, \quad (54.42)$$

where we wrote (see Figure 42.1)

$$h = H + \Delta \eta - \eta_b. \quad (54.43)$$

Our assumed scaling for the bottom topography, (54.23), allows us to drop η_b from this equation. Also note that H is a constant and so it has a zero gradient. Hence, the continuity equation takes the form

$$\frac{\partial \Delta \eta}{\partial t} + (H + \Delta \eta) \nabla \cdot \mathbf{u} + (\mathbf{u} \cdot \nabla) \Delta \eta = 0. \quad (54.44)$$

Introduction of the dimensionless variables leads to

$$H \frac{Ro}{Bu} \left[\frac{1}{T} \frac{\partial}{\partial \hat{t}} + \frac{U}{L} \hat{\mathbf{u}} \cdot \hat{\nabla} \right] \hat{\eta} + \frac{U H}{L} \left[1 + \frac{Ro}{Bu} \hat{\eta} \right] \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (54.45)$$

With the time scale set by advection, $T = L/U$, we have

$$\frac{Ro}{Bu} \frac{D\hat{\eta}}{Dt} + \left[1 + \frac{Ro}{Bu} \hat{\eta} \right] \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (54.46)$$

Equivalently, introducing the inverse Burger number, $F = Bu^{-1}$ renders

$$F Ro \frac{D\hat{\eta}}{Dt} + (1 + F Ro \hat{\eta}) \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (54.47)$$

54.3 Shallow water planetary geostrophy

We now make use of the non-dimensional equations derived in Section 54.2.7 to derive the dynamical equations for planetary geostrophy and quasi-geostrophy. For planetary geostrophy, we drop the parcel acceleration term from the momentum equation (54.41), given that it is order Rossby number smaller than the Coriolis and pressure gradient accelerations. This assumption leads to the geostrophic balance

$$\hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla} \hat{\eta}. \quad (54.48)$$

For the mass continuity equation (54.47), we must make an assumption about the ratio of the Rossby number to the Burger number. For planetary geostrophy, we assume

$$Ro \sim Bu. \quad (54.49)$$

Hence, the full mass continuity equation is retained; no terms are dropped. Since the Rossby number is small, $Ro \sim Bu = (L_d/L)^2 \ll 1$ means the horizontal length scale is much larger than the deformation radius

$$L \gg L_d. \quad (54.50)$$

The dimensional planetary geostrophic equations take the form

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad (54.51a)$$

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} \quad (54.51b)$$

$$\eta = \eta_b + h. \quad (54.51c)$$

As shown in Exercise 54.1, these equations are equivalent to

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad \frac{DQ}{Dt} = 0 \quad \text{with} \quad Q = \frac{f}{h}. \quad (54.52)$$

We encountered this form for the potential vorticity in Section 47.3, in which the shallow water PV, $(f + \zeta)/h$, is approximated by f/h in the small Rossby number limit.

54.4 Shallow water quasi-geostrophy

In this section we develop the quasi-geostrophic equations for a single shallow water fluid layer. We make use of asymptotic methods to derive these equations using the Rossby number as the small parameter.

54.4.1 Quasi-geostrophic scaling

Quasi-geostrophic scaling is based on the following assumptions.

1. $Ro \ll 1$, which is fundamental to geostrophic scaling.
2. $T \sim L/U$; that is, the time scale is determined by advection, which is how time has scaled throughout this chapter.
3. $F^{-1} = Bu \sim 1$, which means that the horizontal scales of motion are on the order of the deformation radius, $L \sim L_d$. From equation (54.33), it furthermore means that undulations of the free surface height scale according to the Rossby number: $\mathcal{H} \sim H Ro$, meaning that the free surface height undulations are small. Finally, when developing the vorticity equation in Section 54.4.4, a unit Froude number means that the advection of vorticity contributes roughly the same as vortex stretching (see equation (54.68)).
4. $|\beta L| \ll |f_o|$, which means that the Coriolis frequency does not vary much from its central value.

The third and fourth assumptions are distinct from planetary geostrophy.

54.4.2 Outlining the asymptotic method

To derive the quasi-geostrophic shallow water model, we employ an asymptotic expansion in the Rossby number and stop at the first nontrivial order. For this purpose, recall the non-dimensional momentum and continuity equations from Section 54.2.7, and make use of the assumed $Bu \sim 1$ scaling

$$Ro \frac{D\hat{\mathbf{u}}}{Dt} + (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = -\hat{\nabla}\hat{\eta} \quad (54.53a)$$

$$F Ro \frac{D\hat{\eta}}{Dt} + (1 + F Ro \hat{\eta}) \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (54.53b)$$

We maintain the dimensionless parameter

$$F \equiv Bu^{-1} \quad (54.54)$$

in the continuity equation for later discussion. Again, for QG it is assumed to have unit scale and so will not play a role in the following asymptotic expansion.

Asymptotic methods are ideally suited for non-dimensional equations since we can unambiguously determine scales via the size of non-dimensional parameters. We here take the Rossby number to be small. It therefore makes sense to perform an asymptotic expansion of the prognostic fields in terms of the Rossby number. There are three prognostic fields, $\hat{u}, \hat{v}, \hat{\eta}$, in which we assume can be written

$$\hat{u} = \hat{u}_0 + Ro \hat{u}_1 + Ro^2 \hat{u}_2 + \dots \quad (54.55a)$$

$$\hat{v} = \hat{v}_0 + Ro \hat{v}_1 + Ro^2 \hat{v}_2 + \dots \quad (54.55b)$$

$$\hat{\eta} = \hat{\eta}_0 + Ro \hat{\eta}_1 + Ro^2 \hat{\eta}_2 + \dots \quad (54.55c)$$

In addition to expanding the prognostic variables, we expand the non-dimensional Coriolis parameter in terms of the Rossby number

$$\hat{\mathbf{f}} = \frac{\mathbf{f}}{f_0} = \frac{(f_0 + \beta y) \hat{\mathbf{z}}}{f_0} \equiv (\hat{f}_0 + Ro \hat{\beta} \hat{y}) \hat{\mathbf{z}}. \quad (54.56)$$

In this equation, $\hat{\mathbf{z}}$ is the unit vector in the vertical, the time scaling in equation (54.36) renders

$$\hat{\beta} \hat{y} = \frac{\beta y}{R_o f_0} = T \beta y, \quad (54.57)$$

and

$$\hat{f}_0 = \frac{f_0}{f_0} = 1. \quad (54.58)$$

Although $\hat{f}_0 = 1$, it is useful to retain this term as a placeholder. The Coriolis expression (54.56), in particular the assumed scaling (54.57), is motivated by assuming the horizontal scales of motion are on the same order as the deformation radius, and that the Coriolis frequency does not vary much from its central value. The full spherical dependence of the Coriolis parameter has been reduced down to a mere constant plus a linear term (i.e., the β -plane approximation discussed in Section 27.2).

The practical goal of asymptotic analysis is to develop a closed set of prognostic equations for functions appearing in the asymptotic expansions (54.55a)-(54.56). 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 requires more complex algebraic manipulations. Hence, pursuit of higher order expansions should be motivated by first determining that the lower order equation set remains physically lacking in something desired by the analyst.

54.4.3 Zeroth order asymptotic equations

At this point the setup has been done, the philosophy exposed, so we are ready to enter the “turn the crank” stage. To do so, we insert the asymptotic expansions (54.55a)-(54.55c) into the non-dimensional partial differential equations (54.53a) and (54.53b). There is a need to pay careful attention to detail here while organizing terms according to the Rossby number power. Since Ro is arbitrarily small, and all non-dimensional fields are order unity regardless their order, the only means to maintain self-consistency is that all terms of equal order in Rossby number balance. This observation is basic to asymptotic methods.

Again, our overall 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 will offer us just the geostrophic balance, which has no prognostic value. Hence, we need to go at least to order Ro^1 , and hopefully no further as the algebraic tedium increases with order. With that anticipation and hope, we only keep track of terms of order Ro^0 and Ro^1 , in which the momentum and continuity equations become

$$Ro \frac{D\hat{\mathbf{u}}_0}{Dt} + (\hat{f}_0 + Ro \hat{\beta} \hat{y}) \hat{\mathbf{z}} \wedge (\hat{\mathbf{u}}_0 + Ro \hat{\mathbf{u}}_1) = -\hat{\nabla}(\hat{\eta}_0 + Ro \hat{\eta}_1) \quad (54.59a)$$

$$F Ro \frac{D\hat{\eta}_0}{Dt} + \hat{\nabla} \cdot \hat{\mathbf{u}}_0 + Ro \hat{\nabla} \cdot \hat{\mathbf{u}}_1 + F Ro \hat{\eta}_0 \hat{\nabla} \cdot \hat{\mathbf{u}}_0 = 0. \quad (54.59b)$$

Terms balancing at order Ro^0 are given by

$$\hat{f}_0 \hat{\mathbf{z}} \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla} \hat{\eta}_0 \quad (54.60a)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}}_0 = 0, \quad (54.60b)$$

with the momentum balance reduced to the f -plane geostrophic balance. Fortunately, these two equations are self-consistent, since the curl of the f -plane geostrophic balance (54.60a) leads to the non-divergence condition (54.60b). Given the non-divergence condition (54.60b), the zeroth order velocity field can be written in terms of a streamfunction

$$\hat{u}_0 = -\frac{\partial \hat{\psi}_0}{\partial \hat{y}} \quad \hat{v}_0 = \frac{\partial \hat{\psi}_0}{\partial \hat{x}} \quad \hat{\zeta}_0 = \hat{\nabla}^2 \hat{\psi}_0, \quad (54.61)$$

where the zeroth order streamfunction is the ratio of the zeroth order surface height to zeroth order Coriolis parameter

$$\hat{\psi}_0 = \frac{\hat{\eta}_0}{\hat{f}_0}, \quad (54.62)$$

and we introduced the zeroth order vorticity, $\hat{\zeta}_0$, which will appear in subsequent steps.

54.4.4 First order asymptotic equations

The zeroth order equations do not render a prognostic equation, for which we need to consider equations at order Ro^1

$$\frac{D_0 \hat{u}_0}{D\hat{t}} + \hat{f}_0 \hat{z} \wedge \hat{u}_1 + \hat{\beta} \hat{y} \hat{z} \wedge \hat{u}_0 = -\hat{\nabla} \hat{\eta}_1 \quad (54.63a)$$

$$F \frac{D_0 \hat{\eta}_0}{D\hat{t}} + \hat{\nabla} \cdot \hat{u}_1 = 0. \quad (54.63b)$$

At this order, the material time derivative makes use of *only* the zeroth order geostrophic horizontal velocity

$$\frac{D_0}{D\hat{t}} = \frac{\partial}{\partial \hat{t}} + \hat{u}_0 \cdot \hat{\nabla}. \quad (54.64)$$

The set of first order equations (54.63a) and (54.63b) are not closed, because evolution of zeroth order terms are functions of first order terms. However, the first order terms can be eliminated using two steps. First, we produce the vorticity equation from the momentum equation; second, we combine the vorticity equation and continuity equation. Although the details are specific to shallow water quasi-geostrophy, similar steps are frequently encountered in other geophysical fluid dynamical systems.

Taking the curl of the momentum equation (54.63a) eliminates the pressure gradient, $\hat{\nabla} \hat{\eta}_1$, thus producing the vorticity equation

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{u}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \hat{\nabla} \cdot \hat{u}_1. \quad (54.65)$$

To this order, vortex stretching on the right hand side arises just from the planetary vorticity, since relative vorticity stretching occurs at a higher order. Note that since $\hat{\beta} \hat{y}$ is time independent, we can write the vorticity equation in the material form

$$\frac{D_0 (\hat{\zeta}_0 + \hat{\beta} \hat{y})}{D\hat{t}} = -\hat{f}_0 \hat{\nabla} \cdot \hat{u}_1. \quad (54.66)$$

As anticipated, we need one more step to close the system, since the evolution of zeroth order vorticity in equations (54.65) and (54.66) is a function of vortex stretching induced by convergence of the first order velocity. We can substitute for the ageostrophic term $\hat{\nabla} \cdot \hat{u}_1$ through use of

the continuity equation (54.63b), thus leading to a prognostic equation involving just zeroth order terms

$$\frac{\partial(\hat{\zeta}_0 + \hat{\beta} \hat{y} - F \hat{f}_0 \hat{\eta}_0)}{\partial \hat{t}} + (\hat{u}_0 \cdot \hat{\nabla})(\hat{\zeta}_0 + \hat{\beta} \hat{y} - F \hat{f}_0 \hat{\eta}_0) = 0, \quad (54.67)$$

which can be written in the material form

$$\frac{D_0}{D\hat{t}} \left[\hat{\zeta}_0 + \hat{\beta} \hat{y} - F \hat{f}_0 \hat{\eta}_0 \right] = 0. \quad (54.68)$$

Finally, we introduce the geostrophic streamfunction $\hat{\psi}_0 = \hat{\eta}_0/\hat{f}_0$ (equation (54.62)) to render

$$\frac{D_0}{D\hat{t}} \left[\hat{\nabla}^2 \hat{\psi}_0 + \hat{\beta} \hat{y} - F \hat{f}_0^2 \hat{\psi}_0 \right] = 0. \quad (54.69)$$

Equations (54.68) and (54.69) are statements of the material conservation of quasi-geostrophic potential vorticity (QGPV), where material conservation is defined by the zeroth order horizontal geostrophic currents (equation (54.64)).

54.4.5 Dimensional equations and quasi-geostrophic PV

The material conservation equation (54.69) represents the culmination of our quest to realize a self-consistent closed prognostic equation via an asymptotic expansion to first order in Rossby number. We now gather the pieces, and in so doing transform the non-dimensional equations into their dimensional form. Since we are not interested in higher order terms, we drop all 0 subscripts, except for the Coriolis parameter. To proceed, invert the process started in Section 54.2.7, so that

$$\hat{t} = T^{-1} t \quad (\hat{x}, \hat{y}) = L^{-1} (x, y) \quad \partial_{\hat{t}} = T \partial_t \quad \hat{\nabla} = L \nabla \quad (54.70a)$$

$$(\hat{u}, \hat{v}) = U^{-1} (u, v) \quad \hat{f}_0 = \frac{f_0}{f_0} \quad \hat{\beta} \hat{y} = \frac{\beta y}{Ro f_0} = \frac{L}{U} \beta y \quad (54.70b)$$

$$\hat{\eta} = \mathcal{H}^{-1} \Delta \eta \quad \mathcal{H} = H F Ro \quad \hat{\zeta} = \frac{L}{U} \zeta = L^2 \nabla^2 \psi, \quad (54.70c)$$

where the quasi-geostrophic streamfunction is

$$\psi = \frac{g \Delta \eta}{f_0}. \quad (54.71)$$

Note that [Vallis \(2017\)](#) defines the streamfunction as

$$\psi_{\text{vallis}} = \frac{g \eta}{f_0} = \frac{g (H + \Delta \eta)}{f_0}, \quad (54.72)$$

which differs by the constant $g H/f_0$. However, there is no difference in the dynamics since a streamfunction is defined only up to a constant. Also note that we take $F \sim 1$ for QG.

We now make use of these relations in the non-dimensional quasi-geostrophic potential vorticity

(QGPV) equation (54.68)

$$\hat{q} = \hat{\zeta}_0 + \hat{\beta} \hat{y} - \hat{f}_0 \hat{\eta}_0 \quad (54.73a)$$

$$= \frac{L}{U} (\zeta + \beta y) - \frac{\Delta\eta}{\mathcal{H}} \quad (54.73b)$$

$$= \frac{L}{U} (\zeta + \beta y) - \frac{\Delta\eta}{H Ro} \quad (54.73c)$$

$$= \frac{L}{U} \left[\zeta + \beta y - \frac{f_0 \Delta\eta}{H} \right] \quad (54.73d)$$

$$= \frac{L}{U} \left[\zeta + \beta y - \frac{g \Delta\eta}{f_0 L_d^2} \frac{1}{L_d^2} \right] \quad (54.73e)$$

$$= \frac{L}{U} [\zeta + \beta y - L_d^{-2} \psi]. \quad (54.73f)$$

Multiplying both sides by f_0 leads to the dimensionful QGPV for the shallow water fluid layer

$$q = Ro f_0 \hat{q} = \zeta + \beta y - L_d^{-2} \psi = \beta y + (\nabla^2 - L_d^{-2}) \psi. \quad (54.74)$$

54.5 Exercises

EXERCISE 54.1: PV CONSERVATION FOR PG

Show that the planetary geostrophic equations

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} \quad \text{with} \quad \eta = \eta_b + h \quad (54.75)$$

are equivalent to

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad \frac{DQ}{Dt} = 0 \quad \text{with} \quad Q = \frac{f}{h}. \quad (54.76)$$

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 54.2: CONSTRAINTS ON STEADY STATE PLANETARY GEOSTROPHIC FLOW

Consider a shallow water fluid satisfying the planetary geostrophic equations developed in Section 54.3. Assume the flow is in steady state.

- (a) In what manner does potential vorticity conservation constrain the velocity field?
- (b) 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)?
- (c) Describe the path of the velocity field for the case where the ocean sea surface height undulations, $\Delta\eta$, are far smaller than undulations in the bottom topography, η_b (see Figure 42.1 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 non-zero Coriolis parameter, show that there is no meridional geostrophic velocity. That is, the flow is zonally aligned.

EXERCISE 54.3: LINEARIZED SHALLOW WATER PV

The potential vorticity for a shallow water layer is given by

$$Q = \frac{\zeta + f}{h}. \quad (54.77)$$

It is materially constant when the flow is inviscid

$$\frac{DQ}{Dt} = 0. \quad (54.78)$$

Suppose that deviations of the free surface height, $\Delta\eta$, from its equilibrium position are small compared to the thickness, H , of the resting layer. Also assume the Rossby number is small so that $|\zeta| \ll |f|$. Consider flow on a β -plane so that $f = f_0 + \beta y$.

- (a) Show that the evolution equation for potential vorticity can be approximated as

$$\frac{D}{Dt} \left[\zeta + \beta y - \frac{f_0 \Delta\eta}{H} \right] = 0. \quad (54.79)$$

Hint: read Section 56.4.1.

- (b) Using f -plane geostrophic balance, obtain an expression for ζ in terms of η .

EXERCISE 54.4: CONSTRAINT ON f -PLANE GEOSTROPHIC FLOW FROM BOTTOM TOPOGRAPHY

Throughout the quasi-geostrophic scaling for the shallow water in Section 54.2, we assumed the bottom topography has a tiny amplitude relative to the undulations of the free surface height (see equation (54.23)). Return to the shallow water quasi-geostrophic scaling, and instead assume the topography undulations are order one relative to the resting fluid thickness, H . With $\eta_b/H \sim 1$, what does this imply for the zeroth order geostrophic flow? Discuss how the geostrophic streamfunction relates to the bottom topography.

55

Continuously stratified planetary geostrophy

READER'S GUIDE FOR THIS CHAPTER

In this chapter, we extend the shallow water discussions from Chapter 54 to derive the continuously stratified planetary geostrophic equations. We make use of stratified geophysical fluid dynamics from Chapters 22 and 29, as well as potential vorticity from Chapter 50. Material in this chapter is used for the continuous quasi-geostrophic discussion in Chapter 56.

55.1	Open threads	836
55.2	Continuously stratified Boussinesq fluid	836
55.2.1	Dimensional parameters	837
55.2.2	Physical dimensions and non-dimensional parameters	838
55.2.3	Choosing the non-dimensional parameters	838
55.2.4	Relating the buoyancy scale to the Coriolis acceleration scale	839
55.2.5	Richardson number	840
55.2.6	The Rossby deformation radius	840
55.2.7	Assumed values for the non-dimensional parameters	841
55.2.8	Non-dimensional Boussinesq equations	842
55.3	Equations for planetary geostrophy	843
55.4	Planetary geostrophic potential vorticity	844
55.4.1	Derivation	844
55.4.2	Impermeability theorem	845
55.4.3	Kinematic fluxes satisfying impermeability	846
55.5	Depth integrated meridional flow	846
55.5.1	PG vorticity equation	847
55.5.2	Bottom kinematics and dynamics	847
55.5.3	Surface kinematics and dynamics	847
55.5.4	Processes leading to depth integrated meridional transport	848
55.6	Depth integrated flow and f/H contours	849
55.6.1	Massaging the depth integrated momentum budget	849
55.6.2	Flow relative to f/H	850
55.6.3	The rigid lid approximation	850
55.6.4	Further study	851
55.7	Exercises	851

55.1 Open threads

- Rossby waves

55.2 Continuously stratified Boussinesq fluid

Our starting point is the adiabatic stratified hydrostatic Boussinesq equations (Section 28.1)

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -\nabla_z \phi \quad (55.1a)$$

$$\frac{\partial \phi}{\partial z} = b \quad (55.1b)$$

$$\frac{Db}{Dt} = 0 \quad (55.1c)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (55.1d)$$

where $\mathbf{v} = (\mathbf{u}, w)$ is the three-dimensional velocity, $b = -g(\rho - \rho_0)/\rho_0$ is the buoyancy, ρ is the density, ρ_0 is a constant reference density, $\phi = \delta p/\rho_0$ is the dynamic pressure (dimensions of $(\text{length})^2 (\text{time})^{-2}$), and $\nabla_z = (\partial_x, \partial_y, 0)$ is the horizontal gradient operator. We separate a background vertical buoyancy profile from the fluctuating buoyancy

$$b = \tilde{b}(z) + b'(x, y, z, t), \quad (55.2)$$

and introduce the corresponding background squared buoyancy frequency

$$N^2 = \frac{\partial \tilde{b}}{\partial z}. \quad (55.3)$$

With this decomposition, the buoyancy equation (55.1c) takes the form

$$\frac{Db'}{Dt} + w N^2 = 0. \quad (55.4)$$

We also introduce an associated decomposition of the hydrostatic pressure

$$\phi = \tilde{\phi}(z) + \phi'(x, y, z, t) \quad (55.5)$$

where $\tilde{\phi}$ is hydrostatically balanced by \tilde{b}

$$\frac{d\tilde{\phi}}{dz} = \tilde{b}, \quad (55.6)$$

and the fluctuating pressure, ϕ' , is hydrostatically balanced by b'

$$\frac{\partial \phi'}{\partial z} = b'. \quad (55.7)$$

Hence, the scale for the fluctuating pressure, Φ , is related to the scale for the fluctuating buoyancy, B , and the depth scale, H

$$\Phi/H = B. \quad (55.8)$$

We return to this scaling relation, and others, in the following.

55.2.1 Dimensional parameters

Following the shallow water discussion in Section 54.2.1, we have the following dimensional parameters for the adiabatic Boussinesq fluid.

- LENGTH SCALES
 - ★ H = depth scale of a typical vertical structure in the fluid (e.g., the depth of the thermocline).
 - ★ L = horizontal/lateral length scale of motions under consideration.
 - ★ R = radius of the planet.
- VELOCITY SCALES
 - ★ U = horizontal velocity scale for fluid parcel motion.
 - ★ W = vertical velocity scale for fluid parcel motion.
- PRESSURE AND BUOYANCY SCALES: Pressure is a contact force, acting on the boundary of an arbitrary fluid region, and buoyancy is a force acting to raise or lower a fluid parcel depending on its density relative to the environment. They have scales given by the following.
 - ★ Φ = scale for pressure fluctuations ϕ' (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 54.2.1, we do not introduce a wave speed since it does not affect the asymptotics. We also dropped the bottom topography scale, assuming it is small for present purposes.

55.2.2 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 no mass since mass is determined by the density (buoyancy) and volume. The Buckingham-II theorem then says there are

$$N_{\text{dimensionless}} = 9 - 2 = 7 \quad (55.9)$$

non-dimensional parameters.

55.2.3 Choosing the non-dimensional parameters

Following the shallow water discussion in Section 54.2.4, we choose the following non-dimensional parameters.

1. VERTICAL TO HORIZONTAL ASPECT RATIO: The ratio of the vertical scale to the horizontal scale defines the aspect ratio

$$\delta_{\text{vertical/horizontal}} = \frac{\text{vertical depth scale}}{\text{horizontal length scale}} = \frac{H}{L}. \quad (55.10)$$

2. RATIO OF HORIZONTAL SCALE TO PLANETARY SCALE: The ratio of the lateral length scale to the planet radius is given by

$$\delta_{\text{horizontal/planet}} = \frac{\text{lateral length scale}}{\text{planetary length scale}} = \frac{L}{R}. \quad (55.11)$$

3. RATIO VERTICAL TO HORIZONTAL VELOCITY SCALE: The ratio of the vertical to horizontal velocity is given by

$$\frac{\text{vertical velocity scale}}{\text{horizontal velocity scale}} = \frac{W}{U}. \quad (55.12)$$

4. HYDROSTATIC NUMBER: The hydrostatic number is the ratio of the pressure gradient scale to the buoyancy scale. For the hydrostatic fluid fluctuations

$$\frac{\Phi}{H} = B, \quad (55.13)$$

where B is the scale for the buoyancy fluctuations. We encountered this relation earlier in equation (55.8).

5. ROSSBY NUMBER: The Rossby number is the ratio of the fluid parcel acceleration scale to the Coriolis acceleration

$$Ro = \frac{\text{parcel acceleration}}{\text{Coriolis acceleration}} = \frac{U}{fL}, \quad (55.14)$$

where we again assume time scales advectively

$$T \sim \frac{L}{U}. \quad (55.15)$$

6. GEOSTROPHIC NUMBER: The ratio of the Coriolis acceleration to the pressure gradient acceleration defines the geostrophic number

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}}. \quad (55.16)$$

The Coriolis acceleration scales as

$$\text{Coriolis acceleration} \sim fU \quad (55.17)$$

whereas the pressure gradient acceleration from the fluctuating pressure, ϕ' , scales as

$$\text{pressure gradient acceleration} \sim \frac{\Phi}{L}, \quad (55.18)$$

so that

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}} = \frac{fU}{(\Phi/L)}. \quad (55.19)$$

7. RATIO FLUCTUATING STRATIFICATION TO BACKGROUND STRATIFICATION: The ratio of the buoyancy frequency arising from the fluctuating buoyancy, B/H , to the background buoyancy frequency, N^2 , is given by

$$\frac{\text{fluctuating buoyancy frequency}}{\text{background buoyancy frequency}} = \frac{B/H}{N^2}. \quad (55.20)$$

55.2.4 Relating the buoyancy scale to the Coriolis acceleration scale

The fluctuating buoyancy (b') and fluctuating pressure (ϕ') have scales related through the hydrostatic balance. From equation (55.13) we have

$$B = \frac{\Phi}{H}. \quad (55.21)$$

Additionally, assuming geostrophic scaling, equation (55.19) means that the fluctuating pressure has a scale related to the Coriolis acceleration scale according to

$$\Phi = fUL. \quad (55.22)$$

Hence, the scale for the fluctuating buoyancy is given by

$$B = \frac{fUL}{H}. \quad (55.23)$$

55.2.5 Richardson number

The *Richardson number* is the non-dimensional ratio of the squared buoyancy frequency to the squared vertical shear of the horizontal velocity

$$Ri = \frac{N^2}{|\partial_z \mathbf{u}|^2}. \quad (55.24)$$

In regions where $Ri \ll 1$, the vertical shear is strong and the flow tends to be unstable to *Kelvin-Helmholz instability*. In these regions, there is enough kinetic energy in the vertical shear to extract potential energy from the stratification, and this extraction process occurs via a dynamical instability. In contrast, for large-scale highly stratified flow, the Richardson number is quite large, with $Ri \sim 100$ common. This is the regime where quasi-geostrophy is relevant.

In our choice for dimensionless parameters, we could choose one determined by the scale for the Richardson number

$$Ri = \frac{N^2}{(U/H)^2}, \quad (55.25)$$

where we set the vertical scale equal to H , the horizontal velocity scale to U , and the squared buoyancy frequency to a scale N^2 . However, the Richardson number can be related to the Rossby and Burger numbers through

$$Bu = \left[\frac{L_d}{L} \right]^2 = \left[\frac{NH}{fL} \right]^2 = \frac{U^2 Ri}{U^2/(Ro)^2} = (Ro)^2 Ri. \quad (55.26)$$

For QG flows, the horizontal length scales, L , 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 (55.26) thus means that the Richardson number scales as

$$Ri \sim (Ro)^{-2} \quad \text{QG flow.} \quad (55.27)$$

For atmospheric flows with a Rossby number order 1/10, QG flow regimes are realized with a Richardson number ~ 100 . For the ocean, the Rossby number can be even smaller, in which case QG flows are characterized by an even larger Richardson number. For planetary geostrophy, the Burger number is small, in which case PG flows are characterized by somewhat smaller Richardson numbers than QG flows.

55.2.6 The Rossby deformation radius

The combined effects of buoyancy and rotation yield the richness of continuously stratified QG motions. Hence, the buoyancy frequency and the Coriolis parameter play central roles in QG theory. The ratio of these two frequencies N/f in regions of nontrivial vertical stratification is typically around 100. Rotational inertial oscillations (usually just called *inertial oscillations*) have about 100 times longer period $T_f = 2\pi/f$ than buoyancy oscillations with period $T_b = 2\pi/N$.

Letting the squared buoyancy frequency N^2 refer to a value typical of a particular flow regime, one can define the Rossby deformation radius

$$L_d = H \frac{N}{f}. \quad (55.28)$$

The ratio f/N appears frequently in rotating/stratified fluids, and is sometimes called the Prandtl ratio

$$\frac{f}{N} = \text{Prandtl ratio.} \quad (55.29)$$

With $H \approx 1$ km and $N/f \approx 100$, the Rossby radius is roughly 100 km. In general, the Rossby radius is a crucial scale in geophysical fluids. For example, it sets the scale for the most unstable baroclinic waves leading to baroclinically unstable flow (see Chapter 6 of [Vallis \(2017\)](#)).

55.2.7 Assumed values for the non-dimensional parameters

We now enumerate the assumed values for the non-dimensional parameters, again largely following the choices made for the shallow water system in Section 54.2.5.

1. SMALL VERTICAL TO HORIZONTAL ASPECT RATIO: The aspect ratio is generally small for large-scale atmospheric and oceanic fluid systems

$$\delta_{\text{vertical/horizontal}} \ll 1. \quad (55.30)$$

This assumption was made when making the hydrostatic approximation (Section 27.3).

2. SMALL OR ORDER ONE RATIO OF HORIZONTAL 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

$$\delta_{\text{horizontal/planet}} \ll 1 \quad \text{quasi-geostrophy} \quad (55.31a)$$

$$\delta_{\text{horizontal/planet}} \sim 1 \quad \text{planetary geostrophy.} \quad (55.31b)$$

3. SMALL RATIO VERTICAL TO HORIZONTAL VELOCITY SCALE: The continuity equation implies

$$\frac{W}{H} = \frac{U}{L}, \quad (55.32)$$

so that

$$W = U \frac{H}{L}. \quad (55.33)$$

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 Ro . We will see that factor naturally appear in the following.

4. UNIT HYDROSTATIC NUMBER: As already noted, the hydrostatic balance (55.1b) means that the scales for a buoyancy fluctuation and pressure fluctuation are related by (see equation (55.8))

$$\Phi = H B. \quad (55.34)$$

5. SMALL ROSSBY NUMBER: The Rossby number is assumed small

$$Ro = \frac{U}{f L} = \frac{1}{f T} \ll 1, \quad (55.35)$$

where we set the time scale according to advection, $T = L/U$.

6. UNIT GEOSTROPHIC NUMBER: The geostrophic number is assumed to be order unity

$$Ge \sim 1, \quad (55.36)$$

which means that the Coriolis acceleration and pressure gradient acceleration scale together

$$f U \sim \frac{\Phi}{L}. \quad (55.37)$$

This scaling is consistent with the momentum equation (55.1a) so long as the Rossby number is small, $Ro \ll 1$.

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

$$\frac{B/H}{N^2} = \frac{\Phi}{H^2 N^2} = \frac{f U L}{H^2 N^2} = \frac{U}{f L} \frac{L^2 f^2}{H^2 N^2} = Ro \frac{L^2}{L_d^2}, \quad (55.38)$$

where we introduced the deformation radius for the continuously stratified system

$$L_d = H \frac{N}{f}. \quad (55.39)$$

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. Notably, since N^2 is a function of depth, we must keep this in mind when returning to dimensional fields, particularly for the QG system in Section 56.2.3.

55.2.8 Non-dimensional Boussinesq equations

Following the shallow water approach in Section 54.2.7, we introduce non-dimensional variables according to

$$t = T \hat{t} \quad (x, y) = L(\hat{x}, \hat{y}) \quad \partial_t = T^{-1} \partial_{\hat{t}} \quad \nabla_z = L^{-1} \hat{\nabla}_z \quad \partial_z = H^{-1} \partial_{\hat{z}} \quad (55.40a)$$

$$(u, v) = U(\hat{u}, \hat{v}) \quad w = W \hat{w} \quad f = f_o \hat{f} \quad \phi' = f_0 U L \hat{\phi} \quad b' = B \hat{b} = \left[\frac{f_0 U L}{H} \right] \hat{b}. \quad (55.40b)$$

For the second equality in the buoyancy scale, we made use of the relation (55.23) to connect the buoyancy fluctuation scale to the Coriolis acceleration scale. We also make use of the following relations between scales

$$T = \frac{L}{U} \quad W = \frac{U H}{L} \quad Ro = \frac{U}{f_0 L} = \frac{1}{T f_0}. \quad (55.41)$$

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 vertical velocity scales according to the continuity equation, $W = U(H/L)$. This continuity scaling for W will be seen to be an over-estimate in the following.

Introducing the above variables and scales into the Boussinesq momentum equation (55.1a) renders

$$\frac{U}{T} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + \frac{U^2}{L} (\hat{\mathbf{u}} \cdot \hat{\nabla}_z) \hat{\mathbf{u}} + \frac{W U}{H} \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} + f_o U (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = -f_0 U \hat{\nabla}_z \hat{\phi}. \quad (55.42)$$

Hence, dividing by $f_o U$ leads to

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}_{\hat{z}}) \hat{\mathbf{u}} + \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} \right] + (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = -\hat{\nabla}_z \hat{\phi}. \quad (55.43)$$

The momentum 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 parcel acceleration. Likewise, the non-dimensional hydrostatic balance is given by

$$\frac{\partial \hat{\phi}}{\partial \hat{z}} = \hat{b}, \quad (55.44)$$

and the non-dimensional continuity equation is

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0. \quad (55.45)$$

The buoyancy equation (55.4) requires a bit more work to non-dimensionalize. The material time derivative takes the form

$$\frac{Db'}{Dt} = \frac{B}{T} \frac{D\hat{b}}{D\hat{t}} = \frac{U}{L} \frac{f_0 U L}{H} \frac{D\hat{b}}{D\hat{t}} = \frac{f_0 U^2}{H} \frac{D\hat{b}}{D\hat{t}}, \quad (55.46)$$

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

$$N^2 w = N^2 W \hat{w} = N^2 U (H/L) \hat{w} = L_d^2 \frac{U f_0^2}{H L} \hat{w}, \quad (55.47)$$

where we introduced the deformation scale, $L_d = HN/f$, from equation (55.39). Bringing these two pieces together leads to

$$Ro F \frac{D\hat{b}}{D\hat{t}} + \hat{w} = 0, \quad (55.48)$$

where we introduced the Burger number

$$Bu = F^{-1} = \left[\frac{L_d}{L} \right]^2. \quad (55.49)$$

55.3 Equations for planetary geostrophy

Just like for the shallow water model in Section 54.3, the planetary geostrophic system for the stratified Boussinesq system is rather simple to derive. For this case, we assume the horizontal scales are large compared to the deformation radius, so that

$$F Ro = Ro/Bu \sim 1, \quad (55.50)$$

or

$$L^2 \sim L_d^2 Ro^{-1}. \quad (55.51)$$

With this scaling, and with the Rossby number small, the momentum equation (55.43) reduces to geostrophic balance. However, the continuity and buoyancy equations retain their unapproximated Boussinesq form. Hence, in dimensional form, the adiabatic planetary geostrophic equations for a stratified Boussinesq fluid take the form

$$\frac{Db'}{Dt} + w N^2 = 0 \quad (55.52a)$$

$$\mathbf{f} \wedge \mathbf{u} = -\nabla_z \phi' \quad (55.52b)$$

$$\frac{\partial \phi'}{\partial z} = b' \quad (55.52c)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (55.52d)$$

We could just as well write these equations in terms of the full buoyancy $b = \tilde{b}(z) + b'$, in which

$$\frac{Db}{Dt} = 0 \quad (55.53a)$$

$$\mathbf{f} \wedge \mathbf{u} = -\nabla_z \phi \quad (55.53b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (55.53c)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (55.53d)$$

Note that the material time derivative in PG makes use of advection by the three components of the velocity field, $\mathbf{v} = (\mathbf{u}, w)$, where the horizontal components are given by the geostrophic balance (55.52b). This situation contrasts with the QG approach, where it is only the horizontal advection that contributes to material time evolution at leading order (Section 56.2).

55.4 Planetary geostrophic potential vorticity

In Section 52.1 we developed the potential vorticity equation for the hydrostatic Boussinesq system in the presence of friction in the horizontal 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

$$f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z \phi + \mathbf{F} \quad \text{frictional geostrophy} \quad (55.54a)$$

$$\frac{\partial b}{\partial t} + \mathbf{u} \cdot \nabla_z b + N^2 w = \dot{b} \quad \text{diabatic buoyancy equation} \quad (55.54b)$$

$$\frac{\partial \phi}{\partial z} = b \quad \text{hydrostatic} \quad (55.54c)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad \text{incompressible} \quad (55.54d)$$

$$N^2 = \frac{\partial b}{\partial z} \quad \text{squared buoyancy frequency.} \quad (55.54e)$$

Assume a linear equation of state so that buoyancy is linearly proportional to potential temperature and/or salinity.

55.4.1 Derivation

Derivation of the PV equation proceeds much like that for the hydrostatic Boussinesq system. The first step requires the PG vorticity equation as determined by taking the curl of the momentum equation (55.54a)

$$\beta v = f \frac{\partial w}{\partial z} + \hat{\mathbf{z}} \cdot (\nabla_z \wedge \mathbf{F}) \quad \text{with} \quad \beta = \frac{\partial f}{\partial y}. \quad (55.55)$$

The frictionless version of the vorticity equation yields the Sverdrup relation

$$\beta v = f \frac{\partial w}{\partial z} \quad \text{Sverdrup balance,} \quad (55.56)$$

whereby meridional advection of planetary vorticity (left hand side) balances vertical vortex stretching (see Section 29.4.4). Next make use of the thermal wind relation in the presence of friction

$$f \frac{\partial \mathbf{u}}{\partial z} = \hat{\mathbf{z}} \wedge \nabla_z b - \frac{\partial(\hat{\mathbf{z}} \wedge \mathbf{F})}{\partial z} \quad (55.57)$$

as well as the identities

$$N^2 \frac{Df}{Dt} = N^2 \beta v \quad (55.58a)$$

$$f \frac{DN^2}{Dt} = f \frac{\partial \dot{b}}{\partial z} - f \nabla b \cdot \frac{\partial \mathbf{v}}{\partial z} \quad (55.58b)$$

$$f \nabla b \cdot \frac{\partial \mathbf{v}}{\partial z} = f N^2 \frac{\partial w}{\partial z} - \frac{\partial(\hat{\mathbf{z}} \wedge \mathbf{F})}{\partial z} \cdot \nabla_z b, \quad (55.58c)$$

to render

$$\frac{D(f N^2)}{Dt} = N^2 \beta v + f \frac{\partial \dot{b}}{\partial z} - f \nabla b \cdot \frac{\partial \mathbf{v}}{\partial z} \quad (55.59a)$$

$$= N^2 \left[f \frac{\partial w}{\partial z} + \hat{\mathbf{z}} \cdot (\nabla_z \wedge \mathbf{F}) \right] + f \frac{\partial \dot{b}}{\partial z} - f N^2 \frac{\partial w}{\partial z} + \frac{\partial(\hat{\mathbf{z}} \wedge \mathbf{F})}{\partial z} \cdot \nabla_z b \quad (55.59b)$$

$$= f \frac{\partial \dot{b}}{\partial z} + \nabla b \cdot (\nabla \wedge \mathbf{F}) \quad (55.59c)$$

$$= \nabla \cdot (f \dot{b} \hat{\mathbf{z}} + \mathbf{F} \wedge \nabla b). \quad (55.59d)$$

We thus identify the planetary geostrophic potential vorticity

$$Q_{PG} = f N^2, \quad (55.60)$$

which is materially invariant in the absence of diabatic processes and friction

$$\frac{DQ_{PG}}{Dt} = 0 \quad \text{when } \dot{b} = 0 \text{ and } \mathbf{F} = 0. \quad (55.61)$$

We can write the general budget equation in the form of an Eulerian flux-form expression

$$\frac{\partial Q_{PG}}{\partial t} + \nabla \cdot \mathbf{J}_{PG} = 0, \quad (55.62)$$

where the potential vorticity flux is given by

$$\mathbf{J}_{PG} = \mathbf{v} Q - \dot{b} f \hat{\mathbf{z}} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A}, \quad (55.63)$$

where \mathbf{A} is an arbitrary gauge function. Comparing to the hydrostatic Boussinesq case from Section 52.1.3, we see that the planetary geostrophic result follows by dropping contributions from the relative vorticity and dropping the inertial acceleration.

55.4.2 Impermeability theorem

Following the discussion in Section 51.1.2, we here verify that the potential vorticity flux vector (55.63) satisfies the impermeability theorem for buoyancy isosurfaces. Setting the gauge function to zero and making use of the identities

$$\mathbf{v}_{PG} \cdot \nabla b = \frac{\mathbf{J}_{PG} \cdot \nabla b}{Q} = \mathbf{v} \cdot \nabla b - \dot{b} = -\frac{\partial b}{\partial t}, \quad (55.64)$$

reveals that

$$\frac{\partial b}{\partial t} + \mathbf{v}_{PG} \cdot \nabla b = 0. \quad (55.65)$$

We conclude that there is zero flux of PV-substance crossing buoyancy isosurfaces, even in the presence of irreversible processes that allow matter and heat to cross buoyancy surfaces.

55.4.3 Kinematic fluxes satisfying impermeability

Following the discussion of impermeability for the Ertel PV in Section 51.1.2, we expose a purely kinematic means to derive the impermeability theorem for the PG PV. This derivation follows by computing the time tendency of the PV

$$\frac{\partial Q}{\partial t} = \frac{\partial}{\partial t} \nabla \cdot (f b \hat{z}) = \nabla \cdot \left[f \frac{\partial b}{\partial t} \hat{z} \right] \equiv -\nabla \cdot \tilde{\mathbf{J}}_{\text{PG}}, \quad (55.66)$$

where

$$\tilde{\mathbf{J}}_{\text{PG}} = -f \frac{\partial b}{\partial t} \hat{z}. \quad (55.67)$$

This form of the PV-substance flux also satisfies impermeability since

$$\tilde{\mathbf{v}}_{\text{PG}} \cdot \nabla b = \frac{\tilde{\mathbf{J}}_{\text{PG}} \cdot \nabla b}{Q} = -\frac{\partial b}{\partial t}, \quad (55.68)$$

so that

$$\frac{\partial b}{\partial t} + \tilde{\mathbf{v}}_{\text{PG}} \cdot \nabla b = 0. \quad (55.69)$$

Note how the flux $\tilde{\mathbf{J}}_{\text{PG}}$ vanishes in the steady state, whereas the steady state form of the flux \mathbf{J}_{PG} does not vanish. Following the discussion in Section 51.2.2, 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 (51.34) renders

$$\mathbf{J}_Q^{\text{marshall PG}} = -\nabla(g z + p/\rho_0) \wedge \nabla b - f \frac{\partial b}{\partial t} \hat{z}. \quad (55.70)$$

This flux differs from $\tilde{\mathbf{J}}_{\text{PG}}$ by a curl

$$\nabla(g z + p/\rho_0) \wedge \nabla b = \nabla \wedge [(g z + p/\rho_0) \nabla b], \quad (55.71)$$

and it also satisfies the impermeability theorem. As discussed in Section 51.4.6, 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 involve any irreversible processes, with $\mathbf{J}_Q^{\text{marshall PG}}$ satisfying these desires.

55.5 Depth integrated meridional flow

We here consider the depth integrated vorticity budget for the planetary geostrophic system and the resulting implications for the meridional flow. These considerations are pertinent to studies of large-scale ocean circulation. We are interested in a vertical stress divergence as well as diabatic processes so that the planetary geostrophic equations are given by

$$\rho_o f (\hat{z} \wedge \mathbf{u}) = -\nabla p - \rho g \hat{z} + \frac{\partial \boldsymbol{\tau}}{\partial z} \quad (55.72a)$$

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (55.72b)$$

$$\frac{Db}{Dt} = \dot{b}. \quad (55.72c)$$

55.5.1 PG vorticity equation

In Section 29.4.2 we derived the vorticity equation for the planetary geostrophic system, whose vertical component is given by equation (29.38)

$$\rho_o \beta v = \frac{\partial}{\partial z} [\rho_o f w + \hat{z} \cdot (\nabla \wedge \boldsymbol{\tau})]. \quad (55.73)$$

Vertical integration from the ocean bottom at $z = -H(x, y)$ to free surface at $z = \eta(x, y, t)$ leads to

$$\rho_o \beta V = \rho_o f [w(\eta) - w(-H)] + \hat{z} \cdot (\nabla \wedge \Delta \boldsymbol{\tau}), \quad (55.74)$$

where

$$\Delta \boldsymbol{\tau} = \boldsymbol{\tau}(\eta) - \boldsymbol{\tau}(-H) \quad (55.75)$$

is the difference in stress applied at the ocean surface and ocean bottom. We now make use of boundary conditions and horizontal momentum equation to write the vertical velocity difference in terms of pressure and boundary undulations. Doing so provides us with insights into processes that drive vertical flows at the boundaries, and thus what leads to depth integrated meridional transport.

55.5.2 Bottom kinematics and dynamics

The bottom kinematic boundary condition (Section 17.4.1) leads to

$$w(-H) = -\mathbf{u}(-H) \cdot \nabla H. \quad (55.76)$$

Evaluating the horizontal momentum equation (29.24a) at the ocean bottom yields

$$\rho_o f \hat{z} \wedge \mathbf{u}(-H) = -\nabla p_b + \frac{\partial \boldsymbol{\tau}(-H)}{\partial z}, \quad (55.77)$$

where $p_b(x, y, t)$ is the bottom pressure. To proceed, we focus exclusively on the geostrophic component of the bottom velocity, which is driven solely by the bottom pressure. In this case the dynamic bottom boundary condition for the horizontal velocity is given by

$$\rho_o f \hat{z} \wedge \mathbf{u}_g(-H) = -\nabla p_b. \quad (55.78)$$

Note that all velocity components vanish at the bottom when imposing a no-slip bottom boundary condition. We thus consider $\mathbf{u}_g(-H)$ as the horizontal geostrophic velocity within the bottom boundary layer arising from the bottom pressure gradient.

55.5.3 Surface kinematics and dynamics

For purposes of large-scale circulation studies using the PG equations, it is generally sufficient to assume a rigid lid upper boundary condition, whereby $w(\eta) = w(0) = 0$. Even so, we find it interesting to present the results here for the free surface case in which there is the possibility of nonzero surface mass fluxes. The surface kinematic boundary condition (Section 19.2) yields

$$w(\eta) = -\frac{Q_m}{\rho_0} + \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta, \quad (55.79)$$

in which the steady state balance is¹

$$w(\eta) = -\frac{Q_m}{\rho_0} + \mathbf{u} \cdot \nabla \eta. \quad (55.80)$$

Evaluating the horizontal momentum equation (29.24a) at the ocean surface renders

$$\rho_o f \hat{\mathbf{z}} \wedge \mathbf{u}(\eta) = -\nabla p_a + \frac{\partial \boldsymbol{\tau}(\eta)}{\partial z}, \quad (55.81)$$

where $p_a(x, y, t)$ is the pressure applied to the ocean surface. Like the bottom, we are only interested in that portion of the horizontal velocity at the ocean boundary driven by surface pressure, in which case²

$$\rho_o f \hat{\mathbf{z}} \wedge \mathbf{u}_g(\eta) = -\nabla p_a. \quad (55.82)$$

55.5.4 Processes leading to depth integrated meridional transport

Plugging expressions (55.82) and (55.78) into the depth integrated balance (55.74) renders the depth integrated meridional transport

$$\rho_o \beta V = -f Q_m + \hat{\mathbf{z}} \cdot [\nabla p_a \wedge \nabla \eta + \nabla p_b \wedge \nabla H + \nabla \wedge \Delta \boldsymbol{\tau}]. \quad (55.83)$$

A large part of the bottom pressure gradient arises from changes in bottom depth. However, that portion of the bottom pressure has no impact on the meridional transport, as seen by writing the bottom pressure as

$$p_b = \rho_0 g H + p'_b \quad (55.84)$$

so that

$$\rho_o \beta V = -f Q_m + \hat{\mathbf{z}} \cdot [\nabla p_a \wedge \nabla \eta + \nabla p'_b \wedge \nabla H + \nabla \wedge \Delta \boldsymbol{\tau}]. \quad (55.85)$$

This equation shows that the steady state depth integrated meridional transport is balanced by four terms.

- SURFACE MASS TRANSPORT: The term $-f Q_m$ arises from mass transport across the ocean surface. It gives rise to a meridional circulation known as the Goldsbrough-Stommel circulation (see [Huang and Schmitt \(1993\)](#) for a review). It is generally sub-dominant to the turbulent boundary flux contributions from wind stresses.
- TURBULENT BOUNDARY STRESSES: The term $\hat{\mathbf{z}} \cdot (\nabla \wedge \Delta \boldsymbol{\tau})$ 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 the planetary geostrophic vorticity budget.

¹Even for transient solutions, the balance in equation (55.79) is largely that found in equation (55.80) since the time tendency $\partial \eta / \partial t$ is about five orders of magnitude smaller than the typical vertical velocity under the planetary geostrophic regime. See Section 3.3 of [Samelson \(2011\)](#) for more details.

²As discussed in Chapter 35, we generally ignore surface tension effects for geophysical fluids at scales larger than a few centimeters to a meter. In the absence of surface tension there is a continuity of tangential stresses across the air-sea boundary (see Chapter 26 or Section 4.10 of [Kundu et al. \(2012\)](#)), in which case $\partial \boldsymbol{\tau}(\eta) / \partial z = 0$, thus leading to equation (55.82).

- ATMOSPHERIC PRESSURE TORQUE: The term

$$\nabla p_a \wedge \nabla \eta = \nabla \wedge (p_a \nabla \eta) = -\nabla \wedge (\eta \nabla p_a) \quad (55.86)$$

arises from differences in lines of constant atmospheric pressure and lines of constant sea level (surface topography). Such misalignments create a torque that drives a depth integrated meridional flow. This torque vanishes when making a rigid lid approximation, which is a sensible choice when focused on just the large-scale flow.

- BOTTOM PRESSURE TORQUE: The term

$$\nabla p'_b \wedge \nabla H = \nabla \wedge (p'_b \nabla H) = -\nabla \wedge (H \nabla p'_b) \quad (55.87)$$

arises from differences in lines of constant bottom pressure and lines of constant bottom topography. Such misalignments create a torque that drives a depth integrated meridional flow. This term vanishes when the bottom topography is flat or when the fluid has a constant density. In more general situations it can be a far larger contribution to the vorticity budget than the turbulent bottom stress, and in some cases it can rival contributions from surface wind stress.

55.6 Depth integrated flow and f/H contours

As shown in Section 54.3, f/h is the potential vorticity (PV) for the shallow water planetary geostrophic system, where h is the thickness of the layer of shallow water fluid. Here, we note that when the flow is steady and the fluid is perfect (i.e., no friction or diffusion), the shallow water velocity is oriented along lines of constant f/h . In the limit where the free surface is quasi-stationary, we find that f/H provides a good approximation to the shallow water PV. The quasi-stationary approximation is a bit more general than the rigid lid introduced in Section 55.5, as it allows for non-zero boundary mass fluxes (e.g., evaporation and precipitation) to balance a non-zero divergence in the depth integrated flow. It is a good approximation for large-scale planetary geostrophic flow, where time changes to the free surface are generally much smaller than typical vertical velocities (see Section 3.3 of [Samelson \(2011\)](#) for more details).

55.6.1 Massaging the depth integrated momentum budget

The depth integrated horizontal momentum equation (55.72a) is given by

$$\rho_o f \hat{z} \wedge \mathbf{U} = - \int_{-H}^{\eta} \nabla_z p \, dz + \Delta\tau \quad (55.88)$$

where $\mathbf{U} = \int_{-H}^{\eta} \mathbf{u} \, dz$ is the depth integrated horizontal velocity. The depth integrated pressure gradient can be written as the sum of boundary terms plus the potential energy per area

$$\int_{-H}^{\eta} p \, dz = \int_{-H}^{\eta} [d(pz) - z \, dp] = p_a \eta + p_b H + \int_{-H}^{\eta} g \rho z \, dz, \quad (55.89)$$

where we used the hydrostatic balance to write $dp = -g \rho dz$, and the potential energy per area of a fluid column is given by

$$\mathcal{P} = \int_{-H}^{\eta} g \rho z \, dz. \quad (55.90)$$

These results then lead to the depth integrated horizontal pressure gradient³

$$\int_{-H}^{\eta} \nabla_z p \, dz = \nabla \left[\int_{-H}^{\eta} p \, dz \right] - p_a \nabla \eta - p_b \nabla H \quad (55.91a)$$

$$= \nabla [p_a \eta + p_b H + \mathcal{P}] - p_a \nabla \eta - p_b \nabla H \quad (55.91b)$$

$$= \eta \nabla p_a + H \nabla p_b + \nabla \mathcal{P}, \quad (55.91c)$$

which in turn renders the depth integrated horizontal momentum balance

$$\rho_o f \hat{z} \wedge \mathbf{U} = -\eta \nabla p_a - H \nabla p_b - \nabla \mathcal{P} + \Delta \boldsymbol{\tau}. \quad (55.92)$$

The momentum balance is here written in terms of pressure form stress on the top and bottom interfaces of the fluid column (see Section 43.2 for more discussion of pressure form stress), the gradient of the potential energy per area, and the difference in turbulent stresses at the top and bottom boundaries.

55.6.2 Flow relative to f/H

Dividing the depth integrated momentum equation (55.92) by the depth H and taking the curl leads to

$$\nabla \cdot (\mathbf{U} f/H) = \hat{z} \cdot [\nabla(\eta/H) \wedge \nabla p_a + \nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)]. \quad (55.93)$$

Performing the chain rule on the left hand side leads to

$$\mathbf{U} \cdot \nabla(f/H) + (f/H) \nabla \cdot \mathbf{U} = \hat{z} \cdot [\nabla(\eta/H) \wedge \nabla p_a + \nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)]. \quad (55.94)$$

In Section 19.3 we show that a column volume budget for an incompressible fluid leads to the free surface time tendency

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}. \quad (55.95)$$

Assuming a steady state then leads to boundary mass fluxes balancing a divergence in the depth integrated flow, in which case

$$\mathbf{U} \cdot \nabla(f/H) = -\frac{f}{H} \frac{Q_m}{\rho_0} + \hat{z} \cdot [\nabla(\eta/H) \wedge \nabla p_a + \nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)]. \quad (55.96)$$

This balance indicates that there are many sources for the depth integrated flow to deviate from contours of constant f/H .

55.6.3 The rigid lid approximation

We consider a special case of a rigid lid flow, in which $\nabla \cdot \mathbf{U} = 0$, thus allowing for the introduction of a streamfunction (with dimensions of $L^3 T$) for the depth integrated flow

$$\mathbf{U} = \hat{z} \wedge \nabla \Psi. \quad (55.97)$$

The balance (55.96) in turn takes the form

$$[\nabla \Psi \wedge \nabla(f/H)] \cdot \hat{z} = [\nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)] \cdot \hat{z}, \quad (55.98)$$

³We write $\nabla_z p$ inside the integral of equation (55.91a), since p is a function of depth. However, we use the more concise ∇ in equations (55.91b) and (55.91c) since all terms are only spatial functions of the horizontal coordinates x, y meaning that ∇ reduces to ∇_z .

where we dropped the boundary mass flux and the free surface term as per the rigid lid approximation. In the absence of boundary stresses, contours of constant f/H serve as streamlines for flow where the potential energy parallels lines of constant H . For the rigid lid case, the potential energy per area is given by $\mathcal{P} = \int_{-H}^0 g \rho z dz$, so that we ignore potential energy associated with the free surface undulations. The special case of a constant density fluid leads to $\mathcal{P} = g \rho H^2/2$ so that $\nabla(1/H) \wedge \nabla\mathcal{P} = 0$. Hence, steady rigid lid flow in a homogeneous fluid aligns with f/H contours.

55.6.4 Further study

Chapter 3 in [Samelson \(2011\)](#) offers more discussion of the large-scale ocean circulation based on material in Sections 55.5 and 55.6

55.7 Exercises

EXERCISE 55.1: RIGID LID AND f/H CONTOURS

Derive equation (55.98) for rigid lid flow.

56

Continuously stratified quasi-geostrophy

READER'S GUIDE FOR THIS CHAPTER

In this chapter we extend the shallow water discussions of quasi-geostrophy in Chapter 54 to continuously stratified fluids. We make use of stratified geophysical fluid dynamics from Chapters 22 and 29, as well as potential vorticity from Chapter 50 and the development of continuously stratified planetary geostrophy in Chapter 55. Material in this chapter is basic to the quasi-geostrophic theory of atmospheric and oceanic flows.

56.1	Open threads	854
56.2	Quasi-geostrophy	854
56.2.1	Zeroth order asymptotic equations	855
56.2.2	First order asymptotic equations	856
56.2.3	Dimensional QG-PV equation	857
56.2.4	Properties of the steady state flow field	858
56.2.5	Constant background buoyancy frequency	859
56.2.6	Buoyancy advection at the boundaries	860
56.3	Dimensions of various forms for potential vorticity	860
56.4	Connecting QG-PV to Ertel PV	861
56.4.1	Shallow water layer	861
56.4.2	Continuously stratified hydrostatic Boussinesq fluid	862
56.5	Energetics	863
56.5.1	Kinetic energy	863
56.5.2	Available potential energy	865
56.5.3	Energy conversion	865
56.5.4	Scaling APE and KE	865
56.6	Exercises	866

56.1 Open threads

- Rossby waves
- Impermeability theorem
- Synoptic development
- TEM

56.2 Quasi-geostrophy

We proceed much like for the single layer of shallow water fluid in Section 54.4. In particular, quasi-geostrophic scaling from Section 54.4.1 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. For this purpose, recall the non-dimensional momentum and continuity equations from Section 55.2.8

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}_{\hat{z}}) \hat{\mathbf{u}} + \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} \right] + \hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla}_z \hat{\phi} \quad (56.1a)$$

$$\frac{\partial \hat{\phi}}{\partial \hat{z}} = \hat{b} \quad (56.1b)$$

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0 \quad (56.1c)$$

$$Ro F \frac{D\hat{b}}{Dt} + \hat{w} = 0. \quad (56.1d)$$

We expand the prognostic variables in an asymptotic series in Rossby number

$$\hat{u} = \hat{u}_0 + Ro \hat{u}_1 + Ro^2 \hat{u}_2 + \dots \quad (56.2a)$$

$$\hat{v} = \hat{v}_0 + Ro \hat{v}_1 + Ro^2 \hat{v}_2 + \dots \quad (56.2b)$$

$$\hat{w} = \hat{w}_0 + Ro \hat{w}_1 + Ro^2 \hat{w}_2 + \dots \quad (56.2c)$$

$$\hat{b} = \hat{b}_0 + Ro \hat{b}_1 + Ro^2 \hat{b}_2 + \dots \quad (56.2d)$$

along with the expansion (54.56) for the Coriolis parameter

$$\hat{\mathbf{f}} = (\hat{f}_0 + Ro \hat{\beta} \hat{y}) \hat{\mathbf{k}}, \quad (56.3)$$

where we write the vertical unit vector as $\hat{\mathbf{k}}$ rather than $\hat{\mathbf{z}}$ to reduce confusion with the dimensionless vertical coordinate $z = \hat{z} H$, and where (equation (54.57))

$$\hat{\beta} \hat{y} = \frac{\beta y}{Ro f_0} = T \beta y. \quad (56.4)$$

Furthermore, under quasi-geostrophic scaling it is important to retain the depth dependence of the Burger number through its dependence on the background stratification $N^2(z)$

$$F(z) = \left[\frac{L}{L_d} \right]^2 = \frac{1}{[N(z)]^2} \left[\frac{L f_0}{H} \right]^2. \quad (56.5)$$

Additionally, the Burger number is order unity since the horizontal length scales are on the order of the deformation radius

$$Bu = F^{-1} \sim 1 \Rightarrow L \sim L_d. \quad (56.6)$$

56.2.1 Zeroth order asymptotic equations

The zeroth order asymptotic equations take the form

$$\hat{\mathbf{f}}_0 \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla}_z \hat{\phi}_0 \quad (56.7a)$$

$$\frac{\partial \hat{\phi}_0}{\partial \hat{z}} = \hat{b}_0 \quad (56.7b)$$

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_0 + \frac{\partial \hat{w}_0}{\partial \hat{z}} = 0 \quad (56.7c)$$

$$\hat{w}_0 = 0. \quad (56.7d)$$

The first equation represents f -plane geostrophy, which means that the horizontal velocity has zero divergence

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_0 = 0. \quad (56.8)$$

Equation (56.7b) means the zeroth order buoyancy determines the zeroth order hydrostatic pressure. Since the horizontal velocity has zero divergence, the continuity equation (56.7c) means that the vertical velocity is depth independent

$$\frac{\partial \hat{w}_0}{\partial \hat{z}} = 0. \quad (56.9)$$

If it vanishes somewhere, such as a solid boundary, then it vanishes everywhere. This is a manifestation of the Taylor-Proudman theorem (see Section 29.4.3). Indeed, a vanishing \hat{w}_0 is required

by the zeroth-order buoyancy equation (56.7d). Hence, the non-dimensional velocity has a nonzero contribution only at order Ro^1

$$\hat{w} = Ro \hat{w}_1 + Ro^2 \hat{w}_2 + \dots, \quad (56.10)$$

thus manifesting the vertical stiffening of fluid columns found in rotating fluids. Hence, the dimensionful vertical velocity has the asymptotic expansion

$$w = W \hat{w} = W Ro (\hat{w}_1 + Ro \hat{w}_2 + \dots), \quad (56.11)$$

so that to leading to order Ro^1

$$\hat{w}_1 = \frac{w}{WRo}. \quad (56.12)$$

Since the zeroth-order velocity is non-divergent, we can introduce a geostrophic streamfunction

$$\hat{u}_0 = -\frac{\partial \hat{\psi}_0}{\partial \hat{y}} \quad \hat{v}_0 = \frac{\partial \hat{\psi}_0}{\partial \hat{x}} \quad \hat{\zeta}_0 = \hat{\nabla}^2 \hat{\psi}_0, \quad (56.13)$$

where the zeroth-order streamfunction is the ratio of the zeroth order pressure to zeroth order Coriolis parameter

$$\hat{\psi}_0 = \frac{\hat{\phi}_0}{\hat{f}_0}. \quad (56.14)$$

Note also that the zeroth-order system satisfies the thermal wind balance

$$\hat{f}_0 \hat{z} \wedge \frac{\partial \hat{u}_0}{\partial \hat{z}} = -\hat{\nabla}_z \hat{b}_0. \quad (56.15)$$

Finally, note that the zeroth order buoyancy is related to the streamfunction through the hydrostatic balance

$$\hat{b}_0 = \frac{\partial \hat{\phi}_0}{\partial \hat{z}} = \hat{f}_0 \frac{\partial \hat{\psi}_0}{\partial \hat{z}}. \quad (56.16)$$

56.2.2 First order asymptotic equations

The zeroth order equations do not render a prognostic equation, for which we need to consider equations at order Ro^1

$$\frac{D_0 \hat{u}_0}{Dt} + \hat{f}_0 \hat{k} \wedge \hat{u}_1 + \hat{\beta} \hat{y} \hat{k} \wedge \hat{u}_0 = -\hat{\nabla}_z \hat{\phi}_1 \quad (56.17a)$$

$$\hat{\nabla}_z \cdot \hat{u}_1 + \frac{\partial \hat{w}_1}{\partial \hat{z}} = 0 \quad (56.17b)$$

$$F \frac{D_0 \hat{b}_0}{Dt} + \hat{w}_1 = 0. \quad (56.17c)$$

At this order, the material time derivative makes use *only* of the zeroth order geostrophic horizontal velocity

$$\frac{D_0}{Dt} = \frac{\partial}{\partial t} + \hat{u}_0 \cdot \hat{\nabla}. \quad (56.18)$$

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 QG potential vorticity equation.

Taking the curl of the momentum equation (56.17a) eliminates the pressure gradient, $\hat{\nabla}\hat{\phi}_1$, thus producing the vorticity equation

$$\frac{\partial\hat{\zeta}_0}{\partial\hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \hat{\nabla}_z \cdot \mathbf{u}_1. \quad (56.19)$$

We make use of the continuity equation (56.17b) to eliminate the horizontal convergence

$$\frac{\partial\hat{\zeta}_0}{\partial\hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = \hat{f}_0 \frac{\partial\hat{w}_1}{\partial\hat{z}}. \quad (56.20)$$

The right hand side represents the contribution to vorticity evolution from stretching by planetary rotation. We can now eliminate the vertical velocity through the buoyancy equation (56.17c). When doing so, it is important to keep the $F(z)$ depth dependence according to equation (56.5), with this depth dependence arising from the background stratification. The resulting vorticity equation is

$$\frac{\partial\hat{\zeta}_0}{\partial\hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \frac{\partial}{\partial\hat{z}} \left[F \frac{D_0 \hat{b}_0}{D\hat{t}} \right]. \quad (56.21)$$

We now use the identity

$$\frac{\partial}{\partial\hat{z}} \left[F(\hat{z}) \frac{D_0 \hat{b}_0}{D\hat{t}} \right] = \frac{\partial}{\partial\hat{z}} \left[F \left(\frac{\partial}{\partial\hat{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla} \right) \hat{b}_0 \right] \quad (56.22a)$$

$$= \frac{D_0}{D\hat{t}} \left[\frac{\partial}{\partial\hat{z}} (F \hat{b}_0) \right] + F \frac{\partial\hat{\mathbf{u}}_0}{\partial\hat{z}} \cdot \hat{\nabla}_z \hat{b}_0 \quad (56.22b)$$

$$= \frac{D_0}{D\hat{t}} \left[\frac{\partial}{\partial\hat{z}} (F \hat{b}_0) \right], \quad (56.22c)$$

where we set

$$\frac{\partial\hat{\mathbf{u}}_0}{\partial\hat{z}} \cdot \hat{\nabla}_z \hat{b}_0 = 0 \quad (56.23)$$

since the zeroth-order velocity maintains thermal wind balance (56.15). Bringing terms together then leads to the material conservation equation for quasi-geostrophic potential vorticity

$$\frac{D_0}{D\hat{t}} \left[\hat{\zeta}_0 + \hat{\beta} \hat{y} + \hat{f}_0 \frac{\partial}{\partial\hat{z}} (F \hat{b}_0) \right] = 0. \quad (56.24)$$

56.2.3 Dimensional QG-PV equation

From equation (56.24), we identify the non-dimensional QG-PV

$$\hat{q} = \hat{\zeta} + \hat{\beta} \hat{y} + \hat{f}_0 \frac{\partial(F \hat{b}_0)}{\partial\hat{z}}, \quad (56.25)$$

where we dropped the 0 asymptotic subscript for brevity. Introducing dimensional quantities to the right hand side yields (recall $\hat{f}_0 = 1$ and $\hat{\beta} \hat{y} = T \beta y = (L/U) \beta y$)

$$\hat{q} = \frac{L}{U} [\zeta + \beta y] + \frac{\partial}{\partial z} \left[\frac{H F b'}{B} \right]. \quad (56.26)$$

The scale for the fluctuating buoyancy is given by equation (55.23)

$$B = \frac{f_0 U L}{H}, \quad (56.27)$$

and the inverse Burger number is given by equation (56.5)

$$F(z) = \left[\frac{L}{L_d} \right]^2 = \frac{1}{[N(z)]^2} \left[\frac{L f_0}{H} \right]^2. \quad (56.28)$$

These terms then yield for the non-dimensional PV

$$\hat{q} = \frac{L}{U} (\zeta + \beta y) + \frac{\partial}{\partial z} \left[\frac{H F b'}{B} \right] \quad (56.29a)$$

$$= \frac{L}{U} [\zeta + \beta y] + \frac{1}{Ro} \left[\frac{\partial}{\partial z} \left(\frac{b'}{N^2} \right) \right] \quad (56.29b)$$

$$= \frac{L}{U} [\nabla_z^2 \psi + \beta y] + \frac{f_0}{Ro} \left[\frac{\partial}{\partial z} \left(\frac{1}{N^2} \frac{\partial \psi}{\partial z} \right) \right], \quad (56.29c)$$

where we introduced the QG streamfunction

$$\psi = \frac{\phi'}{f_0} \quad (56.30)$$

for the final equality, and made use of the identities

$$u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x} \quad \zeta = \nabla_z^2 \psi \quad b' = f_0 \left[\frac{\partial \psi}{\partial z} \right]. \quad (56.31)$$

We thus identify the dimensional QG-PV

$$q = Ro f_0 \hat{q} \quad (56.32a)$$

$$= \zeta + \beta y + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] \quad (56.32b)$$

$$= \nabla_z^2 \psi + \beta y + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right]. \quad (56.32c)$$

We can add a constant to q without changing the dynamics. Consequently, some authors like to add f_0 in which case

$$q = \underbrace{\nabla_z^2 \psi}_{\text{relative vorticity}} + \underbrace{f_0 + \beta y}_{\text{planetary vorticity}} + \underbrace{\frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right]}_{\text{stretching by } f}. \quad (56.33)$$

There are three contributions to the QG-PV:

- relative vorticity of the geostrophic flow;
- planetary vorticity due to the rotation of the reference frame;
- stretching due to motion on the rotating planet (see equation (56.20) to be reminded why the third term represents stretching).

56.2.4 Properties of the steady state flow field

The geostrophic velocity takes the following form in terms of the QG streamfunction

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi. \quad (56.34)$$

This equality then allows us to write the following equivalent forms for the material time derivative of QG-PV

$$\frac{Dq}{Dt} = \frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q \quad (56.35a)$$

$$= \frac{\partial q}{\partial t} + (\hat{\mathbf{z}} \wedge \nabla \psi) \cdot \nabla q \quad (56.35b)$$

$$= \frac{\partial q}{\partial t} + (\nabla \psi \wedge \nabla q) \cdot \hat{\mathbf{z}} \quad (56.35c)$$

$$= \frac{\partial q}{\partial t} + J(\psi, q), \quad (56.35d)$$

where the final equality introduced the Jacobian operator

$$J(\psi, q) = \frac{\partial \psi}{\partial x} \frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial q}{\partial x}. \quad (56.36)$$

For a perfect fluid, in which $Dq/Dt = 0$, a steady state (zero Eulerian time derivative) is realized when

$$\mathbf{u} \cdot \nabla q = (\nabla \psi \wedge \nabla q) \cdot \hat{\mathbf{z}} = J(\psi, q) = 0. \quad (56.37)$$

The first expression says that the velocity field is aligned parallel to surfaces of constant q , with the second expression saying the same. We are ensured that these equalities hold if the streamfunction is a function only of the potential vorticity

$$\psi = \psi(q) \Rightarrow J(\psi, q) = 0. \quad (56.38)$$

56.2.5 Constant background buoyancy frequency

Consider the QG-PV for the special case of a constant background buoyancy frequency,

$$N^2 = \text{constant}. \quad (56.39)$$

For this case the QG-PV in equation (56.32c) can be written

$$q = \beta y + \nabla_z^2 \psi + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] \quad (56.40a)$$

$$= \beta y + \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{f_0^2}{N^2} \frac{\partial^2 \psi}{\partial z^2} \quad (56.40b)$$

$$= \beta y + \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial \tilde{z}^2}. \quad (56.40c)$$

For the final equality we introduced the vertical coordinate

$$\tilde{z} = \frac{N}{f} z. \quad (56.41)$$

Since $N/f \gg 1$ for the stably stratified flows considered in QG, we term \tilde{z} a *stretched* vertical coordinate.

56.2.6 Buoyancy advection at the boundaries

We need boundary conditions in order to invert the elliptic QG-PV equation (56.33) to solve for the streamfunction ψ . For lateral boundaries, one may choose periodicity, whereby the boundaries are in effect absent. Alternatively, we may choose to set the normal component of the flow to zero for the inviscid case. The top and bottom boundaries are less trivial and require some care.

Returning to the Boussinesq equations (55.1a)-(??), we focus on the buoyancy equation, written as in equation (??)

$$\frac{Db'}{Dt} + w N^2 = 0, \quad (56.42)$$

where advection is via the horizontal geostrophic currents. Inserting the geostrophic streamfunction

$$b' = f_0 \frac{\partial \psi}{\partial z} \quad (56.43)$$

leads to

$$\frac{\partial}{\partial t} \left[\frac{\partial \psi}{\partial z} \right] + \mathbf{u} \cdot \nabla \left[\frac{\partial \psi}{\partial z} \right] + w N^2 = 0. \quad (56.44)$$

Consider two contributions to nonzero vertical velocity at a boundary. The first arises from slopes in the topography, $\nabla \eta_b \neq 0$. The no-normal flow condition means that at the boundary, the velocity is constrained so that

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \Rightarrow w = \mathbf{u} \cdot \nabla \eta_b, \quad (56.45)$$

where $\hat{\mathbf{n}}$ is the outward normal to the boundary (Section 17.4.1). The second contribution arises from Ekman pumping or suction (Section 31.1). In general, flow in an Ekman layer generates vorticity due to curl in the boundary stresses. We abstract this process by stating that at the boundary of the Ekman layer and the interior flow, the vertical velocity component takes on a value (see Section 14.1.1 of [Vallis \(2006\)](#))

$$w_{\text{Ekman}} = r \nabla_z^2 \psi, \quad (56.46)$$

where r is a length scale proportional to the Ekman layer thickness, and $\nabla_z^2 \psi$ is the quasi-geostrophic vorticity. Bringing these two effects together leads to the boundary condition

$$w = \mathbf{u} \cdot \nabla \eta_b + r \nabla_z^2 \psi. \quad (56.47)$$

Using this expression in the buoyancy equation (56.44) leads to the boundary evolution of buoyancy

$$\frac{\partial}{\partial t} \left[\frac{\partial \psi}{\partial z} \right] + \mathbf{u} \cdot \nabla \left[\frac{\partial \psi}{\partial z} + N^2 \eta_b \right] + N^2 r \nabla_z^2 \psi = 0. \quad (56.48)$$

We say that the QG fluid system is characterized by horizontal advection of PV in the interior, and advection of buoyancy on the boundaries.

56.3 Dimensions of various forms for potential vorticity

The following dimensions are taken by the various forms of potential vorticity seen thus far in these notes.

- Shallow water PV:

$$\frac{\zeta + f}{h} [\equiv] (\text{time} \times \text{length})^{-1}. \quad (56.49)$$

- Entropic Ertel PV based on potential temperature,

$$\frac{\omega_a \cdot \nabla \theta}{\rho} [\equiv] \frac{\text{length}^2 \times \text{temperature}}{\text{time} \times \text{mass}}, \quad (56.50)$$

where θ is the potential temperature.

- Ertel PV with buoyancy in a Boussinesq fluid

$$\omega_a \cdot \nabla b [\equiv] \text{time}^{-3}. \quad (56.51)$$

- Quasi-geostrophic PV

$$q = \nabla_z^2 \psi + \beta y + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] [\equiv] \text{time}^{-1}. \quad (56.52)$$

56.4 Connecting QG-PV to Ertel PV

Equations describing both the PG and QG systems can be encapsulated by the material conservation of PG or QG potential vorticity. We here determine how potential vorticity for PG and QG flows relates to the Ertel PV from Chapter 50.

56.4.1 Shallow water layer

The Ertel PV for a single layer of shallow water fluid is (Section 47.3)

$$Q = \frac{f + \zeta}{h}, \quad (56.53)$$

where

$$h = H + \Delta\eta - \eta_b \quad (56.54)$$

is the thickness of the layer (see Figure 42.1). We now consider the limit as the Rossby number is small, and one of the following two regimes for the thickness.

- PG: free surface height and bottom topography undulations are on the order of the resting depth H .
- QG: free surface height undulations ($\Delta\eta$) are small, and bottom topography deviations are small (η_b can be neglected);

The PG limit is simplest, in which case we merely drop the relative vorticity and keep the full form of the layer thickness

$$Q_{\text{PG}} = \frac{f}{h}. \quad (56.55)$$

The QG limit requires a bit more algebra, whereby

$$Q = \frac{\zeta + f}{h} \quad (56.56a)$$

$$= \frac{\zeta + f}{H(1 + \Delta\eta/H)} \quad (56.56b)$$

$$\approx \frac{1}{H} (\zeta + f) \left[1 - \frac{\Delta\eta}{H} \right] \quad (56.56c)$$

$$\approx \frac{1}{H} \left[\zeta + f_0 + \beta y - f_0 \frac{\Delta\eta}{H} \right]. \quad (56.56d)$$

56.4.2 Continuously stratified hydrostatic Boussinesq fluid

We now consider the connection between Ertel PV and QG-PV for the continuously stratified hydrostatic Boussinesq fluid. For this purpose, make use of the Ertel PV derived in Exercise 50.1, for which

$$Q = \boldsymbol{\omega}_a \cdot \nabla b = \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + (\zeta + f) \frac{\partial b}{\partial z}. \quad (56.57)$$

Now split the buoyancy into a depth dependent background and a deviation from the background

$$b = \tilde{b}(z) + b'(x, y, z, t), \quad (56.58)$$

and write its vertical derivative as

$$\frac{\partial b}{\partial z} = N^2 + \frac{\partial b'}{\partial z}, \quad (56.59)$$

where

$$N^2 = \frac{\partial \tilde{b}}{\partial z} \quad (56.60)$$

is the squared buoyancy frequency for the background buoyancy field. Rather than introducing non-dimensional variables¹, we work more briefly by arranging the terms in the PV according to their QG scaling in terms of the Rossby number

$$Q = [f_0 N^2] + \left[(\beta y + \zeta) N^2 + f_0 \frac{\partial b'}{\partial z} \right] + \left[(\beta y + \zeta) \frac{\partial b'}{\partial z} + \frac{\partial u}{\partial z} \frac{\partial b'}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b'}{\partial x} \right]. \quad (56.61)$$

We drop the third bracket term, as it is order Ro^2 , and write

$$Q = \tilde{Q} + N^2 q_*, \quad (56.62)$$

where

$$\tilde{Q} = f_0 N^2 \quad (56.63)$$

is the f-plane planetary geostrophic PV, and

$$q_* = \beta y + \zeta + \frac{f_0}{N^2} \frac{\partial b'}{\partial z}. \quad (56.64)$$

The material conservation of PV now takes the form

$$N^2 \frac{Dq_*}{Dt} + w \left[1 + \frac{q_*}{f_0} \right] \frac{\partial \tilde{Q}}{\partial z} = 0 \quad (56.65)$$

where advection in the material time derivative operator is now assumed to occur from the horizontal geostrophic velocity. We drop the term q_*/f_0 as it is small, and divide by N^2 , to have

$$\frac{Dq_*}{Dt} + \frac{w}{N^2} \frac{\partial \tilde{Q}}{\partial z} = 0. \quad (56.66)$$

To eliminate the vertical velocity component, we introduce the buoyancy equation

$$\frac{Db'}{Dt} + w N^2 = 0, \quad (56.67)$$

¹SMG: Should introduce non-dimensional variables in future versions of the notes to enhance the analysis from *Vallis (2006)* Section 5.5.1.

so that

$$\frac{Dq_*}{Dt} - \frac{f_0}{N^4} \frac{\partial N^2}{\partial z} \frac{Db'}{Dt} = 0. \quad (56.68)$$

Writing

$$\frac{\partial}{\partial z} \left[\frac{1}{N^2} \right] = -\frac{1}{N^4} \frac{\partial N^2}{\partial z} \quad (56.69)$$

leads to

$$\frac{Dq_*}{Dt} + f_0 \left[\frac{\partial N^{-2}}{\partial z} \right] \frac{Db'}{Dt} = 0. \quad (56.70)$$

Since the material time derivative operator only involves horizontal advection, we can merge these two terms to render

$$\frac{Dq_*}{Dt} + f_0 \left(\frac{\partial N^{-2}}{\partial z} \right) \frac{Db'}{Dt} = \frac{D}{Dt} \left[q_* + f_0 b' \left(\frac{\partial N^{-2}}{\partial z} \right) \right] \quad (56.71a)$$

$$= \frac{D}{Dt} \left[\beta y + \zeta + \frac{f_0}{N^2} \left(\frac{\partial b'}{\partial z} \right) + f_0 b' \left(\frac{\partial N^{-2}}{\partial z} \right) \right] \quad (56.71b)$$

$$= \frac{D}{Dt} \left[\beta y + \zeta + f_0 \frac{\partial}{\partial z} \left(\frac{b'}{N^2} \right) \right] \quad (56.71c)$$

$$= 0. \quad (56.71d)$$

The term inside the bracket is the QG-PV given by equation (56.33).

56.5 Energetics

Consider the QG system with flat top and flat bottom boundaries, and assume for the lateral directions either periodicity or constant streamfunction on solid boundaries. The QG vorticity equation and buoyancy equations are given by

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta = f_0 \frac{\partial w}{\partial z} \quad (56.72a)$$

$$\frac{\partial b}{\partial t} + \mathbf{u} \cdot \nabla b = -w N^2, \quad (56.72b)$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \quad b' = f_0 \frac{\partial \psi}{\partial z} \quad u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x} \quad \zeta = \nabla_z^2 \psi. \quad (56.73)$$

56.5.1 Kinetic energy

The kinetic energy per mass for the total fluid domain (assumed to be a constant volume material domain) is given by the integral

$$\mathcal{K} = \frac{1}{2} \int (\mathbf{u} \cdot \mathbf{u}) dV = \frac{1}{2} \int (\nabla \psi \cdot \nabla \psi) dV, \quad (56.74)$$

and its time derivative is

$$\frac{d\mathcal{K}}{dt} = \int \nabla \psi \cdot \nabla \left[\frac{\partial \psi}{\partial t} \right] dV. \quad (56.75)$$

For this result, we assumed the fluid domain has a constant volume to thus allow the time derivative to move inside the integral without introducing boundary terms. Manipulation renders

$$\frac{d\mathcal{K}}{dt} = \int \nabla\psi \cdot \nabla \left[\frac{\partial\psi}{\partial t} \right] dV \quad (56.76a)$$

$$= \int [\nabla \cdot [\psi \nabla(\partial\psi/\partial t)] - \psi \partial(\nabla_z^2\psi)/\partial t] dV \quad (56.76b)$$

$$= - \int \psi \frac{\partial\zeta}{\partial t} dV, \quad (56.76c)$$

where we dropped the lateral boundary term and introduced relative vorticity. Use of the vorticity equation (56.72a) yields

$$\frac{d\mathcal{K}}{dt} = - \int \psi \frac{\partial\zeta}{\partial t} dV \quad (56.77a)$$

$$= \int \psi [\mathbf{u} \cdot \nabla\zeta - f_0 \partial_z w] dV. \quad (56.77b)$$

The first term vanishes, since

$$\int \psi [\mathbf{u} \cdot \nabla\zeta] dV = \int \psi \nabla \cdot (\mathbf{u}\zeta) dV \quad (56.78a)$$

$$= \int [\nabla \cdot (\psi \mathbf{u}\zeta) - \nabla\psi \cdot \mathbf{u}\zeta] dV \quad (56.78b)$$

$$= 0, \quad (56.78c)$$

where the boundary term vanishes and $\mathbf{u} \cdot \nabla\psi = 0$ since ψ is the streamfunction for the horizontal geostrophic flow. We are thus left with the expression for the kinetic energy evolution

$$\frac{d\mathcal{K}}{dt} = - \int \psi f_0 \frac{\partial w}{\partial z} dV. \quad (56.79)$$

Since the top and bottom are assumed flat, the vertical velocity vanishes on these boundaries (rigid lid top and solid bottom), in which case we can write

$$\frac{d\mathcal{K}}{dt} = - \int \psi f_0 \frac{\partial w}{\partial z} dV \quad (56.80a)$$

$$= - \int f_0 \left[\frac{\partial(w\psi)}{\partial z} - w \frac{\partial\psi}{\partial z} \right] dV \quad (56.80b)$$

$$= \int f_0 w \frac{\partial\psi}{\partial z} dV, \quad (56.80c)$$

where we dropped the boundary term given the rigid boundaries. Making use of

$$b' = f_0 \partial\psi/\partial z \quad (56.81)$$

leads to

$$\frac{d\mathcal{K}}{dt} = \int w b' dV. \quad (56.82)$$

Kinetic energy thus increases when vertical motion is positively correlated with anomalous 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 full fluid system, as discussed in Section 24.1. It is reassuring to see the same behavior in the quasi-geostrophic system.

56.5.2 Available potential energy

Available potential energy was introduced in Section 3.10 of [Vallis \(2006\)](#). Specializing that expression to the QG fluid leads to the available potential energy

$$\mathcal{A} = \frac{1}{2} \int \left[\frac{f_0}{N} \frac{\partial \psi}{\partial z} \right]^2 dV. \quad (56.83)$$

Taking a time derivative leads to

$$\frac{d\mathcal{A}}{dt} = \int \left[\frac{f_0}{N} \right]^2 \frac{\partial \psi}{\partial z} \frac{\partial}{\partial t} \frac{\partial \psi}{\partial z} dV \quad (56.84a)$$

$$= \int \frac{f_0}{N^2} \frac{\partial \psi}{\partial z} [-w N^2 - \nabla \cdot (\mathbf{u} b)] dV, \quad (56.84b)$$

where we used the buoyancy equation (56.72b) for the second equality. Recall that the divergence operator acts just in the horizontal since \mathbf{u} is the horizontal velocity. The second term vanishes, since

$$\int \frac{f_0}{N^2} \frac{\partial \psi}{\partial z} [\nabla \cdot (\mathbf{u} b)] dV = \int \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] \mathbf{u} \cdot \nabla \left[\frac{\partial \psi}{\partial z} \right] dV \quad (56.85a)$$

$$= \frac{1}{2} \int \nabla \cdot \left[\mathbf{u} \left(\frac{f_0}{N} \frac{\partial \psi}{\partial z} \right)^2 \right] dV \quad (56.85b)$$

$$= 0. \quad (56.85c)$$

Consequently, the APE changes according to

$$\frac{d\mathcal{A}}{dt} = - \int w f_0 \frac{\partial \psi}{\partial z} dV = - \int w b' dV. \quad (56.86)$$

56.5.3 Energy conversion

Notice how 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, since as kinetic energy increases through buoyancy work, available potential energy decreases

$$\frac{d(\mathcal{K} + \mathcal{A})}{dt} = 0. \quad (56.87)$$

The buoyancy work conversion term is given by

$$\text{buoyancy work} = \int w f_0 \frac{\partial \psi}{\partial z} dV = \int w b' dV, \quad (56.88)$$

which has the same form as that encountered for the conversion between potential and kinetic energy in the unapproximated equations (Section 24.1)

56.5.4 Scaling APE and KE

The scale for the kinetic energy is given by

$$\mathcal{K} = \frac{1}{2} \int (\nabla \psi \cdot \nabla \psi) dV \sim L^{-2} \Psi V \quad (56.89)$$

and the APE scale is

$$\mathcal{A} = \frac{1}{2} \int \left[\frac{f_0}{N} \frac{\partial \psi}{\partial z} \right]^2 dV \sim L_d^{-2} \Psi V, \quad (56.90)$$

where we wrote Ψ for the streamfunction scale, and V for the domain volume. Taking the ratio yields

$$\frac{\mathcal{K}}{\mathcal{A}} \sim \left[\frac{L_d}{L} \right]^2 = \left[\frac{H}{L} \right]^2 \left[\frac{N}{f_0} \right]^2 = Bu. \quad (56.91)$$

The Burger number is hence the ratio of the kinetic energy scale to the available potential energy scale. A large Burger number means that the horizontal scales of the flow are smaller than the deformation radius, in which case the QG 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 QG scaling), the Burger number is less than unity, in which case the QG dynamics is dominated by available potential energy.

56.6 Exercises

EXERCISE 56.1: QG-PV EVOLUTION WITH VERTICAL FRICTION

The first part of this problem involves elements of the asymptotic method used for deriving the QG equations, only now with the advent of a non-zero friction. Use is made to incorporate the non-dimensionalization detailed in Section 31.1, which provides a detailed discussion of the Ekman number and Ekman layers. The second part of the question makes use of the thermal wind balance to connect vertical viscous momentum transfer to horizontal buoyancy transfer.

- (a) Derive the material evolution equation for QG-PV 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

$$\mathbf{F} = \frac{\partial}{\partial z} \left[\nu \frac{\partial \mathbf{u}}{\partial z} \right], \quad (56.92)$$

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 QG-PV in which we drop relative vorticity (i.e., QG-PV is dominated by planetary vorticity and stretching). Determine the form for the vertical eddy viscosity so that the approximate form of QG-PV is laterally diffused via

$$\frac{Dq^{\text{approx}}}{Dt} = A \nabla_z^2 q^{\text{approx}}, \quad (56.93)$$

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 56.2: TRADITIONAL FORM OF THE QUASI-GEOSTROPHIC ω -EQUATION

As discussed in Section 56.2, the vertical component to the velocity is non-zero only at first order in Rossby number, whereas the zeroth order flow is horizontal and geostrophic. To time step the horizontal geostrophic flow it is not necessary to explicitly compute the vertical velocity. However, the vertical velocity is non-zero and there are cases where it is of interest. In this exercise we derive

the ω -equation for quasi-geostrophic flow, thus providing a diagnostic expression for the vertical velocity. The name for this equation originates from the atmospheric community where ω is the common symbol for transport across pressure surfaces. Here, we make use of the Boussinesq system so that the vertical velocity component is across depth surfaces.

An outline for the derivation of the traditional form for the ω -equation is given in Section 5.4 of [Vallis \(2017\)](#) for the anelastic version of quasi-geostrophy and for the f -plane. Here we work with the Boussinesq system and consider a β -plane. Nonetheless, the solution is nearly the same as in [Vallis \(2017\)](#). Hence, your job throughout this exercise is to fully explain the derivation and show each of the relevant steps.

- (a) From Section 56.2.2, the non-dimensional velocity, continuity, buoyancy and vorticity equations for β -plane flow valid at order Ro^1 are given by

$$\frac{D_0 \hat{\mathbf{u}}_0}{Dt} + \hat{f}_0 \hat{\mathbf{k}} \wedge \hat{\mathbf{u}}_1 + \hat{\beta} \hat{y} \hat{\mathbf{k}} \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla}_z \hat{\phi}_1 \quad (56.94a)$$

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_1 + \frac{\partial \hat{w}_1}{\partial \hat{z}} = 0 \quad (56.94b)$$

$$F \frac{D_0 \hat{b}_0}{Dt} + \hat{w}_1 = 0 \quad (56.94c)$$

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = \hat{f}_0 \frac{\partial \hat{w}_1}{\partial \hat{z}}. \quad (56.94d)$$

Write the dimensional form of these four equations, showing all steps when moving from the dimensionless to dimensional equations. For the horizontal velocity, write

$$\mathbf{u}_g = U \hat{\mathbf{u}}_0 \quad (56.95)$$

for that portion of the geostrophic flow that is horizontally non-divergent, and

$$\mathbf{u}_{ag} = U(\hat{\mathbf{u}} - \hat{\mathbf{u}}_0) \approx Ro U \hat{\mathbf{u}}_1 \quad (56.96a)$$

$$\phi_{ag} = \Phi(\hat{\phi} - \hat{\phi}_0) \approx Ro \Phi \hat{\phi}_1. \quad (56.96b)$$

for an estimate of the ageostrophic portion of the velocity and pressure.

- (b) Write the dimensional buoyancy and vorticity equations using the geostrophic streamfunction. Introduce the Jacobian operator for the advection.
- (c) Cross-multiply the dimensional buoyancy and vorticity equations to eliminate the time derivative, thus revealing a diagnostic equation for the vertical velocity that is valid to order Ro^1 .
- (d) The equation for the vertical velocity takes the form

$$\mathcal{L}w = \sigma, \quad (56.97)$$

where

$$\mathcal{L} = N^2 \nabla_z^2 + f_0^2 \frac{\partial^2}{\partial z^2} \quad (56.98)$$

is a linear partial differential operator and

$$\sigma = f_0 J(\psi, \zeta + \beta y) - \nabla^2 J(\psi, b) \quad (56.99)$$

is a source term. The source is a function of the geostrophic flow and the buoyancy field. For vertically stable stratification, $N^2 > 0$, characterize the differential operator \mathcal{L} according to the elliptic, hyperbolic, or parabolic classes discussed in Section 5.3. What about when $N^2 < 0$? What new physical phenomena do you expect when $N^2 < 0$?

EXERCISE 56.3: HOSKINS' FORM OF THE QUASI-GEOSTROPHIC ω -EQUATION

We here rederive the ω -equation from Exercise 56.2 using methods introduced by [Hoskins et al. \(1978\)](#). It is not necessary to have solved Exercise 56.2 to solve the present exercise.

Hoskins' approach reveals an insightful form for the source function contributing to vertical motion. As in Exercise 56.2, we work with the adiabatic and hydrostatic Boussinesq system (see Section 55.2)

$$\frac{Du}{Dt} - fv = -\frac{\partial \phi}{\partial x} \quad (56.100a)$$

$$\frac{Dv}{Dt} + fu = -\frac{\partial \phi}{\partial y} \quad (56.100b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (56.100c)$$

$$\frac{Db}{Dt} = 0 \quad (56.100d)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (56.100e)$$

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (56.100f)$$

In Exercise 56.2, we derived the ω -equation making use of the buoyancy equation and vorticity equation. [Hoskins et al. \(1978\)](#) worked with the momentum equation rather than the vorticity equation. For this purpose, rather than consider an asymptotic expansion, Hoskins exactly decomposed the horizontal velocity into its geostrophic and ageostrophic components

$$\mathbf{u} = \mathbf{u}_g + \mathbf{u}_{ag}, \quad (56.101)$$

with the geostrophic velocity balancing the horizontal gradient of the full pressure field

$$f\mathbf{u}_g = \hat{\mathbf{z}} \wedge \nabla \phi. \quad (56.102)$$

This definition for \mathbf{u}_g is distinct from that arising from an asymptotic expansion, whereby the geostrophic velocity is the zeroth order term balancing the zeroth order pressure gradient (see Section 56.2.1 or Exercise 56.2). We are generally able to access the full hydrostatic pressure field through knowledge of the buoyancy field, in which case there is no need to make an asymptotic expansion of pressure.

Hoskins' definition for the geostrophic velocity brings the horizontal momentum equations into the rather elegant form

$$\frac{D\mathbf{u}}{Dt} + f\hat{\mathbf{z}} \wedge \mathbf{u}_{ag} = 0, \quad (56.103)$$

with the pressure gradient annihilated since it exactly balances the geostrophic velocity. In this manner, the material evolution of horizontal velocity is determined solely by the ageostrophic Coriolis acceleration. Again, there has been no approximation made thus far. Rather, we have only introduced a strategic decomposition of the velocity field as per Hoskins.

At this point we make the quasi-geostrophic approximation by setting the momentum equation equal to

$$\frac{\partial \mathbf{u}_g}{\partial t} + (\mathbf{u}_g \cdot \nabla_z) \mathbf{u}_g + f\hat{\mathbf{z}} \wedge \mathbf{u}_{ag} = 0 \quad (56.104)$$

and the buoyancy equation equal to

$$\frac{\partial b}{\partial t} + \mathbf{u}_g \cdot \nabla_z b + N^2 w = 0, \quad (56.105)$$

where $N^2(z)$ is a prescribed static background stratification. That is, both the horizontal geostrophic velocity and the buoyancy are advected just by the geostrophic velocity. The buoyancy equation is the same as derived to order Ro^1 using asymptotic methods (Section 56.2.2). However, Hoskins' momentum equation (56.104) has no pressure gradient on the right hand side, whereas an asymptotic approach has contributions from higher order pressure gradients (Section 56.2.2). Hoskins' approach dispenses with such pressure terms by defining the geostrophic velocity using the full pressure field.

- (a) Show that the evolution of horizontal buoyancy gradients by the horizontal geostrophic currents can be written

$$\left[\frac{\partial}{\partial t} + \mathbf{u}_g \cdot \nabla_z \right] |\nabla_z b|^2 = \mathbf{Q} \cdot \nabla_z b. \quad (56.106)$$

Hence, horizontal buoyancy gradients grow in magnitude in regions where the horizontal buoyancy gradient projects positively onto the \mathbf{Q} -vector. Write the expression for the vector \mathbf{Q} . Hint: A general version of this result was derived in Exercise 16.1.

- (b) Show that the quasi-geostrophic ω -equation on an f -plane can be written

$$N^2 \nabla_z^2 w + f_0^2 \frac{\partial^2 w}{\partial z^2} = 2 \nabla_z \cdot \mathbf{Q}. \quad (56.107)$$

We see that the source for vertical motion is the divergence of the \mathbf{Q} -vector. This formula offers useful insight into the origin of vertical motion, with [Hoskins et al. \(1978\)](#) offering examples. For this part of the exercise, you are to fully explain the derivation of equation (56.107) and show each of the relevant steps.

- (c) The equation (56.107) for the vertical velocity takes the form

$$\mathcal{L}w = \sigma, \quad (56.108)$$

where

$$\mathcal{L} = N^2 \nabla_z^2 + f_0^2 \frac{\partial^2}{\partial z^2} \quad (56.109)$$

is a linear partial differential operator and

$$\sigma = 2 \nabla_z \cdot \mathbf{Q} \quad (56.110)$$

is a source term. The source term is a function of the geostrophic flow and the buoyancy field. For vertically stable stratification, $N^2 > 0$, characterize the differential operator \mathcal{L} according to the elliptic, hyperbolic, or parabolic classes discussed in Section 5.3. What about when $N^2 < 0$? What new physical phenomena do you expect when $N^2 < 0$?

Local stability of fronts

We here study the dynamics of flows exhibiting either axial symmetry (rotating column of fluid) or horizontal symmetry in one direction (geostrophic fronts). The analysis of these systems share much in common given that they all possess a material invariant that is fundamental to the stability analysis: angular momentum for the axially symmetric column and potential momentum for the two-dimensional front. Another important assumption for our analysis of fronts concerns the f -plane. Potential momentum is a material invariant on the f -plane but not the β -plane or sphere.

Our study of these systems is motivated by fronts appearing in the atmosphere and ocean, whereby regions of strong baroclinicity are associated with jet-like flows. These flows generally exhibit secondary overturning circulations as well as a variety of instabilities. We are also motivated by the study of the stability of rotating vortices such as Gulf Stream rings and tornadoes.

57.1	Three types of instabilities	872
57.1.1	Elements of the instabilities	872
57.1.2	The nature of the base state and the perturbations	873
57.1.3	Comments	873
57.2	Centrifugal instability of cyclostrophically balanced flow	874
57.2.1	Equations of motion	874
57.2.2	Solid-body motion	875
57.2.3	Stability analysis based on energetic arguments	875
57.2.4	Stability analysis based on parcel arguments	877
57.2.5	Comments	878
57.3	Potential momentum	878
57.3.1	Linear momentum and potential momentum	879
57.3.2	Zonal potential momentum on a β -plane	880
57.3.3	Further reading	880
57.4	Horizontal inertial instability of a geostrophically balanced front	880
57.4.1	Equations of motion and equilibrium state	881
57.4.2	Stability analysis based on energetic arguments	882
57.4.3	Stability analysis based on parcel arguments	883
57.4.4	Comments	884
57.5	Symmetric (isentropic inertial) instability of a baroclinic front	885
57.5.1	Equations of motion and equilibrium state	885
57.5.2	Stability analysis based on parcel arguments	886
57.5.3	Comments	888
57.6	Secondary circulation along baroclinic fronts	888
57.6.1	Hydrostatic and Boussinesq fluid on an f -plane	889

57.6.2	Geostrophic momentum approximation	890
57.6.3	Secondary ageostrophic circulation	891
57.6.4	Ageostrophic overturning circulation for a symmetric front	892
57.6.5	Connection to potential vorticity and symmetric instability	893

57.1 Three types of instabilities

The instabilities studied in this chapter are termed *local* or *parcel*. These terms arise from the ability to identify an unstable flow profile through a local condition on the vorticity or potential vorticity, thus yielding a necessary and sufficient instability condition. The instability condition can be determined by a thought experiment where fluid parcels are displaced while materially conserving certain flow properties. The physical features of the three instabilities considered here are quite similar, thus making it convenient to study them together. Note that by using parcel arguments to develop instability conditions, one might imagine the instabilities also hold in analogous particle mechanics configurations. However, the instabilities we discuss in this chapter are fundamentally fluid dynamical since they require a pressure field to balance either the centrifugal or Coriolis accelerations. Hence, the analogy with particle mechanics does not hold.

- Summarize the perspective given in http://glossary.ametsoc.org/wiki/Inertial_instability

57.1.1 Elements of the instabilities

Here is a telescopic summary of the instabilities studied in this chapter.

- CENTRIFUGAL INSTABILITY: Consider an equilibrium flow state under inviscid cyclostrophic balance. Cyclostrophic balance arises when pressure and centrifugal accelerations are balanced, with centrifugal forces arising from curvature in the fluid parcel trajectory. The angular momentum is materially invariant when the equilibrium state is rotationally symmetric, as in an ideal circular vortex or a rotating circular tank. Flow stability is probed by horizontally displacing a rotationally symmetric circular ring of fluid parcels, with each parcel retaining its original angular momentum. If the parcels are displaced to a position where pressure and centrifugal accelerations further support the displacement, then the base state is unstable to centrifugal instability.
- HORIZONTAL INERTIAL INSTABILITY: Consider an equilibrium flow state under inviscid geostrophic balance on an f -plane. Geostrophic balance arises when pressure and Coriolis accelerations are balanced. The potential momentum is materially invariant when the equilibrium state is symmetric in a horizontal direction, as in a zonally or meridionally symmetric front. Flow stability is probed by horizontally displacing a symmetric line of fluid parcels, with each parcel retaining its original potential momentum. If the parcels are displaced to a position where pressure and Coriolis accelerations further support the displacement, then the base state is unstable to horizontal inertial instability.
- ISENTROPIC INERTIAL (SYMMETRIC) INSTABILITY: Consider an equilibrium flow state under inviscid geostrophic balance in the presence of baroclinicity. Both potential momentum and buoyancy are materially invariant when the equilibrium state is symmetric in a horizontal direction. Flow stability is probed by isentropically displacing a symmetric line of fluid parcels, with each parcel retaining its original potential momentum and buoyancy. If the parcels are displaced to a position where pressure and Coriolis accelerations further support

the displacement, then the base state is unstable to isentropic inertial instability (also called *symmetric* instability).

57.1.2 The nature of the base state and the perturbations

For our study of centrifugal instability, we make use of the shallow water system, whereas we consider the continuously stratified Boussinesq system for the horizontal and isentropic inertial instabilities. When the fluid is continuously stratified and inviscid, all motion occurs along isentropes. However, when probing for centrifugal or horizontal inertial instabilities, we examine stability to perturbations along geopotential surfaces. Such horizontal displacements generally cross isentropic surfaces in a baroclinic fluid and so comprise irreversible perturbations. The isentropic inertial instability analysis in Section 57.5 maintains the adiabatic nature of displacements when probing for instabilities. Even so, these perturbations are not concerned with maintaining a materially invariant potential vorticity. Why?

One hypothesis is that parcel/local instabilities are realized only when parcels probe a direction that breaks one of the invariants. For example, gravitational instability is based on probing the vertical direction across horizontal isentropes. Likewise, entering the wedge of instability for symmetrically unstable flow requires a parcel to leave its constant buoyancy and constant potential momentum surfaces. Granted, the thought experiment involves the parcel maintaining its materially invariant property (e.g., buoyancy, angular momentum, potential momentum) to probe the stability. Yet for a realistic case the material invariant property is immediately mixed with the surrounding environment as the parcel instability ensues. That is, the integrity of the fluid parcels is sacrificed in the process of undergoing the instability. This situation contrasts to the case with wave/global instabilities, where all the material invariants are preserved while the flow realizes its instability. The normal mode instability ensues via constructive interference between linear waves (coherent parcel motion) whose amplitude exponentially grows. There is no reliance on irreversibility to realize wave/global instabilities.

Another point of distinction between parcel/local and wave/global is that for the parcel instability, the parcel is probing a direction without modifying the base state; it tests stability of the base state without modifying that state. The existence of an instability is not dependent on the back reaction of the parcel movement onto the base state. That is, parcel/local instabilities do not arise from coherent interactions. They just arise from the overall instability of the base state to any perturbation that probes the unstable direction. In contrast, for wave/global instability, the perturbation modifies the base state in a way that supports the growth of the perturbation. For example, the two-layer baroclinic instability problem requires interactions between the layers for the instability to grow. Without this interaction between the layers, the layers then decouple and there is no unstable baroclinic wave.

57.1.3 Comments

[Holton \(1992\)](#) distinguishes between *parcel* and *wave* instabilities, whereas [Cushman-Roisin and Beckers \(2011\)](#) use the terms *local* and *global*. Canonical examples of global or wave instabilities are Kelvin-Helmholz and baroclinic. We are not concerned with those instabilities in this chapter. Rather, the three instabilities considered here are examples of local or parcel instabilities.

Because of the rather close similarities between centrifugal and inertial instability, the oceanographic literature often uses the term centrifugal instability when referring to the inertial instability considered here (e.g., see [Thomas et al. \(2013\)](#) and [McWilliams \(2016\)](#)). However, we do not follow that usage since inertial instability is *not* associated with centrifugal accelerations. Rather, it is as-

sociated with Coriolis accelerations. We thus follow the terminology of the atmospheric literature, such as detailed in the texts by [Holton \(1992\)](#) and [Markowski and Richardson \(2010\)](#), which also follows the classical fluid mechanics terminology used in the text by [Drazin and Reid \(1981\)](#). In this manner, we reserve the term *centrifugal instability* for an axisymmetric base state in cyclostrophic balance, and *inertial instability* for a two-dimensional base state in geostrophic balance.

57.2 Centrifugal instability of cyclostrophically balanced flow

Consider a shallow water fluid in a cylindrical tank rotating about its vertical axis and maintaining rotational symmetry (also termed *axisymmetric*). All dynamical fields are thus a function only of the radial distance from the rotational axis. We are interested in questions concerning flow stability as a function of the radial distribution of the angular velocity, $v^\theta(r)$. In particular, we examine stability of cyclostrophically balanced flow, defined by flow whose radial acceleration vanishes so that the radial pressure gradient balances the centrifugal acceleration. We will show that such cyclostrophic flow is stable to rotationally symmetric perturbations so long as the squared angular momentum increases radially. This system provides a useful introduction to studies of the stability of rotating vortices in the ocean and atmosphere, and it establishes analysis methods used for inertial instabilities realized in two-dimensional frontal regions.

Our use of the shallow water layer is based on convenience since we can directly connect this analysis to earlier work in Section 43.5 where we studied the angular momentum of a tank of shallow water. However, we could just as well make use of a fully baroclinic flow without adding any analytical complexity. Notably, the centrifugal instability described here does *not* rely on baroclinic structure. Rather, it arises from the balance/imbalance between centrifugal and pressure forces along a geopotential.¹

57.2.1 Equations of motion

We studied the angular momentum of this rotating shallow water system in Section 43.5, where we made use of a rotating reference frame and polar coordinates (r, θ) measured in the rotating frame. Here, r is the radial position from the rotational axis and θ is the azimuthal angle measured counter-clockwise from the rotating x -axis. We furthermore derived the acceleration in cylindrical-polar coordinates in Section 42.1. Making use of those earlier results allows us to write down the horizontal components to the equations of motion

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta}{\partial r} + r^{-3} (l^z)^2 \quad (57.1a)$$

$$\frac{Dl^z}{Dt} = -g \frac{\partial \eta}{\partial \theta}. \quad (57.1b)$$

In these equations, η is the free surface height for the shallow water layer (see Figure 42.1),

$$l^z = \hat{z} \cdot [\mathbf{r} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] = r(v^\theta + r\Omega) \quad (57.2)$$

is the angular momentum per mass computed about the rotation axis (the z -axis), and we make use of the radial and azimuthal velocity components

$$v^r = \frac{Dr}{Dt} \quad r^{-1} v^\theta = \frac{D\theta}{Dt}. \quad (57.3)$$

¹One may conceive of centrifugal instability in a baroclinic flow where parcel displacements maintain their angular momentum and buoyancy. That analysis would lead to isentropic centrifugal instability, which is directly analogous to the isentropic inertial instability discussed in Section 57.5.

The material evolution of the radial velocity (equation (57.1a)) is affected by the radial pressure gradient plus centrifugal acceleration, whereas the material evolution of angular momentum is affected only by angular gradients in the pressure field. We consider equilibrium states where the radial acceleration vanishes. Such states are said to be in *cyclostrophic balance*, whereby the radial pressure gradient balances the centrifugal acceleration

$$\frac{Dv^r}{Dt} = 0 \implies g \frac{\partial \eta}{\partial r} = \frac{(v^\theta + r \Omega)^2}{r} \quad \text{cyclostrophic balance.} \quad (57.4)$$

Are there unstable profiles of the rotating cylindrical flow in cyclostrophic balance? To answer this question in an analytically tractable manner, we examine stability of a flow state with rotational symmetry. Rotational symmetry means that the angular momentum remains materially constant

$$\frac{\partial \eta}{\partial \theta} = 0 \implies \frac{Dl^z}{Dt} = 0. \quad (57.5)$$

This constraint plays a fundamental role in determining stability of the flow. Angular symmetry also means that the angular velocity is a function only of the radial direction, $v^\theta = v^\theta(r)$.

57.2.2 Solid-body motion

Consider the rotating fluid undergoing solid-body rotation. In this case, the angular momentum per mass is given by

$$l_{\text{solid-body}}^z = r^2 \Omega \quad (57.6)$$

so that the magnitude of the angular momentum increases as the square of the radial distance. To account for the possibility of positive or negative rotational motion, we find it convenient to make use of the square of the angular momentum, which also increases radially for the solid-body motion

$$\frac{d[l^z(r)]^2}{dr} = 4r^3 \Omega^2 > 0. \quad (57.7)$$

As shown in this section, flow is centrifugally unstable if the square of its angular momentum is a decreasing function of its radial distance. Such configurations will spontaneously adjust through *centrifugal instability* towards a state where its squared angular momentum increases radially. The instability is termed “centrifugal” since it is the centrifugal acceleration that “throws outward” the fluid if its squared angular momentum decreases radially, thus bringing the fluid back into a stable state.

57.2.3 Stability analysis based on energetic arguments

As first introduced by Lord Rayleigh, we consider a thought experiment in which two adjacent equal mass circular fluid rings are swapped, one originating from radial position $r = r_1$ and the other at $r = r_2 = r_1 + \Delta r$. Furthermore, assume that the radial velocity vanishes so that the kinetic energy of the rings is due only to their rotational motion. If swapping the rings decreases the net kinetic energy in the base state, then it can be used to fuel an instability.² In this case we say that the flow is *centrifugally unstable*, with this name used since it is the centrifugal acceleration from the circular parcel trajectory that promotes the instability. In general, any curved flow will be exposed to centrifugal instability if swapping parcels reduces the kinetic energy of the base state while maintaining constant angular momentum.

²Gravitational potential energy plays no role here, as we are swapping fluid rings at the same vertical position.

Condition for centrifugal instability

With no radial flow, kinetic energy only arises from angular motion so that a ring of mass $\delta m = \rho \delta V$ and radius r_1 has kinetic energy

$$E(r_1) = (\delta m/2) [v^\theta(r_1) + r_1 \Omega]^2 = (\delta m/2) [l^z(r_1)/r_1]^2. \quad (57.8)$$

The initial kinetic energy for the two rings is thus given by the sum

$$E_{\text{init}} = (\delta m/2) ([l^z(r_1)/r_1]^2 + [l^z(r_2)/r_2]^2). \quad (57.9)$$

The equilibrium state and the perturbation each maintain rotational symmetry. Hence, when swapping their radial positions, the rings each maintain their respective angular momentum. But by changing radial positions their kinetic energy changes, thus leading to the kinetic energy of the swapped state

$$E_{\text{swap}} = (\delta m/2) ([l^z(r_1)/r_2]^2 + [l^z(r_2)/r_1]^2). \quad (57.10)$$

The difference in energy is given by

$$E_{\text{swap}} - E_{\text{init}} = (\delta m/2) ([l^z(r_2)]^2 - [l^z(r_1)]^2)(r_1^{-2} - r_2^{-2}). \quad (57.11)$$

Since $r_2 = r_1 + \Delta r > r_1$, we have a release of kinetic energy ($E_{\text{swap}} - E_{\text{init}} < 0$) if the squared angular momentum decreases upon moving outward

$$\frac{d[l^z(r)]^2}{dr} < 0 \implies \frac{d}{dr} \left[r^4 (\dot{\theta} + \Omega)^2 \right] = \frac{d}{dr} \left[r^2 (v^\theta + r \Omega)^2 \right] < 0. \quad (57.12)$$

Rearrangement leads to the instability condition

$$(\dot{\theta} + \Omega)(\zeta + f) < 0, \quad (57.13)$$

where

$$\zeta = \frac{1}{r} \frac{d(r v^\theta)}{dr} \quad (57.14)$$

is the vertical component to the relative vorticity for axisymmetric flow. Suppose that $\Omega > 0$ (counter-clockwise rotation) and $\dot{\theta} + \Omega > 0$ (counter-clockwise rotation dominated by solid-body). The instability condition thus reduces to $\zeta < -f$, so that flow is unstable if the relative vorticity has an opposite sign to the cylinder's rotation and it has a large enough magnitude to render the angular momentum a decreasing function of radial distance. The opposite holds for $\Omega < 0$ (clockwise rotation) and $\dot{\theta} + \Omega < 0$, in which case instability is realized when $\zeta > -f$. In either case, the instability condition is increasingly difficult to realize when the rotation rate increases.

Notably, if the flow satisfies the condition

$$(\dot{\theta} + \Omega)(\zeta + f) > 0, \quad (57.15)$$

then the flow is stable to axisymmetric perturbations. However, we do not know if the flow is stable to more general perturbations since our analysis only considers axisymmetric perturbations. More analysis is required to determine stability criteria for general perturbations.

57.2.4 Stability analysis based on parcel arguments

We supplement the previous arguments by considering the force balance in the radial momentum equation. We show that an unstable angular velocity profile leads to radial force acting on a parcel that supports the parcel's movement away from its initial position. We again assume rotational symmetry so that the angular momentum is a material invariant. We also consider the equilibrium state with zero radial acceleration so that the radial momentum equation (57.1a) leads to

$$\frac{Dv^r}{Dt} = 0 = -g \frac{\partial \bar{\eta}}{\partial r} + r^{-3} (\bar{l}^z)^2 \quad (57.16)$$

Subtracting this equilibrium state from the full momentum equation (57.1a) leads to an equation for radial acceleration of perturbations about the equilibrium state

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta'}{\partial r} + r^{-3} [(l^z)^2 - (\bar{l}^z)^2], \quad (57.17)$$

where $\eta' = \eta - \bar{\eta}$ is the perturbation surface height.

Probing stability by perturbing the radius of a circular fluid ring

Consider a perturbation realized by moving a circular fluid ring outward from its initial equilibrium state at radius r to a radius $r + \Delta r$. During the expansion of the ring, the angular momentum remains constant due to the rotational symmetry, so that

$$(l^z)^2(r + \Delta r) = (l^z)^2(r) = (\bar{l}^z)^2(r), \quad (57.18)$$

where the second equality holds since we are starting the ring from its equilibrium state. To determine the acceleration at the new radius $r + \Delta r$ we have

$$(l^z)^2(r + \Delta r) - (\bar{l}^z)^2(r + \Delta r) = (\bar{l}^z)^2(r) - (\bar{l}^z)^2(r + \Delta r) \approx -\Delta r \frac{d(\bar{l}^z)^2(r)}{dr}, \quad (57.19)$$

so that the radial momentum equation at $r + \Delta r$ takes on the form

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta'}{\partial r} - \frac{\Delta r}{(r + \Delta r)^3} \frac{d(\bar{l}^z)^2(r)}{dr}. \quad (57.20)$$

We thus see that if the squared angular momentum decreases upon moving the ring to a larger radius, then the second term in equation (57.20) provides a positive radial acceleration, thus supporting the initial outward perturbation. Ignoring the potential for the perturbation pressure gradient, $-g\partial\eta'/\partial r$, to counter-act the acceleration, we are left with the same instability condition (57.12) derived using energetic arguments.

Centrifugal oscillations

Ignoring the perturbation pressure gradient, and introducing a centrifugal frequency

$$\sigma^2(r) \equiv \frac{1}{(r + \Delta r)^3} \frac{d(\bar{l}^z)^2(r)}{dr} \quad (57.21)$$

leads to the oscillator equation for the deviation from the equilibrium radial position

$$\frac{D^2 \Delta r}{Dt^2} + \sigma^2 \Delta r = 0, \quad (57.22)$$

where $u^r = D\Delta r/Dt$. For stable cases with $\sigma^2 > 0$, the parcel oscillates around the equilibrium position with period

$$T_{\text{centrifugal}} = \frac{2\pi}{\sigma}. \quad (57.23)$$

In contrast, for the unstable case with $\sigma^2 < 0$, then Δr grows exponentially.

57.2.5 Comments

Chapter 3 of [Drazin and Reid \(1981\)](#) is the canonical reference for centrifugal instability, where they provide a thorough stability analysis including both axisymmetric and non-axisymmetric perturbations. In our treatment, we also made use of the parcel arguments from Section 3.2 of [Markowski and Richardson \(2010\)](#). Furthermore, [Markowski and Richardson \(2010\)](#) comment on the perturbation pressure gradient in equation (57.20). They note that parcel stability arguments generally ignore changes to the pressure gradient. Stated otherwise, a parcel analysis concerns the equilibrium angular momentum profile and its contribution to movement away from equilibrium. It is not concerned with “back reaction” from pressure perturbations associated with movement of parcels or fluid rings. Such considerations generally require analysis beyond the parcel framework, and thus offer a limit concerning the validity of purely parcel based arguments.

57.3 Potential momentum

A front is a region of enhanced lateral gradients in the buoyancy field (baroclinic front) or sea level (shallow water front). These fronts generally have corresponding currents (jets) arising from geostrophic balance (when off-equator). Figure 57.1 illustrates a baroclinic front that is symmetric in the zonal direction so that the buoyancy field is only a function of latitude, depth, and time, $b = b(y, z, t)$. We likewise assume that all other fields possess zonal symmetry, including pressure and velocity. Fronts can generally be oriented in any direction. Furthermore, on an f -plane there is rotational symmetry in the horizontal plane so that we can orient the coordinate system as desired.

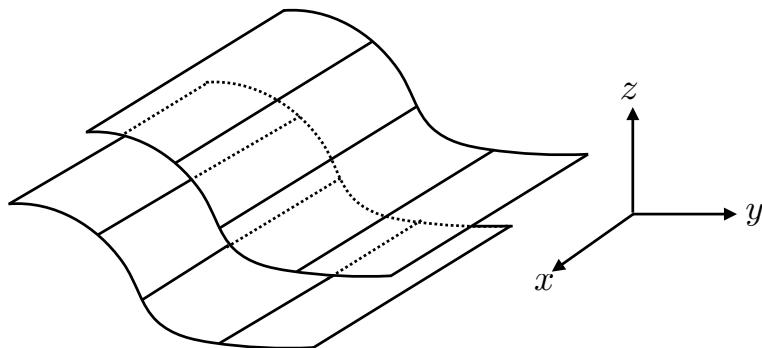


Figure 57.1: Example of a zonally symmetric baroclinic frontal region, showing two iso-buoyancy surfaces with $b = b(y, z, t)$. With $\partial b / \partial y > 0$ as drawn, the corresponding zonal thermal wind shear, $f \partial u / \partial z = -\partial b / \partial y < 0$ is westward; i.e., stronger westward flow with increasing height.

57.3.1 Linear momentum and potential momentum

The horizontal linear momentum per mass in a Boussinesq perfect fluid on an f -plane evolves according to

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z p. \quad (57.24)$$

Since f is a constant, this equation can be written

$$\frac{DM}{Dt} = -\nabla_z p, \quad (57.25)$$

where we introduced the *potential momentum* per mass

$$\mathbf{M} \equiv \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{x} = (u - f y) \hat{\mathbf{x}} + (v + f x) \hat{\mathbf{y}}, \quad (57.26)$$

and noted that $\mathbf{v} = D\mathbf{x}/Dt$. Notably, the potential momentum remains materially invariant in directions where the horizontal pressure gradient vanishes. We described the potential momentum for a point particle in Section 15.2. Elements of the point particle dynamics remain for the fluid though with added features from pressure gradients.

Materially constant zonal potential momentum

For the zonal buoyancy front illustrated in Figure 57.1, we assume that all fields are independent of the zonal direction so that

$$\frac{\partial p}{\partial x} = 0. \quad (57.27)$$

A vanishing zonal pressure gradient means that the *zonal potential momentum* per mass is a Lagrangian invariant

$$\frac{DM^x}{Dt} = 0 \quad \text{where } M^x \equiv u - f y. \quad (57.28)$$

This Lagrangian invariance greatly constrains the flow. For example, consider a fluid parcel at an initial latitude y_0 with zonal velocity u_0 . Movement of the parcel to a new latitude leads to the equality

$$u_0 - f y_0 = y_1 - f y_1, \quad (57.29)$$

so that the zonal velocity at the new latitude is given by

$$u_1 = u_0 - f(y_0 - y_1). \quad (57.30)$$

Motivating the name

The term “potential” is motivated by the same reasoning used for potential temperature. Namely, the zonal potential momentum identifies that amount of zonal momentum that a parcel would acquire if moved from an arbitrary latitude y_1 to a reference latitude y_0 . More specifically, inverting equation (57.30) we see that

$$u_0 = M^x(y_1) + f y_0. \quad (57.31)$$

Hence, the quantity $M^x(y_1)$ is the extra zonal momentum per mass available at the reference latitude, y_0 , upon moving a parcel from y_1 to y_0 . We thus see that the potential momentum is a material invariant in the way that potential temperature is for a perfect fluid. Furthermore, as seen in Section 57.4, meridional gradients of M^x measure the *inertial* stability of a flow configuration in a manner directly analogous to how vertical gradients of potential temperature (or buoyancy) measure gravitational stability.

Meridional potential momentum

There are occasions when a front exhibits meridional symmetry, in which case the Boussinesq perfect fluid equations take the form

$$\frac{Du}{Dt} = fv - \frac{1}{\rho_0} \frac{\partial p}{\partial x} \quad (57.32a)$$

$$\frac{Dv}{Dt} = -fu. \quad (57.32b)$$

In this case the meridional potential momentum is materially invariant

$$\frac{DM^y}{Dt} = 0 \quad \text{where } M^y \equiv v + fx. \quad (57.33)$$

57.3.2 Zonal potential momentum on a β -plane

The f -plane is rotationally invariant about the rotational axis. Correspondingly, we can write the momentum equation in the form (57.25), thus exposing the potential momentum. Material invariance for potential momentum holds along the symmetry direction of an arbitrarily oriented symmetric front.

The β -plane is not rotationally invariant. Rather, it only maintains symmetry along lines of constant latitude (zonal directions). Consequently, only zonally oriented symmetric fronts maintain material invariance of zonal potential momentum. To see this property, write the zonal momentum equation in the form

$$\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - f v = 0, \quad (57.34)$$

where we assumed zonal symmetry ($\partial_x = 0$). Now write the Coriolis parameter in the form

$$\Gamma = f_0 y + \beta y^2 / 2 \implies f = \frac{d\Gamma}{dy}, \quad (57.35)$$

so that the zonal momentum equation takes the form

$$\frac{\partial(u - \Gamma)}{\partial t} + v \frac{\partial(u - \Gamma)}{\partial y} + w \frac{\partial(u - \Gamma)}{\partial z} = 0. \quad (57.36)$$

We thus see that that $M = u - \Gamma$ is materially invariant for this zonally symmetric front.

57.3.3 Further reading

See Section 15.3 for more discussion of potential momentum as it concerns a point particle. We are motivated to use the name *potential momentum* following arguments given on page 51 of [Markowski and Richardson \(2010\)](#).

57.4 Horizontal inertial instability of a geostrophically balanced front

We now examine stability of a geostrophically balanced front in an inviscid Boussinesq fluid. The analysis of centrifugal instability in Section 57.2 is closely emulated, with rotational symmetry replaced by along-front symmetry and angular momentum replaced by potential momentum. We consider both the Rayleigh energetic stability analysis and the parcel analysis. Furthermore, the

perturbations maintain symmetry in the along-front direction, so our perturbations consist of a displaced row of parcels oriented along the front. Stability to more general perturbations, such as those that are not symmetric along the front, is not addressed here.

The results of our analysis are rotationally invariant since the f -plane is rotationally invariant. Hence, we choose to orient the coordinate system based on convenience whereby the x -axis is the along front axis and the y -axis is across the front. Furthermore, their relative orientation is chosen so that $\hat{x} \wedge \hat{y} = \hat{z}$, where \hat{z} is anti-parallel to gravity.

Although we make use of a continuously stratified Boussinesq fluid, the instability is not associated with baroclinicity. Rather, as for the centrifugal case in Section 57.2, it is associated with stability of an equilibrium state to horizontal displacements along geopotential surfaces. In a continuously stratified adiabatic fluid, such horizontal displacements generally cross isentropic surfaces and so comprise irreversible perturbations. So long as the associated mixing of momentum is negligible, we can still make use of material invariance of potential momentum. The isentropic inertial instability analysis in Section 57.5 maintains the adiabatic nature of displacements when probing for instabilities.

57.4.1 Equations of motion and equilibrium state

The horizontal momentum equation for an inviscid Boussinesq fluid on an f -plane is given by

$$\frac{Du}{Dt} + f \hat{z} \wedge \mathbf{u} = -\frac{1}{\rho_0} \nabla_z p. \quad (57.37)$$

In the presence of along-front symmetry, an exact solution to the horizontal momentum equation is given by along-front geostrophic flow and zero cross-front flow

$$u_g = -\frac{1}{f \rho_0} \frac{\partial p}{\partial y} \quad (57.38a)$$

$$v = 0. \quad (57.38b)$$

We examine the stability of this exact equilibrium base state to perturbations aligned with the front. Subtracting the base state solution from the full momentum equation (57.37) leads to

$$\frac{Du}{Dt} = fv \quad (57.39a)$$

$$\frac{Dv}{Dt} = f(u_g - u). \quad (57.39b)$$

We continue to assume along-front symmetry thus eliminating the along-front pressure gradient. Cross-front accelerations are determined by deviations from geostrophy of the along-front velocity. Likewise, along-front accelerations are determined by the Coriolis acceleration arising from a non-zero cross-front velocity. Following the treatment of potential momentum in Section 15.2, we write the along-front momentum equation as the material time derivative of the along-front potential momentum per mass³

$$m = u - fy, \quad (57.40)$$

bringing the suite of perturbation equations to

$$\frac{Dm}{Dt} = 0 \quad (57.41a)$$

$$\frac{Dv}{Dt} = f(u_g - u). \quad (57.41b)$$

³To reduce notational clutter, we write m for the along-front potential momentum rather than M^x . Additionally, we use m rather than M since we reserve M for the Montgomery potential used in Section 57.5.

Material invariance of the along-front potential momentum plays a fundamental role in the stability analysis.

57.4.2 Stability analysis based on energetic arguments

We follow the energetic arguments given in Section 57.2.3 for centrifugal instability of cyclostrophic flow. Here, we focus on the exact solution as given by along-front geostrophic flow and ask whether a swap of two along-front oriented rows releases kinetic energy from the base state. If so, then the base state flow is *inertially unstable* to an along-front symmetric perturbation. In that case, perturbations spontaneously initiate inertial instability to affect a return to an inertially stable state.

Instability condition

The kinetic energy per mass for the along-front geostrophic flow is given by

$$E = (1/2) u_g^2 = (1/2) (m_g + fy)^2, \quad (57.42)$$

where we replaced the geostrophic velocity with the geostrophic potential momentum through equation (57.40). The kinetic energy per mass contained in two equal mass parcels at distinct cross-front positions $y = y_1$ and $y = y_2 = y_1 + \Delta y$ is given by

$$E_{\text{init}} = (1/2) (m_g(y_1) + fy_1)^2 + (1/2) (m_g(y_2) + fy_2)^2. \quad (57.43)$$

Swapping the parcels while leaving the potential momentum unchanged leads to the kinetic energy in the swapped state

$$E_{\text{swap}} = (1/2) (m_g(y_1) + fy_2)^2 + (1/2) (m_g(y_2) + fy_1)^2. \quad (57.44)$$

A bit of algebra leads to the difference in kinetic energy for the two states

$$E_{\text{swap}} - E_{\text{init}} = -f \Delta y \Delta m_g, \quad (57.45)$$

where $\Delta m_g = m_g(y_2) - m_g(y_1)$. A Taylor series computed about the cross-front position $y = y_1$ leads to

$$E_{\text{swap}} - E_{\text{init}} = -f (\Delta y)^2 \frac{dm_g}{dy} \quad (57.46a)$$

$$= -f (\Delta y)^2 \left(\frac{\partial u_g}{\partial y} - f \right) \quad (57.46b)$$

$$= f (\Delta y)^2 (\zeta_g + f), \quad (57.46c)$$

where

$$\zeta_g = -\frac{\partial u_g}{\partial y} \quad (57.47)$$

is the vertical component to the relative geostrophic vorticity for the symmetric base state. Energy is released upon swapping the two rows if the following condition is satisfied

$$-f \frac{\partial m_g}{\partial y} = f (\zeta_g + f) < 0 \quad \text{inertial instability.} \quad (57.48)$$

The first instability condition arises if the cross-front gradient of the geostrophic potential momentum has the same sign as the Coriolis parameter. In this case, the geostrophic potential momentum is increasing along with the planetary rotation. We can further our understanding by considering a case of zero geostrophic velocity, so that the potential momentum is $m = -fy$ and

$$-f(\partial m/\partial y) = f^2 > 0 \quad \text{zero along-front flow.} \quad (57.49)$$

Hence, an inertially unstable base state has an along-front flow that overcomes the contribution to potential momentum from planetary rotation, whereby solid-body motion imposes $\partial m/\partial y = -f$. The second instability condition in equation (57.48) says that the base state is unstable if the absolute geostrophic vorticity, $\zeta_g + f$, has an opposite sign to the planetary rotation. This condition is directly analogous to that derived for centrifugal instability in Section 57.2.3.

Details of the instability condition

The instability condition (57.48) takes the following form for the northern and southern hemispheres. Again, the x -axis is oriented along the front and y -axis is across the front with $\hat{x} \wedge \hat{y} = \hat{z}$.

$$\text{northern hemisphere } (f > 0): \quad \frac{\partial m_g}{\partial y} > 0 \quad \zeta_g < -|f| \quad \frac{\partial u_g}{\partial y} > +|f| \quad (57.50a)$$

$$\text{southern hemisphere } (f < 0): \quad \frac{\partial m_g}{\partial y} < 0 \quad \zeta_g > +|f| \quad \frac{\partial u_g}{\partial y} < -|f|. \quad (57.50b)$$

In both hemispheres, instability arises when the relative geostrophic vorticity is anti-cyclonic and larger in magnitude than the planetary vorticity. Under such conditions, inertial instability allows the flow to readjust toward a state of less extreme relative vorticity, thus returning the flow to a state with absolute vorticity dominated by planetary vorticity. Equivalently, inertial instability arises for flows where $\partial m_g/\partial y > 0$ in the northern hemisphere and $\partial m_g/\partial y < 0$ in the southern hemisphere.

57.4.3 Stability analysis based on parcel arguments

As with any stability analysis based on parcel movement, it is critical to be precise with the thought experiment. We start with an equilibrium base state of along-front geostrophic balance with zero motion in the cross-front direction. We examine the stability of this base state with respect to along-front perturbations of fluid parcels. For this purpose, imagine moving a row of fluid parcels from cross-front position y to position $y + \Delta y$. In general, the displaced row of parcels will not be in geostrophic balance at the new position, thus providing for a non-zero acceleration in the cross-front direction. What is the sign of that acceleration? If the acceleration is directed back to the original position, then the base state is stable and displaced parcels exhibit inertial oscillations. In contrast, the base state is inertially unstable if the acceleration is directed towards further displacement. As we will see, this analysis leads to the same instability condition (57.48) found through energetic arguments. However, the parcel analysis, which considers the conditions for acceleration, offers complementary insights that serve to deepen our understanding.

Mathematical formulation

The cross-front acceleration in the new cross-front position is

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)], \quad (57.51)$$

where $u(y + \Delta y)$ is the along-front velocity of the displaced parcel at the new position. Likewise, $u_g(y + \Delta y)$ is the geostrophic velocity at $y + \Delta y$. The geostrophic velocity $u_g(y + \Delta y)$ determines a Coriolis acceleration at the position $y + \Delta y$ that is balanced by the cross-front pressure gradient at $y + \Delta y$. To determine the sign of the acceleration acting on the displaced parcel, we make use of the material invariance of along-front potential momentum. This invariance means that each parcel carries its potential momentum from the original position

$$m(y) = m_g(y) = u_g(y) - fy \quad (57.52)$$

to the new position. In turn, invariance of along-front potential momentum allows us to determine the along-front velocity of the parcel at the new position in terms of $u_g(y)$

$$m(y + \Delta y) = u(y + \Delta y) - f(y + \Delta y) \quad (57.53a)$$

$$= m(y) \quad (57.53b)$$

$$= u_g(y) - fy, \quad (57.53c)$$

which leads to

$$u(y + \Delta y) = u_g(y) + f\Delta y. \quad (57.54)$$

The cross-front acceleration (57.51) thus takes the form

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)] \quad (57.55a)$$

$$= f[u_g(y + \Delta y) - u_g(y) - f\Delta y] \quad (57.55b)$$

$$\approx f\Delta y \left(\frac{\partial u_g}{\partial y} - f \right) \quad (57.55c)$$

$$= -\Delta y f (\zeta_g + f). \quad (57.55d)$$

As with the energetic arguments in Section 57.4.2, we are left with the instability condition

$$f (\zeta_g + f) < 0 \implies \text{inertially unstable}. \quad (57.56)$$

Summary of the parcel argument

At the initial location in the base state, a parcel under geostrophic balance has its Coriolis acceleration balanced by its pressure acceleration. However, the displaced parcel generally will not be in geostrophic balance at the new location, in which case its Coriolis acceleration does not balance the local pressure gradient. Does the imbalance lead to an acceleration back towards its initial position (oscillation) or further away (exponential growth)? This question is quite general. Along-front symmetry of the base state and the perturbation ensures material invariance of along-front potential momentum. This invariance provides an explicit expression for the acceleration felt by the displaced parcel, thus determining a condition on stability of the base state to the symmetric perturbations. The method of analysis is directly analogous to that applied to the rotating tank of fluid in Section 57.2 for centrifugal instability, as well as for a vertical column of fluid in Section 25.3 for gravitational stability.

57.4.4 Comments

As stated in Section 57.1.3, the rather close similarities between centrifugal and inertial instability has prompted the oceanographic literature to often use the term centrifugal instability when

referring to the inertial instability considered here (e.g., [McWilliams, 2016](#)). We instead prefer the terminology of the atmospheric literature, such as detailed in the texts by [Holton \(1992\)](#) and [Markowski and Richardson \(2010\)](#), which also follows the classical fluid mechanics terminology used in the text by [Drazin and Reid \(1981\)](#). In this manner, we reserve the term *centrifugal instability* for an axisymmetric base state in cyclostrophic balance, and *inertial instability* for a two-dimensional base state in geostrophic balance.

57.5 Symmetric (isentropic inertial) instability of a baroclinic front

Here, we consider stability of a geostrophically balanced baroclinic front on an f -plane. As before, the front exhibits along-front symmetry so that the along-front potential momentum is a material invariant. Additionally, we assume the fluid to be adiabatic, so that buoyancy is also materially invariant. We investigate the stability of a geostrophically balanced along-front flow to symmetric displacements of parcels along a constant buoyancy surface. By construction, this displacement is neutral to gravitational instability since it occurs along a constant buoyancy surface. However, a displaced parcel could still find itself in an unstable position depending on the potential momentum of the environment.

The analysis proceeds analogously to the horizontally inertial instability case of Section 57.4.3, with the displacements here isentropic (constant buoyancy) rather than horizontal. We are thus motivated to refer to the ensuing instability as *isentropic inertial instability*. However, this terminology is not common, with *symmetric instability* far more common.

57.5.1 Equations of motion and equilibrium state

Given the adiabatic nature of the fluid, and the role of baroclinicity, we make use of the buoyancy coordinates for a Boussinesq fluid as detailed in Section 45.1. The horizontal momentum equation is thus given by

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b M \quad (57.57)$$

where ∇_b is a horizontal gradient computed along constant buoyancy surfaces. Furthermore,

$$M = p/\rho_0 - b z \quad (57.58)$$

is the Montgomery potential that contributes the acceleration

$$-\nabla_b M = -\nabla_b(p/\rho_0) + b \nabla_b z. \quad (57.59)$$

The first term arises from pressure gradients along constant buoyancy surfaces, and the second from geopotential gradients. In the presence of along-front symmetry, an exact solution to the horizontal momentum equation is given by along-front geostrophic flow and zero cross-front flow

$$u_g = -\frac{1}{f \rho_0} \left[\frac{\partial M}{\partial y} \right]_b \quad (57.60a)$$

$$v = 0. \quad (57.60b)$$

We examine the stability of this base state to perturbations symmetric in the along-front direction. Subtracting the exact equilibrium solution from the full momentum equation (57.37) leads to

$$\frac{Du}{Dt} = fv \quad (57.61a)$$

$$\frac{Dv}{Dt} = f(u_g - u), \quad (57.61b)$$

where we continue to assume along-front symmetry thus allowing us to drop the along-front gradient of the Montgomery potential. Following the treatment in Sections 15.2 and 57.4.1, we write the along-front momentum equation as the material time derivative of the along-front potential momentum per mass $m = u - fy$ (equation (57.40)), thus bringing the perturbation equations to

$$\frac{Dm}{Dt} = 0 \quad (57.62a)$$

$$\frac{Dv}{Dt} = f(u_g - u). \quad (57.62b)$$

57.5.2 Stability analysis based on parcel arguments

We follow the parcel analysis for inertial instability in Section 57.4.3, starting with an equilibrium base state of along-front geostrophic balance with zero meridional motion. We examine the stability of this state with respect to symmetric perturbations of fluid parcels along a constant buoyancy surface. For this purpose, imagine moving a row of fluid parcels from cross-front position y to position $y + \Delta y$ while maintaining a fixed buoyancy. In general, the displaced row of parcels will not be in geostrophic balance at the new position, thus providing for a non-zero cross-front acceleration at that displaced position. What is the sign of that acceleration?

Mathematical formulation

At the new cross-front position the cross-front acceleration is given by

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)], \quad (57.63)$$

where $u(y + \Delta y)$ is the along-front velocity of the displaced parcel at the new position, and $u_g(y + \Delta y)$ is the geostrophic velocity at that position. To determine the sign of the acceleration acting on the displaced parcel, we make use of the invariance of along-front potential momentum, whereby each parcel carries its potential momentum from the original position

$$m(y) = m_g(y) = u_g(y) - fy \quad (57.64)$$

to the new position. In turn, invariance of along-front potential momentum allows us to determine the parcel's along-front velocity at the new position in terms of $u_g(y)$

$$m(y + \Delta y) = u(y + \Delta y) - f(y + \Delta y) \quad (57.65a)$$

$$= m(y) \quad (57.65b)$$

$$= u_g(y) - fy, \quad (57.65c)$$

which leads to

$$u(y + \Delta y) = u_g(y) + f\Delta y. \quad (57.66)$$

The cross-front acceleration (57.63) thus takes the form

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)] \quad (57.67a)$$

$$= f[u_g(y + \Delta y) - u_g(y) - f\Delta y] \quad (57.67b)$$

$$\approx f\Delta y \left[\left(\frac{\partial u_g}{\partial y} \right)_b - f \right] \quad (57.67c)$$

$$= -\Delta y f (\zeta_g^b + f). \quad (57.67d)$$

In the final equality we introduced the relative geostrophic vorticity as written in buoyancy coordinates (see Section 45.1)

$$\zeta_g^b = \left[\frac{\partial v_g}{\partial x} \right]_b - \left[\frac{\partial u_g}{\partial y} \right]_b = - \left[\frac{\partial u_g}{\partial y} \right]_b. \quad (57.68)$$

Furthermore, we can introduce the Boussinesq Ertel potential vorticity in the form

$$Q = (\zeta^b + f) N^2, \quad (57.69)$$

with

$$N^2 = \frac{\partial b}{\partial z} \quad (57.70)$$

the squared buoyancy frequency. We assume $N^2 > 0$ since we are interested in gravitationally stable flow. Bringing these results together leads to the equivalent expressions for the symmetric instability condition

$$f(\zeta_g^b + f) < 0 \implies \text{symmetrically unstable}. \quad (57.71a)$$

$$f Q_g < 0 \implies \text{symmetrically unstable}. \quad (57.71b)$$

The instability condition (57.71a) is a direct translation of the inertial instability condition (57.56), translating from horizontal displacements to isopycnal displacements.

Relative slope of buoyancy and potential momentum surfaces: Part I

The instability condition (57.71a) can be written as a geometric statement about the relative slopes of buoyancy and potential momentum surfaces. For this purpose we make use of the identity for an along-front symmetric geostrophic flow

$$\left[\frac{\partial m}{\partial y} \right]_b = \left[\frac{\partial u}{\partial y} \right]_b - f = -(\zeta^b + f). \quad (57.72)$$

Consequently, the instability condition (57.71a) is equivalent to

$$f \left[\frac{\partial m_g}{\partial y} \right]_b > 0 \implies \text{symmetrically unstable}. \quad (57.73)$$

This condition is directly analogous to the inertial instability condition (57.48), only here with displacement along an isopycnal rather than a geopotential. In the northern hemisphere, if one moves in the $+\hat{y}$ direction on a constant buoyancy surface and encounters increasing values for the potential momentum, then the flow is symmetrically unstable. Conversely in the southern hemisphere, if one moves in $-\hat{y}$ direction on a constant buoyancy surface and encounters increasing values for the potential momentum, then the flow is symmetrically unstable. We illustrate this situation in Figure 57.2.

Relative slope of buoyancy and potential momentum surfaces: Part II

As another way to write the instability condition (57.73), we make use of the expression (11.68b) to transform the derivative on buoyancy surfaces back to geopotential coordinates, and then introduce

the cross-front derivative of the buoyancy on a potential momentum surface

$$\left[\frac{\partial m_g}{\partial y} \right]_b = \left[\frac{\partial m_g}{\partial y} \right]_z - \frac{\partial b / \partial y}{\partial b / \partial z} \frac{\partial m_g}{\partial z} \quad (57.74a)$$

$$= N^{-2} \hat{x} \cdot (\nabla m_g \wedge \nabla b) \quad (57.74b)$$

$$= -N^{-2} \hat{x} \cdot (\nabla b \wedge \nabla m_g) \quad (57.74c)$$

$$= - \left[\frac{\partial b}{\partial y} \right]_{m_g}. \quad (57.74d)$$

We are thus led to the additional expression for a symmetrically unstable base state

$$f \left[\frac{\partial b}{\partial y} \right]_{m_g} < 0 \implies \text{symmetrically unstable.} \quad (57.75)$$

Hence, a northern hemisphere symmetrically unstable configuration sees a reduction in buoyancy when moving in the $+\hat{y}$ direction along a constant potential momentum surface. This situation is illustrated in Figure 57.2.

Relative slope of buoyancy and potential momentum surfaces: Part III

We offer one final identity for writing the instability condition, pulling us back to the PV version. Here, use the geopotential coordinates to write the following relation between the absolute vorticity and gradient of the potential momentum for the front

$$\omega + f\hat{z} = \hat{x} \wedge \nabla m, \quad (57.76)$$

so that

$$Q = (\omega + f\hat{z}) \cdot \nabla b \quad (57.77a)$$

$$= (\nabla b \wedge \nabla m) \cdot \hat{x} \quad (57.77b)$$

$$= N^2 (\zeta^b + f). \quad (57.77c)$$

The condition for symmetric instability thus can be written

$$f Q_g = -N^2 f \left[\frac{\partial m_g}{\partial y} \right]_b = N^2 f \left[\frac{\partial b}{\partial y} \right]_{m_g} < 0. \quad (57.78)$$

57.5.3 Comments

[Hoskins \(1974\)](#) made the connection between the condition for symmetric instability and potential vorticity. His two-page paper is a classic in brevity and profundity.

57.6 Secondary circulation along baroclinic fronts

We here consider an overturning circulation that arises in symmetric fronts, with the overturning occurring in the cross-front and vertical plane. This *secondary* circulation is ageostrophic. We make use of elements from the symmetric stability analysis in Section 57.5. However, the focus here is not on the stability per se, but instead on deriving an expression for the overturning circulation.

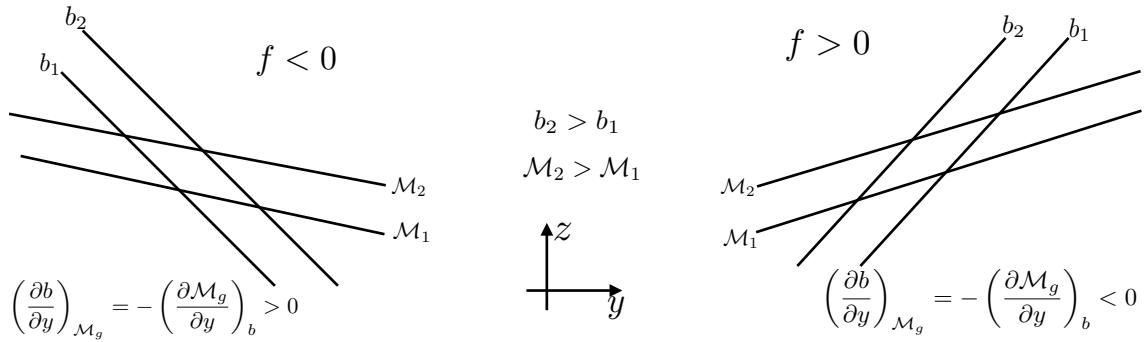


Figure 57.2: Example flow configurations that are symmetrically unstable; i.e., inertially unstable to a symmetric perturbation along a constant buoyancy surface. We show example buoyancy surfaces and potential momentum surfaces for the southern hemisphere (left) and northern hemisphere (right). The instability conditions (57.73) and (57.75) are indicated on the respective panels. In both cases, surfaces of constant buoyancy are more steeply sloped than constant potential momentum surfaces. The x coordinate measures distance in the along-front direction and y measures distance in the cross-front direction, oriented so that $\hat{x} \wedge \hat{y} = \hat{z}$ where \hat{z} is anti-parallel to gravity (\hat{x} is out of the page). A means to quickly judge whether a flow is symmetrically unstable is to note that the wedge region between buoyancy and potential momentum surfaces provides a source of available potential energy. Symmetric instability can feed off the potential energy only when buoyancy surfaces are more steeply sloped than potential momentum surfaces.

Hoskins (1975) provided a major advance in our ability to study secondary overturning circulations along fronts. He did so by introducing the semi-geostrophic (SG) equations. The SG system is balanced (i.e., gravity waves are filtered), just like in quasi-geostrophy. But the SG system is more general than QG, allowing for the study of Rossby numbers of order unity such as occur in regions of ocean submesoscale fronts and atmospheric synoptic fronts. Large magnitude vertical velocities are signatures of order unity Rossby number flow. We here develop elements of the semi-geostrophic frontal equations to study ageostrophic motions (including vertical motion) occurring along fronts in the ocean and atmosphere. We use the language of oceanography as much here follows the ocean submesoscale discussion in the review article by Thomas et al. (2008).

57.6.1 Hydrostatic and Boussinesq fluid on an f -plane

We frame our analysis within the adiabatic and hydrostatic Boussinesq equations (see Section 55.2) on an f -plane

$$\frac{Du}{Dt} - fv = -\frac{\partial \phi}{\partial x} \quad (57.79a)$$

$$\frac{Dv}{Dt} + fu = -\frac{\partial \phi}{\partial y} \quad (57.79b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (57.79c)$$

$$\frac{Db}{Dt} = 0 \quad (57.79d)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0. \quad (57.79e)$$

The geostrophic velocity is given in terms of the pressure field by

$$\mathbf{u}_g = f^{-1} \hat{z} \wedge \nabla_z \phi \implies u_g = -\frac{1}{f} \frac{\partial \phi}{\partial y} \quad v_g = \frac{1}{f} \frac{\partial \phi}{\partial x}. \quad (57.80)$$

The horizontal momentum equations (57.79a)-(57.79b) can thus be written

$$\mathbf{u} = \mathbf{u}_g + \hat{\mathbf{z}} \wedge \mathcal{D}\mathbf{u} \implies u = u_g - \mathcal{D}v \quad v = v_g + \mathcal{D}u, \quad (57.81)$$

where we introduced the dimensionless material time operator

$$\mathcal{D} = \frac{1}{f} \frac{D}{Dt}. \quad (57.82)$$

One step of iteration then leads to

$$u = u_g - \mathcal{D}(v_g + \mathcal{D}u) \quad (57.83a)$$

$$v = v_g + \mathcal{D}(u_g - \mathcal{D}v). \quad (57.83b)$$

Recall that on an f -plane, the geostrophic velocity \mathbf{u}_g is horizontally non-divergent.

57.6.2 Geostrophic momentum approximation

The *geostrophic momentum approximation* assumes the horizontal velocity takes the form

$$\mathbf{u} \approx \mathbf{u}_{gm} = \mathbf{u}_g + \hat{\mathbf{z}} \wedge \mathcal{D}\mathbf{u}_g. \quad (57.84)$$

For simplicity in notation, we drop the “gm” subscripts in the following. The geostrophic momentum approximation holds so long as

$$|u| \gg |\mathcal{D}^2 u| \quad (57.85a)$$

$$|v| \gg |\mathcal{D}^2 v|. \quad (57.85b)$$

Rearranging the geostrophic momentum approximation (57.84) leads to

$$\mathcal{D}\mathbf{u}_g + \hat{\mathbf{z}} \wedge (\mathbf{u} - \mathbf{u}_g) = 0. \quad (57.86)$$

Reintroducing the material time derivative yields

$$\frac{D\mathbf{u}_g}{Dt} + f\hat{\mathbf{z}} \wedge \mathbf{u}_{ag} = 0 \quad (57.87)$$

where \mathbf{u}_{ag} is the ageostrophic velocity based on the geostrophic momentum velocity

$$\mathbf{u}_{ag} = \mathbf{u} - \mathbf{u}_g = \hat{\mathbf{z}} \wedge \mathcal{D}\mathbf{u}_g. \quad (57.88)$$

Hence, for the geostrophic momentum approximation, the material time evolution of the geostrophic velocity is forced by the Coriolis acceleration due to the ageostrophic velocity. Furthermore, the material time derivative for the semi-geostrophic system is given by

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla = \frac{\partial}{\partial t} + (\mathbf{u}_g + \mathbf{u}_{ag}) \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (57.89)$$

with

$$\mathbf{v} = (\mathbf{u}_g + \mathbf{u}_{ag}) + \hat{\mathbf{z}} w. \quad (57.90)$$

57.6.3 Secondary ageostrophic circulation

In Sections 57.4 and 57.5 we focused on the stability of a geostrophically balanced equilibrium with flow along a symmetric front. In addition to the geostrophic flow along the front, there is generally an ageostrophic circulation that circulates in the plane orthogonal to the front; i.e., the cross-front / vertical plane. We here derive a general equation describing this overturning circulation, and then specialize that equation in Section 57.6.4. For that purpose, start from the zonal momentum equation, buoyancy equation, and continuity equation within the semi-geostrophic system

$$\frac{\partial u_g}{\partial t} + (\mathbf{u}_g \cdot \nabla_z) u_g + (\mathbf{u}_{ag} \cdot \nabla_z) u_g + w \frac{\partial u_g}{\partial z} - f v_{ag} = 0 \quad (57.91a)$$

$$\frac{\partial b}{\partial t} + \mathbf{u}_g \cdot \nabla_z b + \mathbf{u}_{ag} \cdot \nabla_z b + w N^2(z) = 0 \quad (57.91b)$$

$$\frac{\partial u_{ag}}{\partial x} + \frac{\partial v_{ag}}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (57.91c)$$

Note that any vertical flow is ageostrophic, so that it is not necessary to place the “ag” subscript on w .

The vertical derivative of the zonal momentum equation and the meridional derivative of the buoyancy equation lead to

$$\frac{\partial}{\partial t} \frac{\partial u_g}{\partial z} + \left(\frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial u_g}{\partial z} + \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial u_g}{\partial z} + \frac{\partial}{\partial z} (w \frac{\partial u_g}{\partial z}) - f \frac{\partial v_{ag}}{\partial z} = 0 \quad (57.92a)$$

$$\frac{\partial}{\partial t} \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial b}{\partial y} + \frac{\partial w}{\partial y} N^2 = 0. \quad (57.92b)$$

We now make use of thermal wind for the geostrophic velocity

$$f \frac{\partial \mathbf{u}_g}{\partial z} = \hat{\mathbf{z}} \wedge \nabla_z b \implies \frac{\partial u_g}{\partial z} = -\frac{1}{f} \frac{\partial b}{\partial y} \quad \frac{\partial v_g}{\partial z} = \frac{1}{f} \frac{\partial b}{\partial x} \quad (57.93)$$

thus rendering

$$\frac{\partial}{\partial t} \frac{\partial b}{\partial y} - f \left(\frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial b}{\partial y} - f \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial b}{\partial y} + \frac{\partial}{\partial z} (w \frac{\partial b}{\partial y}) + f^2 \frac{\partial v_{ag}}{\partial z} = 0 \quad (57.94a)$$

$$\frac{\partial}{\partial t} \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial b}{\partial y} + \frac{\partial w}{\partial y} N^2 = 0. \quad (57.94b)$$

Subtracting equation (57.94b) from equation (57.94a) eliminates the time derivative thus revealing the diagnostic relation

$$-f \left(\frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z \right) u_g - f \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + \frac{\partial}{\partial z} (w \frac{\partial b}{\partial y}) + f^2 \frac{\partial v_{ag}}{\partial z} - \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z \right) b - \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b - \frac{\partial w}{\partial y} N^2 = 0. \quad (57.95)$$

We now make use of thermal wind and horizontal non-divergence for the geostrophic velocity to write (as in the solution to Exercise 56.3)

$$f \frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z u_g = \frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z b = -Q^{(y)}, \quad (57.96)$$

where $Q^{(y)}$ is the meridional component of the geostrophic \mathbf{Q} -vector (see Exercise 56.3)

$$\mathbf{Q} = - \left(\frac{\partial \mathbf{u}_g}{\partial x} \cdot \nabla_z b \right) \hat{\mathbf{x}} - \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z b \right) \hat{\mathbf{y}}. \quad (57.97)$$

Introduction of $Q^{(y)}$ into equation (57.95) yields

$$-f \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + \frac{\partial w}{\partial z} \frac{\partial b}{\partial y} + w \frac{\partial^2 b}{\partial z \partial y} + f^2 \frac{\partial v_{ag}}{\partial z} - \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b - \frac{\partial w}{\partial y} N^2 = -2Q^{(y)}. \quad (57.98)$$

Again making use of thermal wind and $\nabla_z \cdot \mathbf{u}_g = 0$ allows us to write

$$-f\left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z\right)u_g - \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z\right)b = f\frac{\partial u_{ag}}{\partial z}\frac{\partial v_g}{\partial y} - f\frac{\partial v_{ag}}{\partial z}\frac{\partial u_g}{\partial y} + f\frac{\partial u_{ag}}{\partial y}\frac{\partial v_g}{\partial z} + f\frac{\partial v_{ag}}{\partial y}\frac{\partial u_g}{\partial z}. \quad (57.99)$$

The mixed partial derivative of the buoyancy vanishes

$$\frac{\partial^2 b}{\partial z \partial y} = \frac{\partial}{\partial y} \frac{\partial b}{\partial z} = \frac{\partial N^2(z)}{\partial y} = 0, \quad (57.100)$$

which follows since we are assuming a background vertical stratification that is independent of horizontal direction. Bringing these results together into equation (57.98) leads to

$$f\frac{\partial u_{ag}}{\partial z}\frac{\partial v_g}{\partial y} + f\frac{\partial u_{ag}}{\partial y}\frac{\partial v_g}{\partial z} + f\frac{\partial v_{ag}}{\partial y}\frac{\partial u_g}{\partial z} + \frac{\partial w}{\partial z}\frac{\partial b}{\partial y} + f\frac{\partial v_{ag}}{\partial z}\left(f - \frac{\partial u_g}{\partial y}\right) - \frac{\partial w}{\partial y}N^2 = -2Q^{(y)}. \quad (57.101)$$

Another use of thermal wind brings this result to the form

$$f\frac{\partial u_{ag}}{\partial z}\frac{\partial v_g}{\partial y} + f\frac{\partial u_{ag}}{\partial y}\frac{\partial v_g}{\partial z} + f\frac{\partial u_g}{\partial z}\left(\frac{\partial v_{ag}}{\partial y} - \frac{\partial w}{\partial z}\right) + f\frac{\partial v_{ag}}{\partial z}\left(f - \frac{\partial u_g}{\partial y}\right) - \frac{\partial w}{\partial y}N^2 = -2Q^{(y)}. \quad (57.102)$$

This equation provides a relation for the ageostrophic cross-flow and vertical circulation, (v_{ag}, w) , written in terms of the buoyancy field and the geostrophic flow. We next consider flow surrounding a symmetric front, in which case equation (57.102) becomes a diagnostic equation for the ageostrophic overturning streamfunction.

57.6.4 Ageostrophic overturning circulation for a symmetric front

The general result (57.102) is now specialized by assuming the zonal velocity is purely geostrophic (as in a zonal geostrophic front) so that

$$u_{ag} = 0. \quad (57.103)$$

For this flow, the ageostrophic flow in the cross-flow/depth plane is non-divergent

$$\frac{\partial v_{ag}}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (57.104)$$

The diagnostic equation (57.102) now takes on the specialized form for a symmetric front

$$-2f\frac{\partial u_g}{\partial z}\frac{\partial w}{\partial z} + f(f + \zeta_g)\frac{\partial v_{ag}}{\partial z} - \frac{\partial w}{\partial y}N^2 = -2Q^{(y)}, \quad (57.105)$$

where

$$\zeta_g = -\frac{\partial u_g}{\partial y} \quad (57.106)$$

is the vertical component of the geostrophic relative vorticity. Introducing an overturning streamfunction for the cross-flow/vertical ageostrophic circulation

$$v_{ag} = \frac{\partial \psi}{\partial z} \quad w = -\frac{\partial \psi}{\partial y}, \quad (57.107)$$

and using thermal wind brings equation (57.105) into the form

$$(N^2 \partial_{yy} - 2 \partial_y b \partial_{yz} + f(f + \zeta_g) \partial_{zz})\psi = -2Q^{(y)}. \quad (57.108)$$

Equation (57.108) is useful for the study of ageostrophic ($Ro \sim 1$) dynamics along a front in which there is an ageostrophic overturning circulation in response to geostrophic forcing from $Q^{(y)}$.

57.6.5 Connection to potential vorticity and symmetric instability

The partial differential equation (57.108) can be written

$$\mathcal{K}\psi = -2Q^{(y)}, \quad (57.109)$$

where

$$\mathcal{K} = N^2 \partial_{yy} - 2 \partial_y b \partial_{zy} + f(f + \zeta_g) \partial_{zz} \quad (57.110)$$

is a linear partial differential operator that is a function of the geostrophic flow and the buoyancy. Following the considerations in Section 5.4, we know that this operator is elliptic if the following inequality holds

$$\left(\frac{\partial b}{\partial y}\right)^2 - N^2 f(f + \zeta_g) < 0. \quad (57.111)$$

We can relate the ellipticity condition (57.111) to the Ertel potential vorticity for the Boussinesq geostrophic flow. For this purpose, write the geostrophic vorticity as

$$\boldsymbol{\omega}_g = \nabla \wedge \mathbf{u}_g \quad (57.112a)$$

$$= -\hat{x} \frac{\partial v_g}{\partial z} + \hat{y} \frac{\partial u_g}{\partial z} + \hat{z} \left(\frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \right) \quad (57.112b)$$

$$= -\frac{1}{f} \left(\hat{x} \frac{\partial b}{\partial x} + \hat{y} \frac{\partial b}{\partial y} \right) + \hat{z} \left(\frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \right). \quad (57.112c)$$

If we assume the front is zonally symmetric, then the geostrophic vorticity takes the form

$$\boldsymbol{\omega}_g^{2d} = -\frac{1}{f} \frac{\partial b}{\partial y} \hat{y} - \hat{z} \frac{\partial u_g}{\partial y}, \quad (57.113)$$

in which case the Ertel potential vorticity (for the geostrophic and Boussinesq flow) takes the form

$$q_g^{2d} = \nabla b \cdot (\boldsymbol{\omega}_g + f\hat{z}) \quad (57.114a)$$

$$= -\frac{1}{f} \left(\frac{\partial b}{\partial y} \right)^2 + N^2 \left(f - \frac{\partial u_g}{\partial y} \right). \quad (57.114b)$$

$$= -\frac{1}{f} \left(\frac{\partial b}{\partial y} \right)^2 + N^2 (f + \zeta_g). \quad (57.114c)$$

Ellipticity of the PDE (57.108) is thus assured so long as

$$fq_g^{2d} = -\left(\frac{\partial b}{\partial y}\right)^2 + N^2 f(f + \zeta_g) > 0. \quad (57.115)$$

The PDE (57.5) transitions to a hyperbolic system when $fq_g^{2d} < 0$, which is the condition for symmetric instability detailed in Section 57.5.

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.
- Acheson, D., *Elementary Fluid Dynamics*, Oxford Applied Mathematics and Computing Science Series, Oxford, Oxford, 1990.
- 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, 2004.
- Adcroft, A., and R. Hallberg, On methods for solving the oceanic equations of motion in generalized vertical coordinates, *Ocean Modelling*, 11, 224–233, 2006.
- 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.
- 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 the Earth System, submitted to JAMES*, 2019.
- 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.
- Andrews, D., and M. McIntyre, An exact theory of nonlinear waves on a Lagrangian-mean flow, *Journal of Fluid Mechanics*, 89, 609–646, 1978a.
- Andrews, D. G., and M. E. McIntyre, On wave action and its relatives, *Journal of Fluid Mechanics*, 89, 647–664, 1978b.
- Anstey, J., and L. Zanna, A deformation-based parametrization of ocean mesoscale eddy Reynolds stresses, *Ocean Modelling*, 112, 99–111, 2017.

- Apel, J. R., *Principles of Ocean Physics, International Geophysics Series*, vol. 38, Academic Press, London, 1987.
- Arbic, B., S. T. Garner, R. W. Hallberg, and H. L. Simmons, The accuracy of surface elevations in forward global barotropic and baroclinic tide models, *Deep Sea Research*, 51, 3069–3101, 2004.
- Aris, R., *Vectors, Tensors and the Basic Equations of Fluid Mechanics*, Dover Publishing, New York, 1962.
- Armi, L., Effects of variations in eddy diffusivity on property distributions in the oceans, *Journal of Marine Research*, 37, 515–530, 1979.
- Batchelor, G. K., *An Introduction to Fluid Dynamics*, Cambridge University Press, Cambridge, England, 1967.
- 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, 1998.
- 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, 2000.
- Bennett, A., *Lagrangian Fluid Dynamics*, Cambridge University Press, Cambridge, UK, 2006.
- Bleck, R., Finite difference equations in generalized vertical coordinates. Part I: Total energy conservation, *Contributions to Atmospheric Physics*, 51, 360–372, 1978.
- Bleck, R., An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates, *Ocean Modelling*, 4, 55–88, 2002.
- 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.
- Brown, E., *Waves, tides, and shallow-water processes*, 227 pp., The Open University, Milton Keys, UK, 1999.
- Bryan, K., A numerical method for the study of the circulation of the world ocean, *Journal of Computational Physics*, 4, 347–376, 1969.
- Bühler, O., *Waves and mean flows*, 2nd ed., Cambridge University Press, Cambridge, UK, doi: 10.1017/CBO9781107478701, 2014.
- Callen, H. B., *Thermodynamics and an Introduction to Thermostatics*, John Wiley and Sons, New York, 493 + xvi pp, 1985.
- Chaikin, P. M., and T. C. Lubensky, *Principles of Condensed Matter Physics*, Cambridge University Press, Cambridge, United Kingdom, 1995.
- Chandrasekhar, S., *Hydrodynamic and Hydromagnetic Stability*, Dover Publications, New York, 654 pp, 1961.

- 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.
- Chen, X., N. Andronova, B. van Leer, J. Penner, J. Boyd, C. Jablonowski, and S.-J. Lin, A control-volume model of the compressible Euler equations with a vertical Lagrangian coordinate, *Monthly Weather Review*, 141, 2526–2544, 2013.
- 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.
- Courant, R., and D. Hilbert, *Methods of Mathematical Physics Volume I*, Wiley-Interscience, New York, 1953.
- Courant, R., and D. Hilbert, *Methods of Mathematical Physics Volume II: Partial Differential Equations*, Wiley-Interscience, 1962.
- Cox, M. D., Isopycnal diffusion in a z -coordinate ocean model, *Ocean Modelling*, 74, 1–5, 1987.
- Cushman-Roisin, B., and J.-M. Beckers, *Introduction to Geophysical Fluid Dynamics*, Academic Press, Amsterdam, 828, 2011.
- Cushman-Roisin, B., Subduction, in *Dynamics of the oceanic surface mixed-layer*, pp. 181–196, Hawaii Institute of Geophysical Special Publications, 1987.
- 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.
- Davis, R. E., Diapycnal mixing in the ocean: equations for large-scale budgets, *Journal of Physical Oceanography*, 24, 777–800, 1994.
- DeGroot, S. R., and P. Mazur, *Non-Equilibrium Thermodynamics*, Dover Publications, New York, 510 pp, 1984.
- DeSzeoke, R. A., and A. F. Bennett, Microstructure fluxes across density surfaces, *Journal of Physical Oceanography*, 23, 2254–2264, 1993.
- Drazin, P., and R. Johnson, *Solitons: an introduction*, Cambridge University Press, Cambridge, UK, 226 pp, 1989.
- Drazin, P., and W. Reid, *Hydrodynamic stability*, Cambridge University Press, Cambridge, UK, 527 pp, 1981.
- Duchateau, P., and D. Zachmann, *Partial differential equations*, Schaum's Outline Series in Mathematics, McGraw-Hill, New York, 1986.
- Dukowicz, J. K., and J. R. Baumgardner, Incremental remapping as a transport/advection algorithm, *Journal of Computational Physics*, 160, 310–335, 2000.
- Dukowicz, J. K., and R. D. Smith, Stochastic theory of compressible turbulent fluid transport, *Physics of Fluids*, 9, 3523–3529, 1997.

Dunne, J. P., J. G. John, R. W. Hallberg, S. M. Griffies, E. N. Shevliakova, R. J. Stouffer, J. P. Krasting, L. A. Sentman, P. C. D. Milly, S. L. Malyshev, A. J. Adcroft, W. Cooke, K. A. Dunne, M. J. Harrison, H. Levy, B. L. Samuels, M. Spelman, M. Winton, A. T. Wittenberg, P. J. Phillips, and N. Zadeh, GFDLs ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics, *Journal of Climate*, 25, 6646–6665, 2012.

Eckart, C., An analysis of the stirring and mixing processes in incompressible fluids, *Journal of Marine Research*, 7, 265–275, 1948.

Einstein, A., Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen, *Annalen der Physik (in German)*, 322, 549–560, 1905.

Ertel, H., Ein neuer hydrodynamicscher wirbelsatz, *Meteorol. Z.*, 59, 271–281, 1942.

Falkovich, G., *Fluid Mechanics: A short course for physicists*, Cambridge University Press, 167pp, 2011.

Farrell, W., and J. Clark, On postglacial sea level, *Geophysical Journal of the Royal Astronomical Society*, 46, 646–667, 1976.

Feistel, R., Equilibrium thermodynamics of seawater revisited, *Progress in Oceanography*, 31, 101–179, 1993.

Ferrari, R., J. C. McWilliams, V. M. Canuto, and M. Dubovikov, Parameterization of eddy fluxes near oceanic boundaries, *Journal of Climate*, 21, 2770–2789, 2008.

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, 2010.

Ferreira, D., and J. Marshall, Formulation and implementation of a residual-mean ocean circulation model, *Ocean Modelling*, 13, 86–107, 2006.

Fetter, A. L., and J. D. Walecka, *Theoretical Mechanics of Particles and Continua*, McGraw-Hill Book Company, New York, 570 pp, 1980.

Fox-Kemper, B., R. Ferrari, and J. Pedlosky, A note on the indeterminacy of rotational and divergent eddy fluxes, *Journal of Physical Oceanography*, 33, 478–483, 2003.

Garrett, C., Turbulent dispersion in the ocean, *Progress in Oceanography*, 70, 113–125, 2006.

Gent, P. R., and J. C. McWilliams, Isopycnal mixing in ocean circulation models, *Journal of Physical Oceanography*, 20, 150–155, 1990.

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, 1995.

Gill, A., *Atmosphere-Ocean Dynamics, International Geophysics Series*, vol. 30, Academic Press, London, 662 + xv pp, 1982.

Gnanadesikan, A., A global model of silicon cycling: Sensitivity to eddy parameterization and dissolution, *Global Biogeochemical Cycles*, 13, 199–220, 1999.

- Graham, F., and T. McDougall, Quantifying the nonconservative production of Conservative Temperature, potential temperature, and entropy, *Journal of Physical Oceanography*, 43, 838–862, 2013.
- Greatbatch, R. J., A note on the representation of steric sea level in models that conserve volume rather than mass, *Journal of Geophysical Research*, 99, 12,767–12,771, 1994.
- 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, 1990.
- Gregg, M. C., Entropy generation in the ocean by small-scale mixing, *Journal of Physical Oceanography*, 14, 688–711, 1984.
- 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*, p. in review, 2019.
- Griffies, S. M., The Gent-McWilliams skew-flux, *Journal of Physical Oceanography*, 28, 831–841, 1998.
- Griffies, S. M., *Fundamentals of Ocean Climate Models*, Princeton University Press, Princeton, USA, 518+xxxiv pages, 2004.
- Griffies, S. M., and A. J. Adcroft, Formulating the equations for ocean models, in *Ocean Modeling in an Eddying Regime, Geophysical Monograph*, vol. 177, edited by M. Hecht and H. Hasumi, pp. 281–317, American Geophysical Union, 2008.
- 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.
- 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, 1998.
- 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, 2000a.
- Griffies, S. M., R. C. Pacanowski, and R. W. Hallberg, Spurious diapycnal mixing associated with advection in a z -coordinate ocean model, *Monthly Weather Review*, 128, 538–564, 2000b.
- Griffies, S. M., R. Pacanowski, M. Schmidt, and V. Balaji, Tracer conservation with an explicit free surface method for z -coordinate ocean models, *Monthly Weather Review*, 129, 1081–1098, 2001.
- Griffies, S. M., A. Biastoch, C. W. Böning, F. Bryan, G. Danabasoglu, E. Chassignet, M. H. England, R. Gerdes, H. Haak, R. W. Hallberg, W. Hazeleger, J. Jungclaus, W. G. Large, G. Madec, A. Pirani, B. L. Samuels, M. Scheinert, A. S. Gupta, C. A. Severijns, H. L. Simmons, A. M. Treguier, M. Winton, S. Yeager, and J. Yin, Coordinated Ocean-ice Reference Experiments (COREs), *Ocean Modelling*, 26, 1–46, doi:10.1016/j.ocemod.2008.08.007, 2009.

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. Illicak, 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. Volodire, 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.

Groeskamp, S., R. P. Abernathey, and A. Klocker, Water mass transformation by cabbeling and thermobaricity, *Geophysical Research Letters*, doi:10.1002/2016GL070860, 2016.

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.

Hallberg, R., A thermobaric instability in Lagrangian vertical coordinate ocean models, *Ocean Modelling*, 8, 227–300, 2005.

Hallberg, R., and A. Adcroft, Reconciling estimates of the free surface height in lagrangian vertical coordinate ocean models with mode-split time stepping, *Ocean Modelling*, 29, 15–26, 2009.

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.

Hallberg, R. W., Stable split time stepping schemes for large-scale ocean modeling, *Journal of Computational Physics*, 135, 54–65, 1997.

Haney, R. L., On the pressure gradient force over steep topography in sigma-coordinate ocean models, *Journal of Physical Oceanography*, 21, 610–619, 1991.

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.

Haynes, P. H., and M. E. McIntyre, On the conservation and impermeability theorems for potential vorticity, *Journal of Atmospheric Sciences*, 47, 2021–2031, 1990.

Hildebrand, F., *Advanced Calculus for Applications*, Prentice-Hall Publishers, Englewood Cliffs, New Jersey, 1976.

Hirt, C., A. Amsden, and J. Cook, An arbitrary Lagrangian-Eulerian computing method for all flow speeds, *Journal of Computational Physics*, 135, 203–216, 1997.

Holloway, G., Eddy transport of thickness and momentum in layer and level models, *Journal of Physical Oceanography*, 27, 1153–1157, 1997.

Holloway, G., and P. Rhines, Angular momenta of modeled ocean gyres, *Journal of Geophysical Research*, 27, 843–846, 1991.

Holton, J. R., *An Introduction to Dynamic Meteorology*, Academic Press, San Diego, USA, 507 pp, 1992.

- Hoskins, B., The role of potential vorticity in symmetric stability and instability, *Quarterly Journal of the Royal Meteorological Society*, 100, 480–482, 1974.
- Hoskins, B., The geostrophic momentum approximation and the semi-geostrophic equations, *Journal of Atmospheric Sciences*, 32, 233–242, 1975.
- Hoskins, B., I. Draghici, and H. Davies, A new look at the ω equation, *Quarterly Journal of the Royal Meteorological Society*, 104, 31–38, 1978.
- Hoskins, B. J., and I. N. James, *Fluid Dynamics of the Midlatitude Atmosphere*, Wiley Blackwell, Chichester, UK, 2014.
- Huang, K., *Statistical Mechanics*, John Wiley and Sons, New York, 493 pp, 1987.
- Huang, R. X., Real freshwater flux as a natural boundary condition for the salinity balance and thermohaline circulation forced by evaporation and precipitation, *Journal of Physical Oceanography*, 23, 2428–2446, 1993.
- Huang, R. X., and R. W. Schmitt, Goldsbrough-stommel circulation of the world oceans, *Journal of Physical Oceanography*, 23, 1277–1284, 1993.
- 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.
- 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, 196pp, 2010.
- Jackson, J. D., *Classical Electrodynamics*, John Wiley and Sons, New York, USA, 848 pp, 1975.
- Killworth, P. D., On the parameterization of eddy transfer Part I: Theory, *Journal of Marine Research*, 55, 1171–1197, 1997.
- Killworth, P. D., D. Stainforth, D. J. Webb, and S. M. Paterson, The development of a free-surface Bryan-Cox-Semtner ocean model, *Journal of Physical Oceanography*, 21, 1333–1348, 1991.
- Klocker, A., and R. Abernathey, Global patterns of mesoscale eddy properties and diffusivities, *Journal of Physical Oceanography*, 44, 1030–1046, 2014.
- 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, 2010.
- Kundu, P., I. Cohen, and D. Dowling, *Fluid Mechanics*, Academic Press, 921 + xxiv pp, 2012.
- Kushner, P. J., and I. M. Held, Potential vorticity thickness fluxes and wave-mean flow interaction, *Journal of Atmospheric Sciences*, 56, 948–958, 1999.
- Landau, L. D., and E. M. Lifshitz, *Mechanics*, Pergamon Press, Oxford, UK, 170 pp, 1976.
- Landau, L. D., and E. M. Lifshitz, *Fluid Mechanics*, Pergamon Press, Oxford, UK, 539 pp, 1987.
- 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, 1994.

- 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.
- 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, 2012.
- Lilly, J., Kinematics of a fluid ellipse in a linear flow, *Fluids*, 3(16), doi:10.3390/fluids3010016, 2018.
- 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, 1997.
- Lin, S.-J., A vertically lagrangian finite-volume dynamical core for global models, *Monthly Weather Review*, 132, 2293–2307, 2004.
- 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, 2013.
- Marion, J. B., and S. T. Thornton, *Classical Dynamics of Particles and Systems*, Harcourt Brace Jovanovich, San Diego, USA, 602 pp, 1988.
- Markowski, P., and Y. Richardson, *Mesoscale Meteorology in Midlatitudes*, Wiley-Blackwell Publishers, Oxford, UK, 2010.
- Marshall, D., Vertical fluxes of potential vorticity and the structure of the thermocline, *Journal of Physical Oceanography*, 30, 3102–3112, 2000.
- Marshall, J., and R. A. Plumb, *Atmosphere, Ocean, and Climate Dynamics*, 1st ed., Academic Press, Amsterdam, 319pp, 2008.
- Marshall, J., D. Jamous, and J. Nilsson, Reconciling thermodynamic and dynamic methods of computation of water-mass transformation rates, *Deep-Sea Research I*, 46, 545–572, 1999.
- 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.
- 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.
- McDougall, T. J., Neutral surfaces, *Journal of Physical Oceanography*, 17, 1950–1967, 1987a.
- McDougall, T. J., Thermobaricity, cabbeling, and water-mass conversion, *Journal of Geophysical Research*, 92, 5448–5464, 1987b.
- McDougall, T. J., Potential enthalpy: a conservative oceanic variable for evaluating heat content and heat fluxes, *Journal of Physical Oceanography*, 33, 945–963, 2003.
- McDougall, T. J., and R. Feistel, What causes the adiabatic lapse rate, *Deep-Sea Research*, 50, 1523–1535, 2003.

- McDougall, T. J., and D. R. Jackett, On the helical nature of neutral trajectories in the ocean, *Progress in Oceanography*, 20, 153–183, 1988.
- McDougall, T. J., and D. R. Jackett, The thickness of the ocean in $s - \theta - p$ space and the implications for mean diapycnal advection, *Journal of Physical Oceanography*, 37, 1714–1732, 2007.
- 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, 2001.
- 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.
- McDougall, T. J., J. A. Church, and D. R. Jackett, Does the nonlinear equation of state impose an upper bound on the buoyancy frequency?, *Journal of Marine Research*, 61, 745–764, 2003.
- McDougall, T. J., S. Groeskamp, and S. M. Griffies, On geometric aspects of interior ocean mixing, *Journal of Physical Oceanography*, 44, 2164–2175, 2014.
- McIntosh, P. C., and T. J. McDougall, Isopycnal averaging and the residual mean circulation., *Journal of Physical Oceanography*, 26, 1655–1660, 1996.
- McWilliams, J., *Fundamentals of Geophysical Fluid Dynamics*, Cambridge University Press, Cambridge, Cambridge, UK, 2006.
- McWilliams, J. C., The emergence of isolated coherent vortices in turbulent flow, *Journal of Fluid Mechanics*, 146, 21–43, 1984.
- McWilliams, J. C., Submesoscale currents in the ocean, *Proceedings of the Royal Society, A472*, doi:10.1098/rspa.2016.0117, 2016.
- 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.
- Middleton, J. F., and J. W. Loder, Skew fluxes in polarized wave fields, *Journal of Physical Oceanography*, 19, 68–76, 1989.
- Mitrovica, J. X., M. E. Tamisiea, J. L. Davis, and G. A. Milne, Recent mass balance of polar ice sheets inferred from patterns of global sea-level change, *Nature*, 409, 1026–1029, 2001.
- Moffatt, H., Transport effects associated with turbulence with particular attention to the influence of helicity, *Reports on Progress in Physics*, 46, 621–664, 1983.
- Moffatt, H., Helicity and singular structures in fluid dynamics, *Proceedings of the National Academy of Science*, 111, 3663–3670, 2014.
- Morse, P. M., and H. Feshbach, *Methods of Theoretical Physics Part I and II*, McGraw-Hill Book Company, New York, 1953.
- Müller, P., Ertel's potential vorticity theorem in physical oceanography, *Reviews of Geophysics*, 33, 67–97, 1995.

- Müller, P., *The Equations of Oceanic Motions*, 1st ed., Cambridge University Press, Cambridge, 302pp, 2006.
- Müller, P., and C. Garrett, From stirring to mixing in a stratified ocean, *Oceanography*, 15, 12–19, 2002.
- Munk, W., and C. Wunsch, Abyssal recipes II: Energetics of tidal and wind mixing, *Deep-Sea Research*, 45, 1977–2010, 1998.
- Munk, W. H., Abyssal recipes, *Deep-Sea Research*, 13, 707–730, 1966.
- 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.
- Naveira Garabato, A., E. Frajka-Williams, C. Spingys, A. Legg, K. Polzin, A. Forryan, E. Abrahamsen, C. Buckingham, S. Griffies, S. McPhail, K. Nicholls, L. Thomas, and M. 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.
- Nurser, A. G., and S. M. Griffies, Relating diffusive surface salinity fluxes to boundary freshwater and salt fluxes, *Journal of Physical Oceanography*, *in revision*, 2019.
- 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.
- Nurser, A. G., and M.-M. Lee, Isopycnal averaging at constant height. Part I: Relating to the residual streamfunction in Eulerian space, *Journal of Physical Oceanography*, 34, 2740–2755, 2004b.
- Olbers, D. J., J. Willebrand, and C. Eden, *Ocean Dynamics*, 1st ed., Springer, Berlin, Germany, 704 pages, 2012.
- 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.
- Pedlosky, J., *Geophysical Fluid Dynamics*, 2nd ed., Springer-Verlag, Berlin Heidelberg New York, 710 + xv pp, 1987.
- Plumb, R. A., Eddy fluxes of conserved quantities by small-amplitude waves, *Journal of Atmospheric Sciences*, 36, 1699–1704, 1979.
- Polton, J., and D. Marshall, Overturning cells in the Southern Ocean and subtropical gyres, *Ocean Science*, 3, 17–30, 2007.
- 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.
- Pugh, D. T., *Tides, surges, and mean sea-level*, 472 pp., John Wiley and Sons, 1987.
- Redi, M. H., Oceanic isopycnal mixing by coordinate rotation, *Journal of Physical Oceanography*, 12, 1154–1158, 1982.

- Reichl, L. E., *A Modern Course in Statistical Physics*, John Wiley and Sons, New York, 822 pp, 1987.
- Reif, F., *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York, 1965.
- Rhines, P. B., Basic dynamics of the large-scale geostrophic circulation, in *WHOI 1982 Summer Study Program*, Woods Hole Oceanographic Institute, 1982.
- Rhines, P. B., and W. R. Young, Homogenization of potential vorticity in planetary gyres, *Journal of Fluid Mechanics*, 122, 347–367, 1982.
- Rintoul, S. R., The global influence of localized dynamics in the Southern Ocean, *Nature*, 558, 209–218, doi:10.1038/s41586-018-0182-3, 2018.
- Rintoul, S. R., and A. C. Naveira Garabato, Dynamics of the Southern Ocean 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. 471–492, Academic Press, 2013.
- Rintoul, S. R., C. W. Hughes, and D. Olbers, The Antarctic Circumpolar Current system, in *Ocean Circulation and Climate, 1st Edition, International Geophysics Series*, vol. 103, edited by G. Siedler, J. Gould, and J. Church, pp. 271–301, Academic Press, 2001.
- Roekel, L. V., 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 the Earth System*, doi:10.1029/2018ms001336, 2018.
- Salmon, R., *Lectures on Geophysical Fluid Dynamics*, Oxford University Press, Oxford, England, 378 + xiii pp., 1998.
- Samelson, R., *The Theory of Large-Scale Ocean Circulation*, Cambridge University Press, Cambridge, UK, 193 pp., 2011.
- Schär, C., A generalization of Bernoulli's Theorem, *Journal of the Atmospheric Sciences*, 50, 1437–1443, doi:10.1029/2003JC001823, 1993.
- Schey, H., *Div, grad, curl and all that: an informal text on vector calculus*, W.W. Norton and Company, Inc., 176 pp, 2004.
- Schmitt, R. W., Double diffusion in oceanography, *Annual Review of Fluid Mechanics*, 26, 255–285, 1994.
- Schutz, B. F., *A First Course in General Relativity*, Cambridge University Press, Cambridge, UK, 392 pp, 1985.
- 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.
- Shchepetkin, A., and J. McWilliams, The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, *Ocean Modelling*, 9, 347–404, 2005.

Smagorinsky, J., General circulation experiments with the primitive equations: I. The basic experiment, *Monthly Weather Review*, 91, 99–164, 1963.

Smagorinsky, J., Some historical remarks on the use of nonlinear viscosities, in *Large Eddy Simulation of Complex Engineering and Geophysical Flows*, edited by B. Galperin and S. A. Orszag, pp. 3–36, Cambridge University Press, 1993.

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.

Smith, K. S., and G. K. Vallis, The scales and equilibration of midocean eddies: freely evolving flow, *Journal of Physical Oceanography*, 31, 554–570, 2001.

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.

Smith, R. D., The primitive equations in the stochastic theory of adiabatic stratified turbulence, *Journal of Physical Oceanography*, 29, 1865–1880, 1999.

Spiegel, M., *Theory and Problems of Vector Analysis*, Schaum's Outline Series, McGraw-Hill International Book Company, New York, 1974.

Stacey, M. W., S. Pond, and Z. P. Nowak, A numerical model of the circulation in Knight Inlet, British Columbia, Canada, *Journal of Physical Oceanography*, 25, 1037–1062, 1995.

Stakgold, I., *Boundary value problems of mathematical physics, volume I*, SIAM, Philadelphia, 340 pp, 2000a.

Stakgold, I., *Boundary value problems of mathematical physics, volume II*, SIAM, Philadelphia, 408 pp, 2000b.

Stanley, G. J., Neutral surface topology, *Ocean Modelling*, doi:10.1016/j.ocemod.2019.01.008, 2019.

Starr, V. P., A quasi-Lagrangian system of hydrodynamical equations, *Journal of Meteorology*, 2, 227–237, 1945.

Stewart, R., *An Introduction to Physical Oceanography*, 345 pp., Texas A& M, College Station, Texas, USA, 2008.

Straub, D. N., On thermobaric production of potential vorticity in the ocean, *Tellus*, 51A, 314–325, 1999.

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.

Talley, L. D., G. L. Pickard, W. J. Emery, and J. H. Swift, *Descriptive Physical Oceanography*, 6th ed., Elsevier, 555pp, 2011.

Taylor, G., Diffusion by continuous movements, *Proceedings of the London Mathematical Society*, 20, 196–212, 1921.

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.

- 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.
- Thorne, K., and R. Blandford, *Modern Classical Physics*, Princeton University Press, Princeton, USA, 1511 + xl pp, 2017.
- Tomczak, M., and J. S. Godfrey, *Regional Oceanography: An Introduction*, Pergamon Press, Oxford, England, 422 + vii pp, 1994.
- Treguier, A. M., I. M. Held, and V. D. Larichev, On the parameterization of quasi-geostrophic eddies in primitive equation ocean models, *Journal of Physical Oceanography*, 27, 567–580, 1997.
- Vallis, G. K., *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*, 1st ed., Cambridge University Press, Cambridge, 745 + xxv pp, 2006.
- Vallis, G. K., *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*, 2nd ed., Cambridge University Press, Cambridge, 946 + xxv pp, 2017.
- Vallis, G. K., *Essentials of Atmospheric and Oceanic Dynamics*, 1st ed., Cambridge University Press, Cambridge, 2019.
- van Heijst, G., Dynamics of vortices in rotating and stratified flows, in *Fronts, Waves, and Vortices in Geophysical Flows*, Lecture notes in Physics 805, p. 192, Springer, 2010.
- van Sebille, E., S. M. Griffies, R. Abernathey, T. Adams, P. Berloff, A. Biastoch, B. Blanke, E. Chassignet, Y. Cheng, C. Cotter, E. Deleersnijder, K. Döös, H. Drake, S. Drijfhout, S. Gary, A. Heemink, J. Kjellsson, I. Koszalka, M. Lange, C. Lique, G. MacGilchrist, R. Marsh, G. M. Adame, R. McAdam, F. Nencioli, C. Paris, M. Piggott, J. Polton, S. Rühs, S. Shah, M. Thomas, J. Wang, P. Wolfram, L. Zanna, and J. Zika, Lagrangian ocean analysis: fundamentals and practices, *Ocean Modelling*, 121, 49–75, doi:10.1016/j.ocemod.2017.11.008, 2018.
- Veronis, G., Large scale ocean circulation, *Advances in Applied Mechanics*, 13, 2–92, 1973.
- Walin, G., A theoretical framework for the description of estuaries, *Tellus*, 29(2), 128–136, doi: 10.1111/j.2153-3490.1977.tb00716.x, 1977.
- Walin, G., On the relation between sea-surface heat flow and thermal circulation in the ocean, *Tellus*, 34, 187–195, 1982.
- Wallace, J., and P. Hobbs, *Atmospheric Science: An Introductory Survey*, Academic Press, 2006.
- 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.
- Weiss, J., The dynamics of enstrophy transfer in two-dimensional hydrodynamics, *Physica D*, 273–294, 1991.
- Welander, P., Studies of the general development of motion in a two-dimensional ideal fluid, *Tellus*, 7, 141–156, 1955.

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.

White, L., and A. Adcroft, A high-order finite volume remapping scheme for nonuniform grids: The piecewise quartic method (PQM), *Journal of Computational Physics*, 227, 7394–7422, 2008.

White, L., A. Adcroft, and R. W. Hallberg, High-order regriddingremapping schemes for continuous isopycnal and generalized coordinates in ocean models, *Journal of Computational Physics*, 228, doi:10.1016/j.jcp.2009.08.016, 2009.

Williams, J., and S. Elder, *Fluid Physics for Oceanographers and Physicists: an introduction to incompressible flow*, Pergamon Press, Oxford, 1989.

Wolfe, C., Approximations to the oceans residual circulation in arbitrary tracer coordinates, *Ocean Modelling*, 75, 20–35, doi:10.1016/j.ocemod.2013.12.004, 2014.

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.

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, 2008.