

A HANDBOOK FOR GEOPHYSICAL FLUID MECHANICS WITH SPECIAL TOPICS IN PHYSICAL OCEANOGRAPHY

Stephen M. Griffies
Princeton University Program in Atmospheric and Oceanic Sciences
smg@princeton.edu

Draft from April 8, 2019

COPYRIGHT ©2019 BY STEPHEN M. GRIFFIES

ALL RIGHTS RESERVED

GRIFFIES, STEPHEN M., 1962-

A HANDBOOK FOR GEOPHYSICAL FLUID MECHANICS WITH SPECIAL TOPICS IN PHYSICAL OCEANOGRAPHY
/ STEPHEN M. GRIFFIES.

THIS BOOK WAS COMPOSED IN L^AT_EX.

Contents

PREFACE	vii
Part I. Mathematical tools 3	
1 Linear partial differential equations	7
2 Cartesian tensor algebra	15
3 Cartesian tensor calculus	27
4 General tensors in brief	43
5 General tensor algebra [†]	49
6 General tensor calculus [†]	59
7 Orthogonal coordinates [†]	67
8 Generalized vertical coordinates [†]	81
9 Tracer coordinates [†]	101
Part II. Particle mechanics in a rotating frame 107	
10 Particle kinematics	109
11 Particle dynamics	127
12 Symmetries and conservation laws	133
Part III. Fluid kinematics 145	
13 The continuum hypothesis [†]	147
14 Kinematics of fluid motion	155
15 Mass conservation	185
16 Incompressible flow	211
17 Material fluid objects [†]	221
18 Kinematics with generalized vertical coordinates [†]	231

Part IV. Thermodynamics and buoyancy	247
19 Thermodynamics	249
20 Buoyancy acceleration	275
21 Ocean buoyancy[†]	287
 Part V. Tracers	 299
22 Advection and diffusion	301
23 Tracer kinematics[†]	325
24 Ocean tracer subgrid scale transport[†]	357
25 Ocean tracer analysis[†]	379
 Part VI. Geophysical fluid dynamics	 381
26 Momentum and energy dynamics	383
27 Approximate dynamical equations	397
28 Stress in fluids	413
29 Oceanic Boussinesq fluid	423
30 Rapidly rotating fluids	435
31 Dynamics with generalized vertical coordinates[†]	453
32 Space-time dependent gravity[†]	477
 Part VII. The shallow water system	 489
33 Formulation of shallow water models	491
34 Shallow water dynamics	507
35 Gravity waves and geostrophic adjustment	523
 Part VIII. Vorticity and potential vorticity	 537
36 Shallow water vorticity and potential vorticity	539
37 Vorticity and circulation	549

38 Potential vorticity	571
39 Ocean potential vorticity [†]	583
40 Angular momentum, vorticity, and strain [†]	601
Part IX. Balanced models	609
41 Two-dimensional barotropic flows	611
42 Shallow water PG and QG	617
43 Continuously stratified PG and QG	633
44 Local stability of fronts [†]	661

Contents

PREFACE

Synopsis of this book

Geophysical fluid mechanics (GFM) is a scientific discipline focusing on physical processes and emergent phenomena exhibited by fluid motion on rotating and gravitating planets. This book presents mathematics, kinematics, dynamics, thermodynamics, and tracer mechanics as part of formulating a variety of physical models of use for GFM. We study perfect fluids, which only consider reversible processes, as well as in real fluids, which include multiple constituents and involve irreversible processes such as mixing. We also encounter specialized topics from physical oceanography, thus exemplifying how geophysical fluid mechanics can be of use for understanding ocean processes.

The presentation is largely deductive with a degree of rigor and generality realized through the use of physically informed mathematical analysis. The material is presented at the level of a senior undergraduate student of physics, engineering, or applied mathematics. Mastery of the material requires a working knowledge of classical Newtonian mechanics, equilibrium thermodynamics, continuum field theory (electromagnetism, solid mechanics, or fluid mechanics), vector calculus, tensor analysis, and partial differential equations. Most of these pre-requisites are developed within this book. Hence, for readers finding the waters rather deep, patience and persistence will be rewarded.

Two pillars of geophysical fluid mechanics

There are two general pillars to GFM, with the first comprising the physical and mathematical formulation of the conceptual models used to garner insight into rotating and stratified fluids. This pillar forms the focus of this book. The second pillar comprises a study of phenomena emerging from the conceptual models, such as waves, instabilities, turbulence, and general circulation. We emphasize the physical foundations of theoretical models thus facilitating an understanding of the mechanics of emergent phenomena. Pillar one material is more than sufficient to fill a book. Hence, we leave the exploration of emergent phenomena for other treatments. Correspondingly, phenomenology is introduced only so far as motivating theoretical developments.

Organization of this book

This book is organized into parts according to their particular focus, with chapters within parts split into two tiers. Each chapter starts with a brief overview and reader's guide pointing to dependencies of the material in that chapter. Tier-I chapters focus on topics required for a basic

understanding of the subject and offer exercises to test that understanding. This material can be suitable for a first semester graduate course. Tier-II chapters (and a few sections) are denoted by a dagger[†]. They offer monograph-style special topics that further the foundations and exemplify applications. Portions of Tier-II material are geared toward topics in physical oceanography, although much of that material remains of a general nature.

Emphasizing the synergy between physics and maths

This book grew from class notes developed for a first semester graduate course in geophysical fluid mechanics at Princeton University. This 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 the milestone books by [Vallis \(2017\)](#) and [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. I thus offer the following personal motivation for pursuing that evolution.

Geophysical fluid mechanics is a discipline within the broader pantheon of theoretical physics, providing a mechanistic framework for studies in physical oceanography, atmospheric science, climate dynamics, and planetary circulations. Typical treatments eschew certain of the foundational physical and mathematical concepts. Conversely, those treatments that dive into the foundations often assume prior knowledge well beyond most practitioners. This perception motivated the writing of 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, providing practitioners a theoretical platform for research and teaching thus prompting 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 book is an attempt to exemplify that synergy.

Caveats

This book is incomplete, with a variety of topics remaining to be included (e.g., Hamilton’s Principle, Rossby waves, baroclinic instability). More effort is also needed to unify notation, build an index, extend the exercises, enhance figures, improve discussions, correct errors, and add missing topics. Feedback is solicited.

A sample course syllabus

The following offers a sample syllabus for 24 lectures of order 80-minutes each. This syllabus has formed the basis for a one-semester graduate level course on geophysical fluid mechanics at Princeton University (AOS 571). Students are expected to have read the relevant material prior to class lectures. Homework exercises will be assigned as they arise during the lectures.

- LECTURE 1: Introduction to the course and overview of mathematical tools from Part I of the book. Some students will need to spend more time on this material than others in order to ensure sufficient maths baseline for the course.
- LECTURE 2: Particle mechanics: kinematics of a particle moving around a sphere, including position, velocity, and acceleration in terms of Cartesian and spherical coordinates (Chapter 10).
- LECTURE 3: Particle mechanics: finish kinematics of particle moving around a gravitating sphere (Chapter 10); dynamics of particle around a gravitating sphere, including Newton's equation of motion, rotating reference frame, Coriolis and centrifugal accelerations, and the gravitational geopotential (Chapter 11).
- LECTURE 4: Particle mechanics: conservation laws for particle moving around a gravitating sphere, including mechanical energy and angular momentum (Chapter 12).
- LECTURE 5: Fluid kinematics: Eulerian and Lagrangian descriptions, Galilean invariance, material time derivative, flow lines (Chapter 14).
- LECTURE 6: Fluid kinematics: Stokes drift (Chapter 14), mass conservation (Chapter 15).
- LECTURE 7: Fluid kinematics: mass conservation, kinematic boundary conditions (Chapter 15).
- LECTURE 8: Fluid kinematics: incompressible flow kinematics, including streamfunctions, area and volume conservation, meridional-depth overturning circulation (Chapter 16).
- LECTURE 9: Thermodynamics and tracers: develop thermodynamics for fluid flows (Chapter 19).

- LECTURE 10: Thermodynamics and tracers: more on thermodynamics (Chapter 19) and stratification (Chapter 20)
- LECTURE 11: Thermodynamics and tracers: advection/diffusion (Chapter 22).
- LECTURE 12: Geophysical fluid dynamics: dynamical equations for rotating and stratified fluids: Navier-Stokes in a rotating frame, primitive equations (Chapter 26).
- LECTURE 13: Geophysical fluid dynamics: approximate dynamical equations for rotating and stratified fluids: Primitive equations, tangent plane approximations, Boussinesq approximation, hydrostatic balance (Chapters 27 and 29).
- LECTURE 14: Geophysical fluid dynamics: diagnostic relations for rotating fluids (Chapter 30).
- LECTURE 15: Geophysical fluid dynamics: planetary geostrophy, Taylor-Proudman, and thermal wind (Chapter 30), then start shallow water formulation (Chapter 33).
- LECTURE 16: Shallow water model: formulation of shallow water model (Chapter 33), and start shallow water dynamics (Chapter 34).
- LECTURE 17: Shallow water model: finish shallow water dynamics (Chapter 34).
- LECTURE 18: Shallow water model: gravity waves and geostrophic adjustment (Chapter 35).
- LECTURE 19: Vorticity and potential vorticity: vorticity and potential vorticity for the shallow water system (Chapter 36) and start general discussion of vorticity (Chapter 37).
- LECTURE 20: Vorticity and potential vorticity: vorticity in a continuously stratified fluid (Chapter 37).
- LECTURE 21: Vorticity and potential vorticity: potential vorticity in a continuously stratified fluid (Chapter 38).
- LECTURE 22: Balanced models: Two-dimensional barotropic model (Chapter 41); Buckingham's Π theorem and the non-dimensional shallow water system; planetary geostrophy and quasi-geostrophy in the shallow water system (Chapter 42).
- LECTURE 23: Balanced models: continuously stratified quasi-geostrophy (Chapter 43).
- LECTURE 24: Catch-up and class review.

Part I

Mathematical tools

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 some mathematical topics with a focus on how they are useful for GFM. The experienced reader may choose to skim this material whereas the novice may encounter topics requiring careful study.

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 to compare with experiments.

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 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. An adept use of tensor analysis reveals how maths can inform the physics.

- There is a duality in fluid kinematics between Eulerian and Lagrangian descriptions of motion. To develop an understanding of this duality we make use of tensor analysis to facilitate a 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 Cartesian and spherical coordinates.
- Rotating laboratory fluids move in a circular tank. Consequently, cylindrical polar coordinates are the preferred choice for respecting the symmetry. We thus 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 parcel according to its buoyancy rather than its height. This “isopycnal” vertical coordinate choice leads to a non-orthogonal coordinate description of the fluid motion. Transforming between a Cartesian and an isopycnal coordinate description requires the mathematical precision of general tensors.

SUMMARY OF THE CHAPTERS IN THIS PART

Certain chapters in this part of the book are essential for nearly all subsequent chapters, whereas other chapters target the aficionados and serve somewhat limited, albeit interesting, 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. For those who do not penetrate these pages, we provide “as needed” summaries of salient points in later chapters.

- PARTIAL DIFFERENTIAL EQUATIONS: Chapter 1 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 the behavior of the fluid.
- CARTESIAN TENSOR ALGEBRA: Chapter 2 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 3 extends the algebraic ideas from Chapter 2 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.
- GENERAL TENSOR INTRODUCTION: Chapter 4 provides an introduction to general tensor analysis and its applications to geophysical fluids. The discussion is accessible to anyone who has read Chapter 2, 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 5 and 6.
- GENERAL TENSOR ALGEBRA: Chapter 5 extends the Cartesian tensor algebra from Chapter 2 to allow for the use of arbitrary, or general, coordinates. The material in this chapter is essential for those wishing to understand the mathematics underlying non-Cartesian coordinates, such as spherical and isopycnal coordinates.
- GENERAL TENSOR CALCULUS: Chapter 6 extends the Cartesian tensor calculus from Chapter 3 to the case of general coordinates. Again, the material in this chapter is essential for those wishing to understand the mathematics underlying non-Cartesian coordinates, such as spherical and isopycnal coordinates.

- EXAMPLE ORTHOGONAL COORDINATE SYSTEMS: Chapter 7 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 8 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.
- TRACER COORDINATES: Chapter 9 offers a reference for the mathematics of tracer coordinates. These coordinates can be used to describe the circulation of fluids in tracer space, with a tracer coordinate description requiring the full power of general tensor analysis.

1

Linear partial differential equations

Fluid mechanics is a nonlinear field theory. Consequently, its mathematical description involves nonlinear partial differential equations (PDEs). Nonetheless, on occasion it is useful to approximate a nonlinear PDE by its linear cousin. For example, a gently perturbed fluid responds by exhibiting wave fluctuations. Waves provide but one example where it is useful to understand salient properties of linear PDEs, thus motivating this chapter. For simplicity, this chapter makes use of Cartesian coordinates.

READER'S GUIDE TO THIS CHAPTER

The study of linear PDEs encountered in mathematical physics, even when restricted to fluid mechanics, is a vast endeavour. Our treatment is extremely terse and meant only to introduce basic features. One can find many resources devoted to the subject throughout the physics, engineering, and applied mathematics literature, with chapter 8 of *Hildebrand (1976)* offering a useful starting point. We here assume a basic knowledge of ordinary differential equations and partial differential calculus. We make use of many features from this chapter throughout the remainder of the book.

1.1	Open threads	7
1.2	The advection equation	8
1.2.1	Constant advection velocity	8
1.2.2	Specifying the arbitrary functions resulting from PDEs	8
1.2.3	Further reading	9
1.3	Characteristic curves for first order PDEs	9
1.3.1	General formulation	9
1.3.2	Examples	10
1.4	Second order PDEs	11
1.4.1	Wave equation	12
1.4.2	Heat or diffusion equation	12
1.4.3	Laplace's equation and Poisson's equation	13

1.1 Open threads

- Work through some analytic solutions to the diffusion equation.

- Introduce Green's function method.

1.2 The advection equation

Consider a tracer concentration, C , which for our present purposes is a scalar field that is a function of space and time. As derived in Section 22.4, the tracer concentration in the absence of diffusion satisfies the advection equation

$$\frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C = 0. \quad (1.1)$$

The highest derivatives in both space and time are first order, indicating that the advection equation is a first order PDE. It is a nonlinear PDE for *active* tracers such as temperature, with active tracers defined as those that affect the velocity through changes to pressure. In contrast, the advection equation is linear for *passive* tracers (e.g., dust), defined as those tracers that do not significantly alter velocity.

1.2.1 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,

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} = 0 \quad (1.2)$$

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

$$C(x, t) = \Gamma(x - Ut) \quad (1.3)$$

is a general solution to equation (1.2), where Γ is an arbitrary differentiable function. Verification of this result is readily found by noting

$$\frac{\partial C}{\partial x} = \frac{d\Gamma(x - Ut)}{d(x - Ut)} \frac{\partial(x - Ut)}{\partial x} = \Gamma' \quad (1.4a)$$

$$\frac{\partial C}{\partial t} = \frac{d\Gamma(x - Ut)}{d(x - Ut)} \frac{\partial(x - Ut)}{\partial t} = -U \Gamma'. \quad (1.4b)$$

The functional dependence $x - Ut$ reveals that as time progresses as if $U > 0$, an observer that moves in the positive \hat{x} direction with a speed U maintains a constant argument to the solution. This behaviour reflects the expected result that the tracer concentration is being transported without changing its structure through advection with a speed U .¹ We illustrate this behaviour in Figure 1.1.

1.2.2 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

$$C(x, t = 0) = C_0 \sin x, \quad (1.5)$$

¹Some disciplines refer to the linear advection equation (??), with constant advection speed, as the one-way wave equation or the square-root of the wave equation.

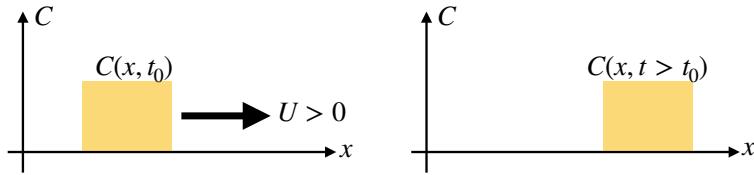


Figure 1.1: Illustrating the advection of a scalar field from a constant advection velocity $\mathbf{v} = U\hat{x}$ with $U > 0$. The initially square pulse of tracer is translated, unchanged, by the constant advection velocity.

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

$$C(x, t) = C_0 \sin(x - Ut). \quad (1.6)$$

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

1.2.3 Further reading

We further develop both the mathematics and physics of the advection equation in Sections 22.4 and 22.5.

1.3 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

$$P(x, t, \psi) \frac{\partial \psi}{\partial x} + Q(x, t, \psi) \frac{\partial \psi}{\partial t} = R(x, t, \psi), \quad (1.7)$$

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.

1.3.1 General formulation

For the first order PDE (1.7) assume there is a functional relation

$$\Upsilon(x, t, \psi) = \text{constant} \quad (1.8)$$

that determines ψ consistent with the PDE (1.7). 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

$$\frac{d\Upsilon}{dt} = 0 = \frac{\partial \Upsilon}{\partial \psi} \frac{\partial \psi}{\partial t} + \frac{\partial \Upsilon}{\partial t} \quad (1.9a)$$

$$\frac{d\Upsilon}{dx} = 0 = \frac{\partial \Upsilon}{\partial \psi} \frac{\partial \psi}{\partial x} + \frac{\partial \Upsilon}{\partial x}. \quad (1.9b)$$

So long as $\partial\Upsilon/\partial\psi \neq 0$ then the first order PDE (1.7) takes on the equivalent form

$$P \frac{\partial\Upsilon}{\partial x} + Q \frac{\partial\Upsilon}{\partial t} + R \frac{\partial\Upsilon}{\partial\psi} = 0. \quad (1.10)$$

Considering the ordered triple, (P, Q, R) , as components to a vector in (x, t, ψ) space, then equation (1.10) 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

$$\Upsilon(x, t, z) = \text{constant}. \quad (1.11)$$

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

$$\frac{d\mathbf{r}}{ds} = \hat{x} \frac{dx}{ds} + \hat{t} \frac{dt}{ds} + \hat{z} \frac{d\psi}{ds}, \quad (1.12)$$

where \hat{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

$$P = \mu \frac{dx}{ds} \quad Q = \mu \frac{dt}{ds} \quad R = \mu \frac{d\psi}{ds} \quad (1.13)$$

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

$$\frac{dx}{P} = \frac{dt}{Q} = \frac{d\psi}{R}. \quad (1.14)$$

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

1.3.2 Examples

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

$$U \frac{\partial\psi}{\partial x} + \frac{\partial\psi}{\partial t} = 0, \quad (1.15)$$

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

$$\frac{dx}{U} = \frac{dt}{1}, \quad (1.16)$$

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

$$x - Ut = \alpha \quad (1.17)$$

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

$$U \frac{\partial\psi}{\partial x} + \frac{\partial\psi}{\partial t} = R. \quad (1.18)$$

The two ODEs defining the characteristic curve are

$$\frac{dx}{U} = \frac{dt}{1} = \frac{d\psi}{R}. \quad (1.19)$$

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

$$\Gamma[x - Ut, \psi - Rt] = 0, \quad (1.20)$$

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

$$\psi = f(x - Ut) + Rt \quad (1.21)$$

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

$$x \frac{\partial \psi}{\partial x} + t \frac{\partial \psi}{\partial t} = \psi, \quad (1.22)$$

in which the ODEs determining the characteristics are given by

$$\frac{dx}{x} = \frac{dt}{t} = \frac{d\psi}{\psi}. \quad (1.23)$$

We are thus led to the relations

$$\frac{t}{x} = \alpha_1 \quad \frac{\psi}{x} = \alpha_2. \quad (1.24)$$

Hence, the general solution of the PDE (1.22) is given by

$$\Gamma(t/x, \psi/x) = 0 \Rightarrow \psi = x F(t/x) \quad (1.25)$$

for an arbitrary smooth function F .

1.4 Second order PDEs

There are occasions when we encounter second order PDEs in fluid mechanics, a general form of which in one space dimension is given by

$$A \frac{\partial^2 \psi}{\partial x^2} + B \frac{\partial^2 \psi}{\partial x \partial t} + C \frac{\partial^2 \psi}{\partial t^2} = \mathcal{F}. \quad (1.26)$$

For linear PDEs, A, B, C are arbitrary functions of space and time that are independent of ψ and \mathcal{F} 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 $\mathcal{F} = 0$).

The terms involving second derivatives in equation (1.26) 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

$$\text{PDE form} = \begin{cases} \text{hyperbolic} & B^2 - 4AC > 0 \\ \text{elliptic} & B^2 - 4AC < 0 \\ \text{parabolic} & B^2 - 4AC = 0. \end{cases} \quad (1.27)$$

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

$$\psi(x, t) = f(mx + t). \quad (1.28)$$

Plugging into the 2nd order PDE (1.26) with $\mathcal{F} = 0$ leads to

$$Am^2 + Bm + C = 0. \quad (1.29)$$

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.

1.4.1 Wave equation

The hyperbolic case with $B^2 - 4AC > 0$ contains two real characteristics. The canonical example of a hyperbolic PDE is the linear homogeneous wave equation

$$\frac{\partial^2 \psi}{\partial t^2} - U^2 \frac{\partial^2 \psi}{\partial x^2} = 0. \quad (1.30)$$

Solutions have the general form of a moving wave in both directions (the two characteristics)

$$\psi(x, t) = F(x - Ut) + G(x + Ut), \quad (1.31)$$

where F and G are differentiable functions whose form is determined by the initial conditions. Note that we can factor the differential operator into the form

$$(\partial_t - U \partial_x)(\partial_t + U \partial_x)\psi = 0. \quad (1.32)$$

Consequently, if either one of the linear first-order PDEs are satisfied

$$(\partial_t - U \partial_x)\psi = 0 \quad (1.33a)$$

$$(\partial_t + U \partial_x)\psi = 0 \quad (1.33b)$$

then ψ will satisfy the full wave equation. These first-order PDEs are the one-dimensional advection equations considered in Section 1.3, 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.

1.4.2 Heat or diffusion equation

The parabolic case $B^2 - 4AC = 0$ contains a single real characteristic. The canonical example is the heat or diffusion equation

$$\frac{\partial \psi}{\partial t} = \kappa \frac{\partial^2 \psi}{\partial x^2}, \quad (1.34)$$

where $\kappa > 0$ is the diffusivity (see Section 22.2). Solutions generally act to decay their initial condition towards zero; to reduce the amplitude of all extrema; and to correspondingly not introduce any extrema (see Section 22.3).

1.4.3 Laplace's equation and Poisson's equation

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

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0, \quad (1.35)$$

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 solutions to the diffusion equation in two space dimensions. Furthermore, as discussed in Section 3.2.2, solutions to Laplace's equation are known as *harmonic functions*. Another common elliptic equation is Poisson's equation, which results from adding a source to Laplace's equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = \mathcal{F}. \quad (1.36)$$

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 values. Strictly speaking this behaviour 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.² 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.

²A 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.

2

Cartesian tensor algebra

READER'S GUIDE TO THIS CHAPTER

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.

2.1	Introduction to tensors and tensor fields	16
2.2	Points and vectors	16
2.3	Distance and the scalar product	17
2.3.1	Distance between points	17
2.3.2	Magnitude of a vector and the scalar product	18
2.4	Vector product	19
2.4.1	Basis vector orientation and the Levi-Civita tensor	19
2.4.2	Orthogonality relations between cross products	19
2.4.3	Vector product of arbitrary vectors	20
2.4.4	Geometric interpretation of the vector product	20
2.4.5	Generalization to arbitrary vectors	21
2.5	Measuring volume	21
2.5.1	Volume defined by three vectors	21
2.5.2	Cartesian volume element for integration	22
2.5.3	n -space volumes and the Levi-Civita tensor	22
2.6	Example vector identities using the Levi-Civita tensor	23
2.6.1	Double vector product	23
2.6.2	Scalar product of two vector products	23
2.7	Transforming the coordinate representations	23
2.7.1	Inverse transformation	24
2.7.2	Orthogonal transformation	24
2.7.3	Geometric interpretation of orthogonal transformations	25
2.7.4	Transforming the coordinate representation of a vector	25
2.7.5	Form invariance of the scalar product	25
2.7.6	Transforming the coordinate representation of a second order tensor	26
2.7.7	Importance of distinguishing between tensors and matrices	26

2.1 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 give us nearly all of the formalism necessary to study general tensors in geophysical fluid mechanics.

2.2 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 shown in Figure 2.1. We write this coordinate representation as

$$\mathcal{P} \mapsto \vec{P} = \hat{x} P_1 + \hat{y} P_2 + \hat{z} P_3. \quad (2.1)$$

Vectors are denoted by an arrow. The right hand side of equation (2.1) 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, $(\hat{x}, \hat{y}, \hat{z})$. 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 sometimes find it useful to make use of alternative notations in which the position vector is written

$$\vec{P} = \mathbf{P} \quad (2.2a)$$

$$= \hat{x} P_1 + \hat{y} P_2 + \hat{z} P_3 \quad (2.2b)$$

$$= \sum_{a=1}^3 \vec{e}_a P_a \quad (2.2c)$$

$$= \vec{e}_a P_a. \quad (2.2d)$$

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

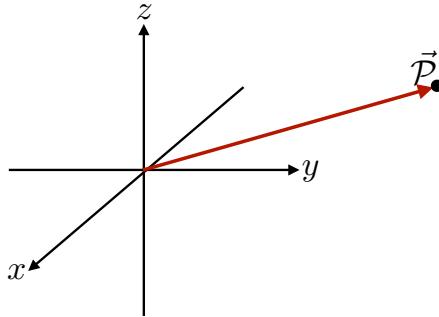


Figure 2.1: An arbitrary point in space, \mathcal{P} has an objective existence independent of our subjective choice of coordinates to describe its position. We here represent its position with respect to the origin of a right-handed Cartesian coordinate system. The Cartesian representation of the position vector is $\mathcal{P} \mapsto \vec{P} = \hat{x}P_1 + \hat{y}P_2 + \hat{z}P_3$, with the Cartesian basis vectors the normalized unit vectors $(\hat{x}, \hat{y}, \hat{z})$ 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, with each such coordinate systems having distinct coordinate representations.

The first equality introduced the boldface notation, which we commonly use for the representation of vectors in Cartesian coordinates. Notably, the boldface notation is less convenient with the general tensors of Chapter 4. Equation (2.2c) introduced a generic notation for the basis vectors

$$\vec{e}_1 = \hat{x} \quad \vec{e}_2 = \hat{y} \quad \vec{e}_3 = \hat{z}. \quad (2.3)$$

Equation (2.2d) introduced the Einstein summation convention in which repeated indices are summed over their range, thus allowing us to drop the summation symbol.

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

$$P_a = (P_1, P_2, P_3) \quad \vec{e}_a = (\vec{e}_1, \vec{e}_2, \vec{e}_3) = (\hat{x}, \hat{y}, \hat{z}). \quad (2.4)$$

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 (2.3). Furthermore, a normalized vector can change only through rotation since by definition it remains of unit norm (see Section 3.1.4). Note that for the general tensors of Chapter 4, the most convenient basis vectors are not necessarily normalized.

2.3 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.

2.3.1 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

$$\mathcal{P} \mapsto \vec{P} = \vec{e}_a P_a \quad (2.5a)$$

$$\mathcal{P} + d\mathcal{P} \mapsto \vec{P} + d\vec{x} = \vec{e}_a (P_a + dx_a). \quad (2.5b)$$

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

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

$$[\text{distance}(\mathcal{P}, \mathcal{P} + d\mathcal{P})]^2 = (\vec{P} + d\vec{x} - \vec{P}) \cdot (\vec{P} + d\vec{x} - \vec{P}) \quad (2.6a)$$

$$= (\mathbf{P} + d\mathbf{x} - \mathbf{P}) \cdot (\mathbf{P} + d\mathbf{x} - \mathbf{P}) \quad (2.6b)$$

$$= dx_a dx_b (\vec{e}_a \cdot \vec{e}_b) \quad (2.6c)$$

$$= dx_a dx_b \delta_{ab} \quad (2.6d)$$

$$= dx_a dx_a \quad (2.6e)$$

$$= (dx_1)^2 + (dx_2)^2 + (dx_3)^2. \quad (2.6f)$$

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

$$\vec{e}_a \cdot \vec{e}_b = \delta_{ab} = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.7)$$

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 5.1 we introduce alternative representations for the metric based on the use of non-Cartesian coordinates and non-Euclidean manifolds.

2.3.2 Magnitude of a vector and the scalar product

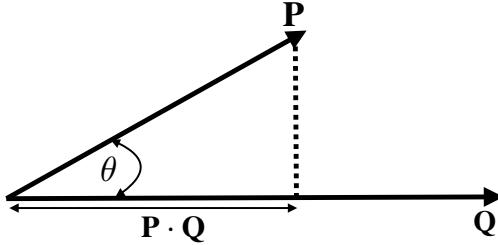


Figure 2.2: Illustrating the geometry associated with forming the scalar product between two vectors, $\mathbf{P} \cdot \mathbf{Q} = |\mathbf{P}| |\mathbf{Q}| \cos \theta$.

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

$$|\mathbf{P}|^2 = \mathbf{P} \cdot \mathbf{P} = P_a P_b (\vec{e}_a \cdot \vec{e}_b) = P_a P_a = (P_1)^2 + (P_2)^2 + (P_3)^2. \quad (2.8)$$

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

$$\mathbf{P} \cdot \mathbf{Q} = P_a Q_b (\vec{e}_a \cdot \vec{e}_b) = P_a Q_a. \quad (2.9)$$

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

$$\cos \theta \equiv \frac{\mathbf{P} \cdot \mathbf{Q}}{|\mathbf{P}| |\mathbf{Q}|} = \frac{P_a Q_a}{\sqrt{P_a P_a} \sqrt{Q_a Q_a}}. \quad (2.10)$$

We illustrate this equation in Figure 2.2. It is useful to convince oneself that this definition is consistent with $-1 \leq \cos \theta \leq 1$.

2.4 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.

2.4.1 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

$$\vec{e}_a \wedge \vec{e}_b = \epsilon_{abc} \vec{e}_c. \quad (2.11)$$

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

$$\epsilon_{123} = 1 \quad (2.12a)$$

$$\epsilon_{abc} = \begin{cases} 1, & \text{even permutation of } abc \text{ (123, 312, 231)} \\ -1, & \text{odd permutation of } abc \text{ (321, 132, 213)} \\ 0, & \text{all other } abc. \end{cases} \quad (2.12b)$$

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

$$\epsilon_{abc} = -\epsilon_{bac} = -\epsilon_{acb}, \quad (2.13)$$

whereas cycling indices (an even permutation) preserves the sign

$$\epsilon_{abc} = \epsilon_{cab} = \epsilon_{bca}. \quad (2.14)$$

2.4.2 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

$$\vec{e}_a \cdot (\vec{e}_a \wedge \vec{e}_b) = \vec{e}_a \cdot \epsilon_{abc} \vec{e}_c \quad \text{definition} \quad (2.15a)$$

$$= \epsilon_{abc} \vec{e}_a \cdot \vec{e}_c \quad \text{rearrangement} \quad (2.15b)$$

$$= \epsilon_{cba} \vec{e}_c \cdot \vec{e}_a \quad \text{relabel } a \text{ to } c \text{ and } c \text{ to } a \quad (2.15c)$$

$$= -\epsilon_{abc} \vec{e}_c \cdot \vec{e}_a \quad cba \text{ is an odd permutation of } abc \quad (2.15d)$$

$$\Rightarrow \vec{e}_a \cdot (\vec{e}_a \wedge \vec{e}_b) = 0. \quad (2.15e)$$

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 .

To fully digest step (2.15c) it can be useful to reintroduce the summation symbol so that

$$\epsilon_{abc} \vec{e}_a \cdot \vec{e}_c = \sum_{a=1}^3 \sum_{c=1}^3 \epsilon_{abc} \vec{e}_a \cdot \vec{e}_c \quad \text{summation symbols reintroduced} \quad (2.16a)$$

$$= \sum_{c=1}^3 \sum_{a=1}^3 \epsilon_{cba} \vec{e}_c \cdot \vec{e}_a \quad \text{swap } a \text{ and } c \quad (2.16b)$$

$$= \epsilon_{cba} \vec{e}_c \cdot \vec{e}_a \quad \text{reintroduce summation convention} \quad (2.16c)$$

Additionally, to digest step (2.15d) we step through the permutations

$$\epsilon_{cba} = -\epsilon_{bca} \quad \text{swap } c \text{ with } b \text{ to pick up a minus sign} \quad (2.17a)$$

$$= \epsilon_{bac} \quad \text{swap } c \text{ with } a \text{ to pick up a minus sign} \quad (2.17b)$$

$$= -\epsilon_{abc} \quad \text{swap } b \text{ with } a \text{ to pick up a minus sign.} \quad (2.17c)$$

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 2.2).

2.4.3 Vector product of arbitrary vectors

The expression (2.11) for the vector product of two basis vectors renders the vector product of arbitrary vectors.

$$\mathbf{P} \wedge \mathbf{Q} = P_a \vec{e}_a \wedge Q_b \vec{e}_b \quad (2.18a)$$

$$= P_a Q_b \vec{e}_a \wedge \vec{e}_b \quad (2.18b)$$

$$= P_a Q_b \epsilon_{abc} \vec{e}_c \quad (2.18c)$$

$$= (P_2 Q_3 - P_3 Q_2) \vec{e}_1 + (P_3 Q_1 - P_1 Q_3) \vec{e}_2 + (P_1 Q_2 - P_2 Q_1) \vec{e}_3. \quad (2.18d)$$

We can write the vector product as a determinant

$$\mathbf{P} \wedge \mathbf{Q} = \det \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \end{bmatrix}. \quad (2.19)$$

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

$$\mathbf{P} \cdot (\mathbf{P} \wedge \mathbf{Q}) = (P_d \vec{e}_d) \cdot (P_a Q_b \epsilon_{abc} \vec{e}_c) \quad (2.20a)$$

$$= P_c P_a Q_b \epsilon_{abc} \quad (2.20b)$$

$$= 0, \quad (2.20c)$$

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

2.4.4 Geometric interpretation of the vector product

The expression (2.18d) leads to the identity

$$|\mathbf{P} \wedge \mathbf{Q}|^2 = |\mathbf{P}|^2 |\mathbf{Q}|^2 - (\mathbf{P} \cdot \mathbf{Q})^2 \quad (2.21a)$$

$$= |\mathbf{P}|^2 |\mathbf{Q}|^2 (1 - \cos^2 \theta), \quad (2.21b)$$

$$= |\mathbf{P}|^2 |\mathbf{Q}|^2 \sin^2 \theta, \quad (2.21c)$$

where we used the scalar product expression (2.10) 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

$$\text{area}(\mathbf{P}, \mathbf{Q}) = |\mathbf{P}| |\mathbf{Q}| \sin \theta = |\mathbf{P} \wedge \mathbf{Q}|. \quad (2.22)$$

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

$$\mathbf{P} \wedge \mathbf{Q} = \hat{\mathbf{n}} \text{area}(\mathbf{P}, \mathbf{Q}) = \hat{\mathbf{n}} |\mathbf{P}| |\mathbf{Q}| \sin \theta, \quad (2.23)$$

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 2.3.

2.4.5 Generalization to arbitrary vectors

Thus far our 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.

2.5 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.

2.5.1 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.

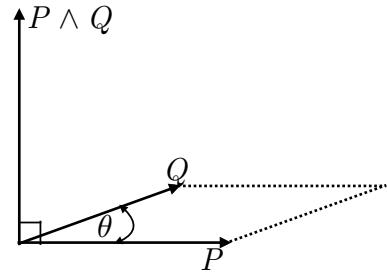


Figure 2.3: The magnitude for the vector product between two vectors is given by the product of their magnitudes and the sine of the angle between them, $|\mathbf{P} \wedge \mathbf{Q}| = |\mathbf{P}| |\mathbf{Q}| \sin \theta$. This magnitude equals to the area of the parallelogram subtended by the two vectors. The vector product is directed perpendicular to the plane determined by the two vectors and oriented according to the right hand rule.

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

$$(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R} = (P_a \vec{e}_a \wedge Q_b \vec{e}_b) \cdot R_d \vec{e}_d \quad (2.24a)$$

$$= P_a Q_b (\vec{e}_a \wedge \vec{e}_b) \cdot R_d \vec{e}_d \quad (2.24b)$$

$$= P_a Q_b (\epsilon_{abc} \vec{e}_c) \cdot \vec{e}_d R_d \quad (2.24c)$$

$$= P_a Q_b \epsilon_{abc} (\vec{e}_c \cdot \vec{e}_d) R_d \quad (2.24d)$$

$$= P_a Q_b \epsilon_{abc} \delta_{cd} R_d \quad (2.24e)$$

$$= P_a Q_b \epsilon_{abc} R_c \quad (2.24f)$$

$$= R_c P_a Q_b \epsilon_{abc} \quad (2.24g)$$

$$= R_a P_b Q_c \epsilon_{bca} \quad (2.24h)$$

$$= R_a P_b Q_c \epsilon_{abc} \quad (2.24i)$$

$$= (\mathbf{R} \wedge \mathbf{P}) \cdot \mathbf{Q}. \quad (2.24j)$$

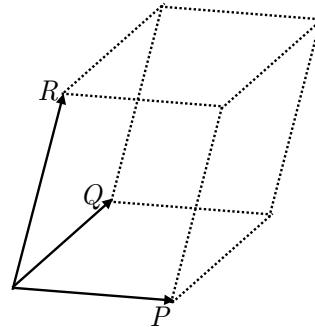


Figure 2.4: Three linearly independent position vectors determine a volume given by $|(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}| = |(\mathbf{R} \wedge \mathbf{P}) \cdot \mathbf{Q}| = |(\mathbf{Q} \wedge \mathbf{R}) \cdot \mathbf{P}|$

We thus have the geometric result illustrated in Figure 2.4

$$\text{volume}(\mathbf{P}, \mathbf{Q}, \mathbf{R}) = |(\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R}| = |(\mathbf{R} \wedge \mathbf{P}) \cdot \mathbf{Q}| = |(\mathbf{Q} \wedge \mathbf{R}) \cdot \mathbf{P}|. \quad (2.25)$$

2.5.2 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

$$\mathbf{P} = \hat{x} dx \quad \mathbf{Q} = \hat{y} dy \quad \mathbf{R} = \hat{z} dz, \quad (2.26)$$

in which case the volume element is

$$dV = (\mathbf{P} \wedge \mathbf{Q}) \cdot \mathbf{R} = dx dy dz (\hat{x} \wedge \hat{y}) \cdot \hat{z} = dx dy dz. \quad (2.27)$$

This expression for the volume element could have been written down without the formalism of a vector triple product. However, in Chapter 4 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.

2.5.3 n -space volumes and the Levi-Civita tensor

We combine the geometric specification of the vector product as a means to measure area, (2.23), with the algebraic specification (2.18d) by writing

$$\text{2-volume} = \epsilon(\mathbf{P}, \mathbf{Q}) = \epsilon_{ab} P_a Q_b = \det \begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix}, \quad (2.28)$$

where ϵ_{ab} is the totally anti-symmetric 2×2 matrix with Cartesian components

$$\epsilon_{ab} = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{21} & \epsilon_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (2.29)$$

In words, the first equality in equation (2.28) 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

$$\text{3-volume} = \epsilon(\mathbf{P}, \mathbf{Q}, \mathbf{R}) = \epsilon_{abc} P_a Q_b R_c = \det \begin{bmatrix} P_1 & Q_1 & R_1 \\ P_2 & Q_2 & R_2 \\ P_3 & Q_3 & R_3 \end{bmatrix}. \quad (2.30)$$

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

$$\text{surface area} = \epsilon(\ , \mathbf{Q}, \mathbf{R}). \quad (2.31)$$

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

2.6 Example vector identities using the Levi-Civita tensor

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

2.6.1 Double vector product

Consider the double vector product

$$\mathbf{P} \wedge (\mathbf{Q} \wedge \mathbf{R}) = P_a Q_b R_c \vec{e}_a \wedge (\vec{e}_b \wedge \vec{e}_c) \quad (2.32a)$$

$$= P_a Q_b R_c \vec{e}_a \wedge (\epsilon_{bcd} \vec{e}_d) \quad (2.32b)$$

$$= P_a Q_b R_c \epsilon_{bcd} \epsilon_{ade} \vec{e}_e \quad (2.32c)$$

$$= -P_a Q_b R_c \epsilon_{bcd} \epsilon_{aed} \vec{e}_e. \quad (2.32d)$$

Through explicit substitution, we can verify that the product $\epsilon_{bcd} \epsilon_{aed}$ equals to

$$\epsilon_{bcd} \epsilon_{aed} = \delta_{ba} \delta_{ce} - \delta_{be} \delta_{ca}. \quad (2.33)$$

This identity then leads to

$$\epsilon_{bcd} \epsilon_{aed} \vec{e}_e = \delta_{ba} \vec{e}_c - \delta_{ca} \vec{e}_b \quad (2.34)$$

so that

$$\mathbf{P} \wedge (\mathbf{Q} \wedge \mathbf{R}) = -P_a Q_b R_c (\delta_{ba} \vec{e}_c - \delta_{ca} \vec{e}_b) \quad (2.35a)$$

$$= -(\mathbf{P} \cdot \mathbf{Q}) \mathbf{R} + (\mathbf{P} \cdot \mathbf{R}) \mathbf{Q}. \quad (2.35b)$$

2.6.2 Scalar product of two vector products

We make further use of the Levi-Civita identity (2.33) to write

$$(\mathbf{P} \wedge \mathbf{Q}) \cdot (\mathbf{R} \wedge \mathbf{S}) = (\epsilon_{abc} p_a q_b) (\epsilon_{dec} R_d s_e) \quad (2.36a)$$

$$= p_a q_b R_d s_e \epsilon_{abc} \epsilon_{dec} \quad (2.36b)$$

$$= p_a q_b R_d s_e (\delta_{ad} \delta_{be} - \delta_{ae} \delta_{bd}) \quad (2.36c)$$

$$= (\mathbf{P} \cdot \mathbf{R}) (\mathbf{Q} \cdot \mathbf{S}) - (\mathbf{P} \cdot \mathbf{S}) (\mathbf{Q} \cdot \mathbf{R}). \quad (2.36d)$$

2.7 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

$$\vec{e}_{\bar{a}} = \mathcal{R}_{\bar{a}b} \vec{e}_b. \quad (2.37)$$

This expression introduced components to the transformation matrix

$$\mathcal{R}_{\bar{a}\bar{b}} = \begin{bmatrix} \mathcal{R}_{\bar{1}\bar{1}} & \mathcal{R}_{\bar{1}\bar{2}} & \mathcal{R}_{\bar{1}\bar{3}} \\ \mathcal{R}_{\bar{2}\bar{1}} & \mathcal{R}_{\bar{2}\bar{2}} & \mathcal{R}_{\bar{2}\bar{3}} \\ \mathcal{R}_{\bar{3}\bar{1}} & \mathcal{R}_{\bar{3}\bar{2}} & \mathcal{R}_{\bar{3}\bar{3}} \end{bmatrix}. \quad (2.38)$$

In Cartesian tensor analysis the transformation is assumed to be independent of space (it is a function of space and time for the general tensors considered in Chapter 5). 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.

2.7.1 Inverse transformation

Assuming the transformation is invertible leads to the inverse transformation

$$\vec{e}_a = (\mathcal{R}^{-1})_{a\bar{b}} \vec{e}_{\bar{b}}. \quad (2.39)$$

As a self-consistency check we combine this relation with equation (2.37) thus rendering

$$\vec{e}_a = (\mathcal{R}^{-1})_{a\bar{b}} \vec{e}_{\bar{b}} = (\mathcal{R})_{a\bar{b}}^{-1} (\mathcal{R}_{\bar{b}c} \vec{e}_c). \quad (2.40)$$

This relation holds since

$$(\mathcal{R}^{-1})_{a\bar{b}} \mathcal{R}_{\bar{b}c} = \delta_{ac}, \quad (2.41)$$

or as a matrix identity

$$\mathcal{R}^{-1} \mathcal{R} = I. \quad (2.42)$$

2.7.2 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

$$\delta_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} \quad (2.43a)$$

$$= \mathcal{R}_{\bar{a}a} \vec{e}_a \cdot \mathcal{R}_{\bar{b}b} \vec{e}_b \quad (2.43b)$$

$$= \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}b} \vec{e}_a \cdot \vec{e}_b \quad (2.43c)$$

$$= \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}b} \delta_{ab} \quad (2.43d)$$

$$= \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}a} \quad (2.43e)$$

$$= \mathcal{R}_{\bar{a}a} (\mathcal{R}^T)_{\bar{a}\bar{b}}, \quad (2.43f)$$

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

$$(\mathcal{R}^T)_{a\bar{b}} = \mathcal{R}_{\bar{b}a}. \quad (2.44)$$

Written as a matrix equation we see that

$$\mathcal{R} \mathcal{R}^T = I. \quad (2.45)$$

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

$$\mathcal{R}^{-1} = \mathcal{R}^T. \quad (2.46)$$

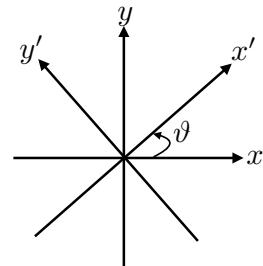


Figure 2.5: Counter-clockwise rotation of horizontal Cartesian axes through an angle ϑ .

2.7.3 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 2.5 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*

$$\mathcal{R}_{\bar{a}a} = \begin{bmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{bmatrix} = \begin{bmatrix} \cos \vartheta & \cos(\pi/2 - \vartheta) \\ \cos(\pi/2 + \vartheta) & \cos \vartheta \end{bmatrix} = \begin{bmatrix} \vec{e}_1 \cdot \vec{e}_1 & \vec{e}_1 \cdot \vec{e}_2 \\ \vec{e}_2 \cdot \vec{e}_1 & \vec{e}_2 \cdot \vec{e}_2 \end{bmatrix}. \quad (2.47)$$

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

$$\mathcal{R}_{\bar{a}a} = \begin{bmatrix} \vec{e}_1 \cdot \vec{e}_1 & \vec{e}_1 \cdot \vec{e}_2 & \vec{e}_1 \cdot \vec{e}_3 \\ \vec{e}_2 \cdot \vec{e}_1 & \vec{e}_2 \cdot \vec{e}_2 & \vec{e}_2 \cdot \vec{e}_3 \\ \vec{e}_3 \cdot \vec{e}_1 & \vec{e}_3 \cdot \vec{e}_2 & \vec{e}_3 \cdot \vec{e}_3 \end{bmatrix}. \quad (2.48)$$

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

2.7.4 Transforming the coordinate representation of a vector

We introduced the transformation (2.39) 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

$$\mathbf{P} = P_a \vec{e}_a \quad (2.49a)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} \vec{e}_{\bar{a}} \quad (2.49b)$$

$$\equiv P_{\bar{a}} \vec{e}_{\bar{a}}, \quad (2.49c)$$

where we defined the transformation of the vector components

$$P_{\bar{a}} = P_a (\mathcal{R}^T)_{a\bar{a}} = \mathcal{R}_{\bar{a}a} P_a. \quad (2.50)$$

2.7.5 Form invariance of the scalar product

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

$$\mathbf{P} \cdot \mathbf{Q} = P_a Q_a \quad (2.51)$$

is form invariant

$$\mathbf{P} \cdot \mathbf{Q} = P_a \vec{e}_a \cdot Q_b \vec{e}_b \quad (2.52a)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} Q_b (\mathcal{R}^T)_{b\bar{b}} (\vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}}) \quad (2.52b)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} Q_b (\mathcal{R}^T)_{b\bar{b}} \delta_{\bar{a}\bar{b}} \quad (2.52c)$$

$$= P_a (\mathcal{R}^T)_{a\bar{a}} Q_b (\mathcal{R}^T)_{b\bar{a}} \quad (2.52d)$$

$$= P_{\bar{a}} Q_{\bar{a}}. \quad (2.52e)$$

We return to form invariance in Section 6.1, where it is referred to as *general covariance* in the context of general tensor analysis.

2.7.6 Transforming the coordinate representation of a second order tensor

The diffusion tensor is introduced in Chapter 22 and the stress tensor in Chapter 26. These tensors are second order, with second order tensors having a coordinate representation given by

$$\mathcal{T} = T_{ab} \vec{e}_a \vec{e}_b, \quad (2.53)$$

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

$$\mathcal{T} = T_{ab} \vec{e}_a \vec{e}_b \quad (2.54a)$$

$$= T_{ab} (\mathcal{R}^T)_{a\bar{a}} \vec{e}_{\bar{a}} (\mathcal{R}^T)_{b\bar{b}} \vec{e}_{\bar{b}} \quad (2.54b)$$

$$= T_{ab} (\mathcal{R}^T)_{a\bar{a}} (\mathcal{R}^T)_{b\bar{b}} \vec{e}_{\bar{a}} \vec{e}_{\bar{b}} \quad (2.54c)$$

$$\equiv T_{\bar{a}\bar{b}} \vec{e}_{\bar{a}} \vec{e}_{\bar{b}}. \quad (2.54d)$$

The final equality introduced the transformed tensor components

$$T_{\bar{a}\bar{b}} = T_{ab} (\mathcal{R}^T)_{a\bar{a}} (\mathcal{R}^T)_{b\bar{b}} = T_{ab} \mathcal{R}_{\bar{a}a} \mathcal{R}_{\bar{b}b}. \quad (2.55)$$

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

2.7.7 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 2.1) 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.

2.8 Exercises

EXERCISE 2.1: PRODUCT OF SYMMETRIC MATRICES

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

EXERCISE 2.2: 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: $A_{mn} S_{nm} = 0$.

3

Cartesian tensor calculus

READER'S GUIDE TO THIS CHAPTER

In this chapter, we build from the Cartesian tensor algebra of Chapter 2 to develop elements of Cartesian tensor calculus. The material in this chapter is used throughout this book. The reader may refer to chapter 6 of *Hildebrand (1976)* for a more thorough treatment.

3.1	The gradient of a scalar field	28
3.1.1	Direction of steepest ascent	28
3.1.2	Tangent to an isosurface	28
3.1.3	Normal to an isosurface	29
3.1.4	Unit vectors change only by rotation	29
3.1.5	Showing that $\delta\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 0$	30
3.2	The divergence of a vector field	30
3.2.1	Divergence of a scalar field times a vector field	31
3.2.2	Laplacian of a scalar field	31
3.3	The curl of a vector field	31
3.3.1	Computing the curl	31
3.3.2	Curl-free vector fields	32
3.3.3	Curl-free and divergence-free fields	33
3.3.4	Identities involving the curl	33
3.4	Path integral of a scalar function	35
3.5	Path integral of a vector function	36
3.5.1	Circulation	37
3.5.2	Circulation example	37
3.5.3	Fundamental theorem of calculus	37
3.6	Stokes' theorem	38
3.6.1	Statement of Stokes' theorem	38
3.6.2	Stokes' theorem for a rectangular region	38
3.7	Divergence theorem (Gauss's law)	39
3.7.1	An example rectangular volume	39
3.7.2	Divergence theorem for scalar fields	40
3.7.3	First and second form of Green's theorem	40
3.8	Exercises	41

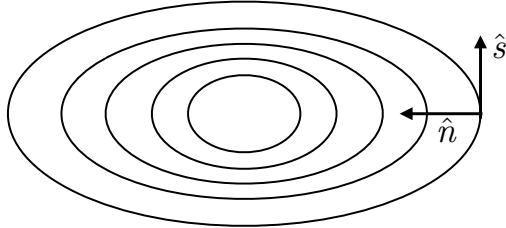


Figure 3.1: Contours of a scalar field $\psi(x, y)$, with values increasing toward the center. At any point in space, $\nabla\psi$ points in the direction of steepest increase (ascent) and orients the normal vector $\hat{n} = |\nabla\psi|^{-1} \nabla\psi$. The unit tangent vector, \hat{s} , points in a direction tangent to a ψ isosurface so that it follows the surface of constant ψ and it orthogonal to the tangent: $\hat{n} \cdot \hat{s} = 0$.

3.1 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

$$\psi(\mathbf{x} + d\mathbf{x}) = \psi(\mathbf{x}) + \frac{\partial\psi}{\partial x_1} dx_1 + \frac{\partial\psi}{\partial x_2} dx_2 + \frac{\partial\psi}{\partial x_3} dx_3 + \mathcal{O}(d\mathbf{x} \cdot d\mathbf{x}) \quad (3.1a)$$

$$\approx [1 + dx_a \partial_a] \psi(\mathbf{x}), \quad (3.1b)$$

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

$$\partial_a = \frac{\partial}{\partial x_a}. \quad (3.2)$$

We can introduce the gradient operator according to

$$\nabla = \vec{e}_a \partial_a = \hat{x} \partial_x + \hat{y} \partial_y + \hat{z} \partial_z \quad (3.3)$$

in which case

$$\psi(\mathbf{x} + d\mathbf{x}) \approx (1 + d\mathbf{x} \cdot \nabla) \psi(\mathbf{x}). \quad (3.4)$$

3.1.1 Direction of steepest ascent

Using the approximate relation (3.4), and the geometric expression (2.10) for the scalar product, renders

$$\psi(\mathbf{x} + d\mathbf{x}) - \psi(\mathbf{x}) \approx |d\mathbf{x}| |\nabla\psi| \cos \theta, \quad (3.5)$$

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 3.1). The opposite direction is that of *steepest descent*, where $\theta = \pi$.

3.1.2 Tangent to an isosurface

Consider a family of isosurfaces defined by points satisfying

$$\psi(\mathbf{x}) = \text{constant} \quad (3.6)$$

Figure 3.1 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

$$\psi(\mathbf{x} + \hat{\mathbf{s}} ds) - \psi(\mathbf{x}) = 0, \quad (3.7)$$

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

$$\hat{\mathbf{s}} \cdot \nabla \psi = \frac{\partial \psi}{\partial s} = 0. \quad (3.8)$$

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.

3.1.3 Normal to an isosurface

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

$$\hat{\mathbf{n}} = |\nabla \psi|^{-1} \nabla \psi. \quad (3.9)$$

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

$$\hat{\mathbf{n}} \cdot \nabla \psi = \frac{\partial \psi}{\partial n} = |\nabla \psi|. \quad (3.10)$$

For the spherically symmetric example,

$$\hat{\mathbf{n}} = \frac{\mathbf{x}}{|\mathbf{x}|} = \hat{\mathbf{r}}, \quad (3.11)$$

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

$$\hat{\mathbf{n}} \cdot \nabla \psi = \frac{\partial \psi}{\partial r} \quad \text{spherically symmetric } \psi. \quad (3.12)$$

3.1.4 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

$$\hat{\mathbf{m}} \cdot \hat{\mathbf{m}} = 1. \quad (3.13)$$

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 3.2). We see this property through considering an arbitrary change, symbolized by δ , in which

$$0 = \delta(1) = \delta(\hat{\mathbf{m}} \cdot \hat{\mathbf{m}}) = 2 \hat{\mathbf{m}} \cdot \delta \hat{\mathbf{m}}. \quad (3.14)$$

In Section 10.5, we formally show that the constraint

$$\delta \hat{\mathbf{m}} \cdot \hat{\mathbf{m}} = 0 \quad (3.15)$$

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

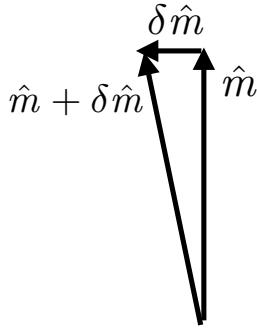


Figure 3.2: The infinitesimal change to a unit vector is itself perpendicular to the unit vector: $\delta\hat{m} \cdot \hat{m} = 0$.

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

As an illustration of the constraint (3.15), let us verify that it holds for the special case of a unit normal vector (3.9) defined according to surfaces of constant scalar field

$$\hat{n} = |\nabla\psi|^{-1} \nabla\psi. \quad (3.16)$$

The proof follows first by writing

$$\delta\hat{n} = |\nabla\psi|^{-1} [\delta(\nabla\psi) - \hat{n} \delta|\nabla\psi|], \quad (3.17)$$

so that

$$|\nabla\psi| \hat{n} \cdot \delta\hat{n} = \hat{n} \cdot \delta(\nabla\psi) - \delta|\nabla\psi| \quad (3.18a)$$

$$= \frac{\nabla\psi \cdot \delta(\nabla\psi)}{|\nabla\psi|} - \delta|\nabla\psi| \quad (3.18b)$$

$$= \frac{|\nabla\psi| \delta|\nabla\psi|}{|\nabla\psi|} - \delta|\nabla\psi| \quad (3.18c)$$

$$= 0. \quad (3.18d)$$

The last step made use of the identity

$$\delta(|\nabla\psi|) = \delta(\sqrt{\nabla\psi \cdot \nabla\psi}) = \frac{1}{2\sqrt{\nabla\psi \cdot \nabla\psi}} \delta(\nabla\psi \cdot \nabla\psi) = \frac{\nabla\psi \cdot \delta(\nabla\psi)}{|\nabla\psi|}. \quad (3.19)$$

3.2 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

$$\text{div}(\mathbf{F}) = \nabla \cdot \mathbf{F} = \partial_a F_a \begin{cases} > 0 \Rightarrow \text{diverging vector field,} \\ < 0 \Rightarrow \text{converging vector field.} \end{cases} \quad (3.20)$$

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 3.3). 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 15).

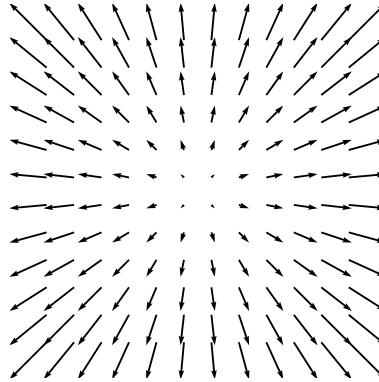


Figure 3.3: A vector field with a non-zero horizontal divergence. With $\mathbf{F} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$ the field diverges from the origin with a spatially constant divergence $\nabla \cdot \mathbf{F} = 2$.

3.2.1 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

$$\nabla \cdot (\phi \mathbf{F}) = \partial_a(\phi F_a) \quad (3.21a)$$

$$= \partial_a(\phi) F_a + \phi \partial_a F_a \quad (3.21b)$$

$$= \mathbf{F} \cdot \nabla \phi + \phi \nabla \cdot \mathbf{F}. \quad (3.21c)$$

3.2.2 Laplacian of a scalar field

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

$$\nabla^2 \psi = \nabla \cdot \nabla \psi. \quad (3.22)$$

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

$$\nabla^2 \psi = 0 \quad \text{harmonic function.} \quad (3.23)$$

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.

3.3 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 37).

3.3.1 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 2.4, the curl is specified by both a

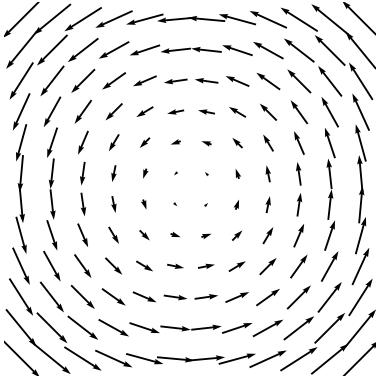


Figure 3.4: A horizontal vector field with a constant curl and zero divergence: $\mathbf{F} = -y \hat{\mathbf{x}} + x \hat{\mathbf{y}}$, $\nabla \wedge \mathbf{F} = 2 \hat{\mathbf{z}}$.

magnitude and a direction

$$\text{curl}(\mathbf{F}) = \nabla \wedge \mathbf{F} \quad (3.24a)$$

$$= \vec{e}_a \partial_a \wedge \vec{e}_b F_b \quad (3.24b)$$

$$= \vec{e}_a \wedge \vec{e}_b \partial_a F_b + \vec{e}_a F_b \wedge \partial_b \vec{e}_a \quad (3.24c)$$

$$= \epsilon_{abc} \vec{e}_c \partial_a F_b \quad (3.24d)$$

$$= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \hat{\mathbf{x}} + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) \hat{\mathbf{y}} + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \hat{\mathbf{z}}. \quad (3.24e)$$

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 (2.11) for the cross product of basis vectors. We can express the curl from equation (3.24e) as a determinant

$$\nabla \wedge \mathbf{F} = \det \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ F_1 & F_2 & F_3 \end{bmatrix}. \quad (3.25)$$

The horizontal vector field $\mathbf{F} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}}$ shown in Figure 3.3 has zero curl yet non-zero divergence. Figure 3.4 shows another vector field, $\mathbf{F} = -y \hat{\mathbf{x}} + x \hat{\mathbf{y}}$, with zero divergence yet nonzero curl $\nabla \wedge \mathbf{F} = 2 \hat{\mathbf{z}}$. As seen in Section 37.1.2, this vector field corresponds to the velocity due to solid-body motion on a rotating planet, with its curl determining the planetary vorticity.

3.3.2 Curl-free vector fields

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

$$\nabla \wedge \mathbf{F} = 0. \quad (3.26)$$

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 35.1). We illustrate a curl-free vector field in Figure 3.5, where the scalar potential is given by $\psi = \sin x \sin y$.

¹Basis vectors corresponding to non-Cartesian coordinates are spatially dependent (see Chapters 4 and 6), thus making this step invalid for general tensors. We will find a “fix” for this step in Section 6.9 by defining the *covariant curl operator*.

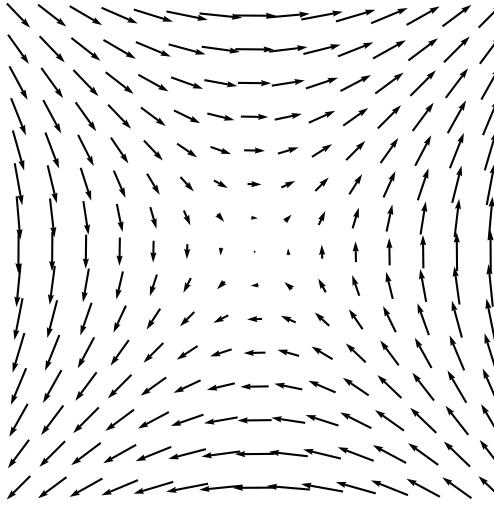


Figure 3.5: A horizontal vector field with a zero curl, where $\mathbf{F} = -\nabla\psi$ with $\psi = \sin x \sin y$.

The curl of a gradient vanishes

$$\nabla \wedge \nabla\psi = \vec{e}_a \partial_a \wedge \vec{e}_b \partial_b \psi \quad (3.27a)$$

$$= (\vec{e}_a \wedge \vec{e}_b) \partial_a \partial_b \psi \quad (3.27b)$$

$$= 0, \quad (3.27c)$$

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

$$\mathbf{F} = -\nabla\psi \quad \text{scalar potential.} \quad (3.28)$$

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 11.1 and Chapter 32).

3.3.3 Curl-free and divergence-free fields

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

$$\nabla \wedge \mathbf{F} = 0 \Rightarrow \mathbf{F} = -\nabla\psi. \quad (3.29)$$

The divergence-free property means that ψ is a harmonic function (Section 3.2.2)

$$\nabla \cdot \nabla\psi = \nabla^2\psi = 0. \quad (3.30)$$

The velocity field arising from a linear non-rotating gravity wave (Section 35.1) in a Boussinesq fluid (Section 29.1) maintains zero vorticity and zero divergence. Furthermore, curl-free and divergence-free velocity fields are commonly encountered in aerodynamics.

3.3.4 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 4, 5, and 6, 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

$$\nabla \cdot (\nabla \times \mathbf{F}) = \partial_a(\epsilon_{abc} \partial_b F_c) \quad (3.31a)$$

$$= \epsilon_{abc} \partial_a \partial_b F_c \quad (3.31b)$$

$$= 0. \quad (3.31c)$$

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

$$\nabla \cdot (\mathbf{F} \wedge \mathbf{E}) = \mathbf{E} \cdot (\nabla \wedge \mathbf{F}) - \mathbf{F} \cdot (\nabla \wedge \mathbf{E}) \quad (3.32)$$

through the following manipulations

$$\nabla \cdot (\mathbf{F} \wedge \mathbf{E}) = \vec{e}_a \partial_a (F_b \vec{e}_b \wedge E_c \vec{e}_c) \quad (3.33a)$$

$$= \vec{e}_a \cdot (\vec{e}_b \wedge \vec{e}_c) \partial_a (F_b E_c) \quad (3.33b)$$

$$= \epsilon_{abc} \partial_a (F_b E_c) \quad (3.33c)$$

$$= F_b \epsilon_{abc} \partial_a E_c + E_c \epsilon_{abc} \partial_a F_b \quad (3.33d)$$

$$= -\mathbf{F} \cdot (\nabla \wedge \mathbf{E}) + \mathbf{E} \cdot (\nabla \wedge \mathbf{F}). \quad (3.33e)$$

Curl of a scalar times a vector

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

$$\nabla \wedge (\psi \mathbf{F}) = \vec{e}_a \partial_a \wedge \psi \vec{e}_b F_b \quad (3.34a)$$

$$= (\vec{e}_a \wedge \vec{e}_b) \partial_a (\psi F_b) \quad (3.34b)$$

$$= \epsilon_{abc} \vec{e}_c (\psi \partial_a F_b + F_b \partial_a \psi) \quad (3.34c)$$

$$= \psi \nabla \wedge \mathbf{F} + \nabla \psi \wedge \mathbf{F}. \quad (3.34d)$$

Curl of a cross product

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

$$\nabla \wedge (\mathbf{F} \wedge \mathbf{E}) = \vec{e}_a \partial_a \wedge (\vec{e}_b F_b \wedge \vec{e}_c E_c) \quad (3.35a)$$

$$= \vec{e}_a \wedge (\vec{e}_b \wedge \vec{e}_c) \partial_a (F_b E_c) \quad (3.35b)$$

$$= \vec{e}_a \wedge (\epsilon_{bcd} \vec{e}_d) \partial_a (F_b E_c) \quad (3.35c)$$

$$= \epsilon_{ade} \epsilon_{bcd} \vec{e}_e \partial_a (F_b E_c) \quad (3.35d)$$

$$= -\epsilon_{aed} \epsilon_{bcd} \partial_a (F_b E_c) \quad (3.35e)$$

$$= -(\delta_{ab} \delta_{ec} - \delta_{ac} \delta_{eb}) \vec{e}_e \partial_a (F_b E_c) \quad (3.35f)$$

$$= (-\delta_{ba} \vec{e}_c + \delta_{ca} \vec{e}_b) \partial_a (F_b E_c) \quad (3.35g)$$

$$= \mathbf{F} (\nabla \cdot \mathbf{E}) + (\mathbf{E} \cdot \nabla) \mathbf{F} - \mathbf{E} (\nabla \cdot \mathbf{F}) - (\mathbf{F} \cdot \nabla) \mathbf{E}. \quad (3.35h)$$

Curl of a curl

A special case of the identity (3.35h) allows us to write the curl of a curl as

$$\nabla \wedge (\nabla \wedge \mathbf{F}) = \vec{e}_a \wedge (\vec{e}_b \wedge \vec{e}_c) \partial_a \partial_b F_c \quad (3.36a)$$

$$= (-\delta_{ba} \vec{e}_c + \delta_{ca} \vec{e}_b) \partial_a \partial_b F_c \quad (3.36b)$$

$$= \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}. \quad (3.36c)$$

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 37.7.1). Here, we aim to show that

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = \boldsymbol{\omega} \wedge \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2, \quad (3.37)$$

where

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (3.38)$$

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

$$\boldsymbol{\omega} \wedge \mathbf{v} = (\nabla \wedge \mathbf{v}) \wedge \mathbf{v} \quad \text{insert } \boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (3.39a)$$

$$= (\vec{e}_a \partial_a \wedge \vec{e}_b v_b) \wedge \vec{e}_c v_c \quad \text{Cartesian representation of } \mathbf{v} \text{ and } \nabla \quad (3.39b)$$

$$= (\vec{e}_a \wedge \vec{e}_b) \wedge \vec{e}_c (\partial_a v_b) v_c \quad \text{rearrange} \quad (3.39c)$$

$$= \epsilon_{abd} (\vec{e}_d \wedge \vec{e}_c) (\partial_a v_b) v_c \quad \text{first cross product expanded} \quad (3.39d)$$

$$= \epsilon_{abd} \epsilon_{dce} \vec{e}_e (\partial_a v_b) v_c \quad \text{second cross product expanded} \quad (3.39e)$$

$$= \epsilon_{abd} \epsilon_{ced} \vec{e}_e (\partial_a v_b) v_c \quad \text{arrange indices to prepare for next step} \quad (3.39f)$$

$$= (\delta_{ac} \delta_{be} - \delta_{ae} \delta_{bc}) \vec{e}_e (\partial_a v_b) v_c \quad \text{use identity (2.33)} \quad (3.39g)$$

$$= \vec{e}_a v_c \partial_c v_a - \vec{e}_a v_c \partial_a v_c \quad \text{contract the Kronecker deltas} \quad (3.39h)$$

$$= \vec{e}_a [(\mathbf{v} \cdot \nabla) v_a - \partial_a \mathbf{v}^2/2] \quad \text{re-express as Cartesian tensor} \quad (3.39i)$$

$$= (\mathbf{v} \cdot \nabla) \mathbf{v} - \nabla[\mathbf{v} \cdot \mathbf{v}/2] \quad \text{rearrange.} \quad (3.39j)$$

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

3.4 Path integral of a scalar function

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

$$\mathcal{I} = \int_{\mathcal{C}} \psi(\varphi) d\varphi. \quad (3.40)$$

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

$$\mathcal{C} \mapsto \mathbf{x}(\varphi) = \hat{\mathbf{x}} x(\varphi) + \hat{\mathbf{y}} y(\varphi) + \hat{\mathbf{z}} z(\varphi) \quad (3.41)$$

so that the path integral can be written

$$\mathcal{I} = \int_{\mathcal{C}} \psi(\varphi) d\varphi = \int_{\mathcal{C}} \psi[\mathbf{x}(\varphi)] d\varphi. \quad (3.42)$$

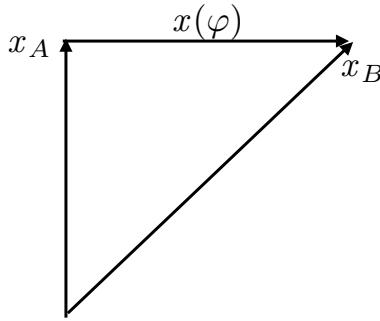


Figure 3.6: A linear path, $\mathbf{x}(\varphi)$ extending from \mathbf{x}_A to \mathbf{x}_B that is parameterized by a non-dimensional parameter $\varphi \in [0, 1]$ via $\mathbf{x}(\varphi) = \mathbf{x}_A + (\mathbf{x}_B - \mathbf{x}_A)\varphi$. Alternatively it can be parameterized by the arc-length along the path via $\mathbf{x}(s) = \mathbf{x}_A + \hat{\mathbf{s}} s$ with $s \in [0, L]$.

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

$$ds = \sqrt{d\mathbf{x} \cdot d\mathbf{x}} = ds \sqrt{\frac{d\mathbf{x}}{ds} \cdot \frac{d\mathbf{x}}{ds}} = ds |\mathbf{x}'(s)|, \quad (3.43)$$

so that the path integral takes the form

$$\mathcal{I} = \int_C \psi[\mathbf{x}(s)] |\mathbf{x}'(s)| ds. \quad (3.44)$$

Linear path example

As a specific example, consider a line between two points, \mathbf{x}_A and \mathbf{x}_B , as in Figure 3.6. We can parameterize the line using a dimensionless parameter φ according to

$$\mathbf{x}(\varphi) = \mathbf{x}_A + (\mathbf{x}_B - \mathbf{x}_A)\varphi \quad \varphi \in [0, 1]. \quad (3.45)$$

Alternatively, we can parameterize using the arc length

$$\mathbf{x}(s) = \mathbf{x}_A + \hat{\mathbf{s}} s \quad s \in [0, L], \quad (3.46)$$

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

$$\hat{\mathbf{s}} = \frac{\mathbf{x}'(s)}{|\mathbf{x}'(s)|} = \frac{\mathbf{x}_B - \mathbf{x}_A}{|\mathbf{x}_B - \mathbf{x}_A|}. \quad (3.47)$$

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

3.5 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

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C \mathbf{F} \cdot \frac{d\mathbf{x}}{ds} ds \quad (3.48)$$

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

3.5.1 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

$$\text{circulation of vector field} = \oint_C \mathbf{F} \cdot d\mathbf{x}. \quad (3.49)$$

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

3.5.2 Circulation example

Consider the vector field $\mathbf{F} = -y \hat{\mathbf{x}} + x \hat{\mathbf{y}}$ shown in Figure 3.4. 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 (3.48) thus takes the form

$$\mathbf{F} \cdot \frac{d\mathbf{x}(\varphi)}{d\varphi} = r(y \sin \varphi + x \cos \varphi) = r^2. \quad (3.50)$$

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

$$\oint_C \mathbf{F} \cdot d\mathbf{x} = 2\pi r^2 = 2 \text{ area}. \quad (3.51)$$

This result has application for geophysical fluids when computing the vorticity induced by the rotating planet (see Section 37.9.3).

3.5.3 Fundamental theorem of calculus

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

$$\psi(\mathbf{x}_A) - \psi(\mathbf{x}_B) = \int_{\mathbf{x}_A}^{\mathbf{x}_B} d\psi = \int_{\mathbf{x}_A}^{\mathbf{x}_B} \nabla\psi \cdot d\mathbf{x}. \quad (3.52)$$

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

$$\oint_C d\psi = \oint_C \nabla\psi \cdot d\mathbf{x} = 0. \quad (3.53)$$

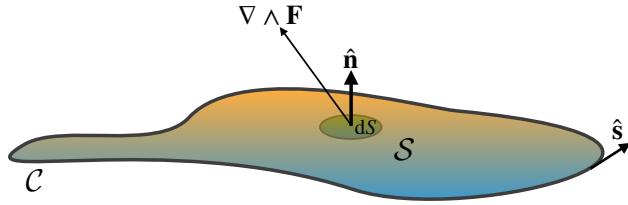


Figure 3.7: Illustrating the geometry of Stokes' Theorem. The unit normal $\hat{\mathbf{n}}$ points outward from the surface, \mathcal{S} , with $(\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}}$ the projection of the curl of a vector field onto the surface normal. The outer boundary of the area, \mathcal{C} , is traversed counterclockwise following a tangent vector $\hat{\mathbf{s}}$ when computing the circulation.

3.6 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 37).

3.6.1 Statement of Stokes' theorem

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

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{x} = \int_{\mathcal{S}} (\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS, \quad (3.54)$$

where $\hat{\mathbf{n}}$ is a unit vector normal to the surface, and dS is the infinitesimal surface area element. The orientation of the outward normal is determined by the right hand rule according to the circulation direction.

3.6.2 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

$$(\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{z}} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}, \quad (3.55)$$

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

$$\int_{\mathcal{S}} (\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS = \int_0^{L_x} \int_0^{L_y} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy. \quad (3.56)$$

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

$$\int_0^{L_x} \int_0^{L_y} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy = \int_0^{L_y} F_2(x, y) \Big|_{x=0}^{x=L_x} dy - \int_0^{L_x} F_1(x, y) \Big|_{y=0}^{y=L_y} dx \quad (3.57a)$$

$$= \int_0^{L_x} F_1(x, 0) dx + \int_0^{L_y} F_2(L_x, y) dy + \int_{L_x}^0 F_1(x, L_y) dx + \int_{L_y}^0 F_2(0, y) dy \quad (3.57b)$$

$$= \oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{x}. \quad (3.57c)$$

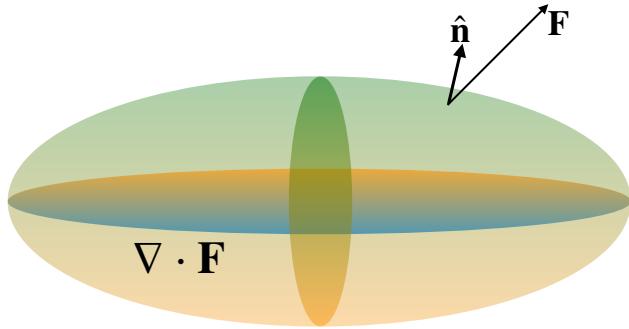


Figure 3.8: Illustrating the geometry of Gauss's law or the divergence theorem for an ellipsoidal volume. The outward unit normal $\hat{\mathbf{n}}$ is projected onto the vector field \mathbf{F} via the scalar product, $\mathbf{F} \cdot \hat{\mathbf{n}}$. Integrating this scalar product over the closed surface S yields the same result as computing the volume integral of the divergence, $\nabla \cdot \mathbf{F}$, over the region bounded by the closed surface.

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.

3.7 Divergence theorem (Gauss's law)

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

$$\int_V \nabla \cdot \mathbf{F} dV = \int_S \mathbf{F} \cdot \hat{\mathbf{n}} dS, \quad (3.58)$$

where $\hat{\mathbf{n}}$ is the outward normal to the boundary surface and dS is the surface area element.

3.7.1 An example rectangular volume

To build intuition for Gauss's law, consider a rectangular volume with dimensions $L_x \times L_y \times L_z$. The volume integral on the left hand side of equation (3.58) gives

$$\int_V \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \right) dx dy dz. \quad (3.59)$$

Focusing on just the leftmost term, integration in x gives

$$\int_V \frac{\partial F_1}{\partial x} dx dy dz = \int_{y=0}^{y=L_y} \int_{z=0}^{z=L_z} [F_1(L_x, y, z) - F_1(0, y, z)] dy dz \quad (3.60a)$$

$$= \int_{S_1+S_2} \mathbf{F} \cdot \mathbf{n} dS, \quad (3.60b)$$

where S_1 is the rectangle's face with normal $\hat{\mathbf{x}}$ and S_2 is the rectangle's face with normal $-\hat{\mathbf{x}}$. Repeating this procedure on the other terms in equation (3.59) gives the area integrated flux (i.e., the transport) through the full boundary. To verify the theorem for a general volume V , we take the approach we used to prove Stokes' theorem. First, divide up the volume up 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.

3.7.2 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 (3.58) yields

$$\int_S \phi \mathbf{c} \cdot \hat{\mathbf{n}} dS = \int_V \nabla \cdot (\phi \mathbf{c}) dV. \quad (3.61)$$

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

$$\mathbf{c} \cdot \left[\int_S \phi \hat{\mathbf{n}} dS - \iint_V \nabla \phi dV \right] = 0. \quad (3.62)$$

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

$$\int_S \phi \hat{\mathbf{n}} dS = \int_V \nabla \phi dV. \quad (3.63)$$

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 26.1.3 when formulating the contribution of stresses to the motion of a fluid element.

3.7.3 First and second form of Green's theorem

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

$$\mathbf{F} = \psi \nabla \phi, \quad (3.64)$$

with ψ and ϕ scalar fields. Substitution into the divergence theorem (3.58) leads to

$$\int_S \psi \frac{\partial \phi}{\partial n} dS = \int_V [\nabla \psi \cdot \nabla \phi + \psi \nabla^2 \phi] dV \quad \text{first form of Green's theorem.} \quad (3.65)$$

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

$$\int_S \left[\psi \frac{\partial \phi}{\partial n} - \phi \frac{\partial \psi}{\partial n} \right] dS = \int_V [\psi \nabla^2 \phi - \phi \nabla^2 \psi] dV \quad \text{second form of Green's theorem.} \quad (3.66)$$

Finally, by setting $\phi = 1$ we have

$$\int_S \frac{\partial \psi}{\partial n} dS = \int_V \nabla^2 \psi dV \iff \int_S \nabla \psi \cdot \hat{\mathbf{n}} dS = \int_V \nabla \cdot \nabla \psi dV. \quad (3.67)$$

3.8 Exercises

EXERCISE 3.1: PRACTICE WITH THE LAPLACIAN

Show that the Laplacian of the function

$$\psi = \frac{z x^2}{r^2} \quad (3.68)$$

is given by

$$\nabla^2 \psi = \frac{2 z (r^2 - 5 x^2)}{r^4} \quad (3.69)$$

where

$$r^2 = x^2 + y^2 + z^2. \quad (3.70)$$

Perform the proof using both Cartesian coordinates as well as spherical coordinates (see Figure 7.1), making use of the following expressions for Laplacian operator acting on a scalar field

$$\nabla^2 \psi(x, y, z) = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \quad (3.71a)$$

$$\nabla^2 \psi(\lambda, \phi, r) = \frac{1}{r^2 \cos \phi} \left[\frac{1}{\cos \phi} \frac{\partial^2 \psi}{\partial \lambda^2} + \frac{\partial}{\partial \phi} \left(\cos \phi \frac{\partial \psi}{\partial \phi} \right) + \cos \phi \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) \right]. \quad (3.71b)$$

EXERCISE 3.2: CONSERVATIVE VECTOR FIELD AND SCALAR POTENTIAL

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

$$\mathbf{F} = 2 x z \hat{x} + 2 y^2 \hat{y} + (x^2 + 2 y^2 z - 1) \hat{z}. \quad (3.72)$$

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)

$$\psi = -[x^2 z + (y z)^2 - z]. \quad (3.73)$$

EXERCISE 3.3: BELTRAMI FLOW

Beltrami flow is defined by a velocity and vorticity field satisfying

$$\nabla \cdot \mathbf{v} = 0 \quad (3.74a)$$

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} = \lambda \mathbf{v} \quad (3.74b)$$

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

$$\mathbf{v} = (A \sin z + C \cos y) \hat{x} + (B \sin x + A \cos z) \hat{y} + (C \sin y + B \cos x) \hat{z}, \quad (3.75)$$

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

$$u = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \quad v = \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \quad w = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \quad (3.76)$$

EXERCISE 3.4: PRACTICE WITH PATH INTEGRALS

Consider the vector field

$$\mathbf{F} = x y^2 \hat{x} + 2 \hat{y} + x \hat{z}. \quad (3.77)$$

Let L be a path parameterized by

$$x = c t \quad y = c/t \quad z = d \quad t \in [1, 2], \quad (3.78)$$

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

$$\int_L \mathbf{F} dt = c^3 \ln 2 \hat{\mathbf{x}} + 2 \hat{\mathbf{y}} + \frac{3c}{2} \hat{\mathbf{z}} \quad (3.79a)$$

$$\int_L \mathbf{F} dy = -\frac{3c^4}{8} \hat{\mathbf{x}} - c \hat{\mathbf{y}} - c^2 \ln 2 \hat{\mathbf{z}} \quad (3.79b)$$

$$\int_L \mathbf{F} \cdot d\mathbf{x} = c^4 \ln 2 - c, \quad (3.79c)$$

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 3.5: STOKES' THEOREM ON A PLANE

Show that

$$I = \oint_C [y(4x^2 + y^2) dx + x(2x^2 + 3y^2) dy] = \frac{\pi}{2} b a^3 \quad (3.80)$$

when integrating around the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (3.81)$$

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

$$\int_{\pi}^0 \sin^2(2\phi) d\phi = -\frac{\pi}{2}. \quad (3.82)$$

EXERCISE 3.6: PRACTICE WITH THE DIVERGENCE THEOREM

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

$$\mathbf{F} = \frac{\alpha \mathbf{r}}{(r^2 + a^2)^{3/2}}, \quad (3.83)$$

where α and a are constants. First, compute the flux of this \mathbf{F} through a spherical surface of radius $|\mathbf{r}| = a\sqrt{3}$. That is, compute

$$\Phi = \int_{|\mathbf{r}|=a\sqrt{3}} \mathbf{F} \cdot d\mathbf{r} = \frac{3\pi\alpha\sqrt{3}}{2}. \quad (3.84)$$

Next, show that this flux is equal to the integral of the divergence over the volume of the sphere

$$\Phi = \int_{|\mathbf{r}|=a\sqrt{3}} \nabla \cdot \mathbf{F} dV. \quad (3.85)$$

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

$$\int_V \nabla \cdot \mathbf{F} dV = \int_{\partial V} \mathbf{F} \cdot \hat{\mathbf{n}} dS, \quad (3.86)$$

where $\hat{\mathbf{n}}$ is the outward normal on the boundary ∂V .

4

General tensors in brief

In the study of physics, we aim to uncover objective statements about how the physical universe works. 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 begin the process of furthering our skills with such tools with a focus on general tensors.

READER'S GUIDE TO THIS CHAPTER

This chapter offers an introduction to general tensor analysis. The general notions discussed here are of use even when not concerned with the details that are taken up in Chapters 5 and 6.

4.1	General tensors and geophysical fluid mechanics	43
4.2	Points and coordinates	44
4.2.1	The time parameter and the time coordinate	44
4.2.2	The importance of index placement	45
4.3	Example coordinate descriptions	45
4.3.1	Eulerian coordinates	45
4.3.2	Isopycnal coordinates	45
4.3.3	Lagrangian or material coordinates	46
4.3.4	Tracer coordinates	46
4.4	The velocity vector and basis vectors	46
4.4.1	Coordinate representation of the velocity vector	47
4.4.2	Basis vectors	47
4.5	Notational conventions	47
4.5.1	Placement of tensor labels	47
4.5.2	Einstein summation convention	48
4.5.3	The boldface notation	48

4.1 General tensors and geophysical fluid mechanics

The Cartesian tensors described in Chapters 2 and 3 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. 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.

4.2 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 14.1). 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.

4.2.1 The time parameter and the time 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 4.2.2), then $\partial/\partial t$ is perhaps a mixture of Lagrangian and Eulerian or perhaps neither.

4.2.2 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 4.3.

4.3 Example coordinate descriptions

We here offer a few examples of coordinates used for describing geophysical fluid systems.

4.3.1 Eulerian coordinates

The Cartesian coordinates for a point are written

$$\xi^a = (x, y, z) \quad \text{Cartesian}, \quad (4.1)$$

whereas for spherical coordinates we write (see Figure 7.1)

$$\xi^a = (r, \lambda, \phi) \quad \text{spherical} \quad (4.2)$$

and polar cylindrical coordinates

$$\xi^a = (r, \lambda, z) \quad \text{cylindrical}. \quad (4.3)$$

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 14.2.

As shown in Section 4.4, 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 2 and 3). 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.

4.3.2 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

$$\xi^a = (x, y, b) \quad \text{isopycnal coordinates}, \quad (4.4)$$

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 ideal fluid flow (flow absent irreversible

processes such as mixing). Hence, all fluid particle motion occurs on surfaces of constant b . Under such ideal 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.

4.3.3 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 14.2.

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 ideal fluid motion. In this manner, we refer to the isopycnal coordinates as “quasi-Lagrangian” since its vertical coordinate follows the vertical position of an ideal fluid parcel whereas its horizontal coordinates are Eulerian.

4.3.4 Tracer coordinates

Consider a trio of linearly independent tracer concentrations $C^a = C^a(x, y, z, t)$ (we introduce tracers in Section 15.4). 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

$$\xi^a = (C^1, C^2, C^3). \quad (4.5)$$

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.

4.4 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

$$\vec{v}(\tau) = \lim_{\Delta\tau \rightarrow 0} \frac{\mathcal{P}(\tau + \Delta\tau/2) - \mathcal{P}(\tau - \Delta\tau/2)}{\Delta\tau} \quad (4.6a)$$

$$= \frac{d\mathcal{P}(\tau)}{d\tau}. \quad (4.6b)$$

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. 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.

4.4.1 Coordinate representation of the velocity vector

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

$$\mathcal{P}(\tau) = \mathcal{P}[\xi^a(\tau)], \quad (4.7)$$

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

$$\vec{v}(\tau) = \frac{d\mathcal{P}(\tau)}{d\tau} \quad (4.8a)$$

$$= \frac{\partial \mathcal{P}}{\partial \xi^a} \frac{d\xi^a}{d\tau} \quad (4.8b)$$

$$\equiv \vec{e}_a v^a. \quad (4.8c)$$

The expansion coefficients

$$v^a = \frac{d\xi^a}{d\tau} \quad (4.9)$$

provide a representation the velocity vector $\vec{v}(\tau)$ within the coordinate system ξ^a .

4.4.2 Basis vectors

For each number v^a there is a corresponding basis vector \vec{e}_a defined by

$$\vec{e}_a = \frac{\partial \mathcal{P}}{\partial \xi^a}. \quad (4.10)$$

The basis vectors are generally a function of position and time, although they are constant for Cartesian coordinates.

4.5 Notational conventions

We here introduce notational conventions that help to simplify many of the manipulations (“index gymnastics”) encountered with general tensors.

4.5.1 Placement of tensor labels

As indicated in Section 4.2.2, 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 (4.10) 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.

4.5.2 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

$$\vec{v}(\tau) = \sum_{a=1}^3 \vec{e}_a v^a = \vec{e}_a v^a. \quad (4.11)$$

This rule generalizes that used for Cartesian tensors in Chapter 2. For general tensors, contraction is between a lower and an upper label. We return to such contractions in Section 5.2, where we show that the contraction of a vector and a one-form renders a scalar.

4.5.3 The boldface notation

Motivated by notation common in conventional vector analysis, we often find it useful to organize the velocity vector representation v^a into an ordered list (v^1, v^2, v^3) and to use the boldface notation

$$\boldsymbol{v} = (v^1, v^2, v^3). \quad (4.12)$$

Likewise, we organize the basis vectors according to

$$\boldsymbol{\vec{e}} = (\vec{e}_1, \vec{e}_2, \vec{e}_3) \quad (4.13)$$

and the coordinates as

$$\boldsymbol{\xi} = (\xi^1, \xi^2, \xi^3). \quad (4.14)$$

With this notation the velocity vector representation from equation (4.8c) takes on the form

$$\vec{v} = \vec{e}_a v^a = \boldsymbol{\vec{e}} \cdot \boldsymbol{v}. \quad (4.15)$$

Likewise, a trajectory can be represented in terms of a chosen set of coordinates according to

$$\boldsymbol{\mathcal{P}} = \vec{e}_a \xi^a = \boldsymbol{\vec{e}} \cdot \boldsymbol{\xi}. \quad (4.16)$$

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.

5

General tensor algebra[†]

In this chapter we develop the algebra of general tensors. General tensor algebra is very similar to Cartesian tensors (Chapter 2), 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 6, with applications to the mathematics of generalized vertical coordinates discussed in Chapter 8. We assume here a mastery of Cartesian tensor algebra in Chapter 2.

5.1	The metric tensor and coordinate transformations	50
5.1.1	Cartesian coordinates in Euclidean space	50
5.1.2	The metric as a symmetric second order tensor	50
5.1.3	Coordinate representation of the metric tensor	51
5.1.4	Transforming the representation of the metric tensor	51
5.1.5	Finite distance between points	52
5.2	One-forms	52
5.2.1	Coordinate representation of a one-form	52
5.2.2	Basis one-forms and the orthogonality relation	52
5.2.3	Metric as a mapping between vectors and one-forms	53
5.2.4	Transformation of the coordinate representation	53
5.3	Scalar product	54
5.4	The volume element and Jacobian of transformation	54
5.4.1	Jacobian of transformation	54
5.4.2	Relating the Jacobian to the determinant of the metric	55
5.5	The permutation symbol and the determinant	55
5.5.1	Connecting the permutation symbol to the determinant	55
5.5.2	Further identities satisfied by the determinant	56
5.6	The Levi-Civita tensor and the volume element	56
5.6.1	General coordinate representation of the Levi-Civita tensor	57
5.6.2	The Levi-Civita tensor and the volume element	57
5.7	Cross product and biorthogonality relation	57

5.1 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.

5.1.1 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 2.3)

$$ds^2 = d\vec{x} \cdot d\vec{x} = \vec{e}_a \cdot \vec{e}_b dx^a dx^b = \delta_{ab} dx^a dx^b. \quad (5.1)$$

In this expression,

$$(ds)^2 \equiv ds^2 \quad (5.2)$$

is the squared infinitesimal arc-length separating the two points. The Cartesian representation of the Kronecker symbol, δ_{ab} , is symmetric

$$\delta_{ab} = \delta_{ba}, \quad (5.3)$$

and vanishes when $a \neq b$ and is unity when $a = b$.

5.1.2 The metric as a symmetric second order tensor

As defined by equation (5.1), 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

$$\text{distance}(\vec{P}, \vec{Q}) = \sqrt{\mathcal{G}(\vec{P}, \vec{Q})}. \quad (5.4)$$

Here, \mathcal{G} is the metric tensor with coordinate representation g_{ab} and \vec{P}, \vec{Q} are infinitesimally close vectors with coordinate representations

$$\vec{P} = \xi^a \vec{e}_a \quad \vec{Q} = \vec{P} + d\xi^a \vec{e}_a. \quad (5.5)$$

Equation (5.4) indicates that the metric tensor takes two vectors as argument and produces a scalar. Furthermore, since

$$\text{distance}(\vec{P}, \vec{Q}) = \text{distance}(\vec{Q}, \vec{P}) \geq 0, \quad (5.6)$$

the metric tensor is a symmetric and positive tensor that produces zero only when $\vec{P} = \vec{Q}$.

5.1.3 Coordinate representation of the metric tensor

Given the geometric expression (5.4) for the metric, we determine its representation in an arbitrary coordinate system by considering the squared distance between the coordinate basis vectors

$$\text{distance}(\vec{e}_a, \vec{e}_b) = \sqrt{\mathcal{G}(\vec{e}_a, \vec{e}_b)}. \quad (5.7)$$

This relation determines the coordinate components of the metric tensor

$$\mathcal{G}(\vec{e}_a, \vec{e}_b) \equiv g_{ab}. \quad (5.8)$$

Furthermore, in Euclidean space this relation is written

$$g_{ab} = \vec{e}_a \cdot \vec{e}_b. \quad (5.9)$$

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 $a = b$.

5.1.4 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,

$$ds^2 = \delta_{ab} d\xi^a d\xi^b \quad (5.10a)$$

$$= \delta_{ab} \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} \frac{\partial \xi^b}{\partial \xi^{\bar{b}}} d\xi^{\bar{a}} d\xi^{\bar{b}} \quad (5.10b)$$

$$\equiv \delta_{ab} \Lambda_{\bar{a}}^a \Lambda_{\bar{b}}^b d\xi^{\bar{a}} d\xi^{\bar{b}} \quad (5.10c)$$

$$\equiv g_{\bar{a}\bar{b}} d\xi^{\bar{a}} d\xi^{\bar{b}}, \quad (5.10d)$$

where

$$g_{\bar{a}\bar{b}} = \delta_{ab} \Lambda_{\bar{a}}^a \Lambda_{\bar{b}}^b \quad (5.11)$$

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

$$\Lambda^a_{\bar{a}} = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}}. \quad (5.12)$$

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.

¹We generally follow the convention of displacing the lower index to the right to help keep track of which index refers to the column.

5.1.5 Finite distance between points

Once the metric is determined, the distance along a curve between two finitely separated points is given by the integration

$$\begin{aligned} L &= \int \sqrt{ds^2} \\ &= \int_{\varphi_1}^{\varphi_2} \left| g_{ab} \frac{d\xi^a}{d\varphi} \frac{d\xi^b}{d\varphi} \right|^{1/2} d\varphi, \end{aligned} \quad (5.13)$$

where φ is a parameter specifying the curve (e.g., the arc length as in Section 3.4), and $\varphi_{1,2}$ are the curve's endpoints.

5.2 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 (5.4))

$$\text{distance}(\vec{A}, \vec{B}) = \sqrt{\mathcal{G}(\vec{A}, \vec{B})}. \quad (5.14)$$

What if the metric only eats one vector? The resulting geometric object is known as a one-form

$$\tilde{A} \equiv \mathcal{G}(\vec{A},), \quad (5.15)$$

with the tilde used to distinguish a one-form from a vector.

5.2.1 Coordinate representation of a one-form

We can determine the coordinate representation of a one-form by eating a basis vector

$$\tilde{A}(\vec{e}_b) = \mathcal{G}(\vec{A}, \vec{e}_b) \quad (5.16a)$$

$$= \mathcal{G}(A^a \vec{e}_a, \vec{e}_b) \quad (5.16b)$$

$$= \mathcal{G}(\vec{e}_a, \vec{e}_b) A^a \quad (5.16c)$$

$$= g_{ab} A^a. \quad (5.16d)$$

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

$$A_b = g_{ab} A^a. \quad (5.17)$$

5.2.2 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

$$\mathcal{G}(\tilde{e}^a, \vec{e}_b) = \tilde{e}^a \cdot \vec{e}_b = \delta_b^a, \quad (5.18)$$

where

$$\delta_b^a = g^{ac} g_{cb} \quad (5.19)$$

are components to the Kronecker delta tensor, taking the value of unity when $a = b$ and zero otherwise

$$\delta_a^a = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5.20)$$

It is only for Cartesian coordinates that we have

$$\delta_c^a = g^{ab} \delta_{bc} \quad \text{Cartesian coordinates,} \quad (5.21)$$

which follows since $g^{ab} = \delta^{ab}$ in Cartesian coordinates.

5.2.3 Metric as a mapping between vectors and one-forms

We can contract the expression (5.17) with components of the inverse metric tensor, g^{ab} , to render

$$g^{ab} A_b = g^{ab} g_{bc} A^c \quad (5.22a)$$

$$= \delta_c^a A^c \quad (5.22b)$$

$$= A^a. \quad (5.22c)$$

This identity, as well as equation (5.17), 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 (5.17) 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.

5.2.4 Transformation of the coordinate representation

The transformation matrix (5.12) 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

$$\vec{F}(\tilde{e}^a) = F^a. \quad (5.23)$$

Now consider another coordinate system with basis one-forms $\tilde{e}^{\bar{a}}$, so that the vector has a representation

$$\vec{F}(\tilde{e}^{\bar{a}}) = F^{\bar{a}}. \quad (5.24)$$

Transforming the basis one-form using the transformation matrix leads to

$$F^{\bar{a}} = \vec{F}(\tilde{e}^{\bar{a}}) = \vec{F}(\Lambda_a^{\bar{a}} \tilde{e}^a) = \Lambda_a^{\bar{a}} \vec{F}(\tilde{e}^a) = \Lambda_a^{\bar{a}} F^a. \quad (5.25)$$

Transformation of an arbitrary one-form representation takes place with the inverse transformation matrix

$$F_{\bar{a}} = \tilde{F}(\tilde{e}_{\bar{a}}) = \tilde{F}(\Lambda_{\bar{a}}^a \tilde{e}_a) = \Lambda_{\bar{a}}^a \tilde{F}(\tilde{e}_a) = \Lambda_{\bar{a}}^a F_a. \quad (5.26)$$

5.3 Scalar product

In Section 2.3.2 we defined the scalar product between two Cartesian vectors. The natural generalization is given by

$$\vec{P} \cdot \vec{Q} = P^a Q^b \vec{e}_a \cdot \vec{e}_b = P^a Q^b g_{ab} = P^a Q_a = P_b Q^b, \quad (5.27)$$

where the second equality made use of the metric tensor coordinate representation given by equation (5.9). 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

$$\tilde{P}(\vec{Q}) = \tilde{P}(Q^a \vec{e}_a) = Q^a \tilde{P}(\vec{e}_a) = Q^a P_a, \quad (5.28)$$

which equals to

$$\vec{Q}(\tilde{P}) = \vec{Q}(P_a \tilde{e}^a) = P_a \vec{Q}(\tilde{e}^a) = P_a Q^a. \quad (5.29)$$

The scalar product is invariant to coordinate changes, as seen through

$$\vec{Q}(\tilde{P}) = \vec{Q}(P_a \tilde{e}^a) = P_a Q^a = \vec{Q}(P_{\bar{a}} \tilde{e}^{\bar{a}}) = P_{\bar{a}} Q^{\bar{a}}. \quad (5.30)$$

The invariance is also revealed by working just with the coordinate representations and introducing the transformation matrix elements

$$P_a Q^a = (\Lambda^{\bar{a}}_a P_{\bar{a}}) (\Lambda^a_{\bar{b}} Q^{\bar{b}}) = \Lambda^{\bar{a}}_a \Lambda^a_{\bar{b}} P_{\bar{a}} Q^{\bar{b}} = \delta^{\bar{a}}_{\bar{b}} P_{\bar{a}} Q^{\bar{b}} = P_{\bar{a}} Q^{\bar{a}}. \quad (5.31)$$

5.4 The volume element and Jacobian of transformation

Recall from Section 2.5.2 that we derived an expression for the volume of an infinitesimal region of Euclidean space \mathbb{R}^3 using Cartesian coordinates

$$dV = dx dy dz (\hat{x} \wedge \hat{y}) \cdot \hat{z} = dx dy dz. \quad (5.32)$$

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.

5.4.1 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

$$d\xi^1 d\xi^2 d\xi^3 = \frac{\partial(\xi^1, \xi^2, \xi^3)}{\partial(\xi^{\bar{1}}, \xi^{\bar{2}}, \xi^{\bar{3}})} d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}} \quad (5.33a)$$

$$= \det(\Lambda^a_{\bar{a}}) d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}, \quad (5.33b)$$

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.

5.4.2 Relating the Jacobian to the determinant of the metric

Recall the expression (5.10d) for the transformation of the metric

$$g_{\bar{a}\bar{b}} = \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} g_{ab}. \quad (5.34)$$

We can write this expression as a matrix equation

$$\bar{\mathcal{G}} = \Lambda^T \mathcal{G} \Lambda \quad (5.35)$$

where Λ^T is the transposed matrix. This equation is valid upon taking determinants of both sides so that

$$\det(\bar{\mathcal{G}}) = \det(\Lambda^T \mathcal{G} \Lambda) \quad (5.36a)$$

$$= \det(\Lambda^T) \det(\mathcal{G}) \det(\Lambda) \quad (5.36b)$$

$$= [\det(\Lambda)]^2 \det(\mathcal{G}). \quad (5.36c)$$

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,

$$\det(\Lambda^a_{\bar{a}}) = \frac{\sqrt{\det(g_{\bar{a}\bar{b}})}}{\sqrt{\det(g_{ab})}} = \frac{\sqrt{\det(\bar{\mathcal{G}})}}{\sqrt{\det(\mathcal{G})}}. \quad (5.37)$$

We are thus led to the equivalent expressions for the volume element

$$dV \equiv \sqrt{\det(\mathcal{G})} d\xi^1 d\xi^2 d\xi^3 = \sqrt{\det(\bar{\mathcal{G}})} d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}. \quad (5.38)$$

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

$$\det(\Lambda^a_{\bar{a}}) = \sqrt{\det(g_{\bar{a}\bar{b}})} \quad \text{unbarred coordinates are Cartesian.} \quad (5.39)$$

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.

5.5 The permutation symbol and the determinant

As discussed in Section 2.4.1, 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.

5.5.1 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

$$\det(\Lambda^a_{\bar{a}}) = \Lambda^1_{\bar{1}} \Lambda^2_{\bar{2}} - \Lambda^1_{\bar{2}} \Lambda^2_{\bar{1}}. \quad (5.40)$$

Introducing the permutation symbol ϵ_{ab} allows us to write this expression in a more tidy manner

$$\det(\Lambda^a_{\bar{a}}) = \epsilon_{ab} \Lambda^a_{\bar{1}} \Lambda^b_{\bar{2}} \quad (5.41)$$

with

$$\epsilon_{12} = 1 \quad \epsilon_{21} = -1. \quad (5.42)$$

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 2.4.1 when discussing the vector cross product, in which case the permutation symbol is

$$\epsilon_{123} = 1 \quad (5.43a)$$

$$\epsilon_{abc} = \begin{cases} 0 & \text{if any two labels are the same,} \\ 1 & \text{if } a, b, c \text{ is an even permutation of } 1, 2, 3, \\ -1 & \text{if } a, b, c \text{ is an odd permutation of } 1, 2, 3. \end{cases} \quad (5.43b)$$

Likewise, the determinant of the transformation matrix takes the form

$$\det(\Lambda^a_{\bar{a}}) = \frac{\partial(\xi^1, \xi^2, \xi^3)}{\partial(\bar{\xi}^1, \bar{\xi}^2, \bar{\xi}^3)} = \frac{\partial \boldsymbol{\xi}}{\partial \bar{\boldsymbol{\xi}}} = \epsilon_{abc} \Lambda^a_{\bar{1}} \Lambda^b_{\bar{2}} \Lambda^c_{\bar{3}}. \quad (5.44)$$

5.5.2 Further identities satisfied by the determinant

The following identity in two dimensions can be readily verified through enumeration

$$\epsilon_{ab} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} = \epsilon_{\bar{a}\bar{b}} \det(\Lambda^a_{\bar{a}}). \quad (5.45)$$

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

$$\frac{1}{2} \epsilon^{\bar{a}\bar{b}} \epsilon_{ab} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} = \det(\Lambda^a_{\bar{a}}), \quad (5.46)$$

where we used

$$\epsilon^{\bar{a}\bar{b}} \epsilon_{\bar{a}\bar{b}} = \epsilon^{\bar{1}\bar{2}} \epsilon_{\bar{1}\bar{2}} + \epsilon^{\bar{2}\bar{1}} \epsilon_{\bar{2}\bar{1}} = 2. \quad (5.47)$$

The three dimensional version takes the form

$$\epsilon_{abc} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} = \epsilon_{\bar{a}\bar{b}\bar{c}} \det(\Lambda^a_{\bar{a}}), \quad (5.48)$$

along with

$$\frac{1}{3!} \epsilon^{\bar{a}\bar{b}\bar{c}} \epsilon_{abc} \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} = \det(\Lambda^a_{\bar{a}}). \quad (5.49)$$

5.6 The Levi-Civita tensor and the volume element

The metric tensor introduced in Section 5.1 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 2.5.3.

5.6.1 General coordinate representation of the Levi-Civita tensor

The relations (5.45) and (5.48) 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 2) 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*

$$\varepsilon_{abc} = \sqrt{\det(\mathcal{G})} \epsilon_{abc}. \quad (5.50)$$

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

$$\Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} \varepsilon_{abc} = \Lambda^a_{\bar{a}} \Lambda^b_{\bar{b}} \Lambda^c_{\bar{c}} \sqrt{\det(\mathcal{G})} \epsilon_{abc} \quad (5.51a)$$

$$= \sqrt{\det(\mathcal{G})} \epsilon_{\bar{a}\bar{b}\bar{c}} \det(\Lambda^a_{\bar{a}}) \quad (5.51b)$$

$$= \sqrt{\det(\mathcal{G})} \epsilon_{\bar{a}\bar{b}\bar{c}} \quad (5.51c)$$

$$= \varepsilon_{\bar{a}\bar{b}\bar{c}}, \quad (5.51d)$$

where equations (5.37) and (5.45) were used. Therefore, ε_{abc} transforms as components to a third order covariant tensor. Likewise,

$$\varepsilon^{abc} = \frac{\epsilon^{abc}}{\sqrt{\det(\mathcal{G})}} \quad (5.52)$$

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.

5.6.2 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

$$\varepsilon(\vec{e}_1 d\xi^1, \vec{e}_2 d\xi^2, \vec{e}_3 d\xi^3) = d\xi^1 d\xi^2 d\xi^3 \varepsilon(\vec{e}_1, \vec{e}_2, \vec{e}_3) \quad (5.53a)$$

$$= d\xi^1 d\xi^2 d\xi^3 \varepsilon_{123} \quad (5.53b)$$

$$= d\xi^1 d\xi^2 d\xi^3 \sqrt{\det(\mathcal{G})} \epsilon_{123} \quad (5.53c)$$

$$= dV, \quad (5.53d)$$

where we used equation (5.38) for the final step. This result means that geometrically, the Levi-Civita tensor measures the volume defined by three vectors

$$\varepsilon(\vec{A}, \vec{B}, \vec{C}) = \text{volume}(\vec{A}, \vec{B}, \vec{C}). \quad (5.54)$$

This interpretation accords with the Cartesian coordinate discussion of the Levi-Civita tensor in Section 2.5.3.

5.7 Cross product and biorthogonality relation

The cross product of two Cartesian basis vectors yields the third, so that

$$\hat{x} \wedge \hat{y} = \hat{z} \quad \text{cyclic.} \quad (5.55)$$

The coordinate invariant generalization of this relation is given by the biorthogonality relation

$$\vec{e}_a \wedge \vec{e}_b \equiv \varepsilon_{abc} \tilde{e}^c. \quad (5.56)$$

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

$$\vec{P} \wedge \vec{Q} = P^a Q^b \vec{e}_a \wedge \vec{e}_b \quad (5.57a)$$

$$= P^a Q^b \varepsilon_{abc} \tilde{e}^c. \quad (5.57b)$$

6

General tensor calculus[†]

READER'S GUIDE TO THIS CHAPTER

In this chapter, we generalize the Cartesian vector calculus of Chapter 3 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. This chapter assumes understanding of the general tensor algebra in Chapter 5. 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. Also note that much here stems from similar material given in Chapter 21 of *Griffies* (2004).

6.1	General covariance	59
6.2	The covariant derivative operator	60
6.3	Covariant derivative of a vector	60
6.3.1	Derivative of a vector	61
6.3.2	An alternative derivation	61
6.3.3	Christoffel symbols are not components of a tensor	62
6.4	Covariant derivative of a one-form	62
6.5	Covariant derivative of the metric	62
6.6	Christoffel symbols in terms of the metric	62
6.7	Covariant divergence of a vector	63
6.7.1	Contraction of the Christoffel symbols	63
6.7.2	Exponential of the determinant	63
6.8	Covariant Laplacian of a scalar	64
6.9	Covariant curl of a vector	64
6.10	The Lie derivative	64
6.11	Divergence theorem	64
6.12	Stokes' theorem	65

6.1 General covariance

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. The “extend that relation” step requires the equations to respect *general covariance*.

Operationally, general covariance means that all tensor indices are properly matched and each derivative is covariant. In this chapter, we provide the details needed to understand general covariance. The elegance and power rendered by general covariance is the key reason that tensor analysis is ubiquitous in theoretical physics.

Respecting general covariance means that an equation retains the same form under general coordinate changes. That is, general covariance equals form invariance. To ensure an equation respects general covariance requires us to understand certain properties of general tensors. We start by noting that if two tensors have equal components in a given coordinate basis, then they have equal components in all bases and as such are identical. In particular, if all components of a tensor vanish in one set of coordinates, they vanish in all coordinates. Furthermore, if an equation consists of tensors combined only by *permissible tensor operations*, and if the equation is true in one basis, then it is true in any coordinate basis. A permissible tensor operation is an operation that produces components of new tensors. We list here some of the permissible tensor operations that are used in these notes.

1. Multiplication of a tensor by a scalar produces a new tensor of the same type.
2. Addition of components of two tensors of the same type gives components of a new tensor of the same type.
3. Covariant differentiation of a tensor increases by one the type of a tensor. For example, the covariant derivative of a scalar produces a vector and the covariant derivative of a vector produces a second order tensor.

For the partial differential equations of geophysical fluid mechanics, covariant differentiation is the key to general covariance.

6.2 The covariant derivative operator

Application of the chain rule leads to the transformation of the partial derivative operator

$$\partial_{\bar{a}} = \frac{\partial}{\partial \xi^{\bar{a}}} = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} \frac{\partial}{\partial \xi^a} = \Lambda^a_{\bar{a}} \partial_a. \quad (6.1)$$

Contracting the partial derivative with the basis of one-forms renders the form invariant expression of the gradient

$$\text{grad}(\psi) = \nabla \psi = \tilde{e}^a \partial_a \psi = \tilde{e}^{\bar{a}} \partial_{\bar{a}} \psi. \quad (6.2)$$

We thus define the covariant derivative operator

$$\nabla = \tilde{e}^a \partial_a. \quad (6.3)$$

Equation (6.2) provides the expression for the covariant derivative when acting on a scalar field, which is more commonly known as the gradient of the scalar.

6.3 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

$$\nabla \vec{F} = (\tilde{e}^b \partial_b) (F^a \vec{e}_a). \quad (6.4)$$

6.3.1 Derivative of a vector

The chain rule leads to the expression for the partial derivative operator acting on a vector field

$$\partial_b \vec{F} = \partial_b(\vec{e}_a F^a) \quad \text{coordinate representation of the vector } \vec{F} \quad (6.5a)$$

$$= (\partial_b F^a) \vec{e}_a + F^a \partial_b \vec{e}_a \quad \text{chain rule} \quad (6.5b)$$

$$= (\partial_b F^a) \vec{e}_a + F^a \Gamma_{ba}^c \vec{e}_c \quad \text{define Christoffel symbols} \quad (6.5c)$$

$$= (\partial_b F^a + F^c \Gamma_{bc}^a) \vec{e}_a \quad \text{reorganize} \quad (6.5d)$$

$$= (\nabla_b F^a) \vec{e}_a \quad \text{define covariant derivative acting on vector component.} \quad (6.5e)$$

In the third equality we introduced the *Christoffel symbols*

$$\partial_b \vec{e}_a = \Gamma_{ba}^c \vec{e}_c. \quad (6.6)$$

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

$$\nabla_b F^a = \partial_b F^a + \Gamma_{bc}^a F^c. \quad (6.7)$$

Contracting $\partial_b \vec{F}$ with the basis one-form \tilde{e}^b leads to

$$\nabla \vec{F} = (\tilde{e}^b \partial_b) \vec{F} = (\tilde{e}^b \nabla_b F^a) \vec{e}_a. \quad (6.8)$$

6.3.2 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

$$\partial_b \vec{F} = \lim_{\Delta \rightarrow 0} \frac{\vec{F}(\vec{P} + \Delta \vec{e}_b) - \vec{F}(\vec{P})}{\Delta}, \quad (6.9)$$

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

$$\partial_b \vec{F} = (\partial_b F^a) \vec{e}_a \quad \text{Cartesian coordinates.} \quad (6.10)$$

However, for general coordinates both the vector components and the basis vectors are spatially dependent, in which case

$$\vec{F}(\vec{P} + \Delta \vec{e}_b) - \vec{F}(\vec{P}) = [F^a + \Delta \partial_b F^b] [\vec{e}_a + \Delta \partial_b \vec{e}_a] - F^a \vec{e}_a \quad (6.11a)$$

$$= \Delta \partial_b (F^a \vec{e}_a) + \mathcal{O}(\Delta^2). \quad (6.11b)$$

This is the same result as found in the first step of the chain rule used in equation (6.5a). Following through that derivation then leads to the same coordinate expression for the covariant derivative acting on a vector field.

6.3.3 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 6.1, 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.

6.4 Covariant derivative of a one-form

The gradient acting on the product of a one-form and a vector is given by

$$\nabla(\tilde{E} \cdot \vec{F}) = \tilde{e}^b \partial_b (E_a F^a). \quad (6.12)$$

Expanding the partial derivative yields

$$\partial_b(E_a F^a) = F^a \partial_b E_a + E_a \partial_b F^a \quad (6.13a)$$

$$= F^a \partial_b E_a + E_a (\nabla_b F^a - \Gamma_{bc}^a F^c) \quad (6.13b)$$

$$= F^a (\partial_b E_a - \Gamma_{ba}^c E_c) + E_a \nabla_b F^a \quad (6.13c)$$

$$\equiv F^a \nabla_b E_a + E_a \nabla_b F^a. \quad (6.13d)$$

The last equality defines the covariant derivative when acting on the components to a one form

$$\nabla_b E_a = \partial_b E_a - \Gamma_{ba}^c E_c, \quad (6.14)$$

which leads to

$$\nabla \tilde{E} = (\tilde{e}^b \partial_b) \tilde{E} = (\tilde{e}^b \nabla_b E_a) \tilde{e}^a. \quad (6.15)$$

6.5 Covariant derivative of the metric

When written in Cartesian coordinates, the covariant derivative of the metric for Euclidean space vanishes,

$$\nabla g_{ab} = \nabla \delta_{ab} = 0, \quad (6.16)$$

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.

6.6 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 (6.16)) then leads to the identity

$$0 = \nabla_c g_{ab} = \partial_c g_{ab} - \Gamma_{ca}^d g_{db} - \Gamma_{cb}^d g_{ad}. \quad (6.17)$$

We can solve this equation for the Christoffel symbols

$$\Gamma_{ab}^c = \frac{1}{2} g^{cd} (\partial_b g_{da} + \partial_a g_{db} - \partial_d g_{ab}). \quad (6.18)$$

This expression exhibits the symmetry property of the lower two indices on the Christoffel symbols

$$\Gamma_{ab}^c = \Gamma_{ba}^c. \quad (6.19)$$

6.7 Covariant divergence of a vector

The covariant divergence of the components to a vector results in a scalar

$$\nabla_a F^a = \partial_a F^a + \Gamma_{ab}^a F^b. \quad (6.20)$$

We now bring this expression into a form more convenient for practical calculations.

6.7.1 Contraction of the Christoffel symbols

Expression (6.18) for the Christoffel symbols yields for the contraction

$$\Gamma_{ab}^a = \frac{1}{2} g^{ad} (\partial_b g_{da} + \partial_a g_{db} - \partial_d g_{ab}) = \frac{1}{2} g^{ad} \partial_b g_{ad} \quad (6.21)$$

where symmetry of the metric tensor and its inverse was used.

6.7.2 Exponential of the determinant

For any symmetric positive definite matrix such as the metric tensor we can write

$$\det(A) = e^{\ln \det(A)} \quad \text{simple identity} \quad (6.22a)$$

$$= e^{\ln(\prod_i \Lambda_i)} \quad \text{determinant related to product of eigenvalues} \quad (6.22b)$$

$$= e^{\sum_i \ln \Lambda_i} \quad \text{simple identity} \quad (6.22c)$$

$$= e^{\text{Tr}(\ln A)} \quad \text{sum of eigenvalues related to trace of matrix.} \quad (6.22d)$$

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

$$\partial_c \ln \det(A) = \partial_c [\text{Tr}(\ln A)] \quad (6.23a)$$

$$= \text{Tr}(\partial_c \ln A) \quad (6.23b)$$

$$= \text{Tr}(A^{-1} \partial_c A). \quad (6.23c)$$

With A now set equal to the metric tensor \mathcal{G} with components g_{ab} , this result yields

$$\partial_c \ln \det(\mathcal{G}) = g^{ab} \partial_c g_{ab} \quad (6.24)$$

which in turn yields for the contracted Christoffel symbol

$$\Gamma_{ac}^a = \partial_c \ln \sqrt{\det(\mathcal{G})}. \quad (6.25)$$

This result brings the covariant divergence of a vector to the form

$$\nabla_a F^a = \partial_a F^a + F^a \partial_a \ln \sqrt{\det(\mathcal{G})} \quad (6.26a)$$

$$= \frac{1}{\sqrt{\mathcal{G}}} \partial_a [\sqrt{\mathcal{G}} F^a]. \quad (6.26b)$$

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 $\sqrt{\det(\mathcal{G})}$.

6.8 Covariant Laplacian of a scalar

Making use of equation (8.72) with

$$F^a = g^{ab} \partial_b \psi \quad (6.27)$$

leads to the covariant Laplacian of a scalar field

$$\nabla_a (g^{ab} \partial_b \psi) = \frac{1}{\sqrt{\mathcal{G}}} \partial_a [\sqrt{\mathcal{G}} g^{ab} \partial_b \psi]. \quad (6.28)$$

This expression is fundamental to the evolution of scalar fields under the impacts from diffusion (Chapter 22).

6.9 Covariant curl of a vector

The Levi-Civita tensor $\varepsilon_{abc} = \sqrt{\det(\mathcal{G})} \epsilon_{abc}$ from Section 5.6 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

$$\text{curl}(\vec{F}) = \varepsilon^{abc} (\nabla_b F_c) \vec{e}_a = \varepsilon_{abc} (\nabla^b F^c) \vec{e}^a. \quad (6.29)$$

This expression simplifies by making use of equation (6.14) 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

$$\text{curl}(\vec{F}) = \varepsilon^{abc} (\partial_b F_c) \vec{e}_a = \varepsilon^{abc} [\partial_b (g_{cd} F^d)] \vec{e}_a. \quad (6.30)$$

The second equality made use of the equality $F_c = g_{cd} F^d$.

6.10 The Lie derivative

As in Setion 14.2.1 of [Thorne and Blandford \(2017\)](#).

6.11 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 6.1, 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 (5.38)

$$dV = \sqrt{\det(\mathcal{G})} d\xi^1 d\xi^2 d\xi^3, \quad (6.31)$$

multiplied by the covariant divergence (8.72). Hence, the volume integral of the divergence is given by

$$\int_V (\nabla_a F^a) dV = \int_V \partial_a [\sqrt{\det(\mathcal{G})} F^a] d\xi^1 d\xi^2 d\xi^3 = \int_{\partial V} F^a \hat{n}_a dS, \quad (6.32)$$

where \hat{n} is the outward normal one-form for the boundary and \hat{n}_a are its covariant components.

6.12 Stokes' theorem

The Cartesian form of Stokes' Theorem from Section 3.6 is generalized in a manner similar to the divergence theorem

$$\oint_C \vec{F} \cdot d\vec{x} = \int_S \text{curl}(\vec{F}) \cdot \hat{n} dS. \quad (6.33)$$

For the circulation on the left hand side we have

$$\vec{F} \cdot d\vec{x} = F^a \vec{e}_a \cdot \vec{e}_b dx^b = F_b dx^b = F_{\bar{b}} d\xi^{\bar{b}}. \quad (6.34)$$

For the curl on the right hand side we have

$$\text{curl}(\vec{F}) \cdot \hat{n} = \varepsilon^{abc} (\partial_b F_c) \vec{e}_a \cdot \hat{n} = \varepsilon^{abc} (\partial_b F_c) \hat{n}_a = \varepsilon^{\bar{a}\bar{b}\bar{c}} (\partial_{\bar{b}} F_{\bar{c}}) \hat{n}_{\bar{a}}, \quad (6.35)$$

thus leading to the expression of Stokes' theorem in arbitrary coordinates

$$\oint_C F_{\bar{b}} d\xi^{\bar{b}} = \int_S \varepsilon^{\bar{a}\bar{b}\bar{c}} (\partial_{\bar{b}} F_{\bar{c}}) \hat{n}_{\bar{a}} dS. \quad (6.36)$$

Orthogonal coordinates[†]

READER'S GUIDE TO THIS CHAPTER

This chapter compiles mathematical results for particular orthogonal coordinate choices, providing explicit examples using the machinery of general tensor analysis. Many results discussed here are used throughout this book. A thorough treatment can also be found in Section 21.11 of *Griffies (2004)*.

7.1	Cartesian coordinates	67
7.1.1	The basics	68
7.1.2	Summary of Cartesian coordinate expressions	68
7.2	Spherical coordinates	69
7.2.1	Transforming between Cartesian and spherical coordinates	70
7.2.2	Basis vectors	70
7.2.3	Basis one-forms	71
7.2.4	Position and velocity	71
7.2.5	Metric tensor	71
7.2.6	Components of a vector field	72
7.2.7	Differential operators	73
7.2.8	Summary of spherical coordinate expressions	73
7.3	Cylindrical-polar coordinates	73
7.3.1	Transforming between Cartesian and cylindrical-polar coordinates	75
7.3.2	Basis vectors	75
7.3.3	Basis one-forms	75
7.3.4	Position and velocity	76
7.3.5	Metric tensor	76
7.3.6	Components of a vector field	77
7.3.7	Differential operators	77
7.3.8	Summary of cylindrical coordinate expressions	78
7.4	General orthogonal coordinates	78

7.1 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 2 and 3.

7.1.1 The basics

We start by expressing the trajectory as

$$\mathcal{P}(\tau) = \vec{e}_1 x(\tau) + \vec{e}_2 y(\tau) + \vec{e}_3 z(\tau) = \hat{\mathbf{x}} x(\tau) + \hat{\mathbf{y}} y(\tau) + \hat{\mathbf{z}} z(\tau) = \vec{\mathbf{x}}(\tau), \quad (7.1)$$

with the basis vectors written

$$\vec{e}_1 = \hat{\mathbf{x}} \quad \vec{e}_2 = \hat{\mathbf{y}} \quad \vec{e}_3 = \hat{\mathbf{z}}. \quad (7.2)$$

The boldface notation is used for the position vector in the final equality of equation (7.1), with the boldface commonly used throughout this book. Notably, the orthogonal unit vectors for Cartesian coordinates are normalized so that

$$\vec{e}_1 \cdot \vec{e}_1 = \vec{e}_2 \cdot \vec{e}_2 = \vec{e}_3 \cdot \vec{e}_3 = 1. \quad (7.3)$$

Furthermore, the basis vectors are identical to the basis one-forms

$$\vec{e}_1 = \tilde{e}^1 = \hat{\mathbf{x}} \quad \vec{e}_2 = \tilde{e}^2 = \hat{\mathbf{y}} \quad \vec{e}_3 = \tilde{e}^3 = \hat{\mathbf{z}}. \quad (7.4)$$

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

$$\vec{v}(\tau) = \frac{d\mathcal{P}}{d\tau} \quad (7.5a)$$

$$= \frac{d\vec{\mathbf{x}}}{d\tau} \quad (7.5b)$$

$$= \vec{e}_1 \frac{dx(\tau)}{d\tau} + \vec{e}_2 \frac{dy(\tau)}{d\tau} + \vec{e}_3 \frac{dz(\tau)}{d\tau} \quad (7.5c)$$

$$= \hat{\mathbf{x}} v^1(\tau) + \hat{\mathbf{y}} v^2(\tau) + \hat{\mathbf{z}} v^3(\tau) \quad (7.5d)$$

$$= \vec{v}(\tau). \quad (7.5e)$$

7.1.2 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.

$$\mathbf{x} = (x^1, x^2, x^3) = (x, y, z) \quad \text{Cartesian coordinates} \quad (7.6)$$

$$\mathbf{F} = \hat{\mathbf{x}} F^1 + \hat{\mathbf{y}} F^2 + \hat{\mathbf{z}} F^3 = \hat{\mathbf{x}} F_1 + \hat{\mathbf{y}} F_2 + \hat{\mathbf{z}} F_3 \quad \text{covariant = contravariant} \quad (7.7)$$

$$\frac{\partial}{\partial x^a} = \partial_{x^a} \quad \text{or} \quad (\partial_x, \partial_y, \partial_z) \quad \text{partial derivative operator} \quad (7.8)$$

$$\nabla = \hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y + \hat{\mathbf{z}} \partial_z \quad \text{gradient operator} \quad (7.9)$$

$$\nabla_z = \hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y \quad \text{horizontal gradient operator} \quad (7.10)$$

$$\nabla \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y + \partial_z F_z \quad \text{divergence of a vector} \quad (7.11)$$

$$\nabla_z \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y \quad \text{horizontal divergence of a vector} \quad (7.12)$$

$$(\nabla \wedge \mathbf{F})_a = \epsilon_{abc} \partial_b F_c \quad \text{components to curl of a vector} \quad (7.13)$$

$$\int \nabla \cdot \mathbf{F} dV = \int \mathbf{F} \cdot \hat{\mathbf{n}} dS \quad \text{divergence (or Gauss's) theorem} \quad (7.14)$$

$$\oint_C \mathbf{F} \cdot d\mathbf{x} = \int_S (\nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad \text{Stokes' theorem.} \quad (7.15)$$

7.2 Spherical coordinates

We now consider spherical coordinates defined by Figure 7.1 and related to Cartesian coordinates through

$$x = r \cos \phi \cos \lambda \quad (7.16a)$$

$$y = r \cos \phi \sin \lambda \quad (7.16b)$$

$$z = r \sin \phi. \quad (7.16c)$$

The radial coordinate

$$r = \sqrt{x^2 + y^2 + z^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}} \quad (7.17)$$

measures the distance from the center of the sphere to position of the particle. The spherical angle coordinates

$$0 \leq \lambda \leq 2\pi \quad \text{longitude} \quad (7.18)$$

$$-\pi/2 \leq \phi \leq \pi/2 \quad \text{latitude} \quad (7.19)$$

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

$$(x, y, z) = (\xi^1, \xi^2, \xi^3) \equiv \xi^a \quad (7.20)$$

$$(\lambda, \phi, r) = (\bar{\xi}^1, \bar{\xi}^2, \bar{\xi}^3) \equiv \xi^{\bar{a}}. \quad (7.21)$$

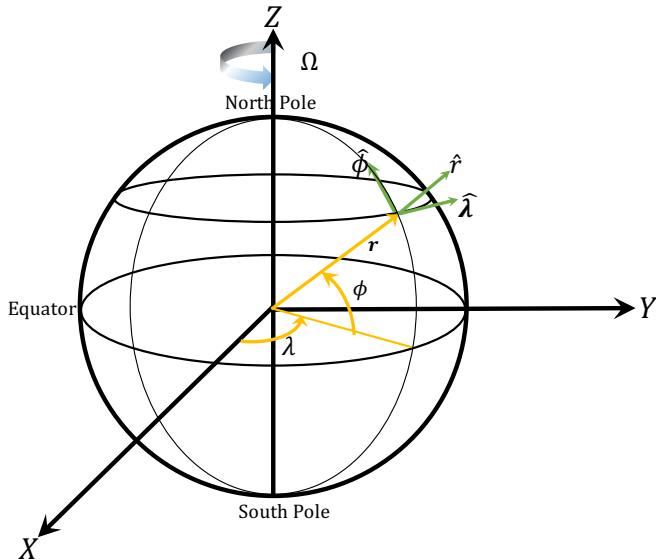


Figure 7.1: This schematic illustrates the geometry and notation for motion around a rotating sphere of radius R . For geophysical applications, the sphere rotates counter-clockwise when looking down from the north polar axis with angular speed Ω . The Cartesian triad of orthonormal basis vectors, $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$, points along the orthogonal axes. 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 point in the radial direction (positive away from the center).

7.2.1 Transforming between Cartesian and spherical coordinates

Following the general discussion in Section 5.1.4, 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

$$d\xi^a = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} d\xi^{\bar{a}} = \Lambda_{\bar{a}}^a d\xi^{\bar{a}}. \quad (7.22)$$

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 (7.16a)-(7.16c), this transformation matrix is given by

$$\Lambda_{\bar{a}}^a = \begin{bmatrix} \partial \xi^1 / \partial \xi^{\bar{1}} & \partial \xi^1 / \partial \xi^{\bar{2}} & \partial \xi^1 / \partial \xi^{\bar{3}} \\ \partial \xi^2 / \partial \xi^{\bar{1}} & \partial \xi^2 / \partial \xi^{\bar{2}} & \partial \xi^2 / \partial \xi^{\bar{3}} \\ \partial \xi^3 / \partial \xi^{\bar{1}} & \partial \xi^3 / \partial \xi^{\bar{2}} & \partial \xi^3 / \partial \xi^{\bar{3}} \end{bmatrix} = \begin{bmatrix} -r \cos \phi \sin \lambda & -r \sin \phi \cos \lambda & \cos \phi \cos \lambda \\ r \cos \phi \cos \lambda & -r \sin \phi \sin \lambda & \cos \phi \sin \lambda \\ 0 & r \cos \phi & \sin \phi \end{bmatrix}. \quad (7.23)$$

The determinant of the transformation (Jacobian) is given by

$$\det(\Lambda_{\bar{a}}^a) = r^2 \cos \phi. \quad (7.24)$$

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

$$\Lambda_{\bar{a}}^a = \frac{1}{r^2 \cos \phi} \begin{bmatrix} -r \sin \lambda & r \cos \lambda & 0 \\ -r \cos \phi \sin \phi \cos \lambda & -r \cos \phi \sin \phi \sin \lambda & r \cos^2 \phi \\ r^2 \cos^2 \phi \cos \lambda & r^2 \cos^2 \phi \sin \lambda & r^2 \cos \phi \sin \phi \end{bmatrix}. \quad (7.25)$$

7.2.2 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

$$\vec{e}_{\bar{a}} = \Lambda_{\bar{a}}^a \vec{e}_a. \quad (7.26)$$

The transformation matrix (7.23) leads to

$$\vec{e}_\lambda = r \cos \phi (-\hat{x} \sin \lambda + \hat{y} \cos \lambda) \quad (7.27a)$$

$$\vec{e}_\phi = r (-\hat{x} \sin \phi \cos \lambda - \hat{y} \sin \phi \sin \lambda + \hat{z} \cos \phi) \quad (7.27b)$$

$$\vec{e}_r = \hat{x} \cos \phi \cos \lambda + \hat{y} \cos \phi \sin \lambda + \hat{z} \sin \phi. \quad (7.27c)$$

We can introduce the orthonormal unit vectors through

$$\vec{e}_\lambda = r \cos \phi \hat{\lambda} \quad \vec{e}_\phi = r \hat{\phi} \quad \vec{e}_r = \hat{r}, \quad (7.28)$$

so that

$$\hat{\lambda} = -\hat{x} \sin \lambda + \hat{y} \cos \lambda \quad (7.29a)$$

$$\hat{\phi} = -\hat{x} \cos \lambda \sin \phi - \hat{y} \sin \lambda \sin \phi + \hat{z} \cos \phi \quad (7.29b)$$

$$\hat{r} = \hat{x} \cos \lambda \cos \phi + \hat{y} \sin \lambda \cos \phi + \hat{z} \sin \phi \quad (7.29c)$$

along with the inverse relations

$$\hat{x} = -\hat{\lambda} \sin \lambda - \hat{\phi} \cos \lambda \sin \phi + \hat{r} \cos \lambda \cos \phi \quad (7.30a)$$

$$\hat{y} = \hat{\lambda} \cos \lambda - \hat{\phi} \sin \lambda \sin \phi + \hat{r} \sin \lambda \cos \phi \quad (7.30b)$$

$$\hat{z} = \hat{\phi} \cos \phi + \hat{r} \sin \phi. \quad (7.30c)$$

7.2.3 Basis one-forms

Since spherical coordinates are orthogonal, we can readily derive the one-form basis through inverting the vector basis

$$\tilde{e}^\lambda = (r \cos \phi)^{-1} \hat{\lambda} \quad \tilde{e}^\phi = r^{-1} \hat{\phi} \quad \tilde{e}^r = \hat{r}, \quad (7.31)$$

which satisfy the orthogonality relation with the basis vectors (Section 5.2.2)

$$\tilde{e}^{\bar{b}} \cdot \vec{e}_{\bar{a}} = \delta^{\bar{b}}_{\bar{a}}. \quad (7.32)$$

7.2.4 Position and velocity

In spherical coordinates, the position of a point is fully specified by the radial position

$$\mathcal{P}(\tau) = r \vec{e}_r. \quad (7.33)$$

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

$$\vec{v}(\tau) = \frac{d\mathcal{P}}{d\tau} \quad (7.34a)$$

$$= \vec{e}_r \frac{dr}{d\tau} + r \frac{d\vec{e}_r}{d\tau} \quad (7.34b)$$

$$= \vec{e}_r \frac{dr}{d\tau} + r \frac{\partial \vec{e}_r}{\partial \lambda} \frac{d\lambda}{d\tau} + r \frac{\partial \vec{e}_r}{\partial \phi} \frac{d\phi}{d\tau} \quad (7.34c)$$

$$\equiv \vec{e}_r \frac{dr}{d\tau} + \vec{e}_\lambda \frac{d\lambda}{d\tau} + \vec{e}_\phi \frac{d\phi}{d\tau} \quad (7.34d)$$

$$= \vec{e}_r v^r + \vec{e}_\lambda v^\lambda + \vec{e}_\phi v^\phi. \quad (7.34e)$$

To reach this result we made use of the identities satisfied by the spherical basis vectors

$$\vec{e}_\lambda = r \frac{\partial \vec{e}_r}{\partial \lambda} \quad \vec{e}_\phi = r \frac{\partial \vec{e}_r}{\partial \phi}. \quad (7.35)$$

7.2.5 Metric tensor

The metric tensor for spherical coordinates takes on the diagonal form

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} (r \cos \phi)^2 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (7.36)$$

as does the spherical representation of the inverse metric tensor

$$g^{\bar{a}\bar{b}} = \tilde{e}^{\bar{a}} \cdot \tilde{e}^{\bar{b}} = \begin{bmatrix} (r \cos \phi)^{-2} & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (7.37)$$

Volume element and Levi-Civita tensor

The square root of the determinant of the metric tensor written in spherical coordinates (from equation (7.36)) is given by

$$\sqrt{\det(\bar{G})} = r^2 \cos \phi \quad (7.38)$$

so that the volume element is

$$dV = r^2 \cos \phi dr d\lambda d\phi. \quad (7.39)$$

The covariant Levi-Civita tensor has the spherical representation

$$\varepsilon_{\bar{a}\bar{b}\bar{c}} = (r^2 \cos \phi) \epsilon_{\bar{a}\bar{b}\bar{c}}. \quad (7.40)$$

Cross product of basis vectors

As a check on the formalism for cross products, let us verify the relation (5.56) for the cross product of two basis vectors using spherical coordinates

$$\vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}} \implies \vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = (r^2 \cos \phi) \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}}. \quad (7.41)$$

Making use of the spherical coordinate basis vectors and one-forms renders

$$\vec{e}_r \wedge \vec{e}_{\lambda} = (r \cos \phi) (\hat{r} \wedge \hat{\lambda}) = (r \cos \phi) \hat{\phi} = (r^2 \cos \phi) \tilde{e}^{\phi} = \varepsilon_{r\lambda\phi} \tilde{e}^{\phi} \quad (7.42a)$$

$$\vec{e}_{\lambda} \wedge \vec{e}_{\phi} = (r^2 \cos \phi) (\hat{\lambda} \wedge \hat{\phi}) = (r^2 \cos \phi) \hat{r} = (r^2 \cos \phi) \tilde{e}^r = \varepsilon_{\lambda\phi r} \tilde{e}^r \quad (7.42b)$$

$$\vec{e}_{\phi} \wedge \vec{e}_r = r (\hat{\phi} \wedge \hat{r}) = r \hat{\lambda} = (r^2 \cos \phi) \tilde{e}^{\lambda} = \varepsilon_{\phi r \lambda} \tilde{e}^{\lambda}. \quad (7.42c)$$

To reach these results we made use of the cross products for the spherical coordinate unit vectors

$$\hat{r} \wedge \hat{\lambda} = \hat{\phi} \quad \hat{\lambda} \wedge \hat{\phi} = \hat{r} \quad \hat{\phi} \wedge \hat{r} = \hat{\lambda}. \quad (7.43)$$

7.2.6 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

$$F^{\bar{1}} = (r \cos \phi)^{-1} [-F^x \sin \lambda + F^y \cos \lambda] \quad (7.44a)$$

$$F^{\bar{2}} = r^{-1} [-F^x \sin \phi \cos \lambda - F^y \sin \phi \sin \lambda + F^z \cos \phi] \quad (7.44b)$$

$$F^{\bar{3}} = F^x \cos \phi \cos \lambda + F^y \cos \phi \sin \lambda + F^z \sin \phi. \quad (7.44c)$$

Making use of the spherical unit vector (7.29a)-(7.29c) leads to the identities

$$(r \cos \phi) F^{\bar{1}} = \hat{\lambda} \cdot \mathbf{F} \quad (7.45a)$$

$$r F^{\bar{2}} = \hat{\phi} \cdot \mathbf{F} \quad (7.45b)$$

$$F^{\bar{3}} = \hat{r} \cdot \mathbf{F}. \quad (7.45c)$$

7.2.7 Differential operators

In spherical coordinates the gradient operator $\nabla = \tilde{e}^a \partial_a$ takes on the form

$$\nabla = \hat{\lambda} (r \cos \phi)^{-1} \partial_\lambda + \hat{\phi} r^{-1} \partial_\phi + \hat{r} \partial_r \quad (7.46)$$

and the covariant divergence of a vector field is given by

$$\nabla_{\bar{a}} F^{\bar{a}} = (r^2 \cos \phi)^{-1} \partial_{\bar{a}} [r^2 \cos \phi F^{\bar{a}}] \quad (7.47a)$$

$$= (r^2 \cos \phi)^{-1} \left(\partial_\lambda [r^2 \cos \phi F^{\bar{1}}] + \partial_\phi [r^2 \cos \phi F^{\bar{2}}] + \partial_r [r^2 \cos \phi F^{\bar{3}}] \right) \quad (7.47b)$$

$$= \frac{1}{r \cos \phi} \frac{\partial (\hat{\lambda} \cdot \mathbf{F})}{\partial \lambda} + \frac{1}{r \cos \phi} \frac{\partial (\hat{\phi} \cdot \mathbf{F} \cos \phi)}{\partial \phi} + \frac{1}{r^2} \frac{\partial (\hat{r} \cdot \mathbf{F} r^2)}{\partial r} \quad (7.47c)$$

The covariant curl (Section 6.9) takes the form

$$(\text{curl } \vec{F})^{\bar{1}} = (r^2 \cos \phi)^{-1} [\partial_\phi F^{\bar{3}} - \partial_r (r^2 F^{\bar{2}})] \quad (7.48a)$$

$$(\text{curl } \vec{F})^{\bar{2}} = (r^2 \cos \phi)^{-1} [\partial_r (r^2 \cos^2 \phi F^{\bar{1}}) - \partial_\lambda F^{\bar{3}}] \quad (7.48b)$$

$$(\text{curl } \vec{F})^{\bar{3}} = (r^2 \cos \phi)^{-1} [\partial_\lambda (r^2 F^{\bar{2}}) - \partial_\phi (r^2 \cos^2 \phi F^{\bar{1}})], \quad (7.48c)$$

which can be written in the more conventional form (e.g., equation (2.33) of [Vallis \(2017\)](#))

$$r \cos \phi (\text{curl } \vec{F})^{\bar{1}} = \frac{1}{r} \left[\frac{\partial (\hat{r} \cdot \mathbf{F})}{\partial \phi} - \frac{\partial (r \hat{\phi} \cdot \mathbf{F})}{\partial r} \right] \quad (7.49a)$$

$$r (\text{curl } \vec{F})^{\bar{2}} = \frac{1}{r} \left[\frac{\partial (r \hat{\lambda} \cdot \mathbf{F})}{\partial r} - \frac{1}{\cos \phi} \frac{\partial (\hat{r} \cdot \mathbf{F})}{\partial \lambda} \right] \quad (7.49b)$$

$$(\text{curl } \vec{F})^{\bar{3}} = \frac{1}{r \cos \phi} \left[\frac{\partial (\hat{\phi} \cdot \mathbf{F})}{\partial \lambda} - \frac{\partial (\cos \phi \hat{\lambda} \cdot \mathbf{F})}{\partial \phi} \right]. \quad (7.49c)$$

7.2.8 Summary of spherical coordinate expressions

We here summarize the spherical coordinate version of some common mathematical operators.

$$(\lambda, \phi, r) = (x^{\bar{1}}, x^{\bar{2}}, x^{\bar{3}}) \quad \text{spherical coordinates} \quad (7.50)$$

$$(r \cos \phi) F^{\bar{1}} = \hat{\lambda} \cdot \mathbf{F} \quad r F^{\bar{2}} = \hat{\phi} \cdot \mathbf{F} \quad F^{\bar{3}} = \hat{r} \cdot \mathbf{F} \quad \text{vector components} \quad (7.51)$$

$$\nabla = \hat{\lambda} (r \cos \phi)^{-1} \partial_\lambda + \hat{\phi} r^{-1} \partial_\phi + \hat{r} \partial_r \quad \text{gradient} \quad (7.52)$$

$$\nabla_{\bar{a}} F^{\bar{a}} = \frac{1}{r \cos \phi} \frac{\partial (\hat{\lambda} \cdot \mathbf{F})}{\partial \lambda} + \frac{1}{r \cos \phi} \frac{\partial (\hat{\phi} \cdot \mathbf{F} \cos \phi)}{\partial \phi} + \frac{1}{r^2} \frac{\partial (\hat{r} \cdot \mathbf{F} r^2)}{\partial r} \quad \text{divergence} \quad (7.53)$$

$$(\nabla \wedge \mathbf{F})_{\bar{a}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} F^{\bar{c}} \quad \text{see equations (7.49a) -- (7.49c)} \quad \text{curl of a vector} \quad (7.54)$$

7.3 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 7.2, here focusing on cylindrical-polar coordinates as shown in Figure 7.2. Our task here is somewhat simpler than for the spherical coordinates since the vertical or axial position, z , remains unchanged from its Cartesian value.

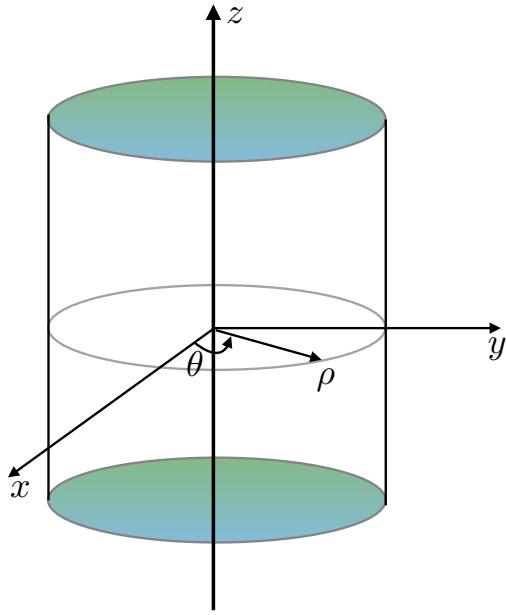


Figure 7.2: This schematic illustrates the geometry and notation for cylindrical-polar coordinates used to describe motion in a rotating laboratory tank. The Cartesian triad of orthonormal basis vectors, $(\hat{x}, \hat{y}, \hat{z})$ points along the orthogonal axes. The cylindrical-polar triad of orthonormal 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

$$x = \rho \cos \theta \equiv \xi^1 \cos \xi^2 \quad (7.55a)$$

$$y = \rho \sin \theta \equiv \xi^1 \sin \xi^2 \quad (7.55b)$$

$$z = \xi^3. \quad (7.55c)$$

The radial coordinate for cylindrical-polar coordinates

$$\rho = \sqrt{x^2 + y^2} \quad (7.56)$$

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

$$(x, y, z) = (\xi^1, \xi^2, \xi^3) \equiv \xi^a \quad (7.57)$$

$$(\rho, \theta, z) = (\xi^1, \xi^2, \xi^3) \equiv \xi^{\bar{a}}. \quad (7.58)$$

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.

7.3.1 Transforming between Cartesian and cylindrical-polar coordinates

The coordinate relation (7.55a)-(7.55c) leads to the transformation matrix

$$\Lambda_{\bar{a}}^a = \begin{bmatrix} \partial\xi^1/\partial\xi^{\bar{1}} & \partial\xi^1/\partial\xi^{\bar{2}} & \partial\xi^1/\partial\xi^{\bar{3}} \\ \partial\xi^2/\partial\xi^{\bar{1}} & \partial\xi^2/\partial\xi^{\bar{2}} & \partial\xi^2/\partial\xi^{\bar{3}} \\ \partial\xi^3/\partial\xi^{\bar{1}} & \partial\xi^3/\partial\xi^{\bar{2}} & \partial\xi^3/\partial\xi^{\bar{3}} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7.59)$$

and the inverse transformation is given by

$$\Lambda^{\bar{a}}_a = \frac{1}{\rho} \begin{bmatrix} \rho\cos\theta & \rho\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & \rho \end{bmatrix}. \quad (7.60)$$

The determinant of the transformation (Jacobian) is given by

$$\det(\Lambda_{\bar{a}}^a) = \rho, \quad (7.61)$$

which vanishes along the vertical axis, where the transformation is singular.

7.3.2 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 (7.59) leads to

$$\vec{e}_{\rho} = \hat{x}\cos\theta + \hat{y}\sin\theta \quad (7.62a)$$

$$\vec{e}_{\theta} = \rho(-\hat{x}\sin\theta + \hat{y}\cos\theta) \quad (7.62b)$$

$$\vec{e}_{\bar{z}} = \hat{z}. \quad (7.62c)$$

We sometimes make use of the following orthonormal unit vectors $(\hat{\rho}, \hat{\theta}, \hat{z})$

$$\vec{e}_{\rho} = \hat{\rho} \quad \vec{e}_{\theta} = \rho\hat{\theta} \quad \vec{e}_{\bar{z}} = \hat{z} \quad (7.63)$$

along with the inverse relations

$$\hat{x} = \hat{\rho}\cos\theta - \hat{\theta}\sin\theta \quad (7.64a)$$

$$\hat{y} = \hat{\rho}\sin\theta + \hat{\theta}\cos\theta \quad (7.64b)$$

$$\hat{z} = \hat{z}. \quad (7.64c)$$

7.3.3 Basis one-forms

Since cylindrical-polar coordinates are orthogonal, we can readily derive the one-form basis through inverting the vector basis

$$\tilde{e}^{\rho} = \hat{\rho} \quad \tilde{e}^{\theta} = \rho^{-1}\hat{\theta} \quad \tilde{e}^{\bar{z}} = \hat{z}, \quad (7.65)$$

which satisfy the orthogonality relation (Section 5.2.2)

$$\tilde{e}^{\bar{b}} \cdot \vec{e}_{\bar{a}} = \delta_{\bar{a}}^{\bar{b}}. \quad (7.66)$$

7.3.4 Position and velocity

In cylindrical-polar coordinates, the position of a point is specified by the radial position plus the vertical position

$$\mathcal{P}(\tau) = \rho \vec{e}_\rho + z \vec{e}_{\bar{z}}. \quad (7.67)$$

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

$$\vec{v}(\tau) = \frac{d\mathcal{P}}{d\tau} \quad (7.68a)$$

$$= \vec{e}_\rho \frac{d\rho}{d\tau} + \rho \frac{d\vec{e}_\rho}{d\tau} + \vec{e}_{\bar{z}} \frac{dz}{d\tau} \quad (7.68b)$$

$$= \vec{e}_\rho \frac{d\rho}{d\tau} + \rho \frac{\partial \vec{e}_\rho}{\partial \theta} \frac{d\theta}{d\tau} + \vec{e}_{\bar{z}} \frac{dz}{d\tau} \quad (7.68c)$$

$$= \vec{e}_\rho \frac{d\rho}{d\tau} + \vec{e}_\theta \frac{d\theta}{d\tau} + \vec{e}_{\bar{z}} \frac{dz}{d\tau} \quad (7.68d)$$

$$= \vec{e}_\rho v^\rho + \vec{e}_\theta v^\theta + \vec{e}_{\bar{z}} v^{\bar{z}}. \quad (7.68e)$$

To reach this result we made use of the identity

$$\vec{e}_\theta = r \frac{\partial \vec{e}_\rho}{\partial \theta}. \quad (7.69)$$

7.3.5 Metric tensor

The metric tensor for cylindrical-polar coordinates takes on the diagonal form

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (7.70)$$

as does the inverse metric tensor

$$g^{\bar{a}\bar{b}} = \vec{e}^{\bar{a}} \cdot \vec{e}^{\bar{b}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^{-2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (7.71)$$

Volume element and Levi-Civita tensor

The square root of the determinant of the metric tensor written in cylindrical-polar coordinates (from equation (7.70)) is given by

$$\sqrt{\det(\mathcal{G})} = \rho \quad (7.72)$$

so that the volume element is

$$dV = \rho d\rho d\theta dz. \quad (7.73)$$

The covariant Levi-Civita tensor has the cylindrical-polar representation

$$\varepsilon_{\bar{a}\bar{b}\bar{c}} = \rho \epsilon_{\bar{a}\bar{b}\bar{c}}. \quad (7.74)$$

Cross product of basis vectors

As a check on the formalism for cross products, let us verify the relation (5.56) for the cross product of two basis vectors using spherical coordinates

$$\vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}} \implies \vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \rho \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}}. \quad (7.75)$$

Making use of the cylindrical-polar coordinate basis vectors and one-forms renders

$$\vec{e}_\rho \wedge \vec{e}_\theta = \rho (\hat{\rho} \wedge \hat{\theta}) = \rho \tilde{e}^{\bar{z}} = \epsilon_{\rho\theta\bar{z}} \tilde{e}^{\bar{z}} \quad (7.76a)$$

$$\vec{e}_\theta \wedge \vec{e}_{\bar{z}} = \rho (\hat{\theta} \wedge \hat{z}) = \rho \hat{\rho} = \rho \tilde{e}^\rho = \epsilon_{\theta\bar{z}\rho} \tilde{e}^\rho \quad (7.76b)$$

$$\vec{e}_{\bar{z}} \wedge \vec{e}_\rho = \hat{z} \wedge \hat{\rho} = \hat{\theta} = \rho \tilde{e}^\theta = \epsilon_{\bar{z}\rho\theta} \tilde{e}^\theta. \quad (7.76c)$$

To reach these results we made use of the cross products for the unit vectors

$$\hat{\rho} \wedge \hat{\theta} = \hat{z} \quad \hat{\theta} \wedge \hat{z} = \hat{\rho} \quad \hat{z} \wedge \hat{\rho} = \hat{\theta}. \quad (7.77)$$

7.3.6 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

$$F^{\bar{1}} = F^x \cos \theta + F^y \sin \theta \quad (7.78a)$$

$$F^{\bar{2}} = \rho^{-1} [-F^x \sin \theta + F^y \cos \theta] \quad (7.78b)$$

$$F^{\bar{3}} = F^z. \quad (7.78c)$$

Introducing the cylindrical-polar unit vectors (7.63) leads to

$$F^{\bar{1}} = \hat{\rho} \cdot \mathbf{F} \quad (7.79a)$$

$$\rho F^{\bar{2}} = \hat{\theta} \cdot \mathbf{F} \quad (7.79b)$$

$$F^{\bar{3}} = \hat{z} \cdot \mathbf{F}. \quad (7.79c)$$

7.3.7 Differential operators

In cylindrical-polar coordinates, the gradient operator $\nabla = \tilde{e}^a \partial_a$ takes on the form

$$\nabla = \hat{\rho} \frac{\partial}{\partial \rho} + \frac{\hat{\theta}}{\rho} \frac{\partial}{\partial \theta} + \hat{z} \frac{\partial}{\partial z} \quad (7.80)$$

and the covariant divergence of a vector field is given by

$$\nabla_{\bar{a}} F^{\bar{a}} = \rho^{-1} \partial_{\bar{a}} (\rho F^{\bar{a}}) \quad (7.81a)$$

$$= \rho^{-1} \left(\partial_\rho [\rho F^{\bar{1}}] + \partial_\theta [\rho F^{\bar{2}}] + \partial_z [\rho F^{\bar{3}}] \right) \quad (7.81b)$$

$$= \frac{1}{\rho} \frac{\partial (\rho \hat{\rho} \cdot \mathbf{F})}{\partial \rho} + \frac{1}{\rho} \frac{\partial (\hat{\theta} \cdot \mathbf{F})}{\partial \theta} + \frac{\partial (\hat{z} \cdot \mathbf{F})}{\partial z}. \quad (7.81c)$$

The covariant curl (Section 6.9) takes the form

$$(\text{curl} \vec{F})^{\bar{1}} = \rho^{-1} [\partial_{\theta} F^{\bar{3}} - \partial_{\bar{z}} (\rho^2 F^{\bar{2}})] \quad (7.82a)$$

$$(\text{curl} \vec{F})^{\bar{2}} = \rho^{-1} [\partial_{\bar{z}} F^{\bar{1}} - \partial_{\rho} F^{\bar{3}}] \quad (7.82b)$$

$$(\text{curl} \vec{F})^{\bar{3}} = \rho^{-1} [\partial_{\rho} (\rho^2 F^{\bar{2}}) - \partial_{\theta} F^{\bar{1}}], \quad (7.82c)$$

which can be written in the more conventional form

$$(\text{curl} \vec{F})^{\bar{1}} = \frac{1}{\rho} \frac{\partial(\hat{\mathbf{z}} \cdot \mathbf{F})}{\partial \theta} - \frac{\partial(\hat{\boldsymbol{\theta}} \cdot \mathbf{F})}{\partial z} \quad (7.83a)$$

$$\rho (\text{curl} \vec{F})^{\bar{2}} = \frac{\partial(\hat{\boldsymbol{\rho}} \cdot \mathbf{F})}{\partial z} - \frac{\partial(\hat{\mathbf{z}} \cdot \mathbf{F})}{\partial z} \quad (7.83b)$$

$$(\text{curl} \vec{F})^{\bar{3}} = \frac{1}{\rho} \frac{\partial(\rho \hat{\boldsymbol{\theta}} \cdot \mathbf{F})}{\partial \rho} - \frac{1}{\rho} \frac{\partial(\hat{\boldsymbol{\rho}} \cdot \mathbf{F})}{\partial \theta}. \quad (7.83c)$$

7.3.8 Summary of cylindrical coordinate expressions

We here summarize the cylindrical coordinate version of some common mathematical operators.

$$(\rho, \theta, z) = (x^{\bar{1}}, x^{\bar{2}}, x^{\bar{3}}) \quad \text{cylindrical coordinates} \quad (7.84)$$

$$F^{\bar{1}} = \hat{\boldsymbol{\rho}} \cdot \mathbf{F} \quad \rho F^{\bar{2}} = \hat{\boldsymbol{\theta}} \cdot \mathbf{F} \quad F^{\bar{3}} = \hat{\mathbf{z}} \cdot \mathbf{F} \quad \text{vector components} \quad (7.85)$$

$$\nabla = \hat{\boldsymbol{\rho}} \frac{\partial}{\partial \rho} + \frac{\hat{\boldsymbol{\theta}}}{\rho} \frac{\partial}{\partial \theta} + \hat{\mathbf{z}} \frac{\partial}{\partial z} \quad \text{gradient} \quad (7.86)$$

$$\nabla_{\bar{a}} F^{\bar{a}} = \frac{1}{\rho} \frac{\partial(\rho \hat{\boldsymbol{\rho}} \cdot \mathbf{F})}{\partial \rho} + \frac{1}{\rho} \frac{\partial(\hat{\boldsymbol{\theta}} \cdot \mathbf{F})}{\partial \theta} + \frac{\partial(\hat{\mathbf{z}} \cdot \mathbf{F})}{\partial z}. \quad \text{divergence} \quad (7.87)$$

$$(\nabla \wedge \mathbf{F})_{\bar{a}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} F^{\bar{c}} \quad \text{see equations (7.82a) -- (7.82c).} \quad \text{curl of a vector} \quad (7.88)$$

7.4 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

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} h_{\bar{1}} & 0 & 0 \\ 0 & h_{\bar{2}} & 0 \\ 0 & 0 & h_{\bar{3}} \end{bmatrix}, \quad (7.89)$$

where $h_{\bar{a}} > 0$ are “stretching” functions. The corresponding volume element is expressed as

$$dV = h_{\bar{1}} h_{\bar{2}} h_{\bar{3}} d\xi^{\bar{1}} d\xi^{\bar{2}} d\xi^{\bar{3}}. \quad (7.90)$$

These *generalized orthogonal curvilinear coordinates* have a corresponding orthogonal set of basis vectors

$$\vec{e}_{\bar{a}} = h_{\bar{a}} \hat{\mathbf{e}}_{(\bar{a})} \quad \text{no implied sum.} \quad (7.91)$$

The objects $\hat{\mathbf{e}}_{(\bar{a})}$ are the dimensionless unit directions. The corresponding one-form basis is given by

$$\tilde{e}^{\bar{a}} = (h_{\bar{a}})^{-1} \hat{\mathbf{e}}_{(\bar{a})}. \quad (7.92)$$

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 7.2 and cylindrical-polar results from Section 7.3. A thorough examination of these coordinates is found in Section 21.11 of *Griffies* (2004).

8

Generalized vertical coordinates[†]

Generalized vertical coordinates (GVCs) are non-orthogonal coordinates commonly used to describe stratified geophysical fluids. 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 an increased need to understand their mathematics. Much of the material on GVCs in this book is based on Chapter 6 of [Griffies \(2004\)](#) and [Young \(2012\)](#).

READER'S GUIDE TO THIS CHAPTER

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 for the orthogonal coordinates considered in Chapter 7. Particular chapters directly relying on material in this chapter include the fluid kinematics discussed in Chapter 18, the tracer equation diffusion and stirring operators discussed in Chapter 24, and the dynamics discussed in Chapter 31. These related chapters are all part of the book's Tier-II material.

8.1	Introducing GVCs	82
8.1.1	Relating Cartesian and generalized vertical coordinates	83
8.1.2	A common confusion	84
8.2	Spatial basis vectors	84
8.2.1	A few more points on the transformation matrix	85
8.2.2	Expressions for the basis vectors	85
8.3	Basis one-forms	85
8.3.1	A few more points on the inverse transformation matrix	86
8.3.2	GVC basis one-forms	86
8.3.3	Verifying the orthogonality relation	86
8.4	Showing that $\Lambda^{\bar{a}}_a \Lambda^a_{\bar{b}} = \delta^{\bar{a}}_{\bar{b}}$ and $\Lambda^a_{\bar{a}} \Lambda^{\bar{a}}_b = \delta^a_b$	86
8.5	Triple product identities	87
8.6	Position vector	88
8.7	Transforming components of a first order tensor	89
8.8	Velocity	90
8.8.1	Contravariant components to the velocity	90
8.8.2	Covariant components to the velocity vector	91
8.8.3	Introducing the material time derivative	91
8.8.4	Equivalence to the Cartesian velocity representation	91
8.8.5	Comments	91
8.9	Metric tensor	92
8.9.1	Jacobian of transformation	92
8.9.2	Covariant and contravariant representations	92
8.10	Volume element and the Levi-Civita tensor	93
8.11	Cross product of basis vectors	93
8.12	Partial derivative operators	93
8.12.1	Analytical derivation	94
8.12.2	Geometrical derivation	94
8.13	Material time derivative	95
8.14	Divergence of a vector and the divergence theorem	95
8.15	The diffusion operator	96
8.16	Vorticity	97
8.16.1	The components	98
8.16.2	Transforming from Cartesian coordinates	98
8.17	Circulation	98

8.1 Introducing GVCs

In this chapter we present the mathematics of generalized vertical coordinates (GVC), with Figure 8.1 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 (30.1a) and the mass continuity equation discussed in Section 18.8.2. 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 isopycnal 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 31.4. The ALE method offers the means to evolve the model state even without an explicit specification of the vertical coordinate.

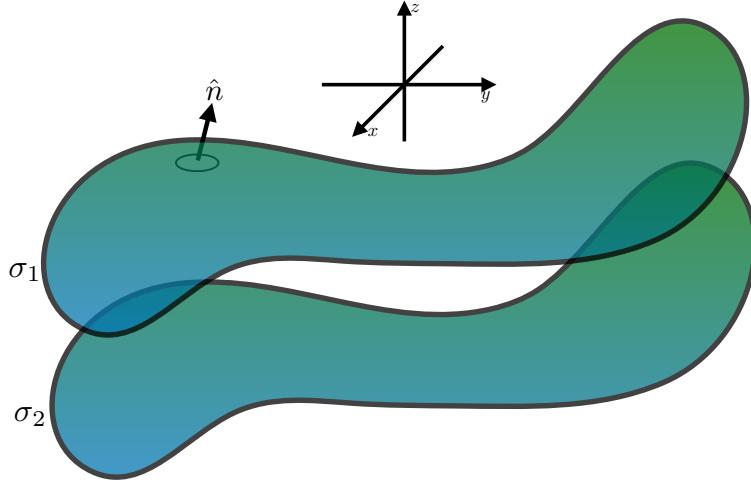


Figure 8.1: This schematic illustrates the geometry of two surfaces of constant generalized vertical coordinate $\sigma(x, y, z, t) = \sigma_1$ and $\sigma(x, y, z, t) = \sigma_2$, here showing patches on two such surfaces. The surfaces are generally undulating in space and time yet are assumed to monotonically 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 .

8.1.1 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 the GVC 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 8.12.

To help keep track of the two coordinate systems, we write the time coordinate and spatial Cartesian coordinates according to

$$\xi^\alpha = (\xi^0, \xi^a) = (\xi^0, \xi^1, \xi^2, \xi^3) = (t, x, y, z) \text{ with } \alpha = 0, 1, 2, 3, \text{ and } a = 1, 2, 3. \quad (8.1)$$

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 8.12).

The corresponding generalized vertical coordinates are denoted with an overbar

$$\bar{\xi}^{\bar{\alpha}} = (\bar{\xi}^0, \bar{\xi}^1, \bar{\xi}^2, \bar{\xi}^3) = (\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (8.2)$$

The one-to-one coordinate transformation between Cartesian and GVC coordinates is written

$$\xi^{\bar{0}} = \xi^0 \iff \bar{t} = t \quad (8.3a)$$

$$\xi^{\bar{1}} = \xi^1 \iff \bar{x} = x \quad (8.3b)$$

$$\xi^{\bar{2}} = \xi^2 \iff \bar{y} = y \quad (8.3c)$$

$$\xi^{\bar{3}} = \sigma(t, x, y, z). \quad (8.3d)$$

Since the coordinate transformation is invertible, we can define the inverse to equation (8.3d), thus providing an expression for the vertical position of a given GVC surface

$$\xi^3 = \xi^3(\xi^{\bar{a}}) = z(\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (8.4)$$

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

$$\xi^0 = \xi^{\bar{0}} \quad (8.5a)$$

$$\xi^1 = \xi^{\bar{1}} \quad (8.5b)$$

$$\xi^2 = \xi^{\bar{2}} \quad (8.5c)$$

$$\xi^3 = z(\bar{t}, \bar{x}, \bar{y}, \sigma). \quad (8.5d)$$

8.1.2 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 for 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, after we have developed our GVC brain muscle we are more relaxed when the meaning is clear.

8.2 Spatial basis vectors

Making use of the tensor formalism from Chapters 5 and 6, consider the transformation of the Cartesian basis vectors into their corresponding GVC representation. This transformation takes on the form

$$\vec{e}_{\bar{a}} = \Lambda_{\bar{a}}^a \vec{e}_a, \quad (8.6)$$

where the transformation matrix takes the form

$$\Lambda_{\bar{a}}^a = \begin{bmatrix} \partial x / \partial \bar{x} & \partial x / \partial \bar{y} & \partial x / \partial \sigma \\ \partial y / \partial \bar{x} & \partial y / \partial \bar{y} & \partial y / \partial \sigma \\ \partial z / \partial \bar{x} & \partial z / \partial \bar{y} & \partial z / \partial \sigma \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial z / \partial \bar{x} & \partial z / \partial \bar{y} & \partial z / \partial \sigma \end{bmatrix}. \quad (8.7)$$

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.

8.2.1 A few more points on the transformation matrix

To further detail how to produce elements of the transformation matrix (8.7), 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

$$\Lambda_{\bar{a}}^1 = [\begin{array}{ccc} [\partial x/\partial \bar{x}]_{\bar{y}, \sigma} & [\partial x/\partial \bar{y}]_{\bar{x}, \sigma} & [\partial x/\partial \sigma]_{\bar{x}, \bar{y}} \end{array}] \quad (8.8)$$

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).

8.2.2 Expressions for the basis vectors

Use of the transformation matrix (8.7) renders the spatial components of the GVC basis vectors

$$\vec{e}_{\bar{1}} = \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial \bar{x}) \quad (8.9a)$$

$$\vec{e}_{\bar{2}} = \hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial \bar{y}) \quad (8.9b)$$

$$\vec{e}_{\bar{3}} = \hat{\mathbf{z}} (\partial z / \partial \sigma). \quad (8.9c)$$

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 8.2 illustrates the basis vectors.

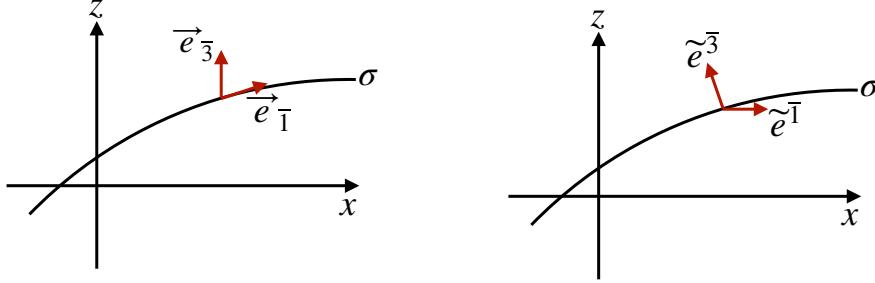


Figure 8.2: Illustrating the basis vectors (left panel) and basis one-forms (right panel) for generalized vertical coordinates. The $\vec{e}_{\bar{3}}$ basis vector is vertical whereas $\vec{e}_{\bar{1}}$ and $\vec{e}_{\bar{2}}$ lie within the tangent plane to the σ surface. As a complement, the basis 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.

8.3 Basis one-forms

The basis one-forms are obtained by transforming from Cartesian into GVCs through use of the inverse transformation

$$\tilde{e}^{\bar{a}} = \Lambda_a^{\bar{a}} \tilde{e}^a, \quad (8.10)$$

where the inverse transformation matrix takes the form

$$\Lambda_{\bar{a}}^{\bar{a}} = \begin{bmatrix} \partial \bar{x}/\partial x & \partial \bar{x}/\partial y & \partial \bar{x}/\partial z \\ \partial \bar{y}/\partial x & \partial \bar{y}/\partial y & \partial \bar{y}/\partial z \\ \partial \sigma/\partial x & \partial \sigma/\partial y & \partial \sigma/\partial z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial \sigma/\partial x & \partial \sigma/\partial y & \partial \sigma/\partial z \end{bmatrix}. \quad (8.11)$$

As for the transformation matrix (8.7), 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.

8.3.1 A few more points on the inverse transformation matrix

When computing elements of the inverse transformation matrix (8.11), it is crucial to ensure that the proper variables are held fixed. For example, consider the top row where we compute

$$\bar{\Lambda}_{\bar{a}}^{\bar{1}} = [[\partial \bar{x}/\partial x]_{y,z} \quad [\partial \bar{x}/\partial y]_{x,z} \quad [\partial \bar{x}/\partial z]_{x,y}]. \quad (8.12)$$

Just as for the transformation matrix (8.8), since $x = \bar{x}$, all but the first element vanish in equation (8.12). 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}).

8.3.2 GVC basis one-forms

Use of the inverse transformation matrix (8.11) renders the spatial components of the GVC basis one-forms

$$\tilde{e}^{\bar{1}} = \hat{x} \quad (8.13a)$$

$$\tilde{e}^{\bar{2}} = \hat{y} \quad (8.13b)$$

$$\tilde{e}^{\bar{3}} = \tilde{e}^a \partial_a \sigma = \hat{x} (\partial \sigma / \partial x) + \hat{y} (\partial \sigma / \partial y) + \hat{z} (\partial \sigma / \partial z) = \nabla \sigma. \quad (8.13c)$$

The left panel of Figure 8.2 illustrates the basis one-forms.

8.3.3 Verifying the orthogonality relation

The basis one-forms satisfy the orthogonality relation (5.18) with the basis vectors

$$\tilde{e}^{\bar{a}} \cdot \tilde{e}_{\bar{b}} = \delta_{\bar{b}}^{\bar{a}}. \quad (8.14)$$

This identity is trivial to verify for all $\bar{a} = 1, 2, 3$.

8.4 Showing that $\Lambda_{\bar{a}}^{\bar{a}} \Lambda_{\bar{b}}^a = \delta_{\bar{b}}^{\bar{a}}$ and $\Lambda_{\bar{a}}^a \Lambda_{\bar{b}}^{\bar{a}} = \delta_{\bar{b}}^a$

We present two arguments to verify that the matrix (8.11) is indeed the inverse of the matrix (8.7). Both rely on writing the coordinate transformation as a composite function

$$\xi^{\bar{a}} = \xi^{\bar{a}}(\xi^a) = \xi^{\bar{a}}[\xi^a(\xi^{\bar{b}})]. \quad (8.15)$$

Taking partial derivatives and using the chain rule renders

$$\delta_{\bar{b}}^{\bar{a}} = \frac{\partial \xi^{\bar{a}}}{\partial \xi^{\bar{b}}} = \frac{\partial \xi^{\bar{a}}}{\partial \xi^a} \frac{\partial \xi^a}{\partial \xi^{\bar{b}}} = \Lambda_a^{\bar{a}} \Lambda_b^a \quad (8.16)$$

as well as

$$\delta_b^a = \frac{\partial \xi^a}{\partial \xi^b} = \frac{\partial \xi^a}{\partial \xi^{\bar{a}}} \frac{\partial \xi^{\bar{a}}}{\partial \xi^b} = \Lambda_{\bar{a}}^a \Lambda_b^{\bar{a}} \quad (8.17)$$

8.5 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 8.4

$$\sigma = \sigma(t, x, y, z) = \sigma[t, x, y, z(\bar{t}, \bar{x}, \bar{y}, \sigma)], \quad (8.18)$$

so that the chain rule leads to the differential increment

$$d\sigma = dt \left[\frac{\partial \sigma}{\partial t} \right]_{x,y,z} + dx \left[\frac{\partial \sigma}{\partial x} \right]_{t,y,z} + dy \left[\frac{\partial \sigma}{\partial y} \right]_{t,x,z} + dz \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y}. \quad (8.19)$$

Likewise, writing $z = z[\bar{t}, \bar{x}, \bar{y}, \sigma]$ leads to the differential increment dz

$$dz = d\bar{t} \left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} + d\bar{x} \left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} + d\bar{y} \left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} + d\sigma \left[\frac{\partial z}{\partial \sigma} \right]_{\bar{t}, \bar{x}, \bar{y}}. \quad (8.20)$$

We note the identities

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \sigma} \right]_{\bar{t}, \bar{x}, \bar{y}} = 1 \quad d\bar{t} = dt \quad d\bar{x} = dx \quad d\bar{y} = dy, \quad (8.21)$$

which follow since $t = \bar{t}$, $x = \bar{x}$, and $y = \bar{y}$. Substituting equation (8.20) into equation (8.19) and making use of the identities (8.21) yields

$$\begin{aligned} 0 &= dt \left[\left[\frac{\partial \sigma}{\partial t} \right]_{x,y,z} + \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} \right] \\ &\quad + dx \left[\left[\frac{\partial \sigma}{\partial x} \right]_{t,y,z} + \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} \right] + dy \left[\left[\frac{\partial \sigma}{\partial y} \right]_{t,x,z} + \left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} \right]. \end{aligned} \quad (8.22)$$

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

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} = - \left[\frac{\partial \sigma}{\partial t} \right]_{x,y,z} \quad (8.23a)$$

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} = - \left[\frac{\partial \sigma}{\partial x} \right]_{t,y,z} \quad (8.23b)$$

$$\left[\frac{\partial \sigma}{\partial z} \right]_{t,x,y} \left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} = - \left[\frac{\partial \sigma}{\partial y} \right]_{t,x,z} \quad (8.23c)$$

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

$$\left[\frac{\partial z}{\partial \bar{t}} \right]_{\bar{x}, \bar{y}, \sigma} = - \frac{[\partial\sigma/\partial t]_{x,y,z}}{[\partial\sigma/\partial z]_{t,x,y}} = - \left[\frac{\partial\sigma}{\partial t} \right]_{x,y,z} \left[\frac{\partial z}{\partial\sigma} \right]_{t,x,y} \quad (8.24a)$$

$$\left[\frac{\partial z}{\partial \bar{x}} \right]_{\bar{t}, \bar{y}, \sigma} = - \frac{[\partial\sigma/\partial x]_{t,y,z}}{[\partial\sigma/\partial z]_{t,x,y}} = - \left[\frac{\partial\sigma}{\partial x} \right]_{t,y,z} \left[\frac{\partial z}{\partial\sigma} \right]_{t,x,y} \quad (8.24b)$$

$$\left[\frac{\partial z}{\partial \bar{y}} \right]_{\bar{t}, \bar{x}, \sigma} = - \frac{[\partial\sigma/\partial y]_{t,x,z}}{[\partial\sigma/\partial z]_{t,x,y}} = - \left[\frac{\partial\sigma}{\partial y} \right]_{t,x,z} \left[\frac{\partial z}{\partial\sigma} \right]_{t,x,y}. \quad (8.24c)$$

Since $t = \bar{t}$, $x = \bar{x}$, and $y = \bar{y}$ we can write these identities in the more succinct form

$$\left[\frac{\partial z}{\partial \bar{t}} \right]_\sigma = - \frac{[\partial\sigma/\partial t]_z}{[\partial\sigma/\partial z]} \quad (8.25a)$$

$$\left[\frac{\partial z}{\partial \bar{x}} \right]_\sigma = - \frac{[\partial\sigma/\partial x]_z}{[\partial\sigma/\partial z]} \quad (8.25b)$$

$$\left[\frac{\partial z}{\partial \bar{y}} \right]_\sigma = - \frac{[\partial\sigma/\partial y]_z}{[\partial\sigma/\partial z]}. \quad (8.25c)$$

These identities are quite useful for manipulating equations involving GVCs. In particular, equations (8.25b) and (8.25c) provide alternate expressions for the slope of constant σ surfaces relative to the horizontal plane (see Section 8.12).

8.6 Position vector

We are familiar with locating a point in space using Cartesian coordinates as in Figure 2.1. What about specifying the position using GVCs? We can do so by making use of the basis vectors (8.9a)-(8.9c) so that the position of an arbitrary point in space is given by

$$\mathcal{P} = \xi^{\bar{a}} \vec{e}_{\bar{a}} \quad (8.26a)$$

$$= \bar{x} [\hat{\mathbf{x}} + (\partial z / \partial \bar{x}) \hat{\mathbf{z}}] + \bar{y} [\hat{\mathbf{y}} + (\partial z / \partial \bar{y}) \hat{\mathbf{z}}] + \sigma (\partial z / \partial \sigma) \hat{\mathbf{z}} \quad (8.26b)$$

$$= \hat{\mathbf{x}} \bar{x} + \hat{\mathbf{y}} \bar{y} + \hat{\mathbf{z}} [\bar{x} (\partial z / \partial \bar{x}) + \bar{y} (\partial z / \partial \bar{y}) + \sigma (\partial z / \partial \sigma)] \quad (8.26c)$$

$$= \hat{\mathbf{x}} \bar{x} + \hat{\mathbf{y}} \bar{y} + \hat{\mathbf{z}} \xi^{\bar{a}} \partial_{\bar{a}} z. \quad (8.26d)$$

We identify the following properties as a means to help understand these expressions, with Figure 8.3 offering a schematic.

- The form (8.26b) 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 8.3. 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

$$\mathcal{P} = \bar{x} \hat{\mathbf{x}} + \hat{\mathbf{z}} [\bar{x} (\partial z / \partial \bar{x}) + \sigma (\partial z / \partial \sigma)] \quad (8.27a)$$

$$= \bar{x} \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial \sigma) [\bar{x} (\partial \sigma / \partial z)_x (\partial z / \partial \bar{x})_\sigma + \sigma] \quad (8.27b)$$

$$= \bar{x} \hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial \sigma) [-\bar{x} (\partial \sigma / \partial x)_z + \sigma], \quad (8.27c)$$

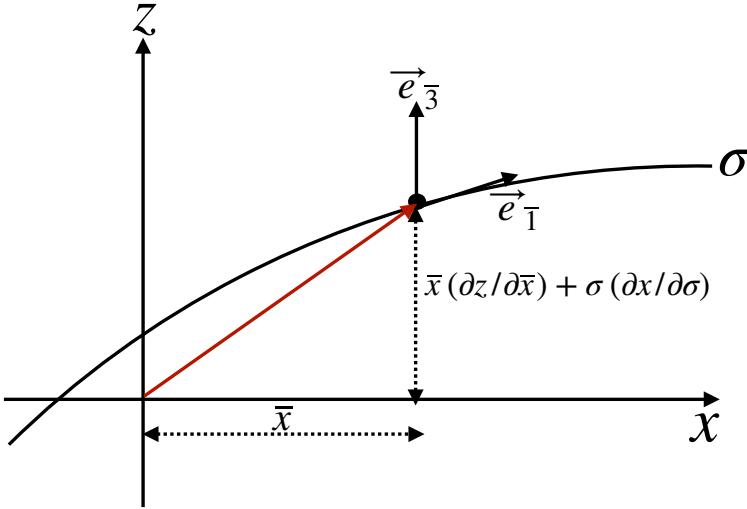


Figure 8.3: The position of a point in space as represented using GVCs following equation (8.27a). For this example, $\bar{y} = 0$ so that the horizontal position is determined by the coordinate $\bar{x} = x$, whereas the vertical position is determined by $\bar{x}(\partial z/\partial \bar{x}) + \sigma(\partial z/\partial \sigma)$.

where we used the triple product identity (8.25b) 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

$$\mathcal{P} \cdot \tilde{e}^{\bar{b}} = \xi^{\bar{a}} \tilde{e}_{\bar{a}} \cdot \tilde{e}^{\bar{b}} = \xi^{\bar{b}}. \quad (8.28)$$

This result follows from the orthogonality relation (5.18). So the projection of the position vector onto a basis one-form picks out the corresponding coordinate value.

- Equation (8.4) provides the spatial dependence for the vertical position of the surface of constant GVC

$$z = z(\xi^{\bar{a}}). \quad (8.29)$$

At any particular time instance we can perform a Taylor series about a reference depth z_0 , so that

$$z(\xi^{\bar{a}}) \approx z_0 + \xi^{\bar{a}} \partial_{\bar{a}} z. \quad (8.30)$$

We can thus write the position (8.26d) in the form

$$\mathcal{P} = \hat{x} \bar{x} + \hat{y} \bar{y} + \hat{z} [z - z_0]. \quad (8.31)$$

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.

8.7 Transforming components of a first order tensor

Consider a vector field \vec{F} with Cartesian representation

$$\vec{F} = \mathbf{F} = F^a \vec{e}_a = F^x \hat{x} + F^y \hat{y} + F^z \hat{z}. \quad (8.32)$$

The corresponding GVC components are related through the transformation matrix

$$F^{\bar{a}} = \Lambda_{\bar{a}}^{\bar{a}} F^a. \quad (8.33)$$

Making use of the transformation matrix (8.11) yields the relations between GVC components and Cartesian components

$$F^{\bar{1}} = F^1 \quad F^{\bar{2}} = F^2 \quad F^{\bar{3}} = \nabla \sigma \cdot \mathbf{F}, \quad (8.34)$$

where we wrote

$$\nabla \sigma \cdot \mathbf{F} = (\partial \sigma / \partial x) F^1 + (\partial \sigma / \partial y) F^2 + (\partial \sigma / \partial z) F^3. \quad (8.35)$$

The vector field thus can be represented in GVC coordinates as

$$\vec{F} = F^{\bar{a}} \vec{e}_{\bar{a}} = F^1 \vec{e}_{\bar{1}} + F^2 \vec{e}_{\bar{2}} + (\nabla \sigma \cdot \mathbf{F}) \vec{e}_{\bar{3}}. \quad (8.36)$$

Similarly, the covariant components transform as $F_{\bar{a}} = \Lambda_{\bar{a}}^a F_a$, where use of the inverse transformation matrix (8.11) renders

$$F_{\bar{1}} = F_1 + \frac{\partial z}{\partial \bar{x}} F_3 \quad F_{\bar{2}} = F_2 + \frac{\partial z}{\partial \bar{y}} F_3 \quad F_{\bar{3}} = \frac{\partial z}{\partial \sigma} F_3, \quad (8.37)$$

and the expression for the vector field

$$\vec{F} = F_{\bar{a}} \vec{e}^{\bar{a}} = [F_1 + (\partial z / \partial \bar{x}) F_3] \vec{e}^{\bar{1}} + [F_2 + (\partial z / \partial \bar{y}) F_3] \vec{e}^{\bar{2}} + (\partial z / \partial \sigma) F_3 \vec{e}^{\bar{3}}. \quad (8.38)$$

Recall also that for Cartesian coordinates the contravariant and covariant components to a vector are identical: $F^a = F_a$.

8.8 Velocity

We now make use of the results from Section 8.7 to represent the velocity vector, considering both covariant and contravariant representations. As for the position vector detailed in Section 8.6, 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 8.8.4 we verify that the transformation formalism indeed respects this equivalence, with the GVC representations equivalent to the Cartesian representation

$$\vec{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}}. \quad (8.39)$$

8.8.1 Contravariant components to the velocity

Following Section 8.7, we have the contravariant velocity components

$$v^{\bar{1}} = u \quad v^{\bar{2}} = v \quad v^{\bar{3}} = \mathbf{v} \cdot \nabla \sigma. \quad (8.40)$$

Use of the basis vectors (8.9a)-(8.9c) then leads to

$$\vec{v} = v^{\bar{a}} \vec{e}_{\bar{a}} \quad (8.41a)$$

$$= u \vec{e}_{\bar{x}} + v \vec{e}_{\bar{y}} + (\mathbf{v} \cdot \nabla \sigma) \vec{e}_{\bar{\sigma}} \quad (8.41b)$$

$$= u [\hat{\mathbf{x}} + (\partial z / \partial \bar{x}) \hat{\mathbf{z}}] + v [\hat{\mathbf{y}} + (\partial z / \partial \bar{y}) \hat{\mathbf{z}}] + (\mathbf{v} \cdot \nabla \sigma) (\partial z / \partial \sigma) \hat{\mathbf{z}}. \quad (8.41c)$$

8.8.2 Covariant components to the velocity vector

The covariant velocity components are given by

$$v_{\bar{1}} = u + \frac{\partial z}{\partial \bar{x}} w \quad v_{\bar{2}} = v + \frac{\partial z}{\partial \bar{y}} w \quad v_{\bar{3}} = \frac{\partial z}{\partial \sigma} w. \quad (8.42)$$

The one-form basis (8.13a)–(8.13c) thus leads to the velocity vector

$$\vec{v} = v_{\bar{a}} \tilde{e}^{\bar{a}} = [u + (\partial z / \partial \bar{x}) w] \hat{x} + [v + (\partial z / \partial \bar{y}) w] \hat{y} + w (\partial z / \partial \sigma) \nabla \sigma. \quad (8.43)$$

8.8.3 Introducing the material time derivative

The material evolution for the generalized vertical coordinate can be written

$$\frac{D\sigma}{Dt} = \frac{\partial \sigma}{\partial t} + \mathbf{v} \cdot \nabla \sigma = \dot{\sigma}, \quad (8.44)$$

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. In contrast, if the GVC is an isopycnal, then $\dot{\sigma} \neq 0$ arises from irreversible processes such as mixing. Using the expression (8.44) in the velocity vector expression (8.41c) leads to

$$\vec{v} = u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] + (\mathbf{v} \cdot \nabla \sigma) (\partial z / \partial \sigma) \hat{z}. \quad (8.45a)$$

$$= u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] + (\dot{\sigma} - \partial \sigma / \partial t) (\partial z / \partial \sigma) \hat{z} \quad (8.45b)$$

$$= u \hat{x} + v \hat{y} + [\partial z / \partial \bar{t} + \mathbf{u} \cdot \nabla_{\sigma} z + (\partial z / \partial \sigma) \dot{\sigma}] \hat{z}, \quad (8.45c)$$

where the final equality made use of the triple product (8.24a): $(\partial \sigma / \partial t) (\partial z / \partial \sigma) = -\partial z / \partial \bar{t}$. 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

$$\vec{v} = u [\hat{x} + (\partial z / \partial \bar{x}) \hat{z}] + v [\hat{y} + (\partial z / \partial \bar{y}) \hat{z}] \quad \text{if } \partial_t \sigma = 0 \text{ and } \dot{\sigma} = 0. \quad (8.46)$$

However, in general there are transient fluctuations and material changes to σ so that $\mathbf{v} \cdot \nabla \sigma \neq 0$.

8.8.4 Equivalence to the Cartesian velocity representation

Use of the triple product identities (8.25b)–(8.25c) allows us to manipulate both expressions (8.41c) and (8.43) to recover the Cartesian expression

$$\vec{v} = u \hat{x} + v \hat{y} + w \hat{z}. \quad (8.47)$$

Another way to see this identity is to note that in equation (8.45c), the vertical component is an expression for the material time derivative of the vertical position

$$w = \frac{Dz}{Dt} = \frac{\partial z}{\partial \bar{t}} + \mathbf{u} \cdot \nabla_{\sigma} z + \frac{\partial z}{\partial \sigma} \dot{\sigma}. \quad (8.48)$$

We derive this identity in Section 18.4 where we discuss further kinematic results using GVCs.

8.8.5 Comments

In Chapter 18 we have far more to say on the GVC expressions for the velocity vector.

8.9 Metric tensor

The GVC representation of the metric tensor for measuring the distance between two points in space (Section 5.1) is given by

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} 1 + (\partial z / \partial \bar{x})^2 & (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) & (\partial z / \partial \bar{x})(\partial z / \partial \sigma) \\ (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) & 1 + (\partial z / \partial \bar{y})^2 & (\partial z / \partial \bar{y})(\partial z / \partial \sigma) \\ (\partial z / \partial \bar{x})(\partial z / \partial \sigma) & (\partial z / \partial \bar{y})(\partial z / \partial \sigma) & (\partial z / \partial \sigma)^2 \end{bmatrix}, \quad (8.49)$$

and the GVC representation of the inverse metric tensor is given by the somewhat simpler form

$$g^{\bar{a}\bar{b}} = \tilde{e}^{\bar{a}} \cdot \tilde{e}^{\bar{b}} = \begin{bmatrix} 1 & 0 & \partial \sigma / \partial x \\ 0 & 1 & \partial \sigma / \partial y \\ \partial \sigma / \partial x & \partial \sigma / \partial y & |\nabla \sigma|^2 \end{bmatrix}. \quad (8.50)$$

Proof that $g^{\bar{a}\bar{b}} g_{\bar{b}\bar{c}} = \delta_{\bar{c}}^{\bar{a}}$ requires use of the triple product identities (8.25b) and (8.25c).

8.9.1 Jacobian of transformation

The determinant of the GVC representation of the metric tensor (8.49) is

$$\det(g_{\bar{a}\bar{b}}) = (\partial z / \partial \sigma)^2 \quad (8.51)$$

so that the Jacobian of transformation (Section 5.4) is the specific thickness

$$\frac{\partial(x, y, z)}{\partial(\bar{x}, \bar{y}, \sigma)} = \frac{\partial z}{\partial \sigma}. \quad (8.52)$$

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

$$\frac{\partial z}{\partial \sigma} = \frac{\partial z}{\partial p} = -\frac{1}{\rho g}, \quad (8.53)$$

which is indeed single-signed since the mass density is always positive. We formally introduce the Jacobian of transformation in Section 8.9.1.

8.9.2 Covariant and contravariant representations

The metric tensor allows us to convert between the covariant and contra-variant representations of a vector via the identity (Section 5.2.3)

$$F_{\bar{a}} = g_{\bar{a}\bar{b}} F^{\bar{b}}. \quad (8.54)$$

²We derive the hydrostatic balance in Section 27.2.

We use triple product identities (8.25b)-(8.25c) to verify that this relation agrees with the transformation matrix approach detailed in Section 8.7. For example,

$$F_{\bar{1}} = g_{\bar{1}\bar{b}} F^{\bar{b}} \quad (8.55a)$$

$$= [1 + (\partial z / \partial \bar{x})^2] F^{\bar{1}} + (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) F^{\bar{2}} + (\partial z / \partial \bar{x})(\partial z / \partial \sigma) F^{\sigma} \quad (8.55b)$$

$$= [1 + (\partial z / \partial \bar{x})^2] F^1 + (\partial z / \partial \bar{x})(\partial z / \partial \bar{y}) F^2 + (\partial z / \partial \bar{x})(\partial z / \partial \sigma) \nabla \sigma \cdot \mathbf{F} \quad (8.55c)$$

$$= F^1 + (\partial z / \partial \bar{x}) F^3 \quad (8.55d)$$

$$= F_1 + (\partial z / \partial \bar{x}) F_3, \quad (8.55e)$$

where the final equality holds since $F^1 = F_1$ and $F^3 = F_3$ for Cartesian tensor components.

8.10 Volume element and the Levi-Civita tensor

The square root of the determinant of the metric tensor (8.49) is

$$\sqrt{\det(g_{\bar{a}\bar{b}})} = \partial z / \partial \sigma \quad (8.56)$$

so that the volume element (Section 5.4) is

$$dV = (\partial z / \partial \sigma) dx dy d\sigma. \quad (8.57)$$

The covariant Levi-Civita tensor (Section 5.6) has the GVC representations

$$\varepsilon_{\bar{a}\bar{b}\bar{c}} = (\partial z / \partial \sigma) \epsilon_{\bar{a}\bar{b}\bar{c}} \quad \varepsilon^{\bar{a}\bar{b}\bar{c}} = (\partial z / \partial \sigma)^{-1} \epsilon^{\bar{a}\bar{b}\bar{c}} \quad (8.58)$$

ϵ is the permutation symbol introduced in Section 2.4.1 with its components independent of coordinate representation.

8.11 Cross product of basis vectors

We now verify the relation (5.56) for the cross product of two basis vectors using GVCs

$$\vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = \varepsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}} \implies \vec{e}_{\bar{a}} \wedge \vec{e}_{\bar{b}} = (\partial z / \partial \sigma) \epsilon_{\bar{a}\bar{b}\bar{c}} \tilde{e}^{\bar{c}}. \quad (8.59)$$

Making use of the basis vectors from Section 8.2 and the basis one-forms from Section 8.3 renders

$$\vec{e}_{\bar{x}} \wedge \vec{e}_{\bar{y}} = \hat{\mathbf{z}} - \hat{\mathbf{x}}(\partial z / \partial \bar{x}) - \hat{\mathbf{y}}(\partial z / \partial \bar{y}) = (\partial z / \partial \sigma) \nabla \sigma = \varepsilon_{\bar{x}\bar{y}\sigma} \tilde{e}^{\sigma} \quad (8.60a)$$

$$\vec{e}_{\bar{y}} \wedge \vec{e}_{\sigma} = \hat{\mathbf{x}}(\partial z / \partial \sigma) = \varepsilon_{\bar{y}\sigma\bar{x}} \tilde{e}^{\bar{x}} \quad (8.60b)$$

$$\vec{e}_{\sigma} \wedge \vec{e}_{\bar{x}} = \hat{\mathbf{y}}(\partial z / \partial \sigma) = \varepsilon_{\sigma\bar{x}\bar{y}} \tilde{e}^{\bar{y}}. \quad (8.60c)$$

8.12 Partial derivative operators

We here consider the partial derivative operators and their transformation between coordinate systems. These identities are used throughout GVC calculus.

8.12.1 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

$$\partial_{\bar{t}} = \partial_t + (\partial z / \partial \bar{t}) \partial_z \quad \partial_{\bar{x}} = \partial_x + (\partial z / \partial \bar{x}) \partial_z \quad \partial_{\bar{y}} = \partial_y + (\partial z / \partial \bar{y}) \partial_z \quad \partial_\sigma = (\partial z / \partial \sigma) \partial_z. \quad (8.61)$$

We can make use of the triple product identities (8.25b) and (8.25c) to express the slope of a constant GVC surface in two equivalent manners

$$\nabla_\sigma z = - \frac{\nabla_z \sigma}{\partial \sigma / \partial z} \quad (8.62)$$

where we introduced the standard shorthand notation

$$\nabla_\sigma = \hat{x} \frac{\partial}{\partial \bar{x}} + \hat{y} \frac{\partial}{\partial \bar{y}} \quad \nabla_z = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y}. \quad (8.63)$$

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

$$\nabla_\sigma = \nabla_z + (\nabla_\sigma z) \partial_z \equiv \nabla_z + \mathbf{S} \partial_z, \quad (8.64)$$

where we introduced the notation for the slope of the σ surface relative to the horizontal plane

$$\mathbf{S} = \nabla_\sigma z = - \frac{\nabla_z \sigma}{\partial \sigma / \partial z}. \quad (8.65)$$

8.12.2 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 8.4. This feature of the horizontal derivative operator is a key aspect of the GVCs' non-orthogonality.

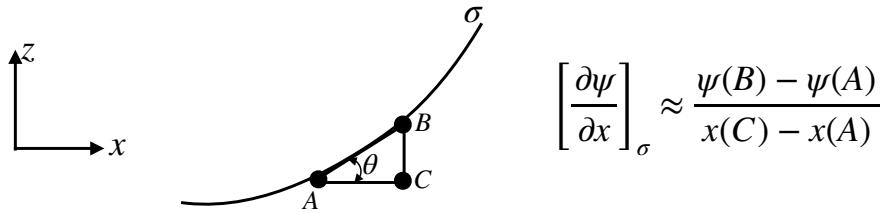


Figure 8.4: A surface of constant generalized vertical coordinate, σ , along with a local tangent plane with a slope $\tan \theta$ with respect to the horizontal plane. This figure illustrates the identities (8.67a)-(8.67d), with these identities relating a lateral derivative taken along the GVC surface to horizontal and vertical derivatives taken along orthogonal Cartesian axes.

Consider the geometry shown in Figure 8.4, which shows a generalized vertical coordinate surface (constant σ surface) along with a sample tangent plane with a slope

$$S^{(x)} = \frac{\text{rise}}{\text{run}} = \tan \theta = \frac{z(B) - z(C)}{x(C) - x(A)} \approx \left[\frac{\partial z}{\partial x} \right]_\sigma = - \frac{(\partial \sigma / \partial x)_z}{(\partial \sigma / \partial z)} \quad (8.66)$$

relative to the horizontal. We readily verify the following identities based on finite difference operations for an arbitrary function

$$\left[\frac{\partial \psi}{\partial x} \right]_{\sigma} \approx \frac{\psi(B) - \psi(A)}{x(C) - x(A)} \quad (8.67a)$$

$$= \frac{\psi(C) - \psi(A)}{x(C) - x(A)} + \frac{\psi(B) - \psi(C)}{x(C) - x(A)} \quad (8.67b)$$

$$= \frac{\psi(C) - \psi(A)}{x(C) - x(A)} + \left[\frac{z(B) - z(C)}{x(C) - x(A)} \right] \frac{\psi(B) - \psi(C)}{z(B) - z(C)} \quad (8.67c)$$

$$= \left[\frac{\partial \psi}{\partial x} \right]_z + S^{(x)} \left[\frac{\partial \psi}{\partial z} \right]_x. \quad (8.67d)$$

Taking the continuum limit then leads to the relations between horizontal derivatives computed on constant σ surfaces to those computed on constant z surfaces

$$\left[\frac{\partial}{\partial x} \right]_{\sigma} = \left[\frac{\partial}{\partial x} \right]_z + \left[\frac{\partial z}{\partial x} \right]_{\sigma} \frac{\partial}{\partial z} \quad (8.68a)$$

$$\left[\frac{\partial}{\partial y} \right]_{\sigma} = \left[\frac{\partial}{\partial y} \right]_z + \left[\frac{\partial z}{\partial y} \right]_{\sigma} \frac{\partial}{\partial z}, \quad (8.68b)$$

which can be written in the shorthand vector notation

$$\nabla_{\sigma} = \hat{x} \left[\frac{\partial}{\partial x} \right]_{\sigma} + \hat{y} \left[\frac{\partial}{\partial y} \right]_{\sigma} = \nabla_z + (\nabla_{\sigma} z) \partial_z. \quad (8.69)$$

8.13 Material time derivative

Making use of the relations for the partial derivative operators in Section 8.12 allows us to write the material time derivative in the following equivalent forms

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (8.70a)$$

$$= \left[\frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + \frac{D\sigma}{Dt} \frac{\partial}{\partial \sigma} \quad (8.70b)$$

$$= \left[\frac{\partial}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} + \dot{\sigma} \frac{\partial}{\partial \sigma}. \quad (8.70c)$$

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

$$(u, v) = \frac{D(x, y)}{Dt}. \quad (8.71)$$

We detail more properties of the material time derivative operator in Section 18.4.

8.14 Divergence of a vector and the divergence theorem

Making use of the general expression (6.7) for the covariant divergence of a vector renders the GVC expression

$$\nabla_{\bar{a}} F^{\bar{a}} = [\det(g_{\bar{a}\bar{b}})]^{-1/2} \partial_{\bar{a}} \left[[\det(g_{\bar{a}\bar{b}})]^{1/2} F^{\bar{a}} \right] = (\partial z / \partial \sigma)^{-1} \partial_{\bar{a}} [(\partial z / \partial \sigma) F^{\bar{a}}]. \quad (8.72)$$

Recall that the GVC vector components, $F^{\bar{a}}$, are related to the Cartesian components in equation (8.34), and the GVC components of the partial derivative operator, $\partial_{\bar{a}}$, are related to the Cartesian operator in equation (8.63).

When making use of the divergence theorem (Section 6.11), we require the product of the volume element and the covariant divergence. For GVCs this product takes on the form

$$(\nabla_{\bar{a}} F^{\bar{a}}) dV = \partial_{\bar{a}} [(\partial z / \partial \sigma) F^{\bar{a}}] d\bar{x} d\bar{y} d\sigma, \quad (8.73)$$

which reduces to a boundary integral when integrating over a volume.

8.15 The diffusion operator

As an explicit example of the covariant divergence operator (8.72), we here consider the diffusion operator discussed in Chapter 22. The derivation here recovers much of what we just discussed in Section 8.14, yet we make use of a bit less tensor formalism though at the cost of more algebra.

The diffusion operator is the convergence of the diffusive flux

$$\mathcal{R} = -\nabla \cdot \mathbf{J}, \quad (8.74)$$

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 8.12

$$-\mathcal{R} = \nabla \cdot \mathbf{J} \quad (8.75a)$$

$$= \nabla_z \cdot \mathbf{J}^h + \partial_z J^z \quad (8.75b)$$

$$= (\nabla_{\sigma} - \nabla_{\sigma} z \partial_z) \cdot \mathbf{J}^h + (\sigma_z) \partial_{\sigma} J^z \quad (8.75c)$$

$$= \sigma_z \left[z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}^h + (\hat{\mathbf{z}} \partial_{\sigma} - \nabla_{\sigma} z \partial_{\sigma}) \cdot \mathbf{J} \right] \quad (8.75d)$$

$$= \sigma_z \left[\nabla_{\sigma} \cdot (z_{\sigma} \mathbf{J}^h) - \mathbf{J}^h \cdot \nabla_{\sigma} (z_{\sigma}) + \partial_{\sigma} J^z - \partial_{\sigma} (\nabla_{\sigma} z \cdot \mathbf{J}) + \mathbf{J} \cdot \partial_{\sigma} (\nabla_{\sigma} z) \right] \quad (8.75e)$$

$$= \sigma_z \left[\nabla_{\sigma} \cdot (z_{\sigma} \mathbf{J}^h) + \partial_{\sigma} J^z - \partial_{\sigma} (\nabla_{\sigma} z \cdot \mathbf{J}^h) \right] \quad (8.75f)$$

$$= \sigma_z \left(\nabla_{\sigma} \cdot (\partial_{\sigma} z \mathbf{J}^h) + \partial_{\sigma} [(\hat{\mathbf{z}} - \nabla_{\sigma} z) \cdot \mathbf{J}] \right) \quad (8.75g)$$

$$= \sigma_z \left[\nabla_{\sigma} \cdot (z_{\sigma} \mathbf{J}^h) + \partial_{\sigma} (z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}) \right], \quad (8.75h)$$

where we used

$$z_{\sigma} \nabla_{\sigma} = \hat{\mathbf{z}} - \nabla_{\sigma} z \quad (8.76)$$

to reach the final equality, and made use of the shorthand

$$z_{\sigma} = \frac{\partial z}{\partial \sigma} \quad \sigma_z = \frac{\partial \sigma}{\partial z}. \quad (8.77)$$

Making use of the coordinate transformations in Section 8.7 for vector components reveals that the expression (8.75h) is identical to equation (8.72) derived using formal tensor methods. Likewise, when multiplying by the volume element

$$dV = dx dy dz = dx dy z_{\sigma} d\sigma, \quad (8.78)$$

we are led to

$$-\mathcal{R} dV = \left[\nabla_\sigma \cdot (z_\sigma \mathbf{J}^h) + \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] dx dy d\sigma, \quad (8.79)$$

which is identical to the expression (8.73). Now note that the increment $d\sigma$ commutes with the horizontal operator ∇_σ , so that

$$-\mathcal{R} dV = \left[\nabla_\sigma \cdot (d\sigma z_\sigma \mathbf{J}^h) + d\sigma \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] dx dy \quad (8.80a)$$

$$= \frac{1}{dz} \left[\nabla_\sigma \cdot (d\sigma z_\sigma \mathbf{J}^h) + d\sigma \partial_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] dx dy dz \quad (8.80b)$$

$$= \frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \delta_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right] dx dy h^\sigma, \quad (8.80c)$$

where we introduced the layer thickness

$$h^\sigma = dz = z_\sigma d\sigma = \frac{\partial z}{\partial \sigma} d\sigma \quad (8.81)$$

and the non-dimensional differential operator

$$\delta_\sigma \equiv d\sigma \frac{\partial}{\partial \sigma}. \quad (8.82)$$

Cancelling the volume element on both sides leads to the diffusion operator

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \delta_\sigma (z_\sigma \nabla \sigma \cdot \mathbf{J}) \right], \quad (8.83)$$

which is of the form commonly found in the ocean modeling literature. Notably, the thickness weighted flux $h^\sigma \mathbf{J}^h$ is in 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 . Furthermore, the flux $z_\sigma \nabla \sigma \cdot \mathbf{J}$ is commonly referred to as the dia-surface subgrid scale flux.

For the special case where the diffusive flux has zero component parallel to $\nabla \sigma$, then the diffusion operator reduces to

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) \right] \quad \text{if } \nabla \sigma \cdot \mathbf{J} = 0. \quad (8.84)$$

The neutral diffusion operator of Section 24.3.3 is an example of such an operator, with σ in this case given by the locally referenced potential density.

8.16 Vorticity

As detailed in Chapter 37, vorticity is the curl of the velocity

$$\vec{\omega} = \text{curl}(\vec{v}), \quad (8.85)$$

where the curl has components (Section 6.9)

$$\text{curl}(\vec{v}) = \tilde{e}_a \varepsilon^{abc} \partial_b v_c = \tilde{e}_{\bar{a}} \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}}. \quad (8.86)$$

8.16.1 The components

We identify the contravariant components of the vorticity via

$$\omega^{\bar{a}} = \varepsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}} = (\partial z / \partial \sigma)^{-1} \epsilon^{\bar{a}\bar{b}\bar{c}} \partial_{\bar{b}} v_{\bar{c}} \quad (8.87)$$

where we made use of equation (8.58) to introduce the permutation symbol. Expanding the components leads to

$$\omega^{\bar{1}} = (\partial \sigma / \partial z) (\partial_{\bar{2}} v_{\bar{3}} - \partial_{\bar{3}} v_{\bar{2}}) \quad (8.88a)$$

$$\omega^{\bar{2}} = (\partial \sigma / \partial z) (\partial_{\bar{3}} v_{\bar{1}} - \partial_{\bar{1}} v_{\bar{3}}) \quad (8.88b)$$

$$\omega^{\bar{3}} = \omega^\sigma = (\partial \sigma / \partial z) (\partial_{\bar{1}} v_{\bar{2}} - \partial_{\bar{2}} v_{\bar{1}}). \quad (8.88c)$$

8.16.2 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

$$\omega^{\bar{a}} = \Lambda_a^{\bar{a}} \omega^a, \quad (8.89)$$

where ω^a are the Cartesian components

$$\boldsymbol{\omega} = \hat{x} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) + \hat{y} \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right) + \hat{z} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right). \quad (8.90)$$

Making use of the transformation matrix $\Lambda_a^{\bar{a}}$ from equation (8.11) yields (as in Section 8.7)

$$\omega^{\bar{x}} = \omega^x = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \quad \omega^{\bar{y}} = \omega^y = \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \quad \omega^\sigma = \boldsymbol{\omega} \cdot \nabla \sigma. \quad (8.91)$$

Note that for isopycnal coordinates in a Boussinesq fluid, ω^σ equals to the potential vorticity when the vorticity is the absolute vorticity (Section 39.5). That is, the potential vorticity is the isopycnal component of the absolute vorticity.

8.17 Circulation

The circulation (Section 37.2) is given by the closed oriented path integral of the velocity projected into the direction of the path

$$C \equiv \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} \quad (8.92)$$

where $d\mathbf{r}$ is the vector line element along the path and ∂S is the closed path defining the boundary to a two-dimensional surface S . Stokes' Theorem from Section 3.6 leads to the identity (see also Section 37.2.1)

$$C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int_S (\nabla \wedge \mathbf{v}) \cdot \hat{\mathbf{n}} dS = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (8.93)$$

where $\hat{\mathbf{n}}$ is the outward normal vector orienting the area element dS according to the right-hand rule applied to the bounding circuit. These results are all written in a generally covariant manner (Section 6.1) 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

$$\hat{\mathbf{n}} = \frac{\nabla\sigma}{|\nabla\sigma|} \quad (8.94)$$

is the outward normal and

$$\boldsymbol{\omega} \cdot \hat{\mathbf{n}} = \frac{\omega^\sigma}{|\nabla\sigma|} \quad (8.95)$$

where $\omega^\sigma = \boldsymbol{\omega} \cdot \nabla\sigma$ (equation (8.91)). So long as the vertical stratification remains non-zero ($\partial\sigma/\partial z \neq 0$) we can write the area factor in the form

$$\frac{dS}{|\nabla\sigma|} = \frac{dS}{\sqrt{(\partial\sigma/\partial x)^2 + (\partial\sigma/\partial y)^2 + (\partial\sigma/\partial z)^2}} \quad (8.96a)$$

$$= \frac{dS}{|\partial\sigma/\partial z| \sqrt{[(\partial\sigma/\partial x)/(\partial\sigma/\partial z)]^2 + [(\partial\sigma/\partial y)/(\partial\sigma/\partial z)]^2 + 1}} \quad (8.96b)$$

$$= \frac{dS}{|\partial\sigma/\partial z| \sqrt{1 + \tan^2 \theta}} \quad (8.96c)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| |\cos \theta| dS \quad (8.96d)$$

$$= \left| \frac{\partial z}{\partial \sigma} \right| dA. \quad (8.96e)$$

The equality (8.96c) introduces the angle, θ , between the boundary surface and the horizontal plane as in Figure 8.4. The squared slope of this surface given by

$$\tan^2 \theta = \frac{\nabla_z \sigma \cdot \nabla_z \sigma}{(\partial\sigma/\partial z)^2} = \nabla_\sigma z \cdot \nabla_\sigma z. \quad (8.97)$$

The equality (8.96d) made use of a trigonometric identity, and the equality (8.96e) introduced the horizontal projection of the area,

$$dA = |\cos \theta| dS. \quad (8.98)$$

Bringing these results together leads to the expression for circulation around a closed loop on a constant σ surface

$$C_{\sigma-\text{surface}} = \int_S (\boldsymbol{\omega} \cdot \nabla\sigma) |\partial z/\partial \sigma| dA. \quad (8.99)$$

9

Tracer coordinates[†]

In this chapter we generalize the GVC considerations of Chapter 8 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 14). 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 7 and the generalized vertical coordinates of Chapter 8.

READER'S GUIDE TO THIS CHAPTER

This chapter assumes a working knowledge of material in the generalized vertical coordinate Chapter 8 as well as the treatment of general tensors in Chapters 5 and 6. No other chapter makes use of this chapter.

9.1	Incomplete threads	101
9.2	Relating Cartesian coordinates and tracer coordinates	102
9.3	The transformation matrix and its inverse	102
9.4	Determinant of the transformation	103
9.5	Basis vectors	103
9.6	Basis one-forms	104
9.7	Trajectory or position vector	104
9.8	Velocity	104
9.9	Metric tensor	105
9.10	Vorticity	105
9.11	Comments	105

9.1 Incomplete threads

- Equations of motion
- Mass conservation and the tracer equations

- Diffusion equation
- Potential vorticity
- Angular momentum
- Eddy and mean flow in tracer space

9.2 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

$$\xi^{\bar{0}} = \xi^0 \iff \bar{t} = t \quad (9.1a)$$

$$\xi^{\bar{1}} = C^{\bar{1}}(\xi^a, t) \quad (9.1b)$$

$$\xi^{\bar{2}} = C^{\bar{2}}(\xi^a, t) \quad (9.1c)$$

$$\xi^{\bar{3}} = C^{\bar{3}}(\xi^a, t). \quad (9.1d)$$

The inverse functional relation yields the position of a point in space in the form

$$t = \xi^0 = \xi^{\bar{0}} \iff t = \bar{t} \quad (9.2a)$$

$$x = \xi^1 = \xi^1(C^{\bar{a}}, \bar{t}) \quad (9.2b)$$

$$y = \xi^2 = \xi^2(C^{\bar{a}}, \bar{t}) \quad (9.2c)$$

$$z = \xi^3 = \xi^3(C^{\bar{a}}, \bar{t}) \quad (9.2d)$$

As for the GVC case of Chapter 8, 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.

9.3 The transformation matrix and its inverse

We can write the coordinate transformation from ξ^a to $\xi^{\bar{a}}$ in the generic form

$$\xi^{\bar{a}} = \xi^{\bar{a}}(\xi^a), \quad (9.3)$$

and the inverse transformation as

$$\xi^a = \xi^a(\xi^{\bar{a}}). \quad (9.4)$$

For the spatial coordinates, the coordinate transformation matrix is comprised of the partial derivatives

$$\Lambda^a_{\bar{a}} = \begin{bmatrix} (\partial\xi^1/\partial\xi^{\bar{1}})_{\bar{2},\bar{3}} & (\partial\xi^1/\partial\xi^{\bar{2}})_{\bar{1},\bar{3}} & (\partial\xi^1/\partial\xi^{\bar{3}})_{\bar{1},\bar{2}} \\ (\partial\xi^2/\partial\xi^{\bar{1}})_{\bar{2},\bar{3}} & (\partial\xi^2/\partial\xi^{\bar{2}})_{\bar{1},\bar{3}} & (\partial\xi^2/\partial\xi^{\bar{3}})_{\bar{1},\bar{2}} \\ (\partial\xi^3/\partial\xi^{\bar{1}})_{\bar{2},\bar{3}} & (\partial\xi^3/\partial\xi^{\bar{2}})_{\bar{1},\bar{3}} & (\partial\xi^3/\partial\xi^{\bar{3}})_{\bar{1},\bar{2}} \end{bmatrix} = \begin{bmatrix} \partial x/\partial C^{\bar{1}} & \partial x/\partial C^{\bar{2}} & \partial x/\partial C^{\bar{3}} \\ \partial y/\partial C^{\bar{1}} & \partial y/\partial C^{\bar{2}} & \partial y/\partial C^{\bar{3}} \\ \partial z/\partial C^{\bar{1}} & \partial z/\partial C^{\bar{2}} & \partial z/\partial C^{\bar{3}} \end{bmatrix}. \quad (9.5)$$

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

$$\Lambda^{\bar{a}}_a = \begin{bmatrix} (\partial\xi^{\bar{1}}/\partial\xi^1)_{2,3} & (\partial\xi^{\bar{1}}/\partial\xi^2)_{1,3} & (\partial\xi^{\bar{1}}/\partial\xi^3)_{1,2} \\ (\partial\xi^{\bar{2}}/\partial\xi^1)_{2,3} & (\partial\xi^{\bar{2}}/\partial\xi^2)_{1,3} & (\partial\xi^{\bar{2}}/\partial\xi^3)_{1,2} \\ (\partial\xi^{\bar{3}}/\partial\xi^1)_{2,3} & (\partial\xi^{\bar{3}}/\partial\xi^2)_{1,3} & (\partial\xi^{\bar{3}}/\partial\xi^3)_{1,2} \end{bmatrix} = \begin{bmatrix} \partial C^{\bar{1}}/\partial x & \partial C^{\bar{1}}/\partial y & \partial C^{\bar{1}}/\partial z \\ \partial C^{\bar{2}}/\partial x & \partial C^{\bar{2}}/\partial y & \partial C^{\bar{2}}/\partial z \\ \partial C^{\bar{3}}/\partial x & \partial C^{\bar{3}}/\partial y & \partial C^{\bar{3}}/\partial z \end{bmatrix}. \quad (9.6)$$

To verify that the matrix (9.6) is indeed the inverse of the matrix (9.5), we follow the method used in Section 8.4 by writing the coordinate transformation as a composite function

$$\xi^{\bar{a}} = \xi^{\bar{a}}(\xi^a) = \xi^{\bar{a}}[\xi^a(\xi^{\bar{b}})]. \quad (9.7)$$

Taking partial derivatives and using the chain rule thus renders

$$\delta^{\bar{a}}_b = \frac{\partial\xi^{\bar{a}}}{\partial\xi^b} = \frac{\partial\xi^{\bar{a}}}{\partial\xi^a} \frac{\partial\xi^a}{\partial\xi^b} \quad (9.8)$$

as well as

$$\delta^a_b = \frac{\partial\xi^a}{\partial\xi^b} = \frac{\partial\xi^a}{\partial\xi^{\bar{b}}} \frac{\partial\xi^{\bar{b}}}{\partial\xi^b}. \quad (9.9)$$

9.4 Determinant of the transformation

The determinant of the transformation matrix (9.5) is given by

$$\det(\Lambda^{\bar{a}}_a) = \left(\frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} \wedge \frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \right) \cdot \frac{\partial\mathbf{x}}{\partial C^{\bar{3}}} = \left(\frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \wedge \frac{\partial\mathbf{x}}{\partial C^{\bar{3}}} \right) \cdot \frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} = \left(\frac{\partial\mathbf{x}}{\partial C^{\bar{3}}} \wedge \frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} \right) \cdot \frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \quad (9.10)$$

and the determinant of the inverse transformation matrix (9.6) is

$$\det(\Lambda^a_{\bar{a}}) = \left(\frac{\partial\mathbf{C}}{\partial x} \wedge \frac{\partial\mathbf{C}}{\partial y} \right) \cdot \frac{\partial\mathbf{C}}{\partial z} = \left(\frac{\partial\mathbf{C}}{\partial y} \wedge \frac{\partial\mathbf{C}}{\partial z} \right) \cdot \frac{\partial\mathbf{C}}{\partial x} = \left(\frac{\partial\mathbf{C}}{\partial z} \wedge \frac{\partial\mathbf{C}}{\partial x} \right) \cdot \frac{\partial\mathbf{C}}{\partial y}. \quad (9.11)$$

For convenience we introduced a shorthand for the three tracer coordinates

$$\mathbf{C} = (C^{\bar{1}}, C^{\bar{2}}, C^{\bar{3}}) \quad (9.12)$$

and wrote

$$\vec{x} = \mathbf{x} = \hat{\mathbf{x}} x + \hat{\mathbf{y}} y + \hat{\mathbf{z}} z \quad (9.13)$$

for the Cartesian representation of the position vector.

9.5 Basis vectors

We transform the basis vectors from Cartesian into tracer coordinates through the transformation

$$\vec{e}_{\bar{a}} = \Lambda^a_{\bar{a}} \vec{e}_a. \quad (9.14)$$

Use of the transformation matrix (9.5) renders the tracer coordinate basis vectors

$$\vec{e}_{\bar{1}} = \frac{\partial\mathbf{x}}{\partial C^{\bar{1}}} \quad (9.15a)$$

$$\vec{e}_{\bar{2}} = \frac{\partial\mathbf{x}}{\partial C^{\bar{2}}} \quad (9.15b)$$

$$\vec{e}_{\bar{3}} = \frac{\partial\mathbf{x}}{\partial C^{\bar{3}}}. \quad (9.15c)$$

9.6 Basis one-forms

The basis one-forms are obtained by transforming from Cartesian coordinates to tracer coordinates through use of the inverse transformation

$$\tilde{e}^{\bar{a}} = \Lambda_{\bar{a}}^{\bar{a}} \tilde{e}^a. \quad (9.16)$$

Use of the inverse transformation matrix (8.11) renders the one-form basis in terms of tracer coordinates

$$\tilde{e}^{\bar{1}} = \hat{x} \frac{\partial C^{\bar{1}}}{\partial x} + \hat{y} \frac{\partial C^{\bar{1}}}{\partial y} + \hat{z} \frac{\partial C^{\bar{1}}}{\partial z} = \nabla C^{\bar{1}} \quad (9.17a)$$

$$\tilde{e}^{\bar{2}} = \hat{x} \frac{\partial C^{\bar{2}}}{\partial x} + \hat{y} \frac{\partial C^{\bar{2}}}{\partial y} + \hat{z} \frac{\partial C^{\bar{2}}}{\partial z} = \nabla C^{\bar{2}} \quad (9.17b)$$

$$\tilde{e}^{\bar{3}} = \hat{x} \frac{\partial C^{\bar{3}}}{\partial x} + \hat{y} \frac{\partial C^{\bar{3}}}{\partial y} + \hat{z} \frac{\partial C^{\bar{3}}}{\partial z} = \nabla C^{\bar{3}}. \quad (9.17c)$$

Use of the chain rule relations (9.8) and (9.9) ensure that the basis one-forms satisfy the orthogonality relation (5.18) with the basis vectors

$$\tilde{e}^{\bar{a}} \cdot \tilde{e}_{\bar{b}} = \delta_{\bar{a}}^{\bar{b}}. \quad (9.18)$$

9.7 Trajectory or position vector

The basis vectors (9.15a)-(9.15c) lead to the tracer coordinate representation of the position of a point in the fluid

$$\mathcal{P} = \xi^{\bar{a}} \tilde{e}_{\bar{a}} \quad (9.19a)$$

$$= C^{\bar{1}} \frac{\partial \mathbf{x}}{\partial C^{\bar{1}}} + C^{\bar{2}} \frac{\partial \mathbf{x}}{\partial C^{\bar{2}}} + C^{\bar{3}} \frac{\partial \mathbf{x}}{\partial C^{\bar{3}}} \quad (9.19b)$$

$$= (\mathbf{C} \cdot \nabla_C) \mathbf{x}. \quad (9.19c)$$

Following the isopycnal coordinate representation of the trajectory given by equation (8.31), we can provide a Taylor series interpretation of the tracer trajectory (9.19c).

9.8 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

$$v^{\bar{a}} = \lambda_{\bar{a}}^{\bar{a}} v^a, \quad (9.20)$$

with the inverse transformation matrix, $\lambda_{\bar{a}}^{\bar{a}}$, given by equation (9.6). The result of the calculation yields

$$\vec{v} = (\mathbf{v} \cdot \nabla) C^{\bar{1}} \frac{\partial \mathbf{x}}{\partial C^{\bar{1}}} + (\mathbf{v} \cdot \nabla) C^{\bar{2}} \frac{\partial \mathbf{x}}{\partial C^{\bar{2}}} + (\mathbf{v} \cdot \nabla) C^{\bar{3}} \frac{\partial \mathbf{x}}{\partial C^{\bar{3}}} \quad (9.21)$$

where we have the Cartesian inner product

$$\mathbf{v} \cdot \nabla C^{\bar{a}} = u (\partial C^{\bar{a}} / \partial x) + v (\partial C^{\bar{a}} / \partial y) + w (\partial C^{\bar{a}} / \partial z). \quad (9.22)$$

9.9 Metric tensor

The tracer coordinate representation of the metric tensor is given by

$$g_{\bar{a}\bar{b}} = \vec{e}_{\bar{a}} \cdot \vec{e}_{\bar{b}} = \begin{bmatrix} \frac{\partial \mathbf{x}}{\partial C^1} \cdot \frac{\partial \mathbf{x}}{\partial C^1} & \frac{\partial \mathbf{x}}{\partial C^1} \cdot \frac{\partial \mathbf{x}}{\partial C^2} & \frac{\partial \mathbf{x}}{\partial C^1} \cdot \frac{\partial \mathbf{x}}{\partial C^3} \\ \frac{\partial \mathbf{x}}{\partial C^2} \cdot \frac{\partial \mathbf{x}}{\partial C^1} & \frac{\partial \mathbf{x}}{\partial C^2} \cdot \frac{\partial \mathbf{x}}{\partial C^2} & \frac{\partial \mathbf{x}}{\partial C^2} \cdot \frac{\partial \mathbf{x}}{\partial C^3} \\ \frac{\partial \mathbf{x}}{\partial C^3} \cdot \frac{\partial \mathbf{x}}{\partial C^1} & \frac{\partial \mathbf{x}}{\partial C^3} \cdot \frac{\partial \mathbf{x}}{\partial C^2} & \frac{\partial \mathbf{x}}{\partial C^3} \cdot \frac{\partial \mathbf{x}}{\partial C^3} \end{bmatrix}, \quad (9.23)$$

and the tracer coordinate representation of the inverse metric tensor is given by

$$g^{\bar{a}\bar{b}} = \vec{e}^{\bar{a}} \cdot \vec{e}^{\bar{b}} = \begin{bmatrix} \nabla C^1 \cdot \nabla C^1 & \nabla C^1 \cdot \nabla C^2 & \nabla C^1 \cdot \nabla C^3 \\ \nabla C^2 \cdot \nabla C^1 & \nabla C^2 \cdot \nabla C^2 & \nabla C^2 \cdot \nabla C^3 \\ \nabla C^3 \cdot \nabla C^1 & \nabla C^3 \cdot \nabla C^2 & \nabla C^3 \cdot \nabla C^3 \end{bmatrix}. \quad (9.24)$$

Proof that $g^{\bar{a}\bar{b}} g_{\bar{b}\bar{c}} = \delta_{\bar{c}}^{\bar{a}}$ requires use of the chain rule relations (9.8) and (9.9).

9.10 Vorticity

Following from the discussion of vorticity in generalized vertical coordinates (Section 8.16), we determine the tracer coordinate representation through transforming $\omega^{\bar{a}} = \Lambda_a^{\bar{a}} \omega^a$ 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 (9.6). The resulting expression is given by

$$\omega^{\bar{1}} = \boldsymbol{\omega} \cdot \nabla C^{\bar{1}} \quad (9.25a)$$

$$\omega^{\bar{2}} = \boldsymbol{\omega} \cdot \nabla C^{\bar{2}} \quad (9.25b)$$

$$\omega^{\bar{3}} = \boldsymbol{\omega} \cdot \nabla C^{\bar{3}}, \quad (9.25c)$$

revealing that the tracer coordinate representation of the vorticity is obtained by projecting the Cartesian representation onto the directions perpendicular to the tracer surfaces.

9.11 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 38).

Part II

Particle mechanics in a rotating frame

In this part of the book, we develop the Newtonian mechanics of a particle moving around a gravitating sphere as viewed in a rotating reference frame. A particular example is the motion of a satellite moving around the earth in a near geostationary orbit. We encounter important concepts that later appear in geophysical fluid mechanics, such as trajectories, linear momentum, angular momentum, forces, non-inertial accelerations, and spherical coordinates.

10

Particle kinematics

In this chapter we 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. This motion offers us a useful introduction to features of motion in a spherical geometry as observed from a non-inertial reference frame. We thus encounter the suite of non-inertial accelerations (centrifugal and Coriolis) that also appear for fluid motion in a rotating frame.

READER'S GUIDE TO THIS CHAPTER

This chapter makes use of basic features of both Cartesian and general tensor algebra as presented in Chapters 2, 4, 5, and Section 7.2. 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 11 focused on particle dynamics on the sphere, and Chapter 26 considering fluid dynamics on the sphere.

10.1	The rotating earth	110
10.2	Reference frames and non-inertial accelerations	110
10.3	A few points from tensor algebra	112
10.3.1	Why we need general tensors	112
10.3.2	The coordinate representation of a vector	112
10.4	Galilean invariance for particle motion	113
10.5	Rotationally generated changes to a vector	114
10.5.1	Change in direction: brief derivation	114
10.5.2	Change in direction: detailed derivation for planar motion	115
10.6	The velocity vector	115
10.6.1	Coordinate velocity	115
10.6.2	Changes to the basis vectors	116
10.7	Inertial acceleration and its decomposition	116
10.8	Coordinate representation of the position vector	117
10.9	Coordinate representation of the velocity vector	117
10.9.1	Planetary Cartesian coordinate representation	118
10.9.2	Spherical coordinate representation	118
10.9.3	Axial angular momentum	119
10.10	Cartesian representation of the acceleration vector	119

10.10.1	Cartesian coordinate representation	119
10.10.2	Summary of the Cartesian acceleration contributions	120
10.11	Spherical representation of the acceleration vector	121
10.11.1	Decomposing the acceleration	122
10.11.2	Spherical coordinate acceleration	122
10.11.3	Metric acceleration	123
10.11.4	Centrifugal acceleration	123
10.11.5	Coriolis acceleration	123
10.11.6	Coriolis acceleration for large-scale motions	124
10.11.7	Further reading	124
10.12	Exercises	125

10.1 The rotating earth

The rotation of the planet is a defining feature of motion of the atmosphere and ocean. Motion described from within a rotating reference frame introduces a great deal of richness, and complexity, relative to a non-rotating frame. We will be developing experience with rotating physics throughout this book. Here, we note some of the basic details of the earth's rotation.

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 10.1). Other astronomical motions can be neglected for geophysical fluid dynamics. 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

$$\Omega = \frac{2\pi + 2\pi/365.24}{86400\text{s}} = \left[\frac{\pi}{43082} \right] \text{s}^{-1} = 7.2921 \times 10^{-5} \text{ s}^{-1}. \quad (10.1)$$

The earth's angular velocity, both its direction and its magnitude, is assumed constant in time for purposes of geophysical fluid dynamics (see Section 10.1)

$$\frac{d\vec{\Omega}}{dt} = 0. \quad (10.2)$$

10.2 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 10.4). 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, 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 10.2 illustrates the case for the position vector of a particle relative to the origin of a rotating sphere.

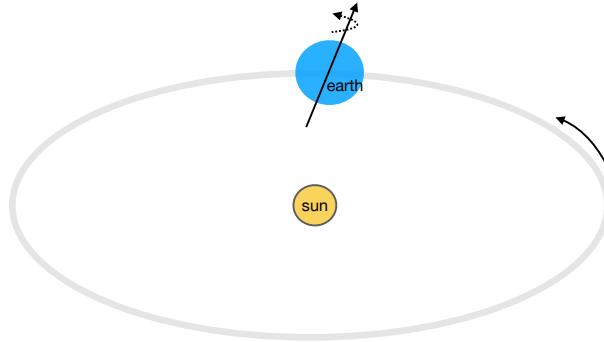


Figure 10.1: The angular velocity of the earth arises from both the spin about the polar axis and the orbit of the planet around the sun. This angular velocity determines the strength of the Coriolis acceleration and the centrifugal acceleration.

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.

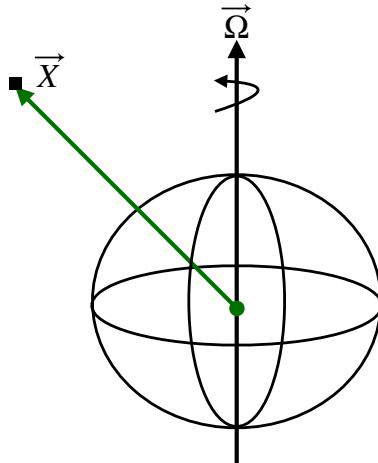


Figure 10.2: The position vector, $\vec{X}(t)$, for a particle moving around a rotating sphere with coordinate origins at the center of the sphere. The rotation axis is through the north pole, with angular velocity vector $\vec{\Omega}$. The sphere rotates in a positive right hand sense through the north polar axis (counter clockwise from above). The rotating frame has a “solid-body” velocity $\vec{U}_{\text{solid}} = \vec{\Omega} \wedge \vec{X}$ (equation (10.25)). 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.

10.3 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 the earlier material.

10.3.1 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 27.3). However, we make routine use of spherical coordinates when describing geophysical motion, and cylindrical-polar coordinates for studies of rotating tank experiments (see Exercise 27.4 and Section 33.3). Finally, we use generalized vertical coordinates in the description of stratified flows (Chapters 8, 18, and 31). 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.

10.3.2 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

$$\vec{X} = \sum_{a=1}^3 \xi^a \vec{e}_a = \xi^a \vec{e}_a, \quad (10.3)$$

where the Einstein summation convention is defined by the final equality. In this equation,

$$\vec{e}_a = (\vec{e}_1, \vec{e}_2, \vec{e}_3) \quad (10.4)$$

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 7.2.2). Note that we commonly make use of the boldface notation for a vector

$$\mathbf{X} = \vec{X}. \quad (10.5)$$

The basis vectors in equation (10.3) 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 (called a *one-form*) \tilde{e}^a . Duality here is defined using the familiar (Euclidean) inner product

$$\vec{e}_a \cdot \tilde{e}^b = \delta_a^b, \quad (10.6)$$

with δ_a^b the Kronecker delta tensor

$$\delta_a^b = \begin{cases} 1 & \text{if } b = a \\ 0 & \text{if } b \neq a. \end{cases} \quad (10.7)$$

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 (Chapters 7, 8, and 9).

10.4 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. We prove this statement by contradiction. Assume there is a relative acceleration between the two frames, and assume one frame is inertial. Without loss of generality, let the inertial frame be at rest. Then the other frame is accelerating. However, inertial frames are not accelerating, so the other frame cannot be inertial.

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, as doing so offers us practice for the less trivial case of a rotating reference frame.

Mathematically, the Galilean transformation is given by

$$\bar{t} = t \quad (10.8)$$

$$\bar{\mathbf{X}} = \mathbf{X} + \mathbf{U}t, \quad (10.9)$$

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} . The particle velocity in the moving reference frame is given by

$$\bar{\mathbf{V}} = \frac{d\bar{\mathbf{X}}}{dt} = \frac{d\mathbf{X}}{dt} + \frac{d(\mathbf{U}t)}{dt} = \mathbf{V} + \mathbf{U}, \quad (10.10)$$

where we set

$$\frac{d\mathbf{U}}{dt} = 0 \quad (10.11)$$

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

$$\bar{\mathbf{A}} = \frac{d^2\bar{\mathbf{X}}}{dt^2} = \frac{d\mathbf{V}}{dt} = \mathbf{A}. \quad (10.12)$$

The accelerations are indeed identical. As seen in Section 14.5, Galilean invariance for fluid motion provides a richer statement than it does for a point particle.

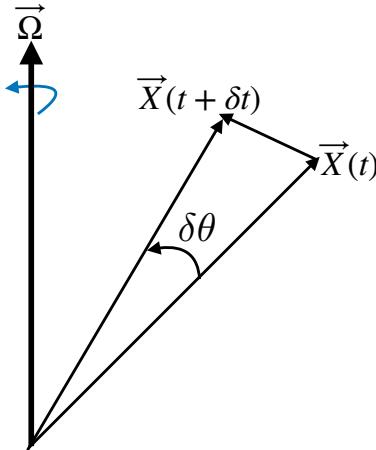


Figure 10.3: The change in a vector under a pure rotation leaves the vector magnitude unchanged, $|\mathbf{X}(t)| = |\mathbf{X}(t + \delta t)|$. Only the vector direction changes, here shown to be $\delta\theta = \vec{\Omega} \cdot \delta t$. Infinitesimal changes generated by the angular velocity $\vec{\Omega}$ lead to the vector differences $\mathbf{X}(t + \delta t) - \mathbf{X}(t) = \delta t \vec{\Omega} \wedge \mathbf{X}(t)$.

10.5 Rotationally generated changes to a vector

How does a vector change under a solid-body rotation such as that shown in Figure 10.3? 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

$$|\mathbf{X}(t)| = |\mathbf{X}(t + \delta t)|. \quad (10.13)$$

The condition (10.13) can be written as

$$\frac{d(\mathbf{X} \cdot \mathbf{X})}{dt} = 0, \quad (10.14)$$

which leads to the constraint

$$\mathbf{X} \cdot \frac{d\mathbf{X}}{dt} = 0. \quad (10.15)$$

That is, the velocity generated by a pure rotation is itself perpendicular to the position. We encountered this result in Section 3.1.4 when showing that unit vectors can only change through rotations.

10.5.1 Change in direction: brief derivation

Referring to Figure 10.3, 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

$$\mathbf{X}(t + \delta t) - \mathbf{X}(t) = \delta t \vec{\Omega} \wedge \mathbf{X}(t). \quad (10.16)$$

Dividing by δt leads to

$$\frac{d\mathbf{X}}{dt} = \vec{\Omega} \wedge \mathbf{X} \quad \text{pure rotation.} \quad (10.17)$$

Note that this evolution satisfies the constraint (10.15) since $\mathbf{X} \cdot (\vec{\Omega} \wedge \mathbf{X}) = 0$, meaning that the magnitude of the vector remains fixed.

10.5.2 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 $\boldsymbol{\Omega}$ 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

$$\delta\theta = |\boldsymbol{\Omega}| \delta t \quad (10.18)$$

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

$$\delta s = |\mathbf{X}(t)| \delta\theta = |\mathbf{X}(t)| |\boldsymbol{\Omega}| \delta t. \quad (10.19)$$

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

$$\mathbf{X}(t + \delta t) - \mathbf{X}(t) = \boldsymbol{\Omega} \wedge \mathbf{X}(t) \delta t. \quad (10.20)$$

Dividing through by δt and taking the limit $\delta t \rightarrow 0$ gives

$$\frac{d\mathbf{X}}{dt} = \boldsymbol{\Omega} \wedge \mathbf{X} \quad \text{pure rotation.} \quad (10.21)$$

The proof for the general 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.

10.6 The velocity vector

The inertial velocity is the time derivative of the position vector

$$\mathbf{V} = \frac{d\mathbf{X}}{dt}. \quad (10.22)$$

In general, both the coordinate representation and the basis vectors are time dependent, so that the inertial velocity has two contributions, one from the time dependence of the coordinates and another from the basis vectors

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = \frac{d(\xi^a \vec{e}_a)}{dt} = \frac{d\xi^a}{dt} \vec{e}_a + \xi^a \frac{d\vec{e}_a}{dt}. \quad (10.23)$$

10.6.1 Coordinate velocity

The first term on the right hand side of equation (10.23) is the velocity as measured within the rotating reference frame using the chosen coordinates

$$\mathbf{V}_{\text{coord}} \equiv \frac{d\xi^a}{dt} \vec{e}_a. \quad (10.24)$$

10.6.2 Changes to the basis vectors

The second term on the right hand side of equation (10.23) 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 10.5)

$$\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X}. \quad (10.25)$$

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.

10.7 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

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2\mathbf{X}}{dt^2}. \quad (10.26)$$

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 (10.26), and making use of the chain rule, leads to

$$\mathbf{A} = \frac{d}{dt} \frac{d\mathbf{X}}{dt} \quad (10.27a)$$

$$= \frac{d}{dt} \frac{d(\xi^a \vec{e}_a)}{dt} \quad (10.27b)$$

$$= \frac{d}{dt} \left[\frac{d\xi^a}{dt} \vec{e}_a + \xi^a \frac{d\vec{e}_a}{dt} \right] \quad (10.27c)$$

$$= \frac{d^2\xi^a}{dt^2} \vec{e}_a + 2 \frac{d\xi^a}{dt} \frac{d\vec{e}_a}{dt} + \xi^a \frac{d^2\vec{e}_a}{dt^2}. \quad (10.27d)$$

The first term on the right hand side is the acceleration of the coordinate representation as measured in the rotating reference frame

$$\mathbf{A}_{\text{coord}} \equiv \frac{d^2\xi^a}{dt^2} \vec{e}_a. \quad (10.28)$$

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 (10.27d) 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.

10.8 Coordinate representation of the position vector

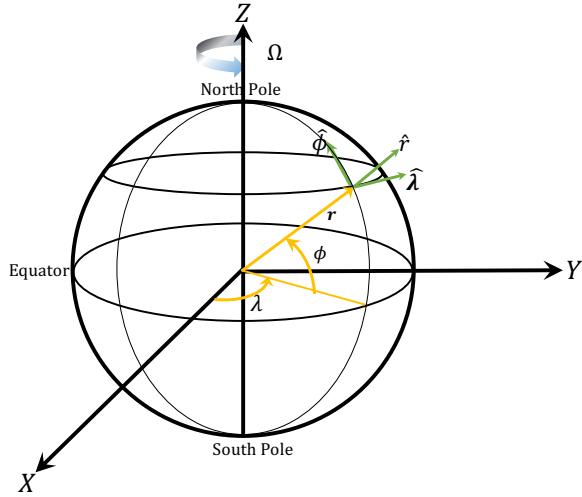


Figure 10.4: Geometry and notation for motion around a rotating sphere. For geophysical applications, the sphere rotates counter-clockwise when looking down from the north polar axis with angular speed Ω . The planetary Cartesian triad of orthonormal basis vectors, $(\hat{x}, \hat{y}, \hat{z})$ points 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 7.2 relating Cartesian and spherical coordinates and as furthermore defined by Figure 10.4. 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

$$\mathbf{X} = x \hat{x} + y \hat{y} + z \hat{z} \quad (10.29a)$$

$$= (r \cos \phi \cos \lambda) \hat{x} + (r \cos \phi \sin \lambda) \hat{y} + (r \sin \phi) \hat{z} \quad (10.29b)$$

$$= r \hat{r} \quad (10.29c)$$

$$= |\mathbf{X}| \hat{r}. \quad (10.29d)$$

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.

10.9 Coordinate representation of the velocity vector

As seen in Section 10.6, the inertial velocity vector has a coordinate representation written as

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = \frac{d(\xi^a \vec{e}_a)}{dt} = \frac{d\xi^a}{dt} \vec{e}_a + \xi^a \frac{d\vec{e}_a}{dt}. \quad (10.30)$$

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.

10.9.1 Planetary Cartesian coordinate representation

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 27.3 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

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}, \quad (10.31)$$

so that the solid-body velocity only has components in the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ directions

$$\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X} = \Omega (-\hat{\mathbf{x}} y + \hat{\mathbf{y}} x). \quad (10.32)$$

The inertial velocity thus takes the form

$$\mathbf{V} = \left[-y \Omega + \frac{dx}{dt} \right] \hat{\mathbf{x}} + \left[x \Omega + \frac{dy}{dt} \right] \hat{\mathbf{y}} + \frac{dz}{dt} \hat{\mathbf{z}} \quad (10.33a)$$

$$= \mathbf{V}_{\text{Cartesian}} + \boldsymbol{\Omega} \wedge \mathbf{X}, \quad (10.33b)$$

where we defined the Cartesian velocity vector

$$\mathbf{V}_{\text{Cartesian}} \equiv \frac{dx}{dt} \hat{\mathbf{x}} + \frac{dy}{dt} \hat{\mathbf{y}} + \frac{dz}{dt} \hat{\mathbf{z}}. \quad (10.34)$$

10.9.2 Spherical coordinate representation

The position vector in the spherical coordinate representation is given by

$$\mathbf{X} = r \hat{\mathbf{r}}. \quad (10.35)$$

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 this change by taking the time derivative of $\hat{\mathbf{r}}$ as given by equation (7.29c)

$$\frac{d\hat{\mathbf{r}}}{dt} = \frac{d}{dt} [\hat{\mathbf{x}} \cos \lambda \cos \phi + \hat{\mathbf{y}} \sin \lambda \cos \phi + \hat{\mathbf{z}} \sin \phi] \quad (10.36a)$$

$$= \cos \phi \left[\frac{d\lambda}{dt} + \Omega \right] \hat{\lambda} + \frac{d\phi}{dt} \hat{\phi}. \quad (10.36b)$$

Consequently, the inertial velocity is given by

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} \quad (10.37a)$$

$$= \frac{d(r \hat{\mathbf{r}})}{dt} \quad (10.37b)$$

$$= \frac{dr}{dt} \hat{\mathbf{r}} + r \frac{d\hat{\mathbf{r}}}{dt} \quad (10.37c)$$

$$= r_{\perp} \left[\frac{d\lambda}{dt} + \Omega \right] \hat{\lambda} + r \frac{d\phi}{dt} \hat{\phi} + \frac{dr}{dt} \hat{\mathbf{r}} \quad (10.37d)$$

$$= (u + r_{\perp} \Omega) \hat{\lambda} + v \hat{\phi} + w \hat{\mathbf{r}} \quad (10.37e)$$

$$= \mathbf{V}_{\text{spherical}} + \mathbf{U}_{\text{solid}}. \quad (10.37f)$$

In this equation we introduced the spherical velocity vector

$$\mathbf{V}_{\text{spherical}} = r_{\perp} \frac{d\lambda}{dt} \hat{\lambda} + r \frac{d\phi}{dt} \hat{\phi} + \frac{dr}{dt} \hat{r} \quad (10.38a)$$

$$= u \hat{\lambda} + v \hat{\phi} + w \hat{r}, \quad (10.38b)$$

where

$$u = r_{\perp} \frac{d\lambda}{dt} \quad v = r \frac{d\phi}{dt} \quad w = \frac{dr}{dt}, \quad (10.39)$$

are components to the spherical velocity vector, and with

$$r_{\perp} = r \cos \phi \quad (10.40)$$

the distance to the polar axis. We also noted that the solid-body velocity takes on the spherical coordinate form

$$\mathbf{U}_{\text{solid}} = \boldsymbol{\Omega} \wedge \mathbf{X} = r_{\perp} \boldsymbol{\Omega} \hat{\lambda}. \quad (10.41)$$

10.9.3 Axial angular momentum

As seen in Section 12.6, the zonal component of the inertial velocity equals to the axial angular momentum per unit mass

$$L^z = m \hat{\lambda} \cdot \mathbf{V} = m (u + r_{\perp} \Omega). \quad (10.42)$$

As the distance to the rotational axis, r_{\perp} 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. This conservation law offers a very important constraint on the particle trajectory (Section 12.6).

10.10 Cartesian representation of the acceleration vector

The inertial acceleration vector is given by the second time derivative of the position vector

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2 \mathbf{X}}{dt^2}. \quad (10.43)$$

We here consider its representation using Cartesian coordinates (x, y, z) and the Cartesian basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3) = (\hat{x}, \hat{y}, \hat{z})$.

10.10.1 Cartesian coordinate representation

For our study of geophysical fluid motion, we assume the planetary angular velocity, $\boldsymbol{\Omega}$, is a constant in time

$$\frac{d\boldsymbol{\Omega}}{dt} = 0. \quad (10.44)$$

Making use of the results from Section 10.5 leads to

$$\frac{d\vec{e}_a}{dt} = \boldsymbol{\Omega} \wedge \vec{e}_a \quad (10.45)$$

and

$$\frac{d^2\vec{e}_a}{dt^2} = \frac{d}{dt}(\boldsymbol{\Omega} \wedge \vec{e}_a) \quad (10.46a)$$

$$= \boldsymbol{\Omega} \wedge \frac{d\vec{e}_a}{dt} \quad (10.46b)$$

$$= \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \vec{e}_a), \quad (10.46c)$$

which leads to the inertial acceleration

$$\mathbf{A} = \frac{d}{dt} \frac{d\mathbf{X}}{dt} \quad (10.47a)$$

$$= \frac{d^2\xi^a}{dt^2} \vec{e}_a + 2 \frac{d\xi^a}{dt} (\boldsymbol{\Omega} \wedge \vec{e}_a) + \xi^a \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \vec{e}_a) \quad (10.47b)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} + \boldsymbol{\Omega} \wedge (\boldsymbol{\Omega} \wedge \mathbf{X}) \quad (10.47c)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \Omega^2(x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) \quad (10.47d)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} + \nabla\Phi_{\text{centrifugal}}. \quad (10.47e)$$

The first term is the Cartesian acceleration

$$\mathbf{A}_{\text{Cartesian}} = \frac{d^2x}{dt^2} \hat{\mathbf{x}} + \frac{d^2y}{dt^2} \hat{\mathbf{y}} + \frac{d^2z}{dt^2} \hat{\mathbf{z}}, \quad (10.48)$$

which is the coordinate acceleration measured in the rotating frame using Cartesian coordinates. The second term on the right hand side is minus the Coriolis acceleration. It plays a fundamental role in geophysical fluid mechanics, with much more discussion later. The third term is minus the centrifugal acceleration, which is the gradient of a potential

$$-\Phi_{\text{centrifugal}} \equiv \frac{\Omega^2 r_\perp^2}{2} = \frac{\Omega^2 (x^2 + y^2)}{2}. \quad (10.49)$$

10.10.2 Summary of the Cartesian acceleration contributions

For the purpose of formulating the equation of motion in the rotating terrestrial frame, we write the rotating frame acceleration as

$$\mathbf{A}_{\text{Cartesian}} = \mathbf{A} + \mathbf{A}_{\text{Coriolis}} + \mathbf{A}_{\text{centrifugal}} \quad (10.50a)$$

$$= \mathbf{A} - 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \nabla\Phi_{\text{centrifugal}} \quad (10.50b)$$

and identify 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

$$\mathbf{A}_{\text{Coriolis}} = -2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} \quad (10.51a)$$

$$= -2\Omega \hat{\mathbf{z}} \wedge \mathbf{V}_{\text{Cartesian}} \quad (10.51b)$$

$$= -2\Omega \left[-\frac{dy}{dt} \hat{\mathbf{x}} + \frac{dx}{dt} \hat{\mathbf{y}} \right], \quad (10.51c)$$

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 therefore plays a central role in geophysical fluid dynamics. Note that it has components only in the $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ plane. This geometry is to be expected since the Coriolis acceleration arises from rotation about the polar $\hat{\mathbf{z}}$ axis.

- CENTRIFUGAL: The centrifugal acceleration

$$\mathbf{A}_{\text{centrifugal}} = -\nabla\Phi_{\text{centrifugal}} \quad (10.52a)$$

$$= \Omega^2 \mathbf{r}_\perp \quad (10.52b)$$

$$= \Omega^2 (x \hat{\mathbf{x}} + y \hat{\mathbf{y}}) \quad (10.52c)$$

is another term arising from the rotating reference frame, with components only in the $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ plane. Because the centrifugal acceleration can be written as the gradient of a scalar potential

$$\Phi_{\text{centrifugal}} = -(\Omega^2/2)(x^2 + y^2) = -\frac{(r_\perp \Omega)^2}{2}, \quad (10.53)$$

the centrifugal acceleration can be combined with the gravitational acceleration in the equation of motion (Section 11.1). The resulting “effective gravity” leads to a force that is modified relative to the central gravitational field of the non-rotating spherical planet. We detail these points in Section 11.1.2.

10.11 Spherical representation of the acceleration vector

The spherical representation of the inertial velocity is given by equation (10.37f)

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = (u + r_\perp \Omega) \hat{\lambda} + v \hat{\phi} + w \hat{r} = \mathbf{V}_{\text{sphere}} + r_\perp \Omega \hat{\lambda}, \quad (10.54)$$

where we introduced the spherical velocity from equation (10.39)

$$\mathbf{V}_{\text{sphere}} \equiv u \hat{\lambda} + v \hat{\phi} + w \hat{r}. \quad (10.55)$$

We will also make use of the notation for the zonal component of the inertial velocity,

$$u_I = u + r_\perp \Omega. \quad (10.56)$$

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

$$\mathbf{A} = \frac{d}{dt} (u_I \hat{\lambda} + v \hat{\phi} + w \hat{r}) \quad (10.57a)$$

$$= \frac{du_I}{dt} \hat{\lambda} + \frac{dv}{dt} \hat{\phi} + \frac{dw}{dt} \hat{r} + u_I \frac{d\hat{\lambda}}{dt} + v \frac{d\hat{\phi}}{dt} + w \frac{d\hat{r}}{dt}. \quad (10.57b)$$

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 (7.2.2), a bit of algebra yields the time derivatives

$$\frac{d\hat{\lambda}}{dt} = \left(\Omega + \frac{d\lambda}{dt} \right) (\hat{\phi} \sin \phi - \hat{r} \cos \phi) \quad (10.58)$$

$$\frac{d\hat{\phi}}{dt} = -\hat{\lambda} \left(\Omega + \frac{d\lambda}{dt} \right) \sin \phi - \hat{r} \dot{\phi} \quad (10.59)$$

$$\frac{d\hat{r}}{dt} = \hat{\lambda} \left(\frac{d\lambda}{dt} + \Omega \right) \cos \phi + \frac{d\phi}{dt} \hat{\phi}. \quad (10.60)$$

We are thus led to the inertial acceleration components

$$\mathbf{A} \cdot \hat{\lambda} = \left[\frac{du_I}{dt} + \left(\frac{d\lambda}{dt} + \Omega \right) (w \cos \phi - v \sin \phi) \right] \quad (10.61a)$$

$$\mathbf{A} \cdot \hat{\phi} = \left[\frac{dv}{dt} + \left(\frac{d\lambda}{dt} + \Omega \right) u_I \sin \phi + w \frac{d\phi}{dt} \right] \quad (10.61b)$$

$$\mathbf{A} \cdot \hat{r} = \left[\frac{dw}{dt} - \left(\frac{d\lambda}{dt} + \Omega \right) u_I \cos \phi - v \frac{d\phi}{dt} \right]. \quad (10.61c)$$

Use of the identities

$$u = r_{\perp} \frac{d\lambda}{dt} \quad u_I = u + r_{\perp} \Omega \quad \frac{du_I}{dt} = \frac{du}{dt} + \Omega (w \cos \phi - v \sin \phi) \quad (10.62)$$

and some reorganization leads to

$$\begin{aligned} \mathbf{A} = & \hat{\lambda} \left[\frac{du}{dt} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) \right] \\ & + \hat{\phi} \left[\frac{dv}{dt} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi + r_{\perp} \Omega^2 \sin \phi \right] \\ & + \hat{r} \left[\frac{dw}{dt} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi - r_{\perp} \Omega^2 \cos \phi \right]. \end{aligned} \quad (10.63)$$

10.11.1 Decomposing the acceleration

We now decompose the inertial acceleration (10.63) into the following terms

$$\mathbf{A} = \mathbf{A}_{\text{sphere}} + \mathbf{A}_{\text{metric}} - \mathbf{A}_{\text{Coriolis}} - \mathbf{A}_{\text{centrifugal}}. \quad (10.64)$$

We chose signs so that in the rotating frame the acceleration is written

$$\underbrace{\mathbf{A}_{\text{sphere}} + \mathbf{A}_{\text{metric}}}_{\text{net spherical acceleration}} = \mathbf{A} + \mathbf{A}_{\text{Coriolis}} + \mathbf{A}_{\text{centrifugal}}. \quad (10.65)$$

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 written using spherical coordinates. The Coriolis and centrifugal terms are the new terms that arise from rotation.

10.11.2 Spherical coordinate acceleration

The spherical coordinate acceleration is given by the time change in the spherical velocity components

$$\mathbf{A}_{\text{sphere}} = \frac{du}{dt} \hat{\lambda} + \frac{dv}{dt} \hat{\phi} + \frac{dw}{dt} \hat{r}. \quad (10.66)$$

This term has no contribution from changes to the spherical unit vectors.

10.11.3 Metric acceleration

The metric acceleration arises from changes to the spherical unit vectors due to our use of spherical coordinates. It is given by

$$\mathbf{A}_{\text{metric}} = \hat{\lambda} \left(\frac{u(w - v \tan \phi)}{r} \right) + \hat{\phi} \left(\frac{v w + u^2 \tan \phi}{r} \right) - \hat{r} \left(\frac{u^2 + v^2}{r} \right) \quad (10.67\text{a})$$

$$= \hat{\lambda} \left(\frac{u(w \cos \phi - v \sin \phi)}{r \cos \phi} \right) + \hat{\phi} \left(\frac{v w \cos \phi + u^2 \sin \phi}{r \cos \phi} \right) - \hat{r} \left(\frac{u^2 + v^2}{r} \right) \quad (10.67\text{b})$$

$$= \frac{1}{r} [u \tan \phi (\hat{r} \wedge \mathbf{V}_{\text{sphere}}) + w \mathbf{U}_{\text{sphere}} - \hat{r} \mathbf{U}_{\text{sphere}} \cdot \mathbf{U}_{\text{sphere}}], \quad (10.67\text{c})$$

where we wrote the horizontal (angular) and vertical (radial) components of the spherical velocity according to

$$\mathbf{V}_{\text{sphere}} = \mathbf{U}_{\text{sphere}} + \hat{r} w = \hat{\lambda} u + \hat{\phi} v + \hat{r} w. \quad (10.68)$$

For purposes of developing a kinetic energy budget, note that

$$\mathbf{V}_{\text{sphere}} \cdot \mathbf{A}_{\text{metric}} = 0, \quad (10.69)$$

so that

$$\mathbf{V} \cdot \mathbf{A}_{\text{metric}} = \Omega u r_{\perp} (w - v \tan \phi) = \Omega u r (w \cos \phi - v \sin \phi). \quad (10.70)$$

10.11.4 Centrifugal acceleration

The spherical coordinate representation of the centrifugal acceleration is given by

$$\mathbf{A}_{\text{centrifugal}} = -\nabla \Phi_{\text{centrifugal}} \quad (10.71\text{a})$$

$$= \Omega^2 (x \hat{\mathbf{x}} + y \hat{\mathbf{y}}) \quad (10.71\text{b})$$

$$= r_{\perp} \Omega^2 (\hat{\phi} \sin \phi - \hat{r} \cos \phi). \quad (10.71\text{c})$$

For purposes of developing a kinetic energy budget, note that

$$\mathbf{V} \cdot \mathbf{A}_{\text{centrifugal}} = \mathbf{V}_{\text{sphere}} \cdot \mathbf{A}_{\text{centrifugal}} = -\Omega^2 r \cos \phi (w \cos \phi - v \sin \phi). \quad (10.72)$$

10.11.5 Coriolis acceleration

The spherical coordinate representation of the Coriolis acceleration makes use of

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}} = \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi) \quad (10.73)$$

to reach the form

$$\mathbf{A}_{\text{Coriolis}} = -2 \boldsymbol{\Omega} \wedge \mathbf{V}_{\text{sphere}} \quad (10.74\text{a})$$

$$= -2 \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi) \wedge \mathbf{V}_{\text{sphere}} \quad (10.74\text{b})$$

$$= -2 \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi) \wedge (u \hat{\lambda} + v \hat{\phi} + w \hat{r}) \quad (10.74\text{c})$$

$$= -2 \Omega [\hat{\lambda} (w \cos \phi - v \sin \phi) + \hat{\phi} u \sin \phi - \hat{r} u \cos \phi]. \quad (10.74\text{d})$$

Note that

$$\mathbf{V}_{\text{sphere}} \cdot \mathbf{A}_{\text{Coriolis}} = 0, \quad (10.75)$$

so that

$$\mathbf{V} \cdot \mathbf{A}_{\text{Coriolis}} = -2\Omega^2 r \cos \phi (w \cos \phi - v \sin \phi). \quad (10.76)$$

For future uses, we introduce a common shorthand for the angular rotation by defining

$$\mathbf{f} = 2\Omega \sin \phi \hat{\mathbf{r}} \quad (10.77)$$

$$\mathbf{f}^* = 2\Omega \cos \phi \hat{\boldsymbol{\phi}} \quad (10.78)$$

so that the Coriolis acceleration takes the form

$$\mathbf{A}_{\text{Coriolis}} = -(\mathbf{f} + \mathbf{f}^*) \wedge \mathbf{V}_{\text{sphere}}. \quad (10.79)$$

As discussed in Section 10.11.6, large-scale geophysical motions predominantly feel the radial component of the earth's rotation, in which case $\mathbf{A}_{\text{Coriolis}} \approx -\mathbf{f} \wedge \mathbf{V}_{\text{sphere}}$.

10.11.6 Coriolis acceleration for large-scale motions

Let us again write the Coriolis acceleration in equation (10.74d), only now underlining two terms

$$\mathbf{A}_{\text{Coriolis}} = -2\Omega \left[\hat{\boldsymbol{\lambda}} (\underline{w \cos \phi} - \underline{v \sin \phi}) + \hat{\boldsymbol{\phi}} \underline{u \sin \phi} - \hat{\mathbf{r}} \underline{u \cos \phi} \right]. \quad (10.80)$$

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 $\hat{\boldsymbol{\lambda}}$ component. Dropping these two terms results in the form for the Coriolis acceleration used for large-scale dynamics, such as when considering the hydrostatic primitive equations for geophysical fluids (Section 27.1)

$$\mathbf{A}_{\text{Coriolis}}^{\text{large-scale}} \equiv -2\Omega \sin \phi \left(-\hat{\boldsymbol{\lambda}} v + \hat{\boldsymbol{\phi}} u \right) \quad (10.81a)$$

$$\equiv -f \hat{\mathbf{r}} \wedge \mathbf{V}_{\text{sphere}}. \quad (10.81b)$$

For the last equality we introduced the Coriolis parameter

$$f \equiv 2\Omega \sin \phi. \quad (10.82)$$

As illustrated in Figure 10.5, 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

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}} = \Omega (\hat{\boldsymbol{\phi}} \cos \phi + \hat{\mathbf{r}} \sin \phi) \approx \Omega \sin \phi \hat{\mathbf{r}} = \mathbf{f}/2. \quad (10.83)$$

10.11.7 Further reading

Section 3.5 of [Apel \(1987\)](#) offers an insightful presentation of the Coriolis acceleration.

¹The term $\hat{\mathbf{r}} (2\Omega u \cos \phi)$ is called the Eötvös correction in the study of marine gravity.

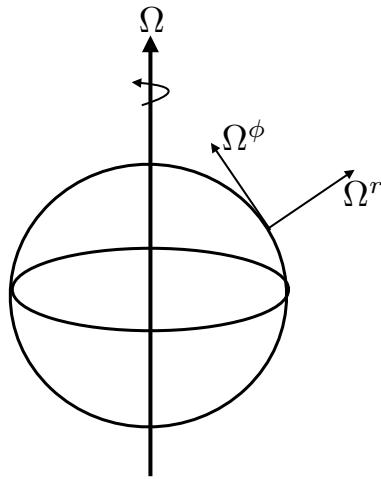


Figure 10.5: This figure illustrates the two components of the earth's rotational velocity, $\Omega = \Omega \hat{z} = \Omega (\hat{\phi} \cos \phi + \hat{r} \sin \phi)$. The local vertical component, $\Omega \approx \Omega \sin \phi \hat{r}$, is the most important component for large scale geophysical fluid dynamics.

10.12 Exercises

EXERCISE 10.1: WORKING THROUGH THE SPHERICAL ACCELERATION

Convince yourself that the spherical form of the acceleration given by equation (10.63) is indeed correct.

EXERCISE 10.2: VELOCITY AND ACCELERATION IN CYLINDRICAL-POLAR COORDINATES

In Section 7.3 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 34.5) 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

$$\mathbf{A} = \mathbf{A}_{\text{cylindrical-polar}} - \mathbf{A}_{\text{centrifugal}} - \mathbf{A}_{\text{Coriolis}}, \quad (10.84)$$

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 10.3: VELOCITY PROJECTED ONTO ACCELERATION

The kinetic energy per mass of a particle is given by

$$\mathcal{K} = \mathbf{V} \cdot \mathbf{V} / 2, \quad (10.85)$$

where \mathbf{V} is the inertial velocity of a particle. In an inertial reference frame it is trivial to show that

$$\frac{1}{2} \frac{d\mathcal{K}}{dt} = \mathbf{V} \cdot \mathbf{A} \quad (10.86)$$

through use of the chain rule, where $\mathbf{A} = \dot{\mathbf{V}}$ is the inertial acceleration. Verify that this identity also holds in the rotating reference frame. For simplicity you can make use of Cartesian coordinates.

11

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 mechanics as viewed in this non-inertial rotating frame. Doing so provides a useful introduction to rotating dynamics that will serve us when moving onto geophysical fluids.

READER'S GUIDE TO THIS CHAPTER

We make use of the mathematics and kinematics introduced for the particle motion in Chapter 10. Later chapters on rotating dynamics makes use of elements considered here, such as the Coriolis acceleration and details of the spherical coordinates.

11.1	Gravitational force and potential energy	127
11.1.1	Newton's Gravitational Law	128
11.1.2	Effective gravitational force from the geopotential	129
11.1.3	Further reading	129
11.2	Newton's second law	129
11.2.1	Cartesian coordinate representation	129
11.2.2	Spherical coordinate representation	130
11.2.3	Geopotential coordinates for slightly oblate spheriods	130
11.2.4	Further reading	132
11.3	Exercises	132

11.1 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.

11.1.1 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

$$P = m \Phi_e, \quad (11.1)$$

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

$$\Phi_e = -\frac{G M}{r} \quad (11.2)$$

where M is the mass of the sphere and

$$G = 6.674 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2} = 6.674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2} \quad (11.3)$$

is Newton's gravitational constant. The gradient of the gravitational potential gives the gravitational acceleration

$$\mathbf{g}_e = -\nabla \Phi_e = -\frac{G M}{r^2} \hat{\mathbf{r}} \quad (11.4)$$

so that the gravitational force acting on the particle is given by

$$\mathbf{F}_{\text{gravity}} = m \mathbf{g}_e = -m \nabla \Phi_e. \quad (11.5)$$

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

$$\mathbf{g}_e = -g_e \hat{\mathbf{r}}, \quad (11.6)$$

where

$$g_e = \frac{G M_e}{R_e^2} \approx 9.8 \text{ m s}^{-2}. \quad (11.7)$$

To reach this value, we assumed a sphere of mass equal to the earth mass

$$M_e = 5.977 \times 10^{24} \text{ kg}, \quad (11.8)$$

and radius

$$R_e = 6.371 \times 10^6 \text{ m} \quad (11.9)$$

determined so that the sphere has the same volume as the earth. The corresponding gravitational potential is given by

$$\Phi_e = g_e r, \quad (11.10)$$

and the gravitational potential energy is

$$m \Phi_e = m g_e r. \quad (11.11)$$

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 32.

11.1.2 Effective gravitational force from the geopotential

We can combine the potential for the centrifugal acceleration as given by equation (10.52c) with the gravitational potential (11.2), thus resulting in the *geopotential*

$$\Phi = r [g_e - \mathbf{U}_{\text{solid}}^2/(2r)]. \quad (11.12)$$

We estimate the contribution from the centrifugal term by making use of terrestrial values, in which $R = R_e = 6.371 \times 10^6$ m (equation (11.9)), and from Section 10.1

$$\Omega_e = 7.292 \times 10^{-5} \text{ s}^{-1}. \quad (11.13)$$

The centrifugal term is its largest at the equator, $\phi = 0$, where

$$\frac{\mathbf{U}_{\text{solid}}^2}{2R_e} \approx 0.017 \text{ m s}^{-2}, \quad (11.14)$$

so that the ratio of the gravitational to centrifugal accelerations is (at most)

$$\frac{g_e}{\mathbf{U}_{\text{solid}}^2/(2R_e)} = \frac{M_e G / R_e^2}{\Omega_e^2 R_e / 2} \approx 576. \quad (11.15)$$

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 11.2.3.

11.1.3 Further reading

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 \(1982\)](#).

11.2 Newton's second law

Newton's second law says that in an inertial reference frame, the time changes to the linear momentum of a system arise 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 second law says

$$m \mathbf{A} = -m \nabla\Phi_e. \quad (11.16)$$

We now move to the rotating reference frame of terrestrial observers, where Coriolis and centrifugal accelerations appear.

11.2.1 Cartesian coordinate representation

The inertial acceleration using Cartesian coordinate is given by equation (10.50b)

$$\mathbf{A} = \mathbf{A}_{\text{Cartesian}} - \mathbf{A}_{\text{Coriolis}} - \mathbf{A}_{\text{centrifugal}} \quad (11.17a)$$

$$= \mathbf{A}_{\text{Cartesian}} + 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} + \nabla\Phi_{\text{centrifugal}}, \quad (11.17b)$$

so that the rotating frame Cartesian equation of motion is given by

$$\mathbf{A}_{\text{Cartesian}} = -\nabla\Phi_e - 2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \nabla\Phi_{\text{centrifugal}} \quad (11.18a)$$

$$= -2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{Cartesian}} - \nabla\Phi, \quad (11.18b)$$

where the geopotential is the sum of the gravitational and centrifugal potentials (equation (11.12))

$$\Phi = \Phi_e + \Phi_{\text{centrifugal}}. \quad (11.19)$$

We can write the equation of motion in component form by exposing the indices (Cartesian tensors) and using a dot for time derivative

$$\ddot{X}_a + 2\epsilon_{abc}\boldsymbol{\Omega}_b \dot{X}_c = -\partial_a\Phi. \quad (11.20)$$

We can write this equation in the standard vector form

$$\frac{d^2\mathbf{X}}{dt^2} + 2\boldsymbol{\Omega} \wedge \dot{\mathbf{X}} = -\nabla\Phi, \quad (11.21)$$

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 26).

11.2.2 Spherical coordinate representation

We now make use of the acceleration written in spherical coordinates as given in Section 10.11

$$\mathbf{A}_{\text{sphere}} + \mathbf{A}_{\text{metric}} = \mathbf{A}_{\text{Coriolis}} + \mathbf{A} + \mathbf{A}_{\text{centrifugal}} \quad (11.22a)$$

$$= -2\boldsymbol{\Omega} \wedge \mathbf{V}_{\text{sphere}} - \nabla\Phi. \quad (11.22b)$$

Notice that the effective gravitational force is not a central force due to the contribution from the centrifugal acceleration. We see this more explicitly by using the equations in Section 10.11 to write the spherical equations

$$\dot{u} + \frac{u(w - v \tan\phi)}{r} + 2\Omega(w \cos\phi - v \sin\phi) = 0 \quad (11.23)$$

$$\dot{v} + \frac{v w + u^2 \tan\phi}{r} + 2\Omega u \sin\phi = -r_\perp \Omega^2 \sin\phi \quad (11.24)$$

$$\dot{w} - \frac{u^2 + v^2}{r} - 2\Omega u \cos\phi = r_\perp \Omega^2 \cos\phi - g_e. \quad (11.25)$$

11.2.3 Geopotential coordinates for slightly oblate spheroids

The radius of a sphere that best fits the volume of the earth is given by

$$R_e = 6.367 \times 10^6 \text{ m}. \quad (11.26)$$

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

$$R_{\text{equator}} = 6.378 \times 10^6 \text{ m} \quad (11.27)$$

$$R_{\text{pole}} = 6.357 \times 10^6 \text{ m}, \quad (11.28)$$

with a corresponding ratio

$$1 - \frac{R_{\text{pole}}}{R_{\text{equator}}} \approx 0.3\%. \quad (11.29)$$

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 fluid 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 11.1. Absorbing the centrifugal term into an effective gravitational potential then leads to the effective gravitational acceleration vector

$$-\nabla\Phi = -g \hat{\mathbf{r}}, \quad (11.30)$$

with g the effective gravitational acceleration. Using this convention, the particle equations of motion take the following form

$$\dot{u} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) = 0 \quad (11.31)$$

$$\dot{v} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi = 0 \quad (11.32)$$

$$\dot{w} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi = -g. \quad (11.33)$$

Notably, the effective gravitational acceleration only impacts the radial equation of motion.

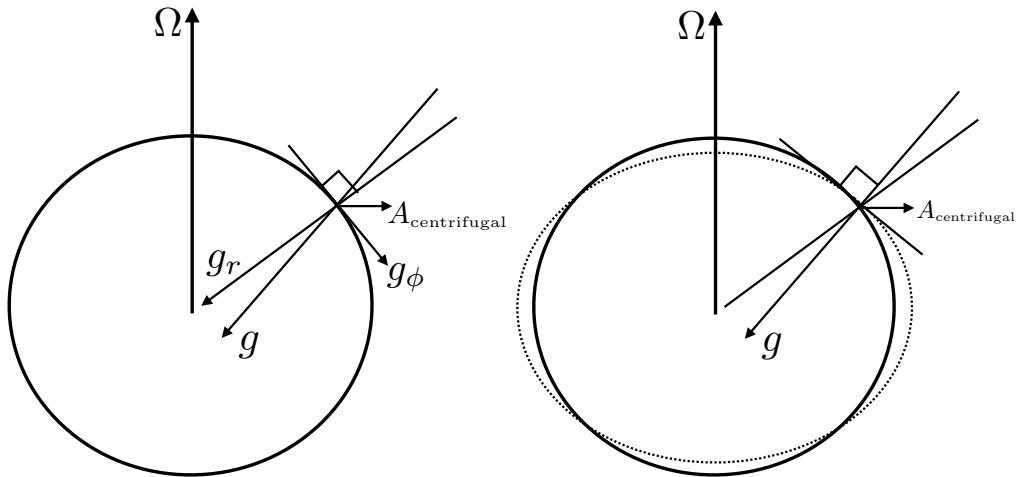


Figure 11.1: This figure illustrates the geopotential vertical coordinate system used to study geophysical fluids. The left panel shows the non-central nature of the effective gravitational force \mathbf{g} , which is given by the sum of the gravitational acceleration \mathbf{g}_r plus centrifugal acceleration $\mathbf{A}_{\text{centrifugal}}$. The 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 (11.29)).

11.2.4 Further reading

Section 2.2 of [Vallis \(2017\)](#) offers a thorough discussion of the spherical coordinate equations of motion for a fluid. 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.

11.3 Exercises

EXERCISE 11.1: GEOMETRY OF CONSTANT GEOPOTENTIAL SURFACES

In the geopotential (equation (11.12)), the squared solid-body velocity is given by $\mathbf{U}_{\text{solid}}^2 = (\Omega r \cos \phi)^2$.

- (A) Sketch surfaces of constant geopotential.
- (B) Discuss the geometry of these surfaces.
- (C) For terrestrial values of g_e and Ω , what is the ratio of the circumference around a geopotential at $\phi = 0$ (equator) relative to around $\phi = \pi/2$ (poles).

EXERCISE 11.2: 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 \(2017\)](#).

12

Symmetries and conservation laws

We here explore consequences for particle motion due to various conservation laws resulting from symmetries of the particle system. These ideas are fundamental to how we garner a qualitative understanding of motion, whereby it is often more useful to know the conservation properties of the motion than the trajectory.

READER'S GUIDE TO THIS CHAPTER

This chapter is an extension of the particle dynamics discussed in Chapter 11. We make use of the ideas introduced in this chapter when considering the motion of a geophysical fluid in Chapter 26.

12.1	Trajectories and dynamical constraints	134
12.1.1	Connecting symmetries to conservation laws	134
12.1.2	Further reading	134
12.2	Potential momentum	135
12.2.1	Potential momentum and motion along geopotential surfaces	135
12.2.2	Summary of dynamical constraints arising from geometric symmetries	135
12.3	Inertial oscillations	136
12.3.1	Oscillator equation	137
12.3.2	Particle trajectory and velocity	137
12.3.3	Period of the inertial oscillations	137
12.3.4	Comments and further reading	138
12.4	Kinetic energy	138
12.4.1	Cartesian expression	138
12.4.2	Kinetic energy using spherical coordinates: Part I	138
12.4.3	Kinetic energy using spherical coordinates: Part II	139
12.5	Mechanical energy conservation	139
12.5.1	General expression for kinetic energy evolution	139
12.5.2	Conservation of mechanical energy	139
12.5.3	Comments and further reading	140
12.6	Axial angular momentum conservation	140
12.6.1	Angular momentum	141
12.6.2	Conservation of axial angular momentum	141
12.7	Coriolis acceleration and angular momentum conservation	142
12.7.1	Easterly acceleration induced by meridionally poleward motion	142
12.7.2	Easterly acceleration induced by radially inward motion	143

12.7.3 Zonal acceleration implied by angular momentum conservation . . .	143
12.7.4 Nearly horizontal motions	144

12.1 Trajectories and dynamical constraints

From Newton's second 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 thus encapsulates all dynamical information about the moving particle. However, it is often difficult to unpack that information to understand the nature of the motion. That is, knowledge of the trajectory may not be 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.

12.1.1 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 uncover conservation laws through the connection between conservation laws and symmetries. For example, does the physical system remain unchanged when shifting the origin of time? If so, then there is an energy conservation law. 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 may be 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 26). Additional conservation properties also arise that are unique to the continuum, with conservation of potential vorticity the most notable one for geophysical fluids (Chapter 38).

12.1.2 Further reading

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 Space Time lecture](#).

12.2 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.

12.2.1 Potential momentum and motion along geopotential surfaces

We generally assume that the rotation of the planet is a constant in time, so that the Cartesian equation of motion (11.21) can be written

$$\frac{d}{dt} \left(\dot{\mathbf{X}} + 2\boldsymbol{\Omega} \wedge \mathbf{X} \right) = -\nabla\Phi. \quad (12.1)$$

This form suggests we introduce the *potential momentum* per mass

$$\mathbf{M} \equiv \frac{d\mathbf{X}}{dt} + 2\boldsymbol{\Omega} \wedge \mathbf{X} = \hat{\mathbf{x}}(u - 2\Omega y) + \hat{\mathbf{x}}(v + 2\Omega x) + \hat{\mathbf{z}}w, \quad (12.2)$$

in which case the momentum equation is written

$$\frac{d\mathbf{M}}{dt} = -\nabla\Phi. \quad (12.3)$$

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 (12.3) leads to

$$\frac{d(\hat{\mathbf{s}} \cdot \mathbf{M})}{dt} = 0. \quad (12.4)$$

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 a constant geopotential. Noether's Theorem (Section 12.1.1) 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.

To help understand the dynamical constraint imposed by constant potential momentum, consider a thought experiment where a particle with potential momentum \mathbf{M} moves 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.

12.2.2 Summary of 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.

Conservation of linear momentum from spatial translation symmetry

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

$$\mathbf{A} = 0. \quad (12.5)$$

When viewed from a rotating frame using Cartesian coordinates, a vanishing inertial acceleration means that the Cartesian acceleration balances Coriolis and centrifugal accelerations

$$\ddot{\mathbf{X}} = -2\boldsymbol{\Omega} \wedge \dot{\mathbf{X}} - \nabla\Phi_{\text{centrifugal}}. \quad (12.6)$$

This equation can be written in terms of the potential momentum

$$\frac{d\mathbf{M}}{dt} = -\nabla\Phi_{\text{centrifugal}}. \quad (12.7)$$

Note that if the motion occurs along a surface of constant centrifugal potential, then the potential momentum remains constant.

Conservation of potential momentum from spatial translation symmetry along a geopotential

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 (12.4))

$$\frac{d(\hat{\mathbf{s}} \cdot \mathbf{M})}{dt} = 0. \quad (12.8)$$

A geopotential is a two-dimensional surface so that this conservation law corresponds to two dynamical constraints.

Conservation of angular momentum from axial rotation symmetry

As detailed in Section 12.6, the angular momentum computed with respect to the axis of rotation is a constant of the motion. This constant arises from rotational symmetry of the system about the rotational axis. The axial angular momentum conservation takes the form

$$\frac{dL^z}{dt} = 0, \quad (12.9)$$

where the axial angular momentum is

$$L^z = m r_{\perp}^2 (\dot{\lambda} + \Omega), \quad (12.10)$$

$r_{\perp} = \sqrt{x^2 + y^2}$ is the distance from the rotation axis (the *moment arm*), and λ measures the angle in the counter-clockwise direction from the positive x -axis.

12.3 Inertial oscillations

In Section 27.3 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

$$f = 2\Omega \sin \phi_0 \quad (12.11)$$

to a constant, where ϕ_0 is a chosen latitude. Consequently, a particle moving on the f -plane maintains a constant horizontal potential momentum

$$\frac{dM_x}{dt} = \frac{d(u - f y)}{dt} = 0 \quad (12.12a)$$

$$\frac{dM_y}{dt} = \frac{d(v + f x)}{dt} = 0, \quad (12.12b)$$

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.

12.3.1 Oscillator equation

Taking the time derivative of the zonal equation (12.12a) and using the meridional equation (12.12b) leads to

$$\ddot{u} - f \dot{v} = \ddot{u} + f^2 u = 0. \quad (12.13)$$

Similar manipulations for the meridional velocity equation render the free oscillator equation for each component of the horizontal velocity

$$\frac{d^2 \mathbf{U}}{dt^2} + f^2 \mathbf{U} = 0. \quad (12.14)$$

12.3.2 Particle trajectory and velocity

Time integrating the equation of motion (12.14) renders the particle trajectory and its velocity

$$\mathbf{X}(t) = \frac{U}{f} \left[\hat{\mathbf{x}} \sin(ft) + \hat{\mathbf{y}} \cos(ft) \right] \quad (12.15a)$$

$$\mathbf{U}(t) = U \left[\hat{\mathbf{x}} \cos(ft) - \hat{\mathbf{y}} \sin(ft) \right], \quad (12.15b)$$

where U is the particle speed (which is a constant) and we assumed the initial conditions

$$\mathbf{X}(0) = \frac{U}{f} \hat{\mathbf{y}} \quad (12.16a)$$

$$\mathbf{U}(0) = U \hat{\mathbf{x}}. \quad (12.16b)$$

Motion is circular with radius

$$R = |U| f^{-1}. \quad (12.17)$$

For $f > 0$ (northern hemisphere) motion occurs in the clockwise direction, whereas southern hemisphere motion occurs in the counter-clockwise direction. Consequently, particle motion undergoing inertial oscillations occurs in an anti-cyclonic sense (opposite to the sense of the rotating reference frame). Finally, note that the potential momentum for these oscillations vanishes since

$$\mathbf{M}(t) = \mathbf{U}(t) + f \hat{\mathbf{z}} \wedge \mathbf{X}(t) = 0. \quad (12.18)$$

Adding an arbitrary constant to the initial position makes the potential momentum equal to a nonzero constant.

12.3.3 Period of the inertial oscillations

Motions that satisfy equation (12.14) are termed *inertial oscillations*. These simple harmonic oscillatory motions have a period given by

$$T_{\text{inertial}} = \frac{2\pi}{f} = \frac{11.97}{|\sin \phi_0|} \text{ hour}, \quad (12.19)$$

where we set $\Omega = 7.292 \times 10^{-5} \text{s}^{-1}$ (equation (10.1)). The period of inertial oscillations is smallest at the poles, where the latitude $\phi_0 = \pm\pi/2$ and $T_{\text{smallest}} \approx 12$ hour. At the equator, $\phi_0 = 0$ and no oscillatory motion is available.

12.3.4 Comments and further reading

Inertial oscillations of fluid parcels are described by the above constant potential momentum equations 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 much 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.

12.4 Kinetic energy

Kinetic energy is the energy contained in the motion of the particle with respect to an inertial reference frame

$$\mathcal{K} = \frac{m}{2} \mathbf{V} \cdot \mathbf{V} \quad (12.20)$$

12.4.1 Cartesian expression

With the inertial velocity given in the Cartesian form by

$$\mathbf{V} = \mathbf{V}_{\text{Cartesian}} + \mathbf{U}_{\text{solid}}, \quad (12.21)$$

the kinetic energy takes the form

$$\mathcal{K} = \frac{m}{2} (\mathbf{V}_{\text{Cartesian}} \cdot \mathbf{V}_{\text{Cartesian}} + 2 \mathbf{V}_{\text{Cartesian}} \cdot \mathbf{U}_{\text{solid}} + \mathbf{U}_{\text{solid}} \cdot \mathbf{U}_{\text{solid}}) \quad (12.22)$$

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.

12.4.2 Kinetic energy using spherical coordinates: Part I

To expose spherical symmetry of the physical system, we express the kinetic energy in terms of the spherical coordinates defined in Figure 10.4. Doing so for the solid body velocity leads to equation (10.41)

$$\mathbf{U}_{\text{solid}} = \Omega r \cos \phi (-\sin \lambda \hat{\mathbf{x}} + \cos \lambda \hat{\mathbf{y}}). \quad (12.23)$$

Likewise, the velocity components measured in the rotating frame are given by

$$\dot{X} = \frac{d(r \cos \phi \cos \lambda)}{dt} = \dot{r} \cos \phi \cos \lambda - r \dot{\phi} \sin \phi \cos \lambda - r \dot{\lambda} \cos \phi \sin \lambda \quad (12.24a)$$

$$\dot{Y} = \frac{d(r \cos \phi \sin \lambda)}{dt} = \dot{r} \cos \phi \sin \lambda - r \dot{\phi} \sin \phi \sin \lambda + r \dot{\lambda} \cos \phi \cos \lambda \quad (12.24b)$$

$$\dot{Z} = \frac{d(r \sin \phi)}{dt} = \dot{r} \sin \phi + r \dot{\phi} \cos \phi. \quad (12.24c)$$

Bringing terms together then leads to the kinetic energy in the form

$$\mathcal{K} = (m/2) \left[(\dot{r}^2 + r^2 \dot{\phi}^2 + \dot{\lambda}^2 r^2 \cos^2 \phi) + (2 \Omega r^2 \dot{\lambda} \cos^2 \phi) + (\Omega r \cos \phi)^2 \right]. \quad (12.25)$$

12.4.3 Kinetic energy using spherical coordinates: Part II

An alternative means for deriving the kinetic energy in equation (12.25) makes use of the spherical coordinate form of the inertial velocity given by equation (10.37f), in which case

$$\mathbf{V} = (u + r_{\perp} \Omega) \hat{\lambda} + v \hat{\phi} + w \hat{r}, \quad (12.26)$$

so that

$$\mathcal{K} = \frac{m}{2} [(u + r_{\perp} \Omega)^2 + v^2 + w^2], \quad (12.27)$$

where $r_{\perp} = r \cos \phi$. Additionally, as discussed in Section 12.6, the axial angular momentum is given by

$$L^z = m (u + r_{\perp} \Omega). \quad (12.28)$$

Consequently, the kinetic energy can be written

$$\mathcal{K} = \frac{(L^z)^2}{2m} + \frac{m}{2} (v^2 + w^2). \quad (12.29)$$

12.5 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 continuous fluid discussed in Chapter 26.

12.5.1 General expression for kinetic energy evolution

The kinetic energy for a particle of mass m is given by

$$\mathcal{K} = \frac{m}{2} \mathbf{V} \cdot \mathbf{V}. \quad (12.30)$$

The time derivative of the kinetic energy is therefore given by

$$\frac{d\mathcal{K}}{dt} = m \mathbf{V} \cdot \frac{d\mathbf{V}}{dt} \quad (12.31a)$$

$$= m \mathbf{V} \cdot \mathbf{A} \quad (12.31b)$$

$$= -m \mathbf{V} \cdot \nabla \Phi_e, \quad (12.31c)$$

where we made use of the inertial acceleration given by equation (11.16).

12.5.2 Conservation of mechanical energy

The gravitational potential is given by (equation (11.10))

$$\Phi_e = g_e r, \quad (12.32)$$

so that

$$\frac{d\mathcal{K}}{dt} = -m \mathbf{V} \cdot \nabla \Phi_e = -m g_e \dot{r}. \quad (12.33)$$

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 (11.11))

$$\frac{d\mathcal{P}_e}{dt} = m g_e \dot{r}, \quad (12.34)$$

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

$$\frac{d(\mathcal{K} + \mathcal{P}_e)}{dt} = 0, \quad (12.35)$$

where the mechanical energy is the sum of the inertial kinetic energy plus the gravitational potential energy

$$m = \mathcal{K} + \mathcal{P}_e \quad (12.36a)$$

$$= \frac{m}{2} \mathbf{V} \cdot \mathbf{V} + m \Phi_e \quad (12.36b)$$

$$= \frac{m}{2} [(u + r_{\perp} \Omega)^2 + v^2 + w^2] + m g_e r \quad (12.36c)$$

$$= \frac{(L^z)^2}{2m} + \frac{m}{2} (v^2 + w^2) + m g_e r. \quad (12.36d)$$

12.5.3 Comments and further reading

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. That choice plays no role in the evolution of the objective properties such as the kinetic energy and gravitational potential 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 \(1988\)](#).

12.6 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 Exercise 26.1).

12.6.1 Angular momentum

The angular momentum of the particle, computed with respect to the origin of the sphere, is given by

$$\mathbf{L} = m \mathbf{X} \wedge \mathbf{V}. \quad (12.37)$$

This is the moment of the linear momentum computed relative to the origin. We write the angular momentum computed along the polar axis as

$$L^z = \mathbf{L} \cdot \hat{\mathbf{z}} \quad (12.38a)$$

$$= m (\mathbf{X} \wedge \mathbf{V}) \cdot \hat{\mathbf{z}} \quad (12.38b)$$

$$= m (\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{V} \quad (12.38c)$$

$$= m r \cos \phi (\hat{\lambda} \cdot \mathbf{V}) \quad (12.38d)$$

$$= m r_{\perp} (\hat{\lambda} \cdot \mathbf{V}). \quad (12.38e)$$

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)

$$r_{\perp} = r \cos \phi. \quad (12.39)$$

In deriving equation (12.38e), we made use of the identity

$$\hat{\mathbf{z}} \wedge \mathbf{X} = r_{\perp} \hat{\lambda}, \quad (12.40)$$

which is useful for proving that axial angular momentum is a constant of the motion.

We now write the axial angular momentum in equation (12.38e) in terms of the rotating frame quantities. To do so, introduce the inertial velocity written using spherical coordinates according to equation (10.37f), which yields

$$L^z = m r_{\perp} (\hat{\lambda} \cdot \mathbf{V}) \quad (12.41a)$$

$$= m r_{\perp}^2 (\dot{\lambda} + \Omega) \quad (12.41b)$$

$$= m r_{\perp} (u + r_{\perp} \Omega). \quad (12.41c)$$

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.

12.6.2 Conservation of axial angular momentum

The time derivative of the axial angular momentum is given by

$$\frac{dL^z}{dt} = m \frac{d}{dt} [(\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{V}] \quad (12.42a)$$

$$= m (\hat{\mathbf{z}} \wedge \mathbf{V}) \cdot \mathbf{V} + m (\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{A} \quad (12.42b)$$

$$= m (\hat{\mathbf{z}} \wedge \mathbf{X}) \cdot \mathbf{A} \quad (12.42c)$$

$$= m r_{\perp} \hat{\lambda} \cdot \mathbf{A}. \quad (12.42d)$$

To reach this result we noted that the polar axis direction, $\hat{\mathbf{z}}$, is time independent, and we used the identity (12.40) for the final step. The inertial acceleration arises just from the central-force gravitational field (equation (11.16))

$$\mathbf{A} = -\nabla \Phi_e = -g_e \hat{\mathbf{r}}. \quad (12.43)$$

Since $\hat{\lambda} \cdot \hat{r} = 0$, we have axial angular momentum conservation

$$\frac{dL^z}{dt} = 0. \quad (12.44)$$

12.7 Coriolis acceleration and angular momentum conservation

Axial angular momentum conservation greatly constrains rotational motion. These constraints in turn offer the means for predicting what will happen to the particle if its motion is modified. We explore in this section how these constraints manifest through considering a few thought experiments. These ideas have direct relevance to the motion of a fluid on a smooth planet with no lateral boundaries.

To motivate the discussion, rewrite the axial angular momentum as

$$L^z = m r_{\perp}^2 (\dot{\lambda} + \Omega) = m r_{\perp} (u + r_{\perp} \Omega). \quad (12.45)$$

Axial angular momentum conservation says that it is not possible to change u or $\dot{\lambda}$ without also changing r_{\perp} . More precisely, fixing L^z leads to the constrained changes

$$\delta L^z = 0 \Rightarrow \frac{\delta(\dot{\lambda} + \Omega)}{\dot{\lambda} + \Omega} = -\frac{2 \delta r_{\perp}}{r_{\perp}}. \quad (12.46)$$

Furthermore, since $r_{\perp} = r \cos \phi$, the axial distance can change by either the radial position or the meridional position:

$$\delta r_{\perp} = (\cos \phi) \delta r - (r \sin \phi) \delta \phi, \quad (12.47)$$

so that

$$\frac{\delta(\dot{\lambda} + \Omega)}{\dot{\lambda} + \Omega} = -\left(\frac{2[(\cos \phi) \delta r - (r \sin \phi) \delta \phi]}{r_{\perp}} \right). \quad (12.48)$$

12.7.1 Easterly acceleration induced by meridionally poleward motion

Perturb the position of the particle in a poleward direction while maintaining a constant radial position ($\delta r = 0$). Poleward motion corresponds to a poleward velocity, $v \neq 0$. As it moves towards the poles, the particle reduces its distance from the rotation axis ($\delta r_{\perp} < 0$). To conserve axial angular momentum, this poleward motion induces a positive zonal acceleration, $\ddot{\lambda} > 0$. For poleward motion in the northern hemisphere, the induced zonal acceleration deflects the particle to the right. In the southern hemisphere, poleward motion induces a zonal acceleration that deflects the particle to the left.

The orientation of the deflection is the same as that arising from the Coriolis acceleration (rightward in the northern hemisphere and leftward in the southern). We see this correspondence from the zonal momentum equation (11.31)

$$\dot{u} + \frac{u(w - v \tan \phi)}{r} = -2\Omega(w \cos \phi - \underline{v} \sin \phi). \quad (12.49)$$

In the northern hemisphere ($\sin \phi > 0$), the underlined term gives rise to a rightward Coriolis acceleration ($\dot{u} > 0$) when moving poleward ($v > 0$). We can thus interpret the Coriolis acceleration as arising from axial angular momentum conservation.

12.7.2 Easterly acceleration induced by radially inward motion

Perform the same thought experiment as above, only now perturb the position radially inward at a fixed latitude ($\delta r < 0$ and $\delta\phi = 0$). This vertical/radial motion reduces r_\perp since the particle gets closer to the rotation axis. Just as before, axial angular momentum conservation induces a positive zonal acceleration, $\ddot{\lambda} > 0$, as the particle moves toward the center of the sphere. The corresponding Coriolis acceleration is given by the underlined vertical velocity term

$$\dot{u} + \frac{u(w - v \tan \phi)}{r} = -2\Omega(\underline{w \cos \phi} - v \sin \phi). \quad (12.50)$$

When $w < 0$ as per radially inward motion, there is a corresponding positive zonal acceleration, $\dot{u} > 0$.

12.7.3 Zonal acceleration implied by 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 of the particle. We can derive the expression by taking the time derivative of the first form of the axial angular momentum in equation (12.45), in which case

$$\frac{1}{m} \frac{dL^z}{dt} = \frac{d[(r \cos \phi)^2 (\dot{\lambda} + \Omega)]}{dt} \quad (12.51a)$$

$$= 2(\dot{r} \cos \phi - r \dot{\phi} \sin \phi)(\dot{\lambda} r \cos \phi + r \Omega \cos \phi) + (r \cos \phi)^2 \ddot{\lambda} \quad (12.51b)$$

$$= 2(w \cos \phi - v \sin \phi)(u + r \Omega \cos \phi) + (r \cos \phi)^2 \ddot{\lambda}, \quad (12.51c)$$

where we introduced the (u, v, w) velocity components according to equation (10.39). With the zonal velocity $u = \dot{\lambda} r \cos \phi$, we have

$$r \cos \phi \ddot{\lambda} = \dot{u} + \frac{u}{r \cos \phi} (v \sin \phi - w \cos \phi), \quad (12.52)$$

so that equation (12.51c) thus takes the form

$$\frac{1}{m} \frac{dL^z}{dt} = 2(w \cos \phi - v \sin \phi)(u + r \Omega \cos \phi) + (r \cos \phi)^2 \ddot{\lambda} \quad (12.53a)$$

$$= (w \cos \phi - v \sin \phi)(u + 2r\Omega \cos \phi) + \dot{u}r \cos \phi. \quad (12.53b)$$

Setting $dL^z/dt = 0$ and rearranging then leads to a prognostic equation for the zonal velocity component; i.e., an equation for the zonal acceleration

$$\frac{du}{dt} = \left[\frac{u}{r \cos \phi} + 2\Omega \right] (v \sin \phi - w \cos \phi). \quad (12.54)$$

The same result can be obtained by performing the time derivative on the second form of the axial angular momentum in equation (12.45), in which case

$$\frac{1}{m} \frac{dL^z}{dt} = \frac{d[u r \cos \phi + \Omega(r \cos \phi)^2]}{dt} \quad (12.55a)$$

$$= \dot{u}r \cos \phi + u\dot{r} \cos \phi - u r \dot{\phi} \sin \phi + 2\Omega r \cos \phi(\dot{r} \cos \phi - r \dot{\phi} \sin \phi). \quad (12.55b)$$

Again, setting $dL^z/dt = 0$ and rearranging leads to the zonal velocity equation (12.54).

12.7.4 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.

The first assumption corresponds to the hydrostatic approximation (Section 27.2), and the second to the shallow fluid approximation (Section 27.1.2). Dropping the vertical velocity component to the kinetic energy leads to

$$\mathcal{K} \approx \frac{m}{2} [u^2 + v^2 + 2\Omega r_{\perp} u + (\Omega r_{\perp})^2]. \quad (12.56)$$

The second assumption means that the axial angular momentum takes the approximate form

$$L^z \approx m R_{\perp} (u + \Omega R_{\perp}) = m R_{\perp}^2 (\dot{\lambda} + \Omega), \quad (12.57)$$

where

$$r = R + z \approx R \quad (12.58)$$

$$R_{\perp} = R \cos \phi. \quad (12.59)$$

We comment further on the approximate form for the angular momentum given by equation (12.57). Namely, vertical motions at constant latitude no longer alter the angular momentum, since such movement does not alter than moment arm R_{\perp} . Hence, the only way to exchange angular momentum between the particle and the sphere is through meridional movement of the particle. Additionally, the general form of the zonal acceleration (12.54) is modified so that axial angular momentum conservation renders

$$\frac{du}{dt} = \frac{u v \tan \phi}{R} + f v, \quad (12.60)$$

where we introduced the Coriolis parameter $f = 2\Omega \sin \phi$. In effect, we dropped the vertical velocity component, w , from the general form of the acceleration (12.54). Correspondingly, the meridional momentum equation takes the form

$$\frac{dv}{dt} = -\frac{u^2 \tan \phi}{R} - f u. \quad (12.61)$$

Part III

Fluid kinematics

Kinematics is concerned with the intrinsic properties of motion, whereas dynamics is concerned with causes for the motion. 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).

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. Having two descriptions of the same motion provides a synergy that is missing with either alone. It is thus very useful to have skills at moving between the two, 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 14.

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 15 and 16. There are two special topic chapters that bookend this part. Chapter 13 provides a resume of the continuum hypothesis that forms the basis for ignoring molecular degrees of freedom in our treatment of fluid motion. Chapter 17 explores the Lagrangian kinematics of material lines, areas, and volumes.

13

The continuum hypothesis[†]

A fluid deforms continuously when applying a force so that a fluid has no preferred shape. Ordinary gases and liquids are the canonical examples of fluids. 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. For these reasons we pursue a phenomenological approach that makes use of conservation laws describing the motion of a continuous fluid media. In this chapter we present salient points concerning the use of a continuum description of fluid flow. It is offered for those curious about the physical assumptions underlying the treatment of fluids as continuous media.

READER'S GUIDE TO THIS CHAPTER

Elements of this chapter borrow from treatments of kinetic theory of gases in books such as [Reif \(1965\)](#) and [Huang \(1987\)](#). However, prior exposure to these treatments is not necessary. No other chapters depend explicitly on material in this chapter.

13.1	The brief version	147
13.1.1	Length scales	148
13.1.2	Continuous fields rather than discrete molecules	149
13.2	The longer version	149
13.2.1	A mole and Avogadro's number	149
13.2.2	Ideal gas law	150
13.2.3	Molecular mean free path	151
13.2.4	Root mean square molecular speed	151
13.2.5	Time between molecular collisions	152
13.2.6	Macroscopically small and microscopically large	152
13.2.7	Whence a rigorous treatment?	152
13.2.8	Further reading	153

13.1 The brief version

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

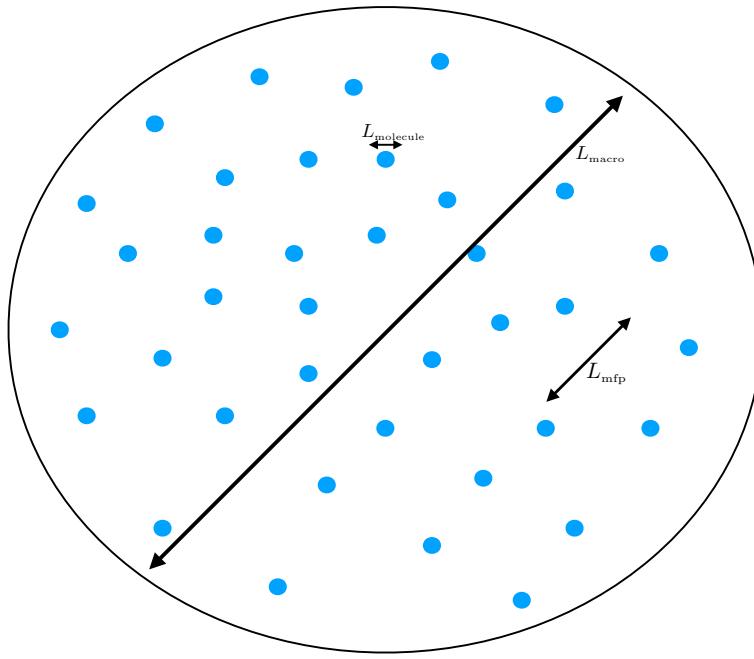


Figure 13.1: Schematic to illustrate the three length scales considered when making the continuum hypothesis. The blue circles represent molecules with diameter L_{molecule} . On average, molecules are separated by the mean free path, $L_{\text{mfp}} \approx 1000 L_{\text{molecule}}$. The smallest macroscopic 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!

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 13.2, although a full discussion is outside the subject of fluid mechanics.

13.1.1 Length scales

A fluid mechanical description focuses on fluid regions that are macroscopically small (e.g., $L_{\text{macro}} \sim 10^{-3} \text{ m}$) yet microscopically large (e.g., $L_{\text{macro}} \gg L_{\text{mfp}} \sim 10^{-7} \text{ m}$, 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 13.2) 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 13.1 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.

13.1.2 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 $L_{\text{mfp}}/v_{\text{rms}}$. 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

$$\text{Kn} = \frac{L_{\text{mfp}}}{L_{\text{macro}}}.$$
 (13.1)

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

$$\text{Kn} \ll 1.$$
 (13.2)

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

$$L_{\text{molecule}} \ll L \ll L_{\text{macro}}.$$
 (13.3)

As noted in Section 13.2.6, $L_{\text{macro}} \sim 10^{-3}$ m, which is much larger than the molecular mean free path.

13.2 The longer version

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, the remainder of this section offers an outline of 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 a topic outside our scope. In Section 13.2.8, we provide literature pointers for those wishing more rigor concerning the continuum limit of matter.

13.2.1 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

$$A_v = 6.022 \times 10^{23} \text{ mole}^{-1}.$$
 (13.4)

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

$$M_{\text{air}} = 0.22 * 32 \text{ g mole}^{-1} + 0.78 * 28 \text{ g mole}^{-1} \approx 28.8 \text{ g mole}^{-1}. \quad (13.5)$$

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

$$M_{\text{water}} = 2 * 1 \text{ g mole}^{-1} + 16 \text{ g mole}^{-1} = 18 \text{ g mole}^{-1}. \quad (13.6)$$

13.2.2 Ideal gas law

The ideal gas law is given by

$$pV = nRT, \quad (13.7)$$

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

$$R = 8.314 \text{ J mole}^{-1} \text{ K}^{-1} = 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1}, \quad (13.8)$$

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

$$\text{J} = \text{kg m}^2 \text{ s}^{-2}. \quad (13.9)$$

Use of the ideal gas law (13.7) 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

$$V = \frac{nRT}{p} \quad (13.10\text{a})$$

$$= \frac{(1 \text{ mole})(8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1})(273.15 \text{ K})}{101.325 \times 10^3 \text{ kg m}^{-1} \text{ s}^{-2}} \quad (13.10\text{b})$$

$$\approx 2.25 \times 10^{-2} \text{ m}^3, \quad (13.10\text{c})$$

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

$$\text{Pa} = \text{N m}^{-2} = \text{kg m}^{-1} \text{ s}^{-2}. \quad (13.11)$$

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

$$n_{\text{gas}} = \frac{\text{number per mole}}{\text{volume per mole}} \quad (13.12\text{a})$$

$$= \frac{A_v}{V} \quad (13.12\text{b})$$

$$= \frac{6.022 \times 10^{23}}{2.25 \times 10^{-2} \text{ m}^3} \quad (13.12\text{c})$$

$$= 2.68 \times 10^{25} \text{ m}^{-3}. \quad (13.12\text{d})$$

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.

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

$$\rho_{\text{air}} = \frac{M_{\text{air}}}{V} = \frac{28.8 \times 10^{-3} \text{ kg}}{2.25 \times 10^{-2} \text{ m}^3} = 1.28 \text{ kg m}^{-3}, \quad (13.13)$$

where we set $M_{\text{air}} = 28.8 \times 10^{-3}$ kg according to equation (13.5). 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).

13.2.3 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

$$L_{\text{mfp}} = \frac{1}{\pi \sqrt{2} n_v d^2} \quad (13.14)$$

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

$$d_{\text{molecule air}} \approx 2 \times 10^{-10} \text{ m}. \quad (13.15)$$

The mean free path for air molecules is thus given by

$$L_{\text{mfp}} = \frac{1}{\pi \sqrt{2} n_v d^2} \quad (13.16a)$$

$$= \frac{1}{\pi \sqrt{2} (2.68 \times 10^{25} \text{ m}^{-3}) (2 \times 10^{-10} \text{ m})^2} \quad (13.16b)$$

$$= 2 \times 10^{-7} \text{ m}. \quad (13.16c)$$

The mean free path for an air molecule is roughly 1000 times larger than the molecular diameter (e.g., Figure 13.1).

13.2.4 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

$$v_{\text{rms}} = \sqrt{\frac{3p}{\rho}} = \sqrt{\frac{3RT}{M}}. \quad (13.17)$$

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

$$v_{\text{rms}} = \sqrt{\frac{3p_{\text{stand}}}{\rho_{\text{air}}}} \quad (13.18a)$$

$$= \sqrt{\frac{3(101.325 \times 10^3 \text{ kg m}^{-1} \text{ s}^{-2})}{1.28 \text{ kg m}^{-3}}} \quad (13.18b)$$

$$= 487 \text{ m s}^{-1}. \quad (13.18c)$$

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!

13.2.5 Time between molecular collisions

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

$$t_{\text{collision}} = \frac{L_{\text{mfp}}}{v_{\text{rms}}} \quad (13.19)$$

The corresponding time for air is given by

$$t_{\text{air}} = \frac{2 \times 10^{-7} \text{ m}}{487 \text{ m s}^{-1}} = 4.1 \times 10^{-10} \text{ s}. \quad (13.20)$$

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

13.2.6 Macroscopically small and microscopically large

For environmental measurements of the atmosphere and ocean, or for conventional measurements in fluid dynamics 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

$$L_{\text{macro}} = 10^{-3} \text{ m}. \quad (13.21)$$

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}^3 contains

$$N_{\text{air molecules}} = V n_{\text{gas}} = (10^{-3} \text{ m})^3 (2.68 \times 10^{25} \text{ m}^{-3}) \approx 3 \times 10^{16} \text{ air molecules}. \quad (13.22)$$

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

$$M_{\text{water}} = \rho_{\text{water}} V = (1000 \text{ kg m}^{-3}) (10^{-9} \text{ m}^3) = 10^{-6} \text{ kg}. \quad (13.23)$$

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

$$N_{\text{water molecules}} = \left(\frac{10^{-6} \text{ kg}}{0.018 \text{ kg mole}^{-1}} \right) \times 6.022 \times 10^{23} \text{ molecules mole}^{-1} = 3 \times 10^{19} \text{ water molecules}. \quad (13.24)$$

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.

13.2.7 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.

13.2.8 Further reading

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 of 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.

14

Kinematics of fluid motion

Fluid motion is very complex. Hence, it is useful to avail oneself with more than one means for describing this 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 of 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. However, a Lagrangian description offers useful insights into the theoretical foundations of the subject. *Salmon (1998)* provides an elegant and accessible treatment of Lagrangian fluid mechanics, and *Bennett (2006)* provides a thorough 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 that in 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 in this book.

14.1	A variety of fluid systems	156
14.1.1	Defining the fluid systems	156
14.1.2	Comments	157
14.2	Lagrangian and Eulerian reference frames	158
14.3	Material and position coordinates	159
14.3.1	Trajectories of fluid particles	159
14.3.2	Material coordinates	159
14.3.3	Transforming between material and position coordinates	160
14.3.4	Transformation matrix and its Jacobian	161
14.3.5	A discrete algorithm to compute the transformation matrix	161
14.3.6	Jacobian as the ratio of volumes	162
14.4	Lagrangian and Eulerian time derivatives	163
14.4.1	Infinitesimal space-time increment of a function	164
14.4.2	Total time derivative of a function	164

14.4.3 Eulerian: evolution measured in the laboratory frame	164
14.4.4 Lagrangian: evolution measured in the material frame	165
14.4.5 Lagrangian time derivative formulated from the material frame	165
14.4.6 Sample material time derivative operations	166
14.5 Galilean invariance	168
14.5.1 Galilean transformation	168
14.5.2 Transformation of the material time derivative	168
14.6 Invariance of the material time derivative	169
14.6.1 Invariance based on definition of the material time derivative	170
14.6.2 Invariance for a rotating reference frame	170
14.6.3 Comments	172
14.7 Fluid flow lines	172
14.7.1 Material pathlines from fluid particle trajectories	172
14.7.2 Fluid streamlines and streamtubes	173
14.7.3 Emphasizing that streamlines are not pathlines in unsteady flow . . .	174
14.7.4 Fluid streaklines	174
14.7.5 An analytic example of flow lines	175
14.7.6 Further reading	177
14.8 Stokes drift	178
14.8.1 Formulation of Stokes drift	178
14.8.2 Stokes drift in the upper ocean	180
14.9 Exercises	183

14.1 A variety of fluid systems

The continuum hypothesis (Chapter 13) 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.

14.1.1 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 14.1). A fluid particle is uniquely specified by its material coordinate and time (see Section 14.3.1 for material coordinates). 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. Furthermore, fluid particle trajectories define the Lagrangian reference frame (Section 14.2).

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 14.1). A material parcel maintains a fixed mass, a fixed matter content, and a fixed specific entropy. Hence, it does not exchange matter or entropy with other fluid parcels. Furthermore, it does not experience irreversible exchange

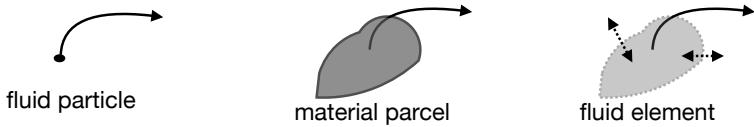


Figure 14.1: Schematic for the fluid systems used in our considerations of fluid kinematics. Left panel: fluid particle that tests the fluid flow field, \mathbf{v} , without altering it. Fluid particles have no mass and no extension. Their trajectories define the Lagrangian reference frame. Middle panel: 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 15.4), 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.

of momentum arising from friction. Its only interaction with adjacent fluid parcels is through reversible mechanical exchanges due 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.

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.

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 14.1). 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 15.4 and 22.2). As detailed in Section 15.4, diffusive matter exchange leaves the mass of the fluid element unchanged since the fluid element velocity is determined by its center of mass.

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 matter, mechanical, and thermodynamic exchanges with the surrounding environment.

14.1.2 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 15.1). 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 15 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 15.4), 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.

14.2 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 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 10).

The Eulerian and Lagrangian approaches nicely complement one another. For example, the Lagrangian approach lends itself to fruitful physical insights since we can borrow freely from point particle mechanics. In contrast, the Eulerian approach is often more straightforward when developing numerical methods for simulations. In general, we advocate the free use of both Eulerian and Lagrangian methods. 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.

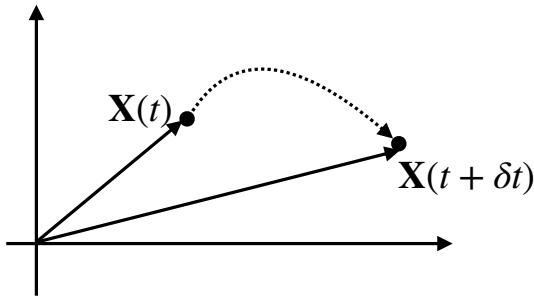


Figure 14.2: Sample trajectory of a fluid particle with endpoints $\mathbf{X}(t)$ and $\mathbf{X}(t + \delta t)$. The trajectory passes through the point $\mathbf{x} = \mathbf{X}(t)$ at time t and $\mathbf{x} + \delta\mathbf{x} = \mathbf{X}(t + \delta t)$ at time $t + \delta t$. Eulerian kinematics describes 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.

14.3 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.

14.3.1 Trajectories of fluid particles

In describing the motion of a classical point particle (Chapter 10), 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

$$\mathbf{x} = \mathbf{X}(t) \quad \text{point particle.} \quad (14.1)$$

A sample trajectory is show in Figure 14.2. 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

$$\mathbf{x} = \mathbf{X}(n, t) \quad n = 1, N \text{ point particles.} \quad (14.2)$$

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

$$\mathbf{x} = \mathbf{X}(\mathbf{a}, t) \quad \text{continuum of matter.} \quad (14.3)$$

The vector \mathbf{a} is referred to as the *material* coordinate. It labels a point of matter within the continuum fluid.

14.3.2 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

$$\mathbf{a} = \mathbf{X}(t = t_0). \quad (14.4)$$

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 20). 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

$$\mathbf{a} = (X, Y, \theta)_{t=t_0}, \quad (14.5)$$

where we write θ as a measure of the specific entropy (or potential temperature). 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 special topics Chapters 8, 18, and 31.

14.3.3 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

$$\mathbf{x} = \mathbf{X}(\mathbf{a}, t), \quad (14.6)$$

which is written in component form as¹

$$x^m = X^m(a^i, t). \quad (14.7)$$

In fluid mechanics, the relation (14.6) 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 (14.6)²

$$\mathbf{a} = \mathbf{A}(\mathbf{x}, t). \quad (14.8)$$

In this equation, \mathbf{A} is the inverse function that specifies the material coordinate \mathbf{a} given (\mathbf{x}, t) . Single-valued 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.

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 (14.8) should not be confused with the acceleration, also written as \mathbf{A} elsewhere in this book. We will not have much use for equation (14.8), thus minimizing the opportunity for confusion.

14.3.4 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 7.2.1, 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 14.5 and 14.6. 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 (14.6) 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

$$F_i^m = \frac{\partial X^m}{\partial a^i} \equiv \begin{bmatrix} \partial X^1 / \partial a^1 & \partial X^1 / \partial a^2 & \partial X^1 / \partial a^3 \\ \partial X^2 / \partial a^1 & \partial X^2 / \partial a^2 & \partial X^2 / \partial a^3 \\ \partial X^3 / \partial a^1 & \partial X^3 / \partial a^2 & \partial X^3 / \partial a^3 \end{bmatrix}. \quad (14.9)$$

As defined by equation (14.9), the upper label, m , denotes the row and the lower label, i , is the column. As seen in Section 17.2, the transformation tensor 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

$$\det(F_i^m) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \det \begin{bmatrix} \partial X^1 / \partial a^1 & \partial X^1 / \partial a^2 & \partial X^1 / \partial a^3 \\ \partial X^2 / \partial a^1 & \partial X^2 / \partial a^2 & \partial X^2 / \partial a^3 \\ \partial X^3 / \partial a^1 & \partial X^3 / \partial a^2 & \partial X^3 / \partial a^3 \end{bmatrix}. \quad (14.10)$$

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.

14.3.5 A discrete algorithm to compute the transformation matrix

To help further our understanding of the transformation matrix (14.9), 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

$$\mathbf{X}(t) = X^1(t) \hat{\mathbf{x}} + X^2(t) \hat{\mathbf{y}}, \quad (14.11)$$

and use a Cartesian representation for the material coordinate

$$\mathbf{a} = a^1 \hat{\mathbf{x}} + a^2 \hat{\mathbf{y}}. \quad (14.12)$$

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

$$\mathbf{x}(e, f) = x(e, f) \hat{\mathbf{x}} + y(e, f) \hat{\mathbf{y}}. \quad (14.13)$$

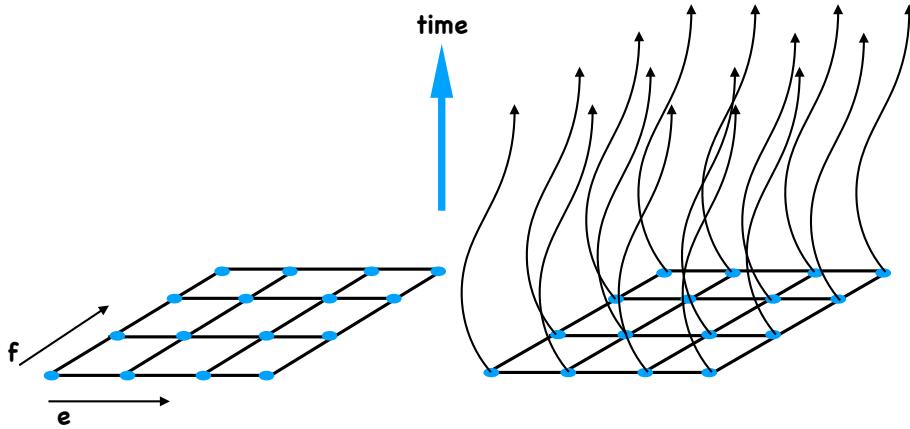


Figure 14.3: Illustrating the computational algorithm of Section 14.3.5 used to approximate the transformation matrix (also known as the deformation tensor) F_i^m . The left panel shows the two-dimensional grid with nodal points defining the initial positions for fluid particles. Each position on the 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.

Initialize fluid particles at each of the lattice grid points,

$$\mathbf{X}(e, f; t = 0) = \mathbf{x}(e, f) = \mathbf{a}(e, f), \quad (14.14)$$

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 14.3. 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 tensor

$$F_i^m = \begin{bmatrix} F_1^1 & F_2^1 \\ F_1^2 & F_2^2 \end{bmatrix} \approx \begin{bmatrix} \frac{X^1(e+1,f;t)-X^1(e-1,f;t)}{X^1(e+1,f;0)-X^1(e-1,f;0)} & \frac{X^1(e,f+1;t)-X^1(e,f-1;t)}{X^2(e,f+1;0)-X^2(e,f-1;0)} \\ \frac{X^2(e+1,f;t)-X^2(e-1,f;t)}{X^2(e+1,f;0)-X^2(e-1,f;0)} & \frac{X^2(e,f+1;t)-X^2(e,f-1;t)}{X^2(e,f+1;0)-X^2(e,f-1;0)} \end{bmatrix}. \quad (14.15)$$

If the grid is regular in both directions, then the initial positions have a separation given by the grid spacing in which case

$$F_i^m \approx \begin{bmatrix} \frac{X^1(e+1,f;t)-X^1(e-1,f;t)}{\Delta} & \frac{X^1(e,f+1;t)-X^1(e,f-1;t)}{\Delta} \\ \frac{X^2(e+1,f;t)-X^2(e-1,f;t)}{\Delta} & \frac{X^2(e,f+1;t)-X^2(e,f-1;t)}{\Delta} \end{bmatrix}. \quad (14.16)$$

This algorithm illustrates how the transformation matrix provides a measure of trajectory spreading as fluid particles move away from their initial positions.

14.3.6 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

$$dV(\mathbf{x}) = dx dy dz. \quad (14.17)$$

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

$$dV(\mathbf{a}) = da db dc, \quad (14.18)$$

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 (14.5)).

The two volumes are related by the Jacobian of transformation between the two coordinate systems

$$dV(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} dV(\mathbf{a}). \quad (14.19)$$

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

$$\frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{dV(\mathbf{x})}{dV(\mathbf{a})}. \quad (14.20)$$

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

$$\delta V(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \delta V(\mathbf{a}) \Rightarrow \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{\delta V(\mathbf{x})}{\delta V(\mathbf{a})}. \quad (14.21)$$

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

$$\frac{\partial \mathbf{X}}{\partial \mathbf{X}(0)} = \frac{\delta V(\mathbf{x})}{\delta V(0)}. \quad (14.22)$$

14.4 Lagrangian and Eulerian time derivatives

As noted in Section 14.3.2, 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.

14.4.1 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

$$\Phi = \Phi(\mathbf{x}, t). \quad (14.23)$$

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

$$d\Phi = \Phi(\mathbf{x} + d\mathbf{x}, t + dt) - \Phi(\mathbf{x}, t) \quad (14.24a)$$

$$= dt \frac{\partial \Phi}{\partial t} + d\mathbf{x} \cdot \nabla \Phi. \quad (14.24b)$$

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

$$d\mathbf{x} = \hat{\mathbf{x}} dx + \hat{\mathbf{y}} dy + \hat{\mathbf{z}} dz. \quad (14.25)$$

We ignore higher order terms in equation (14.24b) since the space and time increments are infinitesimal.

14.4.2 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 (14.24b) by the infinitesimal time increment

$$\frac{d\Phi}{dt} = \frac{\partial \Phi}{\partial t} + \frac{d\mathbf{x}}{dt} \cdot \nabla \Phi. \quad (14.26)$$

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 (14.26) holds in general since the velocity of the moving frame is arbitrary. We next specialize to the two common cases in fluid mechanics.

14.4.3 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

$$\text{Eulerian time derivative} = \frac{\partial \Phi(\mathbf{x}, t)}{\partial t}. \quad (14.27)$$

This result follows from specializing the total time derivative in equation (14.26) 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*. When the Eulerian time derivative vanishes, the field is said to be in a *steady state*.

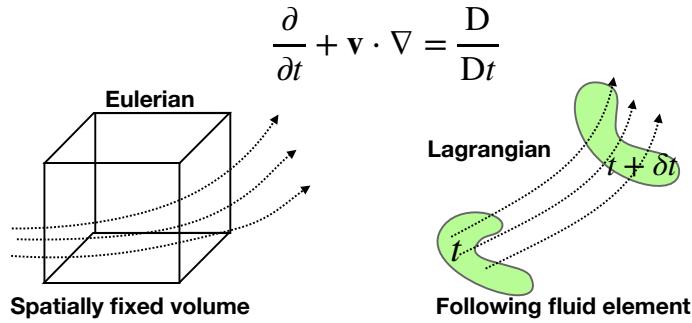


Figure 14.4: Illustrating the distinctions between the Eulerian (laboratory) and Lagrangian (material) reference frames for describing fluid motion. For the Eulerian description we consider a fixed control volume in the laboratory frame and observe fluid moving through the volume. For the Lagrangian 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.

14.4.4 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

$$\text{Lagrangian time derivative} = \frac{D\Phi}{Dt} = \frac{\partial\Phi}{\partial t} + \mathbf{v} \cdot \nabla\Phi. \quad (14.28)$$

The second equality follows by setting $d\mathbf{x}/dt = \mathbf{v}$ in equation (14.26), 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 (14.27), whereas $\mathbf{v} \cdot \nabla$ is referred to as the *advection* operator. Use of the capital D for the material time operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (14.29)$$

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 (14.26).

Equation (14.29) 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 14.4 illustrates the differences between the Eulerian and Lagrangian perspectives.

14.4.5 Lagrangian time derivative formulated from the material frame

Rather than start from the total time derivative (14.26), 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

$$\Phi[\mathbf{X}(\mathbf{a}, t), t] \equiv \Phi^L(\mathbf{a}, t) \quad \Leftarrow \text{sampling } \Phi \text{ on a trajectory } \mathbf{x} = \mathbf{X}(\mathbf{a}, t) \text{ at time } t. \quad (14.30)$$

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

$$\mathbf{v}^L(\mathbf{a}, t) \equiv \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] = \left[\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \right]_{\mathbf{a}}. \quad (14.31)$$

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

$$\left[\frac{\partial \Phi[\mathbf{X}(\mathbf{a}, t), t]}{\partial t} \right]_{\mathbf{a}} = \lim_{\Delta t \rightarrow 0} \left[\frac{\Phi[\mathbf{X}(\mathbf{a}, t + \Delta t/2), t + \Delta t/2] - \Phi[\mathbf{X}(\mathbf{a}, t - \Delta t/2), t - \Delta t/2]}{\Delta t} \right]. \quad (14.32)$$

Expanding the numerator in a Taylor series, and keeping just the leading terms, yields

$$\left[\frac{\partial \Phi[\mathbf{X}(\mathbf{a}, t), t]}{\partial t} \right]_{\mathbf{a}} = \lim_{\Delta t \rightarrow 0} \left[\frac{\Phi[\mathbf{X}(\mathbf{a}, t + \Delta t/2), t + \Delta t/2] - \Phi[\mathbf{X}(\mathbf{a}, t - \Delta t/2), t - \Delta t/2]}{\Delta t} \right] \quad (14.33a)$$

$$= \left[\left(\frac{\partial}{\partial t} \right)_{\mathbf{X}} + \left(\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \right)_{\mathbf{a}} \cdot \nabla \right] \Phi[\mathbf{X}(\mathbf{a}, t), t] \quad (14.33b)$$

$$= \left[\left(\frac{\partial}{\partial t} \right)_{\mathbf{X}} + \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \cdot \nabla \right] \Phi[\mathbf{X}(\mathbf{a}, t), t]. \quad (14.33c)$$

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 (14.28). Even so, it is important to keep in mind the underlying trajectory basis for the material time derivative.

14.4.6 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 advection equation (Section 1.2). Now consider the material time derivative

$$\frac{D(\mathbf{x} - \mathbf{c}t)}{Dt} = \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] (\mathbf{x} - \mathbf{c}t) \quad (14.34a)$$

$$= (\mathbf{v} \cdot \nabla) \mathbf{x} - \frac{\partial(t \mathbf{c})}{\partial t} \quad (14.34b)$$

$$= \mathbf{v}(\mathbf{x}, t) - \mathbf{c}. \quad (14.34c)$$

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 speed, \mathbf{c} , then surfaces of constant Φ are material. However, such constant phase surfaces are generally not material since fluid particles generally do not move with the wave phase velocity.

Materially constant function

Consider a scalar function that remains constant on a material trajectory so that its material time derivative vanishes

$$\frac{D\Phi}{Dt} = 0. \quad (14.35)$$

Material constancy means that at a fixed point in space, the Eulerian time derivative is affected only by advection

$$\frac{\partial\Phi}{\partial t} = -\mathbf{v} \cdot \nabla\Phi. \quad (14.36)$$

As just seen for the phase of a linear wave, $\partial\Phi/\partial t = \mathbf{c}$, so that constant wave phase surfaces are material if they move with velocity $\mathbf{v} = \mathbf{c}$.

We garner geometric insight into relation (14.36) by introducing the unit normal vector to the surface of constant Φ

$$\hat{\mathbf{n}} = \frac{\nabla\Phi}{|\nabla\Phi|}. \quad (14.37)$$

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

$$\frac{\partial\Phi/\partial t}{|\nabla\Phi|} = -\mathbf{v} \cdot \hat{\mathbf{n}}. \quad (14.38)$$

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 15.7.2, 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 (14.26) for the total time derivative, we have the time derivative operator as measured in this non-material moving reference frame

$$\frac{D^{(s)}}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla. \quad (14.39)$$

A function that remains constant within this general moving frame thus satisfies

$$\frac{D^{(s)}\Phi}{Dt} = 0 \Rightarrow \frac{\partial\Phi}{\partial t} = -\mathbf{v}^{(s)} \cdot \Phi. \quad (14.40)$$

Introducing the normal direction $\hat{\mathbf{n}} = |\nabla\Phi|^{-1} \nabla\Phi$ leads to

$$\frac{\partial\Phi/\partial t}{|\nabla\Phi|} = -\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}, \quad (14.41)$$

which is an analog to the material constancy condition (14.38).

14.5 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 10.4. 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 14.6, where we transform space and time derivative operators between an inertial frame and a rotating frame.

14.5.1 Galilean transformation

A Galilean transformation is given by the linear transformation

$$\bar{t} = t \quad (14.42)$$

$$\bar{\mathbf{x}} = \mathbf{x} + \mathbf{U}t \quad (14.43)$$

$$\bar{\mathbf{v}} = \mathbf{v} + \mathbf{U}. \quad (14.44)$$

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

$$t = \bar{t} \quad (14.45)$$

$$\mathbf{x} = \bar{\mathbf{x}} - \mathbf{U}\bar{t} \quad (14.46)$$

$$\mathbf{v} = \bar{\mathbf{v}} - \mathbf{U}. \quad (14.47)$$

14.5.2 Transformation of the material time derivative

We find it useful practice to make use of the transformation matrix formalism used in Section 7.2.1 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

$$\frac{\partial}{\partial x^{\bar{\alpha}}} = \frac{\partial x^\alpha}{\partial \bar{x}^{\bar{\alpha}}} \frac{\partial}{\partial x^\alpha}, \quad (14.48)$$

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

$$\frac{\partial x^{\bar{\alpha}}}{\partial x^\alpha} = \begin{bmatrix} \partial \bar{x}^0 / \partial x^0 & \partial \bar{x}^0 / \partial x^1 \\ \partial \bar{x}^1 / \partial x^0 & \partial \bar{x}^1 / \partial x^1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ U & 1 \end{bmatrix}, \quad (14.49)$$

and the inverse is

$$\frac{\partial x^\alpha}{\partial \bar{x}^{\bar{\alpha}}} = \begin{bmatrix} 1 & 0 \\ -U & 1 \end{bmatrix}. \quad (14.50)$$

Consequently, the Eulerian time derivative as measured in the moving frame is given by

$$\frac{\partial}{\partial \bar{x}^0} = \frac{\partial x^0}{\partial \bar{x}^0} \frac{\partial}{\partial x^0} + \frac{\partial x^1}{\partial \bar{x}^0} \frac{\partial}{\partial x^1} \quad (14.51)$$

$$= \frac{\partial}{\partial \bar{x}^0} - U \frac{\partial}{\partial x^1} \quad (14.52)$$

$$= \frac{\partial}{\partial t} - U \frac{\partial}{\partial x}. \quad (14.53)$$

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

$$\frac{\partial}{\partial \bar{x}^1} = \frac{\partial x^0}{\partial \bar{x}^1} \frac{\partial}{\partial x^0} + \frac{\partial x^1}{\partial \bar{x}^1} \frac{\partial}{\partial x^1} \quad (14.54)$$

$$= \frac{\partial}{\partial \bar{x}^1}, \quad (14.55)$$

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

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (14.56a)$$

$$= \frac{\partial}{\partial \bar{t}} + \mathbf{U} \cdot \bar{\nabla} + (\bar{\mathbf{v}} - \mathbf{U}) \cdot \bar{\nabla} \quad (14.56b)$$

$$= \frac{\partial}{\partial \bar{t}} + \bar{\mathbf{v}} \cdot \bar{\nabla} \quad (14.56c)$$

$$= \frac{D}{D\bar{t}}, \quad (14.56d)$$

where we used the shorthand

$$\bar{\mathbf{v}} \cdot \bar{\nabla} = \bar{u} \frac{\partial}{\partial \bar{x}^1} + \bar{v} \frac{\partial}{\partial \bar{x}^2} + \bar{w} \frac{\partial}{\partial \bar{x}^3}. \quad (14.57)$$

Even though the material time derivative operator is form invariant under an arbitrary Galilean transformation, its individual pieces are separately modified.

14.6 Invariance of the material time derivative

In our discussion of Galilean invariance in Section 14.5, 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.

14.6.1 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

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (14.58)$$

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 form of the material time derivative remains form invariant when changing reference frames.

14.6.2 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 10.4. 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

$$\hat{\mathbf{x}} = \hat{\mathbf{x}} \cos \theta - \hat{\mathbf{y}} \sin \theta \quad (14.59a)$$

$$\hat{\mathbf{y}} = \hat{\mathbf{x}} \sin \theta + \hat{\mathbf{y}} \cos \theta \quad (14.59b)$$

$$\hat{\mathbf{z}} = \hat{\mathbf{z}}, \quad (14.59c)$$

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

$$\theta = \Omega t. \quad (14.60)$$

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)

$$\bar{t} = t \quad (14.61a)$$

$$\bar{x} = x \cos \theta - y \sin \theta \quad (14.61b)$$

$$\bar{y} = x \sin \theta + y \cos \theta \quad (14.61c)$$

$$\bar{z} = z. \quad (14.61d)$$

The inverse transformation can be easily found

$$t = \bar{t} \quad (14.62a)$$

$$x = \bar{x} \cos \theta + \bar{y} \sin \theta \quad (14.62b)$$

$$y = -\bar{x} \sin \theta + \bar{y} \cos \theta \quad (14.62c)$$

$$z = \bar{z}. \quad (14.62d)$$

We are now prepared to make use of the transformation formalism used for the Galilean transformation in Section 14.5, as well as in Section 7.3 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

$$\frac{\partial x^{\bar{\alpha}}}{\partial x^{\alpha}} = \begin{bmatrix} \partial x^{\bar{0}}/\partial x^0 & \partial x^{\bar{0}}/\partial x^1 & \partial x^{\bar{0}}/\partial x^2 & \partial x^{\bar{0}}/\partial x^3 \\ \partial x^{\bar{1}}/\partial x^0 & \partial x^{\bar{1}}/\partial x^1 & \partial x^{\bar{1}}/\partial x^2 & \partial x^{\bar{1}}/\partial x^3 \\ \partial x^{\bar{2}}/\partial x^0 & \partial x^{\bar{2}}/\partial x^1 & \partial x^{\bar{2}}/\partial x^2 & \partial x^{\bar{2}}/\partial x^3 \\ \partial x^{\bar{3}}/\partial x^0 & \partial x^{\bar{3}}/\partial x^1 & \partial x^{\bar{3}}/\partial x^2 & \partial x^{\bar{3}}/\partial x^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\Omega \bar{y} & \cos \theta & -\sin \theta & 0 \\ \Omega \bar{x} & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (14.63)$$

Similarly, the inverse transformation is given by

$$\frac{\partial x^{\alpha}}{\partial x^{\bar{\alpha}}} = \begin{bmatrix} \partial x^0/\partial x^{\bar{0}} & \partial x^0/\partial x^{\bar{1}} & \partial x^0/\partial x^{\bar{2}} & \partial x^0/\partial x^{\bar{3}} \\ \partial x^1/\partial x^{\bar{0}} & \partial x^1/\partial x^{\bar{1}} & \partial x^1/\partial x^{\bar{2}} & \partial x^1/\partial x^{\bar{3}} \\ \partial x^2/\partial x^{\bar{0}} & \partial x^2/\partial x^{\bar{1}} & \partial x^2/\partial x^{\bar{2}} & \partial x^2/\partial x^{\bar{3}} \\ \partial x^3/\partial x^{\bar{0}} & \partial x^3/\partial x^{\bar{1}} & \partial x^3/\partial x^{\bar{2}} & \partial x^3/\partial x^{\bar{3}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \Omega y & \cos \theta & \sin \theta & 0 \\ -\Omega x & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (14.64)$$

The derivative operators transform according to

$$\frac{\partial}{\partial x^{\alpha}} = \frac{\partial x^{\bar{\alpha}}}{\partial x^{\alpha}} \frac{\partial}{\partial x^{\bar{\alpha}}}, \quad (14.65)$$

in which case

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \bar{t}} + (\boldsymbol{\Omega} \wedge \bar{\boldsymbol{x}}) \cdot \bar{\nabla} \quad (14.66a)$$

$$\frac{\partial}{\partial x} = \cos \theta \frac{\partial}{\partial \bar{x}} + \sin \theta \frac{\partial}{\partial \bar{y}} \quad (14.66b)$$

$$\frac{\partial}{\partial y} = -\sin \theta \frac{\partial}{\partial \bar{x}} + \cos \theta \frac{\partial}{\partial \bar{y}} \quad (14.66c)$$

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial \bar{z}}. \quad (14.66d)$$

The velocity vector components transform according to

$$v^{\alpha} = \frac{\partial x^{\alpha}}{\partial x^{\bar{\alpha}}} v^{\bar{\alpha}}, \quad (14.67)$$

so that

$$v^0 = v^{\bar{0}} \quad (14.68a)$$

$$u = \Omega y + \bar{u} \cos \theta + \bar{v} \sin \theta \quad (14.68b)$$

$$v = -\Omega x - \bar{u} \sin \theta + \bar{v} \cos \theta \quad (14.68c)$$

$$w = \bar{w}. \quad (14.68d)$$

Bringing these result together leads to the transformation of the horizontal advection operator

$$u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} = (\bar{\mathbf{u}} - \boldsymbol{\Omega} \wedge \bar{\mathbf{x}}) \cdot \bar{\nabla}. \quad (14.69)$$

Combining this result with the transformed Eulerian time derivative leads to the material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (14.70a)$$

$$= \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \quad (14.70b)$$

$$= \frac{\partial}{\partial t} + (\boldsymbol{\Omega} \wedge \bar{\mathbf{x}}) \cdot \bar{\nabla} + (\bar{\mathbf{u}} - \boldsymbol{\Omega} \wedge \bar{\mathbf{x}}) \cdot \bar{\nabla} + \bar{w} \frac{\partial}{\partial z} \quad (14.70c)$$

$$= \frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \bar{\nabla}. \quad (14.70d)$$

As advertised, the operator is form invariant under time dependent transformations to a non-inertial reference frame.

14.6.3 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. In so doing, it can become a tool of use for exploration where the answer is not *a priori* known.

14.7 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.

14.7.1 Material pathlines from fluid particle trajectories

As introduced in Section 14.3.1, a fluid particle traces out a *trajectory* as it moves through space (Figure 14.2). 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

$$\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (14.71a)$$

$$\mathbf{X}(\mathbf{a}, t = 0) = \mathbf{a}, \quad (14.71b)$$

where

$$\mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] = \mathbf{v}^L(\mathbf{a}, t) \quad (14.72)$$

is Lagrangian velocity of the fluid particle (see equation (14.31)), 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 22). 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.

14.7.2 Fluid streamlines and streamtubes

Streamlines are curves whose tangent is instantaneously parallel to the fluid velocity field. Let

$$d\mathbf{x} = \hat{\mathbf{x}} dx + \hat{\mathbf{y}} dy + \hat{\mathbf{z}} dz \quad (14.73)$$

be an infinitesimal increment along a streamline written using Cartesian coordinates. The family of streamlines at a given time t satisfy the tangent constraint

$$\mathbf{v} \wedge d\mathbf{x} = 0, \quad (14.74)$$

which is equivalent to

$$\frac{dx}{u(\mathbf{x}, t)} = \frac{dy}{v(\mathbf{x}, t)} = \frac{dz}{w(\mathbf{x}, t)}. \quad (14.75)$$

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

$$\frac{\partial \mathbf{X}(s; \mathbf{a}, t)}{\partial s} = \mathbf{v}[\mathbf{X}(s; \mathbf{a}, t), t] \quad (14.76a)$$

$$\mathbf{X}(s = 0; \mathbf{a}, t) = \mathbf{a}. \quad (14.76b)$$

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 14.5). Hence, at each time instance, streamtube sides are parallel to the velocity vector. Furthermore, when the flow is steady (time independent velocity field), streamlines are material pathlines. A streamtube is therefore a material tube for steady flow, in which case no fluid particles cross the streamtube boundary.

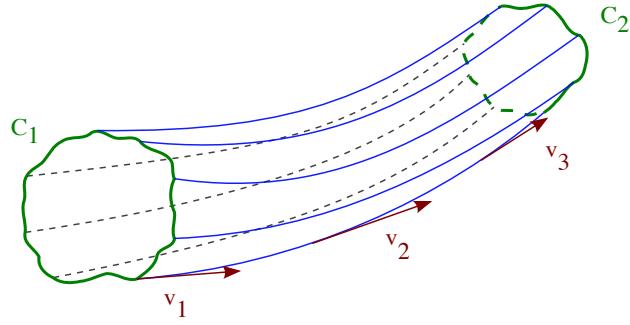


Figure 14.5: This image shows an example streamtube. The side boundaries of a streamtube consist of streamlines, so that at each instance the tangent at each point of a streamtube is parallel to the velocity at that point (see equation (14.74)). A cross-section of the streamtube forms a curve C . Streamlines are pathlines for steady flow, whereas they differ for unsteady flows. That is, for unsteady flows, particle trajectories generally cross through the streamtube boundary.

14.7.3 Emphasizing that streamlines are not pathlines in unsteady flow

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 (14.76a) is not equal to the time, t , used to compute fluid particle trajectories in equation (14.71a). 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

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \mathbf{v}^{\text{line}} \cdot \hat{\mathbf{n}} \quad \text{material lines,} \quad (14.77)$$

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.

14.7.4 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 14.6). 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 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

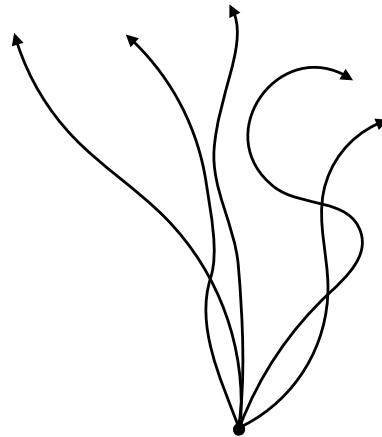


Figure 14.6: This image shows a suite of trajectories that emanate from a single point. Common realizations include the paths of fluid particles that leave from a chimney, or the smoke from a point source. A streakline is defined as the accumulation of positions at time t of particles that passed through the common point at some earlier time $s < t$.

coordinate of the particle would be given by $\mathbf{a}(\mathbf{y}, s)$ (see equation (14.8) relating the 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

$$\mathbf{x} = \mathbf{X}[\mathbf{a}(\mathbf{y}, s), t] \quad 0 \leq s \leq t. \quad (14.78)$$

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 $t = 0$

$$\frac{\partial \mathbf{X}(s; \mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(s; \mathbf{a}, t), t] \quad (14.79a)$$

$$\mathbf{X}(t = s; \mathbf{a}, t) = \mathbf{a}. \quad (14.79b)$$

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.

14.7.5 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

$$u = \frac{x}{\tau + t} \quad (14.80a)$$

$$v = \frac{y}{\tau} \quad (14.80b)$$

$$w = 0, \quad (14.80c)$$

where $\tau > 0$ is a constant with the dimensions of time.

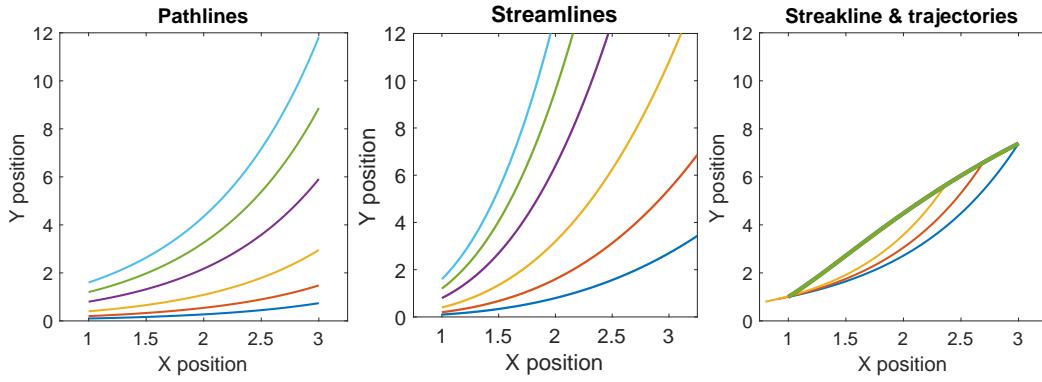


Figure 14.7: Left panel: sample pathlines $X(t) = X_0(1+t/\tau)$ and $Y(t) = Y_0 e^{t/\tau}$ (see equations (14.82a) and (14.82b)) during times $t \in [0, 2]$. The trajectories drawn here all start at $X_0 = 1$ and set the parameter $\tau = 1$. Note that those pathlines with $X_0 = 0$ remain on the y-axis, 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 (14.85a) and (14.85b)). 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 (14.88a) and (14.88b). 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

$$\frac{dX(t)}{dt} = \frac{X(t)}{\tau + t} \quad (14.81a)$$

$$\frac{dY(t)}{dt} = \frac{Y(t)}{\tau} \quad (14.81b)$$

$$\frac{dZ(t)}{dt} = 0, \quad (14.81c)$$

which are found to be

$$X(t) = X_0(1 + t/\tau) \quad (14.82a)$$

$$Y(t) = Y_0 e^{t/\tau} \quad (14.82b)$$

$$Z(t) = Z_0, \quad (14.82c)$$

where $\mathbf{X}(t = 0) = \mathbf{X}_0$. Sample trajectories are shown in Figure 14.7 over time $t \in [0, 2]$. We can eliminate time to yield a curve in the horizontal (x, y) plane

$$y = Y_0 e^{(x-X_0)/X_0}. \quad (14.83)$$

Streamlines

Streamlines are determined by solving the differential equations

$$\frac{dX(s; t)}{ds} = \frac{X(s; t)}{\tau + t} \quad (14.84a)$$

$$\frac{dY(s; t)}{ds} = \frac{Y(s; t)}{\tau} \quad (14.84b)$$

$$\frac{dZ(s; t)}{ds} = 0, \quad (14.84c)$$

where time, t , is a fixed parameter whereas the pseudo-time, s , is varied. Integration renders the streamlines

$$X(s; t) = X_0 e^{s/(\tau+t)} \quad (14.85a)$$

$$Y(s; t) = Y_0 e^{s/\tau} \quad (14.85b)$$

$$Z(s; t) = Z_0. \quad (14.85c)$$

Sample streamlines are shown in Figure 14.7. Note that we can eliminate the pseudo-time s to render a curve in the horizontal (x, y) plane

$$y = Y_0 \left(\frac{x}{X_0} \right)^{(\tau+t)/\tau} \quad (14.86a)$$

$$z = Z_0. \quad (14.86b)$$

Streaklines

For streaklines, invert the trajectory expressions (14.82a)-(14.82b) to find the material coordinates $\mathbf{a}(y, s)$ in the form

$$a_1 = \frac{y_1}{1 + s/\tau} \quad (14.87a)$$

$$a_2 = y_2 e^{-s/\tau} \quad (14.87b)$$

$$a_3 = y_3. \quad (14.87c)$$

We next evaluate the trajectory expressions (14.82a)-(14.82b) with \mathbf{a} as the initial positions to find the streaklines

$$X(s; \mathbf{a}, t) = \frac{y_1 (1 + t/\tau)}{1 + s/\tau} \quad (14.88a)$$

$$Y(s; \mathbf{a}, t) = y_2 e^{(t-s)/\tau} \quad (14.88b)$$

$$Z(s; \mathbf{a}, t) = y_3. \quad (14.88c)$$

Figure 14.7 illustrates the streakline for a particular point $(X, Y) = (1, 1)$.

14.7.6 Further reading

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.

14.8 Stokes drift

The distinction between Lagrangian and Eulerian appears quite clearly when considering the movement of fluid particles in an inhomogeneous wave field. Due to inhomogeneities in the waves, particles oscillate between regions where the drift in one direction does not match that in the other direction. There is generally a net drift of particles known as “Stokes drift”. Formulating the maths of Stokes drift offers a means to explore the differences between averages formed at a point in space (Eulerian mean) versus averages formed 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 23.

14.8.1 Formulation of Stokes drift

Consider a three-dimensional particle trajectory written in Cartesian coordinates,

$$\mathbf{X}(\mathbf{a}, t) = X(\mathbf{a}, t) \hat{\mathbf{x}} + Y(\mathbf{a}, t) \hat{\mathbf{y}} + Z(\mathbf{a}, t) \hat{\mathbf{z}}. \quad (14.89)$$

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 14.7.1, the particle trajectory is determined by time integrating the particle velocity (also known as the “flow map”)

$$\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (14.90)$$

so that

$$\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, 0) + \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t'), t'] dt'. \quad (14.91)$$

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

$$v^n[\mathbf{X}(\mathbf{a}, t), t] \approx v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot [\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)] \quad (14.92a)$$

$$= v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \frac{d\mathbf{X}(\mathbf{a}, t')}{dt'} dt' \quad (14.92b)$$

$$= v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t'), t'] dt', \quad (14.92c)$$

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 (14.92c).

- 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

$$\epsilon = \frac{|\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)|}{\Lambda}. \quad (14.93)$$

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 (14.92c) 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

$$v^n[\mathbf{X}(\mathbf{a}, t), t] \approx v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt', \quad (14.94)$$

with rearrangement leading to

$$v^n[\mathbf{X}(\mathbf{a}, t), t] - v^n[\mathbf{X}(\mathbf{a}, 0), t] \approx \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt'. \quad (14.95)$$

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 (14.95) 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 (14.95) when time averaged over a wave period, which we write as

$$v_{(S)}^n[\mathbf{X}(\mathbf{a}, 0), t] = \overline{v^n[\mathbf{X}(\mathbf{a}, t), t] - v^n[\mathbf{X}(\mathbf{a}, 0), t]} \quad (14.96a)$$

$$\approx \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt'. \quad (14.96b)$$

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

$$v_{(S)}^n(\mathbf{x}, t) \approx \nabla v^n(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt'. \quad (14.97)$$

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. We again see the presence of Stokes drift when considering the generalized Lagrangian mean in Section 23.2.4.



Figure 14.8: 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.

14.8.2 Stokes drift in the upper ocean

The canonical example of Stokes drift occurs in the near surface ocean, where surface gravity waves 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

$$\frac{dX}{dt} = U \sin(\omega t) \quad (14.98a)$$

$$\frac{dZ}{dt} = U \cos(\omega t), \quad (14.98b)$$

where $U > 0$ is the speed of the circular particle motion. We have moved into a frame where the waves are stationary, which makes the maths a bit simpler. Figure 14.9 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

$$X(t) - X_o = -\frac{U_o}{\omega} [\cos(\omega t) - 1] \quad (14.99a)$$

$$Z(t) - Z_o = \frac{U_o}{\omega} \sin(\omega t), \quad (14.99b)$$

where the initial position at time $t = 0$ is

$$\mathbf{X}(t = t_o) = \mathbf{X}_o, \quad (14.100)$$

and the center of the circle is

$$\mathbf{X}_{\text{center}} = \left[X_o + \frac{U_o}{\omega} \right] \hat{\mathbf{x}} + Z_o \hat{\mathbf{z}}. \quad (14.101)$$

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. In turn, we expect there to be a drift in the zonal direction introduced by the vertical 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

$$U \approx U_o + R(Z - Z_o) \quad (14.102)$$

where R has units of inverse time and is given by

$$R = \left[\frac{dU}{dZ} \right]_{Z=Z_o}. \quad (14.103)$$

The Taylor series is valid so long as the vertical trajectories maintain the inequality

$$|R| |Z - Z_o| \ll U_o, \quad (14.104)$$

which says that the vertical shear is small

$$|R| \ll \frac{U_o}{|Z - Z_o|}. \quad (14.105)$$

We use the Taylor series expansion (14.102) to solve for the vertical trajectory as determined by

$$\frac{d(Z - Z_o)}{dt} = [U_o + R(Z - Z_o)] \cos(\omega t). \quad (14.106)$$

Rearrangement leads to

$$\int_{Z_o}^Z \frac{d(Z - Z_o)}{U_o + R(Z - Z_o)} = \int_0^t \cos(\omega t) dt. \quad (14.107)$$

The left hand side integral can be computed by changing variables

$$\Sigma = U_o + R(Z - Z_o) \quad (14.108a)$$

$$d\Sigma = R d(Z - Z_o), \quad (14.108b)$$

so that equation (14.107) becomes

$$\int_{U_o}^{\Sigma} \frac{d\Sigma}{\Sigma} = R \int_0^t \cos(\omega t) dt. \quad (14.109)$$

Performing the integrals and evaluating the end points renders

$$\ln \left[1 + \frac{R}{U_o} (Z - Z_o) \right] = \frac{R \sin(\omega t)}{\omega}, \quad (14.110)$$

which yields the exponential solution

$$1 + \frac{R}{U_o} (Z - Z_o) = e^{(R/\omega) \sin(\omega t)} \Rightarrow Z - Z_o = \frac{U_o}{R} \left(-1 + e^{(R/\omega) \sin(\omega t)} \right). \quad (14.111)$$

The vertical particle position is seen to oscillate around its initial position Z_o .

We next consider the zonal particle position, in which case

$$\frac{d(X - X_o)}{dt} = U_o \left[1 + \frac{R}{U_o} (Z - Z_o) \right] \sin(\omega t) \quad (14.112a)$$

$$= U_o e^{(R/\omega) \sin(\omega t)} \sin(\omega t) \quad (14.112b)$$

where we used equation (14.111) for the vertical trajectory. To make progress, we expand the exponential assuming the ratio of inverse time scales, R/ω , is small

$$|R/\omega| \ll 1. \quad (14.113)$$

In this limit, the vertical trajectory retains its unperturbed form (14.99b), and the zonal trajectory satisfies

$$\frac{d(X - X_o)}{dt} \approx U_o \sin(\omega t) \left[1 + \frac{R}{\omega} \sin(\omega t) \right], \quad (14.114)$$

where we dropped terms of order $(R/\omega)^2$. We can understand the scaling in equation (14.113) by noting that the period for the circular motion is given by

$$\tau_{\text{circle}} = \frac{2\pi}{\omega}. \quad (14.115)$$

The inverse time R introduces a time scale for the drift, defined according to

$$\tau_{\text{drift}} = \frac{2\pi}{|R|}. \quad (14.116)$$

A small ratio $|R/\omega|$ thus implies

$$|R/\omega| = \tau_{\text{circle}}/\tau_{\text{drift}} \ll 1. \quad (14.117)$$

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 (14.114) yields

$$\frac{d(X - X_o)}{dt} = U_o \sin(\omega t) \left[1 + \frac{R}{\omega} \sin(\omega t) \right] \quad (14.118a)$$

$$= U_o \sin(\omega t) + \frac{U_o R}{2\omega} [1 - \cos(2\omega t)], \quad (14.118b)$$

which integrates to

$$X - X_o = \left(\frac{U_o}{\omega} \right) \left[1 - \cos(\omega t) - \frac{R \sin(2\omega t)}{4\omega} + \frac{R t}{2} \right] \quad (14.119a)$$

$$= \underbrace{\left(\frac{U_o}{\omega} \right) [1 - \cos(\omega t)]}_{\text{homogeneous}} + \underbrace{\frac{U_o R t}{2\omega}}_{\text{Stokes drift}} - \underbrace{\frac{U_o R \sin(2\omega t)}{4\omega^2}}_{\text{higher harmonic}} + \mathcal{O}(R/\omega)^2. \quad (14.119b)$$

The leading order term is the homogeneous motion given by equation (14.99a). The next term is the Stokes drift, followed by a higher order harmonic and then further terms on the order of

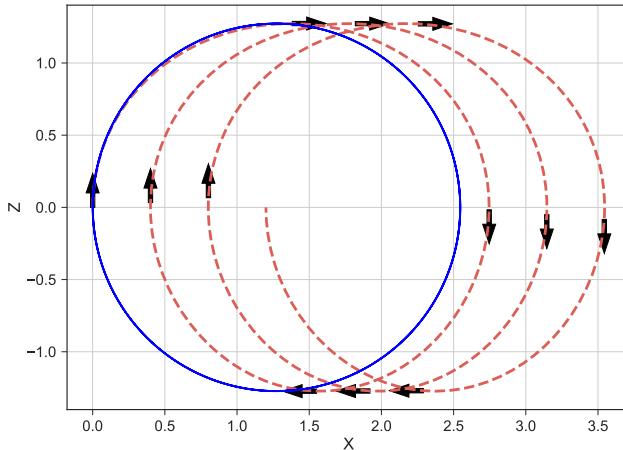


Figure 14.9: Example trajectories of fluid particles undergoing Stokes drift for short surface ocean waves. Particle motion is clockwise in the $x - z$ plane. For homogeneous waves, there is zero Stokes drift with circular trajectories given by equations (14.99a) and (14.99b), as depicted here by the 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 (14.111) for the vertical component and equation (14.119b) 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.

$(R/\omega)^2$. There is no vertical Stokes drift to this order in (R/ω) , so that the Stokes drift velocity is given by

$$\left[\frac{\mathbf{X} - \mathbf{X}_o}{t} \right]^{\text{drift}} = \frac{RU_o}{2\omega} \hat{\mathbf{x}}. \quad (14.120)$$

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 14.9 for an illustration based on a particular choice for the dimensional parameters.

14.9 Exercises

EXERCISE 14.1: 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

$$\frac{D(\partial_m \Phi)}{Dt} = -\partial_m \mathbf{v} \cdot \nabla \Phi. \quad (14.121)$$

For example, if $D\Phi/Dt = 0$, then the zonal partial derivative $\partial_x \Phi$ has a material time derivative given by

$$\frac{D(\partial \Phi / \partial x)}{Dt} = -\frac{\partial \mathbf{v}}{\partial x} \cdot \nabla \Phi. \quad (14.122)$$

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?

EXERCISE 14.2: STOKES DRIFT FOR ONE-DIMENSIONAL MONOCHROMATIC WAVE

Consider a one-dimensional monochromatic longitudinal wave with velocity

$$u = u_o \sin(kx - \omega t), \quad (14.123)$$

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

$$\epsilon = \frac{u_o}{c} = \frac{u_o k}{\omega} = \frac{u_o T}{\lambda} \quad (14.124)$$

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 (14.97).

15

Mass conservation

We assume that mass is neither created nor destroyed anywhere within the fluid domain. This assumption holds independent of the forces acting on the fluid, and so it forms a topic within fluid kinematics. 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 the boundaries. We examine a variety of fluid systems: infinitesimal and finite; moving (Lagrangian) and static (Eulerian); homogeneous and multi-component.

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 5 and 6. We presume an understanding of the Eulerian and Lagrangian kinematic descriptions detailed in Chapter 14. Nearly all material in this chapter is used throughout the remainder of the book.

15.1	Material fluid parcels	186
15.1.1	Lagrangian expression for mass conservation	186
15.1.2	Alternative derivation based on the Jacobian	187
15.1.3	Summary of parcel kinematic equations	187
15.2	Eulerian fluid regions	187
15.2.1	Differential expression	187
15.2.2	Finite volume expression	188
15.2.3	Arbitrary Eulerian region	188
15.3	Material fluid regions	189
15.3.1	Evolution of volume	189
15.3.2	Mass conservation	190
15.3.3	Mass conservation using Lagrangian methods	190
15.3.4	Reynolds Transport Theorem	192
15.4	Fluid elements and the tracer equation	192
15.4.1	Material regions and material parcels for each constituent	192
15.4.2	Total mass conservation	193
15.4.3	Revisiting the fluid element	194
15.4.4	The tracer equation	194
15.4.5	Compatibility between mass continuity and the tracer equation	195

15.4.6	Comments	196
15.5	Passive tracers	196
15.6	Budgets for arbitrary fluid regions	196
15.6.1	Extensive and intensive properties	196
15.6.2	General form of the finite domain integral	197
15.6.3	Static domain	198
15.6.4	Leibniz-Reynolds Transport Theorem	198
15.6.5	Revisiting Reynolds Transport Theorem	199
15.7	Kinematic boundary conditions	201
15.7.1	Static material surface	202
15.7.2	Moving material surface	203
15.7.3	Dynamic and permeable surface	204
15.8	Exercises	207

15.1 Material fluid parcels

We here derive the differential expressions for mass conservation of a material parcel within a Lagrangian reference frame. The differential expressions for volume and density arise as a corollary.

15.1.1 Lagrangian expression for mass conservation

The mass of an infinitesimal fluid parcel is written

$$\delta M = \rho \delta V, \quad (15.1)$$

where δV is the parcel volume and

$$\rho = \frac{\delta M}{\delta V} \quad (15.2)$$

is the mass density. By definition, the parcel has a constant mass, so that its material time derivative vanishes

$$\frac{D(\delta M)}{Dt} = 0. \quad (15.3)$$

Equation (15.3) 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

$$\frac{D(\delta M)}{Dt} = \frac{D(\rho \delta V)}{Dt} \quad (15.4a)$$

$$= \delta M \left[\frac{1}{\rho} \frac{D\rho}{Dt} + \frac{1}{\delta V} \frac{D(\delta V)}{Dt} \right]. \quad (15.4b)$$

Making use of equation (15.21) derived in Section 15.3.1 for the material evolution of the parcel volume then leads to

$$\frac{1}{\delta M} \frac{D(\delta M)}{Dt} = \frac{1}{\rho} \frac{D\rho}{Dt} + \nabla \cdot \mathbf{v}. \quad (15.5)$$

Setting $D(\delta M)/Dt = 0$ then leads to the mass conservation equation, also known as the continuity equation

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}. \quad (15.6)$$

The parcel volume contracts in regions where the velocity converges (we prove that property in Section 15.3.1). As seen by the mass continuity equation (15.6), such regions are also where the parcel density increases. The opposite occurs for regions where the velocity diverges.

15.1.2 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 (17.51c)), which then leads to the material evolution for the fluid parcel mass

$$\frac{D}{Dt} [\rho \delta V(\mathbf{x})] = \frac{D}{Dt} \left[\rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \delta V(\mathbf{a}) \right] \quad (15.7a)$$

$$= \left[\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right] \left[\frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right] \delta V(\mathbf{a}) \quad (15.7b)$$

$$= \left[\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right] \delta V(\mathbf{x}). \quad (15.7c)$$

We recover the mass conservation equation (15.6) when noting that the mass of a material parcel is constant.

15.1.3 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

$$\frac{D(\delta M)}{Dt} = 0 \quad \text{parcel mass} \quad (15.8a)$$

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v} \quad \text{parcel volume} \quad (15.8b)$$

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v} \quad \text{parcel density.} \quad (15.8c)$$

Note that the parcel volume equation is derived in Section 15.3.1 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 ($-\nabla \cdot \mathbf{v} < 0$).

15.2 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.

15.2.1 Differential expression

The Eulerian form of mass continuity results from introducing the Eulerian expression for the material time derivative operator (equation (14.28)) into the mass continuity equation (15.6). The resulting *flux-form* Eulerian mass continuity equation is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (15.9)$$

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

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}). \quad (15.10)$$

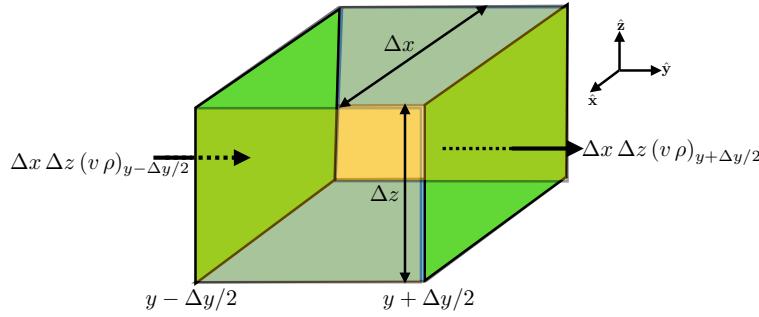


Figure 15.1: A finite sized cube (also known as a cell) of fixed dimensions and position (an Eulerian cube) used to formulate the Eulerian form of mass conservation. We highlight two cell faces with area $\Delta x \Delta z$ and with meridional mass transport crossing the faces given by $\Delta x \Delta z (v \rho)_{y-\Delta y/2}$ 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.

15.2.2 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 15.1. The mass contained within the cube is given by

$$m = \rho \Delta V = \rho \Delta x \Delta y \Delta z, \quad (15.11)$$

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 15.1, the accumulation of mass within the cube through this transport is determined by the difference in mass transport crossing the two adjacent cell faces

$$\text{mass change from meridional transport} = (\Delta x \Delta z) [(v \rho)_{y-\Delta y/2} - (v \rho)_{y+\Delta y/2}]. \quad (15.12)$$

Expanding the difference into a Taylor series leads to

$$\text{mass change from meridional transport} \approx -(\Delta x \Delta y \Delta z) \frac{\partial(v \rho)}{\partial y}. \quad (15.13)$$

The same analysis for the zonal and vertical directions leads to the mass budget for the cube

$$\frac{\partial m}{\partial t} = -\Delta V \left[\frac{\partial(u \rho)}{\partial x} + \frac{\partial(v \rho)}{\partial y} + \frac{\partial(w \rho)}{\partial z} \right]. \quad (15.14)$$

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 (15.9), $\partial \rho / \partial t = -\nabla \cdot (\rho \mathbf{v})$.

15.2.3 Arbitrary Eulerian region

The previous discussion for a cube can be generalized by making use of the Divergence theorem (Section 3.7). For that purpose, consider an arbitrary static and simply closed region within the

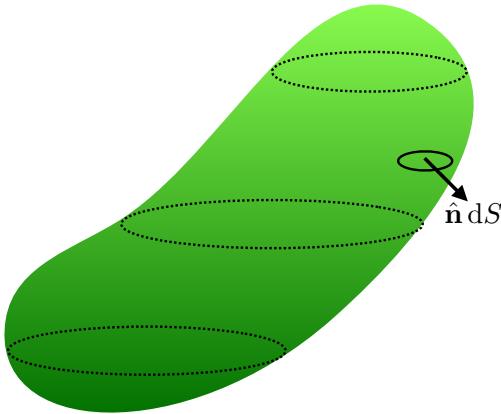


Figure 15.2: An arbitrarily shaped simply closed region, \mathcal{R} , within the fluid. If the region is fixed in space, then it represents a general Eulerian region for considering mass budgets. A surface area element, dS , is oriented according to the outward normal, \hat{n} .

fluid such as in Figure 15.2. Integrating the continuity equation (15.9) over that region leads to

$$\int_{\mathcal{R}} \frac{\partial \rho}{\partial t} dV = - \int_{\mathcal{R}} \nabla \cdot (\rho \mathbf{v}) dV. \quad (15.15)$$

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

$$\frac{\partial}{\partial t} \int_{\mathcal{R}} \rho dV = - \int_{\partial \mathcal{R}} \rho \mathbf{v} \cdot \hat{n} dS, \quad (15.16)$$

where \hat{n} is the outward normal vector along the boundary of the region, and dS 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).

15.3 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 15.2. 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 15.3.4.

15.3.1 Evolution of volume

Consider a finite material region, \mathcal{R} , whose volume is given by the integral

$$V = \int_{\mathcal{R}} dV, \quad (15.17)$$

with dV the volume element. The volume changes according to the motion of the fluid particles fixed to the boundary of the material region. The material region expands when the flow moves outward, and contracts when the flow moves inward. These statements take on the following mathematical expression

$$\frac{D}{Dt} \int_{\mathcal{R}} dV = \int_{\partial\mathcal{R}} \mathbf{v} \cdot \hat{\mathbf{n}} dS, \quad (15.18)$$

where $\hat{\mathbf{n}}$ is the outward normal on the region boundary, dS is the area element on the boundary, and

$$\mathbf{v} \cdot \hat{\mathbf{n}} dS = \text{volume transport (volume per time) at the boundary } \partial\mathcal{R}. \quad (15.19)$$

Use of Gauss's Law then leads to the equivalent expression

$$\frac{D}{Dt} \int_{\mathcal{R}} dV = \int_{\mathcal{R}} \nabla \cdot \mathbf{v} dV. \quad (15.20)$$

We now take the limit as the material region becomes a material parcel, in which case we recover the differential expression

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v}. \quad (15.21)$$

This equation is also derived in Section 17.4.1 using different methods.

15.3.2 Mass conservation

The mass of material parcels over a finite material region is given by

$$M = \int_{\mathcal{R}} \rho dV. \quad (15.22)$$

As a material fluid region, it maintains a constant mass as it moves through the fluid

$$\frac{D}{Dt} \int_{\mathcal{R}} \rho dV = 0. \quad (15.23)$$

In the limit that the material region becomes infinitesimally small, the region mass conservation statement (15.23) becomes the parcel mass conservation statement (15.3)

$$\frac{D(\delta M)}{Dt} = \frac{D(\rho \delta V)}{Dt} = 0. \quad (15.24)$$

15.3.3 Mass conservation using Lagrangian methods

Rather than take the limit as the finite material region \mathcal{R} becomes infinitesimal, we develop some formalism that specifies how to move the material time derivative across the integral over the finite sized material region in equation (15.23). 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)

$$\int_{\mathcal{R}} \rho dV = \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} \rho dx dy dz \equiv \int_{\mathcal{R}(\mathbf{x})} \rho dV(\mathbf{x}), \quad (15.25)$$

where the second equality defines the notation where the region bounds and volume element are specified using the \mathbf{x} -space representation. 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

$$\int_{x_1(t)}^{x_2(t)} \rho \, dx = \int_{x_1(t)}^{x_2(t)} \rho \frac{\partial X}{\partial a} \, da = \int_{a_1}^{a_2} \rho \frac{\partial X}{\partial a} \, da. \quad (15.26)$$

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

$$\int_{\mathcal{R}(\mathbf{x})} \rho \, dV(\mathbf{x}) = \int_{\mathcal{R}(\mathbf{a})} \rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \, dV(\mathbf{a}). \quad (15.27)$$

We now make use of the equality (15.27) 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

$$\frac{D}{Dt} \left[\int_{\mathcal{R}(\mathbf{x})} \rho \, dV(\mathbf{x}) \right] = \frac{D}{Dt} \left[\int_{\mathcal{R}(\mathbf{a})} \rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \, dV(\mathbf{a}) \right] \quad (15.28a)$$

$$= \int_{\mathcal{R}(\mathbf{a})} \frac{D}{Dt} \left[\rho \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right] \, dV(\mathbf{a}) \quad (15.28b)$$

$$= \int_{\mathcal{R}(\mathbf{x})} \left[\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right] \, dV(\mathbf{x}) \quad (15.28c)$$

$$= \int_{\mathcal{R}(\mathbf{x})} \left[\frac{1}{\rho} \frac{D\rho}{Dt} + \nabla \cdot \mathbf{v} \right] \rho \, dV(\mathbf{x}). \quad (15.28d)$$

When expressing the integral using \mathbf{a} -space coordinates, the integral bounds are static, thus allowing us to move the material time derivative inside of the integral sign to reach the second equality. The third equality made use of equation (15.7c) and converted back to \mathbf{x} -space. As the region \mathcal{R} has a materially constant mass, we recover the mass continuity equation (15.6) by setting the integrand in equation (15.28d) to zero. This technique of moving between Eulerian (\mathbf{x} -space) and Lagrangian (\mathbf{a} -space) representations is commonly used in fluid mechanics, and

15.3.4 Reynolds Transport Theorem

Manipulations leading to the mass conservation statement (15.28d) can be generalized by considering the material time derivative of a mass-weighted field

$$\frac{D(\psi \rho dV)}{Dt} = \frac{D\psi}{Dt} \rho dV + \psi \frac{D(\rho dV)}{Dt} \quad (15.29a)$$

$$= \rho dV \left[\frac{D\psi}{Dt} + \frac{\psi D\rho}{\rho} + \psi \nabla \cdot \mathbf{v} \right] \quad (15.29b)$$

$$= dV \left[\frac{\partial(\rho \psi)}{\partial t} + \nabla \cdot (\rho \psi \mathbf{v}) \right]. \quad (15.29c)$$

The first equality used the product rule, which holds for material time derivatives. Mass conservation means that the material derivative $D(\rho dV)/Dt = 0$. However, we choose to write mass conservation in the form of equation (15.7c), which allows us to introduce the flux-form Eulerian expression after replacing the material time derivative with its Eulerian form from equation (14.28).

Following the discussion in Section 15.3.3, we can extend the material parcel result (15.29c) to a finite size material region. The result is known as Reynolds Transport Theorem

$$\frac{D}{Dt} \int_{\mathcal{R}} \psi \rho dV = \int_{\mathcal{R}} \frac{D\psi}{Dt} \rho dV \quad (15.30a)$$

$$= \int_{\mathcal{R}} \left[\frac{\partial(\rho \psi)}{\partial t} + \nabla \cdot (\rho \psi \mathbf{v}) \right] dV \quad (15.30b)$$

$$= \int_{\mathcal{R}} \frac{\partial(\rho \psi)}{\partial t} dV + \int_{\partial\mathcal{R}} \rho \psi \mathbf{v} \cdot \hat{\mathbf{n}} dS, \quad (15.30c)$$

where the final equality follows from Gauss's Law. 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 15.7.2 when discussing the kinematics of a moving material surface.

15.4 Fluid elements and the tracer equation

As defined in Section 14.1, 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*.

15.4.1 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,

$$\rho^{(n)} = V^{-1} M^{(n)}, \quad (15.31)$$

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

$$\frac{D^{(n)} \rho^{(n)}}{Dt} = -\rho^{(n)} \nabla \cdot \mathbf{v}^{(n)} \quad \text{for each of the } n = 1, N \text{ constituents,} \quad (15.32)$$

where the constituent material time derivative is given by

$$\frac{D^{(n)}}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^{(n)} \cdot \nabla. \quad (15.33)$$

We thus have N statements of mass conservation corresponding to each constituent material fluid parcel moving according to the velocity $\mathbf{v}^{(n)}$.

15.4.2 Total mass conservation

The component mass continuity equation (15.32) takes on the Eulerian form

$$\frac{\partial \rho^{(n)}}{\partial t} + \nabla \cdot (\rho^{(n)} \mathbf{v}^{(n)}) = 0. \quad (15.34)$$

Summing over all constituents leads to the continuity equation for the total mass

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (15.35)$$

where the total mass density and barycentric velocity are given by

$$\rho = \sum_{n=1}^N \rho^{(n)} \quad \mathbf{v} = \rho^{-1} \sum_{n=1}^N \rho^{(n)} \mathbf{v}^{(n)}. \quad (15.36)$$

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

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \quad (15.37)$$

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.

¹There is no implied summation in equation (15.32).

15.4.3 Revisiting the fluid element

The mass continuity equation (15.37) 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

$$\delta M = \sum_{n=1}^N \delta M^{(n)}. \quad (15.38)$$

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.

15.4.4 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 (15.34)

$$\left[\frac{\partial}{\partial t} + \mathbf{v}^{(n)} \cdot \nabla \right] \rho^{(n)} = -\rho^{(n)} \nabla \cdot \mathbf{v}^{(n)} \quad (15.39)$$

and insert the barycentric velocity

$$\left[\frac{\partial}{\partial t} + (\mathbf{v} - \mathbf{v} + \mathbf{v}^{(n)}) \cdot \nabla \right] \rho^{(n)} = -\rho^{(n)} \nabla \cdot [\mathbf{v} - \mathbf{v} + \mathbf{v}^{(n)}]. \quad (15.40)$$

Rearrangement leads to

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \rho^{(n)} = -\rho^{(n)} \nabla \cdot \mathbf{v} - \nabla \cdot [\rho^{(n)} (\mathbf{v}^{(n)} - \mathbf{v})], \quad (15.41)$$

which can be written

$$\frac{D\rho^{(n)}}{Dt} = -\rho^{(n)} \nabla \cdot \mathbf{v} - \nabla \cdot \mathbf{J}^{(n)}, \quad (15.42)$$

where we defined the constituent mass flux

$$\mathbf{J}^{(n)} = \rho^{(n)} (\mathbf{v}^{(n)} - \mathbf{v}). \quad (15.43)$$

The dimensions of $\mathbf{J}^{(n)}$ are mass of constituent n per time per area.

The material mass conservation equation (15.42) takes on the Eulerian form

$$\frac{\partial \rho^{(n)}}{\partial t} + \nabla \cdot (\mathbf{v} \rho^{(n)}) = -\nabla \cdot \mathbf{J}^{(n)}. \quad (15.44)$$

Introducing the tracer concentration $C^{(n)}$ according to

$$C^{(n)} = \frac{\rho^{(n)}}{\rho} = \frac{\delta M^{(n)}}{\delta M} = \frac{\text{mass of constituent } n \text{ in fluid element}}{\text{mass of fluid element}}, \quad (15.45)$$

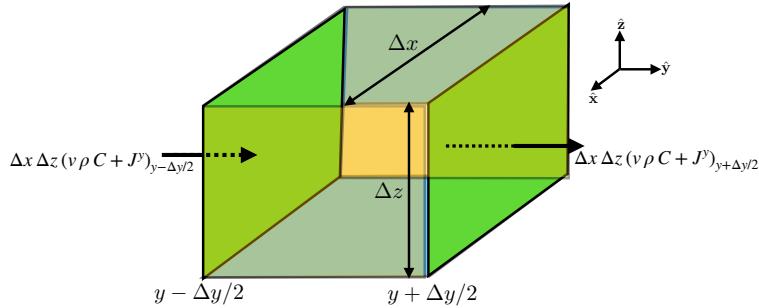


Figure 15.3: A finite sized cube as in Figure 15.1, here used to illustrate the budget of tracer mass over an Eulerian region. In addition to the advective flux of tracer moving with the barycentric velocity, \mathbf{v} , there is a diffusive flux, \mathbf{J} , that contributes to the transfer of tracer across the cell face.

leads to the tracer flux

$$\mathbf{J}^{(n)} = \rho C^{(n)} (\mathbf{v}^{(n)} - \mathbf{v}), \quad (15.46)$$

the Eulerian flux-form tracer budget

$$\frac{\partial(\rho C^{(n)})}{\partial t} + \nabla \cdot [\mathbf{v} \rho C^{(n)} + \mathbf{J}^{(n)}] = 0, \quad (15.47)$$

and the corresponding material form of the tracer equation

$$\rho \frac{DC^{(n)}}{Dt} = -\nabla \cdot \mathbf{J}^{(n)}. \quad (15.48)$$

In Figure 15.3 we illustrate the contributions to the tracer evolution as viewed in the Eulerian flux-form tracer equation (15.47).

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

$$\rho C^{(n)} \mathbf{v}^{(n)} = \rho C^{(n)} (\mathbf{v}^{(n)} - \mathbf{v} + \mathbf{v}) = \mathbf{J}^{(n)} + \rho C^{(n)} \mathbf{v}. \quad (15.49)$$

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 15.5.

15.4.5 Compatibility between mass continuity and the tracer equation

By construction, the Eulerian flux-form of the tracer equation (15.47) is compatible with the flux-form continuity equation

$$\frac{\partial(\rho C^{(n)})}{\partial t} + \nabla \cdot [\rho \mathbf{v} C^{(n)} + \mathbf{J}^{(n)}] = 0 \iff \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (15.50)$$

Compatibility is manifest by summing the tracer equation over all constituents and using the identities

$$\sum_{n=1}^N C^{(n)} = 1 \quad \sum_{n=1}^N \mathbf{J}^{(n)} = 0. \quad (15.51)$$

Furthermore, through use of the barycentric velocity (15.36), 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 $\sum_{n=1}^N \mathbf{J}^{(n)} = 0$.

15.4.6 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 22.2). Correspondingly, the mass of trace matter in a fluid element is altered in the presence of tracer concentration gradients.

15.5 Passive tracers

As defined in equation (15.45), 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 (15.36) 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 14.1 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 25.2) and time scales for transport between fluid regions (Section 25.3).

In Section 15.7.2 as well as in Chapter 17, 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.

15.6 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 here synthesize these earlier presentations by considering mass budgets over an arbitrary finite sized domain. The resulting mass equations form the basis for matter budget analyses used in geophysical fluid mechanics.

15.6.1 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 19.

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?

In the following, we let ψ represent an intensive fluid property, so that $\psi \rho dV$ is the corresponding extensive property

$$\psi = \text{intensive fluid property} \quad (15.52\text{a})$$

$$\psi \rho dV = \text{extensive fluid property}. \quad (15.52\text{b})$$

For example, if ψ is the tracer concentration in an element of seawater (i.e., mass of tracer per mass of seawater), then the corresponding extensive property, $\psi \rho dV$, is the mass of tracer in the seawater element. If ψ is a component of the velocity vector, then the corresponding extensive property, $\psi \rho dV$, 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

$$\rho \frac{D\psi}{Dt} = -\nabla \cdot \mathbf{J} \quad \iff \quad \frac{\partial(\rho\psi)}{\partial t} + \nabla \cdot (\rho\psi \mathbf{v} + \mathbf{J}) = 0, \quad (15.53)$$

where \mathbf{J} is a diffusive flux such as that associated with the tracer equation derived in Section 15.4.4.

15.6.2 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

$$\mathcal{I}[\mathcal{R}(t), t] = \int_{\mathcal{R}(t)} \psi \rho dV \equiv \int_{\mathcal{R}(t)} \varphi dV, \quad (15.54)$$

where we introduced the shorthand

$$\varphi = \rho \psi. \quad (15.55)$$

The integrand in equation (15.54) 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 15.3) or a constant mass fluid region open to the exchange of matter with the surroundings (Section 15.4). Here we make no *a priori* assumption about the region.

The total time derivative of \mathcal{I} can be written as

$$\frac{d\mathcal{I}}{dt} = \left[\frac{\partial \mathcal{I}}{\partial t} \right]_{\mathcal{R}} + \frac{d\mathcal{R}}{dt} \left[\frac{\partial \mathcal{I}}{\partial \mathcal{R}} \right]_t. \quad (15.56)$$

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, weighted by the

dependence of the integral on the region itself. That is, how the integral evolves 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 (14.26).

15.6.3 Static domain

The simplest case is when the domain is static, in which case

$$\frac{d\mathcal{G}}{dt} = \left[\frac{\partial \mathcal{G}}{\partial t} \right]_{\mathcal{R}} = \frac{\partial}{\partial t} \left[\int_{\mathcal{R}} \psi \rho dV \right] = \int_{\mathcal{R}} \left[\frac{\partial(\rho \psi)}{\partial t} \right] dV. \quad (15.57)$$

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 (15.56) vanishes. Furthermore, since the domain is static, the volume increment 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 15.1, 15.2, and 15.3.

15.6.4 Leibniz-Reynolds Transport Theorem

Now allow the domain boundaries to be time dependent, so that both terms in the total time derivative in equation (15.56) contribute. We in turn derive a most general expression of scalar conservation over an arbitrary region, with the result termed the *Leibniz-Reynolds Transport Theorem*.

Derivation of the transport theorem

Consider a one-dimensional case and make use of the chain rule from differential calculus. 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

$$\frac{d}{dt} \left[\int_{x_1(t)}^{x_2(t)} \varphi(x, t) dx \right] = \int_{x_1(t)}^{x_2(t)} \frac{\partial \varphi}{\partial t} dx + \frac{d}{dt} \left[\int_{x_1(t)}^{x_2(t)} \right] \varphi(x, t) dx \quad (15.58a)$$

$$= \int_{x_1(t)}^{x_2(t)} \frac{\partial \varphi}{\partial t} dx + \frac{dx_2(t)}{dt} \varphi(x_2, t) - \frac{dx_1(t)}{dt} \varphi(x_1, t) \quad (15.58b)$$

$$= \int_{x_1(t)}^{x_2(t)} \left[\frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial x} \left(\varphi \frac{dx}{dt} \right) \right] dx. \quad (15.58c)$$

As a matter of convenience we brought the boundary terms back inside the integral for the final equality.

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 15.7). Its generalization to a three-dimensional integral is straightforward

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \varphi dV \right] = \int_{\mathcal{R}} \left[\frac{\partial \varphi}{\partial t} + \nabla \cdot \left(\varphi \frac{d\mathbf{x}}{dt} \right) \right] dV = \int_{\mathcal{R}} \frac{\partial \varphi}{\partial t} dV + \int_{\partial\mathcal{R}} \varphi \frac{d\mathbf{x}}{dt} \cdot \hat{\mathbf{n}} dS, \quad (15.59)$$

where the second equality made use of Gauss's Law to transfer the volume integral into a boundary integral.

We emphasize that the boundary term in equation (15.59) projects out just the normal component to the boundary velocity. We thus never need to have any information about the tangential component to that velocity. Hence, the boundary can be exhibiting an arbitrary motion in its tangential direction without altering the result (15.59).

Transport theorem for a scalar field

The identity (15.59) is known as the Leibniz-Reynolds Transport Theorem. It is a result of great practical utility for fluid mechanics since we are often interested in the evolution of fluid properties within an arbitrary moving domain. We here display a form of the theorem for scalar fields that proves quite useful for budget analyses over moving regions. A generalization to the vector linear momentum is provided in Section 28.5.

For this purpose, make use of the Eulerian flux-form of the scalar conservation equation (15.53) so that the transport theorem is written

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \psi dV \right] = - \int_{\partial\mathcal{R}} [\rho \psi (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} dS, \quad (15.60)$$

where we wrote the velocity of the boundary as

$$\mathbf{v}^{(b)} = \frac{d\mathbf{x}}{dt}. \quad (15.61)$$

The transport theorem (15.60) 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.

In Figure 15.4 we illustrate the transport theorem (15.60) 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 18 and their dynamics in Chapter 31. In particular, in Section 18.3 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.

15.6.5 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 15.3.4. The following results are mere special cases of the general expression (15.60). Even so, it is useful pedagogically to work through the special cases.

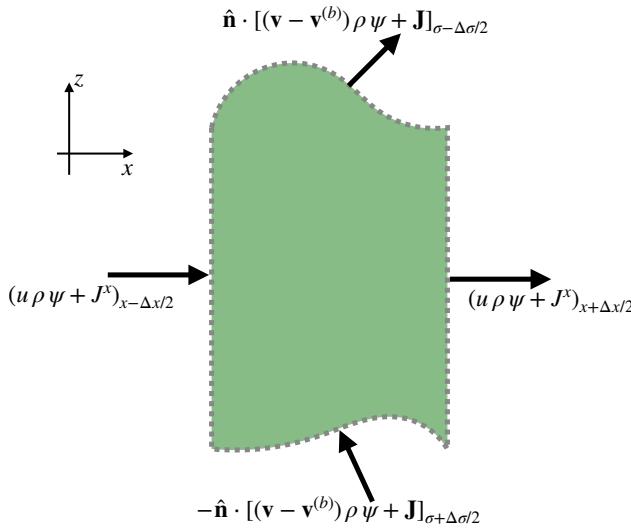


Figure 15.4: This image illustrates the contributions to the Leibnitz-Reynolds Transport Theorem (15.60). The theorem is applied to a domain corresponding to a numerical model grid cell with the top and bottom interfaces defined by generalized vertical coordinates of Chapters 8, 18, and 31. In particular, 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.

Reynolds Transport Theorem

Let us apply the result (15.59) 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 (15.59) is given by the fluid velocity. We thus have

$$\frac{D}{Dt} \left[\int_{\mathcal{R}} \varphi dV \right] = \int_{\mathcal{R}} \left[\frac{\partial \varphi}{\partial t} + \nabla \cdot (\mathbf{v} \varphi) \right] dV = \int_{\mathcal{R}} \left[\frac{D\varphi}{Dt} + \varphi \nabla \cdot \mathbf{v} \right] dV, \quad (15.62)$$

where the second equality introduced the material time derivative (14.28). Note that we wrote D/Dt for the time derivative acting on the integral in the first equality. This usage is consistent with that for material regions in Section 15.3, with here the time derivative computed following the barycentric velocity.

Equation (15.62) is the Reynolds Transport Theorem. The derivation is more general than that in Section 15.3.4, 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 permeable boundaries arising in the presence of diffusion.

Alternative form of Reynolds Transport Theorem

We can put the Reynolds Transport Theorem (15.62) into another useful form by reintroducing $\varphi = \rho\psi$ and making use of mass continuity

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{v}. \quad (15.63)$$

Doing so yields the rather tidy result

$$\frac{D}{Dt} \left[\int_{\mathcal{R}} \psi \rho dV \right] = \int_{\mathcal{R}} \left[\frac{D\varphi}{Dt} + \varphi \nabla \cdot \mathbf{v} \right] dV \quad (15.64a)$$

$$= \int_{\mathcal{R}} \left[\frac{D(\rho\psi)}{Dt} + \rho\psi \nabla \cdot \mathbf{v} \right] dV \quad (15.64b)$$

$$= \int_{\mathcal{R}} \left[\psi \left(\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right) + \rho \frac{D\psi}{Dt} \right] dV \quad (15.64c)$$

$$= \int_{\mathcal{R}} \frac{D\psi}{Dt} \rho dV. \quad (15.64d)$$

Heuristically, this result follows since ρdV is a constant when following the flow, so that passage of the material time derivative across the integral only picks up the derivative of ψ .

We can take the result (15.64d) one more step by inserting the material form of the scalar conservation equation (15.53) so that

$$\frac{D}{Dt} \left[\int_{\mathcal{R}} \psi \rho dV \right] = - \int_{\partial\mathcal{R}} \mathbf{J} \cdot \hat{\mathbf{n}} dS, \quad (15.65)$$

which is a special case of the general transport theorem (15.60) 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 15.4.4. We extend this result to linear momentum in Section 28.5.

15.7 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 along 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.

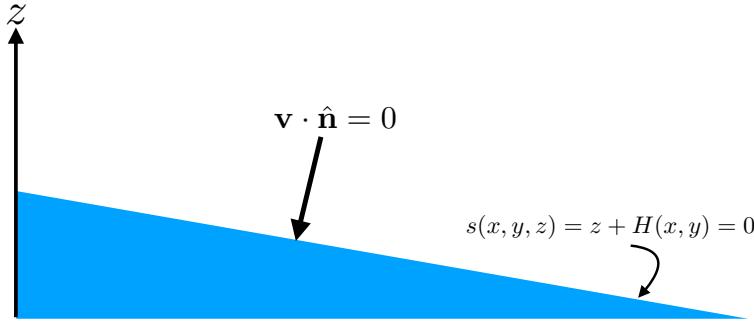


Figure 15.5: Illustrating the no-normal flow boundary condition maintained for a solid boundary, on which $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$ (equation (15.66)). When the solid boundary denotes the ocean bottom, then the position of the interface is written $s(x, y, z) = z + H(x, y) = 0$ (equation (15.67)).

15.7.1 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 15.5)

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{no-flux condition on static material boundary.} \quad (15.66)$$

Recall our discussion of streamlines in Section 14.7.2, 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 (15.66). Doing so helps to develop a geometric formalism especially useful for the more complicated moving boundary conditions in Sections 15.7.2 and 15.7.3. For this purpose, introduce a coordinate expression for the boundary according to

$$s(x, y, z) = z + H(x, y) = 0 \quad \text{static material boundary,} \quad (15.67)$$

with $z = -H(x, y)$ the vertical position of the boundary. The outward normal vector at the boundary is thus given by

$$\hat{\mathbf{n}} = -\frac{\nabla s}{|\nabla s|} = -\frac{\nabla(z + H)}{|\nabla(z + H)|} = -\frac{\hat{\mathbf{z}} + \nabla H}{\sqrt{1 + (\nabla H)^2}}. \quad (15.68)$$

Consequently, the no-flux boundary condition (15.66) takes the form

$$w + \mathbf{u} \cdot \nabla H = 0 \quad \text{at } z = -H(x, y), \quad (15.69)$$

where the velocity is decomposed into its horizontal and vertical components,

$$\mathbf{v} = (\mathbf{u}, w). \quad (15.70)$$

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 $w = 0$.

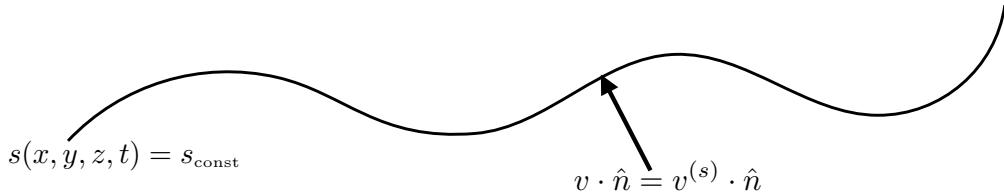


Figure 15.6: Illustrating the boundary condition for a moving material surface, on which $\mathbf{v} \cdot \hat{\mathbf{n}} = \mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}$ (equation (15.71)). For many cases, we can specify the surface by the value of a function that is a constant on the surface: $s(x, y, z, t) = s_{\text{const}}$ (equation (15.72)), in which case the normal direction is given by $\hat{\mathbf{n}} = |\nabla s|^{-1} \nabla s$ (equation (15.73)). For a multi-component fluid, \mathbf{v} is the barycentric velocity so that the material surface allows matter to be exchanged across it in the presence of diffusion. Hence, the moving material boundary is here a generalization of the boundary for a fluid element discussed in Section 15.4.

15.7.2 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

$$(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{moving material boundary condition.} \quad (15.71)$$

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 15.6. Furthermore, as a Corollary to the boundary condition (15.71), 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 14.7.2 and 14.7.3).

Specialized expression of the boundary condition

Now specialize the kinematic condition (15.71) to the case of a material surface specified by a function that takes on a constant value on the surface

$$s(\mathbf{x}, t) = s_{\text{const}}. \quad (15.72)$$

Correspondingly, the surface normal vector is given by

$$\hat{\mathbf{n}} = |\nabla s|^{-1} \nabla s. \quad (15.73)$$

From Section 14.4.6, we know that a point fixed on an arbitrary surface has a velocity that satisfies (see equation (14.41))

$$\frac{\partial s}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla s = 0 \quad \text{on an iso-surface } s(\mathbf{x}, t) = s_0. \quad (15.74)$$

Use of the identity

$$\frac{\partial s}{\partial t} = \frac{Ds}{Dt} - \mathbf{v} \cdot \nabla s \quad (15.75)$$

renders

$$\frac{Ds}{Dt} - \mathbf{v} \cdot \nabla s + \mathbf{v}^{(s)} \cdot \nabla s = \frac{Ds}{Dt} + (\mathbf{v}^{(s)} - \mathbf{v}) \cdot \nabla s = 0. \quad (15.76a)$$

Since $(\mathbf{v}^{(s)} - \mathbf{v}) \cdot \nabla s = 0$ from the boundary condition (15.71), we are left with the material constancy condition

$$\frac{Ds}{Dt} = \frac{\partial s}{\partial t} + \mathbf{v} \cdot \nabla s = 0 \quad \text{on material surface } s(\mathbf{x}, t) = s_0. \quad (15.77)$$

Consequently, no net matter crosses a surface of constant s as long as s is materially constant. This is a very important kinematic property that reappears in many forms throughout this book.

Boundary condition for a material interface

The expression (15.77) 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

$$s(x, y, z, t) = z - \eta(x, y, t) = s_o. \quad (15.78)$$

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 (15.77) thus takes the form

$$\frac{Ds}{Dt} = \frac{D(z - \eta)}{Dt} = 0. \quad (15.79)$$

Hence, the vertical velocity component at the interface equals to the material time derivative of the interface displacement

$$\frac{Dz}{Dt} = \frac{D\eta}{Dt} \Rightarrow w = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta \quad \text{material b.c. at interface } z = \eta(x, y, t). \quad (15.80)$$

This boundary condition can be equivalently written in the form

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \frac{\partial \eta / \partial t}{\sqrt{1 + |\nabla \eta|^2}}, \quad (15.81)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{-\nabla \eta + \hat{\mathbf{z}}}{\sqrt{1 + |\nabla \eta|^2}} \quad (15.82)$$

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.

15.7.3 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

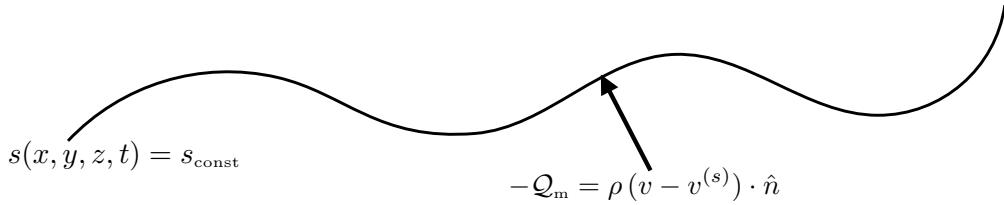


Figure 15.7: Illustrating the boundary condition for a moving permeable surface, such as the ocean free surface. On this surface, the boundary condition states that $\rho(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} dS = -\mathcal{Q}_m$ (equation (15.83)). In the special case of an ocean free surface with no overturns, this boundary condition reduces to equation (15.96).

we allow for a generally non-zero net transport (mass per time). We write this transport condition as

$$\rho(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} dS = -\mathcal{Q}_m dS \quad \text{moving non-material boundary condition.} \quad (15.83)$$

In this equation, dS 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 (15.74) for $\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}}$ holds for a point on an arbitrary surface so that

$$\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = -\frac{\partial s / \partial t}{|\nabla s|}. \quad (15.84)$$

Furthermore, the projection of the barycentric fluid velocity onto the normal direction can be written

$$\frac{Ds}{Dt} = \frac{\partial s}{\partial t} + \mathbf{v} \cdot \nabla s \Rightarrow \mathbf{v} \cdot \hat{\mathbf{n}} = \frac{1}{|\nabla s|} \left(\frac{Ds}{Dt} - \frac{\partial s}{\partial t} \right). \quad (15.85)$$

Bringing these results together leads to

$$\rho(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} dS = \frac{\rho dS}{|\nabla s|} \frac{Ds}{Dt}. \quad (15.86)$$

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 15.7.2).

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

$$\frac{\partial s}{\partial z} \neq 0, \quad (15.87)$$

so that we can further massage the boundary condition (15.86) by writing the area factor in the form

$$\frac{dS}{|\nabla s|} = \frac{dS}{\sqrt{(\partial s/\partial x)^2 + (\partial s/\partial y)^2 + (\partial s/\partial z)^2}} \quad (15.88a)$$

$$= \frac{dS}{|\partial s/\partial z| \sqrt{[(\partial s/\partial x)/(\partial s/\partial z)]^2 + [(\partial s/\partial y)/(\partial s/\partial z)]^2 + 1}} \quad (15.88b)$$

$$= \frac{dS}{|\partial s/\partial z| \sqrt{1 + \tan^2 \theta}} \quad (15.88c)$$

$$= \left| \frac{\partial z}{\partial s} \right| |\cos \theta| dS \quad (15.88d)$$

$$= \left| \frac{\partial z}{\partial s} \right| dA. \quad (15.88e)$$

The equality (15.88c) introduced the angle, θ , between the boundary surface and the horizontal plane. The squared slope of this surface given by

$$\tan^2 \theta = \frac{\nabla_z s \cdot \nabla_z s}{(\partial s/\partial z)^2} = \nabla_s z \cdot \nabla_s z \quad (15.89)$$

with

$$\nabla_z = \hat{x} \left[\frac{\partial}{\partial x} \right]_{y,z} + \hat{y} \left[\frac{\partial}{\partial y} \right]_{x,z} \quad (15.90)$$

the horizontal gradient operator on constant z surfaces, and

$$\nabla_s = \hat{x} \left[\frac{\partial}{\partial x} \right]_{y,s} + \hat{y} \left[\frac{\partial}{\partial y} \right]_{x,s} \quad (15.91)$$

the horizontal gradient operator on constant s surfaces.² The equality (15.88d) made use of a trigonometric identity, and the equality (15.88e) introduced the horizontal projection of the area,

$$dA = |\cos \theta| dS. \quad (15.92)$$

These results bring the kinematic boundary condition (15.86) into the form

$$\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{n} dS = -Q_m dS \quad (15.93a)$$

$$= \rho \frac{Ds}{Dt} \left| \frac{\partial z}{\partial s} \right| dA \quad (15.93b)$$

$$\equiv -Q_m dA. \quad (15.93c)$$

As defined, the flux Q_m is the net mass per time per horizontal area crossing the boundary surface

$$Q_m = -\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{n} \left(\frac{dS}{dA} \right) = -\rho \frac{Ds}{Dt} \left| \frac{\partial z}{\partial s} \right|. \quad (15.94)$$

The minus sign is a convention that we motivate through the ocean free surface example in the following.

²We discuss such generalized vertical coordinates in Chapter 8.

Kinematic boundary condition at the ocean free surface

Consider the ocean free surface located at

$$s(x, y, z, t) = z - \eta(x, y, t) = 0 \quad \text{ocean free surface.} \quad (15.95)$$

For this boundary, $\partial s / \partial z = 1$, so that the boundary condition (15.93b) takes on the form

$$\rho \left[\frac{D(z - \eta)}{Dt} \right] = -Q_m \Rightarrow w + \rho^{-1} Q_m = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta. \quad (15.96)$$

To motivate the sign convention in equation (15.93c), 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

$$s = \sigma(x, y, z, t) \quad (15.97)$$

for a particular isopycnal σ . The mass transport crossing the isopycnal is written

$$Q_m = \rho \frac{D\sigma}{Dt} \left| \frac{\partial z}{\partial \sigma} \right| \equiv \rho w^{(\sigma)}, \quad (15.98)$$

where we introduced the *diapycnal velocity*

$$w^{(\sigma)} \equiv \frac{D\sigma}{Dt} \left| \frac{\partial z}{\partial \sigma} \right|. \quad (15.99)$$

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.

15.8 Exercises

EXERCISE 15.1: CENTER OF MASS MOTION

Consider a material fluid region, \mathcal{R} , with constant mass written as

$$M = \int_{\mathcal{R}} \rho dV. \quad (15.100)$$

- (a) Show that the centre of mass for the region moves with the region's total linear momentum

$$\frac{D}{Dt} \left[\frac{1}{M} \int_{\mathcal{R}} \rho \mathbf{x} dV \right] = \frac{1}{M} \int_{\mathcal{R}} \rho \mathbf{v} dV. \quad (15.101)$$

- (b) Show that the time change in the linear momentum for the region is given by

$$\frac{D}{Dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}} \frac{D\mathbf{v}}{Dt} \rho dV. \quad (15.102)$$

EXERCISE 15.2: KINEMATIC OCEAN FREE SURFACE EQUATION

Vertically integrate the mass continuity equation 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 the bottom and surface kinematic boundary conditions to derive the following kinematic expression for the free surface time tendency

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho(\eta)} - \nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right] - \int_{-H}^{\eta} \frac{1}{\rho} \frac{D\rho}{Dt} dz, \quad (15.103)$$

where $\rho(\eta) = \rho(z = \eta)$ is the ocean density evaluated at the ocean free surface.

EXERCISE 15.3: 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 Exercise 15.2, but we come at the problem differently and arrive at a slightly different (though equivalent) form for column mass balance. Figure 15.8 provides a schematic of the setup. This is a long problem to state, but the solution is actually quite minimal.

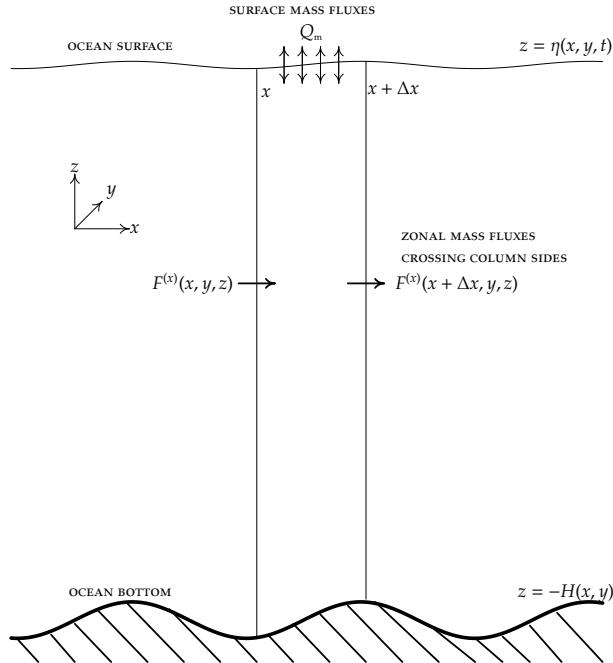


Figure 15.8: A longitudinal-vertical slice of ocean fluid from the surface at $z = \eta(x, y, t)$ to bottom at $z = -H(x, y)$. The horizontal boundaries of the column x and $x + \Delta x$ are static, meaning that the horizontal cross-sectional area for the fluid column is time independent. The ocean bottom at the solid-earth boundary, $z = -H(x, y)$, is also static with no mass crossing this surface. The ocean surface at $z = \eta(x, y, t)$ is time dependent with mass flux Q_m crossing this surface.

The mass within an arbitrary fluid region is given by

$$m = \int \rho dV. \quad (15.104)$$

Consider the fluid mass within the column shown in Figure 15.8. 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

$$m = \iint_{\text{column}} dx dy \int_{-H(x,y)}^{\eta(x,y,t)} \rho dz, \quad (15.105)$$

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

$$\frac{dm}{dt} = - \iint_{\text{surface}} \rho \Delta \mathbf{v} \cdot \hat{\mathbf{n}} dS, \quad (15.106)$$

where $\hat{\mathbf{n}}$ is the outward normal to the surface of the fluid region, dS 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

$$\Delta \mathbf{v} = \mathbf{v} - \mathbf{v}^{(s)} \quad (15.107)$$

is the velocity of the fluid relative to the velocity of the boundary.

Mass transported in the zonal direction ($\hat{\mathbf{x}}$) that crosses the column's vertical boundary at x is given by

$$F^{(x)}(x, t) = \int dy \int_{-H(x,y)}^{\eta(x,y,t)} u(x, y, z, t) \rho(x, y, z, t) dz. \quad (15.108)$$

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 ($\hat{\mathbf{x}}, \hat{\mathbf{y}}$). 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

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} \rho dz \right] = -\nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} \rho dz \right] + Q_m. \quad (15.109)$$

In this expression, Q_m is the mass transport entering the ocean through the surface, per horizontal area, so that

$$\iint Q_m dx dy = - \iint \rho \Delta \mathbf{v} \cdot \hat{\mathbf{n}} dS \quad \text{at } z = \eta. \quad (15.110)$$

In words, the mass budget in equation (15.109) 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 (15.109)?

EXERCISE 15.4: 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

$$M = \int_{\mathcal{R}} \rho dV. \quad (15.111)$$

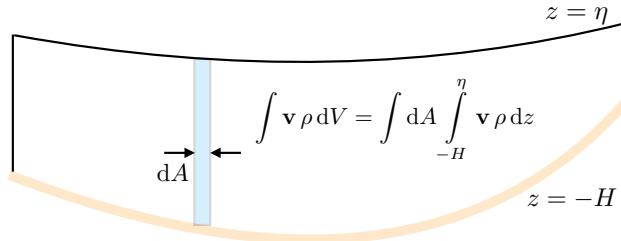


Figure 15.9: Cross-section of the integration region for Exercise 15.4, with the region extending from the ocean bottom at $z = -H(x, y)$ and the free surface at $z = \eta(x, y, t)$. An infinitesimal column is shown with cross-sectional area dA , extending from the bottom to the surface. The cross-sectional 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 (15.113).

Show that the time change in the linear momentum for this ocean basin is given by

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}} \frac{D\mathbf{v}}{Dt} \rho dV. \quad (15.112)$$

This result is identical to that derived in Exercise 15.1. Rather than just repeating the solution method used there, make explicit use of Leibniz's theorem, the kinematic boundary condition detailed in Section 15.7.2, and mass conservation.

Hint: Refer to Figure 15.9 for a schematic of the integration where we have expanded the volume integral into the form

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \frac{d}{dt} \left[\int dA \int_{-H}^{\eta} \rho \mathbf{v} dz \right], \quad (15.113)$$

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 15.7.3 (in particular equation (15.92)). Unlike the formulation in Exercise 15.1, 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.

16

Incompressible flow

In this chapter, we specialize the general kinematics from Chapters 14 and 15 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 5 and 6. We presume an understanding of the kinematics of mass conservation from Chapter 15. This is a relatively brief chapter, and yet it introduces many concepts and tools of use for the remainder of the book.

16.1	Introduction to incompressible flow	211
16.2	Kinematic boundary conditions	212
16.3	Streamfunction for two-dimensional flow	212
16.3.1	Streamfunction isolines are streamlines	213
16.3.2	Streamfunction is constant on material boundaries	213
16.3.3	Transport between two points	213
16.4	Vector streamfunction for three-dimensional flow	214
16.5	Evolution of material volume and area	215
16.6	Meridional-depth overturning circulation	215
16.7	Gauge symmetry	218
16.8	Exercises	218

16.1 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 29.1), whereby the volume of a fluid element is constant. Recalling the expression

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v} \quad (16.1)$$

from Section 17.4.1, we see that a constant volume for a fluid element constrains the velocity field to be non-divergent

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = 0 \Rightarrow \nabla \cdot \mathbf{v} = 0 \quad \text{incompressible.} \quad (16.2)$$

A slightly less onerous constraint arises from the anelastic approximation, whereby

$$\nabla \cdot (\rho \mathbf{v}) = 0. \quad (16.3)$$

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.

16.2 Kinematic boundary conditions

For incompressible flow, there are slight modifications to the compressible boundary conditions detailed in Section 15.7. 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 (15.83) to a volume transport condition

$$\rho_0 (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} dS = -Q_m dS. \quad \text{moving non-material boundary condition.} \quad (16.4)$$

Correspondingly, the kinematic boundary condition (15.96) applied at the ocean free surface takes on the form

$$\rho_0 \frac{D(z - \eta)}{Dt} = -Q_m \Rightarrow w + \frac{Q_m}{\rho_0} = \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta. \quad (16.5)$$

16.3 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

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi = -\hat{\mathbf{x}} \frac{\partial \psi}{\partial y} + \hat{\mathbf{y}} \frac{\partial \psi}{\partial x}. \quad (16.6)$$

The constraint $\nabla \cdot \mathbf{u} = 0$ is satisfied since the partial derivative operators commute

$$\frac{\partial^2 \psi}{\partial x \partial y} = \frac{\partial^2 \psi}{\partial y \partial x}. \quad (16.7)$$

We refer to ψ as the *streamfunction*, with this name motivated by the following considerations.

16.3.1 Streamfunction isolines are streamlines

At any fixed time, the total differential of the streamfunction is

$$d\psi = \frac{\partial\psi}{\partial x} dx + \frac{\partial\psi}{\partial y} dy \quad (16.8a)$$

$$= v dx - u dy, \quad (16.8b)$$

where the second equality follows from equation (16.6). Instantaneous lines along which ψ is a constant satisfy

$$d\psi = 0 \Rightarrow \frac{dx}{u} = \frac{dy}{v}. \quad (16.9)$$

Furthermore, the normal direction to constant ψ lines

$$\hat{\mathbf{n}} = \frac{\nabla\psi}{|\nabla\psi|} = \frac{v \hat{\mathbf{x}} - u \hat{\mathbf{y}}}{|\mathbf{u}|} \quad (16.10)$$

is normal to the velocity

$$\mathbf{u} \cdot \nabla\psi = u v - v u = 0. \quad (16.11)$$

Consequently, at each time instance, lines of constant ψ are streamlines (see Section 14.7.2 for discussion of streamlines). This property in turn motivates the name *streamfunction*.

16.3.2 Streamfunction is constant on material boundaries

As a corollary to the results from Section 16.3.1, 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 (16.11). We can also see it from

$$0 = \mathbf{u} \cdot \mathbf{n} = (\hat{\mathbf{z}} \wedge \nabla\psi) \cdot \hat{\mathbf{n}} = (\hat{\mathbf{n}} \wedge \hat{\mathbf{z}}) \cdot \nabla\psi = \hat{\mathbf{t}} \cdot \nabla\psi, \quad (16.12)$$

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.

16.3.3 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

$$\psi(\mathbf{x}_1) - \psi(\mathbf{x}_2) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} d\psi \quad (16.13a)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \left(dx \frac{\partial\psi}{\partial x} + dy \frac{\partial\psi}{\partial y} \right) \quad (16.13b)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \nabla\psi \cdot d\mathbf{x} \quad (16.13c)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \nabla\psi \cdot \hat{\mathbf{t}} ds. \quad (16.13d)$$

For the final equality we wrote

$$d\mathbf{x} = \hat{\mathbf{t}} ds, \quad (16.14)$$

where

$$ds = |d\mathbf{x}| \quad (16.15)$$

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

$$\hat{\mathbf{t}} = \hat{\mathbf{n}} \wedge \hat{\mathbf{z}}, \quad (16.16)$$

which then allows us to write

$$\psi(\mathbf{x}_1) - \psi(\mathbf{x}_2) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} \nabla \psi \cdot (\hat{\mathbf{n}} \wedge \hat{\mathbf{z}}) ds \quad (16.17a)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} (\hat{\mathbf{z}} \wedge \nabla \psi) \cdot \hat{\mathbf{n}} ds \quad (16.17b)$$

$$= \int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{u} \cdot \hat{\mathbf{n}} ds. \quad (16.17c)$$

The final equality is an expression for the net area transport of fluid normal to the curve.¹ 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

$$|\mathbf{u}| = |\nabla \psi|. \quad (16.18)$$

16.4 Vector streamfunction for three-dimensional flow

A three-dimensional non-divergent velocity, $\nabla \cdot \mathbf{v} = 0$, can be specified by a vector streamfunction

$$\mathbf{v} = \nabla \wedge \Psi. \quad (16.19)$$

The constraint $\nabla \cdot \mathbf{v} = 0$ is trivially satisfied since the divergence of the curl vanishes

$$\nabla \cdot (\nabla \wedge \Psi) = 0. \quad (16.20)$$

In a three-dimensional fluid, the volume transport across a surface is defined by

$$\mathcal{T}(\mathcal{S}) = \int_{\mathcal{S}} \mathbf{v} \cdot \hat{\mathbf{n}} dS, \quad (16.21)$$

¹For two-dimensional flow, the units of transport are squared length per time.

where $\hat{\mathbf{n}}$ is the outward unit normal vector on the surface. Introducing the vector streamfunction and making use of Stokes' Theorem (Section 3.6) then leads to

$$\mathcal{T}(\mathcal{S}) = \int_{\mathcal{S}} \mathbf{v} \cdot \hat{\mathbf{n}} dS \quad (16.22a)$$

$$= \int_{\mathcal{S}} (\nabla \wedge \Psi) \cdot \hat{\mathbf{n}} dS \quad (16.22b)$$

$$= \oint_{\partial\mathcal{S}} \Psi \cdot \hat{\mathbf{t}} ds. \quad (16.22c)$$

Hence, the transport through the surface depends only on the vector streamfunction on the perimeter of the surface. If the transport through the surface vanishes, then on the perimeter the vector streamfunction can be written as the gradient of an arbitrary scalar, $\Psi = \nabla\phi$, since

$$\oint_{\partial\mathcal{S}} \Psi \cdot \hat{\mathbf{t}} ds = \oint_{\partial\mathcal{S}} \nabla\phi \cdot \hat{\mathbf{t}} ds = \oint_{\partial\mathcal{S}} \nabla\phi \cdot d\mathbf{x} = 0. \quad (16.23)$$

16.5 Evolution of material volume and area

As shown by equation (16.2), the volume of a material parcel remains materially constant in an incompressible flow. Correspondingly, a material fluid region maintains a constant volume

$$\frac{D}{Dt} \int_{\mathcal{V}} dV = \int_{\mathcal{V}} \frac{D(\delta V)}{Dt} \quad (16.24a)$$

$$= \int_{\mathcal{V}} (\nabla \cdot \mathbf{v}) dV \quad (16.24b)$$

$$= 0. \quad (16.24c)$$

Likewise, following from the material area element equation (17.43), the area of a material region in a two-dimensional incompressible flow remains materially constant

$$\frac{D}{Dt} \int_{\mathcal{S}} dA = \int_{\mathcal{S}} \frac{D(\delta A)}{Dt} \quad (16.25a)$$

$$= \int_{\mathcal{S}} (\nabla \cdot \mathbf{u}) dV \quad (16.25b)$$

$$= 0. \quad (16.25c)$$

This area preservation property is illustrated in Figure 16.1, in which a two-dimensional flow is seen to deform a black/white grid, yet to retain a fixed area.

16.6 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

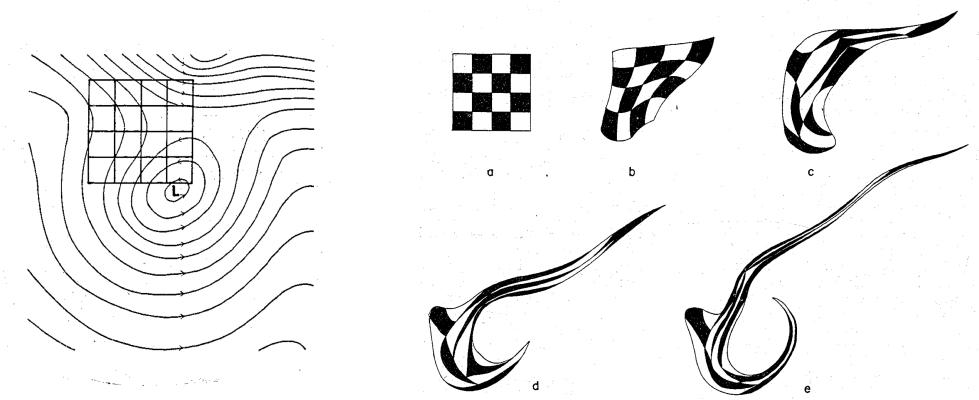


Figure 16.1: This figure is taken from Figure 2 of [Welander \(1955\)](#). It shows the deformation and rotation of a material area in two-dimensional non-divergent flow (left panel) after 6, 12, 24, and 36 hours from a model simulation. As discussed in Section 16.5, the area of a material region remains fixed in two-dimensional non-divergent flow.

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

$$\mathcal{V}^\rho = \int_{x_1}^{x_2} \rho v \, dx \quad \mathcal{W}^\rho = \int_{x_1}^{x_2} \rho w \, dx. \quad (16.26)$$

We can go even further for incompressible flow, or for steady state compressible flow, in which case we can introduce a streamfunction for the meridional-depth circulation. For definiteness, consider incompressible flow and the zonal integrated velocity

$$\mathcal{V} = \int_{x_1}^{x_2} v \, dx \quad \mathcal{W} = \int_{x_1}^{x_2} w \, dx. \quad (16.27)$$

Taking the meridional derivative of the meridional transport leads to

$$\frac{\partial \mathcal{V}}{\partial y} = \frac{\partial}{\partial y} \left[\int_{x_1}^{x_2} v \, dx \right] \quad (16.28a)$$

$$= \int_{x_1}^{x_2} \frac{\partial v}{\partial y} \, dx \quad (16.28b)$$

$$= - \int_{x_1}^{x_2} \left[\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right] \, dx \quad (16.28c)$$

$$= - \int_{x_1}^{x_2} \frac{\partial w}{\partial z} \, dx \quad (16.28d)$$

$$= - \frac{\partial}{\partial z} \left[\int_{x_1}^{x_2} w \, dx \right] \quad (16.28e)$$

$$= - \frac{\partial \mathcal{W}}{\partial z}, \quad (16.28f)$$

where we moved derivatives across the integral sign due to the orthogonality of the (x, y, z) coordinates. These results show that the two-dimensional zonally integrated transport is non-divergent

$$\frac{\partial \mathcal{V}}{\partial y} + \frac{\partial \mathcal{W}}{\partial z} = 0. \quad (16.29)$$

Consequently, we can introduce a meridional-depth streamfunction

$$\Psi(y, z, t) = - \int_{-H_{\max}}^z \mathcal{V} \, dz \quad (16.30)$$

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

$$\frac{\partial \Psi}{\partial z} = -\mathcal{V} \quad (16.31)$$

and

$$\frac{\partial \Psi}{\partial y} = - \frac{\partial}{\partial y} \left[\int_{-H_{\max}}^z \mathcal{V} \, dz \right] \quad (16.32a)$$

$$= - \int_{-H_{\max}}^z \frac{\partial \mathcal{V}}{\partial y} \, dz \quad (16.32b)$$

$$= \int_{-H_{\max}}^z \frac{\partial \mathcal{W}}{\partial z} \, dz \quad (16.32c)$$

$$= \mathcal{W}. \quad (16.32d)$$

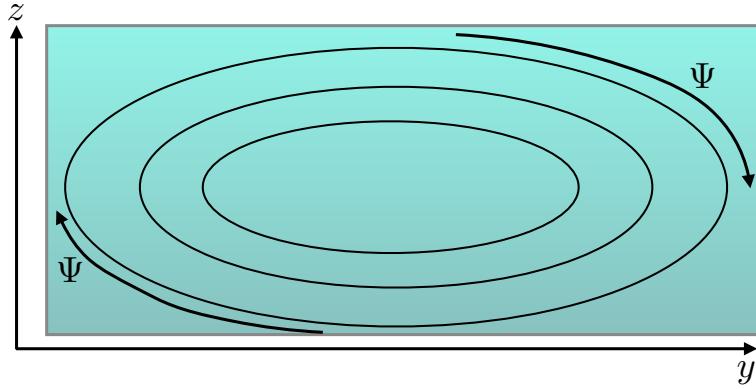


Figure 16.2: An idealized rendition of the meridional-depth overturning circulation found in both the atmosphere and ocean. Shown here are streamlines for the zonally integrated flow between two solid boundaries or over a zonally periodic domain. The flow is assumed to be non-divergent, as per equation (16.29). 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.

To reach this result, we made use of the non-divergent condition (16.29), and set

$$\mathcal{W}(z = -H_{\max}) = 0. \quad (16.33)$$

An idealized version of the meridional-depth circulation is shown in Figure 16.2.

16.7 Gauge symmetry

For the 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, note that the streamfunction is arbitrary up to a constant, k , since

$$\psi' = \psi + k \Rightarrow \mathbf{u}' = \mathbf{u}. \quad (16.34)$$

So the value of the streamfunction at a particular point no absolute meaning. Rather, only the difference in streamfunction between two points is unambiguous. The ability to add a constant to the streamfunction is termed a *gauge symmetry*. Gauge symmetries are common throughout physics, with the example here similar to those appearing in electrodynamics.

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

$$\mathbf{v} = \nabla \wedge \Psi. \quad (16.35)$$

Gauge symmetry manifests through the ability to add the gradient of an arbitrary function to Ψ without altering \mathbf{v} :

$$\Psi' = \Psi + \nabla \lambda \Rightarrow \mathbf{v}' = \mathbf{v}. \quad (16.36)$$

16.8 Exercises

EXERCISE 16.1: STREAMLINES FOR CELLULAR FLOW

Sketch the velocity field for this streamfunction

$$\psi(x, y) = A \sin(k x) \sin(l y), \quad (16.37)$$

where (k, l) are the zonal and meridional wavenumbers.

EXERCISE 16.2: KINEMATIC FREE SURFACE EQUATION FOR INCOMPRESSIBLE FLOW

Vertically integrate the incompressibility constraint 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 the bottom and surface kinematic boundary conditions for an incompressible fluid to derive the following kinematic expression for the free surface time tendency

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right]. \quad (16.38)$$

Thus conclude that for the special case of a steady state with zero boundary flux, the depth integrated flow is non-divergent

$$\nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} dz \right] = 0 \quad \text{if } Q_m = 0 \text{ and } \partial \eta / \partial t = 0. \quad (16.39)$$

EXERCISE 16.3: TRANSPORT AND CIRCULATION

Consider a two dimensional velocity field, $\mathbf{u} = u \hat{x} + v \hat{y}$, that has both a zero divergence and a zero curl

$$\nabla \cdot \mathbf{u} = \nabla \wedge \mathbf{u} = 0. \quad (16.40)$$

- (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 16.4: 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 16.5: 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 16.6: NET VOLUME TRANSPORT ACROSS AN ARBITRARY SURFACE

Consider a non-divergent flow in a container with static sides/bottom. Draw an arbitrary static surface within the fluid from one side of the container to the other. Integrate the volume transport over the surface, $\int \mathbf{v} \cdot \hat{\mathbf{n}} dS$. Show that this transport vanishes. That is, the net transport across the surface is zero. Specialize this result to a horizontal surface, so to show that there is zero area integrated vertical transport across the surface, $\int w dx dy = 0$.

EXERCISE 16.7: SOLID BODY ROTATION

Consider a velocity field corresponding to a time-independent solid-body rotation on a plane

$$\mathbf{u} = \Omega \hat{z} \wedge \mathbf{x} = \Omega (-y \hat{x} + x \hat{y}), \quad (16.41)$$

where Ω is a constant rotation rate.

- (a) Compute the relative vorticity $\nabla \wedge \mathbf{u}$.
- (b) Compute the streamfunction $\mathbf{u} = \nabla \wedge (\hat{z} \psi)$. 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 16.8: ALTERNATIVE FORM OF MERIDIONAL-DEPTH STREAMFUNCTION

In equation (16.30), we introduced the meridional-depth overturning streamfunction

$$\Psi(y, z, t) = - \int_{-H_{\max}}^z \mathcal{V} dz. \quad (16.42)$$

Show that an alternative streamfunction is given by

$$\Gamma(y, z, t) = \int_{y_s}^y \mathcal{W} dy, \quad (16.43)$$

where y_s is a constant latitude just south of the southern-most latitude in the domain.

EXERCISE 16.9: VOLUME TRANSPORT THROUGH STREAMTUBE ENDS

Recall our discussion of streamtubes in Section 14.7.2 (see in particular Figure 14.5). For a steady non-divergent three-dimensional flow, show that the volume transport (volume per time) through the two streamtube ends balances

$$\int_{\mathcal{A}_1} \mathbf{v} \cdot \hat{\mathbf{n}} dS + \int_{\mathcal{A}_2} \mathbf{v} \cdot \hat{\mathbf{n}} dS = 0, \quad (16.44)$$

where $\hat{\mathbf{n}}$ is the outward normal at the two end caps \mathcal{A}_1 and \mathcal{A}_2 . Hence, 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.

EXERCISE 16.10: AREA AVERAGE OF FREE SURFACE TIME TENDENCY

Consider a single homogeneous (constant density) fluid layer 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 $h = H + \eta$. 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.

17

Material fluid objects[†]

We here develop kinematic equations describing the movement of material fluid lines, material areas, and material volumes. The discussion is restricted to Cartesian tensors in order to reduce the math overhead. Consequently, all tensor labels are downstairs with no distinction between covariant and contravariant. Generalizations are straightforward following the rules of general covariance detailed in Chapters 5 and 6.

17.1	Increments in material and position space	221
17.1.1	Eulerian/position space differential	222
17.1.2	Lagrangian/material space differential	222
17.1.3	Duality between Eulerian and Lagrangian perspectives	222
17.2	Evolution of a material line element	223
17.2.1	Deformation gradient tensor	223
17.2.2	Cauchy-Green strain tensor	224
17.2.3	Material evolution of a line element	224
17.2.4	Velocity gradient tensor	224
17.2.5	Stretching and tilting of a material line element	225
17.2.6	Rate of strain tensor	225
17.2.7	Rotation tensor	226
17.2.8	Comments and further reading	227
17.3	Evolution of a material area element	227
17.3.1	Material area in three-dimensional flow	227
17.3.2	Material area in two-dimensional flow	228
17.4	Material parcel volume and the Jacobian of transformation	228
17.4.1	Material parcel volume	229
17.4.2	Evolution of the Jacobian of transformation	229
17.5	Exercises	230

17.1 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.

17.1.1 Eulerian/position space differential

In Section 14.4.1, 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

$$d\Phi(\mathbf{x}, t) = \Phi(\mathbf{x} + d\mathbf{x}, t) - \Phi(\mathbf{x}, t) \quad (17.1a)$$

$$= (d\mathbf{x} \cdot \nabla)\Phi. \quad (17.1b)$$

The operator

$$d\mathbf{x} \cdot \nabla = dx_m \frac{\partial}{\partial x_m} \quad (17.2)$$

is a scalar since it remains form invariant when switching to another set of Cartesian position coordinates.¹

17.1.2 Lagrangian/material space differential

Consider the same function Φ evaluated on a material particle trajectory, and write this “Lagrangian” function as

$$\Phi^L(\mathbf{a}, t) = \Phi[\mathbf{X}(\mathbf{a}, t), t]. \quad (17.3)$$

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

$$\delta\Phi^L(\mathbf{a}, t) = \Phi[\mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t), t] - \Phi[\mathbf{X}(\mathbf{a}, t), t] \quad (17.4a)$$

$$= \Phi^L(\mathbf{a} + \delta\mathbf{a}, t) - \Phi^L(\mathbf{a}, t) \quad (17.4b)$$

$$= (\delta\mathbf{a} \cdot \nabla_{\mathbf{a}})\Phi^L(\mathbf{a}, t). \quad (17.4c)$$

The operator

$$\delta\mathbf{a} \cdot \nabla_{\mathbf{a}} = \delta a_j \frac{\partial}{\partial a_j} \quad (17.5)$$

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.

17.1.3 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

$$\Phi^L(\mathbf{a}, t) = \Phi(\mathbf{x}, t) \quad \text{if } \mathbf{X}(\mathbf{a}, t) = \mathbf{x}. \quad (17.6)$$

¹This form invariance also holds when using curvilinear coordinates if we make use of the general tensor analysis formalism of Chapter 10.

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 10.

Likewise, if the infinitesimal increment in space, $\delta\mathbf{x}$, equals to the vector increment of the two fluid particles,

$$\delta\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, t), \quad (17.7)$$

then the functional increments are identical

$$\delta\Phi^L(\mathbf{a}, t) = \delta\Phi(\mathbf{x}, t) \quad \text{if } \delta\mathbf{X}(\mathbf{a}, t) = \delta\mathbf{x}, \quad (17.8)$$

where

$$\delta\Phi(\mathbf{x}, t) = \Phi(\mathbf{x} + \delta\mathbf{x}, t) - \Phi(\mathbf{x}, t). \quad (17.9)$$

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.

17.2 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.

17.2.1 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 17.1). At time t , the vector displacement between these two particles is written

$$\delta\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, t). \quad (17.10)$$

Expanding this expression to leading order yields

$$\delta\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, t) \approx (\delta\mathbf{a} \cdot \nabla_{\mathbf{a}}) \mathbf{X}(\mathbf{a}, t), \quad (17.11)$$

where again $\nabla_{\mathbf{a}}$ is the gradient operator acting on the material coordinates. Writing this equation in component form leads to

$$\delta X_m = \delta a_j \frac{\partial X_m}{\partial a_j}. \quad (17.12)$$

As in Chapter 14, 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

$$F_{mj} \equiv \frac{\partial X_m}{\partial a_j} \quad (17.13)$$

appearing in equation (17.12) form elements of the transformation matrix linking position space to material space. We already encountered this tensor in Section 14.3.3 (see equation (14.9)). In the continuum mechanics literature, the tensor (17.13) is known as the *displacement gradient tensor* or the *deformation gradient tensor*.

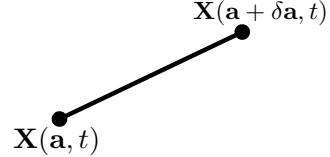


Figure 17.1: The ends of an infinitesimal material line element are defined by the trajectories of two fluid particles, $\mathbf{X}(\mathbf{a}, t)$ and $\mathbf{X}(\mathbf{a} + \delta\mathbf{a}, t)$. All points between the endpoints are part of the material line.

17.2.2 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

$$\delta \mathbf{X} \cdot \delta \mathbf{X} = \frac{\partial \mathbf{X}}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} \delta a_i \delta a_j \equiv C_{ij} \delta a_i \delta a_j. \quad (17.14)$$

Algebraically, this expression is a quadratic form, and the symmetric tensor

$$C_{ij} \equiv \frac{\partial \mathbf{X}}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} = F_{mi} F_{mj} \quad (17.15)$$

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.

17.2.3 Material evolution of a line element

Now consider the material time derivative of the material line element

$$\frac{\partial [\delta \mathbf{X}(\mathbf{a}, t)]}{\partial t} = \frac{\partial \mathbf{X}(\mathbf{a} + \delta \mathbf{a}, t)}{\partial t} - \frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \quad (17.16a)$$

$$= \mathbf{v}^L(\mathbf{a} + \delta \mathbf{a}, t) - \mathbf{v}^L(\mathbf{a}, t) \quad (17.16b)$$

$$\equiv \delta \mathbf{v}^L(\mathbf{a}, t). \quad (17.16c)$$

In these equations, we introduced the Lagrangian velocity

$$\mathbf{v}^L(\mathbf{a}, t) = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (17.17)$$

as per equation (14.31) and the discussion in Section 17.1.2.

As for the line element manipulations in Section 17.2.1, we can massage the expression (17.16c) by performing a Taylor series expansion and truncating to leading order

$$\frac{\partial [\delta \mathbf{X}(\mathbf{a}, t)]}{\partial t} = \delta \mathbf{v}^L(\mathbf{a}, t) = (\delta \mathbf{a} \cdot \nabla_{\mathbf{a}}) \mathbf{v}^L(\mathbf{a}, t). \quad (17.18)$$

Alternatively, we can choose to evaluate this expression using an Eulerian perspective (see Section 17.1.3), in which case

$$\frac{D(\delta \mathbf{x})}{Dt} = \delta \mathbf{v}(\mathbf{x}, t) = (\delta \mathbf{x} \cdot \nabla_{\mathbf{x}}) \mathbf{v}(\mathbf{x}, t). \quad (17.19)$$

17.2.4 Velocity gradient tensor

Writing the Eulerian result (17.19) in component form leads to

$$\frac{D(\delta x_m)}{Dt} = \delta x_n \frac{\partial v_m}{\partial x_n}. \quad (17.20)$$

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

$$\frac{\partial v_m}{\partial x_n} = \frac{1}{2} \left(\frac{\partial v_m}{\partial x_n} + \frac{\partial v_n}{\partial x_m} \right) + \frac{1}{2} \left(\frac{\partial v_m}{\partial x_n} - \frac{\partial v_n}{\partial x_m} \right) \quad (17.21a)$$

$$\equiv S_{mn} + A_{mn}, \quad (17.21b)$$

where³

$$S_{mn} = \frac{1}{2} \left(\frac{\partial v_m}{\partial x_n} + \frac{\partial v_n}{\partial x_m} \right) = S_{nm} \quad \text{rate of strain tensor} \quad (17.22a)$$

$$A_{mn} = \frac{1}{2} \left(\frac{\partial v_m}{\partial x_n} - \frac{\partial v_n}{\partial x_m} \right) = -A_{nm} \quad \text{rotation tensor.} \quad (17.22b)$$

As seen in the following, these tensors affect the line element evolution in very distinct manners.

17.2.5 Stretching and tilting of a material line element

Consider a line element that is initially aligned with the vertical axis

$$\delta \mathbf{x}_{t=0} = \hat{z} \delta Z_0. \quad (17.23)$$

Consequently, the initial evolution of this material line element takes on the form

$$\underbrace{\frac{D(\delta x)}{Dt}}_{\text{tilting}} = \delta Z_0 \left(\frac{\partial u}{\partial z} \right) \quad \underbrace{\frac{D(\delta y)}{Dt}}_{\text{tilting}} = \delta Z_0 \left(\frac{\partial v}{\partial z} \right) \quad \underbrace{\frac{D(\delta z)}{Dt}}_{\text{stretching}} = \delta Z_0 \left(\frac{\partial w}{\partial z} \right). \quad (17.24)$$

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 37. 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.

17.2.6 Rate of strain tensor

Recall the expression (17.14) for the squared length of a line element

$$\delta \mathbf{X} \cdot \delta \mathbf{X} = \frac{\partial \mathbf{X}}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} \delta a_i \delta a_j. \quad (17.25)$$

Its material time derivative is given by

$$\left(\frac{\partial (\delta \mathbf{X} \cdot \delta \mathbf{X})}{\partial t} \right)_a = 2 \frac{\partial \mathbf{v}^L}{\partial a_i} \cdot \frac{\partial \mathbf{X}}{\partial a_j} \delta a_i \delta a_j. \quad (17.26)$$

³The rate of strain tensor, 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 (17.13).

We can express this result using Eulerian \mathbf{x} -coordinates by making use of the duality described in Section 17.1.3, which leads to

$$\frac{\partial \mathbf{v}^L}{\partial a_i} \delta a_i = \frac{\partial \mathbf{v}}{\partial x_n} \delta x_n \quad (17.27a)$$

$$\frac{\partial \mathbf{X}}{\partial a_j} \delta a_j = \delta \mathbf{x}, \quad (17.27b)$$

so that

$$\frac{D(\delta \mathbf{x} \cdot \delta \mathbf{x})}{Dt} = 2 \frac{\partial v_m}{\partial x_n} \delta x_n \delta x_m. \quad (17.28)$$

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

$$\frac{1}{2} \frac{D(\delta \mathbf{x} \cdot \delta \mathbf{x})}{Dt} = S_{mn} \delta x_n \delta x_m. \quad (17.29)$$

Consequently, the rate of strain tensor, 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 (17.29) 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 (17.29).

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 fluid flow strains. As shown in Section 17.4, 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 field

$$S_{mm} = \nabla \cdot \mathbf{v}. \quad (17.30)$$

17.2.7 Rotation tensor

The rotation tensor

$$A_{mn} = \frac{1}{2} \left[\frac{\partial v_m}{\partial x_n} - \frac{\partial v_n}{\partial x_m} \right] \quad (17.31)$$

is anti-symmetric: $A_{mn} = -A_{nm}$. Its components are related to the vorticity vector $\boldsymbol{\omega} = \nabla \wedge \mathbf{v}$ according to

$$A_{mn} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \quad (17.32)$$

The contribution of the rotation matrix to material evolution of the line element is given by

$$\left[\frac{D(\delta x_m)}{Dt} \right]_{\text{rot}} = A_{mn} \delta x_n \Rightarrow \left[\frac{D(\delta \mathbf{x})}{Dt} \right]_{\text{rot}} = \frac{1}{2} (\boldsymbol{\omega} \wedge \delta \mathbf{x}). \quad (17.33)$$

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 10.5). We thus conclude that the anti-symmetric rotation tensor, 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).

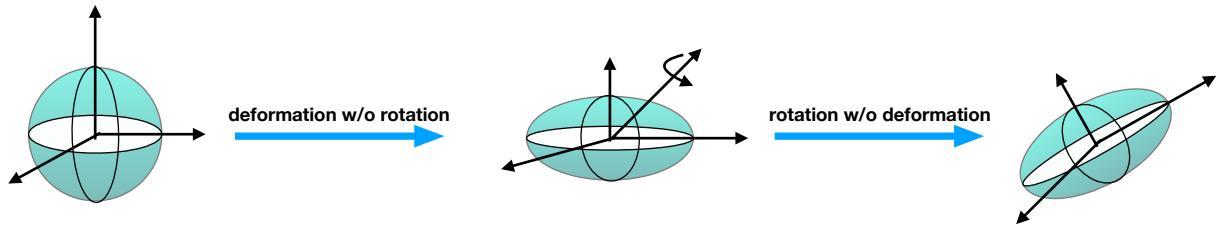


Figure 17.2: Schematic illustrating the decomposition of the manners that fluid flow can modify a material volume. First the sphere can be deformed without rotation, with this process encompassed by the rate of strain tensor, S_{mn} . Next it can be rigidly rotated without changing its shape, as encompassed by the rotation tensor, 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.

17.2.8 Comments and further reading

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 (17.20) to read

$$v_m(\mathbf{x}, t) = v_m(\mathbf{x}_0, t) + S_{mn} \delta x_n + A_{mn} \delta x_n, \quad (17.34)$$

which can be written in the dyadic form

$$\mathbf{v}(\mathbf{x}, t) = \mathbf{v}(\mathbf{x}_0, t) + \mathbf{S} \cdot \delta \mathbf{x} + \mathbf{A} \cdot \delta \mathbf{x}. \quad (17.35)$$

Figure 17.2 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. \(2012\)](#).

17.3 Evolution of a material area element

We extend the discussion of material line elements in Section 17.2 to a material area element such as that shown in Figure 17.3. We consider area elements in both three-dimensional and two-dimensional flows.

17.3.1 Material area in three-dimensional flow

Following from the geometric interpretation of the vector product in Section 2.4.4, we here define a material area element by (see Figure 17.3)

$$\delta \mathbf{S} = \delta \mathbf{A} \wedge \delta \mathbf{B} \Rightarrow \delta S_m = \epsilon_{mnp} \delta A_n \delta B_p \quad (17.36)$$

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

$$\hat{\mathbf{n}} \cdot \delta \mathbf{S} = \hat{\mathbf{n}} \cdot (\delta \mathbf{A} \wedge \delta \mathbf{B}). \quad (17.37)$$

The evolution of the material area element is given by

$$\frac{D(\delta S)}{Dt} = \frac{D(\delta A)}{Dt} \wedge \delta B + \delta A \wedge \frac{D(\delta B)}{Dt} \quad (17.38a)$$

$$= [(\delta A \cdot \nabla) v] \wedge \delta B + \delta A \wedge [(\delta B \cdot \nabla) v], \quad (17.38b)$$

where the second equality made use of the line element evolution equation (17.19). To proceed we expose indices and make use of some tensor identities

$$\frac{D(\delta S_m)}{Dt} = \epsilon_{mnp} [(\delta A_q \partial_q) v_n] \delta B_p + \epsilon_{mnp} \delta A_n [(\delta B_q \partial_q) v_p] \quad (17.39a)$$

$$= \epsilon_{mnp} [\delta A_q \delta B_p \partial_q v_n + \delta A_n \delta B_q \partial_q v_p] \quad (17.39b)$$

$$= \epsilon_{mnp} \partial_q v_n [\delta A_q \delta B_p - \delta A_p \delta B_q] \quad (17.39c)$$

$$= \epsilon_{mnp} \partial_q v_n \epsilon_{rqp} \delta S_r \quad (17.39d)$$

$$= (\delta_{mr} \delta_{nq} - \delta_{mq} \delta_{nr}) \partial_q v_n \delta S_r \quad (17.39e)$$

$$= (\nabla \cdot v) \delta S_m - (\partial_m v) \cdot \delta S. \quad (17.39f)$$

To reach this result we made use of the following identities

$$\delta A_q \delta B_p - \delta A_p \delta B_q = \epsilon_{rqp} \delta S_r \quad (17.40a)$$

$$\epsilon_{mnp} \epsilon_{rqp} = \delta_{mr} \delta_{nq} - \delta_{mq} \delta_{nr} \quad (17.40b)$$

where δ_{mn} is the Kronecker (or identity) tensor.

17.3.2 Material area in two-dimensional flow

Now consider a material area element for two-dimensional fluid flow with velocity, $v = (u, v, 0)$, and $\delta A = \hat{x} \delta x$, $\delta B = \hat{y} \delta y$, with zero dependence on z . In this case, the area of an infinitesimal material region is

$$\delta S = (\delta A \wedge \delta B) \cdot \hat{z} = \delta x \delta y, \quad (17.41)$$

and its evolution is given by

$$\frac{D(\delta S)}{Dt} = (\delta B \wedge \hat{z}) \cdot (\delta A \cdot \nabla) u + (\hat{z} \wedge \delta A) \cdot (\delta B \cdot \nabla) u \quad (17.42a)$$

$$= \delta x \delta y \nabla \cdot u, \quad (17.42b)$$

so that

$$\frac{1}{\delta S} \frac{D(\delta S)}{Dt} = \nabla \cdot u. \quad (17.43)$$

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 (17.39f) to the case of two-dimensional flow with no dependence on the vertical direction.

17.4 Material parcel 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

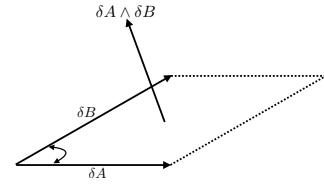


Figure 17.3: A material area defined by the cross product of two material line elements, $\delta S = \delta A \wedge \delta B$. In the special case of $\delta A = \hat{x} \delta x$ and $\delta B = \hat{y} \delta y$, then $\delta S = \delta x \delta y \hat{z}$.

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.

17.4.1 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 17.4). To within a sign the volume is given by

$$\delta V = (\delta \mathbf{A} \wedge \delta \mathbf{B}) \cdot \delta \mathbf{C}. \quad (17.44)$$

Making use of the line element evolution equation (17.19) renders

$$\frac{D(\delta V)}{Dt} = (\delta \mathbf{B} \wedge \delta \mathbf{C}) \cdot (\delta \mathbf{A} \cdot \nabla) \mathbf{v} + (\delta \mathbf{C} \wedge \delta \mathbf{A}) \cdot (\delta \mathbf{B} \cdot \nabla) \mathbf{v} + (\delta \mathbf{A} \wedge \delta \mathbf{B}) \cdot (\delta \mathbf{C} \cdot \nabla) \mathbf{v}. \quad (17.45)$$

Now specialize to the case where the parcel is a parallelepiped oriented according to the coordinate axes

$$\delta \mathbf{A} = \hat{\mathbf{x}} \delta x \quad \delta \mathbf{B} = \hat{\mathbf{y}} \delta y \quad \delta \mathbf{C} = \hat{\mathbf{z}} \delta z, \quad (17.46)$$

so that

$$\delta V = \delta x \delta y \delta z. \quad (17.47)$$

Plugging into equation (17.45) leads to

$$\frac{1}{\delta V} \frac{D(\delta V)}{Dt} = \nabla \cdot \mathbf{v}. \quad (17.48)$$

This result is a three-dimensional generalization of the material area equation derived in Section 17.3.2.

We offer an alternative derivation of equation (17.48) in Section 15.1, 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 ($\nabla \cdot \mathbf{v} < 0$)

$$\nabla \cdot \mathbf{v} > 0 \Rightarrow \text{parcel volume increases in diverging flow} \Rightarrow \text{parcel expands} \quad (17.49a)$$

$$\nabla \cdot \mathbf{v} < 0 \Rightarrow \text{parcel volume decreases in converging flow} \Rightarrow \text{parcel contracts}. \quad (17.49b)$$

17.4.2 Evolution of the Jacobian of transformation

Recall the discussion in Section 14.3.6, 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 (14.21) is given by

$$\delta V(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \delta V(\mathbf{a}) \Rightarrow \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{\delta V(\mathbf{x})}{\delta V(\mathbf{a})}. \quad (17.50)$$

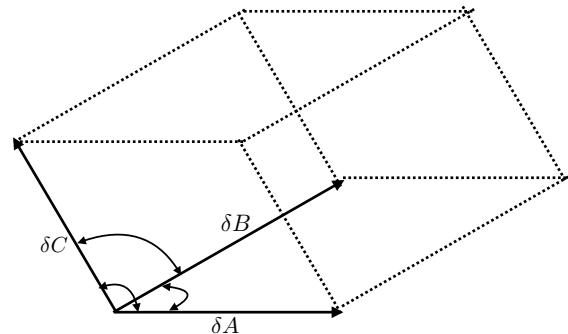


Figure 17.4: A parallelepiped defined by three material line elements, with volume (to within a sign) given by $\delta V = (\delta \mathbf{A} \wedge \delta \mathbf{B}) \cdot \delta \mathbf{C}$. See also the discussion surrounding Figure 2.4.

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

$$\frac{D}{Dt} \frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{1}{\delta V(\mathbf{a})} \frac{D(\delta V(\mathbf{x}))}{Dt} \quad (17.51a)$$

$$= \frac{\delta V(\mathbf{x})}{\delta V(\mathbf{a})} \nabla \cdot \mathbf{v} \quad (17.51b)$$

$$= \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \nabla \cdot \mathbf{v}. \quad (17.51c)$$

The second equality made use of the result (17.48), 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

$$\left(\frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right)^{-1} \frac{D}{Dt} \left(\frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right) = \nabla \cdot \mathbf{v}. \quad (17.52)$$

This equation is identical to the parcel volume equation (17.48), which is expected given the relation between the Jacobian and the parcel volume. In Exercise 17.2, we derive this result using the explicit expression for the Jacobian in terms of the ϵ -tensor.

17.5 Exercises

EXERCISE 17.1: 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

$$\mathbf{v} = \mathbf{U} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (17.53)$$

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 $S_{mn} = 0$, what does that imply about the velocity field? You may also wish to make use of the general decomposition (17.35).

EXERCISE 17.2: EVOLUTION OF THE JACOBIAN USING ϵ -TENSOR GYMNASTICS

There is another way to derive the identity (17.52) 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

$$\frac{\partial \mathbf{X}}{\partial \mathbf{a}} = \frac{1}{3!} \epsilon_{mnp} \epsilon_{ijk} \frac{\partial X_m}{\partial a_i} \frac{\partial X_n}{\partial a_j} \frac{\partial X_p}{\partial a_k}. \quad (17.54)$$

Take the material derivative of this expression and show that we get the same expression as equation (17.52). Hint: make use of the identity

$$\frac{D}{Dt} \frac{\partial X_m}{\partial a_i} = \frac{\partial V_m}{\partial a_i}, \quad (17.55)$$

which holds since the material time derivative is taken with the material coordinates, \mathbf{a} , held fixed.

18

Kinematics with generalized vertical coordinates[†]

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 15.7 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 8. 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, and to derive the GVC version of mass continuity and the tracer equation.

READER'S GUIDE TO THIS CHAPTER

We introduced mathematical properties of generalized vertical coordinates (GVCs) in Chapter 8, including the nature of calculus using these non-orthogonal coordinates. It is essential to have a working knowledge of that material in order to understand material in the present chapter on fluid kinematics. Later in Chapter 31 we detail the dynamical equations using GVCs, with material in that chapter relying on the kinematics presented here.

18.1	Example generalized vertical coordinates	232
18.1.1	Ocean free surface	232
18.1.2	Ocean bottom	232
18.1.3	Ocean mixed layer base	233
18.1.4	Interior generalized vertical coordinate surfaces	233
18.2	Specific thickness	233
18.3	The dia-surface transport	234
18.3.1	Flow normal to the GVC surface	234
18.3.2	Accounting for movement of the surface	235
18.3.3	Cross GVC transport in terms of GVC material evolution	235
18.3.4	Defining the dia-surface transport	235
18.3.5	Expressions for the dia-surface velocity component	236
18.3.6	An alternative definition of dia-surface velocity component	237
18.4	Material time derivative	238
18.5	Vertical velocity and dia-surface velocity	238
18.5.1	Concerning the vertical velocity decomposition	239
18.5.2	Another form of the vertical velocity decomposition	239
18.6	The velocity vector and fluid particle trajectories	240
18.7	Subduction across the mixed layer base	242
18.8	Mass continuity	242
18.8.1	Incompressible fluid	243
18.8.2	Mass continuity using pressure coordinates	244
18.9	Tracer equation	245

18.1 Example generalized vertical coordinates

We here consider some generalized vertical coordinates that will prove of use for our discussion in this chapter.

18.1.1 Ocean free surface

The first surface is the ocean free surface, whose kinematic boundary conditions were derived in Section 15.7.3. 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

$$\sigma(x, y, z, t) = z - \eta(x, y, t) = 0 \quad \text{ocean free surface.} \quad (18.1)$$

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.

18.1.2 Ocean bottom

We may describe the solid Earth lower boundary mathematically by using the time independent expression

$$\sigma(x, y, z) = z + H(x, y) = 0 \quad \text{ocean bottom.} \quad (18.2)$$

As detailed in Section 15.7.1, 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.

18.1.3 Ocean mixed layer base

Let

$$\sigma = z + h^{(\text{mld})}(x, y, t) = 0 \quad (18.3)$$

represent the vertical position of the ocean mixed layer base. The corresponding normal vector is given by

$$\hat{\mathbf{n}}^{(\text{mld})} = \frac{\nabla(z + h^{(\text{mld})})}{|\nabla(z + h^{(\text{mld})})|}. \quad (18.4)$$

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 18.7).

18.1.4 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.

18.2 Specific thickness

As mentioned in Section 8.9.1, 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

$$\frac{\partial z}{\partial \sigma} = z_\sigma, \quad (18.5)$$

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 the for both the ocean bottom and free surface

$$\frac{\partial z}{\partial \sigma} = 1 \quad \text{ocean free surface and bottom.} \quad (18.6)$$

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 $d\sigma$ is given by

$$dz = (\partial z / \partial \sigma) d\sigma. \quad (18.7)$$

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

$$dz = (\partial z / \partial b) db = -\frac{g \rho_o}{N^2} db, \quad (18.8)$$

with

$$N^2 = -\frac{g}{\rho_o} \frac{\partial b}{\partial z} \quad (18.9)$$

the buoyancy frequency (Section 20.3.3) in a Boussinesq fluid (Chapter 29), and ρ_o the constant reference Boussinesq density. For a hydrostatic fluid using pressure as the vertical coordinate, the thickness of a pressure layer is

$$dz = (\partial z / \partial p) dp = -\frac{1}{\rho g} dp, \quad (18.10)$$

where we used the hydrostatic relation (Section 27.2)

$$\frac{\partial p}{\partial z} = -\rho g \quad (18.11)$$

with g the acceleration due to gravity.

18.3 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.

18.3.1 Flow normal to the GVC surface

At an arbitrary point on a surface of constant generalized vertical coordinate (see Figure 18.1), the rate at which fluid moves in the direction normal to the surface is given by

$$\text{RATE OF FLUID FLOW IN DIRECTION } \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}}, \quad (18.12)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla \sigma}{|\nabla \sigma|}, \quad (18.13)$$

is the surface unit normal. For the ocean free surface, $\sigma = z - \eta$, the unit normal takes the form

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla \eta}{\sqrt{1 + |\nabla \eta|^2}} \quad \text{ocean free surface}, \quad (18.14)$$

whereas at the solid Earth bottom, $\sigma = z + H$,

$$\hat{\mathbf{n}} = -\frac{\nabla(z + H)}{|\nabla(z + H)|} = -\frac{\hat{\mathbf{z}} + \nabla H}{\sqrt{1 + |\nabla H|^2}} \quad \text{ocean bottom}. \quad (18.15)$$

Introducing the material time derivative

$$\frac{D\sigma}{Dt} = \frac{\partial \sigma}{\partial t} + \mathbf{v} \cdot \nabla \sigma \quad (18.16)$$

to equation (18.12) leads to the identity

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \frac{1}{|\nabla \sigma|} \left[\frac{D\sigma}{Dt} - \frac{\partial \sigma}{\partial t} \right]. \quad (18.17)$$

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.

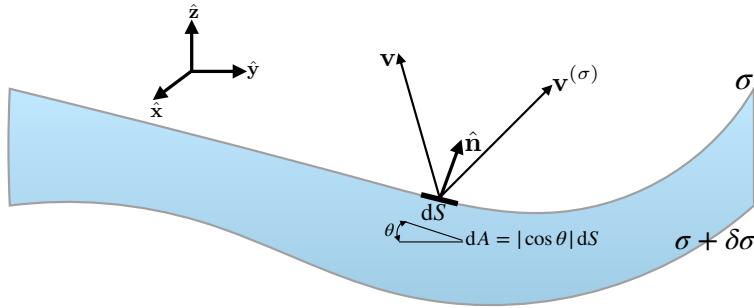


Figure 18.1: Two infinitesimally close surfaces of constant generalized vertical coordinate that form the interfaces of a σ -layer. An upward normal direction \hat{n} is indicated on one of the surfaces. Also shown is an example velocity of a fluid particle v and velocity $v^{(\sigma)}$ of a point living on the surface. Note that kinematics is only concerned with the normal component to the surface velocity, $v^{(\sigma)} \cdot \hat{n}$, as per equation (18.20). Following equation (18.25), 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.

18.3.2 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, $v^{(\sigma)}$, from the velocity of a fluid particle. We are thus led to

$$\text{RATE THAT FLUID CROSSES A MOVING GVC SURFACE} = \hat{n} \cdot (v - v^{(\sigma)}). \quad (18.18)$$

The velocity of the surface, $v^{(\sigma)}$, is the velocity of a point fixed to the σ surface, which in turn means that

$$\frac{\partial \sigma}{\partial t} + v^{(\sigma)} \cdot \nabla \sigma = 0. \quad (18.19)$$

as well as

$$v^{(\sigma)} \cdot \hat{n} = -\frac{1}{|\nabla \sigma|} \frac{\partial \sigma}{\partial t}. \quad (18.20)$$

Hence, the normal component to the velocity of the surface vanishes when the surface is static.

Note that we only need information about $v^{(\sigma)} \cdot \hat{n}$ for our kinematic formulation. Specification of the tangential component of $v^{(\sigma)}$ requires dynamical information specific to the chosen surface.

18.3.3 Cross GVC transport in terms of GVC material evolution

Using expression (18.20) in equation (18.18) leads to the net flux of fluid crossing the GVC surface

$$\hat{n} \cdot (v - v^{(\sigma)}) = \frac{1}{|\nabla \sigma|} \frac{D\sigma}{Dt}. \quad (18.21)$$

The material time derivative of the GVC surface thus vanishes if and only if no fluid crosses the surface. This is a very important result that is used throughout geophysical fluid mechanics.

18.3.4 Defining the dia-surface transport

The area normalizing the volume flux in equation (18.21) is the area dS of an infinitesimal patch on the surface of constant generalized vertical coordinate with outward unit normal \hat{n} . We now follow the trigonometry discussed in Section 8.17 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

$$\frac{dS}{|\nabla\sigma|} = \frac{dS}{\sqrt{(\partial\sigma/\partial x)^2 + (\partial\sigma/\partial y)^2 + (\partial\sigma/\partial z)^2}} \quad (18.22a)$$

$$= \frac{dS}{|\partial\sigma/\partial z| \sqrt{[(\partial\sigma/\partial x)/(\partial\sigma/\partial z)]^2 + [(\partial\sigma/\partial y)/(\partial\sigma/\partial z)]^2 + 1}} \quad (18.22b)$$

$$= \frac{dS}{|\partial\sigma/\partial z| \sqrt{1 + \tan^2 \theta}} \quad (18.22c)$$

$$= \left| \frac{\partial z}{\partial\sigma} \right| |\cos \theta| dS \quad (18.22d)$$

$$= \left| \frac{\partial z}{\partial\sigma} \right| dA. \quad (18.22e)$$

The equality (18.22c) introduced the angle, θ , between the boundary surface and the horizontal plane. The squared slope of this surface given by (see Section 8.12)

$$\tan^2 \theta = \frac{\nabla_z \sigma \cdot \nabla_z \sigma}{(\partial\sigma/\partial z)^2} = \nabla_\sigma z \cdot \nabla_\sigma z. \quad (18.23)$$

The equality (18.22d) made use of a trigonometric identity so that

$$|\cos \theta|^{-1} = |z_\sigma \nabla\sigma|. \quad (18.24)$$

Furthermore, the equality (18.22e) introduced the horizontal projection of the area,

$$dA = |\cos \theta| dS. \quad (18.25)$$

We now introduce the *dia-surface velocity component* for the GVC coordinate

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial\sigma} \frac{D\sigma}{Dt} = z_\sigma \dot{\sigma}, \quad (18.26)$$

which measures the volume of fluid passing through the surface, per unit horizontal area, per unit time

$$w^{(\dot{\sigma})} \equiv \hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \frac{dS}{dA} \quad (18.27)$$

$$= \frac{(\text{VOLUME/TIME}) \text{ FLUID THROUGH SURFACE}}{\text{HORIZONTAL AREA OF SURFACE}}. \quad (18.28)$$

This velocity component is referred to as the dia-surface velocity component as it measures flow rate of fluid through the surface.

18.3.5 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 8.12, allows us to write the dia-surface velocity component in the following

equivalent forms

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} \quad (18.29a)$$

$$= \frac{\partial z}{\partial \sigma} |\nabla \sigma| \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \quad (18.29b)$$

$$= \frac{\partial z}{\partial \sigma} \nabla \sigma \cdot \mathbf{v} - \frac{\partial z}{\partial \sigma} |\nabla \sigma| \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \quad (18.29c)$$

$$= (\hat{z} - \nabla_\sigma z) \cdot \mathbf{v} + \frac{\partial z}{\partial \sigma} \frac{\partial \sigma}{\partial t} \quad (18.29d)$$

$$= (\hat{z} - \nabla_\sigma z) \cdot \mathbf{v} - \frac{\partial z}{\partial t} \quad (18.29e)$$

$$= w - (\partial_t + \mathbf{u} \cdot \nabla_\sigma) z, \quad (18.29f)$$

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 (8.25b) and (8.25c))

$$\nabla_\sigma z = -z_\sigma \nabla_z \sigma \quad (18.30)$$

to express the slope of the σ surface as projected onto the horizontal direction plane, as well as the corresponding identity (8.25a) for the time derivative

$$\left[\frac{\partial z}{\partial t} \right]_\sigma = -\frac{[\partial \sigma / \partial t]_z}{[\partial \sigma / \partial z]} \quad (18.31)$$

The form given by equation (18.29f) directly relates the vertical component to the fluid particle velocity to the dia-surface velocity component

$$w = \frac{Dz}{Dt} \longleftrightarrow w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = w - (\partial_t + \mathbf{u} \cdot \nabla_\sigma) z. \quad (18.32)$$

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

$$w^{(\dot{\sigma})} = w - \mathbf{u} \cdot \nabla_\sigma z \quad \text{static surface}, \quad (18.33)$$

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

$$w^{(\dot{\sigma})} = w = \frac{Dz}{Dt} \quad \text{GVC surface static and flat}, \quad (18.34)$$

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 (18.29f) thus offers a useful means to distinguish w from $w^{(\dot{\sigma})}$.

18.3.6 An alternative definition of dia-surface velocity component

In some literature presentations, the dia-surface velocity component is taken to be

$$w^{\text{dia}} = \hat{n} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = \frac{1}{|\nabla \sigma|} \frac{D\sigma}{Dt}. \quad (18.35)$$

For example, [Groeskamp et al. \(2019\)](#) prefer this definition for watermass analysis. As seen in Section 25.1, the reason is that for watermass transformation analysis we can no longer assume vertically stable stratification for surfaces of constant σ .

18.4 Material time derivative

The expression (18.26) for $w^{(\dot{\sigma})}$ brings the material time derivative operator into the following equivalent forms

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (18.36a)$$

$$= \left[\frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma + \frac{D\sigma}{Dt} \frac{\partial}{\partial \sigma} \quad (18.36b)$$

$$= \left[\frac{\partial}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma + w^{(\dot{\sigma})} \frac{\partial}{\partial z}. \quad (18.36c)$$

Note that the chain-rule means that

$$\frac{\partial}{\partial \sigma} = \frac{\partial z}{\partial \sigma} \frac{\partial}{\partial z}, \quad (18.37)$$

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 (18.36c) 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 (18.36a) 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.

18.5 Vertical velocity and dia-surface velocity

Making use of the material time derivative operator (18.36c) 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

$$w = \frac{Dz}{Dt} = \left[\frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma z + w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \left[-\frac{\partial \sigma}{\partial t} - \mathbf{u} \cdot \nabla_z \sigma + \frac{D\sigma}{Dt} \right], \quad (18.38)$$

and the corresponding expressions for the dia-surface velocity component are given by

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = \frac{\partial z}{\partial \sigma} \left[\frac{\partial \sigma}{\partial t} + \mathbf{u} \cdot \nabla_z \sigma + w \frac{\partial \sigma}{\partial t} \right] = - \left[\frac{\partial z}{\partial t} \right]_\sigma - \mathbf{u} \cdot \nabla_\sigma z + w. \quad (18.39)$$

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 (18.28) 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, $w^{(\dot{\sigma})}$ leads to vertical motion of the σ surface while maintaining a fixed position for the fluid particle.

18.5.1 Concerning the vertical velocity decomposition

The expression

$$w = \left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} \quad (18.40)$$

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 $\nabla_{\sigma} z$.

18.5.2 Another form of the vertical velocity decomposition

Consider the velocity for the surface itself, $\mathbf{v}^{(\sigma)}$, which is defined by

$$\frac{\partial \sigma}{\partial t} + \mathbf{v}^{(\sigma)} \cdot \nabla \sigma = 0. \quad (18.41)$$

Making use of the triple product identities from Section 8.5

$$\frac{\partial z}{\partial \sigma} \nabla \sigma = -\nabla_{\sigma} z + \hat{z} \quad \text{and} \quad \frac{\partial z}{\partial \sigma} \left[\frac{\partial \sigma}{\partial t} \right]_z = - \left[\frac{\partial z}{\partial t} \right]_{\sigma} \quad (18.42)$$

brings equation (18.41) into the form

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} = (\hat{z} - \nabla_{\sigma} z) \cdot \mathbf{v}^{(\sigma)} \implies \hat{z} \cdot \mathbf{v}^{(\sigma)} = \left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u}^{(\sigma)} \cdot \nabla_{\sigma} z, \quad (18.43)$$

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 pass that point

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} = -\mathbf{u}^{(\sigma)} \cdot \nabla_{\sigma} z \quad \text{if } \hat{z} \cdot \mathbf{v}^{(\sigma)} = 0. \quad (18.44)$$

Returning to the general result (18.43) allows us to write

$$\left[\frac{\partial z}{\partial t} \right]_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z = \hat{z} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z. \quad (18.45)$$

Furthermore, return to the fundamental definition of the dia-surface velocity component detailed in Section 18.3, in which we showed that

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt} = \frac{\partial z}{\partial \sigma} \nabla \sigma \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) = (-\nabla_{\sigma} z + \hat{z}) \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}). \quad (18.46)$$

This expression, along with equation (18.45), leads to the rather elaborate decomposition of the vertical velocity component

$$w = \underbrace{\left[\hat{z} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_{\sigma}) z} + \underbrace{\left[\hat{z} \cdot \mathbf{v} - \hat{z} \cdot \mathbf{v}^{(\sigma)} - (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_{\sigma} z \right]}_{w^{(\dot{\sigma})}}. \quad (18.47)$$

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 (18.47). 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

$$w = \underbrace{\left[\hat{z} \cdot \mathbf{v}^{(\sigma)} \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_\sigma)z} + \underbrace{\left[\hat{z} \cdot (\mathbf{v} - \mathbf{v}^{(\sigma)}) \right]}_{w^{(\dot{\sigma})}}. \quad (18.48)$$

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

$$w = 0 \quad (18.49a)$$

$$= \left[\frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma z + w^{(\dot{\sigma})} \quad (18.49b)$$

$$= \underbrace{\left[\hat{z} \cdot \mathbf{v}^{(\sigma)} + (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_\sigma z \right]}_{(\partial_t + \mathbf{u} \cdot \nabla_\sigma)z} + \underbrace{\left[-\hat{z} \cdot \mathbf{v}^{(\sigma)} - (\mathbf{u} - \mathbf{u}^{(\sigma)}) \cdot \nabla_\sigma z \right]}_{w^{(\dot{\sigma})}}. \quad (18.49c)$$

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

$$-w^{(\dot{\sigma})} = \left[\frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma z \quad \text{if } w = 0. \quad (18.50)$$

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.

18.6 The velocity vector and fluid particle trajectories

Recall from Section 18.5 the alternative forms for the vertical velocity component given by equation (18.38). We focus on the form

$$w = \left[\frac{\partial z}{\partial t} \right]_\sigma + \mathbf{u} \cdot \nabla_\sigma z + w^{(\dot{\sigma})} \quad (18.51)$$

so that the velocity vector is written¹

$$\mathbf{v} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}} \quad (18.52a)$$

$$= u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + \left[(\partial z / \partial t)_{\sigma} + \mathbf{u} \cdot \nabla_{\sigma} z + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}} \quad (18.52b)$$

$$= u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_{\sigma}] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_{\sigma}] + \left[(\partial z / \partial t)_{\sigma} + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}}. \quad (18.52c)$$

To help further understand these velocity expressions we consider the following three cases, each of which are illustrated in Figure 18.2.

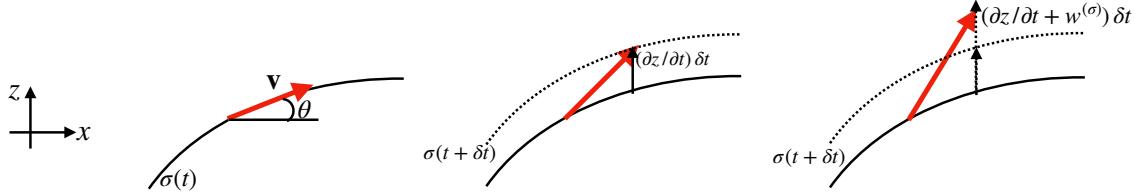


Figure 18.2: This schematic shows the various contributions to the fluid particle velocity (red vector) when written relative to motion of a particular generalized vertical coordinate surface. The fluid particle sits at the tail of the velocity vector at time t and at the head at time $t + \delta t$. 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 (18.54). The slope of the σ -surface in the $\hat{\mathbf{x}}$ -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 (18.55), 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 (18.56). 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).

- 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

$$w \partial \sigma / \partial z = -\mathbf{u} \cdot \nabla_z \sigma \implies w = \mathbf{u} \cdot \nabla_{\sigma} z, \quad (18.53)$$

where we used the triple product identities (8.25b) and (8.25c) for the final equality. The velocity vector thus takes on the form

$$\mathbf{v} = u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_{\sigma}] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_{\sigma}] \quad \partial \sigma / \partial t = 0 \text{ and } D\sigma / Dt = 0. \quad (18.54)$$

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

$$\mathbf{v} = u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_{\sigma}] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_{\sigma}] + (\partial z / \partial t)_{\sigma} \hat{\mathbf{z}} \quad D\sigma / Dt = 0. \quad (18.55)$$

To remain on the moving surface, the fluid particle must move vertically by the extra amount $(\partial z / \partial t)_{\sigma} \delta t \hat{\mathbf{z}}$ relative to the case of a static σ -surface.

¹ As discussed in Section 8.7, we can connect these expressions to the contravariant representation of the velocity vector using GVCs.

- 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, $w^{(\dot{\sigma})} = (\partial z / \partial \sigma) D\sigma / Dt$:

$$\mathbf{v} = u [\hat{\mathbf{x}} + \hat{\mathbf{z}} (\partial z / \partial x)_\sigma] + v [\hat{\mathbf{y}} + \hat{\mathbf{z}} (\partial z / \partial y)_\sigma] + \left[(\partial z / \partial t)_\sigma + w^{(\dot{\sigma})} \right] \hat{\mathbf{z}}. \quad (18.56)$$

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 18.5, 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 18.5.2 with a static particle and moving σ -surface.

18.7 Subduction across the mixed layer base

Consider the GVC (18.3) defined according to the mixed layer base. The dia-surface mass transport across this surface leads us to define the subduction

$$-\mathcal{S}^{(\text{subduction})} \equiv \rho dA \left(\frac{d(z + h^{(\text{mld})})}{dt} \right) \quad \text{at } z = -h^{(\text{mld})}(x, y, t), \quad (18.57)$$

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

$$-\left(\frac{\mathcal{S}^{(\text{subduction})}}{\rho dA} \right) = w + (\partial_t + \mathbf{u} \cdot \nabla) h^{(\text{mld})} \quad \text{at } z = -h^{(\text{mld})}(x, y, t), \quad (18.58)$$

where again we define

$$\mathcal{S}^{(\text{subduction})} > 0 \quad \text{subduction} \quad (18.59)$$

$$\mathcal{S}^{(\text{subduction})} < 0 \quad \text{obduction.} \quad (18.60)$$

This definition of subduction corresponds to that given by [Cushmin-Roisin \(1987\)](#).

18.8 Mass continuity

We here derive the Eulerian equation for mass continuity (15.9) using generalized vertical coordinates. Mass conservation for a fluid parcel states that

$$\rho \delta V = \rho \delta x \delta y \delta z = \rho \delta x \delta y z_\sigma \delta \sigma \quad (18.61)$$

is materially constant, which then leads to the following statement of mass conservation

$$\frac{D}{Dt} (\ln \rho \delta V) = \frac{D(z_\sigma)/Dt}{z_\sigma} + \frac{D \ln \rho}{Dt} + \nabla_\sigma \cdot \mathbf{u} + \partial_\sigma \left(\frac{D\sigma}{Dt} \right). \quad (18.62)$$

Now use the material time derivative in the form (18.36b) to derive the Eulerian expression of mass conservation

$$\frac{\partial(\rho z_\sigma)}{\partial t} + \nabla_s \cdot (\rho z_\sigma \mathbf{u}) + \partial_\sigma (\rho z_\sigma \dot{\sigma}) = 0, \quad (18.63)$$

where we introduced the shorthand

$$\dot{\sigma} = \frac{D\sigma}{Dt}. \quad (18.64)$$

We can furthermore introduce the dia-surface velocity component $w^{(\dot{\sigma})} = z_\sigma D\sigma/Dt$ so that mass continuity takes the form

$$\frac{\partial(\rho z_\sigma)}{\partial t} + \nabla_\sigma \cdot (\rho z_\sigma \mathbf{u}) + \partial_\sigma (\rho w^{(\dot{\sigma})}) = 0. \quad (18.65)$$

Finally, we multiply by the infinitesimal increment $d\sigma$, noting that it commutes with the time and ∇_σ operators, to render the *thickness* equation for a compressible fluid

$$\frac{\partial(\rho h)}{\partial t} + \nabla_\sigma \cdot (\rho h \mathbf{u}) + \delta_\sigma (\rho w^{(\dot{\sigma})}) = 0, \quad (18.66)$$

where

$$h \equiv \frac{\partial z}{\partial \sigma} d\sigma \quad (18.67)$$

is the infinitesimal thickness of a σ -layer, and

$$\delta_\sigma = d\sigma \frac{\partial}{\partial \sigma} \quad (18.68)$$

is the non-dimensional differential σ -operator. We illustrate contributions to the layer mass budget (18.66) in Figure 18.3.

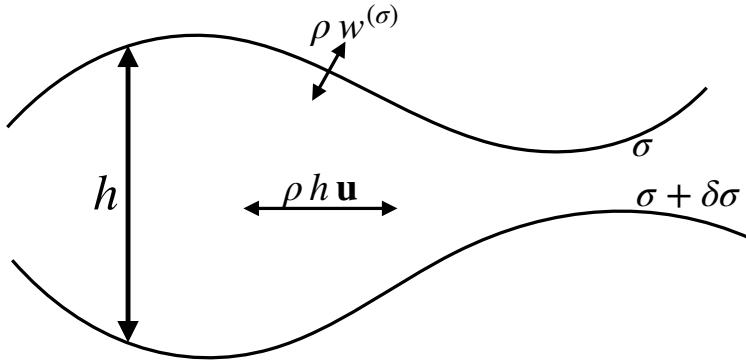


Figure 18.3: Illustrating the terms contributing to changes in layer mass according to the thickness equation (18.66). The layer is shown here with bounding interfaces at σ and $\sigma + \delta\sigma$. Within a layer there is a redistribution due to horizontal advective, whereas matter can cross the layer due to dia-surface transport via $w^{(\dot{\sigma})}$.

18.8.1 Incompressible fluid

Specializing to an incompressible volume conserving fluid (see Chapter 29) yields the incompressible thickness equation

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) + \delta_\sigma w^{(\dot{\sigma})} = 0. \quad (18.69)$$

Further specializing to a fluid without mixing or other irreversible processes means that isopycnal surfaces are material. Choosing σ to be an isopycnal vertical coordinate (see Chapter 31) means that $w^{(\dot{\sigma})} = 0$ so that the thickness equation reduces to

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) = 0 \quad \text{incompressible perfect fluid.} \quad (18.70)$$

18.8.2 Mass continuity using pressure coordinates

We here derive, using two methods, mass continuity for the case of pressure as the vertical coordinate in a hydrostatic fluid.

Method I

The thickness of a hydrostatic pressure layer (equation (18.67)) takes on the following form

$$h = \frac{\partial z}{\partial p} dp = -\frac{dp}{\rho g}, \quad (18.71)$$

in which case the mass continuity equation (18.66) becomes

$$\frac{\partial(dp)}{\partial t} + \nabla_p \cdot (\mathbf{u} dp) + \delta_p (\dot{p}) = 0. \quad (18.72)$$

The partial time derivative vanishes since it is computed by holding pressure fixed so that the pressure increment has a zero time tendency

$$\left[\frac{\partial(dp)}{\partial t} \right]_p = 0. \quad (18.73)$$

Likewise, $\nabla_p (dp) = 0$. Thus, we can divide by dp to render the continuity equation

$$\nabla_p \cdot \mathbf{u} + \frac{\partial \dot{p}}{\partial p} = 0 \quad \text{compressible hydrostatic.} \quad (18.74)$$

This equation is isomorphic to the incompressible continuity equation (for both hydrostatic and non-hydrostatic fluids) written using geopotential coordinates (see Chapter 29)

$$\nabla_z \cdot \mathbf{u} + \frac{\partial \dot{z}}{\partial z} = 0 \quad \text{incompressible,} \quad (18.75)$$

where $w = \dot{z}$ 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, return to the mass conservation discussion in Section 15.1.1, in which we showed that material conservation of mass for a fluid element leads to

$$\frac{D(\rho \delta V)}{Dt} = 0, \quad (18.76)$$

where the volume of the fluid element takes the form

$$\delta V = \delta x \delta y \delta z = \delta x \delta y \left[\frac{\partial z}{\partial p} \right] \delta p = -(\rho g)^{-1} \delta x \delta y \delta p. \quad (18.77)$$

Consequently,

$$0 = \frac{D(\rho \delta V)}{Dt} = g^{-1} \left(\frac{D(\delta x \delta y \delta p)}{Dt} \right), \quad (18.78)$$

so that

$$0 = \frac{1}{\delta x \delta y \delta p} \left(\frac{D(\delta x \delta y \delta p)}{Dt} \right) = \nabla_p \cdot \mathbf{u} + \frac{\partial \dot{p}}{\partial p}. \quad (18.79)$$

The second step made use of the isomorphism between this result and that for equation (17.48) that holds for a geopotential vertical coordinate. This step may require some head-scratching to be convinced of its validity. It is well worth doing so, as it offers an elegant way to derive the evolution of volume using an arbitrary vertical coordinate.

18.9 Tracer equation

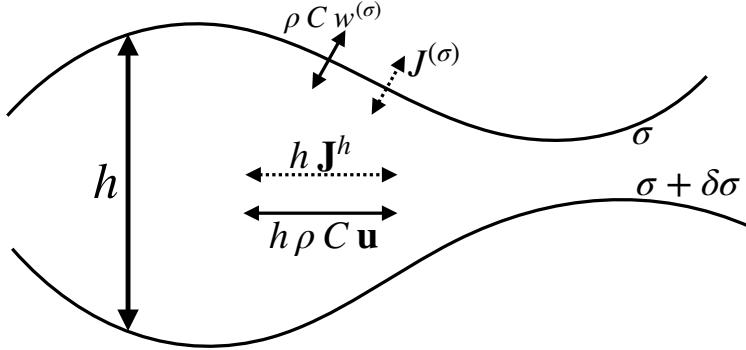


Figure 18.4: Illustrating the terms contributing to changes in layer tracer content according to the tracer equation (18.84). The layer is shown here with bounding interfaces at σ and $\sigma + \delta\sigma$. Within a layer there is a redistribution due to horizontal advective and subgrid scale tracer fluxes, whereas matter can cross the layer due to dia-surface transport via $\rho C w^{(\sigma)}$ and subgrid tracer transport $J^{(\sigma)}$.

The tracer equation from Section 15.4.4 is given by

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}, \quad (18.80)$$

where \mathbf{J} is a subgrid scale flux. Now introduce the material time derivative operator in the form (18.36b) and multiply by the layer thickness to have

$$\rho h \left[\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla_\sigma C + \dot{\sigma} \partial_\sigma C \right] = -h \nabla \cdot \mathbf{J}, \quad (18.81)$$

Making use of the mass conservation equation (18.66) renders

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_\sigma \cdot (h \rho C \mathbf{u}) + \delta_\sigma (\rho C w^{(\dot{\sigma})}) = -h \nabla \cdot \mathbf{J}, \quad (18.82)$$

where again $w^{(\dot{\sigma})} = z_\sigma \dot{\sigma}$ is the dia-surface velocity component. Making use of the expression (8.83) for the subgrid scale operator leads to

$$\frac{\partial(h\rho C)}{\partial t} + \nabla_\sigma \cdot (h\rho C \mathbf{u}) + \delta_\sigma(\rho C w^{(\dot{\sigma})}) = - \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \delta_\sigma(z_\sigma \nabla \sigma \cdot \mathbf{J}) \right]. \quad (18.83)$$

Alternatively, bringing all terms to the left hand side renders

$$\frac{\partial(h\rho C)}{\partial t} + \nabla_\sigma \cdot (h\rho C \mathbf{u} + h \mathbf{J}^h) + \delta_\sigma(\rho C w^{(\dot{\sigma})} + J^{(\sigma)}) = 0 \quad (18.84)$$

where we wrote

$$J^{(\sigma)} = z_\sigma \nabla \sigma \cdot \mathbf{J}. \quad (18.85)$$

We illustrate contributions to the layer tracer budget (18.84) in Figure 18.4.

Part IV

Thermodynamics and buoyancy

In this part of the book we develop thermodynamics and buoyancy. Our treatment of thermodynamics in Chapter 19 is reasonably self-contained, though lacking any discussion of moist thermodynamics appropriate for a realistic atmosphere (see [Vallis \(2017\)](#)). A fluid placed in a gravitational field naturally tends to stratify, with heavier water below lighter water. Such buoyancy stratification is one of the defining features of geophysical fluids (along with rotation). In Chapter 20 we explore various means to measure buoyancy stratification, such as through the buoyancy frequency and neutral directions.

19

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 thermodynamical equilibrium*. This assumption is based on the much shorter equilibration time required for molecular processes as compared to the relatively slow macroscopic processes of interest for fluid mechanics dynamics such as advection, waves, and mixing (e.g., see our discussion of kinetic theory in Section 13.2). 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.

Schematics and clarifications needed

- potential temp versus in situ for the ocean and their evolution
- adiabatic lapse rate for the atmosphere
- potential enthalpy

19.1	Exact and inexact differentials	251
19.1.1	Exact differentials	251
19.1.2	Inexact differentials	251
19.1.3	Integrating factors	252
19.1.4	An example using the velocity field	252
19.1.5	Heuristic physics of exact and inexact differential operations	253
19.2	The First Law of thermodynamics	253
19.2.1	The First Law in its extensive form	253
19.2.2	Mechanical work, heat, and chemical work	254
19.2.3	Fundamental thermodynamic relation	255

19.2.4	A note on partial derivatives	255
19.2.5	Internal energy, homogeneous functions and Gibbs-Duhem	256
19.2.6	Fundamental thermodynamic relation per unit mass	256
19.2.7	The special case of seawater as a binary fluid	257
19.3	Thermodynamic potentials	257
19.3.1	Internal energy	257
19.3.2	Entropy	258
19.3.3	Enthalpy	258
19.3.4	Gibbs potential	259
19.3.5	Combined fundamental thermodynamic relation	259
19.3.6	Alternate functional dependencies	259
19.3.7	Further reading	260
19.4	Specific heat capacity	260
19.5	A simple ideal gas atmosphere	261
19.5.1	Equation of state	261
19.5.2	Internal energy	261
19.5.3	Heat capacity from statistical mechanics	262
19.5.4	Enthalpy	263
19.5.5	Thermal expansion coefficient	263
19.5.6	Fundamental thermodynamic relations	263
19.5.7	Further reading	264
19.6	Adiabatic lapse rate	264
19.6.1	Isentropic rearrangement	264
19.6.2	Thermodynamic formulation	264
19.6.3	Adiabatic lapse rate for pressure changes	265
19.6.4	Adiabatic lapse rate for height changes	265
19.6.5	Adiabatic lapse rate for an ideal gas atmosphere	266
19.6.6	Further reading	266
19.7	Thermodynamics of a moving fluid	267
19.7.1	First Law for a moving fluid element	267
19.7.2	Further reading	269
19.8	Potential temperature	269
19.8.1	Adiabatic temperature changes	269
19.8.2	Defining the potential temperature	269
19.8.3	Potential temperature and specific entropy	270
19.8.4	The First Law in terms of potential temperature	271
19.8.5	Potential temperature for an ideal gas	271
19.8.6	Potential temperature and numerical models	272
19.8.7	Further reading	272
19.9	Exercises	272

READER'S GUIDE FOR THIS CHAPTER

Our general aim is to expose the basics of equilibrium and quasi-equilibrium thermodynamics without extending too deeply into details. Alas, this goal is elusive since a deductive approach involves a plethora of concepts and manipulations. Although aiming to be 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 thoroughly formulated for

fluid mechanics by [DeGroot and Mazur \(1984\)](#) and [Landau and Lifshitz \(1987\)](#).

Material in this chapter will be sprinkled throughout this book in a somewhat sporadic manner. Hence, those with a suitable background in thermodynamics might consider reading this chapter to be optional. Even so, the reader is encouraged to read this chapter if only to be reminded of concepts that will be assumed in other parts of this book.

19.1 Exact and inexact differentials

Thermodynamics makes use of both exact and inexact differentials. We here introduce the mathematics of such differentials prior to entering into the basic physical ideas.

19.1.1 Exact differentials

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

$$d\Phi(\mathbf{x}) = \Phi(\mathbf{x} + d\mathbf{x}) - \Phi(\mathbf{x}) \quad (19.1a)$$

$$= d\mathbf{x} \cdot \nabla \Phi, \quad (19.1b)$$

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

$$\Phi(\mathbf{x}_B) - \Phi(\mathbf{x}_A) = \int_{\mathbf{x}_A}^{\mathbf{x}_B} d\Phi(\mathbf{x}) = \int_{\mathbf{x}_A}^{\mathbf{x}_B} d\mathbf{x} \cdot \nabla \Phi. \quad (19.2)$$

These results are familiar from elementary calculus, with the increment $d\Phi$ given by equation (19.1b) 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

$$\oint d\Phi = 0. \quad (19.3)$$

19.1.2 Inexact differentials

Consider now a general differential expression written as

$$\mathbf{A} \cdot d\mathbf{x} = A dx + B dy + C dz, \quad (19.4)$$

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

$$\nabla \wedge \mathbf{A} = 0 \implies \mathbf{A} = \nabla \Psi, \quad (19.5)$$

in which case we have an exact differential expression

$$\mathbf{A} \cdot d\mathbf{x} = \nabla \Psi \cdot d\mathbf{x} = d\Psi. \quad (19.6)$$

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

$$\nabla \wedge \mathbf{A} = 0 \implies \mathbf{A} \cdot d\mathbf{x} \text{ exact differential.} \quad (19.7)$$

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

$$d\Psi = \mathbf{A} \cdot d\mathbf{x}. \quad (19.8)$$

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

$$\oint d\Psi \neq 0. \quad (19.9)$$

19.1.3 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

$$\nabla \wedge (\mathbf{A} \tau^{-1}) = 0. \quad (19.10)$$

Consequently, we can write

$$\mathbf{A} = \tau \nabla \Phi, \quad (19.11)$$

so that

$$d\Psi = \mathbf{A} \cdot d\mathbf{x} \quad (19.12a)$$

$$= \tau \nabla \Phi \cdot d\mathbf{x} \quad (19.12b)$$

$$= \tau d\Phi. \quad (19.12c)$$

The function τ is known as an *integrating factor*. As seen in Section 19.2.2, 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.

19.1.4 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 37.1) $\boldsymbol{\omega} = \nabla \wedge \mathbf{v}$. 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 circulation¹ around an arbitrary closed loop vanishes (Section 37.2)

$$C \equiv \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0. \quad (19.13)$$

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.

¹The 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 16.3.3.

19.1.5 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.

19.2 The First Law of thermodynamics

As discussed in Chapter 13, 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.

19.2.1 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

$$d\mathcal{I}^e = dQ + dW + dC \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}. \quad (19.14)$$

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 (19.14) 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 19.2.5). As discussed in Section 15.6.1, 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.

19.2.2 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. Quasi-static refers to an idealized situation whereby a thermodynamic system moves from one state to another via an infinite number of intermediate equilibrium states.

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

$$dW = -p dV \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}, \quad (19.15)$$

where p is the pressure applied to the boundaries of the fluid element. The negative sign in the pressure-work relation (19.15) 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 16), 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

$$dQ = T dS^e \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}. \quad (19.16)$$

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

$$dC = \sum_n \mu_n dM_n \quad \text{SI units Joule} = \text{kg m}^2 \text{ s}^{-2}, \quad (19.17)$$

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.

19.2.3 Fundamental thermodynamic relation

Assuming quasi-static processes, substitution of relations (19.15), (19.16), and (19.17) into the First Law (19.14) leads to

$$d\mathcal{J}^e = T dS^e - p dV + \sum_n \mu_n dM_n. \quad (19.18)$$

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 (19.18) 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 (19.18) 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

$$\left[\frac{\partial \mathcal{J}^e}{\partial S^e} \right]_{V, M_n} = T \quad (19.19)$$

$$\left[\frac{\partial \mathcal{J}^e}{\partial V} \right]_{S^e, M_n} = -p \quad (19.20)$$

$$\left[\frac{\partial \mathcal{J}^e}{\partial M_n} \right]_{S^e, V} = \mu_n. \quad (19.21)$$

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 20.2).

19.2.4 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 (19.19)-(19.21) 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.

19.2.5 Internal energy, homogeneous functions and Gibbs-Duhem

The fundamental thermodynamic relation (19.18) indicates that internal energy is naturally considered a function of entropy, volume, and constituent mass

$$\mathcal{J}^e = \mathcal{J}^e(\mathcal{S}^e, V, M_n). \quad (19.22)$$

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 (19.18), the internal energy scales likewise, giving

$$\mathcal{J}^e(\lambda \mathcal{S}^e, \lambda V, \lambda M_n) = \lambda \mathcal{J}^e(\mathcal{S}^e, V, M_n). \quad (19.23)$$

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 (19.19)–(19.21) yields

$$\mathcal{J}^e = T \mathcal{S}^e - p V + \sum_n \mu_n M_n. \quad (19.24)$$

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 (19.18) leads to the *Gibbs-Duhem* relation

$$\mathcal{S}^e dT - V dp + \sum_n M_n d\mu_n = 0. \quad (19.25)$$

19.2.6 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

$$\mathcal{J}^e = M \mathcal{J} \quad (19.26)$$

$$\mathcal{S}^e = M \mathcal{S} \quad (19.27)$$

$$V^e = M \alpha \quad (19.28)$$

$$M_n = M C_n, \quad (19.29)$$

where

$$\alpha = \rho^{-1} \quad (19.30)$$

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 15.4. Substituting the specific quantities (19.26)–(19.29) into the first Gibbs relation (19.18) and using expression (19.24) for the internal energy leads to the fundamental thermodynamic relation in terms of specific thermodynamic quantities

$$d\mathcal{J} = T d\mathcal{S} - p d\alpha + \sum_n \mu_n dC_n. \quad (19.31)$$

We make use of this form of the fundamental thermodynamic relation in the following.

19.2.7 The special case of seawater as a binary fluid

Considering seawater to be a binary system of salt and fresh water, we have

$$C_{\text{salt}} + C_{\text{water}} = 1. \quad (19.32)$$

Introducing this constraint into the first Gibbs relation (19.31) leads to

$$d\mathcal{J} = T dS - p d\alpha + \mu dC, \quad (19.33)$$

where $C = C_{\text{salt}}$ is the concentration of salt, and

$$\mu = \mu_{\text{salt}} - \mu_{\text{water}} \quad (19.34)$$

is the relative chemical potential. The *absolute salinity* S , with units parts per thousand, is related to C_{salt} via

$$S = 1000 C_{\text{salt}}. \quad (19.35)$$

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.

19.3 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.

19.3.1 Internal energy

Recall the fundamental thermodynamic relation (19.33) written for a binary fluid

$$d\mathcal{J} = T dS - p d\alpha + \mu dC. \quad (19.36)$$

Equation (19.36) identifies the specific internal energy, \mathcal{J} , as a function of specific entropy, S , specific volume, α , and matter concentration, C

$$\mathcal{J} = \mathcal{J}(S, \alpha, C). \quad (19.37)$$

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 20.2).

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

$$\left[\frac{\partial \mathcal{I}}{\partial S} \right]_{\alpha, C} = T \quad (19.38)$$

$$\left[\frac{\partial \mathcal{I}}{\partial \alpha} \right]_{S, C} = -p \quad (19.39)$$

$$\left[\frac{\partial \mathcal{I}}{\partial C} \right]_{S, \alpha} = \mu. \quad (19.40)$$

These equations are the intensive form of the extensive equations of state (19.19)-(19.21)

19.3.2 Entropy

Rearrangement of the fundamental thermodynamic relation (19.36) leads to the exact differential for specific entropy

$$dS = \frac{1}{T} d\mathcal{I} + \frac{p}{T} d\alpha - \frac{\mu}{T} dC. \quad (19.41)$$

In this form, specific entropy has the functional dependence

$$S = S(\mathcal{I}, \alpha, C), \quad (19.42)$$

whose knowledge provides yet another form of the fundamental equation of state. This functional dependence, along with equation (19.41), lead to the following thermodynamic equations of state

$$\left[\frac{\partial S}{\partial \mathcal{I}} \right]_{\alpha, C} = \frac{1}{T} \quad (19.43)$$

$$\left[\frac{\partial S}{\partial \alpha} \right]_{\mathcal{I}, C} = \frac{p}{T} \quad (19.44)$$

$$\left[\frac{\partial S}{\partial C} \right]_{\mathcal{I}, \alpha} = -\frac{\mu}{T}. \quad (19.45)$$

19.3.3 Enthalpy

Thus far we have worked only with the fundamental thermodynamic relation (19.36). We now introduce the specific enthalpy

$$\mathcal{H} = \mathcal{I} + p\alpha. \quad (19.46)$$

Use of equation (19.36) leads to the exact differential for enthalpy

$$d\mathcal{H} = d\mathcal{I} + d(p\alpha) \quad (19.47a)$$

$$= T dS - p d\alpha + \mu dC + p d\alpha + \alpha dp \quad (19.47b)$$

$$= T dS + \alpha dp + \mu dC. \quad (19.47c)$$

This equation provides the fundamental thermodynamic relation with enthalpy rather than internal energy. Consequently, the *Legendre transformation* (19.46) renders a functional dependence for enthalpy

$$\mathcal{H} = \mathcal{H}(S, p, C), \quad (19.48)$$

which is yet another fundamental equation of state. This functional dependence is more convenient than that for internal energy, $\mathcal{I}(\mathcal{S}, \alpha, C)$, or for entropy $\mathcal{S}(\mathcal{I}, \alpha, C)$. 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 20.2. 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 19.8, potential temperature is related to specific entropy, making the functional dependence equivalently written as $\mathcal{H}(\theta, p, C)$.

19.3.4 Gibbs potential

The Gibbs potential is defined by the Legendre transformation

$$\mathcal{G} = \mathcal{I} + p\alpha - T\mathcal{S} = \mathcal{H} - T\mathcal{S}. \quad (19.49)$$

Its exact differential is given by

$$d\mathcal{G} = d\mathcal{H} - d(T\mathcal{S}) \quad (19.50a)$$

$$= Td\mathcal{S} + \alpha dp + \mu dC - Td\mathcal{S} - \mathcal{S}dT \quad (19.50b)$$

$$= -\mathcal{S}dT + \alpha dp + \mu dC, \quad (19.50c)$$

where we made use of the fundamental thermodynamic relation (19.47c) written in terms of enthalpy. The Gibbs potential has the functional dependence

$$\mathcal{G} = \mathcal{G}(C, T, p). \quad (19.51)$$

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.

19.3.5 Combined fundamental thermodynamic relation

Bringing together the various forms for the fundamental thermodynamic relations renders

$$d\mathcal{I} + \alpha dp = Td\mathcal{S} + \mu dC = d\mathcal{H} - \alpha dp. \quad (19.52)$$

This relation is useful for fluid mechanics applications.

19.3.6 Alternate functional dependencies

The natural functional dependence for entropy is given by its fundamental equation of state (19.42)

$$\mathcal{S} = \mathcal{S}(\mathcal{I}, \alpha, C). \quad (19.53)$$

Likewise, the internal energy has a corresponding fundamental equation of state (19.37)

$$\mathcal{I} = \mathcal{I}(\mathcal{S}, \alpha, C). \quad (19.54)$$

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

$$\mathcal{S} = \mathcal{S}(C, T, p). \quad (19.55)$$

We do not need to include specific volume, since it is determined through the thermal equation of state discussed in Section 20.2, which yields a relation for the density as a function of temperature, pressure, and concentration (salinity in the ocean, humidity in the atmosphere)

$$\alpha^{-1} = \rho = \rho(C, T, p). \quad (19.56)$$

We make use of the functional dependence (19.55) in Section 19.6 when discussing the lapse rate.

19.3.7 Further reading

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. \(2010\)](#).

19.4 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

$$c_v \equiv \frac{1}{M} \left[\frac{d\Omega}{dT} \right]_{\alpha, C} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1}. \quad (19.57)$$

$$c_p \equiv \frac{1}{M} \left[\frac{d\Omega}{dT} \right]_{p, C} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1}. \quad (19.58)$$

If heating occurs quasi-statically, we can make use of the equation (19.16) relating heat and entropy, applied here in its specific form $M^{-1} d\Omega = T d\mathcal{S}$. The result is a state-function form of the specific heat capacities

$$c_v = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{\alpha, C} \quad (19.59)$$

$$c_p = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{p, C}. \quad (19.60)$$

We can furthermore make use of the fundamental thermodynamic relation (19.31) with specific volume and matter concentration held fixed to write

$$c_v = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{\alpha} = \left[\frac{\partial \mathcal{I}}{\partial T} \right]_{\alpha, C}. \quad (19.61)$$

Likewise, making use of the fundamental thermodynamic relation (19.47c) written in terms of enthalpy leads to the constant pressure heat capacity

$$c_p = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{p, C} = \left[\frac{\partial \mathcal{H}}{\partial T} \right]_{p, C}. \quad (19.62)$$

19.5 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.

19.5.1 Equation of state

An ideal gas satisfies the following equation (see Section 13.2.2)

$$PV = nRT, \quad (19.63)$$

where p is the pressure, V is the volume, n is the number of moles,

$$R = 8.314 \text{ J mole}^{-1} \text{ K}^{-1} = 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1} \quad (19.64)$$

is the gas constant, and T is the absolute temperature (see Section 13.2.2). The number of moles equals to the mass, M , of the gas divided by the mass per mole, M_{mole}

$$n = \frac{M}{M_{\text{mole}}}. \quad (19.65)$$

The mass density, $\rho = M/V$, is thus given by

$$\rho = \frac{p M_{\text{mole}}}{T R} \equiv \frac{p}{T R^M}, \quad (19.66)$$

where

$$R^M = \frac{R}{M_{\text{mole}}} \quad \text{SI units m}^2 \text{ s}^{-2} \text{ K}^{-1} \quad (19.67)$$

is the gas constant normalized by the molar mass for the constituent. The relation (19.66) is known as a *thermal equation of state*, or more succinctly just an equation of state (see Section 20.2 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.

19.5.2 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)

$$\mathcal{I} = \mathcal{I}(T) \quad \text{ideal gas.} \quad (19.68)$$

Consequently, the exact differential of internal energy for an ideal gas is

$$d\mathcal{I} = c_v dT. \quad (19.69)$$

The appearance of c_v , the constant volume specific heat capacity discussed in Section 19.4, arises in order for the ideal gas internal energy to satisfy the general equation (19.61). 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

$$\mathcal{I} = c_v T + \text{constant} \quad \text{simple ideal gas.} \quad (19.70)$$

The arbitrary constant of integration is generally set to zero so that the internal energy vanishes at absolute zero.

19.5.3 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 19.70). 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

$$k_B = 1.3806 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1} \quad (19.71)$$

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,² 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 (13.4))

$$\mathcal{I}_{\text{mole diatomic}} = \frac{5 A_v k_B T}{2} = \frac{5 R T}{2}, \quad (19.72)$$

where the gas constant is given by

$$R = A_v k_B \quad (19.73a)$$

$$= (6.022 \times 10^{23} \text{ mole}^{-1}) (1.3806 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}) \quad (19.73b)$$

$$= 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1}. \quad (19.73c)$$

Finally, dividing by the molar mass for dry air (equation (13.5))

$$M_{\text{air}} = 0.028 \text{ kg mole}^{-1} \quad (19.74)$$

leads to the *simple ideal gas* approximation to the dry air heat capacity

$$c_v = \frac{5 R}{2 M_{\text{air}}} = 742 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1}. \quad (19.75)$$

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.

²At high temperatures, two vibrational degrees of freedom are also excited so that $\mathcal{I}_{\text{molecule}} = 7 k_B T/2$.

19.5.4 Enthalpy

The enthalpy is generally given by equation (19.46)

$$\mathcal{H} = \mathcal{I} + p/\rho. \quad (19.76)$$

For a *simple ideal gas* this expression takes the form

$$\mathcal{H} = \mathcal{I} + p/\rho \quad (19.77a)$$

$$= c_v T + \frac{TR}{M_{\text{mole}}} \quad (19.77b)$$

$$= T [c_v + R^M] \quad (19.77c)$$

where $R^M = R/M_{\text{mole}}$ (equation (19.67)) is the gas constant divided by the molar mass for the gas. Recall that the constant pressure heat capacity is given by equation (19.62)

$$c_p = T \left[\frac{\partial \mathcal{S}}{\partial T} \right]_{p,C} = \left[\frac{\partial \mathcal{H}}{\partial T} \right]_{p,C}. \quad (19.78)$$

Consequently, for a *simple ideal gas* we have

$$c_p = c_v + R^M \quad (19.79)$$

$$\mathcal{H} = c_p T. \quad (19.80)$$

19.5.5 Thermal expansion coefficient

The thermal expansion coefficient measures the relative changes in density as temperature changes at constant pressure and concentration

$$\beta_T = -\frac{1}{\rho} \left[\frac{\partial \rho}{\partial T} \right]_{p,C} \quad (19.81a)$$

$$= \frac{1}{\alpha} \left[\frac{\partial \alpha}{\partial T} \right]_{p,C} \quad (19.81b)$$

The thermal expansion coefficient for an ideal gas is given by

$$\beta_T = \frac{1}{T}, \quad (19.82)$$

so that as temperature rises the thermal expansion reduces.

19.5.6 Fundamental thermodynamic relations

The fundamental thermodynamic relation in terms of internal energy (equation (19.33)) and enthalpy (equation (19.46)) are given by

$$d\mathcal{I} = T d\mathcal{S} - p d\alpha + \mu dC \quad (19.83)$$

$$d\mathcal{H} = T d\mathcal{S} + \alpha dp + \mu dC. \quad (19.84)$$

For a simple ideal gas these relations take the form

$$c_v dT = T d\mathcal{S} - p d\alpha + \mu dC \quad (19.85)$$

$$c_p dT = T d\mathcal{S} + \alpha dp + \mu dC. \quad (19.86)$$

19.5.7 Further reading

Sections 1.5 and 1.7 of [Vallis \(2017\)](#) offer more details and results for a simple ideal gas atmosphere.

19.6 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.

19.6.1 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

$$\hat{\Gamma} \equiv \left[\frac{\partial T}{\partial p} \right]_{C,S}, \quad (19.87)$$

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.

19.6.2 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 (19.55)). Hence, infinitesimal changes in entropy are given by

$$dS = \left[\frac{\partial S}{\partial T} \right]_p dT + \left[\frac{\partial S}{\partial p} \right]_T dp. \quad (19.88)$$

Substituting the definition of heat capacity (19.60) leads to

$$T dS = c_p dT + T \left[\frac{\partial S}{\partial p} \right]_T dp. \quad (19.89)$$

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 (19.33) (with $dC = 0$) leads to

$$T \left[\frac{\partial S}{\partial T} \right]_p = \left[\frac{\partial J}{\partial T} \right]_p + p \left[\frac{\partial \alpha}{\partial T} \right]_p \quad (19.90)$$

as well as

$$T \left[\frac{\partial S}{\partial p} \right]_T = \left[\frac{\partial J}{\partial p} \right]_T + p \left[\frac{\partial \alpha}{\partial p} \right]_T. \quad (19.91)$$

Applying $(\partial/\partial p)_T$ to equation (19.90) and $(\partial/\partial T)_p$ to equation (19.91), and then subtracting, leads to the identity

$$\left[\frac{\partial S}{\partial p} \right]_T = - \left[\frac{\partial \alpha}{\partial T} \right]_p. \quad (19.92)$$

Introducing the thermal expansion coefficient (19.81b) yields an expression for changes in entropy in terms of changes in temperature and pressure

$$T dS = c_p dT - T \left[\frac{\partial \rho^{-1}}{\partial T} \right]_p dp \quad (19.93a)$$

$$= c_p dT - \left[\frac{T \beta_T}{\rho} \right] dp. \quad (19.93b)$$

c_p and β_T are readily measurable *response functions*, thus making equation (19.93b) a useful expression for infinitesimal entropy changes when matter concentration is held constant.

19.6.3 Adiabatic lapse rate for pressure changes

Equation (19.93b) means that the change in temperature associated with motion through a pressure field, with $dS = 0$ and $dC = 0$ can be written

$$\hat{\Gamma} = \left[\frac{\partial T}{\partial p} \right]_{C,S} = \frac{T \beta_T}{\rho c_p}. \quad (19.94)$$

Temperature indeed changes when pressure changes, even though there has been no heat exchanged with the environment.

19.6.4 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 27.2). Hydrostatic balance in a constant gravity field maintains the following relation between the pressure differential increment and the vertical differential increment

$$dp = -g \rho dz. \quad (19.95)$$

Use of the chain rule within the lapse rate expression (19.94) leads to

$$\Gamma = \left[\frac{\partial T}{\partial z} \right]_{C,S} \quad (19.96a)$$

$$= \left[\frac{\partial T}{\partial p} \right]_{C,S} \left[\frac{\partial p}{\partial z} \right] \quad (19.96b)$$

$$= -\rho g \left[\frac{T \beta_T}{\rho c_p} \right] \quad (19.96c)$$

$$= - \left[\frac{g T \beta_T}{c_p} \right]. \quad (19.96d)$$

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.

19.6.5 Adiabatic lapse rate for an ideal gas atmosphere

For an ideal gas (Section 19.5), the thermal expansion coefficient is given by (equation (19.82))

$$\beta_T = \frac{1}{T} \quad (19.97)$$

so that the lapse rates are given by

$$\hat{\Gamma} = \frac{1}{\rho c_p} \quad (19.98)$$

$$\Gamma = -\frac{g}{c_p}. \quad (19.99)$$

The measured specific heat capacity for a dry atmosphere at standard temperature (300 K) is

$$c_p = 1005 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1} \quad (19.100)$$

so that the adiabatic lapse rate for a dry atmosphere is roughly

$$\Gamma_d = -9.8 \text{ K/(1000 m)}. \quad (19.101)$$

Hence, temperature decreases by nearly 10 K when rising 1000 m in a dry and ideal gas atmosphere.

19.6.6 Further reading

[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.

19.7 Thermodynamics of a moving fluid

Recall the fundamental thermodynamic relation (19.33) for a binary fluid such as seawater

$$dJ = T dS - p d\alpha + \mu dC. \quad (19.102)$$

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.³ 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.

19.7.1 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 (19.102), which is the First Law for a quasi-static process, takes the material form

$$\frac{D\mathcal{J}}{Dt} = T \frac{DS}{Dt} - p \frac{D\alpha}{Dt} + \mu \frac{DC}{Dt}. \quad (19.103)$$

This equation is the First Law of thermodynamics for a moving fluid element.

Concerning the transition to a continuum

The expression (19.103) 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 (19.102) are continuous functions of space and time. Furthermore, equation (19.102) provides a relation between exact differentials as detailed in Section 19.1. As exact differentials of continuous fields, we can make use of the space-time increments detailed in Section 14.4.1 to write

$$d\Psi = \Psi(\mathbf{x} + d\mathbf{x}, t + dt) - \Psi(\mathbf{x}, t) \quad (19.104a)$$

$$= dt \partial_t \Psi + d\mathbf{x} \cdot \nabla \Psi, \quad (19.104b)$$

³The *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 14.4.2, we are thus led to the total time derivative for a property following a trajectory $\mathbf{x} = \mathbf{x}(t)$

$$\frac{d\Psi}{dt} = \frac{\partial\Psi}{\partial t} + \frac{d\mathbf{x}}{dt} \cdot \nabla\Psi. \quad (19.105)$$

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 14.4.4

$$\frac{D\Psi}{Dt} = \frac{\partial\Psi}{\partial t} + \mathbf{v} \cdot \nabla\Psi. \quad (19.106)$$

We are thus led to conclude that the expression (19.103) for the First Law in a continuum fluid is an inevitable result of transitioning the equilibrium relation (19.102) to the continuum.

Massaging the First Law

We now massage the result (19.103) to further reveal its connection to the First Law written in the form (19.18). For this purpose, recall that mass conservation as discussed in Section 15.1 means that changes in the volume of a fluid element are related to density changes via

$$\frac{1}{\delta V} \frac{D\delta V}{Dt} = \frac{1}{\alpha} \frac{D\alpha}{Dt}, \quad (19.107)$$

where again $\alpha = \rho^{-1}$ is the specific volume. Hence, equation (19.103) can be written

$$\delta M \frac{D\mathcal{J}}{Dt} = T \delta M \frac{DS}{Dt} - p \frac{D\delta V}{Dt} + \mu \delta M \frac{DC}{Dt}, \quad (19.108)$$

where

$$\delta M = \rho \delta V \quad (19.109)$$

is the mass of the fluid element. Since the mass of the fluid element is constant, equation (19.108) is the fluid element extension of the First Law given by equation (19.18). Alternatively, we can use the further result from mass conservation

$$\frac{1}{\delta V} \frac{D\delta V}{Dt} = \nabla \cdot \mathbf{v} \quad (19.110)$$

to write

$$\rho \frac{D\mathcal{J}}{Dt} = T \rho \frac{DS}{Dt} - p \nabla \cdot \mathbf{v} + \mu \rho \frac{DC}{Dt}. \quad (19.111)$$

We now introduce a commonly used notation for the heating rate

$$T \frac{DS}{Dt} \equiv \dot{Q} \quad (19.112)$$

in which case the First Law for a moving fluid element (19.111) takes the form

$$\frac{D\mathcal{J}}{Dt} = -\alpha p \nabla \cdot \mathbf{v} + \dot{Q} + \mu \frac{DC}{Dt}. \quad (19.113)$$

19.7.2 Further reading

[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 \(2017\)](#).

19.8 Potential temperature

Heating and cooling of the ocean, as well as mass exchange, 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 15.4). 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 19.6).

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.

19.8.1 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 19.6)

$$dT = \hat{\Gamma} dp. \quad (19.114)$$

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. That is the reason to introduce potential temperature.

19.8.2 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 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.

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

$$T = \theta(C, T, p_R) + \int_{p_R}^p \hat{\Gamma}(C, T, p') dp', \quad (19.115)$$

with $\hat{\Gamma}$ the lapse rate defined in terms of pressure changes (equation (19.94)). 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.

19.8.3 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 (19.55))

$$\mathcal{S} = \mathcal{S}(C, T, p) \quad (19.116)$$

leads to the defining identity for potential temperature

$$\mathcal{S}(C, T, p) = \mathcal{S}(C, \theta, p_R). \quad (19.117)$$

This relation directly connects changes in entropy to changes in potential temperature

$$d\mathcal{S} = \left[\frac{\partial \mathcal{S}(C, \theta, p_R)}{\partial \theta} \right]_C d\theta. \quad (19.118)$$

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 (19.118) by recalling that equation (19.93b) relates increments in specific entropy to temperature and pressure

$$T d\mathcal{S} = c_p (dT - \hat{\Gamma} dp), \quad (19.119)$$

where $\hat{\Gamma}$ is the adiabatic lapse rate defined in terms of pressure changes (equation (19.94)), and we set $dC = 0$. Taking the infinitesimal increment (i.e., the differential) of the potential temperature (19.115) leads to

$$dT = d\theta + \hat{\Gamma}(C, T, p) dp + \int_{p_R}^p [d\hat{\Gamma}(C, T, p')] dp'. \quad (19.120)$$

Evaluate this increment at the reference pressure, $p = p_R$, so that the integral vanishes, thus leaving

$$d\theta = dT - \hat{\Gamma}(C, T, p_R) dp. \quad (19.121)$$

We make use of this relation in equation (19.119) to render an expression for the entropy increment in terms of the potential temperature increment

$$d\mathcal{S} = c_p \theta^{-1} d\theta \quad p = p_R \text{ and } dC = 0. \quad (19.122)$$

As part of exercise 19.5, we show that this relation holds for an ideal gas at all pressures, and as part of exercise 19.6, we see that the relation also holds for all pressures in certain liquids.

19.8.4 The First Law in terms of potential temperature

Recall the expression of the First Law for a moving fluid element given by equation (19.113)

$$\frac{D\mathcal{J}}{Dt} = -p \nabla \cdot \mathbf{v} + \dot{Q} + \mu \frac{DC}{Dt}, \quad (19.123)$$

where

$$T \frac{D\mathcal{S}}{Dt} = \dot{Q} \quad (19.124)$$

is the heating rate. From equation (19.122) 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

$$\frac{D\mathcal{S}}{Dt} = \frac{c_p}{\theta} \frac{D\theta}{Dt} \quad (19.125)$$

Since the potential temperature equals to the temperature when $p = p_R$, we have

$$c_p \frac{D\theta}{Dt} = \theta \frac{D\mathcal{S}}{Dt} = \dot{Q} \quad \text{at } p = p_R \text{ and } dC = 0. \quad (19.126)$$

In Exercise 19.5 we consider these relations for an ideal gas.

19.8.5 Potential temperature for an ideal gas

The fundamental thermodynamic relation for a *simple ideal gas* (19.86) takes on the following form for an isentropic change

$$c_p dT = \alpha dp. \quad (19.127)$$

Dividing both sides by temperature and using the ideal gas relation

$$\frac{\alpha}{T} = \frac{R^M}{p} \quad (19.128)$$

leads to

$$c_p d \ln T = R^M d \ln p. \quad (19.129)$$

Since c_p and R^M are constants, we can integrate this relation from the reference pressure to an arbitrary pressure

$$c_p \int_{\theta}^T d \ln T = R^M \int_{p_R}^p d \ln p, \quad (19.130)$$

which renders the explicit expression for the simple ideal gas potential temperature

$$\theta = T \left[\frac{p_R}{p} \right]^{R^M/c_p} \quad (19.131)$$

where

$$c_p = \frac{7 R^M}{2} \quad (19.132)$$

is the constant pressure heat capacity of a simple ideal gas of diatomic molecules (Section 19.5.3).

19.8.6 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 19.6), 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.⁴

In numerical and theoretical models, potential temperature satisfies an advection/diffusion equation analogous to material tracers such as salinity (see Chapter 22). We will make use of potential temperature as the primary field measuring the buoyancy of a fluid in various theoretical models.

19.8.7 Further reading

Much of the material here follows Sections 1.5 and 1.6 of [Vallis \(2017\)](#).

19.9 Exercises

EXERCISE 19.1: DERIVATION OF GIBBS-DUHEM RELATION

Starting from the scaling (19.23) for the internal energy, work through the steps leading to the Gibbs-Duhem relation (19.25).

EXERCISE 19.2: HELMHOLTZ FREE ENERGY

The Helmholtz free energy is defined by the Legendre transformation

$$\mathcal{F} = \mathcal{I} - T\mathcal{S}. \quad (19.133)$$

Show that the fundamental thermodynamic relation in terms of \mathcal{F} is given by

$$d\mathcal{F} = -\mathcal{S}dT - p d\alpha + \mu dC. \quad (19.134)$$

EXERCISE 19.3: 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

$$p\alpha^{c_p/c_v} = \text{constant}, \quad (19.135)$$

where $\alpha = \rho^{-1}$ is the specific volume.

EXERCISE 19.4: 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

$$\frac{\partial T}{\partial z} = -|\Gamma| = \text{constant} \quad (19.136)$$

⁴Further advances ([McDougall, 2003; IOC et al., 2010](#)) show that Conservative Temperature is an even more suitable thermodynamic tracer for the ocean than potential temperature. See Section 1.7.3 of [Vallis \(2017\)](#) for discussion.

has a geopotential height at a pressure p given by

$$z = \frac{T_0}{|\Gamma|} \left[1 - \left(\frac{p_0}{p} \right)^{-R^M |\Gamma|/g} \right], \quad (19.137)$$

where T_0 is the temperature at $z = 0$.

- (b) 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 (19.137) in the appropriate limit.

EXERCISE 19.5: THERMODYNAMIC MANIPULATIONS FOR IDEAL GASES

This question develops some manipulations with the potential temperature.

- (a) Beginning with the expression (19.131) for potential temperature of an ideal gas, show that

$$d\theta = \frac{\theta}{T} \left[dT - \frac{\alpha}{c_p} dp \right]. \quad (19.138)$$

- (b) Given the result (19.138), show that an ideal gas satisfies the following relation

$$T dS = \frac{c_p T}{\theta} d\theta. \quad (19.139)$$

Whereas the relation (19.122) 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 idea gas satisfies the material time relation

$$T \frac{DS}{Dt} = \frac{c_p T}{\theta} \frac{D\theta}{Dt} \Rightarrow \frac{c_p T}{\theta} \frac{D\theta}{Dt} = \dot{Q}. \quad (19.140)$$

EXERCISE 19.6: THERMODYNAMIC MANIPULATIONS FOR A LIQUID

Consider seawater with specific entropy given by (see Section 1.7.2 of [Vallis \(2017\)](#))

$$S(S, T, p) = S_0 + c_{p0} \ln(T/T_o) [1 + \beta_s^* (S - S_o)] - \alpha_o p \left[\beta_T + \beta_T \gamma^* \frac{p}{2} + \beta_T^* (T - T_o) \right], \quad (19.141)$$

and corresponding specific heat capacity at constant pressure

$$c_p(S, T, p) = c_{p0} [1 + \beta_s^* (S - S_o)] - \alpha_o p \beta_T^* T. \quad (19.142)$$

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

$$\theta dS = c_p(S, \theta, p_R) d\theta, \quad (19.143)$$

where θ is the potential temperature and p_R is the corresponding reference pressure. Consequently, we can write for a moving fluid element

$$\dot{Q} = \frac{c_p T}{\theta} \frac{D\theta}{Dt}, \quad (19.144)$$

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 19.5. Hint: Make use of the identity (19.117).

20

Buoyancy acceleration

Buoyancy is an acceleration felt by an object within a fluid due to the acceleration of gravity acting on the object relative to the gravitational acceleration of the fluid displaced by the object. As shown here, this statement of *Archimedes' Principle* is a consequence of hydrostatics. We then discuss the variety of means to measure the density of a fluid, and further examine how to measure the stratification of buoyancy through the buoyancy frequency and neutral directions.

READER'S GUIDE TO THIS CHAPTER

This chapter relies minimally on other chapters, but it is a central feature of all subsequent chapters given the importance of buoyancy in geophysical fluids.

20.1	Buoyancy and Archimedes' Principle	276
20.1.1	Hydrostatic pressure force	276
20.1.2	Vertical buoyancy force	277
20.2	Mass density and its flavors	277
20.2.1	Realistic fluids	277
20.2.2	Equation of state in terms of potential temperature	278
20.2.3	Equation of state in terms of Conservative Temperature	278
20.2.4	Infinitesimal density increments and material time changes	278
20.2.5	Potential density	279
20.2.6	Linear equation of state for the ocean	280
20.2.7	Further reading	280
20.3	Buoyancy stratification	281
20.3.1	Physical ideas underlying neutral directions	281
20.3.2	Comparing density under two forms of displacement	281
20.3.3	Buoyancy frequency	282
20.3.4	Buoyancy frequency and locally referenced potential density	283
20.3.5	Neutral directions	283
20.3.6	Further reading	283
20.4	Revisiting the atmospheric dry adiabatic lapse rate	284
20.5	Exercises	285

20.1 Buoyancy and Archimedes' Principle

Consider an arbitrary rigid object contained within a fluid at rest, such as that shown in Figure 20.1. The object displaces a volume of fluid whose mass is given by the integral of the density over the region,

$$m = \int_{\mathcal{R}} \rho dV. \quad (20.1)$$

What is the force acting on the object?

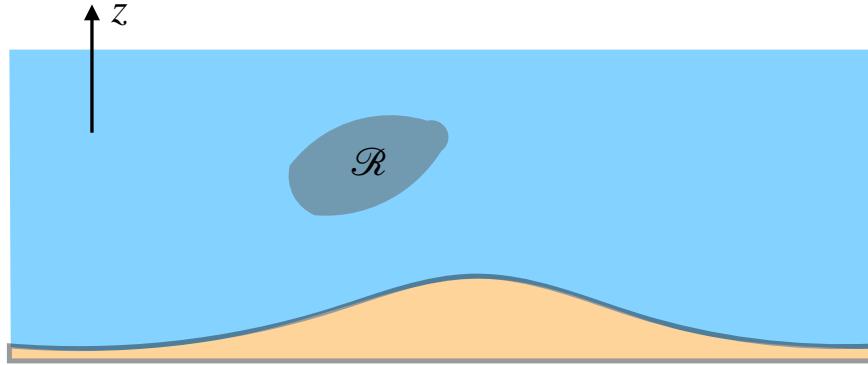


Figure 20.1: An arbitrary rigid object, \mathcal{R} , within a fluid displaces an amount of fluid. The force acting on the object equals to the weight of the fluid displaced by the object. This is Archimedes' Principle, which is a consequence of hydrostatics. The buoyancy force is determined by the volume integrated difference in densities between the solid and fluid, with a positive buoyancy force for objects less dense than the displaced fluid and conversely for denser objects.

20.1.1 Hydrostatic pressure force

As we will discuss in Chapter 28, 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 surface of \mathcal{R} leads to the pressure force acting on the object

$$\mathbf{F}_{\text{pressure}} = - \int_{\partial\mathcal{R}} p \hat{\mathbf{n}} dS, \quad (20.2)$$

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 (3.63) leads to the equivalent expression in terms of the volume integral of the pressure gradient

$$\mathbf{F}_{\text{pressure}} = - \int_{\mathcal{R}} \nabla p dV. \quad (20.3)$$

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, which means that it has a vertical derivative given by

$$\frac{\partial p}{\partial z} = -\rho g. \quad (20.4)$$

We thus see that the force acting on the solid equals to

$$\mathbf{F}_{\text{pressure}} = - \int_{\mathcal{R}} \nabla p \, dV = - \int_{\mathcal{R}} \hat{\mathbf{z}} (\partial p / \partial z) \, dV = \hat{\mathbf{z}} g \int_{\mathcal{R}} \rho \, dV = \hat{\mathbf{z}} g m \quad (20.5)$$

where we assumed a constant gravitational acceleration over \mathcal{R} . The force imparted by the hydrostatic pressure is thus a vertical upward force that equals to the weight of the fluid displaced by the solid. This result is a mathematical expression of Archimedes' Principle.

Note that there are no net horizontal forces acting on the object. The reason is that there are no horizontal pressure gradients since, by assumption, the fluid is at rest and thus experiences no horizontal acceleration. The object experiences horizontal compressive pressure forces, but they balance when integrated around the object so that the object has no horizontal acceleration.

20.1.2 Vertical buoyancy force

The vertical pressure force (20.6) acts to push the object vertically upward against gravity. It therefore appears to reduce the vertical gravitational force when the solid is within the fluid. We say that the object has a *buoyancy* imparted by the fluid that it displaces. Now if the object has a density less than the displaced fluid, then the object experiences a positive buoyancy force relative to the displaced fluid, in which case it rises upward. The converse holds for an object more dense than the displaced fluid, in which case the object sinks.

We introduce a *buoyancy force* acting on an object relative to the displaced fluid as determined by

$$\mathbf{F}_{\text{buoyancy}} = \hat{\mathbf{z}} g \int_{\mathcal{R}} (\rho_{\text{fluid}} - \rho_{\text{object}}) \, dV = \hat{\mathbf{z}} g (m_{\text{fluid}} - m_{\text{object}}). \quad (20.6)$$

The vertical buoyancy force is negative if the mass of the object is larger than that of the fluid it displaces; i.e., it sinks. Conversely, the object rises if less massive. If the densities are equal, then the object is *neutrally buoyant* and thus experiences no net vertical force; i.e., it floats.

20.2 Mass density and its flavors

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 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 19.5.1. We here discuss the equation of state as well as the related flavors of mass density used to study stratified fluids.

20.2.1 Realistic fluids

To allow for general situations of multi-component fluids, we write the *in situ* density as a general function

$$\rho = \rho(S, T, p). \quad (20.7)$$

We made use of the standard oceanographic notation where $S = 1000 C$ is the salinity as defined by equation (19.35). This equation of state can be derived from one of the more general *fundamental equations of state* discussed in Section 19.3.

The thermal equation of state for fluid density is generally not as simple as that for an ideal gas discussed in Section 19.5.1. 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 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 (20.7) 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 19.8).

20.2.2 Equation of state in terms of potential temperature

Equation (19.115) provides a unique relation between potential temperature θ and *in situ* temperature, T , salinity and pressure. Furthermore, as discussed in Section 19.8.6, 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

$$\rho = \rho(S, \theta, p). \quad (20.8)$$

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.

20.2.3 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

$$\rho = \rho(S, \Theta, p). \quad (20.9)$$

For our purposes it is not important to make the distinction between potential temperature and Conservative Temperature. We thus use θ for brevity in this chapter.

20.2.4 Infinitesimal density increments and material time changes

Given the functional dependence for the equation of state written in terms of S, θ, p (equation (20.8)), an infinitesimal density increment is given by

$$d\rho = \left[\frac{\partial \rho}{\partial S} \right] dS + \left[\frac{\partial \rho}{\partial \theta} \right] d\theta + \left[\frac{\partial \rho}{\partial p} \right] dp \quad (20.10a)$$

$$\equiv \rho \beta dS - \rho \alpha d\theta + c_{(s)}^{-2} dp. \quad (20.10b)$$

The second line introduced the following thermodynamic properties of the fluid

$$\beta = \frac{1}{\rho} \left[\frac{\partial \rho}{\partial S} \right]_{\theta,p} \quad \text{haline contraction coefficient} \quad (20.11)$$

$$\alpha = -\frac{1}{\rho} \left[\frac{\partial \rho}{\partial \theta} \right]_{S,p} \quad \text{thermal expansion coefficient} \quad (20.12)$$

$$c_{(s)}^2 = \left[\frac{\partial p}{\partial \rho} \right]_{S,\theta} \quad \text{squared sound speed.} \quad (20.13)$$

The haline contraction coefficient, β , is considered for the ocean, where *haline* refers to salinity.¹

The infinitesimal density increment (20.10b) leads to the expression for the material change

$$\frac{1}{\rho} \frac{D\rho}{Dt} = \beta \frac{DS}{Dt} - \alpha \frac{D\theta}{Dt} + \frac{1}{\rho c_{(s)}^2} \frac{Dp}{Dt}. \quad (20.14)$$

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

$$\frac{D\rho}{Dt} = \frac{1}{c_{(s)}^2} \frac{Dp}{Dt} \quad \text{adiabatic and isohaline changes.} \quad (20.15)$$

20.2.5 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 19.8.3). 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 (19.117))

$$\mathcal{S}(S, T, p) = \mathcal{S}(S, \theta, p^{(r)}). \quad (20.16)$$

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

$$\rho_{\text{pot}} = \rho(S, \theta, p^{(r)}). \quad (20.17)$$

As for the potential temperature, the reference pressure is often taken as that at sea level, though this is not necessary.

¹We also use β in this book to refer to the meridional derivative of the Coriolis parameter: $\beta = \partial f / \partial y$. The distinct contexts for the two uses of β is sufficient to avoid confusion.

Material evolution of potential density

With the definition (20.17), the material evolution of potential density is given by

$$\frac{1}{\rho^{(r)}} \frac{D\rho_{\text{pot}}}{Dt} = \beta^{(r)} \frac{DS}{Dt} - \beta^{(r)} \frac{D\theta}{Dt}, \quad (20.18)$$

where

$$\beta^{(r)} = \frac{1}{\rho^{(r)}} \left[\frac{\partial \rho(S, \theta, p^{(r)})}{\partial S} \right]_{\theta} \quad \text{haline contraction at } p = p^{(r)} \quad (20.19)$$

$$\beta^{(r)} = -\frac{1}{\rho^{(r)}} \left[\frac{\partial \rho(S, \theta, p^{(r)})}{\partial \theta} \right]_S \quad \text{thermal expansion at } p = p^{(r)} \quad (20.20)$$

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 (20.14) and (20.15).

Reference pressures for ρ_{pot} and θ

As defined by equation (20.17), 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 for potential density reference pressures to be distinct from the surface, 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.

20.2.6 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

$$\rho_{\text{pot}} = \rho_0 [1 - \alpha(\theta - \theta_0) + \beta(S - S_0)], \quad (20.21)$$

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.

20.2.7 Further reading

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.

20.3 Buoyancy stratification

In many parts of the ocean and atmosphere, especially those close to the reference pressure, potential temperature (atmosphere) and potential density (ocean) are monotonically stacked in the vertical. Hence, a position in the fluid can be specified by the value of (x, y, θ) (atmosphere) or $(x, y, \rho_{\text{pot}})$ (ocean). Furthermore, many physical processes related to tracer transport naturally occur within, rather than across, such surfaces. Indeed, if there is no mixing or internal sources of heat or mass, then the reversible transport of matter and potential temperature occurs precisely along such layers. We here introduce notions related to these *stratification* surfaces.

20.3.1 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 can modify its thermal, material, and mechanical properties; i.e., its θ , S , and p . Modification of its pressure occurs through acoustic waves, pressure form drag, and other mechanical means for exchanging pressure with adjacent fluid elements. Modification of the thermal and material properties occurs through the exchange of heat and matter with adjacent elements. The exchange of heat and matter only occurs in the presence of irreversible processes such as mixing, whereas mechanical exchanges can occur either reversibly or irreversibly.

In general, the exchange of heat and matter alters the buoyancy of a fluid element relative to its local environment. However, it is possible to modify θ and S during a displacement without modifying its buoyancy. To do so requires that changes in θ that precisely compensate changes in S . The 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.

20.3.2 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, $\rho(\mathbf{x} + d\mathbf{x})$. To leading order, the density difference between the two positions is computed according to

$$d\rho = \rho(\mathbf{x} + d\mathbf{x}) - \rho(\mathbf{x}) \quad (20.22a)$$

$$= d\mathbf{x} \cdot \left[\frac{\partial \rho}{\partial \theta} \nabla \theta + \frac{\partial \rho}{\partial S} \nabla S + \frac{\partial \rho}{\partial p} \nabla p \right] \quad (20.22b)$$

$$= \rho d\mathbf{x} \cdot \left[-\alpha \nabla \theta + \beta \nabla S + \frac{1}{\rho c_{(s)}^2} \nabla p \right]. \quad (20.22c)$$

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

$$(d\rho)_{(\text{no mix})} = \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} - \rho(\mathbf{x}) \quad (20.23a)$$

$$= \rho d\mathbf{x} \cdot \left[\frac{1}{\rho c_{(s)}^2} \nabla p \right]. \quad (20.23b)$$

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

$$\rho(\mathbf{x} + d\mathbf{x}) - \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} = d\rho - (d\rho)_{(\text{no mix})} \quad (20.24a)$$

$$= \rho d\mathbf{x} \cdot [-\alpha \nabla \theta + \beta \nabla S]. \quad (20.24b)$$

20.3.3 Buoyancy frequency

Taking the special case of a vertical displacement yields

$$\rho(z + dz) - \rho(z + dz)_{(\text{no mix})} = \rho dz \left[-\alpha \frac{\partial \theta}{\partial z} + \beta \frac{\partial S}{\partial z} \right]. \quad (20.25)$$

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

$$g [\rho(z + dz) - \rho(z + dz)_{(\text{no mix})}] = g \rho dz \left[-\alpha \frac{\partial \theta}{\partial z} + \beta \frac{\partial S}{\partial z} \right] \equiv -N^2 \rho dz, \quad (20.26)$$

where we defined the squared buoyancy frequency

$$N^2 = g \left[\alpha \frac{\partial \theta}{\partial z} - \beta \frac{\partial S}{\partial z} \right]. \quad (20.27)$$

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*.

20.3.4 Buoyancy frequency and locally referenced potential density

Equation (20.27) 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 (20.17), 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

$$N^2 = -g \left[\frac{\partial \ln \rho_{\text{pot}}}{\partial z} \right]_{p^{(r)}=p} = g \left[\alpha \frac{\partial \theta}{\partial z} - \beta \frac{\partial S}{\partial z} \right]. \quad (20.28)$$

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.

20.3.5 Neutral directions

Rather than specializing to a vertical displacement as for the buoyancy frequency, consider an arbitrary three-dimensional displacement. For this purpose, return to equation (20.24b) to write

$$\rho(\mathbf{x} + d\mathbf{x}) - \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} = \rho(\mathbf{x}) d\mathbf{x} \cdot [-\alpha \nabla \theta + \beta \nabla S] \quad (20.29a)$$

$$= \rho(\mathbf{x}) d\mathbf{x} \cdot \hat{\gamma} | -\alpha \nabla \theta + \beta \nabla S|. \quad (20.29b)$$

The second expression introduced the *dianeutral unit vector*

$$\hat{\gamma} = \frac{\rho_\theta \nabla \theta + \rho_S \nabla S}{|\rho_\theta \nabla \theta + \rho_S \nabla S|} = \frac{-\alpha \nabla \theta + \beta \nabla S}{|-\alpha \nabla \theta + \beta \nabla S|}. \quad (20.30)$$

Displacements, $d\mathbf{x}$, orthogonal to $\hat{\gamma}$ lead to no difference in the density between the environment and the unmixed fluid element

$$d\mathbf{x} \cdot \hat{\gamma} = 0 \Rightarrow \rho(\mathbf{x} + d\mathbf{x}) = \rho(\mathbf{x} + d\mathbf{x})_{(\text{no mix})} \iff \text{neutral displacements.} \quad (20.31)$$

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.

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. Hence, 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 locally defined buoyancy force on the element.

20.3.6 Further reading

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. We have more to say regarding neutral displacements in Section 24.5, where we encounter their nontrivial topology.

20.4 Revisiting the atmospheric dry adiabatic lapse rate

We introduced the adiabatic lapse rate in Section 19.6 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 20.1)

$$N^2 = \frac{g}{\theta} \frac{\partial \theta}{\partial z}. \quad (20.32)$$

The potential temperature for an ideal gas is given by equation (19.131)

$$\theta = T \left[\frac{p^{(r)}}{p} \right]^\varphi \quad (20.33)$$

where

$$\varphi = \frac{R^M}{c_p} \quad (20.34)$$

is a constant for a simple ideal gas. Consequently, the squared buoyancy frequency takes the form

$$g^{-1} N^2 = \frac{\partial \ln \theta}{\partial z} \quad (20.35a)$$

$$= \frac{\partial \ln T}{\partial z} - \varphi \frac{\partial \ln p}{\partial z}. \quad (20.35b)$$

For a hydrostatic fluid with a constant gravitational acceleration, the vertical derivative of pressure is given by

$$\frac{\partial p}{\partial z} = -\rho g, \quad (20.36)$$

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

$$g^{-1} N^2 = \frac{\partial \ln T}{\partial z} + \frac{\varphi g \rho}{p} \quad (20.37a)$$

$$= \frac{1}{T} \frac{\partial T}{\partial z} + \frac{g}{c_p T}, \quad (20.37b)$$

where we used the ideal gas relation $p = \rho T R^M$ for the final step.

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

$$N^2 = 0 \iff \frac{\partial T}{\partial z} = \Gamma_d, \quad (20.38)$$

where (see equation (19.101))

$$\Gamma_d = -\frac{g}{c_p} \approx -9.8 \text{ K/(1000 m)}. \quad (20.39)$$

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

$$\text{stable} \quad N^2 > 0 \iff -\frac{\partial T}{\partial z} < \frac{g}{c_p} \quad (20.40)$$

$$\text{unstable} \quad N^2 < 0 \iff -\frac{\partial T}{\partial z} > \frac{g}{c_p}. \quad (20.41)$$

20.5 Exercises

EXERCISE 20.1: 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 (20.28).

EXERCISE 20.2: 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 20.3: VERTICAL INTEGRAL OF N^2

The expression for squared buoyancy frequency

$$N^2 = -g \left[\frac{\partial \ln \rho_{\text{pot}}}{\partial z} \right]_{p^{(r)}=p} \quad (20.42)$$

makes it tempting to consider its vertical integral according to

$$-g^{-1} \int_{-H}^{\eta} N^2 dz \stackrel{?}{=} [\ln \rho_{\text{pot}}]_{\eta} - [\ln \rho_{\text{pot}}]_{-H}. \quad (20.43)$$

Discuss what is wrong with this equation. Under what conditions is it correct?

21

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 21.1 presumes an understanding of how the tracer equation is derived from the kinematics of mass conserving fluid elements detailed in Section 15.4. Basic notions of thermodynamics (in particular Section 19.8) motivate the use of Conservative temperature (or potential temperature) as a measure of ocean heat transfer.

21.1	Salt and freshwater budgets	288
21.1.1	Mass budgets	288
21.1.2	Kinematic boundary conditons	289
21.1.3	Further reading	291
21.2	Material evolution of <i>in situ</i> density	291
21.2.1	Material changes to pressure	291
21.2.2	Material changes to Θ and S	292
21.2.3	Summary of density changes	292
21.2.4	A synopsis of ocean mixing processes	292
21.3	Boundary fluxes of buoyancy	293
21.3.1	Outlining the boundary fluxes of heat and salt	293
21.3.2	Transport of scalars across the ocean surface	294
21.3.3	Scalar budgets for a surface ocean model grid cell	294
21.3.4	Salt fluxes from sea ice melt and formation	295
21.3.5	Salt and heat fluxes associated with water transport	295
21.3.6	Non-penetrative surface heat fluxes	296
21.3.7	The case of frazil	297
21.3.8	Penetrative shortwave radiation	297
21.3.9	Buoyancy budget for a surface ocean model grid cell	298
21.3.10	Comments	298

21.1 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*

$$S = \frac{\text{mass of salt}}{\text{mass of seawater}} = \frac{\text{mass of salt}}{\text{mass of freshwater} + \text{mass of salt}} \quad (21.1)$$

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

$$F = \frac{\text{mass of freshwater}}{\text{mass of seawater}} = \frac{\text{mass of freshwater}}{\text{mass of freshwater} + \text{mass of salt}} = 1 - S. \quad (21.2)$$

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.

21.1.1 Mass budgets

Following our discussion of the tracer equation in Section 15.4, the mass budget equations for an element of seawater take the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad \text{seawater} \quad (21.3)$$

$$\frac{\partial(\rho S)}{\partial t} + \nabla \cdot (\rho \mathbf{v} S + \mathbf{J}^{(S)}) = 0 \quad \text{salt} \quad (21.4)$$

$$\frac{\partial(\rho F)}{\partial t} + \nabla \cdot (\rho \mathbf{v} F + \mathbf{J}^{(F)}) = 0 \quad \text{freshwater}, \quad (21.5)$$

where ρ is the seawater mass density. Equation (21.3) is the mass budget for seawater and equation (21.4) is the mass budget for salt. The freshwater budget (21.5) is derived by subtracting the salt budget (21.4) from the seawater mass budget (21.3). Hence, only two of the three mass budget equations (21.3)-(21.5) are independent.

We make use of the barycentric velocity in the above conservation laws, where the barycentric velocity is given by

$$\mathbf{v} = S \mathbf{v}^{(S)} + F \mathbf{v}^{(F)}. \quad (21.6)$$

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

$$\rho \left[\frac{\partial S}{\partial t} + \mathbf{v}^{(S)} \cdot \nabla S \right] = \frac{\partial(\rho S)}{\partial t} + \nabla \cdot (\rho \mathbf{v}^{(S)} S) = 0 \quad (21.7a)$$

$$\rho \left[\frac{\partial F}{\partial t} + \mathbf{v}^{(F)} \cdot \nabla F \right] = \frac{\partial(\rho F)}{\partial t} + \nabla \cdot (\rho \mathbf{v}^{(F)} F) = 0. \quad (21.7b)$$

Furthermore, the fluxes $\mathbf{J}^{(S)}$ and $\mathbf{J}^{(F)}$ arise from the difference between the salt and freshwater velocities from the barycentric velocity

$$\mathbf{J}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v}) \quad \mathbf{J}^{(F)} = \rho F (\mathbf{v}^{(F)} - \mathbf{v}). \quad (21.8)$$

These fluxes are generally parameterized by downgradient diffusive fluxes

$$\mathbf{J}^{(S)} = -\rho \kappa_S \nabla S \quad \mathbf{J}^{(F)} = -\rho \kappa_S \nabla F, \quad (21.9)$$

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 24.1, 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

$$\rho \mathbf{v} = \rho S \mathbf{v}^{(S)} + \rho F \mathbf{v}^{(F)}. \quad (21.10)$$

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

$$\rho S \mathbf{v}^{(S)} = \rho S (\mathbf{v}^{(S)} - \mathbf{v}) + \rho S \mathbf{v} = \mathbf{J}^{(S)} + \rho S \mathbf{v} \quad (21.11a)$$

$$\rho F \mathbf{v}^{(F)} = \rho F (\mathbf{v}^{(F)} - \mathbf{v}) + \rho F \mathbf{v} = \mathbf{J}^{(F)} + \rho F \mathbf{v}. \quad (21.11b)$$

The center of mass velocities offer a conceptual framework of use for a formulation of kinematic boundary conditions in Section 21.1.2. 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.

21.1.2 Kinematic boundary conditions

Recall the boundary condition (15.83) derived in Section 15.7.3 for matter crossing the ocean surface

$$\rho (\mathbf{v} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_m, \quad (21.12)$$

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 (21.12), 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

$$\rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_S \quad (21.13a)$$

$$\rho (\mathbf{v}^{(F)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = -\mathcal{Q}_F, \quad (21.13b)$$

where

$$\mathcal{Q}_m = \mathcal{Q}_S + \mathcal{Q}_F \quad (21.14)$$

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 (21.13a) and (21.13b), and using the relation $S + F = 1$, recovers the boundary condition (21.12) 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

$$\rho (\mathbf{v}^{(S)} - \mathbf{v}^{(\eta)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{zero surface salt flux.} \quad (21.15)$$

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 (21.11a) 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 (21.13a) takes the form

$$\mathbf{J}^{(S)} \cdot \hat{\mathbf{n}} = S \mathcal{Q}_F. \quad (21.16)$$

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

$$\mathbf{J}^{(S)} = -\rho \kappa_S \nabla S, \quad (21.17)$$

in which case the boundary condition (21.16) takes the form

$$\rho \kappa_S \nabla S \cdot \hat{\mathbf{n}} = -S \mathcal{Q}_F. \quad (21.18)$$

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.

21.1.3 Further reading

Nurser and Griffies (2019) offer a more detailed discussion of this material.

21.2 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 20.2.2, 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

$$\rho = \rho(\Theta, S, p). \quad (21.19)$$

We formulate the material evolution of density as weighted by the specific volume

$$\nu = \rho^{-1}, \quad (21.20)$$

so that we consider

$$\frac{D \ln \rho}{Dt} = \frac{\partial \ln \rho}{\partial \Theta} \frac{D\Theta}{Dt} + \frac{\partial \ln \rho}{\partial S} \frac{DS}{Dt} + \frac{1}{\rho} \frac{\partial \rho}{\partial p} \frac{Dp}{Dt} \quad (21.21a)$$

$$= -\alpha \frac{D\Theta}{Dt} + \beta \frac{DS}{Dt} + \frac{\omega}{\rho c^2}. \quad (21.21b)$$

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

$$\alpha = - \left[\frac{\partial \ln \rho}{\partial \Theta} \right]_{p,S} \quad \beta = \left[\frac{\partial \ln \rho}{\partial S} \right]_{p,\Theta} \quad c_s^2 = \left[\frac{\partial p}{\partial \rho} \right]_{S,\Theta} \quad \omega = \frac{Dp}{Dt}. \quad (21.22)$$

21.2.1 Material changes to pressure

To garner some exposure to the physics of ω as it appears in equation (21.21), 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 18.3.5)

$$\begin{aligned} w^{(p)} &= \frac{\partial z}{\partial p} \frac{Dp}{Dt} \\ &= -(\rho g)^{-1} \omega. \end{aligned} \quad (21.23)$$

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.

21.2.2 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

$$\rho \frac{D\Theta}{Dt} = -\nabla \cdot \mathbf{J}(\Theta) \quad (21.24a)$$

$$\rho \frac{DS}{Dt} = -\nabla \cdot \mathbf{J}(S). \quad (21.24b)$$

This form for material changes in temperature and salinity then lead to

$$-\alpha \frac{D\Theta}{Dt} + \beta \frac{DS}{Dt} = \nu_\Theta \nabla \cdot \mathbf{J}(\Theta) + \nu_S \nabla \cdot \mathbf{J}(S) \quad (21.25a)$$

$$= \nabla \cdot [\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S)] - [\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S] \quad (21.25b)$$

where again $\nu = \rho^{-1}$ is the specific volume and its partial derivatives are

$$\nu_\Theta = \frac{\partial \nu}{\partial \Theta} = \frac{\alpha}{\rho} \quad \text{and} \quad \nu_S = \frac{\partial \nu}{\partial S} = -\frac{\beta}{\rho}. \quad (21.26)$$

21.2.3 Summary of density changes

Bringing the above results together leads to the density equation

$$\frac{D \ln \rho}{Dt} - \frac{\omega}{\rho c_s^2} = \nabla \cdot [\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S)] - [\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S], \quad (21.27)$$

which has the equivalent form

$$\frac{D\rho}{Dt} - \frac{\omega}{c_s^2} = \nabla \cdot [\alpha \mathbf{J}(\Theta) - \beta \mathbf{J}(S)] - [\mathbf{J}(\Theta) \cdot \nabla \alpha - \mathbf{J}(S) \cdot \nabla \beta]. \quad (21.28)$$

We brought the adiabatic source term from motion across pressure surfaces (Section 21.2.1) 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 21.2.4. The second term on the right hand side of equations (21.27) and (21.28) relates to properties of the locally referenced potential density surface. We study this source term arising from neutral diffusion in Section 24.4, 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 24.1.4).

21.2.4 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 24.3). 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](#)).

21.3 Boundary fluxes of buoyancy

As introduced in Chapter 20, 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.

21.3.1 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.

21.3.2 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 15.4, and specialize to the case of a grid cell with upper boundary the air-sea interface. From equation (15.60), the transport of an arbitrary scalar, ψ , crossing the upper ocean interface is written

$$Q_\psi \, dA = -[\rho \psi (\mathbf{v} - \mathbf{v}^{(b)}) + \mathbf{J}] \cdot \hat{\mathbf{n}} \, dS, \quad (21.29)$$

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 (15.93c) renders the surface kinematic boundary condition

$$\rho (\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} \, dS = -Q_m \, dA \quad (21.30)$$

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

$$\mathbf{J} \cdot \hat{\mathbf{n}} \, dS = -Q_{sgs} \, dA, \quad (21.31)$$

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

$$Q_\psi = Q_m \psi + Q_{sgs}. \quad (21.32)$$

By construction, a positive contribution to Q_ψ increases the amount of ψ within the ocean domain.

21.3.3 Scalar budgets for a surface ocean model grid cell

We now make use of the surface flux expression (21.32) within the finite volume budget equation (15.60) 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

$$\frac{\partial (\rho dz \Theta)}{\partial t} = Q_m \Theta_m + Q_\Theta^{\text{non-pen}} + [Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z)] \quad (21.33)$$

$$\frac{\partial (\rho dz S)}{\partial t} = Q_m S_m + Q_S \quad (21.34)$$

$$\frac{\partial (\rho dz)}{\partial t} = Q_m. \quad (21.35)$$

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 29, 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 21.3.5.
- 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 21.3.5.
- 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 21.3.4.
- $C_p^o Q_\Theta^{\text{non-pen}}$ 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 21.3.6.
- $C_p^o Q_\Theta^{\text{pen}}(z = \eta)$ 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 21.3.8.

21.3.4 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. \(2009\)](#)).

21.3.5 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

$$Q_m \Theta_m = \frac{\mathcal{H}^{\text{liquid runoff}}}{C_p^{\text{liquid runoff}}}, \quad (21.36)$$

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

$$Q_m \Theta_m = \frac{\mathcal{H}^{\text{solid runoff}}}{C_p^{\text{solid runoff}}}, \quad (21.37)$$

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.

21.3.6 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](#))

$$C_p^o Q_\Theta^{\text{non-pen}} = Q_{\text{long}} + Q_{\text{latent}} + Q_{\text{sens}}. \quad (21.38)$$

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

$$\sigma_{\text{SB}} = 5.6734 \times 10^{-8} \text{ W m}^{-2} \text{ }^\circ\text{K}^{-4} \quad (21.39)$$

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

$$H^{\text{vapor}} = 2.5 \times 10^6 \text{ J kg}^{-1}, \quad (21.40)$$

so that

$$Q_{\text{evap}} = H^{\text{vapor}} Q_m^{\text{evap}} \quad (21.41)$$

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

$$H^{\text{fusion}} = 3.34 \times 10^5 \text{ J kg}^{-1}, \quad (21.42)$$

so that

$$Q_{\text{melt}} = -H^{\text{fusion}} Q_m^{\text{frozen precip}}, \quad (21.43)$$

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.

21.3.7 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.

21.3.8 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

$$\text{net shortwave heating of surface grid cell} = C_p^o [Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z)] \quad (21.44)$$

that stays in the surface grid cell.

21.3.9 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

$$-\left[\frac{\rho_o}{g}\right] \frac{\partial b}{\partial t} = \rho_\Theta \frac{\partial \Theta}{\partial t} + \rho_S \frac{\partial S}{\partial t}, \quad (21.45)$$

where we introduced the shorthand

$$\rho_\Theta = \left[\frac{\partial \rho}{\partial \Theta}\right]_{S,p} \quad \rho_S = \left[\frac{\partial \rho}{\partial S}\right]_{\Theta,p} \quad (21.46)$$

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 (21.33) by ρ_Θ and add to the surface salinity equation (21.34) multiplied by ρ_S

$$\rho_\Theta \left[\frac{\partial(\rho dz \Theta)}{\partial t} \right] + \rho_S \left[\frac{\partial(\rho dz S)}{\partial t} \right] = Q_m (\rho_\Theta \Theta_m + \rho_S S_m) + \rho_\Theta [Q_\Theta^{\text{non-pen}} + \delta_k Q_\Theta^{\text{pen}}] + \rho_S Q_S, \quad (21.47)$$

where we introduced the shorthand

$$\delta_k Q_\Theta^{\text{pen}} = Q_\Theta^{\text{pen}}(z = \eta) - Q_\Theta^{\text{pen}}(z = -\Delta z). \quad (21.48)$$

We now use the mass budget (21.35) and introduce the buoyancy tendency according to equation (21.45) to realize an expression for the time tendency of the surface ocean buoyancy

$$(\rho_o/g) \rho dz \frac{\partial b}{\partial t} = Q_m [\rho_\Theta (\Theta - \Theta_m) + \rho_S (S - S_m)] + \rho_\Theta [Q_\Theta^{\text{non-pen}} + \delta_k Q_\Theta^{\text{pen}}] - \rho_S Q_S. \quad (21.49)$$

Introducing the thermal expansion and saline contraction coefficients

$$\alpha = -\frac{1}{\rho} \left[\frac{\partial \rho}{\partial \Theta} \right]_{S,p} \quad \beta = \frac{1}{\rho} \left[\frac{\partial \rho}{\partial S} \right]_{\Theta,p} \quad (21.50)$$

to render

$$dz \frac{\partial b}{\partial t} = \frac{g}{\rho_o} (Q_m [-\alpha (\Theta - \Theta_m) + \beta (S - S_m)] + \alpha (\delta_k Q_\Theta^{\text{pen}} + Q_\Theta^{\text{non-pen}}) - \beta Q_S). \quad (21.51)$$

21.3.10 Comments

The buoyancy flux expression (21.51) 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 watermass transformation as reviewed by [Groeskamp et al. \(2019\)](#) and summarized in Section 25.1.

Part V

Tracers

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.

We devote Chapter 22 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 Appendix ?? 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 22, we encounter some of the canonical partial differential equations appearing in fluid mechanics. Mathematical facets of these equations are introduced in Appendix 1, which offers a synopsis of the linear partial differential equations of mathematical physics.

22

Advection and diffusion

In this chapter we discuss the physical and mathematical aspects of the advection-diffusion equation. This equation is used to describe the evolution of tracers within the ocean and atmosphere. Advection and diffusion have complementary physical and mathematical properties. 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 the advection-diffusion equation for a compressible/non-Boussinesq fluid. For a Boussinesq fluid (Chapter 29), the density factor, ρ , appearing in the non-Boussinesq formulation is set to a constant and thus trivially cancels. Although a Boussinesq fluid is commonly assumed for the ocean, we retain the non-Boussinesq formulation as it provides added generality for applications where compressibility is important.

22.1	Introduction	302
22.2	Diffusion physics	302
22.2.1	Diffusion of matter by random molecular motions	302
22.2.2	Diffusion of matter by random turbulent motions	303
22.2.3	Fick's law for matter diffusion	303
22.2.4	Fourier's law for heat diffusion	305
22.2.5	Newtonian frictional stress and momentum diffusion	305
22.2.6	Further reading	305
22.3	Diffusion maths	306
22.3.1	Sample diffusion tensors	306
22.3.2	Diffusion of concentration powers	307
22.3.3	Global integrals	307
22.3.4	Connecting tracer variance to the diffusion operator [†]	308
22.4	Advection physics	309
22.4.1	The advection equation	309
22.4.2	Eulerian time tendencies from advection	310
22.5	Advection maths	310
22.5.1	Material constancy of C^α	311
22.5.2	Mass transport from the mean, eddy, and residual-mean	311
22.5.3	Advection tracer fluxes and skew tracer fluxes	312
22.5.4	Skew diffusion	312
22.5.5	Further reading	313

22.6 Advection and skewson [†]	313
22.6.1 Choosing a gauge	314
22.6.2 Boundary conditions	315
22.7 Exercises	316

22.1 Introduction

As derived in Section 15.4, the tracer equation takes on the general form

$$\rho \frac{DC}{Dt} = -\nabla \cdot \mathbf{J}(C), \quad (22.1)$$

where \mathbf{J} is a flux that embodies molecular diffusion as well as subgrid scale advection and subgrid scale diffusion (Chapter 24). Advectional transport appears when transforming to an Eulerian or laboratory reference frame, in which case

$$\rho \frac{DC}{Dt} = \frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\mathbf{v} \rho C), \quad (22.2)$$

with $\mathbf{v} \rho C$ the advective flux. 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.

22.2 Diffusion physics

The Continuum Hypothesis summarized in Appendix 13 proposes that a macroscopic description of fluid motion does not require 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.

22.2.1 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 14.1, 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 some dye 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 15.4. 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

¹Recall from Chapter 19 that specific entropy remains materially constant on fluid parcels in the absence of mixing.

²For our purposes, turbulence is characterized by a quasi-random motion of fluid elements.

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.

22.2.2 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. 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.

22.2.3 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 22.1. 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

$$\mathbf{J} = -\kappa \rho \nabla C. \quad (22.3)$$

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*

$$\kappa \quad \text{kinematic diffusivity with SI units m s}^{-2} \quad (22.4)$$

$$\rho \kappa \quad \text{dynamic diffusivity with SI units kg m}^{-2} \text{ s}^{-2}. \quad (22.5)$$

The kinematic diffusivity has units of squared length per time and it sets the efficiency or strength of the diffusion. The diffusive flux (22.3) is known as Fick's law of matter diffusion. It is the most

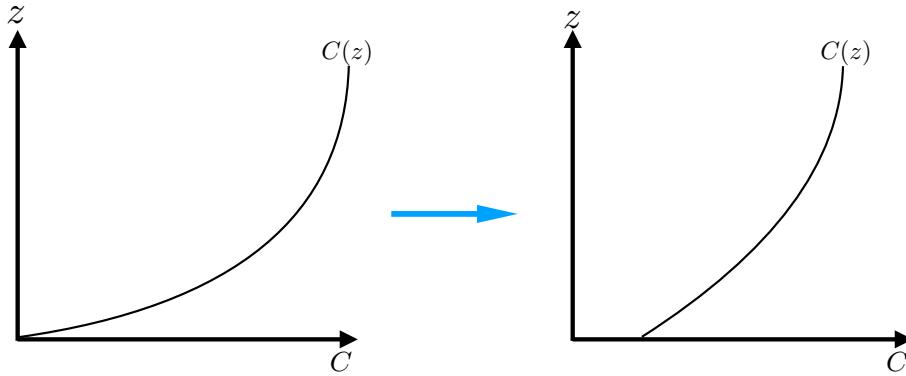


Figure 22.1: Shown here is a line graph illustrating the concentration, C , of a tracer drawn as a function of the space coordinate z , with the left panel showing the concentration at an earlier time than the right panel. Across any arbitrary point, transport of matter through random motions generally 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 deeper regions and low concentration fluid into shallow regions. The concentration is uniform in equilibrium, leading to a flat concentration profile.

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.

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 13.2.3), and the root-mean-square molecular speed, v_{rms} (see Section 13.2.4). 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

$$\rho \frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{J} = \nabla \cdot (\kappa \rho \nabla C). \quad (22.6)$$

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

$$\frac{\partial C}{\partial t} = \rho^{-1} \nabla(\kappa \rho) \cdot \nabla C + \kappa \nabla^2 C. \quad (22.7)$$

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.

22.2.4 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

$$\mathbf{J} = -\gamma \rho \nabla T, \quad (22.8)$$

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.

22.2.5 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 22.2, with shear (i.e., nonzero velocity gradient) in a single direction, Newtonian frictional stress takes the form

$$\tau = \rho \mu \frac{\partial u}{\partial z}, \quad (22.9)$$

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 26 and the nature of stress Chapter 28.

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

$$\text{Pr} = \frac{\mu}{\kappa}. \quad (22.10)$$

Theories for the turbulent Prandtl number are largely empirical in nature, with first principles arguments elusive.

22.2.6 Further reading

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 \(2017\)](#).

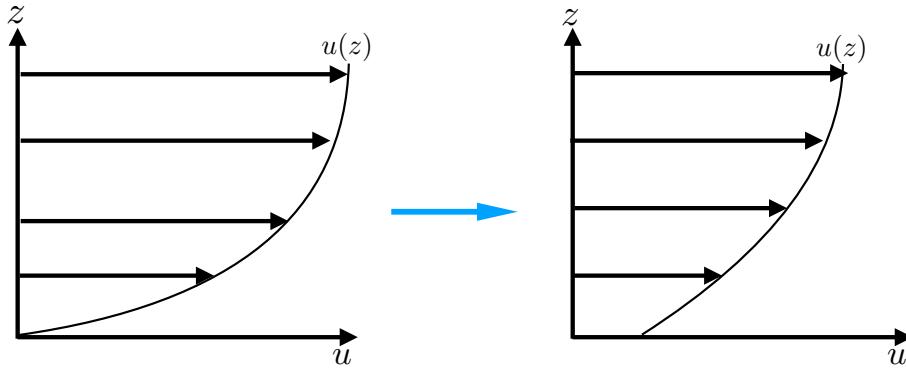


Figure 22.2: Shown here is a line graph illustrating the velocity, u , as a function of the space coordinate z , with the left panel showing the velocity at an earlier time than the right panel. Across any arbitrary point, transport of momentum through random motions generally reduces the magnitude of the velocity gradient; i.e., the diffusive transport leads to a viscous stress that acts to reduce the velocity shear.

22.3 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 20.3). 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

$$J_m = -\rho \mathbb{K}_{mn} \frac{\partial C}{\partial x^n}, \quad (22.11)$$

and the corresponding diffusion equation is

$$\rho \frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{J} = \frac{\partial(\rho \mathbb{K}_{mn} \partial_n C)}{\partial x_m}. \quad (22.12)$$

22.3.1 Sample diffusion tensors

For the isotropic case of molecular diffusion considered in equation (22.3), the diffusion tensor takes on the form

$$\mathbb{K}_{mn} = \kappa \delta_{mn} \quad \text{isotropic diffusion.} \quad (22.13)$$

If we rotate the diffusive fluxes to be along surfaces of constant γ , then

$$\mathbb{K}_{mn} = \kappa (\delta_{mn} - \hat{\gamma}_m \hat{\gamma}_n) \quad \text{rotated diffusion,} \quad (22.14)$$

where

$$\hat{\gamma} = \frac{\nabla \gamma}{|\nabla \gamma|} \quad (22.15)$$

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 20.3.5 for a discussion of neutral directions).

22.3.2 Diffusion of concentration powers

For many applications, it is of interest to determine how diffusion acts on powers of the tracer concentration. For this purpose consider the Eulerian time tendency of C^α , where $\alpha \geq 1$ is a power³

$$\rho \frac{\partial C^\alpha}{\partial t} = \alpha C^{\alpha-1} \rho \frac{\partial C}{\partial t} \quad (22.16a)$$

$$= \alpha C^{\alpha-1} \partial_m \left[\rho \mathbb{K}_{mn} \frac{\partial C}{\partial x^n} \right] \quad (22.16b)$$

$$= \partial_m \left[\rho \mathbb{K}_{mn} \frac{\partial C^\alpha}{\partial x^n} \right] - \rho \alpha (\alpha - 1) C^{\alpha-2} \frac{\partial C}{\partial x^n} \mathbb{K}_{mn} \frac{\partial C}{\partial x^m} \quad (22.16c)$$

$$= -\nabla \cdot \mathbf{J}(C^\alpha) + \alpha (\alpha - 1) C^{\alpha-2} \mathbf{J} \cdot \nabla C. \quad (22.16d)$$

The first term in equation (22.16d) is the convergence of the diffusive flux defined in terms of C^α . It therefore acts to diffuse C^α just like diffusion acts on C . The second term in equation (22.16d) is negative since the diffusion tensor is symmetric and positive-definite

$$\mathbf{J} \cdot \nabla C = -\rho \mathbb{K}_{mn} \frac{\partial C}{\partial x^m} \frac{\partial C}{\partial x^n} < 0. \quad (22.17)$$

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 (22.16d) always acts to reduce the magnitude of C^α towards zero.

22.3.3 Global integrals

We next consider the evolution of global integrals of tracer concentration 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 $\hat{\mathbf{n}}$ is the outward normal at the boundary. Integrating equation (22.16d) over the global domain thus leads to

$$\int \frac{\partial C^\alpha}{\partial t} \rho dV = \alpha (\alpha - 1) \int C^{\alpha-2} \mathbf{J} \cdot \nabla C dV. \quad (22.18)$$

We can bring the time derivative outside of the integral on the left hand side if we additionally assume the boundaries are material, thus rendering the contribution just from diffusion within the fluid interior

$$\frac{d}{dt} \int C^\alpha \rho dV = \alpha (\alpha - 1) \int C^{\alpha-2} \mathbf{J} \cdot \nabla C dV. \quad (22.19)$$

The case of $\alpha = 0$ is trivial, since $C^{\alpha=0} = 1$. The case of $\alpha = 1$ reflects the global conservation of tracer content for the case of insulating and material boundaries. The case of $\alpha = 2$ along with the downgradient orientation of the diffusive flux means that tracer variance is reduced

$$\frac{d}{dt} \int C^2 \rho dV \leq 0. \quad (22.20)$$

Likewise, all even powers of tracer concentration have their global integrated values reduced through downgradient diffusion

$$\frac{d}{dt} \int C^{2n} \rho dV \leq 0 \quad \text{for integers } n \geq 1. \quad (22.21)$$

In contrast, integrals of odd powers of tracer concentration are sign indefinite.

³We do not consider $\alpha < 1$ since there are singularities for C^α in regions of zero tracer concentration.

22.3.4 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

$$\mathcal{F} \equiv -(1/2) \int |\nabla C|^2 \rho d^3x \quad (22.22)$$

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

$$d^3x = dV \quad (22.23)$$

for the volume element. The reason will become apparent at equation (22.30).

Derivative of the diffusion dissipation functional

We introduce the *diffusion dissipation functional*

$$\mathcal{F} = \int \mathcal{L} d^3x \quad (22.24)$$

where the integrand is the quadratic form

$$2\mathcal{L} = \mathbf{J} \cdot \nabla C = -\rho \partial_m C \mathbb{K}_{mn} \partial_n C \leq 0. \quad (22.25)$$

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 variations to the tracer field δC into the dissipation functional

$$\delta\mathcal{F} = \int \left[\delta C \frac{\delta\mathcal{L}}{\delta C} + \delta(\partial_m C) \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right] d^3x. \quad (22.26)$$

Integration by parts on the second term leads to

$$\delta\mathcal{F} = \int \left[\delta C \frac{\delta\mathcal{L}}{\delta C} + \partial_m \left(\delta C \frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) - \delta C \partial_m \left(\frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) \right] d^3x. \quad (22.27)$$

The middle term is a total derivative that integrates to a boundary contribution and the associated *natural boundary condition*

$$\hat{\mathbf{n}} \cdot \frac{\delta\mathcal{L}}{\delta \nabla C} = \hat{\mathbf{n}} \cdot \mathbf{J} = \text{boundary flux}, \quad (22.28)$$

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

$$\delta\mathcal{F} = \int \delta C \left[\frac{\delta\mathcal{L}}{\delta C} - \partial_m \left(\frac{\delta\mathcal{L}}{\delta(\partial_m C)} \right) \right] d^3x. \quad (22.29)$$

Consequently, the functional derivative is given by

$$(d^3y)^{-1} \frac{\delta \mathcal{F}}{\delta C(\mathbf{y})} = \frac{\delta \mathcal{L}}{\delta C} - \partial_m \left[\frac{\delta \mathcal{L}}{\delta (\partial_m C)} \right], \quad (22.30)$$

where d^3y is the volume element at the field point \mathbf{y} . To reach the last step required the identity

$$\frac{\delta C(\mathbf{x})}{\delta C(\mathbf{y})} = d^3y \delta(\mathbf{x} - \mathbf{y}), \quad (22.31)$$

where $\delta(\mathbf{x} - \mathbf{y})$ is the Dirac delta function satisfying

$$\int \delta(\mathbf{x} - \mathbf{y}) d^3y = 1, \quad (22.32)$$

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 (22.31).

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

$$(d^3y)^{-1} \frac{\delta \mathcal{F}}{\delta C(\mathbf{y})} = -\partial_m \left[\frac{\delta \mathcal{L}}{\delta (\partial_m C)} \right] = \partial_m (\rho \mathbb{K}_{mn} \partial_n C). \quad (22.33)$$

The second equality identifies the diffusion operator, thus revealing the connection between the dissipation functional, the diffusion fluxes, and the diffusion operator

$$(d^3y)^{-1} \frac{\delta \mathcal{F}}{\delta C(\mathbf{y})} = -\nabla \cdot \mathbf{J}(\mathbf{y}) = \mathcal{R}. \quad (22.34)$$

22.4 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 22.2, 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.

22.4.1 The advection equation

In the absence of molecular diffusion, 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 also remains constant and thus satisfies the source-free *advection equation*

$$\frac{DC}{Dt} = \frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C = 0. \quad (22.35)$$

The first equality relates the material time derivative to the Eulerian time derivative plus advective transport (see Section 14.4), with \mathbf{v} the velocity of a fluid element. We can convert the *material* form of the advection equation (22.35) into its flux-form by combining with the mass continuity equation (15.9)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (22.36)$$

which yields

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v}) = 0. \quad (22.37)$$

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

$$C(\mathbf{x}, t) = C[\mathbf{X}(0)], \quad (22.38)$$

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.

22.4.2 Eulerian time tendencies from advection

At a point in the fluid, the advection equation (22.35) leads to the Eulerian time tendency for tracer concentration

$$\frac{\partial C}{\partial t} = -\mathbf{v} \cdot \nabla C. \quad (22.39)$$

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 (22.37), the density-weighted tracer concentration (the tracer mass per volume) has an Eulerian time tendency given by the convergence of the advective flux

$$\frac{\partial(\rho C)}{\partial t} = -\nabla \cdot (\rho C \mathbf{v}). \quad (22.40)$$

The tendency vanishes at a point if there is no convergence of tracer mass towards the point.

22.5 Advection maths

We now explore various mathematical properties of the advection equation. For that purpose, recall the mass continuity equation (22.36) and flux-form tracer advection equation (22.40)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (22.41a)$$

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v}) = 0. \quad (22.41b)$$

These equations are manifestly compatible in that the tracer equation (22.41b) reduces to the continuity equation (22.41a) if the tracer concentration is spatially uniform (see Section 15.4.5 for more discussion of compatibility).

22.5.1 Material constancy of C^α

A trivial consequence of the material constancy of tracer concentration is that C^α is also materially constant, for α an arbitrary number. We show this property mathematically by noting that the chain rule holds for a material time derivative, so that

$$\frac{DC^\alpha}{Dt} = n C^{\alpha-1} \frac{DC}{Dt} = 0. \quad (22.42)$$

Likewise, making use of the Eulerian form yields

$$\frac{\partial C^\alpha}{\partial t} + \mathbf{v} \cdot \nabla C^\alpha = \alpha C^{\alpha-1} \left[\frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C \right] = 0. \quad (22.43)$$

Advection thus serves to reversibly transport the tracer concentration without altering any of its powers.

22.5.2 Mass transport from the mean, eddy, and residual-mean

The mass density time tendency

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\mathbf{v} \rho) \quad (22.44)$$

remains unchanged if the advective mass flux, $\rho \mathbf{v}$ (units of mass per time per area), is modified by the addition of a total curl

$$\rho \mathbf{v} \rightarrow \rho \mathbf{v}^\dagger \quad (22.45a)$$

$$= \rho \mathbf{v} + \nabla \wedge (\rho \Psi). \quad (22.45b)$$

As in Section 16.7, the arbitrariness manifest in equation (22.45b) is known as a *gauge symmetry*. Furthermore, when integrated over a fixed region, the additional mass flux, $\nabla \wedge (\rho \Psi)$, leads to no addition nor subtraction of mass from a region. That is, the non-divergent mass flux does not alter the net mass within the fluid.

The non-divergent mass flux

$$\rho \mathbf{v}^* \equiv \nabla \wedge (\rho \Psi) \quad (22.46)$$

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:

$$\mathbf{v} = \text{Eulerian mean velocity} \quad (22.47a)$$

$$\rho \mathbf{v} = \text{Eulerian mean mass flux} \quad (22.47b)$$

$$\mathbf{v}^* = \text{eddy-induced velocity} \quad (22.47c)$$

$$\rho \mathbf{v}^* = \nabla \wedge (\rho \Psi) = \text{eddy-induced mass flux} \quad (22.47d)$$

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^* = \text{residual mean velocity} \quad (22.47e)$$

$$\rho \mathbf{v}^\dagger = \rho (\mathbf{v} + \mathbf{v}^*) = \text{residual mean mass flux}. \quad (22.47f)$$

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.

22.5.3 Advective 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

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho C \mathbf{v}^\dagger) = 0. \quad (22.48)$$

Given the form (22.46) for the eddy mass flux $\rho \mathbf{v}^*$, we can write the advective tracer flux as

$$\rho C \mathbf{v}^\dagger = C(\rho \mathbf{v} + \rho \mathbf{v}^*) \quad (22.49a)$$

$$= C \rho \mathbf{v} + C \nabla \wedge (\rho \Psi) \quad (22.49b)$$

$$= C \rho \mathbf{v} + \nabla \wedge (C \rho \Psi) - \nabla C \wedge \rho \Psi. \quad (22.49c)$$

It is the divergence of the tracer flux that determines the time tendency, in which case the total curl plays no role

$$\nabla \cdot (\rho C \mathbf{v}^\dagger) = \nabla \cdot (\rho C \mathbf{v} + \rho C \mathbf{v}^*) \quad (22.50a)$$

$$= \nabla \cdot (\rho C \mathbf{v} - \nabla C \wedge \rho \Psi). \quad (22.50b)$$

That is, the divergence of the advective mass flux equals to the divergence of the skew tracer flux

$$\underbrace{\nabla \cdot (\rho C \mathbf{v}^*)}_{\text{advective flux divergence}} = \underbrace{\nabla \cdot (-\nabla C \wedge \rho \Psi)}_{\text{skew flux divergence}} \quad (22.51)$$

since the advective flux and skew flux differ by a rotational flux

$$\mathbf{J}^{\text{adv}} = \mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}} \quad (22.52)$$

where

$$\mathbf{J}^{\text{adv}} = C \rho \mathbf{v}^* \quad \mathbf{J}^{\text{skew}} = -\nabla C \wedge \rho \Psi \quad \mathbf{J}^{\text{rot}} = \nabla \wedge (\rho C \Psi). \quad (22.53)$$

Notably, the skew tracer flux is neither upgradient nor downgradient. Rather, it is oriented parallel to iso-surfaces of tracer concentration

$$\nabla C \cdot \mathbf{J}^{\text{skew}} = \nabla C \cdot (-\nabla C \wedge \rho \Psi) = 0. \quad (22.54)$$

This orientation serves as motivation for the name “skew.” Figure 22.3 provides a schematic of the skew tracer fluxes.

22.5.4 Skew diffusion

Introducing tensor labels brings the skew tracer flux into the form

$$J_m^{\text{skew}} = -(\nabla C \wedge \rho \Psi)_m \quad (22.55a)$$

$$= -\epsilon_{mnp} \frac{\partial C}{\partial x^n} \rho \Psi_p \quad (22.55b)$$

$$= -\rho \mathbb{A}_{mn} \frac{\partial C}{\partial x^n}, \quad (22.55c)$$

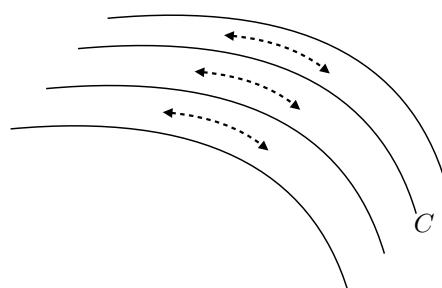


Figure 22.3: Skew fluxes (dashed lines with arrows) for a tracer C are oriented parallel to lines of constant tracer concentration (tracer isolines are the solid lines).

where we defined the anti-symmetric *skew diffusion* tensor

$$\mathbb{A}_{mn} = \epsilon_{mnp} \Psi_p = \begin{bmatrix} 0 & \Psi_3 & -\Psi_2 \\ -\Psi_3 & 0 & \Psi_1 \\ \Psi_2 & -\Psi_1 & 0 \end{bmatrix}. \quad (22.56)$$

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 22.6.2.

In Section 22.6 we pursue the above points to further detail the dual relation between advection and skewson.

22.5.5 Further reading

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.

22.6 Advection and skewson[†]

We introduced skew diffusion in Section 22.5.4 and will again encounter it in Chapter 23. 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 24.1. We here consider facets of advection and skewson for those interested in diving deeper into the details.

22.6.1 Choosing a gauge

Consider an arbitrary divergent-free mass transport

$$\nabla \cdot (\rho \mathbf{v}^*) = 0, \quad (22.57)$$

where the divergent-free constraint is satisfied by introducing a vector streamfunction

$$\rho \mathbf{v}^* = \nabla \wedge (\rho \Psi). \quad (22.58)$$

The streamfunction is arbitrary up to a gauge transformation

$$\rho \Psi' = \rho \Psi + \nabla(\rho \Lambda), \quad (22.59)$$

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

$$\nabla C \wedge [\rho \Psi + \nabla(\rho \Lambda)] = \nabla C \wedge (\rho \Psi) + \nabla \wedge [C \nabla(\rho \Lambda)]. \quad (22.60)$$

and since $\nabla \cdot \nabla \wedge [C \nabla(\rho \Lambda)] = 0$, the flux divergence, $\nabla \cdot \mathbf{J}^{\text{skew}}$, remains unchanged.

Coulomb gauge

Another gauge commonly used in electrodynamics is the *Coulomb gauge*, in which we set

$$\nabla \cdot \Psi = 0 \quad \text{Coulomb gauge.} \quad (22.61)$$

Making use of the curl identity (3.36c) leads to the Poisson equation for the vector potential

$$\nabla^2(\rho \Psi) = -\nabla \wedge (\rho \mathbf{v}^*). \quad (22.62)$$

This equation has a free-space Green's function given by the Coulomb-Ampere expression

$$\rho(\mathbf{x}, t) \Psi(\mathbf{x}, t) = \nabla \wedge \int \frac{\rho(\mathbf{x}', t) \mathbf{v}^*(\mathbf{x}', t)}{4\pi |\mathbf{x} - \mathbf{x}'|} dV', \quad (22.63)$$

where dV' is the volume element for integration over the test points, \mathbf{x}' . We know of no GFM application making use of the Coulomb gauge.

Vertical gauge

One gauge of use for eddy parameterizations (Section 24.1) sets to zero one of the three components of the vector streamfunction. This gauge choice is available since there are only those two independent functional degrees of freedom available from a divergence-free mass transport field. A common choice is the *vertical gauge* in which

$$\Psi_3 = 0 \quad \text{vertical gauge.} \quad (22.64)$$

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 $\nabla(\rho\Lambda) = -\hat{z}\rho\Phi_3$. This choice in turn means that the third component of Ψ is zero.

To further specify the vertical gauge we invert the relations

$$\rho u^* = -\partial_z(\rho\Psi_2) \quad \rho v^* = \partial_z(\rho\Psi_1) \quad \rho w^* = \partial_x(\rho\Psi_2) - \partial_y(\rho\Psi_1) \quad (22.65)$$

to render the vector streamfunction

$$\rho\Psi = \hat{z} \wedge \int_{-H}^z \rho \mathbf{u}^* dz' = \hat{z} \wedge \underline{\mathbf{U}}^{(*\rho)} \quad (22.66)$$

where

$$\underline{\mathbf{U}}^{(*\rho)}(z) = \int_{-H}^z \rho \mathbf{u}^* dz' \quad (22.67)$$

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

$$\rho \mathbb{A}_{mn} = \begin{pmatrix} 0 & 0 & \underline{U}^{(*\rho)} \\ 0 & 0 & \underline{V}^{(*\rho)} \\ -\underline{U}^{(*\rho)} & -\underline{V}^{(*\rho)} & 0 \end{pmatrix}, \quad (22.68)$$

and the corresponding skew, rotational, and advective fluxes are

$$\mathbf{J}^{\text{skew}} = -\underline{\mathbf{U}}^{(*\rho)} \partial_z C + \hat{z} \underline{\mathbf{U}}^{(*\rho)} \cdot \nabla_z C \quad (22.69a)$$

$$\mathbf{J}^{\text{rot}} = \partial_z(C \underline{\mathbf{U}}^{(*\rho)}) - \hat{z} \nabla_z \cdot (C \underline{\mathbf{U}}^{(*\rho)}) \quad (22.69b)$$

$$\mathbf{J}^{\text{adv}} = C (\partial_z \underline{\mathbf{U}}^{(*\rho)}) - \hat{z} C \nabla_z \cdot \underline{\mathbf{U}}^{(*\rho)}. \quad (22.69c)$$

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.

22.6.2 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 15.7), so that \mathbf{v}^* satisfies the no-normal flow condition even on moving boundaries

$$\hat{\mathbf{n}} \cdot \mathbf{v}^* = 0. \quad (22.70)$$

Hence, the advective flux for tracers also has a no-normal boundary condition on all boundaries

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{adv}} = \hat{\mathbf{n}} \cdot \mathbf{v}^* \rho C = 0. \quad (22.71)$$

The corresponding boundary condition for the skew flux is found by inserting the relation (22.52) into the advective flux boundary condition (22.71) to render

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{adv}} = \hat{\mathbf{n}} \cdot [\mathbf{J}^{\text{skew}} + \mathbf{J}^{\text{rot}}] = 0. \quad (22.72)$$

Hence, the skew flux generally has a non-zero normal component at the solid boundaries as determined by the rotational flux

$$\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{skew}} = -\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{rot}}. \quad (22.73)$$

Even so, we may have occasions in which $\hat{\mathbf{n}} \cdot \mathbf{J}^{\text{skew}} = 0$, which is ensured so long as

$$(-\nabla C \wedge \rho \Psi) \cdot \hat{\mathbf{n}} = -(\rho \Psi \wedge \hat{\mathbf{n}}) \cdot \nabla C = 0. \quad (22.74)$$

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 vanishes identically at the boundary. Further details for boundary conditions depend on physical properties of the velocity \mathbf{v}^* . We discuss one example in Section 24.1 as prescribed by the [Gent et al. \(1995\)](#) eddy parameterization scheme.

22.7 Exercises

EXERCISE 22.1: 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

$$\tau_{\text{diffusion}} = \frac{\Delta^2}{\kappa} \quad (22.75)$$

where Δ is the length scale and κ is the kinematic diffusivity (units 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 22.2: ANALYTICAL SOLUTION TO ONE-DIMENSIONAL DIFFUSION EQUATION

Consider a one-dimensional diffusion equation

$$\frac{\partial C}{\partial t} = \kappa \frac{\partial^2 C}{\partial z^2}, \quad (22.76)$$

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

$$\frac{\partial C}{\partial t} = \kappa \frac{\partial^2 C}{\partial z^2}, \quad (22.77a)$$

no-flux boundary condition from part (b) (22.77b)

$$C(z, t=0) = C_0 \cos(Kz), \quad (22.77c)$$

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 22.3: 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

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial z} \left[\kappa \frac{\partial C}{\partial z} \right] \quad (22.78a)$$

$$= \frac{\partial \kappa}{\partial z} \frac{\partial C}{\partial z} + \kappa \frac{\partial^2 C}{\partial z^2}, \quad (22.78b)$$

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

$$\frac{\partial C^*}{\partial t} = \kappa \frac{\partial^2 C^*}{\partial z^2}. \quad (22.79)$$

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

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x^m} \left[K^{mn} \frac{\partial C}{\partial x^n} \right]. \quad (22.80)$$

Now consider an extrema in the tracer field, which is defined by

$$\frac{\partial C^*}{\partial x^n} = 0 \quad \forall n = 1, 2, 3. \quad (22.81)$$

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 22.4: ONE-DIMENSIONAL ADVECTION

Consider the advection equation in one space dimension without boundaries

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = 0 \quad (22.82a)$$

$$C(x, z, t=0) = C_0 \cos(k x) \quad (22.82b)$$

$$u(z, t) = \alpha z \cos(\omega t). \quad (22.82c)$$

The specified zonal velocity is non-divergent, oscillatory in time, and vertically sheared

$$\frac{\partial u}{\partial z} = \alpha \cos(\omega t), \quad (22.83)$$

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 (22.38).

EXERCISE 22.5: SKEW FLUX FOR OCEAN MESOSCALE EDDIES

Consider a middle-latitude mesoscale ocean eddy respecting geostrophic balance (see Section 30.3) on an f -plane (constant Coriolis parameter) and incompressibility. In this case, the horizontal eddy-induced velocity at the ocean surface is non-divergent

$$\mathbf{u}^* = \nabla \wedge \hat{\mathbf{z}} \psi. \quad (22.84)$$

In this equation, the geostrophic streamfunction is given by

$$\psi = -\hat{\mathbf{z}} \frac{g \eta}{f}, \quad (22.85)$$

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 (22.56).
- (b) Determine the skew tracer flux (22.69a).

EXERCISE 22.6: 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.

- (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 22.4. Derive the following expression for the area enclosed by the two contours

$$\mathcal{A} = \int_{C_0}^{C_1} dC \oint \frac{dl}{|\nabla C|}. \quad (22.86)$$

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 $\nabla C \neq 0$.

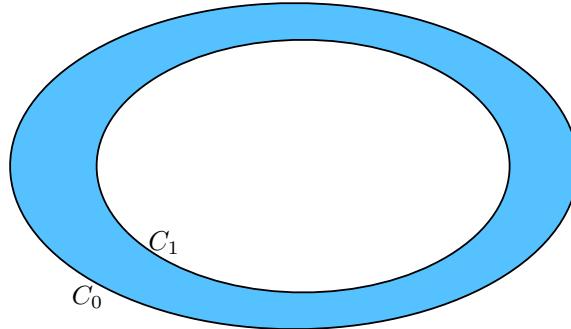


Figure 22.4: Illustrating the area contained between two closed tracer contours, $C_0 \leq C(x, y, t) \leq C_1$. Exercise 22.6 develops some mathematical expressions for integration within this area, with the resulting expressions of use for the analyses of tracer transport.

- (b) As a corollary, show that for

$$\mathcal{A}(C) = \int_{C_0}^C dC' \oint \frac{dl}{|\nabla C'|} \quad (22.87)$$

we have the identity

$$\frac{\partial \mathcal{A}(C)}{\partial C} = \oint \frac{dl}{|\nabla C|}. \quad (22.88)$$

In words, this result means that the 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

$$\frac{\partial}{\partial C} \left[\int \Phi(\mathbf{x}) d\mathcal{A} \right] = \frac{\partial}{\partial C} \left[\int_{C_0}^C dC' \oint \frac{\Phi dl}{|\nabla C'|} \right] \quad (22.89a)$$

$$= \oint \frac{\Phi dl}{|\nabla C|}, \quad (22.89b)$$

with Φ an arbitrary function. This is a truly remarkable identity with many useful applications such as those in [Marshall et al. \(2006\)](#).

EXERCISE 22.7: INTEGRATION BETWEEN TWO SMOOTH TRACER ISO-SURFACES

Consider a volume bounded by two smooth iso-tracer surfaces at an instance in time, such as shown in Figure 22.5. We assume that the three-dimensional gradient of the tracer remains non-zero within the region of interest, $|\nabla \lambda| \neq 0$. This assumption is commonly met for geophysical tracers.

- (a) Derive the following expression for the volume enclosed by the two iso-surfaces

$$V = \int_{\lambda_0}^{\lambda_1} d\lambda \int \frac{dS}{|\nabla \lambda|}. \quad (22.90)$$

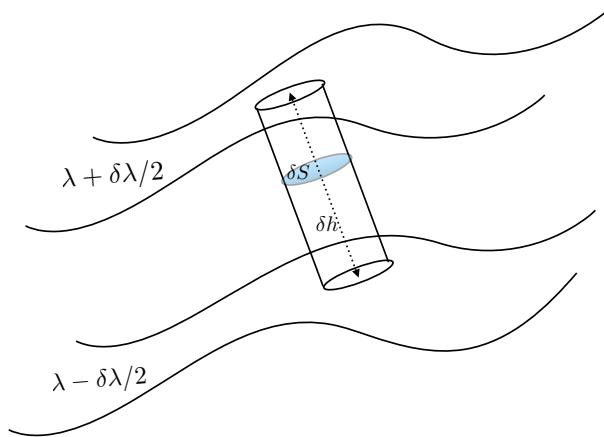


Figure 22.5: This schematic shows a region bounded by two tracer iso-surfaces, $\lambda_0 \leq \lambda(x, y, z, t) \leq \lambda_1$. In this figure we assume the iso-surfaces are monotonically layered in one of the three directions (e.g., stably stratified density in the vertical). However, this assumption is not needed for the results of this exercise. The cylinder region is a representative pillbox extending between the two iso-surfaces, with thickness δh and cross-sectional area δS . The pillbox is oriented according to the normal direction $\hat{n} = |\nabla\lambda|^{-1} \nabla\lambda$, where we assume that $|\nabla\lambda| \neq 0$. The ability to orient the pillbox allows for a unique integration of the volume between the two iso-surfaces even if the surfaces are corrugated or overturning (e.g., unstably stratified density). Exercise 22.7 develops some mathematical expressions for integrations computed between the two iso-surfaces, with the resulting expressions of use for analyses of tracer transport and water mass transformation.

In this expression, dS is the area element for the cross-section of a “pillbox” region extending from one iso-surface to the other (see Figure 22.5), and where $|\nabla\lambda| \neq 0$ by assumption. As expected, regions of large λ gradients have smaller volumes due to the more tightly packed λ surfaces.

- (b) As a corollary, show that

$$\frac{\partial V(\lambda)}{\partial \lambda} = \int \frac{dS}{|\nabla\lambda|}. \quad (22.91)$$

This equation provides a measure of the volume per unit λ .

- (c) Use the above two results to prove the following form of the Fundamental Theorem of Calculus

$$\frac{\partial}{\partial \lambda} \left[\int \Phi(\mathbf{x}) dV \right] = \frac{\partial}{\partial \lambda} \left[\int_{\lambda_0}^{\lambda} d\lambda' \int \frac{\Phi dS}{|\nabla\lambda'|} \right] = \int \frac{\Phi dS}{|\nabla\lambda|}. \quad (22.92)$$

with Φ an arbitrary function.

- (d) Show that the following approximation holds

$$\frac{\partial}{\partial \lambda} \left[\int \Phi(\mathbf{x}) dV \right] \approx \frac{1}{\Delta\lambda} \left[\int_{\lambda-\Delta\lambda/2}^{\lambda+\Delta\lambda/2} d\lambda' \int \frac{\Phi dA}{|\nabla\lambda'|} \right] \equiv \frac{1}{\Delta\lambda} \int_{\lambda-\Delta\lambda/2}^{\lambda+\Delta\lambda/2} \Phi(\mathbf{x}) dV. \quad (22.93)$$

EXERCISE 22.8: 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

$$\langle \mathbf{x} \rangle^C = \frac{\int \mathbf{x} C dV}{\int C dV}, \quad (22.94)$$

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

$$\mathcal{V} = \int dV \quad C = \int C dV. \quad (22.95)$$

Furthermore, when computing the time derivative acting on the integral, make use of the kinematic results from Section 15.6.3, in which for any integrand φ

$$\frac{d}{dt} \int \varphi dV = \int \frac{\partial \varphi}{\partial t} dV. \quad (22.96)$$

- (a) Consider a tracer concentration whose tendency at a point in space is affected only by advection

$$\frac{\partial C}{\partial t} = -\nabla \cdot (\mathbf{v} C), \quad (22.97)$$

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

$$\frac{d\langle \mathbf{x} \rangle^C}{dt} = \langle \mathbf{v} \rangle^C, \quad (22.98)$$

where the tracer center of mass velocity is given by

$$\langle \mathbf{v} \rangle^C = \frac{\int \mathbf{v} C dV}{\int C dV} = \frac{1}{C} \int \mathbf{v} C dV. \quad (22.99)$$

- (b) Consider a tracer concentration whose tendency at a point in space affected only by diffusion

$$\frac{\partial C}{\partial t} = \nabla \cdot (K \nabla C), \quad (22.100)$$

where $K = K(\mathbf{x}) > 0$ is a diffusivity that is assumed to vanish at the domain boundaries. Show that the tracer center of mass drifts up the diffusivity gradient

$$\frac{d\langle \mathbf{x} \rangle^C}{dt} = \langle \nabla K \rangle^C. \quad (22.101)$$

- (c) Consider an initial tracer concentration that is a function only of latitude,

$$C(x, y, z, t=0) = C_0(y), \quad (22.102)$$

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 22.9: 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. The fluid is assumed to have a tracer whose concentration is affected by an irreversible process so that

$$\frac{DC}{Dt} = \dot{C} \neq 0. \quad (22.103)$$

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

$$\langle \mathbf{x} \rangle^C = \frac{\int_{\mathcal{R}} \mathbf{x} C \rho dV}{\int_{\mathcal{R}} C \rho dV}. \quad (22.104)$$

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 15.6.5 since the region has a constant mass. Consequently, we can set

$$\frac{D}{Dt} \int_{\mathcal{R}} \psi \rho dV = \int_{\mathcal{R}} \frac{D\psi}{Dt} \rho dV. \quad (22.105)$$

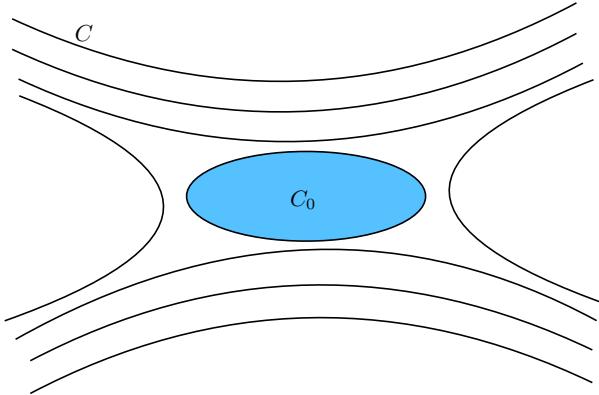


Figure 22.6: In a steady state flow, the tracer concentration within a constant C contour is constant. Diffusion has thus acted to remove all variations in tracer concentration within the region. In this figure, the concentration within the closed region has constant value $C = C_0$, whereas the region with open tracer contours remains non-homogeneous.

EXERCISE 22.10: DIFFUSIVE HOMOGENIZATION OF SCALARS INSIDE CLOSED CONTOURS

The advection-diffusion equation for a tracer concentration is given by

$$\frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho \mathbf{v} C) = -\nabla \cdot \mathbf{J} \quad (22.106)$$

where

$$\mathbf{J} = -\rho \mathbf{K} \cdot \nabla C \quad (22.107)$$

is a downgradient diffusive flux with \mathbf{K} a symmetric positive-definite diffusion tensor. In the steady state, the divergence of the advective tracer flux balances the convergence of the diffusive flux

$$\nabla \cdot (\rho \mathbf{v} C) = -\nabla \cdot \mathbf{J}. \quad (22.108)$$

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 22.6. 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.

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 24).

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 15.4 and the maths and physics of the advection-diffusion equation explored in Chapter 22. We focus most discussion on incompressible flows discussed in Chapter 16 and applicable to the Boussinesq fluid commonly assumed for the ocean (Chapter 29). 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 23.4 and 23.6) are posed using the isopycnal vertical coordinates detailed in Chapter 8 and further pursued in Chapter 31.

23.1	Reynold's decomposition	326
23.2	Basics of the GLM	327
23.2.1	Motivation	328
23.2.2	Length scales and the small parameter	329
23.2.3	Decomposing the particle trajectory	329
23.2.4	GLM and the Stokes mean	330
23.2.5	An example linear wave	331
23.2.6	GLM with a materially constant scalar	333
23.2.7	Further reading	334
23.3	Kinematics of eddy tracer fluxes	334
23.3.1	Particle displacements and eddy tracer fluxes	334
23.3.2	Symmetric and skew-symmetric tracer fluxes	335
23.3.3	Massaging the mean field tracer equation	338
23.3.4	Connection to Stokes drift	338
23.3.5	A linear rotating periodic wave example	339
23.3.6	Further reading	340
23.4	Kinematics of volume transport in isopycnal layers	341
23.4.1	Isopycnal mean	341
23.4.2	Modified mean	342
23.4.3	Transformed residual mean (TRM)	343
23.4.4	Volume conservation and the thickness equation	344
23.4.5	Ensemble mean kinematics in isopycnal coordinates	346
23.4.6	Ensemble mean kinematics in geopotential coordinates	347
23.4.7	Approximate ensemble mean kinematics in geopotential coordinates	347
23.4.8	Comments	349
23.5	Thickness transport for a linear longitudinal wave	349
23.5.1	An undulating fluid layer	349
23.5.2	Stokes drift	350
23.5.3	Linearized thickness perturbations	351
23.5.4	Correlation between thickness and velocity	351
23.5.5	Comments	352
23.6	Mean tracer equation	352
23.6.1	Thickness weighted means	352
23.6.2	Isopycnal mean thickness weighted tracer equation	352
23.6.3	Subgrid scale tracer transport tensor	353
23.6.4	Mean tracer transport beneath a density surface	353
23.6.5	Summary of the tracer parameterization problem	354
23.6.6	Comments	355

23.1 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, $\Phi'(\mathbf{x}, t)$

$$\Phi(\mathbf{x}, t) = \bar{\Phi}(\mathbf{x}, t) + \Phi'(\mathbf{x}, t). \quad (23.1)$$

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”

$$\overline{\Phi'} = 0 \quad (23.2a)$$

$$\overline{\overline{\Phi}} = \overline{\Phi} \quad (23.2b)$$

$$\overline{\gamma \overline{\Phi}} = \gamma \overline{\Phi} \quad \text{for } \gamma \text{ a constant.} \quad (23.2c)$$

Equation (23.2a) says that the mean of an eddy fluctuation vanishes. The equality (23.2b) says that the mean of a mean field returns the mean field. The final equality, (23.2c), 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.

23.2 Basics 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*

$$\text{Lagrangian mean} = \text{Eulerian mean} + \text{Stokes mean}. \quad (23.3)$$

This name is motivated from the *Stokes drift* introduced in Section 14.8, which we again encounter in Section 23.2.4. 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.

23.2.1 Motivation

Consider a materially constant scalar field

$$\frac{D\Phi}{Dt} = \frac{\partial\Phi}{\partial t} + \mathbf{v} \cdot \nabla\Phi = 0. \quad (23.4)$$

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 23.1 leads to the mean field equation

$$\frac{\partial\bar{\Phi}}{\partial t} + \bar{\mathbf{v}} \cdot \nabla\bar{\Phi} = -\overline{\mathbf{v}' \cdot \nabla\Phi'}. \quad (23.5)$$

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

$$\frac{\partial\Phi(\mathbf{a}, t)}{\partial t} = 0. \quad (23.6)$$

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 (23.6), so too does the Lagrangian mean

$$\frac{\partial\bar{\Phi}(\bar{\mathbf{a}}, t)}{\partial t} = 0. \quad (23.7)$$

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

$$\frac{\partial\bar{\Phi}^{(L)}}{\partial t} + \bar{\mathbf{v}}^{(L)} \cdot \nabla\bar{\Phi}^{(L)} = 0. \quad (23.8)$$

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 (23.8), we will motivate the GLM average from the analysis of small amplitude eddying motions.

23.2.2 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

$$|\xi| \ll \lambda \quad \text{small amplitude waves.} \quad (23.9)$$

We thus introduce the small non-dimensional ratio of length scales for the following analysis

$$\alpha = \frac{|\xi|}{\lambda} \ll 1. \quad (23.10)$$

23.2.3 Decomposing the particle trajectory

Recall the discussion of fluid particle trajectories given in Chapter 14. In this description, the trajectory of a particle is determined by integrating the relation between the particle trajectory and the particle velocity

$$\left[\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} \right]_{\mathbf{a}} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t)] \implies \mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, 0) + \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t')] dt', \quad (23.11)$$

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

$$\overline{\xi(\mathbf{x}, t)} = 0. \quad (23.12)$$

Note that the Eulerian mean operator can be any of the operators (or others) satisfying the Reynold's decomposition property discussed in Section 23.1

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 23.1, for which the particle displacement amplitude is much smaller than the wavelength of the disturbance

$$\alpha = \frac{|\xi|}{\lambda} \ll 1. \quad (23.13)$$

In this case the disturbance field is constructed by time integration of the eddy velocity field

$$\left[\frac{\partial \xi(\mathbf{x}, t)}{\partial t} \right]_{\mathbf{x}} = \mathbf{v}'(\mathbf{x}, t) \implies \xi(\mathbf{x}, t) = \int \mathbf{v}'(\mathbf{x}, t') dt'. \quad (23.14)$$

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

$$\nabla \cdot \mathbf{v}' = 0 \implies \nabla \cdot \boldsymbol{\xi} = 0. \quad (23.15)$$

The definition (23.14) 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 (23.11). 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.

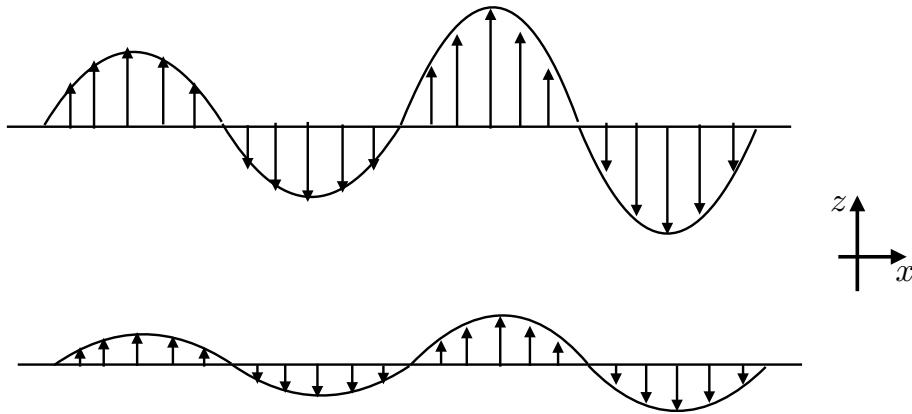


Figure 23.1: Illustrating the displacement of fluid particles at two selective vertical positions due to a linear transverse wave disturbance. The particle positions, $\mathbf{x}^{(\xi)} = \mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t)$, have a disturbance field of the form $\boldsymbol{\xi}(\mathbf{x}, t) = \hat{z} \xi_0(x, z) \sin(kx - \omega t)$, 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 23.2.5), whereas Stokes drift requires particle motion to have a nonzero component in the wave direction (see Figures 14.8 and 14.9). Even so, it does generally produce a Stokes mean for an arbitrary field Φ (Section 23.2.4).

23.2.4 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

$$\underbrace{\overline{\Phi(\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t))}}_{\text{GLM}} \neq \underbrace{\overline{\Phi(\mathbf{x}, t)}}_{\text{Eulerian}}, \quad (23.16)$$

where it is common to make use of the shorthand¹

$$\mathbf{x}^{(\xi)}(\mathbf{x}, t) \equiv \mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t) \quad (23.17)$$

for the instantaneous position of the fluid particle. The average

$$\overline{\Phi}^{(L)}(\mathbf{x}, t) \equiv \overline{\Phi(\mathbf{x} + \boldsymbol{\xi}(\mathbf{x}, t), t)} = \overline{\Phi(\mathbf{x}^{(\xi)})} \quad (23.18)$$

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

$$\overline{\Phi}^{(E)}(\mathbf{x}, t) \equiv \overline{\Phi}(\mathbf{x}, t). \quad (23.19)$$

Following our discussion at the start of Section 23.2, we define the difference between the GLM and Eulerian mean as the Stokes mean

$$\overline{\Phi}^{(S)}(\mathbf{x}, t) = \overline{\Phi}^{(L)}(\mathbf{x}, t) - \overline{\Phi}^{(E)}(\mathbf{x}, t). \quad (23.20)$$

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

$$\Phi(\mathbf{x} + \boldsymbol{\xi}, t) = \Phi(\mathbf{x}, t) + \boldsymbol{\xi} \cdot \nabla \Phi(\mathbf{x}, t) + \frac{1}{2} \xi_m \xi_n \partial_m \partial_n \Phi(\mathbf{x}, t) + \mathcal{O}(\alpha^3). \quad (23.21)$$

The non-dimensional ratio $\alpha = |\boldsymbol{\xi}|/\lambda \ll 1$ was introduced in equation (23.13), 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 (23.21) then leads to an expression for the Stokes mean

$$\overline{\Phi}^{(S)}(\mathbf{x}, t) = \overline{\Phi}^{(L)}(\mathbf{x}, t) - \overline{\Phi}^{(E)}(\mathbf{x}, t) \quad (23.22a)$$

$$= \overline{\boldsymbol{\xi} \cdot \nabla \Phi} + \frac{1}{2} \overline{\xi_m \xi_n \partial_m \partial_n \Phi} + \mathcal{O}(\alpha^3). \quad (23.22b)$$

$$= \overline{\boldsymbol{\xi} \cdot \nabla \Phi'} + \frac{1}{2} \overline{\xi_m \xi_n} \partial_m \partial_n \overline{\Phi}^{(E)} + \mathcal{O}(\alpha^3), \quad (23.22c)$$

where we introduced the Eulerian fluctuation

$$\Phi'(\mathbf{x}, t) = \Phi(\mathbf{x}, t) - \overline{\Phi}^{(E)}(\mathbf{x}, t) \quad (23.23)$$

and all terms on the right hand side of equation (23.22c) are evaluated at (\mathbf{x}, t) . The Stokes drift (Section 14.8) associated with the GLM arises from setting Φ equal to one of the velocity components

$$\overline{v}_p^{(S)} = \overline{\boldsymbol{\xi} \cdot \nabla v'_p} + \frac{1}{2} \overline{\xi_m \xi_n} \partial_m \partial_n \overline{v}_p^{(E)} + \mathcal{O}(\alpha^3). \quad (23.24)$$

23.2.5 An example linear wave

We exemplify the previous discussion by considering the small amplitude linear wave

$$\boldsymbol{\xi} = -\frac{\mathbf{U}(\mathbf{x})}{\omega} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (23.25a)$$

$$\mathbf{v}' = \partial_t \boldsymbol{\xi} = \mathbf{U}(\mathbf{x}) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (23.25b)$$

$$\nabla v'_p = \nabla U_p \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \mathbf{k} U_p \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (23.25c)$$

$$\nabla \cdot \mathbf{v}' = (\nabla \cdot \mathbf{U}) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \mathbf{k} \cdot \mathbf{U} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (23.25d)$$

where \mathbf{U} is the velocity amplitude that is generally a function of space, \mathbf{k} is the wavevector, and

$$T = 2\pi/\omega \quad (23.26)$$

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

$$\bar{\boldsymbol{\xi}}(\mathbf{x}) = \frac{1}{T} \int_0^T \boldsymbol{\xi}(\mathbf{x}, t') dt' = 0 \quad (23.27a)$$

$$\bar{\mathbf{v}}'(\mathbf{x}) = \frac{1}{T} \int_0^T \mathbf{v}'(\mathbf{x}, t') dt' = 0. \quad (23.27b)$$

To maintain a non-divergent eddy velocity at arbitrary times requires

$$\nabla \cdot \mathbf{v}' = 0 \implies \nabla \cdot \mathbf{U} = \mathbf{U} \cdot \mathbf{k} = 0. \quad (23.28)$$

The second condition means that the wave is transverse, so that particle displacements are orthogonal to the wavevector (e.g., Figure 23.1)

Stokes drift

Specializing to the velocity field (23.25b), substituting into the Stokes drift expression (23.24), and making use of an average over a wave period yields

$$\overline{\boldsymbol{\xi} \cdot \nabla v_p'} = \frac{U_p \mathbf{U} \cdot \mathbf{k}}{2\omega} \quad (23.29a)$$

$$\bar{v}_p^{(E)} = 0. \quad (23.29b)$$

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

$$\bar{\mathbf{v}}^{(S)} = \frac{\mathbf{U} (\mathbf{U} \cdot \mathbf{k})}{2\omega} + \mathcal{O}(\alpha^2). \quad (23.30)$$

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

$$\bar{v}^{(S)} = \frac{U^2}{2c} + \mathcal{O}(\alpha^2). \quad (23.31)$$

This result agrees with that derived using Lagrangian trajectories in Section 14.8 (see Exercise 14.2). 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

$$\bar{\Phi}^{(S)}(\mathbf{x}, t) = -\omega^{-1} \mathbf{U} \cdot \overline{\nabla \Phi' \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)} + \mathcal{O}(\alpha^2) \quad (23.32a)$$

$$= -\omega^{-1} \overline{\nabla \cdot (\mathbf{U} \Phi')} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) + \mathcal{O}(\alpha^2), \quad (23.32b)$$

where the second equality made use of the non-divergent nature of the wave field (23.28). 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 23.1. 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.

23.2.6 GLM with a materially constant scalar

Consider a materially constant scalar field, such as a tracer concentration in the absence of mixing and sources

$$\frac{DC}{Dt} = 0. \quad (23.33)$$

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

$$\bar{C}^{(L)}(\mathbf{x}, t) = C(\mathbf{x} + \boldsymbol{\xi}, t). \quad (23.34)$$

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 (23.34) when considering isopycnal kinematics in Sections 23.4 and 23.6.

Relating the particle disturbance field to Eulerian properties of C

There is a frequently used consequence of the identity (23.34) involving the disturbance field, the Eulerian fluctuation

$$C'(\mathbf{x}, t) = C(\mathbf{x}, t) - \bar{C}(\mathbf{x}, t) \quad (23.35)$$

and the Eulerian mean

$$C^{(E)}(\mathbf{x}, t) = \bar{C}(\mathbf{x}, t). \quad (23.36)$$

To derive it, recall the Taylor series expansion (23.21) truncated here to first order accuracy

$$C(\mathbf{x} + \boldsymbol{\xi}, t) = C(\mathbf{x}, t) + \boldsymbol{\xi} \cdot \nabla \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (23.37)$$

Taking the Eulerian mean of both sides renders

$$\overline{C(\mathbf{x} + \boldsymbol{\xi}, t)} = \bar{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (23.38)$$

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 (23.22). From equation (23.34) we know that $C(\mathbf{x} + \boldsymbol{\xi}, t) = \overline{C(\mathbf{x} + \boldsymbol{\xi}, t)}$, so that we can subtract equations (23.37) and (23.38) to find

$$C'(\mathbf{x}, t) = -\boldsymbol{\xi} \cdot \nabla \overline{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2). \quad (23.39)$$

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 23.3. Furthermore, for the isopycnal kinematics in Sections 23.4 and 23.6, we focus on vertical particle displacements, $\boldsymbol{\xi} = \xi \hat{\mathbf{z}}$, in which case the Eulerian fluctuation is given by

$$C'(z, t) = -\xi \partial_z \overline{C}(z, t) + \mathcal{O}(\alpha^2). \quad (23.40)$$

23.2.7 Further reading

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 \(2014\)](#).

23.3 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

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{v} C) = 0, \quad (23.41)$$

and its Eulerian mean is

$$\frac{\partial \overline{C}}{\partial t} + \nabla \cdot (\overline{\mathbf{v}} \overline{C}) = -\nabla \cdot \overline{\mathbf{v}' C'}. \quad (23.42)$$

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, $\overline{\mathbf{v}' C'}$. 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 23.2 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.

23.3.1 Particle displacements and eddy tracer fluxes

Following Section 23.2, we here introduce a particle disturbance vector corresponding to small amplitude eddy fluctuations

$$\partial_t \boldsymbol{\xi}(\mathbf{x}, t) = \mathbf{v}'(\mathbf{x}, t) + \mathcal{O}(\alpha^2) \quad (23.43a)$$

$$\overline{\boldsymbol{\xi}} = 0. \quad (23.43b)$$

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 23.2.6, to leading order we can write the Eulerian fluctuation in terms of the particle displacement (equation (23.39))

$$C'(\mathbf{x}, t) = -\boldsymbol{\xi} \cdot \nabla \overline{C}(\mathbf{x}, t) + \mathcal{O}(\alpha^2) \quad (23.44)$$

Notice that if the particle displacement is oriented along a mean tracer iso-surface, then $\xi \cdot \nabla \bar{C}(\mathbf{x}, t) = 0$ 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

$$\mathbf{v}' C' = -\partial_t \xi (\xi \cdot \nabla) \bar{C} + \mathcal{O}(\alpha^2). \quad (23.45)$$

We unpack this expression for the purpose of characterizing kinematic properties of the eddy tracer flux

23.3.2 Symmetric and skew-symmetric tracer fluxes

From equation (23.45), the m 'th component of the eddy tracer flux is given by

$$v'_m C' = -[(\partial_t \xi_m) \xi_n] \partial_n \bar{C}. \quad (23.46)$$

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

$$2(\partial_t \xi_m) \xi_n = [(\partial_t \xi_m) \xi_n + (\partial_t \xi_n) \xi_m] + [(\partial_t \xi_m) \xi_n - (\partial_t \xi_n) \xi_m] \quad (23.47a)$$

$$= \partial_t(\xi_m \xi_n) + [(\partial_t \xi_m) \xi_n - (\partial_t \xi_n) \xi_m]. \quad (23.47b)$$

Introducing the symmetric and anti-symmetric correlation tensors

$$2\mathbb{K}_{mn} \equiv \overline{\partial_t(\xi_m \xi_n)} \quad (23.48a)$$

$$2\mathbb{A}_{mn} \equiv \overline{(\partial_t \xi_m) \xi_n} - \overline{(\partial_t \xi_n) \xi_m} \quad (23.48b)$$

allows us to write the mean eddy tracer flux

$$\overline{v'_m C'} = -(\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} \quad (23.49)$$

and the mean field tracer equation (23.42)

$$\frac{\partial \bar{C}}{\partial t} + \nabla \cdot (\bar{\mathbf{v}} \bar{C}) = \nabla \cdot [(\mathbb{K} + \mathbb{A}) \cdot \nabla \bar{C}]. \quad (23.50)$$

The right hand side of this equation equals to the convergence of the symmetric and skew-symmetric tracer fluxes³

$$\nabla \cdot [(\mathbb{K} + \mathbb{A}) \cdot \nabla \bar{C}] = -\nabla \cdot (\mathbf{F}^{(sym)} + \mathbf{F}^{(skew)}), \quad (23.51)$$

where

$$\mathbf{F}^{(sym)} = -\mathbb{K} \cdot \nabla \bar{C} \quad (23.52a)$$

$$\mathbf{F}^{(skew)} = -\mathbb{A} \cdot \nabla \bar{C} \quad (23.52b)$$

$$\overline{\mathbf{v}' C'} = \mathbf{F}^{(sym)} + \mathbf{F}^{(skew)}. \quad (23.52c)$$

²See Section 17.2.4 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 (23.52a) is given by

$$F_m^{(\text{sym})} = -\mathbb{K}_{mn} \partial_n \bar{C} = -\frac{1}{2} \overline{\partial_t(\xi_m \xi_n)} \partial_n \bar{C}. \quad (23.53)$$

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 23.3.5). 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 24.3.

The skew, advective, and rotational fluxes

Following our discussion in Section 22.6, we write the skew flux as

$$F_m^{(\text{skew})} = -\mathbb{A}_{mn} \partial_n \bar{C} = -\epsilon_{mnp} \Psi_p \partial_n \bar{C} = -(\nabla \bar{C} \wedge \Psi)_m, \quad (23.54)$$

where we introduced the vector streamfunction (dimensions squared length per time)⁴

$$\Psi = \frac{1}{2} \overline{\partial_t \xi \wedge \xi} = \frac{1}{2} \overline{\mathbf{v}' \wedge \xi}. \quad (23.55)$$

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

$$\mathbf{F}^{(\text{skew})} = -\nabla \bar{C} \wedge \Psi \quad (23.56a)$$

$$= (\nabla \wedge \Psi) \bar{C} - \nabla \wedge (\bar{C} \Psi) \quad (23.56b)$$

$$= \mathcal{U}^{(\Psi)} \bar{C} - \nabla \wedge (\bar{C} \Psi) \quad (23.56c)$$

$$= \mathbf{F}^{(\text{adv})} - \mathbf{F}^{(\text{rot})}, \quad (23.56d)$$

so that the skew flux equals to an advective flux minus a rotational flux. We here introduced the non-divergent velocity

$$\mathcal{U}^{(\Psi)} = \nabla \wedge \Psi \quad (23.57)$$

and the non-divergent rotational flux

$$\mathbf{F}^{(\text{rot})} = \nabla \wedge (\bar{C} \Psi). \quad (23.58)$$

The divergence of the skew flux equals to the divergence of the advective flux

$$\nabla \cdot \mathbf{F}^{(\text{skew})} = \nabla \cdot (\mathbf{F}^{(\text{adv})} - \mathbf{F}^{(\text{rot})}) = \nabla \cdot \mathbf{F}^{(\text{adv})}, \quad (23.59)$$

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 (23.52c), we see that a point measurement of the correlation $\overline{\mathbf{v}' C'}$ provides an estimate of the diffusive and skew diffusive tracer fluxes

$$\overline{\mathbf{v}' C'} = \mathbf{F}^{(\text{sym})} + \mathbf{F}^{(\text{skew})} = -(\mathbb{K} + \mathbb{A}) \cdot \nabla \overline{C}. \quad (23.60)$$

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

$$\overline{\mathbf{v}' C'} = -\mathbb{K} \cdot \nabla \overline{C} - \nabla \overline{C} \wedge \Psi \quad (23.61a)$$

$$= -\mathbb{K} \cdot \nabla \overline{C} - \nabla \wedge (\overline{C} \Psi) + \overline{C} \nabla \wedge \Psi \quad (23.61b)$$

$$= -\mathbb{K} \cdot \nabla \overline{C} - \nabla \wedge (\overline{C} \Psi) + \overline{C} \mathcal{U}^{(\Psi)}. \quad (23.61c)$$

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}

$$\mathcal{T} = \overline{\int_{\mathcal{S}} \mathbf{v} C \cdot \hat{\mathbf{n}} dS} = \int_{\mathcal{S}} \overline{\mathbf{v} C} \cdot \hat{\mathbf{n}} dS = \int_{\mathcal{S}} [\overline{\mathbf{v}} \overline{C} + \overline{\mathbf{v}' C'}] \cdot \hat{\mathbf{n}} dS. \quad (23.62)$$

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 (23.60). Introducing the diffusive, advective, and rotational flux then renders

$$\mathcal{T} = \int_{\mathcal{S}} [\overline{\mathbf{v}} \overline{C} + \mathcal{U}^{(\Psi)} \overline{C} - \nabla \wedge (\overline{C} \Psi) - \mathbb{K} \cdot \nabla \overline{C}] \cdot \hat{\mathbf{n}} dS. \quad (23.63)$$

Use of Stokes' Theorem transforms the rotational term to a line integral around the boundary of the area

$$\mathcal{T} = \int_{\mathcal{S}} [\overline{\mathbf{v}} \overline{C} + \mathcal{U}^{(\Psi)} \overline{C} - \mathbb{K} \cdot \nabla \overline{C}] \cdot \hat{\mathbf{n}} dS - \oint_{\partial \mathcal{S}} \overline{C} \Psi \cdot d\mathbf{l}. \quad (23.64)$$

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

$$\oint_{\partial \mathcal{S}} \overline{C} \Psi \cdot d\mathbf{l} = (1/2) \oint_{\partial \mathcal{S}} \overline{C} (\overline{\mathbf{v}' \wedge \xi}) \cdot d\mathbf{l}. \quad (23.65)$$

We further this interpretation when considering the transport beneath a fluctuating isopycnal surface in Section 23.6.4.

23.3.3 Massaging the mean field tracer equation

We here write the mean tracer equation (23.50) in various forms that can be found throughout the literature. For this purpose, write the right hand side of equation (23.50) in the form

$$\nabla \cdot [(\mathbb{K} + \mathbb{A}) \cdot \nabla \bar{C}] = \partial_m [(\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C}] \quad (23.66a)$$

$$= \partial_m (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} + (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_m \partial_n \bar{C} \quad (23.66b)$$

$$= \partial_m (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} + \mathbb{K}_{mn} \partial_m \partial_n \bar{C}. \quad (23.66c)$$

The final equality follows from the identity

$$\mathbb{A}_{mn} \partial_m \partial_n \bar{C} = 0, \quad (23.67)$$

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 (23.66c) can be interpreted as an advection operator through the action of a non-divergent plus a divergent advection velocity

$$\partial_m (\mathbb{K}_{mn} + \mathbb{A}_{mn}) \partial_n \bar{C} = [\mathcal{U}^{(K)} + \mathcal{U}^{(\Psi)}] \cdot \nabla \bar{C}, \quad (23.68)$$

where⁵

$$\mathcal{U}_n^{(K)} \equiv -\partial_m \mathbb{K}_{mn} \implies \nabla \cdot \mathcal{U}^{(K)} = -\partial_n \partial_m \mathbb{K}_{mn} \neq 0 \quad (23.69a)$$

$$\mathcal{U}_n^{(\Psi)} \equiv -\partial_m \mathbb{A}_{mn} \implies \nabla \cdot \mathcal{U}^{(\Psi)} = \partial_n \partial_m \mathbb{A}_{mn} = 0. \quad (23.69b)$$

Bringing the above results together allows us to write the mean field tracer equation (23.50) in the following equivalent forms

$$\frac{\partial \bar{C}}{\partial t} + [\bar{\mathbf{v}} + \mathcal{U}^{(\Psi)} + \mathcal{U}^{(K)}] \cdot \nabla \bar{C} = \mathbb{K}_{mn} \partial_m \partial_n \bar{C} \quad \text{advective form} \quad (23.70a)$$

$$\frac{\partial \bar{C}}{\partial t} + \nabla \cdot (\bar{\mathbf{v}} + \mathcal{U}^{(\Psi)}) \bar{C} = \nabla \cdot (\mathbb{K} \cdot \nabla \bar{C}) \quad \text{flux form,} \quad (23.70b)$$

where we made use of the identities

$$\nabla \cdot \bar{\mathbf{v}} = 0 \quad (23.71a)$$

$$\nabla \cdot \mathcal{U}^{(\Psi)} = 0 \quad (23.71b)$$

$$\nabla \cdot \mathcal{U}^{(K)} \neq 0. \quad (23.71c)$$

23.3.4 Connection to Stokes drift

From equation (23.24) we have the leading order expression for the Stokes drift

$$\bar{v}_p^{(S)} = \bar{\xi}_n \partial_n \partial_t \bar{\xi}_p + \mathcal{O}(\alpha^2). \quad (23.72)$$

As noted in equation (23.15), 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,

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 (23.69a), whereas they define $\mathcal{U}_n^{(\Psi)} = -\partial_m \mathbb{A}_{mn}$ as in equation (23.69b).

the Stokes drift velocity can be written

$$\bar{v}_p^{(S)} = \overline{\xi_n \partial_n \partial_t \xi_p} \quad (23.73a)$$

$$= \partial_n [(\partial_t \xi_p) \xi_n] \quad (23.73b)$$

$$= \partial_n (\mathbb{K}_{pn} + \mathbb{A}_{pn}) \quad (23.73c)$$

$$= \partial_n (\mathbb{K}_{np} - \mathbb{A}_{np}) \quad (23.73d)$$

$$= -\mathcal{U}_p^{(K)} + \mathcal{U}_p^{(\Psi)}. \quad (23.73e)$$

For the case of periodic waves, the Stokes drift velocity equals to the non-divergent skew velocity

$$\bar{\mathbf{v}}^{(S)} = \mathcal{U}^{(\Psi)} \Rightarrow \nabla \cdot \bar{\mathbf{v}}^{(S)} = 0 \quad \text{periodic waves.} \quad (23.74)$$

More generally, for non-periodic waves the divergent velocity is non-zero so that the Stokes velocity is also divergent

$$\bar{\mathbf{v}}^{(S)} = \mathcal{U}^{(\Psi)} - \mathcal{U}^{(K)} \Rightarrow \nabla \cdot \bar{\mathbf{v}}^{(S)} = -\nabla \cdot \mathcal{U}^{(K)} \neq 0 \quad \text{non-periodic waves.} \quad (23.75)$$

23.3.5 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

$$\boldsymbol{\xi}(\mathbf{x}, t) = \Gamma [\hat{\mathbf{x}} \cos(kx - \omega t) + \hat{\mathbf{y}} \sin(kx - \omega t)] \quad (23.76a)$$

$$\partial_t \boldsymbol{\xi}(\mathbf{x}, t) = \omega \Gamma [\hat{\mathbf{x}} \sin(kx - \omega t) - \hat{\mathbf{y}} \cos(kx - \omega t)], \quad (23.76b)$$

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

$$\boldsymbol{\xi} \cdot \boldsymbol{\xi} = \Gamma^2. \quad (23.77)$$

We are motivated to let the mean operator be a phase average

$$\bar{\phi} = \frac{1}{T} \int_0^T \phi(t) dt, \quad (23.78)$$

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

$$\mathbb{K}_{mn} = \frac{\Gamma^2}{2T} \int_0^T dt \frac{\partial}{\partial t} \begin{bmatrix} \cos^2(kx - \omega t) & \cos(kx - \omega t) \sin(kx - \omega t) \\ \cos(kx - \omega t) \sin(kx - \omega t) & \sin^2(kx - \omega t) \end{bmatrix} \quad (23.79)$$

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.

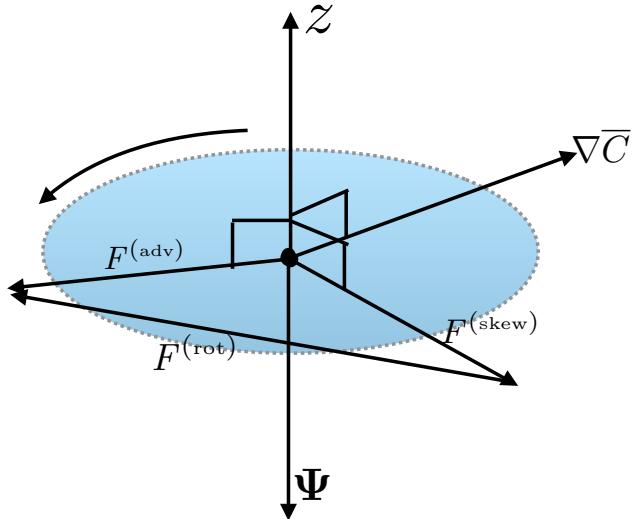


Figure 23.2: Sketch of the various tracer fluxes associated with the polarized displacement vector (23.76a). The particles are moving on the horizontal plane in a circle with time-independent radius Γ . The vector streamfunction (23.81) points in the negative \hat{z} direction. The mean concentration 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.

Skew-symmetric stirring tensor

In contrast, the skew-symmetric tensor has non-zero components

$$\mathbb{A}_{12} = -\mathbb{A}_{21} = \frac{\Gamma^2 \omega}{2} [\sin^2(kx - \omega t) + \cos^2(kx - \omega t)] = \frac{\Gamma^2 \omega}{2}, \quad (23.80)$$

which reflects the counter-clockwise polarization. The corresponding vector streamfunction is vertical

$$\Psi = \frac{\Gamma^2 \omega}{2} \hat{z}, \quad (23.81)$$

and the skew flux is horizontal

$$\mathbf{F}^{(\text{skew})} = \frac{\Gamma^2 \omega}{2} (\hat{z} \wedge \nabla \bar{C}). \quad (23.82)$$

Finally, the advective velocity is given by

$$\nabla \wedge \Psi = \omega \Gamma \nabla \Gamma \wedge \hat{z}, \quad (23.83)$$

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 23.2 offers a schematic of the skew, advective, and rotational fluxes induced by a linear rotating particle wave in the horizontal plane.

23.3.6 Further reading

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.

23.4 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 21.2.4 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 23.6, at which point we also consider eddy correlations between velocity and tracer as in Section 23.3.

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 23.2. 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.

23.4.1 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 ??, 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 8 and 31), with the mean field denoted by

$$\overline{\Phi}^{(\sigma)}(x, y, \sigma, t) \equiv \text{ensemble mean using isopycnal vertical coordinates.} \quad (23.84)$$

This average is straightforward to compute when using isopycnal coordinates, thus producing an isopycnal mean that is a function of the potential density, σ .

Figure 23.3: 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 $z(2, \rho)$.

23.4.2 Modified mean

As a complement to the isopycnal approach in Section 23.4.1, we here introduce the vertical/isopycnal GLM, also known as the modified mean.

Vertical/isopycnal GLM

The discussion in Section 23.2 considered a three-vector particle displacement vector $\xi(\mathbf{x}, t)$. In contrast, we are here interested just in the vertical displacement of an isopycnal layer interface

$$\xi(x, y, \sigma, t) = \hat{z}\xi(x, y, \sigma, t). \quad (23.85)$$

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)

$$z(\sigma) = \bar{z}^{(\sigma)} + \xi(\sigma), \quad (23.86)$$

where

$$\bar{z}^{(\sigma)} = \overline{z(\sigma)}^{(\sigma)} \quad (23.87)$$

is the isopycnal ensemble mean depth, and the displacement field has a zero ensemble mean

$$\overline{\xi(\sigma)}^{(\sigma)} = 0. \quad (23.88)$$

Given the above definitions for the vertical position, we define the *vertical/isopycnal GLM* for an arbitrary function

$$\tilde{\Phi}(x, y, \bar{z}^{(\sigma)}, t) \equiv \overline{\Phi(x, y, \bar{z}^{(\sigma)} + \xi(x, y, \sigma, t), t)}^{(\sigma)}. \quad (23.89)$$

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 (23.89) as the *modified mean*.

Relating the modified mean to the isopycnal mean

Following the general result (23.34), 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

$$\tilde{\sigma}(x, y, \bar{z}^{(\sigma)}, t) = \sigma(x, y, \bar{z}^{(\sigma)} + \xi(x, y, \sigma, t), t). \quad (23.90)$$

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 (23.84)), 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

$$\overline{\Phi}^{(\sigma)}(x, y, \tilde{\sigma}, t) = \tilde{\Phi}(x, y, \bar{z}^{(\sigma)}, t). \quad (23.91)$$

This is a very important identity that will be used in the following.

23.4.3 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

$$\widehat{\Phi}(\sigma) = \frac{\overline{h \Phi}^{(\sigma)}}{\overline{h}^{(\sigma)}}. \quad (23.92)$$

The identity (23.90) then renders

$$\overline{\Phi}^{\#}(x, y, \bar{z}^{(\sigma)}, t) \equiv \widehat{\Phi}(x, y, \tilde{\sigma}, t), \quad (23.93)$$

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

$$\widehat{\mathbf{u}}(x, y, \tilde{\sigma}, t) = \overline{\mathbf{u}}^{\#}(x, y, \bar{z}^{(\sigma)}, t). \quad (23.94)$$

Following the discussion of the vertical gauge in Section 22.6.1 (see in particular equation (22.67)), we are led to define the depth integrated TRM transport

$$\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)}) = \int_{-H}^{\bar{z}^{(\sigma)}} \overline{\mathbf{u}}^{\#}(z) dz = \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \widehat{\mathbf{u}}(\gamma) \overline{h}^{(\gamma)} d\gamma, \quad (23.95)$$

with the second equality following from a change of coordinates from geopotential to isopycnal. We can go further with this expression by writing

$$\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)}) = \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \widehat{\mathbf{u}}(\gamma) \overline{h}^{(\gamma)} d\gamma \quad \text{from equation (23.95)} \quad (23.96a)$$

$$= \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \overline{\mathbf{u}} h^{(\gamma)} d\gamma \quad \text{from equation (23.92)} \quad (23.96b)$$

$$= \int_{\sigma(-H)}^{\sigma(\bar{z}^{(\sigma)} + \xi)} \overline{\mathbf{u}} h^{(\gamma)} d\gamma \quad \text{from equation (23.90).} \quad (23.96c)$$

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

$$\overline{\mathbf{U}}^{\#}(\bar{z}^{(\sigma)}) = \int_{-H}^{\bar{z}^{(\sigma)} + \xi} \overline{\mathbf{u}} dz. \quad (23.97)$$

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 (23.97) can be decomposed into an Eulerian mean plus the correlation of a fluctuation

$$\bar{\mathbf{U}}^\#(\bar{z}^{(\sigma)}) \equiv \bar{\mathbf{U}}(\bar{z}^{(\sigma)}) + \bar{\mathbf{U}}^{\text{qs}}(\bar{z}^{(\sigma)}). \quad (23.98)$$

The first term,

$$\bar{\mathbf{U}}(\bar{z}^{(\sigma)}) = \int_{-H}^{\bar{z}^{(\sigma)}} \mathbf{u} dz \quad (23.99)$$

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

$$\bar{\mathbf{U}}^{\text{qs}}(\bar{z}^{(\sigma)}) \equiv \int_{\bar{z}^{(\sigma)}}^{\bar{z}^{(\sigma)} + \xi} \mathbf{u} dz \quad (23.100)$$

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 23.2)

$$\bar{\mathbf{U}}^{\text{qs}} = \bar{\mathbf{U}}^\# - \bar{\mathbf{U}}. \quad (23.101)$$

As for the traditional Stokes drift discussed in Sections 14.8, and 23.2.5, 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 (22.66), we introduce the TRM vector streamfunction

$$\bar{\Psi}^\# = \bar{\mathbf{U}}^\# \wedge \hat{\mathbf{z}}, \quad (23.102)$$

and the corresponding three-dimensional non-divergent TRM velocity

$$\bar{\mathbf{v}}^\# = \nabla \wedge \bar{\Psi}^\#. \quad (23.103)$$

The vertical component,

$$\bar{w}^\# = \hat{\mathbf{z}} \cdot (\nabla \wedge \bar{\Psi}^\#), \quad (23.104)$$

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, $\bar{\mathbf{U}}^\#$, so the two descriptions in effect make use of the same number of degrees of freedom.

23.4.4 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

$$\delta V = \delta x \delta y \delta z = \delta x \delta y \delta \sigma h, \quad (23.105)$$

where we introduced the specific thickness

$$h = \frac{\partial z}{\partial \sigma}. \quad (23.106)$$

As discussed in Section 8.9.1, 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 20.3.4)

$$N^2 = -\frac{g}{\rho_0} \frac{\partial \sigma}{\partial z} = -\frac{g}{\rho_0 h} \quad (23.107)$$

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 31.2.3)

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) = 0, \quad (23.108)$$

with \mathbf{u} the horizontal velocity field, the time derivative is computed with σ held fixed, and

$$\nabla_\sigma = \nabla_z + \mathbf{S} \frac{\partial}{\partial z} \quad (23.109)$$

is the horizontal derivative operator with σ held fixed and

$$\mathbf{S} = \nabla_\sigma z \quad (23.110)$$

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 22.6.1 (equation (22.64)), 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

$$\frac{\partial \sigma}{\partial t} = \nabla \cdot (\nabla \sigma \wedge \Psi), \quad (23.111)$$

where all derivatives are here taken with fixed Eulerian (geopotential) coordinates, (x, y, z) , and the divergence operator is three-dimensional.

Figure 23.4: 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 $\rho(2, z)$.

23.4.5 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 (23.108) is satisfied by each ensemble member

$$\frac{\partial h}{\partial t} + \nabla_\sigma \cdot (h \mathbf{u}) = 0. \quad (23.112)$$

The ensemble mean computed over these fluid parcels with potential density σ satisfies

$$\partial_t \bar{h}^{(\sigma)} + \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \bar{\mathbf{u}}^{(\sigma)} + \bar{h}' \bar{\mathbf{u}}'^{(\sigma)}) = 0, \quad (23.113)$$

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

$$\partial_t \bar{h}^{(\sigma)} + \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \hat{\mathbf{u}}) = 0. \quad (23.114)$$

In this equation we introduced the thickness weighted isopycnal ensemble mean horizontal velocity

$$\hat{\mathbf{u}} = \frac{\bar{h} \bar{\mathbf{u}}^{(\sigma)}}{\bar{h}^{(\sigma)}} = \bar{\mathbf{u}}^{(\sigma)} + \frac{\bar{h}' \bar{\mathbf{u}}'^{(\sigma)}}{\bar{h}^{(\sigma)}} \equiv \bar{\mathbf{u}}^{(\sigma)} + \mathbf{u}^{\text{bolus}}, \quad (23.115)$$

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

$$\bar{h}^{(\sigma)} \mathbf{u}^{\text{bolus}} = \bar{h}^{(\sigma)} (\hat{\mathbf{u}} - \bar{\mathbf{u}}^{(\sigma)}) = \bar{h}' \bar{\mathbf{u}}'^{(\sigma)} \quad (23.116)$$

arises from the along-isopycnal correlations between specific thickness and horizontal velocity.

Quite conveniently, the mean conservation equation (23.114) takes the *same* mathematical form as the conservation equation (23.112) 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 (23.114) is afforded by use of the Lagrangian vertical coordinate σ .

23.4.6 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 (23.89)), or TRM fields when thickness weighted (equation (23.93)).

Evolution of modified mean density

Following the skew-symmetric formulation from Section 22.6, 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

$$\bar{\mathbf{F}}^\# = -\nabla \tilde{\sigma} \wedge \bar{\Psi}^\#. \quad (23.117)$$

Using the identity $\bar{\Psi}^\# = \bar{\mathbf{U}}^\# \wedge \hat{\mathbf{z}}$, we can write this expression in one of the forms

$$\bar{\mathbf{F}}^\# = -\bar{\mathbf{U}}^\# \partial_z \tilde{\sigma} + \hat{\mathbf{z}} \bar{\mathbf{U}}^\# \cdot \nabla_z \tilde{\sigma} \quad (23.118a)$$

$$= -(\bar{\mathbf{U}}^\# + \hat{\mathbf{z}} \mathbf{S} \cdot \bar{\mathbf{U}}^\#) \partial_z \tilde{\sigma}, \quad (23.118b)$$

where

$$\mathbf{S} = -\frac{\nabla_z \tilde{\sigma}}{\partial_z \tilde{\sigma}} \quad (23.119)$$

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 $z = \bar{z}^{(\sigma)}$,

$$\frac{\partial \tilde{\sigma}}{\partial t} = \nabla \cdot (\nabla \tilde{\sigma} \wedge \bar{\Psi}^\#). \quad (23.120)$$

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 (23.111) satisfied at depth z by a single member of the ensemble.

23.4.7 Approximate ensemble mean kinematics in geopotential coordinates

Equation (23.120) 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 (23.100). We addressed a similar estimation in Section 23.2.4 when discussing the Stokes mean. Here, we expand the TRM transport in a Taylor series about the geopotential $z = \bar{z}^{(\sigma)}$

$$\bar{\mathbf{U}}^\#(z) = \overline{\int_{-H}^{z+\xi} \mathbf{u}(s) \, ds} \quad (23.121a)$$

$$= \bar{\mathbf{U}}(z) + \overline{\mathbf{u} \xi^{(z)}} + \frac{1}{2} \overline{\partial_z \mathbf{u} \xi \xi^{(z)}} + \mathcal{O}(\alpha^3), \quad (23.121b)$$

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 (23.121b) 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

$$\bar{\mathbf{u}}\xi^{(z)} = \bar{\mathbf{u}'}\xi^{(z)}. \quad (23.122)$$

For the second order term, similar considerations lead to

$$\partial_z \bar{\mathbf{u}}\xi\xi^{(z)} \approx \partial_z \bar{\mathbf{u}}^z \xi\xi^{(z)}, \quad (23.123)$$

where neglected terms are third order and higher. Combining these relations leads to the second order accurate expression

$$\bar{\mathbf{U}}^\# \approx \bar{\mathbf{U}} + \bar{\mathbf{u}'}\xi^{(z)} + \frac{1}{2}\bar{\xi}\xi^z \partial_z \bar{\mathbf{u}}^{(z)}. \quad (23.124)$$

The disturbance field

Following the discussion in Section 23.2.6, we here determine the disturbance field, ξ , in terms of fields at constant depth. For this purpose, use the identity (23.90) to give

$$\tilde{\sigma}(z) = \sigma(z + \xi) \quad (23.125a)$$

$$= \sigma(z) + \partial_z \sigma(z) \xi + \frac{1}{2} \partial_{zz} \sigma(z) \xi^2 + \mathcal{O}(\alpha^3). \quad (23.125b)$$

Subtracting the Eulerian mean of equation (23.125b) from the unaveraged equation (23.125b), and noting that $\tilde{\sigma}$ is already a mean field, leads to the second order accurate expression for the deviation

$$\xi = -\sigma'(z)/\partial_z \bar{\sigma}^{(z)} + \mathcal{O}(\alpha^2), \quad (23.126)$$

where

$$\sigma(z) = \bar{\sigma}^{(z)} + \sigma'(z). \quad (23.127)$$

To within the same order, the deviation can be written

$$\xi = -\sigma'(z)/\partial_z \tilde{\sigma}(z) + \mathcal{O}(\alpha^2). \quad (23.128)$$

Approximate quasi-Stokes transport

Substituting the deviation (23.128) into the approximate expression (23.121b) for the TRM transport yields an approximate expression for the Stokes transport

$$\mathbf{U}^{\text{qs}} = -\frac{\bar{\mathbf{u}'}\sigma'^{(z)}}{\partial_z \tilde{\sigma}} + \frac{\bar{\phi}^{(z)} \partial_z \bar{\mathbf{u}}^{(z)}}{(\partial_z \tilde{\sigma})^2} + \mathcal{O}(\alpha^3), \quad (23.129)$$

where

$$\bar{\phi}^{(z)} = \frac{1}{2} \bar{\sigma}' \sigma'^{(z)} \quad (23.130)$$

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 (23.129). We have more to say regarding this parameterization in Section 24.1.

Substituting the deviation (23.128) into the approximate expression (23.125b) yields, to within terms of third order, the relation

$$\tilde{\sigma} = \bar{\sigma}^{(z)} - \partial_z \left[\frac{\bar{\phi}^{(z)}}{\partial_z \bar{\sigma}^{(z)}} \right] + \mathcal{O}(\alpha^3). \quad (23.131)$$

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.

23.4.8 Comments

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 \(2014\)](#).

23.5 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 23.4.5. Indeed, much of our intuition for the general case is based on this relatively simple worked example.⁶

23.5.1 An undulating fluid layer

Figure 23.5 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

$$h(y, t) = h_o + h'(y, t). \quad (23.132)$$

The layer thickness changes in time according to the convergence of the advective transport of thickness

$$\frac{\partial h}{\partial t} = -\nabla \cdot (h \mathbf{u}), \quad (23.133)$$

where the convergence is computed within the layer. We offer a derivation of this volume conservation equation in Section 33.1.3. Even without working through that derivation, the truth of this equation follows from Figure 23.5, 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

$$\frac{\partial h}{\partial t} = -\frac{\partial (h v)}{\partial y}. \quad (23.134)$$

⁶This worked example is based Section 2 of [Lee et al. \(1997\)](#).

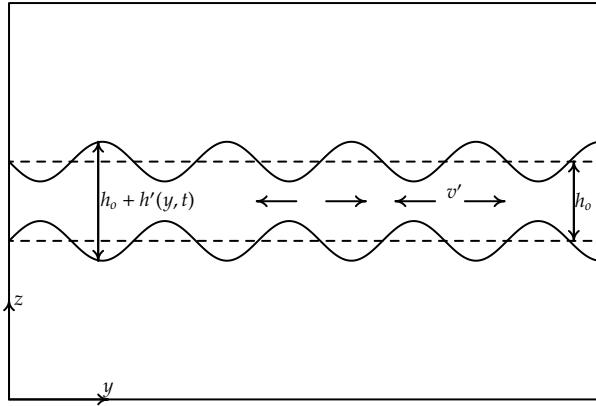


Figure 23.5: Shown here is a single layer of constant density fluid, with resting thickness $h = h_o$ and instantaneous thickness $h = h_o + h'(y, t)$. Associated with the undulations in thickness are fluctuations in the meridional velocity $v' = v_o \sin(k y - \omega t)$, depicted here by the alternating 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.

23.5.2 Stokes drift

Consider a linear wave perturbation in the meridional velocity that propagates in the meridional direction

$$v'(y, t) = v_o \sin(k y - \omega t), \quad (23.135)$$

where k is a constant wave number and ω is a constant frequency. This longitudinal wave is depicted in Figure 23.5. We now follow the general formalism developed in Section 14.8 or equivalently in Section 23.2.4 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

$$\frac{dY}{dt} = v_o \sin(k Y - \omega t), \quad (23.136)$$

where $Y = Y(Y_o, t)$ is the meridional trajectory with initial position Y_o . Following equation (14.95) 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)

$$\frac{dY}{dt} - v(y, t) = v_o^2 k \cos(ky - \omega t) \int_0^t \sin(ky - \omega t') dt' \quad (23.137a)$$

$$= \frac{v_o^2 k}{\omega} (\cos^2(ky - \omega t) - \cos(ky - \omega t) \cos(ky)). \quad (23.137b)$$

Time averaging over a single wave period,

$$T = 2\pi/\omega \quad (23.138)$$

leads to the Stokes drift as per the general expression in equation (14.97)

$$V_{\text{Stokes}} = \frac{v_o^2 k}{2\omega}. \quad (23.139)$$

Introducing the phase speed for the wave $c = \omega/k$ allows us to write the Stokes drift as

$$V_{\text{Stokes}} = \frac{v_o^2}{2c}. \quad (23.140)$$

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.

23.5.3 Linearized thickness perturbations

The velocity and thickness are written in terms of their rest state plus a perturbation due to the wave

$$h = h_o + h' \quad (23.141a)$$

$$v = v', \quad (23.141b)$$

where the velocity vanishes when the wave is absent. The thickness equation (23.134) thus takes the form

$$\frac{\partial h'}{\partial t} + h_o \frac{\partial v'}{\partial y} + v' \frac{\partial h'}{\partial y} = 0. \quad (23.142)$$

Linearizing this equation, and using the wave perturbation (23.135), leads to

$$\frac{\partial h'}{\partial t} + h_o v_o k \cos(ky - \omega t) = 0, \quad (23.143)$$

thus yielding the thickness perturbation

$$h' = h_o \frac{v'}{c}. \quad (23.144)$$

Hence, to leading order, the thickness perturbation is directly proportional to and in phase with the velocity perturbation.

23.5.4 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

$$\overline{h' v'} = \frac{1}{T} \int_0^T h' v' dt \quad (23.145a)$$

$$= \frac{h_o}{c T} \int_0^T v' v' dt \quad (23.145b)$$

$$= \frac{v_o^2 h_o}{c T} \int_0^T \sin^2(ky - \omega t) dt \quad (23.145c)$$

$$= \frac{v_o^2 h_o}{2c} \quad (23.145d)$$

$$= h_o V_{\text{Stokes}}, \quad (23.145e)$$

where we introduced the Stokes drift (23.140) 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.

23.5.5 Comments

The nonzero correlation in equation (23.145e) 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 23.4.5. For the general case, a nonzero bolus velocity (23.116), 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.

23.6 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.

23.6.1 Thickness weighted means

Equation (23.115) 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

$$\Phi(\sigma) = \widehat{\Phi}(\sigma) + \Phi''(\sigma) \quad (23.146a)$$

$$= \frac{\overline{h \Phi}^{(\sigma)}}{\overline{h}^{(\sigma)}} + \Phi''. \quad (23.146b)$$

It follows by definition that

$$\overline{h \Phi''}^{(\sigma)} = 0. \quad (23.147)$$

23.6.2 Isopycnal mean thickness weighted tracer equation

When attaching a tracer to fluid parcels, each member of the ensemble satisfies the isopycnal tracer equation

$$\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla_{\sigma} C = 0. \quad (23.148)$$

Combining the tracer and thickness equations leads to the thickness weighted tracer equation

$$\frac{\partial(h C)}{\partial t} + \nabla_{\sigma} \cdot (h \mathbf{u} C) = 0. \quad (23.149)$$

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

$$\partial_t [h(\widehat{C} + C'')] + \nabla_\sigma \cdot [h(\widehat{\mathbf{u}} + \mathbf{u}'')](\widehat{C} + C'') = 0. \quad (23.150)$$

Taking an ensemble average over fluid elements with the same potential density, and using equation (23.147), yield the mean thickness weighted tracer equation

$$\partial_t (\bar{h}^{(\sigma)} \widehat{C}) + \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \widehat{C} \widehat{\mathbf{u}}) = -\nabla_\sigma \cdot (\bar{h} C'' \mathbf{u}''). \quad (23.151)$$

Now introduce the correlation

$$\bar{h} C'' \mathbf{u}''^{(\sigma)} = \bar{h}^{(\sigma)} \widehat{C'' \mathbf{u}''} \quad (23.152)$$

(see equation (23.146b)), and recall that the mean thickness $\bar{h}^{(\sigma)}$ satisfies the mean thickness equation (23.114). These two points lead to the evolution equation for the mean thickness weighted tracer concentration

$$(\partial_t + \widehat{\mathbf{u}} \cdot \nabla_\sigma) \widehat{C} = -\frac{1}{\bar{h}^{(\sigma)}} \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \widehat{C'' \mathbf{u}''}). \quad (23.153)$$

23.6.3 Subgrid scale tracer transport tensor

The correlation between tracer and velocity found on the right-hand side of the mean thickness weighted tracer equation (23.153) is typically written in terms of a subgrid scale tracer transport tensor

$$\widehat{C'' \mathbf{u}''} = -\mathbb{J} \cdot \nabla_\sigma \widehat{C}. \quad (23.154)$$

This definition leads to the evolution equation

$$(\partial_t + \widehat{\mathbf{u}} \cdot \nabla_\sigma) \widehat{C} = \frac{1}{\bar{h}^{(\sigma)}} \nabla_\sigma \cdot (\bar{h}^{(\sigma)} \mathbb{J} \cdot \nabla_\sigma \widehat{C}). \quad (23.155)$$

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 8.15. However, in addition to symmetric diffusion processes, this operator includes skewed fluxes that lead to skew diffusion as discussed in Section 23.3.2. Whereas the diffusive aspect is commonly parameterized as dia-neutral diffusion and neutral diffusion (Section 24.1), there is no parameterization for the skewed correlations for use in ocean models. We comment further on this situation in Section 24.3.8.

23.6.4 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 $\sigma = \tilde{\sigma}$,

$$\overline{\mathbf{C}}^\#(\bar{z}^{(\sigma)}) = \int_{-H}^{\bar{z}^{(\sigma)} + \xi} C \mathbf{u} dz. \quad (23.156)$$

Setting tracer concentration to unity recovers the expression (23.97) for the TRM transport. Changing coordinates and making use of the tracer correlation tensor renders

$$\overline{C}^\#(\bar{z}^{(\sigma)}) = \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \overline{C} \mathbf{u} \bar{h}^{(\sigma)} d\sigma \quad (23.157a)$$

$$= \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \bar{h}^{(\sigma)} d\sigma (\hat{C} \hat{\mathbf{u}} + \widehat{C'' \mathbf{u}''}) \quad (23.157b)$$

$$= \int_{\sigma(-H)}^{\tilde{\sigma}(\bar{z}^{(\sigma)})} \bar{h}^{(\sigma)} d\sigma (\hat{C} \hat{\mathbf{u}} - \mathbb{J} \cdot \nabla_\sigma \hat{C}) \quad (23.157c)$$

$$= \int_{-H}^{\bar{z}^{(\sigma)}} dz (\hat{C} \hat{\mathbf{u}} - \mathbb{J} \cdot \nabla_\sigma \hat{C}). \quad (23.157d)$$

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.

23.6.5 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

$$\hat{\mathbf{u}}((\sigma)) = \frac{\bar{h} \mathbf{u}^{(\sigma)}}{\bar{h}^{(\sigma)}} = \bar{\mathbf{u}}^{(\sigma)} + \frac{\bar{h}' \mathbf{u}'^{(\sigma)}}{\bar{h}^{(\sigma)}} = \bar{\mathbf{u}}^{(\sigma)} + \mathbf{u}^{\text{bolus}}. \quad (23.158)$$

In addition to the bolus velocity, it is necessary to parameterize the subgrid scale tracer transport tensor

$$\widehat{C'' \mathbf{u}''} = -\mathbb{J} \cdot \nabla_\sigma \hat{C}, \quad (23.159)$$

which generally has symmetric (diffusive) and antisymmetric (stirring) components (Section 23.3).

For a geopotential coordinate description, equation (23.93) 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

$$\hat{C}(x, y, \tilde{\sigma}, t) = \overline{C}^\#(x, y, \bar{z}^{(\sigma)}, t). \quad (23.160)$$

Equation (23.160), and the developed formalism, leads to the mean field tracer equation in geopotential coordinates

$$\partial_t \overline{C}^\# = \nabla \cdot (\nabla \overline{C}^\# \wedge \overline{\Psi}^\#) + R(\overline{C}^\#), \quad (23.161)$$

where $R(\overline{C}^\#)$ is the geopotential coordinate form of the mixing/stirring operator on the right-hand side of equation (23.155). Details for the transformation of the mixing/stirring operator from isopycnal to geopotential coordinates are provided in Section 8.15.

23.6.6 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.

24

Ocean tracer subgrid scale 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 *ocean subgrid scale parameterization problem* is broader and deeper than available from a single chapter. Even so, we synthesize a range of physical and mathematical topics associated with ocean tracers, aiming to develop a platform to further penetrate the vast and growing literature.

READER'S GUIDE FOR THIS CHAPTER

Tracers evolve according to the advection-diffusion equation discussed in Chapter 22 and further unpacked in the tracer kinematics Chapter 23.

- add section on anisotropic neutral diffusion
- add section on anisotropic GM

24.1	Parameterizing eddy induced tracer transport	358
24.1.1	Framework based on tracer variance cascade	358
24.1.2	Density changes from molecular diffusion	359
24.1.3	Mixing from small (or fine) scale processes	359
24.1.4	Mesoscale eddy-induced stirring and mixing	359
24.1.5	Synthesis	362
24.2	Quasi-Stokes induced tracer stirring	363
24.2.1	Gent and McWilliams skewson	363
24.2.2	Local adiabatic dissipation of APE	364
24.2.3	Meridional volume transport	365
24.2.4	Isopycnal thickness diffusion and GM	365
24.3	Neutral diffusion	366
24.3.1	Redi neutral diffusion	367
24.3.2	Small slope neutral diffusion	367
24.3.3	Neutral tangent plane neutral diffusion	368
24.3.4	Neutrality condition	369
24.3.5	Symmetry condition	369
24.3.6	GM skewson plus small slope neutral diffusion	370
24.3.7	Small slope neutral diffusion in generalized vertical coordinates	370
24.3.8	Comments	371
24.4	Cabbeling and thermobaricity	371
24.4.1	Basic manipulations	372
24.4.2	A tidy form	373
24.4.3	Cabbeling	374
24.4.4	Thermobaricity	374
24.4.5	Comments	375
24.5	Neutral helicity	375
24.5.1	Mathematical preliminaries	375
24.5.2	Helical nature of neutral displacements	376
24.5.3	Comments and further reading	377

24.1 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 21.2 to further detail how *in situ* density evolves in the presence of eddy parameterizations.

24.1.1 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.

24.1.2 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

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot [\rho \kappa_\Theta \nabla \Theta] \quad (24.1a)$$

$$\rho \frac{DS}{Dt} = \nabla \cdot [\rho \kappa_S \nabla S], \quad (24.1b)$$

where $\kappa_\Theta > 0$ and $\kappa_S > 0$ are the molecular kinematic diffusivities for Θ and S , respectively. Following equation (21.28), we see that these molecular tracer fluxes lead to the material evolution of *in situ* density

$$\frac{D\rho}{Dt} - \frac{\omega}{c_s^2} = -\alpha \nabla \cdot (\rho \kappa_\Theta \nabla \Theta) + \beta \nabla \cdot (\rho \kappa_S \nabla S) \quad (24.2a)$$

$$= -\nabla \cdot [\rho (\kappa_\Theta \alpha \nabla \Theta - \kappa_S \beta \nabla S)] + \rho (\kappa_\Theta \nabla \Theta \cdot \nabla \alpha - \kappa_S \nabla S \cdot \nabla \beta). \quad (24.2b)$$

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 .

24.1.3 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 (24.1b), 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

$$\frac{D\rho}{Dt} - \frac{\omega}{c_s^2} = -\alpha \nabla \cdot (\rho \kappa \nabla \Theta) + \beta \nabla \cdot (\rho \kappa \nabla S) \quad (24.3a)$$

$$= -\nabla \cdot [\rho \kappa (\alpha \nabla \Theta - \beta \nabla S)] + \rho \kappa (\nabla \Theta \cdot \nabla \alpha - \nabla S \cdot \nabla \beta). \quad (24.3b)$$

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).

24.1.4 Mesoscale eddy-induced stirring and mixing

Stirring from turbulent scales smaller than the grid scale is commonly parameterized by an eddy-induced stirring velocity, 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 24.1.3. 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

$$\rho \frac{DS}{Dt} = \nabla \cdot (\rho \mathbb{M} \cdot \nabla S) \quad (24.4a)$$

$$\rho \frac{D\Theta}{Dt} = \nabla \cdot (\rho \mathbb{M} \cdot \nabla \Theta). \quad (24.4b)$$

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 22, it is useful to decompose the second order transport tensor into the sum of its symmetric and anti-symmetric components

$$\mathbb{M} = \mathbb{K} + \mathbb{A}. \quad (24.5)$$

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

$$\nabla \cdot (\rho \mathbb{A} \cdot \nabla S) = \partial_m (\rho A^{mn} \partial_n S) \quad (24.6a)$$

$$= \partial_m (\rho A^{mn}) \partial_n S + \rho A^{mn} \partial_m \partial_n S \quad (24.6b)$$

$$= -\rho v^{*n} \partial_n S, \quad (24.6c)$$

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

$$\rho A^{mn} \partial_m \partial_n S = 0 \quad (24.7)$$

since A^{mn} is anti-symmetric whereas $\partial_m \partial_n S$ is symmetric. Finally, we introduced a density-weighted eddy-induced velocity

$$\rho v^{*n} = -\partial_m (\rho A^{mn}). \quad (24.8)$$

Importantly, $\rho \mathbf{v}^*$ has a zero divergence, again due to anti-symmetry of A^{mn}

$$\nabla \cdot (\rho \mathbf{v}^*) = \partial_n (\rho v^{*n}) = -\partial_n \partial_m (\rho A^{mn}) = 0. \quad (24.9)$$

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 22.5). 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

$$\rho \frac{D^\dagger S}{Dt} = \nabla \cdot (\rho \mathbb{K} \cdot \nabla S) \quad (24.10a)$$

$$\rho \frac{D^\dagger \Theta}{Dt} = \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta), \quad (24.10b)$$

where the residual mean material time derivative is given by

$$\frac{D^\dagger}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}^\dagger \cdot \nabla \quad (24.11)$$

and the residual mean velocity is

$$\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*. \quad (24.12)$$

Making use of the residual mean velocity then leads to the material evolution of the *in situ* density

$$\frac{D\rho}{Dt} - \frac{1}{c_s^2} \frac{Dp}{Dt} = -\rho \mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S) - \alpha \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbb{K} \cdot \nabla S), \quad (24.13)$$

which can be written in terms of the residual mean material time derivative

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = -\alpha \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbb{K} \cdot \nabla S). \quad (24.14)$$

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 (24.14) can be written

$$-\alpha \nabla \cdot (\rho \mathbb{K} \cdot \nabla \Theta) + \beta \nabla \cdot (\rho \mathbb{K} \cdot \nabla S) = \nabla \cdot [\rho \mathbb{K} \cdot (-\alpha \nabla \Theta + \beta \nabla S)] + \rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S, \quad (24.15)$$

so that the *in situ* density evolves according to

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = -\underbrace{\nabla \cdot [\rho \mathbb{K} \cdot (\alpha \nabla \Theta - \beta \nabla S)]}_{\text{conservative processes}} + \underbrace{\rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S}_{\text{sources from nonlinear EOS processes}}. \quad (24.16)$$

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 24.4.
- CONSERVATIVE PROCESSES AND NEUTRAL DIFFUSION: A linear equation of state is independent of pressure, so that the evolution equation (24.16) takes the form

$$\frac{D^\dagger \rho}{Dt} = -\nabla \cdot [\rho \mathbb{K} \cdot (\alpha \nabla \Theta - \beta \nabla S)]. \quad (24.17)$$

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

$$\mathbb{K} \cdot (\alpha \nabla \Theta - \beta \nabla S) = 0. \quad (24.18)$$

As detailed in Section 24.3, 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

$$\frac{D^\dagger \rho}{Dt} - \frac{1}{c_s^2} \frac{D^\dagger p}{Dt} = \underbrace{\rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S}_{\text{sources from nonlinear EOS processes}}. \quad (24.19)$$

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 (24.18). 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 (24.18). In Section 24.3 we dive into the details of neutral diffusion.

24.1.5 Synthesis

In summary, the proposed evolution equation for *in situ* density in the presence of subgrid scale processes takes the form

$$\frac{D\rho}{Dt} - \underbrace{\frac{1}{c_s^2} \frac{Dp}{Dt}}_{\text{compressibility}} = -\underbrace{\mathbf{v}^* \cdot (-\alpha \nabla \Theta + \beta \nabla S)}_{\text{eddy-induced stirring}} - \underbrace{\alpha \nabla \cdot (\rho \kappa \nabla \Theta) + \beta \nabla \cdot (\rho \kappa \nabla S)}_{\text{small scale mixing and nonlinear EOS}} + \underbrace{\rho \nabla \alpha \cdot \mathbb{K} \cdot \nabla \Theta - \rho \nabla \beta \cdot \mathbb{K} \cdot \nabla S}_{\text{nonlinear EOS processes from eddy mixing}}. \quad (24.20)$$

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 (24.3b) 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 (24.18).

- 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 24.4). 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

$$-\alpha \nabla \cdot (\rho \kappa \nabla \Theta) + \beta \nabla \cdot (\rho \kappa \nabla S) = -\nabla \cdot [\rho \kappa (\alpha \nabla \Theta - \beta \nabla S)] + \rho \kappa (\nabla \alpha \cdot \nabla \Theta - \nabla \beta \cdot \nabla S). \quad (24.21)$$

24.2 Quasi-Stokes induced tracer stirring

As mentioned in Section 23.6.3, 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 23.3, 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 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 23.4, 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.

24.2.1 Gent and McWilliams skewson

[Gent et al. \(1995\)](#) parameterize the divergent part of the quasi-Stokes transport by setting

$$\mathbf{U}^{\text{qs}} = -\kappa \mathbf{S}. \quad (24.22)$$

In this expression, \mathbf{S} is the slope of the potential density surfaces (equation (23.119)), 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

$$\mathbf{v}^* = -\partial_z (\kappa \mathbf{S}) + \hat{z} \nabla_z \cdot (\kappa \mathbf{S}), \quad (24.23)$$

and the antisymmetric stirring tensor is

$$\mathbb{A}_{mn} = \begin{bmatrix} 0 & 0 & -\kappa S_x \\ 0 & 0 & -\kappa S_y \\ \kappa S_x & \kappa S_y & 0 \end{bmatrix}. \quad (24.24)$$

The parameterized skew flux of potential density due to the quasi-Stokes transport is given by

$$\mathbf{F}^{\text{skew}} = -\mathbf{U}^{\text{qs}} \partial_z \sigma + \hat{z} (\mathbf{U}^{\text{qs}} \cdot \nabla_z) \sigma = -\kappa \nabla_z \sigma + \hat{z} S^2 \kappa \partial_z \sigma. \quad (24.25)$$

This parameterization yields horizontal downgradient diffusion of potential density, combined with a vertical upgradient diffusion. 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 (23.100).

24.2.2 Local adiabatic dissipation of APE

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 isopycnal layers, (II) the subgrid operator locally dissipates available potential energy (APE). 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

$$\mathbf{F} = -\nabla\sigma \wedge \Psi, \quad (24.26)$$

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

$$P = g \int \sigma z \, dV, \quad (24.27)$$

where we assume the *in situ* density equals to the potential density as per a linear equation of state (Section 20.2.6). Assuming all boundaries are material and static allows us to focus on the time tendency of potential energy associated with the unknown flux

$$\frac{dP}{dt} = g \int z \frac{\partial\sigma}{\partial t} \, dV \quad (24.28a)$$

$$= -g \int dV (z \nabla \cdot \mathbf{F}) \quad (24.28b)$$

$$= -g \int dV (z \partial_z F^{(z)}) \quad (24.28c)$$

$$= g \int dV F^{(z)}, \quad (24.28d)$$

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

$$F^{(z)} \leq 0, \quad (24.29)$$

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 $\partial_z\sigma < 0$, the following form provides a local APE sink

$$F^{(z)} = \kappa S^2 \frac{\partial\sigma}{\partial z} \quad (24.30)$$

where $\kappa > 0$ is a 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

$$\mathbf{F}^{(h)} = -\kappa \nabla_z \sigma. \quad (24.31)$$

We have thus recovered the skew flux (24.25) 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.

24.2.3 Meridional volume transport

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 across a basin at a particular depth in the form

$$\mathcal{T}(y, z, t) = \int dx \bar{V}^\#(y, z, t) = \int dx (\bar{V} - \kappa S_y). \quad (24.32)$$

We thus see that the parameterized quasi-Stokes transport adds a contribution that scales linearly with basin size, isopycnal slope, and diffusivity,

$$\mathcal{T}^{\text{qs}} \sim L S \kappa. \quad (24.33)$$

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.

24.2.4 Isopycnal thickness diffusion and GM

Recall the ensemble mean thickness equation (23.114) derived in Section 23.4.5

$$\partial_t h + \nabla_\sigma \cdot (h \hat{\mathbf{u}}) = 0, \quad (24.34)$$

where

$$\hat{\mathbf{u}} = \mathbf{u} + \mathbf{u}^{\text{bolus}} \quad (24.35)$$

is the thickness weighted transport velocity affecting evolution of the ensemble mean thickness h . Note that for brevity we here drop the nomenclature $(\)^{(\sigma)}$ used in Section 23.4.5.

Isopycnal correlations of horizontal velocity and layer thickness define the bolus velocity via

$$h \mathbf{u}^{\text{bolus}} = \overline{h' \mathbf{u}'} \quad (24.36)$$

Now consider a downgradient diffusive closure for this correlation

$$h \mathbf{u}^{\text{bolus}} = \overline{h' \mathbf{u}'}^{(\sigma)} \quad (24.37a)$$

$$= -\mathbf{K} \cdot \nabla_\sigma h \quad (24.37b)$$

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

$$\partial_t h + \nabla_\sigma \cdot (h \mathbf{u}) = \nabla_\sigma \cdot (\mathbf{K} \cdot \nabla_\sigma h). \quad (24.38)$$

To make a connection between the thickness diffusion closure (24.37b) and the [Gent et al. \(1995\)](#) closure discussed in Section 24.2.1, note that the specific thickness is the inverse of the vertical derivative of the potential density

$$h = (\partial_z \sigma)^{-1}. \quad (24.39)$$

Correspondingly, using the relation between derivative operators, $\nabla_\sigma = \nabla_z + \mathbf{S} \partial_z$, gives

$$h^{-1} \nabla_\sigma h = -h \nabla_\sigma (1/h) \quad (24.40a)$$

$$= -(\partial_z \sigma)^{-1} (\nabla_z + \mathbf{S} \partial_z) \partial_z \sigma \quad (24.40b)$$

$$= -\frac{\partial_z (\nabla_z \sigma)}{\partial_z \sigma} + \frac{\partial_{zz} \sigma \nabla_z \sigma}{(\partial_z \sigma)^2} \quad (24.40c)$$

$$= \partial_z \mathbf{S}. \quad (24.40d)$$

Consequently, the bolus velocity takes the form

$$\mathbf{u}^{\text{bolus}} = -h^{-1} \mathbf{K} \cdot \nabla_\sigma h = -\mathbf{K} \cdot \partial_z \mathbf{S}. \quad (24.41)$$

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

$$\mathbf{u}^{\text{bolus}} = -\partial_z (\kappa \mathbf{S}) = \mathbf{u}^*, \quad (24.42)$$

where the horizontal component of the [Gent et al. \(1995\)](#) velocity \mathbf{u}^* was identified from equation (24.23). 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.

24.3 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 20.3, 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

$$\mathbf{J} \cdot \hat{\boldsymbol{\gamma}} = 0 \quad (24.43)$$

where (equation (20.30))

$$\hat{\boldsymbol{\gamma}} = \frac{\rho_\theta \nabla \theta + \rho_S \nabla S}{|\rho_\theta \nabla \theta + \rho_S \nabla S|} = \frac{-\alpha \nabla \Theta + \beta \nabla S}{|-\alpha \nabla \Theta + \beta \nabla S|}. \quad (24.44)$$

24.3.1 Redi neutral diffusion

One diffusive flux satisfying the property (24.43) is given by

$$\mathbf{J}^{\text{redi}} = -\rho A [\nabla C - \hat{\gamma} (\hat{\gamma} \cdot \nabla C)], \quad (24.45)$$

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

$$\mathbf{J}^{\text{redi}} \cdot \nabla C = -\rho A [|\nabla C|^2 - (\hat{\gamma} \cdot \nabla C)^2] \leq 0. \quad (24.46)$$

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

$$J^m = -\rho K^{mn} \partial_n C, \quad (24.47)$$

where the Redi diffusion tensor is given by

$$K^{mn} = \frac{A}{1 + S_x^2 + S_y^2} \begin{bmatrix} 1 + S_y^2 & -S_x S_y & S_x \\ -S_x S_y & 1 + S_x^2 & S_y \\ S_x & S_y & S_x^2 + S_y^2 \end{bmatrix} \quad (24.48)$$

with the corresponding neutral diffusion operator given by the three-dimensional flux convergence

$$\mathcal{R}^{\text{redi}} = -\nabla \cdot \mathbf{J}^{\text{redi}} = \partial_m (\rho K^{mn} \partial_n C). \quad (24.49)$$

In the Redi tensor (24.48) we introduced the components of the horizontal vector, $\mathbf{S} = (S_x, S_y, 0)$, with

$$\mathbf{S} = - \begin{bmatrix} -\alpha \nabla_z \Theta + \beta \nabla_z S \\ -\alpha \partial_z \Theta + \beta \partial_z S \end{bmatrix} = \frac{g (-\alpha \nabla_z \Theta + \beta \nabla_z S)}{N^2} \quad (24.50)$$

the slope of the neutral tangent plane relative to the (x, y) horizontal plane, with

$$N^2 = -\frac{g}{\rho_0} \frac{\partial \gamma}{\partial z} = -g [-\alpha \partial_z \Theta + \beta \partial_z S] \quad (24.51)$$

the squared buoyancy frequency (Section 20.3). 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.

24.3.2 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

$$(K^{\text{small}})^{mn} = A \begin{bmatrix} 1 & 0 & S_x \\ 0 & 1 & S_y \\ S_x & S_y & S_x^2 + S_y^2 \end{bmatrix}. \quad (24.52)$$

The corresponding small slope neutral diffusive flux is

$$\mathbf{J}^{\text{small}} = -\rho A [\nabla_{\gamma} + \hat{\mathbf{z}} (\mathbf{S} \cdot \nabla_{\gamma})] C \quad (24.53)$$

where

$$\nabla_{\gamma} = \nabla_z + \mathbf{S} \partial_z \quad (24.54)$$

is the horizontal derivative operator computed on the neutral tangent plane (see equation (8.64)). To show that $\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = 0$, we make use of the identity

$$\hat{\gamma} = \frac{\mathbf{S} - \hat{\mathbf{z}}}{(1 + \mathbf{S} \cdot \mathbf{S})^{1/2}}, \quad (24.55)$$

so that

$$\mathbf{J}^{\text{small}} \cdot \hat{\gamma} = \frac{\mathbf{J}^{\text{small}} \cdot \mathbf{S} - \mathbf{J}^{\text{small}} \cdot \mathbf{S}}{(1 + \mathbf{S} \cdot \mathbf{S})^{1/2}} = 0. \quad (24.56)$$

Furthermore, we confirm that $\mathbf{J}^{\text{small}}$ is oriented down the tracer gradient by noting that

$$\mathbf{J}^{\text{small}} \cdot \nabla C = -\rho A [\nabla_\gamma C \cdot \nabla_z C + (\mathbf{S} \cdot \nabla_\gamma C) \partial_z C] \quad (24.57a)$$

$$= -\rho A [|\nabla_z C|^2 + 2(\mathbf{S} \cdot \nabla_z C) \partial_z C + |\mathbf{S} \partial_z C|^2] \quad (24.57b)$$

$$= -\rho A |\nabla_z C + \mathbf{S} \partial_z C|^2 \quad (24.57c)$$

$$= -\rho A |\nabla_\gamma C|^2 \quad (24.57d)$$

$$\leq 0. \quad (24.57e)$$

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 (24.53) 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

$$\mathcal{R}^{\text{small}} = -\nabla \cdot \mathbf{J}^{\text{small}} = \nabla_z \cdot (\rho A \nabla_\gamma C) + \partial_z (\rho A \mathbf{S} \cdot \nabla_\gamma C). \quad (24.58)$$

24.3.3 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 8.

Following the derivations given in Section 8.15, the neutral diffusive flux in the neutral tangent frame is given by the horizontal flux

$$\mathbf{J}^{\text{ntp}} = -\rho A \nabla_\gamma C. \quad (24.59)$$

This flux is oriented down the tracer gradient as oriented along neutral directions

$$\mathbf{J}^{\text{ntp}} \cdot \nabla_\gamma C = -\rho A |\nabla_\gamma C|^2, \quad (24.60)$$

which is the same as equation (24.57d) for the small slope fluxes. However, as a purely horizontal flux, \mathbf{J}^{ntp} is not oriented along neutral directions

$$\mathbf{J}^{\text{ntp}} \cdot \hat{\gamma} \neq 0. \quad (24.61)$$

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 (8.84)

$$\mathcal{R}^{\text{ntp}} = -\frac{1}{h^\gamma} [\nabla_\gamma \cdot (h^\gamma \mathbf{J}^{\text{ntp}})] = \frac{1}{h^\gamma} [\nabla_\gamma \cdot (h^\gamma \rho A \nabla_\gamma C)], \quad (24.62)$$

where

$$h^\gamma = \frac{\partial z}{\partial \gamma} d\gamma = - \left[\frac{g}{\rho_0 N^2} \right] d\gamma \quad (24.63)$$

measures the thickness of a layer defined by two neutral tangent planes (see equation (8.81)).

As detailed in Section 8.15, \mathcal{R}^{ntp} is identical to the small slope neutral diffusion operator (24.58)

$$\mathcal{R}^{\text{ntp}} = \mathcal{R}^{\text{small}}. \quad (24.64)$$

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 22.3). 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.

24.3.4 Neutrality condition

Given the expression (24.44) for the dianeutral unit vector, $\hat{\gamma}$, it is straightforward to show that the neutral diffusive flux for Conservative Temperature balances that for salinity

$$\alpha \mathbf{J}(\Theta) = \beta \mathbf{J}(S). \quad (24.65)$$

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 (24.45) of [Redi \(1982\)](#), the small slope flux (24.53) of [Gent and McWilliams \(1990\)](#), and the neutral tangent frame neutral diffusive flux (24.60). 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 \(1987\)](#).

24.3.5 Symmetry condition

Since the neutral diffusion tensor is symmetric (as are all diffusion tensors; see Section 22.3), we have

$$\mathbf{J}(\Theta) \cdot \nabla S = -A \rho K^{mn} \partial_n \Theta \partial_m S \quad (24.66a)$$

$$= -A \rho K^{nm} \partial_n S \partial_m \Theta \quad (24.66b)$$

$$= -A \rho K^{mn} \partial_n S \partial_m \Theta \quad (24.66c)$$

$$= \mathbf{J}(S) \cdot \nabla \Theta. \quad (24.66d)$$

This symmetry condition will be useful in our discussion of cabbeling and thermobaricity in Section 24.4.

24.3.6 GM skewson 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 skewson (Section 24.2.1) and small slope neutral diffusion (Section 24.3.2). The combined tracer flux takes the form

$$\mathbf{F} = -A \nabla_z C + (\kappa - A) \mathbf{S} \partial_z C - \hat{\mathbf{z}} [(A + \kappa) \mathbf{S} \cdot \nabla_z C + A S^2 \partial_z C], \quad (24.67)$$

which can be written in terms of a subgrid scale transport tensor ([Griffies, 1998](#))

$$\begin{bmatrix} F^{(x)} \\ F^{(y)} \\ F^{(z)} \end{bmatrix} = \begin{bmatrix} A & 0 & (A - \kappa) S_x \\ 0 & A & (A - \kappa) S_y \\ (A + \kappa) S_x & (A + \kappa) S_y & A S^2 \end{bmatrix} \begin{bmatrix} \partial_x C \\ \partial_y C \\ \partial_z C \end{bmatrix}. \quad (24.68)$$

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

$$\mathbf{F} = -\kappa \nabla_z C - \hat{\mathbf{z}} \kappa (2 \mathbf{S} \cdot \nabla_z C + S^2 C_{,z}). \quad (24.69)$$

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 skewson 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 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.

24.3.7 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 8 and 18. This formulation is relevant for the now common use of generalized vertical coordinates for ocean modeling, with an example algorithm discussed in Section 31.4.

We start by recalling the expression (8.83) for a general diffusion operator written in terms of the generalized vertical coordinate, $\sigma = \sigma(x, y, z, t)$

$$\mathcal{R} = -\frac{1}{h^\sigma} \left[\nabla_\sigma \cdot (h^\sigma \mathbf{J}^h) + \delta_\sigma(z_\sigma \nabla_\sigma \cdot \mathbf{J}) \right], \quad (24.70)$$

where $\delta_\sigma \equiv d\sigma / \partial_\sigma$ is the dimensionless derivative operator, and the thickness of a σ -layer is

$$h^\sigma = dz = z_\sigma d\sigma = \frac{\partial z}{\partial \sigma} d\sigma. \quad (24.71)$$

Now assume the flux, \mathbf{J} , is given by equation (24.53) for small slope neutral diffusion. Transforming to GVCs leads to the horizontal flux component

$$\mathbf{J}^{\text{small, h}} = -\rho A \nabla_\gamma C \quad (24.72\text{a})$$

$$= -\rho A [\nabla_z + (\nabla_\gamma z) \partial_z] C \quad (24.72\text{b})$$

$$= -\rho A [\nabla_\sigma + (-\nabla_\sigma z + \nabla_\gamma z) \partial_z] C \quad (24.72\text{c})$$

$$= -\rho A [\nabla_\sigma + (-\mathbf{S}^{(\sigma/z)} + \mathbf{S}^{(\gamma/z)}) \partial_z] C \quad (24.72\text{d})$$

$$= -\rho A (\nabla_\sigma + \mathbf{S}^{(\gamma/\sigma)} \partial_z) C, \quad (24.72\text{e})$$

where the neutral slopes as shown in Figure 24.1 satisfy the identity

$$\mathbf{S}^{(\sigma/z)} = \mathbf{S}^{(\gamma/z)} - \mathbf{S}^{(\gamma/\sigma)}. \quad (24.73)$$

Furthermore, we made use of the identity (8.64) relating the partial derivative operators

$$\nabla_\gamma = \nabla_z + (\nabla_\gamma z) \partial_z \quad \nabla_z = \nabla_\sigma - (\nabla_\sigma z) \partial_z. \quad (24.74)$$

The horizontal flux (24.72e) 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

$$J^{\text{small, z}} = \mathbf{J}^{\text{small, h}} \cdot \mathbf{S}^{(\gamma/z)} \quad (24.75)$$

takes the form

$$z_\sigma \nabla_\sigma \cdot \mathbf{J}^{\text{small}} = -\mathbf{S}^{(\sigma/z)} \cdot \mathbf{J}^{\text{small, h}} + J^{\text{small, z}} = \mathbf{J}^{\text{small, h}} \cdot \mathbf{S}^{(\gamma/\sigma)}, \quad (24.76)$$

which in turn yields the diffusion operator (24.70)

$$\mathcal{R} = -\frac{1}{h^\sigma} [\nabla_\sigma \cdot (h^\sigma \mathbf{J}^{\text{small, h}}) + \delta_\sigma (\mathbf{J}^{\text{small, h}} \cdot \mathbf{S}^{(\gamma/\sigma)})]. \quad (24.77)$$

In the special case when σ is parallel to the neutral direction so that $\mathbf{S}^{(\gamma/\sigma)} = 0$, the diffusion operator (24.77) reduces to the neutral tangent plane version given by equation (24.62).

24.3.8 Comments

As noted in Section 23.6.3, 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 23.3.2 and [Middleton and Loder \(1989\)](#).

24.4 Cabbeling and thermobaricity

We now return to the density equation (21.28)

$$\frac{D \ln \rho}{Dt} = \nabla \cdot [\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S)] - (\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S) + \frac{\omega}{\rho c^2}. \quad (24.78)$$

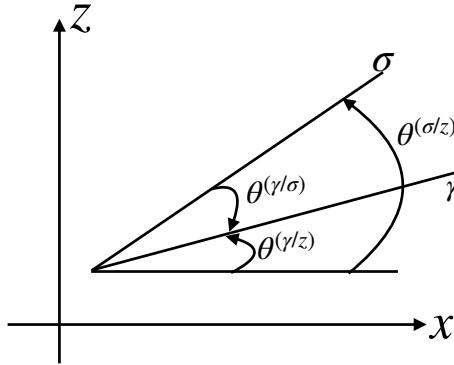


Figure 24.1: Slopes of neutral directions (denoted by γ isolines) relative to both the horizontal plane, $\tan \theta^{(\gamma/z)}$, and relative to a σ -isoline, $\tan \theta^{(\gamma/\sigma)}$. We assume positive angles as measure counter-clockwise relative to the horizontal and relative to the σ -isoline, 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.

We focus here on temperature and salinity fluxes just from neutral diffusion. The neutrality condition (24.65) takes the following form in terms of specific volume

$$\nu_\Theta \mathbf{J}(\Theta) + \nu_S \mathbf{J}(S) = 0. \quad (24.79)$$

Consequently, neutral diffusion affects density evolution only through the source term

$$\left[\frac{D \ln \rho}{Dt} \right]_{\text{ntral diff}} = -\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta - \mathbf{J}(S) \cdot \nabla \nu_S. \quad (24.80)$$

24.4.1 Basic manipulations

As a first step, eliminate the salt flux by using the neutrality condition (24.79) so that

$$\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S = \mathbf{J}(\Theta) \cdot [\nu_S \nabla \nu_\Theta - \nu_\Theta \nabla \nu_S] / \nu_S. \quad (24.81)$$

Next, expand the gradients of the specific volume to write

$$\nabla \nu_\Theta = \nu_{\Theta\Theta} \nabla \Theta + \nu_{\Theta S} \nabla S + \nu_{\Theta p} \nabla p \quad (24.82a)$$

$$\nabla \nu_S = \nu_{SS} \nabla S + \nu_{\Theta S} \nabla \Theta + \nu_{Sp} \nabla p, \quad (24.82b)$$

so that

$$\begin{aligned} \nu_S \nabla \nu_\Theta - \nu_\Theta \nabla \nu_S &= \nabla \Theta (\nu_S \nu_{\Theta\Theta} - \nu_\Theta \nu_{\Theta S}) \\ &\quad + \nabla S (\nu_S \nu_{\Theta S} - \nu_\Theta \nu_{SS}) + \nabla p (\nu_S \nu_{\Theta p} - \nu_\Theta \nu_{Sp}). \end{aligned} \quad (24.83)$$

We again make use of the neutrality condition (24.79), as well as the symmetry condition (24.66d) to write

$$\mathbf{J}(\Theta) \cdot \nabla S (\nu_S \nu_{\Theta S} - \nu_\Theta \nu_{SS}) = -\mathbf{J}(\Theta) \cdot \nabla \Theta \left(\nu_\Theta \nu_{\Theta S} - \nu_{SS} \frac{(\nu_\Theta)^2}{\nu_S} \right). \quad (24.84)$$

Bringing these results together leads to

$$\begin{aligned}\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S &= \mathbf{J}(\Theta) \cdot \nabla p \left[\nu_{\Theta p} - \nu_{pS} \frac{\nu_\Theta}{\nu_S} \right] \\ &\quad + \mathbf{J}(\Theta) \cdot \nabla \Theta \left[\nu_{\Theta \Theta} - 2 \nu_{\Theta S} \frac{\nu_\Theta}{\nu_S} + \nu_{SS} \left(\frac{\nu_\Theta}{\nu_S} \right)^2 \right],\end{aligned}\quad (24.85)$$

which can be written in terms of density partial derivatives as

$$\begin{aligned}\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S &= -\rho^{-2} \mathbf{J}(\Theta) \cdot \nabla p \left[\rho_{\Theta p} - \rho_{pS} \frac{\rho_\Theta}{\rho_S} \right] \\ &\quad - \rho^{-2} \mathbf{J}(\Theta) \cdot \nabla \Theta \left[\rho_{\Theta \Theta} - 2 \rho_{\Theta S} \frac{\rho_\Theta}{\rho_S} + \rho_{SS} \left(\frac{\rho_\Theta}{\rho_S} \right)^2 \right].\end{aligned}\quad (24.86)$$

24.4.2 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

$$\mathcal{T} = \beta \partial_p \left[\frac{\alpha}{\beta} \right] \quad (24.87a)$$

$$= \frac{\partial \alpha}{\partial p} - \frac{\alpha}{\beta} \frac{\partial \beta}{\partial p} \quad (24.87b)$$

$$= \rho \nu_S \partial_p \left[\frac{\nu_\Theta}{\nu_S} \right] \quad (24.87c)$$

$$= -\rho^{-1} \rho_S \partial_p \left[\frac{\rho_\Theta}{\rho_S} \right] \quad (24.87d)$$

$$= -\rho^{-1} \left[\rho_{\Theta p} - \rho_{pS} \left[\frac{\rho_\Theta}{\rho_S} \right] \right], \quad (24.87e)$$

and the *cabbeling* parameter (dimensions of squared inverse temperature)

$$\mathcal{C} = \frac{\partial \alpha}{\partial \Theta} + 2 \frac{\alpha}{\beta} \frac{\partial \alpha}{\partial S} - \left(\frac{\alpha}{\beta} \right)^2 \frac{\partial \beta}{\partial S} \quad (24.88a)$$

$$= -\rho^{-1} \left[\rho_{\Theta \Theta} - 2 \rho_{\Theta S} \left[\frac{\rho_\Theta}{\rho_S} \right] + \rho_{SS} \left[\frac{\rho_\Theta}{\rho_S} \right]^2 \right] \quad (24.88b)$$

$$= \rho \left[\nu_{\Theta \Theta} - 2 \nu_{\Theta S} \left[\frac{\nu_\Theta}{\nu_S} \right] + \nu_{SS} \left[\frac{\nu_\Theta}{\nu_S} \right]^2 \right] \quad (24.88c)$$

to render the very compact result

$$\mathbf{J}(\Theta) \cdot \nabla \nu_\Theta + \mathbf{J}(S) \cdot \nabla \nu_S = \rho^{-1} \mathbf{J}(\Theta) \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta) \quad (24.89)$$

which in turn yields the material evolution of *in situ* density due to neutral diffusion

$$\left[\frac{D\rho}{Dt} \right]_{\text{ntral diff}} = -\mathbf{J}(\Theta) \cdot (\mathcal{T} \nabla p + \mathcal{C} \nabla \Theta). \quad (24.90)$$

24.4.3 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

$$\mathcal{C} = \frac{\partial \alpha}{\partial \Theta} + 2 \frac{\alpha}{\beta} \frac{\partial \alpha}{\partial S} - \left[\frac{\alpha}{\beta} \right]^2 \frac{\partial \beta}{\partial S} \geq 0. \quad (24.91)$$

Given the downgradient nature of the neutral diffusive fluxes, we have

$$\text{Cabbeling} \equiv -\mathcal{C} \mathbf{J}(\Theta) \cdot \nabla \Theta \geq 0, \quad (24.92)$$

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.

24.4.4 Thermobaricity

The thermobaricity parameter

$$\mathcal{T} = \beta \frac{\partial}{\partial p} \left[\frac{\alpha}{\beta} \right] \quad (24.93)$$

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

$$\begin{aligned} \mathcal{T} &= \frac{\partial \alpha}{\partial p} - \frac{\alpha}{\beta} \frac{\partial \beta}{\partial p} \\ &= -\frac{\rho_{\Theta p}}{\rho} + \frac{\rho_{\Theta}}{\rho_S} \frac{\rho_{pS}}{\rho} \end{aligned} \quad (24.94)$$

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)

$$\text{Thermobaricity} \equiv -\mathcal{T} \mathbf{J}(\Theta) \cdot \nabla p \quad (24.95)$$

appearing in equation (24.89). 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.

24.4.5 Comments

Griffies and Greatbatch (2012) discuss the impacts on global mean sea level from thermobaricity and cabbeling as diagnosed from an ocean model. Given that cabbeling always densifies and thermobaricity is also dominated by densification, these processes lead to a general reduction in global mean sea level. *Klocker and McDougall* (2010), *Groeskamp et al.* (2016), and *Groeskamp et al.* (2019) diagnose cabbeling and thermobaricity from observational based measurements, with *Groeskamp et al.* (2019) also offering a more robust numerical method for performing that diagnostic calculation.

Although cabbeling 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 ??). Namely, cabbeling 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, cabbeling 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 cabbeling and thermobaricity. However, the strength of the cabbeling and thermobaricity processes are functions of mesoscale stirring rather than irreversible mixing.

24.5 Neutral helicity

As discussed in Section 20.3, 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 24.3.4. 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*. Consequently, closed paths defined by neutral displacements possess a helical topology.

24.5.1 Mathematical preliminaries

Consider a simply connected smooth surface with outward normal written in the form

$$\hat{\mathbf{n}} = |\mathbf{N}|^{-1} \mathbf{N}. \quad (24.96)$$

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

$$\oint_{\partial\mathcal{S}} \mathbf{N} \cdot \hat{\mathbf{t}} \, dl = 0. \quad (24.97)$$

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 (24.97) to yield

$$\int_S (\nabla \wedge \mathbf{N}) \cdot \hat{\mathbf{n}} dS = 0, \quad (24.98)$$

where dS is the area element in the surface S with outward normal $\hat{\mathbf{n}}$. Since the closed path is arbitrary, the area integral (24.98) vanishes only if the integrand is identically zero. We conclude that for the surface to be simply connected requires that the helicity must vanish

$$\mathcal{H} = \mathbf{N} \cdot (\nabla \wedge \mathbf{N}) = 0 \implies \text{simply connected surface.} \quad (24.99)$$

24.5.2 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

$$\mathbf{N} = -\alpha \nabla \Theta + \beta \nabla S \quad (24.100)$$

rendering the neutral helicity

$$\mathcal{H}_{\text{neutral}} = (-\alpha \nabla \Theta + \beta \nabla S) \cdot \nabla \wedge (-\alpha \nabla \Theta + \beta \nabla S), \quad (24.101)$$

which can be written

$$\mathcal{H}_{\text{neutral}} = (-\alpha \nabla \Theta + \beta \nabla S) \cdot \nabla \wedge (-\alpha \nabla \Theta + \beta \nabla S) \quad (24.102a)$$

$$= -\alpha \nabla \Theta \cdot (\nabla \wedge \beta \nabla S) - \beta \nabla S \cdot (\nabla \wedge \alpha \nabla \Theta) \quad (24.102b)$$

$$= -\alpha \nabla \Theta \cdot (\nabla \beta \wedge \nabla S) - \beta \nabla S \cdot (\nabla \alpha \wedge \nabla \Theta). \quad (24.102c)$$

Expand the gradients of α and β according to

$$\nabla \alpha = \alpha_\Theta \nabla \Theta + \alpha_S \nabla S + \alpha_p \nabla p \quad (24.103a)$$

$$\nabla \beta = \beta_\Theta \nabla \Theta + \beta_S \nabla S + \beta_p \nabla p, \quad (24.103b)$$

so that

$$-\mathcal{H}_{\text{neutral}} = \alpha \nabla \Theta \cdot (\beta_\Theta \nabla \Theta + \beta_p \nabla p) + \beta \nabla S \cdot (\alpha_S \nabla S + \alpha_p \nabla p) \quad (24.104a)$$

$$= \alpha \nabla \Theta \cdot (\nabla p \wedge \nabla S) \beta_p + \beta \nabla S \cdot (\nabla p \wedge \nabla \Theta) \alpha_p \quad (24.104b)$$

$$= \nabla p \cdot (\nabla S \wedge \nabla \Theta) (\alpha \beta_p - \beta \alpha_p). \quad (24.104c)$$

Introducing the thermobaricity parameter from Section 24.4

$$\mathcal{T} = \beta \partial_p (\alpha / \beta) \quad (24.105)$$

renders the tidy result

$$\mathcal{H}_{\text{neutral}} = \beta \mathcal{T} \nabla p \cdot (\nabla S \wedge \nabla \Theta). \quad (24.106)$$

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 2.5)

$$\nabla p \cdot (\nabla S \wedge \nabla \Theta) = \nabla \Theta \cdot (\nabla p \wedge \nabla S) = \nabla S \cdot (\nabla \Theta \wedge \nabla p). \quad (24.107)$$

A nonzero $\mathcal{H}_{\text{neutral}}$ means that a path defined by the accumulation of neutral tangent planes does not close. Rather, they possess a helical structure and the corresponding neutral surfaces are not simply connected.

24.5.3 Comments and further reading

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.

Water masses as bulk regions of the ocean fluid characterized by scalar properties such as Conservative Temperature (Θ), salinity (S), planetary geostrophic potential vorticity, and biogeochemical properties such as dissolved concentrations of carbon, oxygen, and nutrients. Water masses are primed by forcing at select boundary regions and then transformed by irreversible mixing processes in the ocean interior. Water masses provide a coarse grained partitioning of the ocean fluid into conceptually distinct pieces whose origin, evolution, and fate can be measured, studied, and modeled. This approach to describing and understanding the ocean allows one to infer ocean circulation within a water mass *phase space* even without directly measuring the circulation in geographical space. This water mass perspective was comprehensively reviewed by [Groeskamp et al. \(2019\)](#) and it provides the conceptual framework for our survey of ocean tracer mechanics.

In this chapter we focus on methods used to analyze tracer mechanics as revealed from models and observations. These methods include Lagrangian analysis, water mass transformation analysis, stochastic methods, and Green's functions methods.

25.1 Watermass transformation analysis	379
25.2 Tracer transport pathways	379
25.3 Ventilation time scales	379

25.1 Watermass transformation analysis

- Summarize formalism as per [Groeskamp et al. \(2019\)](#)

25.2 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. \(2018\)](#).

25.3 Ventilation time scales

- Summarize formalism from Haine et al review in progress.

Part VI

Geophysical fluid dynamics

We now bring together the many pieces from earlier chapters to develop the basic equations of fluid dynamics for rotating and stratified fluids. That is, we here enter the world of geophysical fluid dynamics.

26

Momentum and energy dynamics

In this chapter, we develop the fluid mechanical equations for momentum and energy 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. They are also based on the application of thermodynamics to a moving fluid (Chapter 19). We make liberal use of results from classical point particle mechanics detailed in Chapter 11 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.

- More on energetics including friction heating
- Develop available internal energy
- More on angular momentum and stress tensor symmetry
- Moist static energy as per Romps' notes

26.1	Continuum fluid equations of motion	384
26.1.1	Body forces	384
26.1.2	Contact forces	384
26.1.3	Equations of motion	385
26.1.4	Spherical coordinate momentum equation	386
26.1.5	Vector-invariant form of the momentum equation	387
26.1.6	Summary of the thermo-hydrodynamical equations	387
26.1.7	Further reading	389
26.2	Mechanical energy budget	389
26.2.1	Kinetic energy	389
26.2.2	Gravitational potential energy	389
26.2.3	Mechanical energy	390
26.3	Total energy budget	390
26.3.1	Mechanical energy plus internal energy	390
26.3.2	Bernoulli function and Bernoulli's theorem	391
26.3.3	Materially constant specific entropy	393
26.3.4	Comments	394
26.4	Moist static energy and atmospheric lapse rate	394

26.4.1 Further reading	394
26.5 Exercises	394

26.1 Continuum fluid equations of motion

We here summarize elements of classical continuum mechanics and in turn apply Newton's second law to derive the equations of motion for a fluid continuum, with application to motion on a rotating and gravitating sphere.

26.1.1 Body forces

Forces acting on an arbitrary volume 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} , multiplied by the mass of the media

$$\mathbf{F}_{\text{body}} = \int \mathbf{f} \rho dV. \quad (26.1)$$

For example, the gravitational force acting on a volume of fluid is given by

$$\mathbf{F}_{\text{gravity}} = \int \mathbf{g} \rho dV, \quad (26.2)$$

where \mathbf{g} is the acceleration of gravity. Likewise, the Coriolis force is given by

$$\mathbf{F}_{\text{Coriolis}} = -2 \int (\boldsymbol{\Omega} \wedge \mathbf{v}) \rho dV. \quad (26.3)$$

26.1.2 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 of the region. The total contact force exerted on the volume V through its boundaries is given by

$$\mathbf{F}_{\text{contact}} = \int \mathbf{T} \cdot \hat{\mathbf{n}} dS, \quad (26.4)$$

where $\hat{\mathbf{n}}$ is the outward normal direction orienting the domain boundary with dS the associated area element, and \mathbf{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 11.

Stresses from friction and pressure

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. In this case, the stress tensor components are given by

$$T^{ab} = \tilde{\tau}^{ab} - p g^{ab}. \quad (26.5)$$

In this equation, p is the pressure, which is a force per unit area. The term g^{ab} is the metric tensor which summarizes the local geometry, and it equals to the Kronecker or unit tensor for Cartesian coordinates in Euclidean space (Section 5.1). The frictional stress tensor is written $\tilde{\tau}^{ab}$. It is also known as the *deviatoric* stress tensor as it represents deviations from the static case when stress is due solely to pressure.

Substitution of the stress tensor (26.5) into the contact force expression (26.4) leads to write

$$\mathbf{F}_{\text{contact}} = \int (\tilde{\tau} \cdot \hat{\mathbf{n}} - p \hat{\mathbf{n}}) dS, \quad (26.6)$$

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, and/or shearing 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.

26.1.3 Equations of motion

The linear momentum of a fluid region is given by

$$\mathbf{P} = \int \rho \mathbf{v} dV. \quad (26.7)$$

Applying Newton's Second Law to the continuum leads to the equation of motion for the region

$$\frac{d}{dt} \int \rho \mathbf{v} dV = \int \rho \mathbf{f} dV + \int \mathbf{T} \cdot \hat{\mathbf{n}} dS. \quad (26.8)$$

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 15.6). Applying Gauss's law (Section 3.7.2) to the area integral yields

$$\frac{d}{dt} \int \rho \mathbf{v} dV = \int (\rho \mathbf{f} + \nabla \cdot \mathbf{T}) dV. \quad (26.9)$$

General form of the equation of motion for a fluid element

Since the volume under consideration is arbitrary, the integral relation (26.9) is satisfied for an arbitrary region. We apply the result to an infinitesimal fluid element in which

$$\frac{D(\rho \mathbf{v} \delta V)}{Dt} = \delta V (\rho \mathbf{f} + \nabla \cdot \mathbf{T}). \quad (26.10)$$

Assuming the mass for the fluid element is constant then leads to the equation of motion

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} + \nabla \cdot \mathbf{T}. \quad (26.11)$$

Momentum equation for a rotating fluid in a gravitational field

We now specialize the momentum equation (26.11) 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 (26.6))

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} - \nabla p + \nabla \cdot \tilde{\boldsymbol{\tau}}. \quad (26.12)$$

Next, move to a rotating terrestrial reference frame and thus expose the Coriolis acceleration and the effective gravitational force (Section 11.2)

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -\rho \nabla \Phi - \nabla p + \nabla \cdot \tilde{\boldsymbol{\tau}}. \quad (26.13)$$

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 10.11). 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 (26.13) 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* equations, though that name is strictly only applicable to the non-rotating case and furthermore with a specific form for the friction operator. We thus refer to it as Newton's Law for a rotating fluid.

26.1.4 Spherical coordinate momentum equation

We now write the spherical coordinate form to the equations of motion. For that purpose, make use of the acceleration as derived in Section 11.2.3 for the point particle, and in particular use geopotential coordinates to measure radial distances. Additionally, the point particle time derivative translates into a material time derivative for fluid elements. We are thus led to

$$\frac{Du}{Dt} + \frac{u(w - v \tan \phi)}{r} + 2\Omega(w \cos \phi - v \sin \phi) = -\frac{1}{\rho r \cos \phi} \frac{\partial p}{\partial \lambda} + F^\lambda \quad (26.14)$$

$$\frac{Dv}{Dt} + \frac{v w + u^2 \tan \phi}{r} + 2\Omega u \sin \phi = -\frac{1}{\rho r} \frac{\partial p}{\partial \phi} + F^\phi \quad (26.15)$$

$$\frac{Dw}{Dt} - \frac{u^2 + v^2}{r} - 2\Omega u \cos \phi = -g - \frac{1}{\rho} \frac{\partial p}{\partial r} + F^r, \quad (26.16)$$

where we introduced the spherical components to the friction acceleration

$$\mathbf{F} = F^\lambda \hat{\boldsymbol{\lambda}} + F^\phi \hat{\boldsymbol{\phi}} + F^r \hat{\boldsymbol{r}}. \quad (26.17)$$

We also note the spherical coordinate form for the gradient operator

$$\nabla = \frac{\hat{\boldsymbol{\lambda}}}{r \cos \phi} \frac{\partial}{\partial \lambda} + \frac{\hat{\boldsymbol{\phi}}}{r} \frac{\partial}{\partial \phi} + \hat{\boldsymbol{r}} \frac{\partial}{\partial r}, \quad (26.18)$$

as well as the material time derivative operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla = \frac{\partial}{\partial t} + \frac{u}{r \cos \phi} \frac{\partial}{\partial \lambda} + \frac{v}{r} \frac{\partial}{\partial \phi} + w \frac{\partial}{\partial r}. \quad (26.19)$$

We can write the spherical momentum equations in a bit more compact form by introducing the spherical coordinate velocity field (see equation (10.38b))

$$\mathbf{v} = \mathbf{u} + \hat{\boldsymbol{r}} w = u \hat{\boldsymbol{\lambda}} + v \hat{\boldsymbol{\phi}} + w \hat{\boldsymbol{r}} \quad (26.20)$$

and the corresponding spherical coordinate acceleration

$$\mathbf{A}_{\text{sphere}} = \frac{D\mathbf{u}}{Dt} \hat{\lambda} + \frac{D\mathbf{v}}{Dt} \hat{\phi} + \frac{D\mathbf{w}}{Dt} \hat{r}. \quad (26.21)$$

We also introduce the expression (10.67c) for the metric acceleration to render

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -\rho \nabla \Phi - \nabla p + \rho \mathbf{F}, \quad (26.22)$$

where we have the acceleration relative to the rotating frame

$$\frac{D\mathbf{v}}{Dt} = \mathbf{A}_{\text{sphere}} + \frac{1}{r} [u \tan \phi (\hat{r} \wedge \mathbf{v}) + w \mathbf{u} - \hat{r} \mathbf{u} \cdot \mathbf{u}]. \quad (26.23)$$

For some purposes it is convenient to combine one piece of the metric acceleration to the Coriolis acceleration to yield

$$\mathbf{A}_{\text{sphere}} + \frac{1}{r} [w \mathbf{u} - \hat{r} \mathbf{u} \cdot \mathbf{u}] + \left(2\boldsymbol{\Omega} + \frac{u \tan \phi \hat{r}}{r} \right) \wedge \mathbf{v} = - \left(\frac{\rho \nabla \Phi + \nabla p}{\rho} \right) + \mathbf{F}. \quad (26.24)$$

26.1.5 Vector-invariant form of the momentum equation

The metric terms appearing in the momentum equation can be cumbersome. As an alternative, we can make use of the identity (equation (3.37)) for the nonlinear self-advection term

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = \boldsymbol{\omega} \wedge \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2, \quad (26.25)$$

where $\boldsymbol{\omega} = \nabla \wedge \mathbf{v}$ is the vorticity (Chapter 37). We derive the corresponding *vector-invariant* form of the momentum equation using Cartesian coordinates and then invoke general covariance (Section 6.1) to extend the result to arbitrary coordinates. Making use of equation (26.25) thus leads to the material acceleration

$$\frac{D\mathbf{v}}{Dt} = \frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega} \wedge \mathbf{v} + \nabla(\mathbf{v} \cdot \mathbf{v})/2 \quad (26.26)$$

so that the momentum equation (26.13) takes on the vector-invariant form

$$\frac{\partial \mathbf{v}}{\partial t} + (2\boldsymbol{\Omega} + \boldsymbol{\omega}) \wedge \mathbf{v} = -\nabla(\Phi + \mathbf{v} \cdot \mathbf{v}/2) + (1/\rho) (-\nabla p + \nabla \cdot \tilde{\tau}). \quad (26.27)$$

26.1.6 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

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -\rho \nabla \Phi - \nabla p + \nabla \cdot \tilde{\tau} \quad \text{momentum} \quad (26.28)$$

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v} \quad \text{mass continuity} \quad (26.29)$$

$$\rho \frac{D\theta}{Dt} = -\nabla \cdot \mathbf{J}(\theta) \quad \text{heat conservation} \quad (26.30)$$

$$\rho \frac{DS}{Dt} = -\nabla \cdot \mathbf{J}(S) \quad \text{matter conservation} \quad (26.31)$$

$$\rho = \rho(S, \theta, p) \quad \text{equation of state.} \quad (26.32)$$

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 identify 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 (26.28). As noted at the end of Section 26.1.3, we write the momentum equation in the form (26.28) 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 22.2 (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 (26.29)). We discussed the many forms of density for the ocean and atmosphere in Section 20.2.
- **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 (19.66) using the density and temperature.
 - For a hydrostatic fluid (Section 27.2), 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 29.2).
- **GEOPOTENTIAL:** The geopotential is specified once the height above an arbitrary reference level is known, as well as the effective gravitational acceleration (Section 11.1.2). For geophysical fluid studies, the reference level is generally taken at the level of a resting sea surface (see Appendix 32). We thus often write the radial coordinate as

$$r = R + z \tag{26.33}$$

where $R = 6.371 \times 10^6$ m is the earth radius (equation (11.9)), and z is the geopotential coordinate measuring the height above sea level.

- **ANGULAR ROTATION:** The earth's angular velocity, $\boldsymbol{\Omega}$, is constant for geophysical fluid studies of concern here. Its value is discussed in Section 10.1.
- **FRICITION:** The friction vector, $\rho\mathbf{F} = \nabla \cdot \tilde{\boldsymbol{\tau}}$, is the divergence of a symmetric and trace-free deviatoric stress tensor, $\tilde{\boldsymbol{\tau}}$ (Section 26.1.2). It is specified in Section 22.2.5 for molecular viscosity. More general forms for the friction vector can be considered for purposes of subgrid-scale modeling.
- **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).

26.1.7 Further reading

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 28.

26.2 Mechanical energy budget

We here develop the evolution equation for the mechanical energy of a fluid on a rotating sphere, starting from the equations of motion (26.14)-(26.16) written in spherical coordinates with a geopotential $\Phi = g z$.

26.2.1 Kinetic energy

Multiplying the zonal momentum equation (26.14) by u , the meridional equation (26.15) by v , and the radial equation (26.16) by w reveals that both the metric acceleration and the Coriolis acceleration drop from the kinetic energy equation. This result is expected since neither terms perform work on a fluid element. Defining the kinetic energy per mass as

$$\mathcal{K} = \frac{\mathbf{v} \cdot \mathbf{v}}{2} \quad (26.34)$$

leads to the material evolution

$$\rho \frac{D\mathcal{K}}{Dt} = -\mathbf{v} \cdot \nabla p - w g \rho + \rho \mathbf{v} \cdot \mathbf{F}. \quad (26.35)$$

The kinetic energy of a fluid element increases in regions where the velocity projects down the pressure gradient, thus having the pressure gradient increase the speed of the fluid element. It is notable that horizontal geostrophic flows (Section 30.3) with

$$f\rho \mathbf{v}_g = \hat{\mathbf{z}} \wedge \nabla p \quad (26.36)$$

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.

26.2.2 Gravitational potential energy

The gravitational potential energy per mass; i.e., the geopotential, for a fluid element is given by

$$\Phi = g z \quad (26.37)$$

so that its material evolution is

$$\rho \frac{D\Phi}{Dt} = w g \rho. \quad (26.38)$$

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.

26.2.3 Mechanical energy

Adding the kinetic and potential energy leads to the mechanical energy equation

$$\rho \frac{Dm}{Dt} = -\mathbf{v} \cdot \nabla p + \rho \mathbf{v} \cdot \mathbf{F}, \quad (26.39)$$

where

$$m = \mathcal{K} + \Phi \quad (26.40)$$

is the mechanical energy per mass for a fluid element. In the absence of pressure and friction we recover the point particle result from Section 12.5. This result is expected since pressure and friction arise from the continuum nature of a fluid and are absent from the dynamics of a point particle.

26.3 Total energy budget

Recall from Section 12.5 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, and as friction dissipates mechanical energy and converts it to heat. We explore these points in this section by combining the mechanical energy budget from Section 26.2 to the internal energy budget from Section 19.7 to study the budget for the total energy of a fluid element.

26.3.1 Mechanical energy plus internal energy

The internal energy per mass, \mathcal{I} , for an adiabatic and constant composition fluid element changes only through pressure work (equation (19.31)), so that

$$\rho \frac{D\mathcal{I}}{Dt} = -p \nabla \cdot \mathbf{v}. \quad (26.41)$$

Hence, the total energy per mass,

$$\mathcal{E} = \mathcal{I} + \mathcal{K}, \quad (26.42)$$

has the material evolution given by

$$\rho \frac{D\mathcal{E}}{Dt} = -\nabla \cdot (p \mathbf{v}). \quad (26.43)$$

Source for total energy of a fluid element

Equation (26.43) reveals that the material time change for the total energy of a fluid element is affected by the convergence of pressure times velocity. So even when incorporating the internal energy and in the absence of dissipation, the fluid element's total energy is not materially constant. The energy source term is fundamental to the continuum. Namely, there is pressure work required for the fluid element to exist within the continuum of other fluid elements. We further describe this mechanical *injection work* in the context of the Bernoulli function in Section 26.3.2.

Eulerian flux-form budget for total energy

Converting the material budget (26.43) into its Eulerian form renders

$$\frac{\partial(\rho \mathcal{E})}{\partial t} + \nabla \cdot [\rho \mathbf{v} (\mathcal{E} + p/\rho)] = 0. \quad (26.44)$$

We thus see that locally the total energy is modified by the advective transport of

$$\mathcal{E} + p/\rho = \mathcal{K} + \mathcal{P} + \mathcal{I} + p\alpha \quad (26.45a)$$

$$= \mathcal{K} + \mathcal{P} + \mathcal{H} \quad (26.45b)$$

$$\equiv \mathcal{B} \quad (26.45c)$$

where

$$\mathcal{H} = \mathcal{I} + p\alpha \quad (26.46)$$

is the enthalpy per mass of the fluid element (Section 19.3.3), and the final equality introduced the Bernoulli function, which is the sum of the enthalpy per mass plus the mechanical energy per mass

$$\mathcal{B} = \mathcal{H} + \mathcal{K} + \mathcal{P} = \mathcal{H} + \mathcal{M}. \quad (26.47)$$

We have more to say in regards to the Bernoulli function in Section 26.3.2. Finally, note that integration over a region with zero boundary transfer of $\mathbf{v} \mathcal{B}$ leads to the conservation of total energy for the region.

26.3.2 Bernoulli function and Bernoulli's theorem

In terms of the Bernoulli function, the total energy equation (26.44) takes the form

$$\frac{\partial(\rho \mathcal{E})}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathcal{B}) = 0. \quad (26.48)$$

This conservation law is distinct from that for trace matter in a perfect fluid, which satisfies the conservation law (see Section 15.4.4)

$$\rho \frac{DC}{Dt} = \frac{\partial(\rho C)}{\partial t} + \nabla \cdot (\rho \mathbf{v} C) = 0. \quad (26.49)$$

That is, the mass per volume of a tracer, ρC , changes at a point due to the advective convergence of the tracer mass per volume onto that point

$$\frac{\partial(\rho C)}{\partial t} = -\nabla \cdot (\rho \mathbf{v} C). \quad (26.50)$$

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

$$\frac{\partial(\rho \mathcal{E})}{\partial t} = -\nabla \cdot (\rho \mathbf{v} \mathcal{B}). \quad (26.51)$$

Why is $\rho \mathcal{E}$ affected by the convergence of $\rho \mathbf{v} \mathcal{B}$ rather than the convergence of $\rho \mathbf{v} \mathcal{E}$?

The mechanical injection 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 $p/\rho = p\alpha$. So what is $p\alpha$? Imagine carving out a unit mass from within a continuous fluid with pressure p and specific volume α , leaving behind a “hole”. The mechanical work required to do so is $p\alpha$. Correspondingly, we interpret $p\alpha$ as the mechanical work required to inject a unit mass of fluid with specific volume α into a region of pressure p . We thus refer to $p\alpha$ as the *injection work*. Recall that the specific enthalphy is the sum of the specific internal energy plus the injection work (equation (26.46)). Hence, specific enthalphy provides a measure of the non-mechanical energy required for a fluid element to exist within a continuum fluid.

Bernoulli's theorem

Consider a fluid flow in steady state (vanishing Eulerian time derivatives). Steady state mass conservation means that

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}) = 0. \quad (26.52)$$

This relation, along with a steady state energy in equation (26.51), means that the steady state velocity field is parallel to contours of constant Bernoulli function

$$\mathbf{v} \cdot \nabla B = 0. \quad (26.53)$$

We thus see that for the perfect fluid to be in steady state flow, the Bernoulli function, which equals the total energy plus the injection work, is constant along streamlines. Hence, as the fluid moves along a streamline, there is an exchange of between the total energy per mass, \mathcal{E} , and the injection work, $p\alpha$, such that their sum remains constant.

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 (26.13) into its vector-invariant form by making use of the vector identity (see Section 3.3.4)²

$$\boldsymbol{\omega} \wedge \mathbf{v} = -\mathcal{K} + (\mathbf{v} \cdot \nabla) \mathbf{v}. \quad (26.54)$$

This identity allows us to eliminate velocity self-advection in favor of the vorticity and kinetic energy per mass

$$\frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega}_a \wedge \mathbf{v} = -\frac{1}{\rho} \nabla p - \nabla m, \quad (26.55)$$

where

$$\boldsymbol{\omega}_a = \boldsymbol{\omega} + 2 \boldsymbol{\Omega} \quad (26.56)$$

is the absolute vorticity (see Chapter 37) and we set the frictional stress tensor, $\tilde{\tau}$, to zero since we are assuming a perfect fluid. The Eulerian time evolution for the kinetic energy per mass is therefore given by

$$\frac{\partial \mathcal{K}}{\partial t} = -\frac{1}{\rho} \mathbf{v} \cdot \nabla p - \mathbf{v} \cdot \nabla m, \quad (26.57)$$

¹This argument follows Section 13.5.4 of [Thorne and Blandford \(2017\)](#).

²We pursue the same manipulations in Section 37.7.1 when deriving the vorticity equation.

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 26.3.3)

$$\frac{D\mathcal{H}}{Dt} = \frac{1}{\rho} \frac{Dp}{Dt}. \quad (26.58)$$

Hence, a steady state perfect fluid maintains the balance

$$\rho(\mathbf{v} \cdot \nabla)\mathcal{H} = (\mathbf{v} \cdot \nabla)p, \quad (26.59)$$

so that the time tendency for the specific kinetic energy is given by

$$\frac{\partial \mathcal{K}}{\partial t} = -\mathbf{v} \cdot \nabla \mathcal{B}. \quad (26.60)$$

We thus see that for a steady state, the Bernoulli function is materially constant since

$$\frac{D\mathcal{B}}{Dt} = \mathbf{v} \cdot \nabla \mathcal{B} = 0 \quad \text{steady state.} \quad (26.61)$$

26.3.3 Materially constant specific entropy

In our discussion of thermodynamics in Chapter 19, 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 this 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 basic physical result.

For that purpose, make use of the fundamental thermodynamic relation in the form of equation (19.47c)

$$d\mathcal{H} = T dS + \alpha dp, \quad (26.62)$$

where we assume a uniform matter composition so to set $dC = 0$. Applying this relation to a material fluid element leads to

$$T\rho \frac{D\mathcal{S}}{Dt} = \rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt}. \quad (26.63)$$

The Legendre transformation for specific enthalpy (equation (19.46)) leads to

$$\rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt} = \rho \frac{D\mathcal{J}}{Dt} - \frac{p}{\rho} \frac{D\rho}{Dt}. \quad (26.64)$$

Use of the continuity equation (15.8c) renders

$$\rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt} = \rho \frac{D\mathcal{J}}{Dt} + p \nabla \cdot \mathbf{v}, \quad (26.65)$$

and further use of the First Law in the form of equation (26.41) yields

$$\rho \frac{D\mathcal{H}}{Dt} - \frac{Dp}{Dt} = 0. \quad (26.66)$$

Making use of this result in equation (26.63) implies that specific entropy is indeed materially constant

$$\frac{D\mathcal{S}}{Dt} = 0. \quad (26.67)$$

The key step in deriving this result sits with equation (26.41) for the First Law, whereby we removed contributions from heating and matter composition changes, again since we are making the perfect fluid assumption. Hence, although a rather circular argument, it is important to verify that the circle indeed closes.

26.3.4 Comments

Our approach to the conservation of total energy is postulatory. That is, we assume there is a total energy for the system that is conserved. 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 \(2017\)](#).

26.4 Moist static energy and atmospheric lapse rate

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 (19.86) for the special case in which entropy changes occur through heating so that

$$T dS = dQ. \quad (26.68)$$

Now assume heating occurs through radiation plus another term whose

$$T dS = dQ \quad (26.69a)$$

$$= L dT + dQ_{\text{rad}}, \quad (26.69b)$$

where L is the latent heat of vaporization.

26.4.1 Further reading

This section is incomplete. The intent is to merge material from Section 1.10 of [Vallis \(2017\)](#) to lecture notes from David Romps.

26.5 Exercises

EXERCISE 26.1: AXIAL ANGULAR MOMENTUM

As in our discussion of a point particle in Section 12.6, the axial angular momentum of a fluid element is given by

$$L^z = (\rho \delta V) r_{\perp} (u + r_{\perp} \Omega) \equiv (\rho \delta V) l^z \quad (26.70)$$

where $l^z = r_{\perp} (u + r_{\perp} \Omega)$ is the angular momentum per unit mass. The distance to the polar rotation axis, $r_{\perp} = r \cos \phi$, is the moment-arm for determining the torques acting on a fluid element.

- (a) Consider a constant mass fluid element in the absence of friction. Show that the zonal momentum equation (26.14) implies that the material evolution of axial angular momentum per mass is given by

$$\frac{Dl^z}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial \lambda}. \quad (26.71)$$

Hint: read section 2.2.7 of [Vallis \(2017\)](#).

- (b) Consider a fluid element at rest in a fluid with zero zonal pressure gradient. Move the fluid element towards the polar rotation axis along a line of constant latitude. What happens to the zonal velocity? Hint: read the discussion of angular momentum for a particle in Section 12.7.

- (c) Consider a fluid element at rest in a fluid with zero zonal pressure gradient. Move the fluid element poleward while keeping the radial position constant. What happens to the zonal velocity? Hint: read Section 12.7
- (d) Give a very brief symmetry argument for why the angular momentum is materially conserved when $\partial p / \partial \lambda = 0$. Hint: recall the discussion of Noether's Theorem in Section 12.1.1.

EXERCISE 26.2: REGIONAL BUDGET FOR AXIAL ANGULAR MOMENTUM

We here extend the considerations from Exercise 26.1 to develop the budget for axial angular momentum over a finite region. We also allow for irreversible momentum stresses within the fluid and at the boundaries. Elements of this exercise are motivated by *Hughes and de Cueves (2001)*.

- (a) Show that the zonal momentum equation (26.14) implies that the flux-form Eulerian evolution of axial angular momentum per mass is given by

$$\frac{\partial(\rho l^z)}{\partial t} + \nabla \cdot (\rho \mathbf{v} l^z) = -\frac{\partial p}{\partial \lambda} + r_{\perp} \rho F^{\lambda}. \quad (26.72)$$

The distance to the polar rotation axis, $r_{\perp} = r \cos \phi$, is the moment-arm for determining the torques acting on a fluid element.

- (b) Vertically integrate the budget (26.72) over a column of ocean fluid to derive the column-integrated angular momentum budget

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} l^z \rho dz \right] + \nabla_z \cdot \left[\int_{-H}^{\eta} l^z \mathbf{u} \rho dz \right] = [l^z Q_m]_{z=\eta} + \int_{-H}^{\eta} \left[-\frac{\partial p}{\partial \lambda} + r_{\perp} \rho F^{\lambda} \right] dz. \quad (26.73)$$

Hint: make use of Leibnitz's Rule (Section 15.6.4), the surface kinematic boundary condition (15.96) for the ocean, and the bottom kinematic boundary condition (15.69).

- (c) Consider a linearized steady state in which we drop time tendencies and advection. Also ignore changes to angular momentum from boundary mass transport, and assume the moment arm $r_{\perp} = r \cos \phi \approx R \cos \phi$. Show that the angular momentum balance reduces to

$$\frac{\partial}{\partial \lambda} \left[\int_{-H}^{\eta} p dz \right] - p_a \frac{\partial \eta}{\partial \lambda} - p_b \frac{\partial H}{\partial \lambda} = R \cos \phi \int_{-H}^{\eta} \rho F^{\lambda} dz. \quad (26.74)$$

- (d) Assume the friction is given just by the vertical transfer of zonal momentum according to

$$\rho F^{\lambda} = \frac{\partial \tau^{\lambda}}{\partial z}, \quad (26.75)$$

where τ^{λ} is the zonal stress. Zonally integrate the linearized angular momentum budget (26.74) between two boundaries where the ocean thickness vanishes (i.e., along a sloping beach), or around a zonally periodic domain. Show that the resulting zonal and depth integrated balance takes the form

$$\int R \cos \phi \left[\frac{p_a}{R \cos \phi} \frac{\partial \eta}{\partial \lambda} + \tau_a^{\lambda} \right] d\lambda = \int R \cos \phi \left[-\frac{p_b}{R \cos \phi} \frac{\partial H}{\partial \lambda} + \tau_b^{\lambda} \right] d\lambda. \quad (26.76)$$

In this equation, τ_a^λ is the zonal stress at the ocean surface imparted through irreversible interactions between the ocean and the winds or sea ice, and τ_b^λ is the zonal stress at the ocean bottom imparted through irreversible interactions between the ocean and the solid-earth. The balance (26.76) is thus between pressure form drag (discussed in Section 34.2) and irreversible stresses acting at the surface and bottom.

There are a variety of interesting limits suggested by the balance (26.76). For example, consider an ocean with no zonal variations in its topography ($\partial H/\partial \lambda = 0$) and zero mass atmosphere ($p_a = 0$). The corresponding angular momentum balance is between the zonal wind stress, τ_a^λ , and zonal bottom stress $-\tau_b^\lambda$. However, when $\partial H/\partial \lambda \neq 0$, then the surface wind stress is balanced by both bottom stress and bottom pressure form drag.

EXERCISE 26.3: FRICTIONAL DISSIPATION FROM VISCOSITY

Assume the friction in the momentum equation takes the form

$$\rho \mathbf{F} = \nabla \cdot (\rho \kappa \nabla \mathbf{v}) = \partial_n (\rho \kappa \partial_n \mathbf{v}), \quad (26.77)$$

with $\kappa > 0$ a scalar kinematic viscosity (generally non-constant). Show that when integrated over the full domain

$$\int \mathbf{F} \cdot \mathbf{v} \rho dV < 0, \quad (26.78)$$

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

$$\rho \mathbf{F} \cdot \mathbf{v} = \rho F_m v_m = \partial_n (\rho \kappa \partial_n v_m) v_m. \quad (26.79)$$

EXERCISE 26.4: CROCCO'S THEOREM

Show that the spatial gradient of the Bernoulli function for a single-component steady perfect fluid can be written

$$\nabla \mathcal{B} = T \nabla \mathcal{S} + \mathbf{v} \wedge \boldsymbol{\omega}_a. \quad (26.80)$$

This result is known as Crocco's Theorem.

Hint: study the discussion in Section 26.3.2 where we showed that the Bernoulli function is constant along a steady flow streamline in a perfect fluid. Also recall from Section 19.3.3 that each differential in the fundamental thermodynamic relation $d\mathcal{H} = T d\mathcal{S} + \alpha dp$ 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

$$d\mathcal{H} = T d\mathcal{S} + \alpha dp \implies \nabla \mathcal{H} = T \nabla \mathcal{S} + \alpha \nabla p. \quad (26.81)$$

27

Approximate dynamical equations

The thermo-hydrodynamical equations (26.28)-(26.32) 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.

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.

27.1	Three approximations for the primitive equations	398
27.1.1	Hydrostatic balance	398
27.1.2	Shallow fluid approximation	398
27.1.3	Traditional approximation	398
27.1.4	Summary of the primitive equations	399
27.1.5	Comments and further reading	400
27.2	Scaling for the hydrostatic balance	400
27.2.1	Preliminaries	400
27.2.2	Scaling relations	401
27.2.3	Stratification effects on hydrostatic scaling	402
27.2.4	An example hydrostatic pressure gradient	405
27.2.5	Further reading	406
27.3	Tangent plane approximations: f -plane and β -plane	406
27.3.1	Tangent plane approximation	406
27.3.2	Traditional approximation and the f -plane	407
27.3.3	β -plane approximation	407
27.4	Evolving the non-hydrostatic fluid forward in time	408
27.4.1	Compressible perfect fluid equations	408
27.4.2	Comments about non-hydrostatic ocean models	410
27.4.3	Further reading	410
27.5	Exercises	410

27.1 Three approximations for 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 are based on the following three approximations.

27.1.1 Hydrostatic balance

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. This balance is also very closely maintained for the larger scales in a moving geophysical fluid (Section 27.2). The hydrostatic balance therefore forms a central position in the study of geophysical fluid dynamics.

Mathematically, the hydrostatic balance represents a balance in the vertical momentum equation (26.16) 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 the these fluids is small relative to the earth's radius. The shallow fluid approximation¹ 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 dis-

¹The shallow fluid approximation is distinct from the *shallow water approximation* treated in Part VII.

cussed in Section 10.11.6

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{Z}} \quad (27.5a)$$

$$= \Omega (\hat{\phi} \cos \phi + \hat{\mathbf{r}} \sin \phi) \quad (27.5b)$$

$$\approx \Omega \sin \phi \hat{\mathbf{r}} \quad (27.5c)$$

$$= \mathbf{f}/2, \quad (27.5d)$$

where

$$\mathbf{f} = (2\Omega \sin \phi) \hat{\mathbf{r}} \quad (27.6)$$

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 27.1).

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

$$\nabla = \frac{\hat{\lambda}}{R \cos \phi} \frac{\partial}{\partial \lambda} + \frac{\hat{\phi}}{R} \frac{\partial}{\partial \phi} + \hat{\mathbf{z}} \frac{\partial}{\partial z}. \quad (27.10)$$

We can write these equations in the succinct form

$$\left(\frac{D}{Dt} + (f + u \tan \phi / R) \hat{\mathbf{z}} \wedge \right) \mathbf{u} = -\rho \nabla \Phi - \nabla p + \mathbf{F}, \quad (27.11)$$

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

$$\frac{D\mathbf{u}}{Dt} = \hat{\lambda} \frac{Du}{Dt} + \hat{\phi} \frac{Dv}{Dt}. \quad (27.13)$$

27.1.5 Comments and further reading

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 remain the basis for nearly all general circulation models of the ocean and atmosphere.

27.2 Scaling for the hydrostatic balance

The hydrostatic balance consists of the vertical momentum equation for a static fluid in a gravitational field

$$\frac{\partial p}{\partial z} = -\rho g. \quad (27.14)$$

Vertically integrating this equation over the depth of the ocean leads to

$$p(z) = p(\eta) + g \int_z^\eta \rho dz, \quad (27.15)$$

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

$$p(z) = g \int_z^{z_{\text{top}}} \rho dz, \quad (27.16)$$

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

$$\alpha_{\text{aspect}} \equiv \frac{H}{L} \ll 1, \quad (27.17)$$

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.

27.2.1 Preliminaries

Consider the vertical momentum equation (27.76) from the tangent plane and Traditional approximations, along with the associated scales for the various terms

$$\frac{Dw}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g \quad (27.18)$$

$$\frac{W}{T} + \frac{U W}{L} + \frac{W W}{H} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g. \quad (27.19)$$

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

$$10^{-6} \text{ m s}^{-2} \sim -\frac{1}{\rho} \frac{\partial p}{\partial z} - g. \quad (27.20)$$

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 $\partial p/\partial z = -\rho g$.

27.2.2 Scaling relations

We now proceed more formally by non-dimensionalizing the non-rotating inviscid oceanic Boussinesq momentum equations

$$\frac{D\mathbf{u}}{Dt} = -\nabla_z \psi \quad (27.21)$$

$$\frac{Dw}{Dt} = -\frac{\partial \psi}{\partial z} + b \quad (27.22)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (27.23)$$

where

$$\nabla_z = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} \quad (27.24)$$

is the horizontal gradient operator. For the present analysis, ignore the potential impacts from vertical stratification (we include stratification in Section 27.2.3). To non-dimensionalize, introduce the dimensional scales (capital letters) and corresponding non-dimensional variables (hat variables) according to

$$(x, y) = L(\hat{x}, \hat{y}) \quad z = H \hat{z} \quad \mathbf{u} = U \hat{\mathbf{u}} \quad w = W \hat{w} \quad (27.25)$$

$$t = T \hat{t} \quad \psi = \Psi \hat{\psi} \quad b = B \hat{b}. \quad (27.26)$$

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

$$T = \frac{L}{U}. \quad (27.27)$$

Secondly, scale the vertical velocity according to the continuity equation

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \Rightarrow W \sim \frac{UH}{L} = U\alpha_{\text{aspect}}. \quad (27.28)$$

Third, scale the pressure according to the non-rotating balance of the horizontal advection and the horizontal pressure gradient

$$\frac{UU}{L} \sim \frac{\Psi}{L} \Rightarrow \Psi \sim U^2. \quad (27.29)$$

Fourth, scale buoyancy according to the hydrostatic balance

$$B \sim \frac{\Psi}{H} = \frac{U^2}{H}. \quad (27.30)$$

Bringing these results together leads to

$$\frac{U}{T} \frac{\partial \hat{u}}{\partial \hat{t}} + \frac{U^2}{L} \hat{u} \frac{\partial \hat{u}}{\partial \hat{x}} + \frac{U^2}{L} \hat{v} \frac{\partial \hat{u}}{\partial \hat{y}} + \frac{UW}{H} \hat{w} \frac{\partial \hat{u}}{\partial \hat{z}} = -\frac{\Psi}{L} \frac{\partial \hat{\psi}}{\partial \hat{x}} \quad (27.31)$$

$$\frac{U}{T} \frac{\partial \hat{v}}{\partial \hat{t}} + \frac{U^2}{L} \hat{u} \frac{\partial \hat{v}}{\partial \hat{x}} + \frac{U^2}{L} \hat{v} \frac{\partial \hat{v}}{\partial \hat{y}} + \frac{UW}{H} \hat{w} \frac{\partial \hat{v}}{\partial \hat{z}} = -\frac{\Psi}{L} \frac{\partial \hat{\psi}}{\partial \hat{y}} \quad (27.32)$$

$$\frac{W}{T} \frac{\partial \hat{w}}{\partial \hat{t}} + \frac{UW}{L} \hat{u} \frac{\partial \hat{w}}{\partial \hat{x}} + \frac{UW}{L} \hat{v} \frac{\partial \hat{w}}{\partial \hat{y}} + \frac{WW}{H} \hat{w} \frac{\partial \hat{w}}{\partial \hat{z}} = -\frac{\Psi}{H} \frac{\partial \hat{\psi}}{\partial \hat{z}} + B \hat{b}. \quad (27.33)$$

Tidying up these relations leads to the non-dimensional oceanic Boussinesq equations

$$\frac{D\hat{u}}{Dt} = -\hat{\nabla} \hat{\psi} \quad (27.34)$$

$$\alpha_{\text{aspect}}^2 \frac{D\hat{w}}{Dt} = -\frac{\partial \hat{\psi}}{\partial \hat{z}} + \hat{b}, \quad (27.35)$$

where we introduced the non-dimensional material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \hat{\nabla} \quad (27.36)$$

and the non-dimensional gradient operator

$$\hat{\nabla} = \frac{\partial}{\partial \hat{x}} + \frac{\partial}{\partial \hat{y}} + \frac{\partial}{\partial \hat{z}}. \quad (27.37)$$

The non-dimensional vertical momentum equation (27.35) reveals that for small aspect ratio flow, the vertical momentum equation reduces to the hydrostatic balance

$$\frac{\partial \hat{\psi}}{\partial \hat{z}} = \hat{b} \quad \text{for } \alpha_{\text{aspect}}^2 \ll 1. \quad (27.38)$$

27.2.3 Stratification effects on hydrostatic scaling

Stratification impacts on how the Boussinesq equations scale. To include stratification requires us to include the thermodynamics equation, which for the Boussinesq system is the evolution of

buoyancy

$$\frac{D\mathbf{u}}{Dt} = -\nabla_z \psi \quad (27.39)$$

$$\frac{Dw}{Dt} = -\frac{\partial \psi}{\partial z} + b \quad (27.40)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (27.41)$$

$$\frac{Db}{Dt} = 0, \quad (27.42)$$

where we ignore irreversible effects on buoyancy. To help isolate the dynamically important portion of stratification, write the buoyancy as

$$b = b'(x, y, z, t) + \tilde{b}(z), \quad (27.43)$$

where $\tilde{b}(z)$, is a static background stratification that is in hydrostatic balance with a corresponding portion of the pressure field

$$\frac{\partial \tilde{\psi}}{\partial z} = \tilde{b}(z). \quad (27.44)$$

The Boussinesq equations thus take the form

$$\frac{D\mathbf{u}}{Dt} = -\nabla_z \psi' \quad (27.45)$$

$$\frac{Dw}{Dt} = -\frac{\partial \psi'}{\partial z} + b' \quad (27.46)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (27.47)$$

$$\frac{Db'}{Dt} = -w N^2, \quad (27.48)$$

where

$$N^2 = \frac{db}{dz} \quad (27.49)$$

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

$$(x, y) = L(\hat{x}, \hat{y}) \quad z = H\hat{z} \quad \mathbf{u} = U\hat{\mathbf{u}} \quad w = W\hat{w} \quad (27.50)$$

$$t = T\hat{t} \quad \psi' = \Psi\hat{\psi}' \quad b' = B\hat{b}' \quad N^2 = \bar{N}^2\hat{N}^2. \quad (27.51)$$

As for the previous case in Section 27.2.2, 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

$$w = W\hat{w} = \epsilon \left[\frac{HU}{L} \right] \hat{w}. \quad (27.52)$$

Inserting into the stratified Boussinesq equations then leads to the non-dimensional system

$$\frac{D\hat{u}}{Dt} = -\hat{\nabla}\hat{\psi}' \quad (27.53)$$

$$\epsilon \alpha_{\text{aspect}}^2 \frac{D\hat{w}}{Dt} = -\frac{\partial \hat{\psi}'}{\partial \hat{z}} + \hat{b}' \quad (27.54)$$

$$\left[\frac{U^2}{N^2 H^2} \right] \frac{D\hat{b}'}{Dt} + \epsilon \hat{N}^2 \hat{w} = 0 \quad (27.55)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}} + \epsilon \frac{\partial \hat{w}}{\partial \hat{z}} = 0 \quad (27.56)$$

where we introduced the non-dimensional material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}} \cdot \hat{\nabla}_z + \epsilon \hat{w} \frac{\partial}{\partial \hat{z}}. \quad (27.57)$$

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

$$\epsilon = \text{Fr}^2 = \frac{U^2}{N^2 H^2}. \quad (27.58)$$

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 $\text{Fr} > 1$ a common indicator of hydraulic instability (see Exercise 33.3 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

$$\text{Ri} = \text{Fr}^{-2} = \frac{N^2 H^2}{U^2}. \quad (27.59)$$

It is a matter of taste whether one works with Fr or Ri .

With this choice the vertical velocity scale is given by

$$W = \text{Fr}^2 \left[\frac{HU}{L} \right]. \quad (27.60)$$

We see that for $\text{Fr} < 1$ that the stratification reduces the scale for the vertical velocity. With this choice for ϵ , the non-dimensional Boussinesq equations now take the form

$$\frac{D\hat{u}}{Dt} = -\hat{\nabla}_z \hat{\psi}' \quad (27.61)$$

$$\text{Fr}^2 \alpha_{\text{aspect}}^2 \frac{D\hat{w}}{Dt} = -\frac{\partial \hat{\psi}'}{\partial \hat{z}} + \hat{b}' \quad (27.62)$$

$$\frac{D\hat{b}'}{Dt} + \hat{N}^2 \hat{w} = 0 \quad (27.63)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}} + \text{Fr}^2 \frac{\partial \hat{w}}{\partial \hat{z}} = 0. \quad (27.64)$$

The condition for hydrostatic balance in a stratified fluid thus takes the form

$$\text{Fr}^2 \alpha_{\text{aspect}}^2 \ll 1. \quad (27.65)$$

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

$$\left| \hat{\nabla} \cdot \hat{\mathbf{u}} \right| = \left| \text{Fr}^2 \frac{\partial \hat{w}}{\partial \hat{z}} \right| \ll \left| \frac{\partial \hat{w}}{\partial \hat{z}} \right|. \quad (27.66)$$

27.2.4 An example hydrostatic pressure gradient

We offer an example of how horizontal differences in density lead to horizontal pressure gradients through the hydrostatic relation. Although somewhat trivial, 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.

For this purpose we consider two adjacent columns of fluid with equal mass but with distinct density, and assume the density in each column is constant throughout the respective columns. We display an schematic in Figure 27.1, where we make the additional assumption that the two fluid columns sit on a flat surface. Since they have equal mass, the hydrostatic pressures at the bottom of the two columns are equal and given by

$$p_{\text{bot}} = g \rho_A (\eta_A + H) = g \rho_B (\eta_B + H), \quad (27.67)$$

where $z = -H$ is the vertical position at the bottom and $z = \eta$ is the surface. The hydrostatic pressure at an arbitrary position within column A is given by

$$p_A(z) = g \rho_A (\eta_A - z) = p_{\text{bot}} - g \rho_A (H + z), \quad (27.68)$$

where we substituted the bottom pressure from equation (27.67) to eliminate the surface height η_A . Likewise, the pressure in column B is given by

$$p_B(z) = g \rho_B (\eta_B - z) = p_{\text{bot}} - g \rho_B (H + z). \quad (27.69)$$

We can now take the difference between the two hydrostatic pressures to find

$$p_B(z) - p_A(z) = g (H + z) (\rho_A - \rho_B) > 0. \quad (27.70)$$

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

$$F_{\text{pressure } B \text{ to } A} = \int_{-H}^{\eta_A} [p_B(z) - p_A(z)] dz = (g/2) (\rho_A - \rho_B) (\eta_A + H)^2. \quad (27.71)$$

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 hydrostatic pressures are zero. 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.

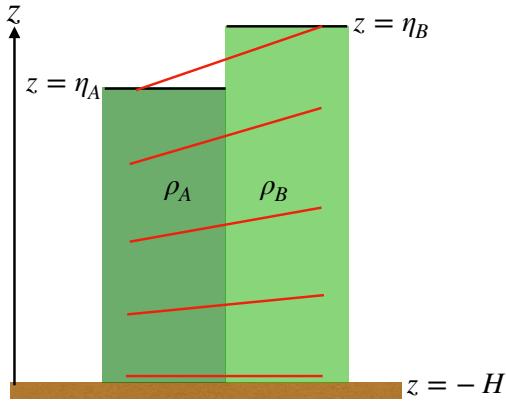


Figure 27.1: Two water columns on a flat bottom with equal mass but different densities with $\rho_A > \rho_B$. We assume the atmosphere above the columns has negligible mass. Furthermore, the cross-sectional area of the two columns are the same so that column B has more volume and so has a greater thickness: $\eta_B > \eta_A$. Since the column masses are the same, the hydrostatic pressures (weight per horizontal area) at the bottom of the two columns are equal: $p_A(z = -H) = p_B(z = -H)$. However, at any position z above the bottom, equation (27.70) 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, which are horizontal next to the bottom but which slope upward to the right moving towards the surface. This configuration provides the essential hydrostatic pressure considerations for the slightly more complex reduced gravity example in Figure 33.4.

27.2.5 Further reading

Section 2.7.4 in [Vallis \(2017\)](#) provides examples of scales over which the hydrostatic relation remains a useful approximation in geophysical fluids.

27.3 Tangent plane approximations: f -plane and β -plane

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.

27.3.1 Tangent plane approximation

Consider a tangent plane located at latitude $\phi = \phi_0$ and introduce a Cartesian set of coordinates according to

$$(x, y, z) = (R \lambda \cos \phi_0, R (\phi - \phi_0), z) \quad (27.72)$$

$$(\hat{x}, \hat{y}, \hat{z}) = (\hat{\lambda}, \hat{\phi}, \hat{r}). \quad (27.73)$$

Use of these Cartesian coordinates leads to the following inviscid (i.e., no friction) equations of motion

$$\frac{Du}{Dt} + 2(\Omega^y w - \Omega^z v) = -\frac{1}{\rho} \frac{\partial p}{\partial x} \quad (27.74)$$

$$\frac{Dv}{Dt} + 2(\Omega^z u - \Omega^x w) = -\frac{1}{\rho} \frac{\partial p}{\partial y} \quad (27.75)$$

$$\frac{Dw}{Dt} + 2(\Omega^x v - \Omega^y u) = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g, \quad (27.76)$$

with rotational vector components

$$\boldsymbol{\Omega} = \Omega (\cos \phi_0 \hat{\mathbf{y}} + \sin \phi_0 \hat{\mathbf{z}}). \quad (27.77)$$

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 27.4 and Section 33.3).

27.3.2 Traditional approximation and the *f*-plane

The Traditional approximation retains only the local vertical component of the rotation vector, resulting in

$$\frac{Du}{Dt} - fv = -\frac{1}{\rho} \frac{\partial p}{\partial x} \quad (27.78)$$

$$\frac{Dv}{Dt} + fu = -\frac{1}{\rho} \frac{\partial p}{\partial y} \quad (27.79)$$

$$\frac{Dw}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g, \quad (27.80)$$

where we introduced the constant Coriolis parameter

$$f = 2\Omega \sin \phi_0 \equiv f_0. \quad (27.81)$$

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.

27.3.3 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

$$f = f_0 + R^{-1} (2\Omega \cos \phi_0) (y - y_0), \quad (27.82)$$

or more simply

$$f = f_0 + \beta y \quad (27.83)$$

$$\beta = \frac{\partial f}{\partial y} = \frac{2 \Omega \cos \phi_0}{R}, \quad (27.84)$$

which constitutes the β -plane approximation.

27.4 Evolving the non-hydrostatic fluid forward in time

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 (26.28)-(26.32). 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.

27.4.1 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 31, and in particular the vertical Lagrangian-remap algorithm in Section 31.4. 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

$$\partial_t(\rho u) + \nabla \cdot [(\rho u) v] = -\partial_x p \quad (27.85)$$

$$\partial_t(\rho v) + \nabla \cdot [(\rho v) v] = -\partial_y p \quad (27.86)$$

$$\partial_t(\rho w) + \nabla \cdot [(\rho w) v] = -\partial_z p - \rho g \quad (27.87)$$

$$\partial_t \rho + \nabla \cdot (\rho v) = 0 \quad (27.88)$$

$$\partial_t(\rho \theta) + \nabla \cdot (\rho \theta v) = 0 \quad (27.89)$$

$$\partial_t(\rho S) + \nabla \cdot (\rho S v) = 0 \quad (27.90)$$

$$\rho = \rho(S, \theta, p). \quad (27.91)$$

As seen in our discussion of scaling for the hydrostatic balance in Section 27.2, 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 perturbation²

$$p = p_h + p_{nh} \quad (27.92)$$

where

$$\frac{\partial p}{\partial z} = -\rho g + \frac{\partial p_{nh}}{\partial z}. \quad (27.93)$$

²We perform a similar decomposition of pressure when writing the Boussinesq momentum equation in Section 29.1.2.

Making use of this split allows us to write the momentum equations (27.85)-(27.87) in the form

$$\partial_t(\rho u) + \partial_x[(\rho u) u] + \partial_y[(\rho u) v] + \partial_z[(\rho u) w] = -\partial_x p \quad (27.94)$$

$$\partial_t(\rho v) + \partial_x[(\rho v) u] + \partial_y[(\rho v) v] + \partial_z[(\rho v) w] = -\partial_y p \quad (27.95)$$

$$\partial_t(\rho w) + \partial_x[(\rho w) u] + \partial_y[(\rho w) v] + \partial_z[(\rho w) w] = -\partial_z p_{nh}. \quad (27.96)$$

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

$$\partial_t(\rho u) = -\partial_x p - \partial_x[(\rho u) u] - \partial_y[(\rho u) v] - \partial_z[(\rho u) w] \quad (27.97)$$

$$\partial_t(\rho v) = -\partial_y p - \partial_x[(\rho v) u] - \partial_y[(\rho v) v] - \partial_z[(\rho v) w] \quad (27.98)$$

$$\partial_t(\rho w) = -\partial_z p_{nh} - \partial_x[(\rho w) u] - \partial_y[(\rho w) v] - \partial_z[(\rho w) w] \quad (27.99)$$

$$\partial_t(\rho \theta) = -\nabla \cdot (\rho \theta \mathbf{v}) \quad (27.100)$$

$$\partial_t(\rho S) = -\nabla \cdot (\rho S \mathbf{v}) \quad (27.101)$$

$$\partial_t \rho = -\nabla \cdot (\rho \mathbf{v}) \quad (27.102)$$

$$\rho = \rho(S, \theta, p). \quad (27.103)$$

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 equations (27.97)-(27.102), 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 (27.97)-(27.99) 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 (27.100)-(27.101) 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 (27.102), and then dividing to get

$$\mathbf{v}(\tau + \Delta\tau) = \frac{(\rho \mathbf{v})(\tau + \Delta\tau)}{\rho(\tau + \Delta\tau)} \quad (27.104)$$

$$\theta(\tau + \Delta\tau) = \frac{(\rho \theta)(\tau + \Delta\tau)}{\rho(\tau + \Delta\tau)} \quad (27.105)$$

$$S(\tau + \Delta\tau) = \frac{(\rho S)(\tau + \Delta\tau)}{\rho(\tau + \Delta\tau)}. \quad (27.106)$$

The updated density also allows us to update the hydrostatic pressure. Finally, the equation of state (27.103) is used to update the full pressure field through

$$\rho(\tau + \Delta\tau) = \rho[S(\tau + \Delta\tau), \theta(\tau + \Delta\tau), p(\tau + \Delta\tau)]. \quad (27.107)$$

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 $\tau + 2\Delta\tau$.

27.4.2 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 19.2). 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.

27.4.3 Further reading

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.

27.5 Exercises

EXERCISE 27.1: PRIMITIVE EQUATIONS AND AXIAL ANGULAR MOMENTUM

The axial angular momentum of a fluid element satisfying the primitive equations is given by

$$L^z = (\rho \delta V) R_\perp (u + R_\perp \Omega) \equiv (\rho \delta V) l^z \quad (27.108)$$

where

$$R_\perp = R \cos \phi \quad (27.109)$$

is the distance from the polar rotation axis to a point on the sphere with radius R , and

$$l^z = R_\perp (u + R_\perp \Omega) \quad (27.110)$$

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

$$\frac{Dl^z}{Dt} = -\frac{1}{\rho} \frac{\partial p}{\partial \lambda}. \quad (27.111)$$

- (b) Move the fluid element towards the polar rotation axis along a line of constant latitude. What happens to the angular momentum of the primitive equation fluid element? Hint: read the discussion of angular momentum for a particle in Section 12.7.
- (c) Give a very brief symmetry argument for why the angular momentum is materially conserved when $\partial p/\partial\lambda = 0$. Hint: recall the discussion of Noether's Theorem in Section 12.1.1.
- (d) Consider the material evolution of 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.

EXERCISE 27.2: 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. Nonetheless, it is sufficient for our purposes. 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. This assumption is sufficient for our purposes.

- (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

$$\int_0^{z_{\text{top}}} (\Phi + \mathcal{I}) \rho dz = \int_0^{z_{\text{top}}} \mathcal{H} \rho dz, \quad (27.112)$$

where

$$\mathcal{H} = p \alpha + \mathcal{I} \quad (27.113)$$

is the enthalpy per mass,

$$\Phi = g z \quad (27.114)$$

is the gravitational potential energy per mass (also known as the *geopotential*) (Section 11.1.2), 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

$$\frac{d\sigma}{dz} = \Pi \frac{d\theta}{dz}, \quad (27.115)$$

where

$$\sigma = \mathcal{H} + \Phi \quad (27.116)$$

is the dry static energy and

$$\Pi = \frac{c_p T}{\theta} \quad (27.117)$$

is the Exner function.

- (c) For an ideal gas atmosphere (either hydrostatic or non-hydrostatic), derive the following expression for the pressure gradient force

$$-\frac{1}{\rho} \nabla p = -\theta \nabla \Pi. \quad (27.118)$$

- (d) For an ideal gas atmosphere (either hydrostatic or non-hydrostatic), derive the following expression for the pressure gradient force

$$-\frac{1}{\rho} \nabla p = -\frac{c_s^2}{\rho \theta} \nabla(\rho \theta). \quad (27.119)$$

- (e) Show that for a hydrostatic atmosphere with an arbitrary equation of state

$$\int_0^{p_s} (\Phi + \mathcal{I}) dp = \int_0^{p_s} \mathcal{H} dp. \quad (27.120)$$

That is, show that the relation in the first part of this problem holds even without making the ideal gas assumption.

EXERCISE 27.3: MASS BALANCE FOR A COLUMN OF HYDROSTATIC FLUID

Return to Exercise 15.3. Show that for a hydrostatic fluid the mass balance for a fluid column (equation (15.109)) takes the form

$$\frac{1}{g} \frac{\partial (p_b - p_a)}{\partial t} = -\nabla \cdot \left[\int_{-H}^{\eta} \mathbf{u} \rho dz \right] + Q_m, \quad (27.121)$$

where

$$p_b = p_a + g \int_{-H}^{\eta} \rho dz \quad (27.122)$$

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 (27.122) into (15.109).

EXERCISE 27.4: 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 11.1). 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 30.2 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

Stress in fluids

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 volume integrating the body force at each point within the region. 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 11.

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. The exchange of dynamical properties, as mediated by mechanical interactions, can be 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 the region. Consequently, the net contact forces acting on a finite region arise just from those contact forces acting at the region boundary. Contact forces communicated through stresses embody the fundamental distinction between the dynamics of a fluid element relative to that of a point particle. We introduced contact forces in Chapters 26 and 27 when deriving the fluid equations of motion and its approximate forms. In this chapter we deepen our understanding of stress by examining certain physical properties and deriving their mathematical expressions.

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 26 and 27. Fluid stresses can be organized into a second order tensor and further decomposed into pressure stresses and viscous stresses. A mathematical and physical understanding of fluid stress apply across the suite of fluid models studied in this book. Throughout this chapter we make use of Cartesian tensors as discussed in Chapters 2 and 3. Nonetheless, the results can be readily generalized for arbitrary coordinates through the rules of general covariance detailed in Chapter 4.

28.1	Cauchy's stress principle and Newton's laws	414
28.1.1	Cauchy's stress principle	414
28.1.2	Local equilibrium of contact stresses	414
28.1.3	Some context for the local equilibrium relation	415
28.2	The stress tensor	416
28.3	Angular momentum and stress tensor symmetry	417

28.4	Flux-form Eulerian momentum equation	418
28.5	Linear momentum for arbitrary regions	419
28.6	Relating stress to pressure and strain	420
28.6.1	Hydrostatic pressure	420
28.6.2	Viscous tensor for a Newtonian fluid	420
28.6.3	Comments and further reading	421
28.7	Exercises	422

28.1 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 28.1). 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.

28.1.1 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 (force per unit area) acting at a point on $\partial\mathcal{R}$. Cauchy's stress principle asserts that the contact stress is a function of the position, time, and boundary normal

$$\boldsymbol{\tau} = \boldsymbol{\tau}(\mathbf{x}, t, \hat{\mathbf{n}}). \quad (28.1)$$

For example, a static fluid in hydrostatic balance has the contact stress vector equal to just the pressure

$$\boldsymbol{\tau} = -p(\mathbf{x}, t) \hat{\mathbf{n}} \quad \text{static fluid in hydrostatic balance.} \quad (28.2)$$

We return to this form of the stress vector in Section 28.6.1.

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, where curvature involves spatial gradients of the normal vector. Nonetheless, we ignore surface tension for most purposes in geophysical fluid mechanics since surface tension is generally negligible for scales larger than a meter.

28.1.2 Local equilibrium of contact stresses

Newton's 2nd law 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

$$\frac{D}{Dt} \int_{\mathcal{R}} \mathbf{v} \rho dV = \int_{\mathcal{R}} \mathbf{f} \rho dV + \int_{\partial\mathcal{R}} \boldsymbol{\tau} dS, \quad (28.3)$$

where $\int_{\mathcal{R}} \mathbf{f} \rho dV$ is the domain integrated body force. 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

$$\lim_{L \rightarrow 0} \frac{1}{L^2} \int_{\partial\mathcal{R}} \tau dS = 0. \quad (28.4)$$

Hence, contact forces at a point in the fluid are locally in equilibrium. As a corollary, contact stresses respecting Cauchy's principle satisfy the relation

$$\tau(\mathbf{x}, t, \hat{\mathbf{n}}) = -\tau(\mathbf{x}, t, -\hat{\mathbf{n}}), \quad (28.5)$$

which is an expression of Newton's third law (action/reaction law) in terms of contact stresses.

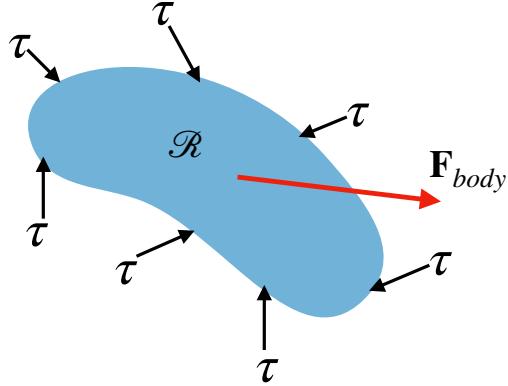


Figure 28.1: Schematic of the net external or body force, \mathbf{F}_{body} , acting on a finite region of fluid, plus the accumulation of internal or contact stresses, τ , acting on the region boundaries. The net body force is determined by a volume integral of the body force at each point within the 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.

28.1.3 Some context for the local equilibrium relation

The local equilibrium relation (28.4), and the corresponding expression of Newton's third law, (28.5), 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. That is, since the contact forces within the domain interior cancel pointwise, the local equilibrium relation (28.4) 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 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 28.2 we illustrate the net contact forces arising just from pressure. Pressure acts solely in a compressive manner as directed along the inward normal to the domain. Note that 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.

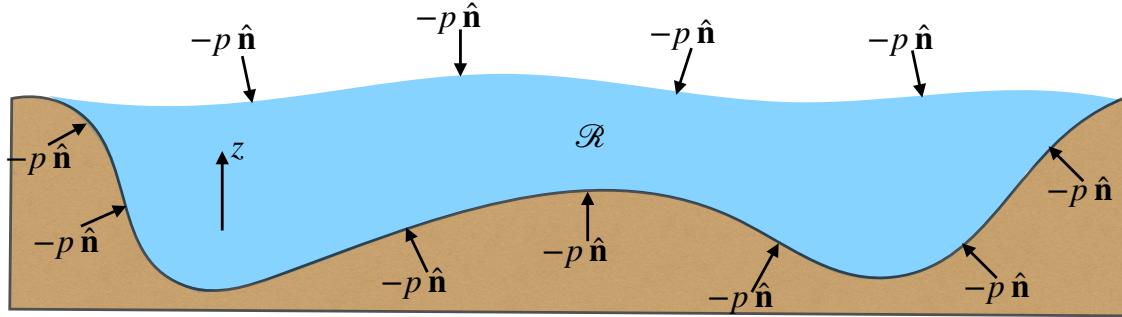


Figure 28.2: Schematic of contact forces from pressure acting on the boundaries to an ocean domain. Pressure forces are directed according to minus the local normal (compressive force). As with all contact forces, the pressure forces acting in the interior of the ocean are locally in equilibrium. Hence, 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.

28.2 The stress tensor

Cauchy's stress principle reduces the complexity for mathematically describing contact stresses. To further pursue that description, consider the tetrahedron fluid region shown in Figure 28.3, 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.

In the limit that the tetrahedron size goes to zero, local equilibrium of the contact forces means that

$$-\sum_{m=1}^3 \boldsymbol{\tau}_{(m)} dA_m + \boldsymbol{\tau}_{\hat{\mathbf{n}}} dA = 0. \quad (28.6)$$

In this equation, $\boldsymbol{\tau}_{(m)} dA_m$ 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

$$dA_m = \hat{n}_m dA, \quad (28.7)$$

so that the local equilibrium relation (28.8) becomes

$$-\sum_{m=1}^3 \tau_{(m)} \hat{n}_m + \tau_{\hat{n}} = 0. \quad (28.8)$$

Now organize the stress components into a 3×3 matrix, T_{mn} , measuring the $n'th$ component of the $m'th$ contact stress $\tau_{(m)}$. Making use of the T_{mn} brings local equilibrium into the form

$$\sum_{m=1}^3 \hat{n}_m T_{mn} = (\tau_{\hat{n}})_n. \quad (28.9)$$

This relation can be written in the suggestive dyadic notation

$$\tau_{\hat{n}} = \hat{n} \cdot \mathbf{T}. \quad (28.10)$$

We thus conclude that the stress vector acting on a surface that is oriented according to a normal vector, \hat{n} , equals to the projection of the stress tensor onto that normal vector. How do we know that 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). Through the quotient rule of tensor analysis, equation (28.9) means that the stresses, T_{mn} , are indeed components to a second order *stress tensor*. As components to a second order tensor, the T_{mn} transform according to the rules developed in Chapter 2 for Cartesian tensors and Chapter 5 for general tensors.

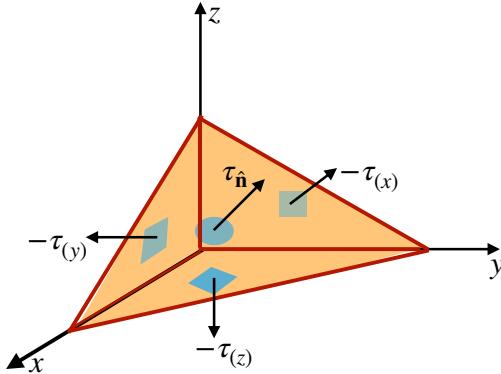


Figure 28.3: Tetrahedron region of fluid with contact stresses acting on the four faces. Note that the stresses are not necessarily directed normal to the faces. Local equilibrium of contact stresses means that the accumulation of these four stresses around the region adds to zero as the region volume goes to zero.

28.3 Angular momentum and stress tensor symmetry

We made use of linear momentum conservation to deduce the local equilibrium property (28.4) 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

$$\mathbf{L} = \rho \delta V (\mathbf{x} \wedge \mathbf{v}), \quad (28.11)$$

and its material time evolution is

$$\frac{D\mathbf{L}}{Dt} = \rho \delta V \mathbf{x} \wedge \frac{D\mathbf{v}}{Dt}, \quad (28.12)$$

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 (26.11)

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} + \nabla \cdot \mathbf{T} \quad (28.13)$$

allows us to write the angular momentum evolution as

$$\frac{D\mathbf{L}}{Dt} = \delta V \mathbf{x} \wedge (\rho \mathbf{f} + \nabla \cdot \mathbf{T}). \quad (28.14)$$

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

$$\left[\frac{DL_m}{Dt} \right]_{\text{contact stress}} = \delta V \epsilon_{mnp} x_n (\nabla \cdot \mathbf{T})_p \quad (28.15a)$$

$$= \delta V \epsilon_{mnp} x_n \partial_q T_{qp} \quad (28.15b)$$

$$= \delta V \epsilon_{mnp} [\partial_q (x_n T_{qp}) - (\partial_q x_n) T_{qp}] \quad (28.15c)$$

$$= \delta V \epsilon_{mnp} [\partial_q (x_n T_{qp}) - T_{np}], \quad (28.15d)$$

where the final equality follows since $\partial_q x_n = \delta_{qn}$. When integrating over a finite region, the first term accounts for torques that may arise from 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 contributes throughout the body of the fluid region. However, if the fluid has no internal torques that lead to a modification of angular momentum, then this second term must vanish. The geophysical fluids of interest do not spontaneously generate angular momentum in the absence of boundary forcing or body forcing. Hence, angular momentum conservation implies the stress tensor is symmetric

$$T_{mn} = T_{nm} \implies \epsilon_{mnp} T_{np} = 0. \quad (28.16)$$

28.4 Flux-form Eulerian momentum equation

We often find it useful to consider Cauchy's form of the momentum equation (28.13) 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 (15.9) so that

$$\rho \frac{D\mathbf{v}}{Dt} = \rho [\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v}] \quad (28.17a)$$

$$= \rho [\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v}] + \mathbf{v} (\partial_t \rho + \nabla \cdot (\rho \mathbf{v})) \quad (28.17b)$$

$$= \partial_t (\rho \mathbf{v}) + \nabla \cdot [(\rho \mathbf{v}) \mathbf{v}]. \quad (28.17c)$$

Consequently, the momentum equation (28.13) takes on the flux-form Eulerian expression

$$\partial_t (\rho \mathbf{v}) + \nabla \cdot [(\rho \mathbf{v}) \mathbf{v}] = \rho \mathbf{f} + \nabla \cdot \mathbf{T}. \quad (28.18)$$

Alternatively, we can move the advection of momentum term onto the right hand side so that

$$\partial_t (\rho \mathbf{v}) = \rho \mathbf{f} + \nabla \cdot [\mathbf{T} - \rho \mathbf{v} \mathbf{v}], \quad (28.19)$$

which takes on the component expression

$$\partial_t(\rho v_m) = \rho f_m + \partial_n[T_{mn} - \rho v_m v_n]. \quad (28.20)$$

We can thus interpret momentum advection as a stress that modifies the linear momentum per volume at a point in space. In the study of turbulent motions the turbulent contribution to $-\rho v_m v_n$ is known as the *Reynolds stress*.

28.5 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 (15.59)

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \varphi dV \right] = \int_{\mathcal{R}} \frac{\partial \varphi}{\partial t} dV + \int_{\partial \mathcal{R}} \varphi \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} dS, \quad (28.21)$$

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 (28.20) leads to

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho v_m dV \right] = \int_{\mathcal{R}} \partial_t(\rho v_m) dV + \int_{\partial \mathcal{R}} (\rho v_m) \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} dS \quad (28.22a)$$

$$= \int_{\mathcal{R}} [\rho f_m + \partial_n(T_{mn} - \rho v_m v_n)] dV + \int_{\partial \mathcal{R}} (\rho v_m) \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} dS \quad (28.22b)$$

$$= \int_{\mathcal{R}} \rho f_m dV + \int_{\partial \mathcal{R}} (T_{mn} - \rho v_m v_n) \hat{\mathbf{n}}_n dS + \int_{\partial \mathcal{R}} (\rho v_m) \mathbf{v}^{(b)} \cdot \hat{\mathbf{n}} dS \quad (28.22c)$$

$$= \int_{\mathcal{R}} \rho f_m dV + \int_{\partial \mathcal{R}} [T_{mn} + \rho v_m (v_n^{(b)} - v_n)] \hat{\mathbf{n}}_n dS. \quad (28.22d)$$

We can write this in a somewhat more suggestive manner through the use of dyadic notation

$$\frac{d}{dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}} \rho \mathbf{f} dV + \int_{\partial \mathcal{R}} [\mathbf{T} + \rho \mathbf{v} (\mathbf{v}^{(b)} - \mathbf{v})] \cdot \hat{\mathbf{n}} dS. \quad (28.23)$$

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.

For the special case of a region moving with the barycentric velocity, we make use of the material time operator, D/Dt , to emphasize the nature of the region's motion. For this case the linear momentum is only affected by body forces as well as contact stresses contained in the stress tensor

$$(\mathbf{v} - \mathbf{v}^{(b)}) \cdot \hat{\mathbf{n}} = 0 \implies \frac{D}{Dt} \left[\int_{\mathcal{R}} \rho \mathbf{v} dV \right] = \int_{\mathcal{R}} \rho \mathbf{f} dV + \int_{\partial \mathcal{R}} \mathbf{T} \cdot \hat{\mathbf{n}} dS. \quad (28.24)$$

This relation is Reynold's transport theorem (Section 15.6.5) as applied to linear momentum.

28.6 Relating stress to pressure and strain

Thus far we have been rather general in our treatment of stress, developing fundamental properties according to the conservation of linear momentum and angular momentum. We now develop *constitutive relations*, which are relations between stress acting between fluid elements and properties of the fluid.

28.6.1 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

$$T_{mn} = -p \delta_{mn} \implies \mathbf{T} \cdot \hat{\mathbf{n}} = -p \hat{\mathbf{n}}, \quad (28.25)$$

where p is the hydrostatic pressure field. For a compressible fluid at rest, we can identify p with the thermodynamic pressure discussed in Chapter 19. 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.

28.6.2 Viscous tensor for a Newtonian fluid

A moving fluid has a more complex stress relation than a static fluid, so that we introduce an extra contribution to the stress tensor

$$T_{mn} = -p \delta_{mn} + \tau_{mn}. \quad (28.26)$$

We refer to τ_{mn} as the *viscous tensor* as it captures the irreversible exchanges of momentum between moving fluid elements arising from the presence of viscosity. The viscous tensor vanishes in a perfect fluid where viscosity is assumed to be zero. It also vanishes when the fluid is static. For the special case that it has a zero trace, $\tau_{mm} = 0$, then it is called the *deviatoric stress tensor*.

To help develop some of the general properties of the viscous 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 14.5) we expect the dynamics to remain unchanged. Since the viscous tensor vanishes when the fluid is static, Galilean invariance implies that the viscous 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 viscous tensor when the fluid locally has uniform motion; i.e., zero gradients in the flow.

Galilean invariance thus leads us to consider a viscous 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 viscous tensor is one that is linearly proportional to the rate of strain tensor introduced in Section 17.2.4. Such fluids are known as *Newtonian fluids*. The linear stress-strain relation takes the same mathematical form as *Hooke's Law* used in the study of elastic solids.

Viscous tensor related to the rate of strain tensor

A *Newtonian fluid* has a viscous tensor that is linearly 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

$$\tau_{mn} = \rho \mu (\partial_m v_n + \partial_n v_m) = 2 \rho \mu S_{mn}, \quad (28.27)$$

with $\mu > 0$ a kinematic viscosity (dimensions of squared length per time) and S_{mn} the components to the rate of strain tensor. Note that with this form of the viscous tensor the trace of the full stress tensor takes the form

$$T_{mm} = -p + 2\rho\mu\nabla \cdot \mathbf{v}. \quad (28.28)$$

For an incompressible fluid (Chapter 16) the trace equals to minus the pressure, in which case τ_{mn} is known as the *deviatoric* stress tensor (deviations from pressure). More generally, the trace deviates from pressure due to the impact from divergence in the velocity.

The viscous force per volume is given by the divergence of the stress tensor

$$\rho F_m = 2\partial_n(\rho\mu S_{mn}). \quad (28.29)$$

Assuming the dynamic viscosity, $\rho\mu$, is constant leads to

$$\rho \mathbf{F} = \rho\mu [\nabla(\nabla \cdot \mathbf{v}) + \nabla^2 \mathbf{v}] \quad \text{if } \rho\mu \text{ is constant.} \quad (28.30)$$

Furthermore, if the flow is incompressible then

$$\rho \mathbf{F} = \rho\mu \nabla^2 \mathbf{v} \quad \text{if } \rho\mu \text{ is constant and } \nabla \cdot \mathbf{v} = 0. \quad (28.31)$$

Viscous stress for a vertically sheared zonal flow

Consider the case of an incompressible velocity that only has a zonal component with a vertical structure (Figure 28.4)

$$\mathbf{v} = u(z)\hat{\mathbf{z}}. \quad (28.32)$$

In this case the only non-zero components to the rate of strain tensor are due to the vertical shear, $S_{13} = S_{31} = \partial_z u/2$. Now consider a horizontal area whose outward normal is parallel to the $\hat{\mathbf{z}}$ direction. The viscous force acting on that area is given by the area integral of the viscous stress

$$\mathbf{F}_{\text{area}} = \int \boldsymbol{\tau} \cdot \hat{\mathbf{n}} dS = \int \boldsymbol{\tau} \cdot \hat{\mathbf{z}} dx dy = \frac{\hat{\mathbf{x}}}{2} \rho\mu A \frac{\partial u}{\partial z}, \quad (28.33)$$

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

$$\left[\frac{\partial(\rho v_m)}{\partial t} \right]_{\text{viscous}} = \partial_n \tau_{nm} \implies \left[\frac{\partial(\rho u)}{\partial t} \right]_{\text{viscous}} = \rho\mu \frac{\partial^2 u}{\partial z^2}, \quad (28.34)$$

so that zonal momentum is preferentially deposited to or removed from regions with high vertical curvature in the zonal velocity.

28.6.3 Comments and further reading

There are far more elaborate stress-strain relations than that from molecular viscosity in equation (28.27). 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](#)).

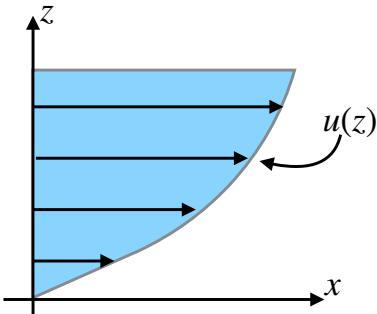


Figure 28.4: Sample profile of zonal velocity possessing a vertical shear: $\mathbf{v} = u(z) \hat{x}$. The resulting zonal viscous stress arises from the nonzero vertical shear.

Further nonlinearities arise in non-Newtonian fluids, which make use of products of the strain tensor for computing stress. 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 as part of their closures, with [Anstey and Zanna \(2017\)](#) offering a compelling approach.

28.7 Exercises

EXERCISE 28.1: STEADY STATE OCEAN FORCE BALANCE

Consider an ocean basin, \mathcal{R} , on a β -plane. The bottom interface separates the ocean from the solid-earth, and the upper interface separates the ocean from a massive atmosphere. 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 bottom stress, τ^{bott} . 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.

29

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.

29.1	The oceanic Boussinesq approximation	423
29.1.1	Density decomposition	423
29.1.2	Momentum equation	424
29.1.3	Mass continuity	424
29.1.4	A non-divergent velocity with density evolution	425
29.1.5	Thermodynamic equation	425
29.1.6	Summary of the oceanic Boussinesq equations	426
29.1.7	Atmospheric analog	426
29.1.8	Further reading	426
29.2	Evolving the Boussinesq fluid forward in time	427
29.3	Hydrostatic Boussinesq ocean model equations	427
29.3.1	Governing equations	427
29.3.2	Material evolution of buoyancy	428
29.3.3	Mechanical forcing and dissipation	428
29.4	Exercises	429

29.1 The oceanic Boussinesq approximation

In this section we derive the oceanic Boussinesq approximation, making use of basics of scaling analysis.

29.1.1 Density decomposition

Decompose density according to

$$\rho(x, y, z, t) = \rho_0 + \delta\rho(x, y, z, t) \quad (29.1)$$

where the deviation density is much smaller than the reference density

$$\delta\rho \ll \rho_0. \quad (29.2)$$

Also write pressure as the sum of a reference pressure, $p_0(z)$, and a deviation $\delta p(x, y, z, t)$,

$$p(x, y, z, t) = p_0(z) + \delta p(x, y, z, t), \quad (29.3)$$

with the reference pressure in hydrostatic balance with the reference density

$$\frac{dp_0}{dz} = -\rho_0 g. \quad (29.4)$$

29.1.2 Momentum equation

With the above density and pressure decompositions, the momentum equation

$$\rho \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] = -\nabla p - \hat{\mathbf{z}} g \rho \quad (29.5)$$

takes the form

$$(\rho_0 + \delta\rho) \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] = -\nabla \delta p - g \delta\rho \hat{\mathbf{z}} - \left[\frac{dp_0}{dz} + \rho_0 g \right] \quad (29.6a)$$

$$= -\nabla \delta p - g \delta\rho \hat{\mathbf{z}}, \quad (29.6b)$$

where we used the hydrostatic balance (29.4) 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 (29.6) by the reference density, and dropping the small term $\delta\rho/\rho_0$ on the left hand, yields the Boussinesq momentum equation

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} = -\nabla \psi + b \hat{\mathbf{z}}. \quad (29.7)$$

We introduced the deviation pressure normalized by the reference density

$$\psi = \frac{\delta p}{\rho_0}, \quad (29.8)$$

as well as the buoyancy

$$b = -\frac{g \delta\rho}{\rho_0} = g \left[1 - \frac{\rho}{\rho_0} \right]. \quad (29.9)$$

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.

29.1.3 Mass continuity

The mass continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}, \quad (29.10)$$

takes the form

$$\frac{D\delta\rho}{Dt} = -(\rho_0 + \delta\rho)(\nabla_z \cdot \mathbf{u} + \partial_z w). \quad (29.11)$$

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

$$\nabla_z \cdot \mathbf{u} = -\partial_z w. \quad (29.12)$$

That is, the velocity field for a Boussinesq fluid is non-divergent

$$\nabla \cdot \mathbf{v} = 0, \quad (29.13)$$

so that the flow is incompressible.

29.1.4 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 29.1.3 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

$$\frac{1}{\delta\rho} \frac{D\delta\rho}{Dt} = -\nabla \cdot \mathbf{v}^d \neq 0. \quad (29.14)$$

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 (29.14).

29.1.5 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 20.2. 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.

29.1.6 Summary of the oceanic Boussinesq equations

The oceanic Boussinesq equations are given by

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} = -\nabla\psi + b\hat{\mathbf{z}} + \mathbf{F} \quad (29.15a)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (29.15b)$$

$$\frac{Db}{Dt} = \dot{b} \quad (29.15c)$$

$$b = -\frac{g\delta\rho}{\rho_0} = -\frac{g(\rho - \rho_0)}{\rho_0} \quad (29.15d)$$

$$\psi = \frac{\delta p}{\rho_0} = \frac{p - p_0(z)}{\rho_0} \quad (29.15e)$$

$$\rho = \rho_0 + \delta\rho(S, \theta, z) \quad (29.15f)$$

$$\frac{dp_0}{dz} = -\rho_0 g. \quad (29.15g)$$

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 20.2 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

$$\rho(S, \theta, p) = \rho(S, \theta, p = -g\rho_0 z) \quad \text{Boussinesq density.} \quad (29.16)$$

29.1.7 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.

29.1.8 Further reading

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.

29.2 Evolving the Boussinesq fluid forward in time

In Section 27.4 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 (29.16))

$$\rho = \rho(S, \theta, p = -g \rho_0 z) \quad \text{Boussinesq density.} \quad (29.17)$$

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 (27.107). 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 (29.2) 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.

29.3 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 29.5, 29.6, 29.7, and 29.8 developing certain properties of this system.

29.3.1 Governing equations

The governing equations for a hydrostatic Boussinesq fluid in a rotating reference frame are given by

$$\frac{Du}{Dt} = f v - \frac{\partial \phi}{\partial x} + \frac{\partial}{\partial z} \left[A \frac{\partial u}{\partial z} \right] \quad (29.18a)$$

$$\frac{Dv}{Dt} = -f u - \frac{\partial \phi}{\partial y} + \frac{\partial}{\partial z} \left[A \frac{\partial v}{\partial z} \right] \quad (29.18b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (29.18c)$$

$$\frac{Db}{Dt} = -\nabla \cdot \mathbf{F} \quad (29.18d)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (29.18e)$$

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (29.18f)$$

Notation corresponds to the oceanic Boussinesq system discussed in Section 29.1. In particular, $b = -g \delta\rho/\rho_o = -g(\rho - \rho_o)/\rho_o$ is the buoyancy relative to the reference state, and $\phi = \delta p/\rho_o = (p - p_o)/\rho_o$ 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.

29.3.2 Material evolution of buoyancy

The material evolution of buoyancy given by equation (29.18d) is affected by the convergence of a subgrid scale (SGS) flux, \mathbf{F} , which we assume takes the form

$$\mathbf{F} = -\kappa \frac{\partial b}{\partial z} \hat{\mathbf{z}} + \mathbf{v}^* b. \quad (29.19)$$

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

$$\kappa = \kappa(x, y, z, t). \quad (29.20)$$

The second term is an advective flux, where the advective velocity, $\mathbf{v}^* = (\mathbf{u}^*, w^*)$, is assumed to be non-divergent

$$\nabla \cdot \mathbf{v}^* = \nabla_z \cdot \mathbf{u}^* + \frac{\partial w^*}{\partial z} = 0. \quad (29.21)$$

The velocity \mathbf{v}^* is commonly termed the *eddy-induced* velocity, with a particular choice for its parameterization discussed in Exercise 29.8.

The boundary conditions for the buoyancy are given by

$$\kappa \frac{\partial b}{\partial z} = Q_b \quad \text{at } z = \eta \quad (29.22a)$$

$$\kappa \frac{\partial b}{\partial z} = 0 \quad \text{at } z = -H \quad (29.22b)$$

$$\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0 \quad \text{all boundaries,} \quad (29.22c)$$

where Q_b is the surface buoyancy flux, and we assume no buoyancy flux at the ocean bottom (i.e., no geothermal heating).

29.3.3 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

$$A = A(x, y, z, t). \quad (29.23)$$

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

$$\rho_o A \frac{\partial u}{\partial z} = \tau^x \quad \text{at } z = \eta \quad (29.24a)$$

$$\rho_o A \frac{\partial v}{\partial z} = \tau^y \quad \text{at } z = \eta \quad (29.24b)$$

$$A \frac{\partial u}{\partial z} = C_D u |\mathbf{u}| \quad \text{at } z = -H(x, y) \quad (29.24c)$$

$$A \frac{\partial v}{\partial z} = C_D v |\mathbf{u}| \quad \text{at } z = -H(x, y). \quad (29.24d)$$

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.

29.4 Exercises

EXERCISE 29.1: 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, m , changes via boundary mass fluxes

$$\frac{dm}{dt} = \mathcal{A} \overline{Q_m}, \quad (29.25)$$

where \mathcal{A} is the ocean surface area and $\overline{Q_m}$ is the area averaged surface mass flux. Global volume of liquid seawater

$$V = \frac{m}{\langle \rho \rangle} \quad (29.26)$$

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, 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

$$\left[\frac{d\bar{\eta}}{dt} \right]_{\text{observed}} \approx 3 \text{ mm yr}^{-1}. \quad (29.27)$$

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

$$\mathcal{T}_{\text{river}} \approx 1.2 \times 10^9 \text{ kg s}^{-1}. \quad (29.28)$$

- (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

$$\mathcal{E}_{\text{bomb}} \approx 6.3 \times 10^{13} \text{ J}. \quad (29.29)$$

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, (29.25). 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)

$$\beta_\theta = -\frac{1}{\rho} \left[\frac{\partial \rho(S, \theta, p)}{\partial \theta} \right]_{S,p} = 2 \times 10^{-4} \text{ K}^{-1}. \quad (29.30)$$

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 29.2: POISSON EQUATION FOR PRESSURE

Consider a perfect Boussinesq fluid that is rotating on a β -plane and thus satisfies the governing equations

$$\frac{D\mathbf{v}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{v} = -\nabla\psi + \hat{\mathbf{z}} b \quad (29.31a)$$

$$\frac{Db}{Dt} = 0 \quad (29.31b)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (29.31c)$$

$$f = f_o + \beta y. \quad (29.31d)$$

Show that the dynamic pressure ψ satisfies the following elliptic boundary value problem (see Section 1.4.3 for discussion of elliptic partial differential equations)

$$-\nabla^2\psi = \nabla \cdot \mathbf{G} \quad \text{within the domain} \quad (29.32a)$$

$$-\hat{\mathbf{n}} \cdot \nabla\psi = \hat{\mathbf{n}} \cdot \mathbf{G} \quad \text{on the domain boundaries.} \quad (29.32b)$$

Determine the vector \mathbf{G} .

EXERCISE 29.3: 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

$$\frac{D\mathbf{v}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{v} = -\nabla\psi + \hat{\mathbf{z}} b \quad (29.33a)$$

$$\frac{Db}{Dt} = 0 \quad (29.33b)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (29.33c)$$

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 29.4: 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

$$\frac{D\mathbf{v}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{v} = -\nabla\psi + \hat{\mathbf{z}} b + \nu \nabla^2 \mathbf{v} \quad (29.34a)$$

$$\frac{Db}{Dt} = Q + \kappa \nabla^2 b \quad (29.34b)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (29.34c)$$

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

$$\nabla^2 \mathbf{v} = -\nabla \wedge \boldsymbol{\omega} \quad \text{if } \nabla \cdot \mathbf{v} = 0 \quad (29.35)$$

to express the viscosity contributions in terms of the vorticity

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v}. \quad (29.36)$$

- (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 29.5: 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 29.3.

- (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 29.6: 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 29.3.

- (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 29.7: 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 29.3.

- (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 29.8: 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 29.3. Namely, we consider the specific form for the parameterized eddy-induced velocity proposed by [Gent et al. \(1995\)](#)

$$\mathbf{u}^* = -\partial_z(B \mathbf{S}) \quad (29.37a)$$

$$w^* = \nabla_z \cdot (B \mathbf{S}) \quad (29.37b)$$

$$\mathbf{S} = -\frac{\nabla_z b}{N^2} \quad (29.37c)$$

$$\mathbf{v}^* \cdot \hat{\mathbf{n}} = 0 \quad \text{at all ocean boundaries.} \quad (29.37d)$$

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

$$\mathbf{v}^* = \nabla \wedge \Psi^*. \quad (29.38)$$

- (b) Show that

$$\int_{-H}^{\eta} \mathbf{u}^* dz = 0. \quad (29.39)$$

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

$$\mathcal{B}^{(y)}(y, t) = \int_{x1}^{x2} dx \int_{-H}^{\eta} b(v + v^*) dz. \quad (29.40)$$

The zonal and vertical integrals are over the full zonal and vertical extent of the ocean domain. Show that the effects from v^* are to reduce the meridional gradients of buoyancy. That is, if buoyancy decreases poleward, then v^* will flux buoyancy poleward to reduce the gradient.

- (d) How does the introduction of \mathbf{v}^* to the buoyancy equation (29.18d) affect the global integrated gravitational potential energy? Discuss.
- (e) How does the introduction of \mathbf{v}^* to the buoyancy equation (29.18d) affect the global integrated available potential energy? Discuss.

30

Rapidly rotating fluids

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. When friction is included, such as within boundary layers, rotationally dominant flows develop Ekman layers, which provide a means to transfer momentum and matter from turbulent boundary layers into the quasi-adiabatic geostrophic interior.

In this chapter, we introduce salient features of geostrophically balanced flow. The associated diagnostic relations involve no time derivatives, and so cannot be used to predict the evolution of the flow state. However, their power for diagnostics is unquestioned as they provide a robust framework for interpreting the large-scale ocean and atmospheric circulation. Throughout our discussion in this chapter we are not explicitly concerned with sphericity, thus enabling the use of Cartesian coordinates.

30.1	Primitive equations	436
30.2	The Rossby number	436
30.2.1	Scaling for the Rossby number	437
30.2.2	Ratio of material acceleration to Coriolis acceleration	437
30.2.3	Ratio of local time tendency to Coriolis acceleration	437
30.2.4	Rossby number for a kitchen sink	438
30.2.5	Rossby number for a Gulf Stream ring	438
30.2.6	Further reading	438
30.2.7	Exercise	438
30.3	Geostrophic balance	438
30.3.1	Geostrophic balance is distinctly fluid mechanical	439
30.3.2	Geostrophic relation in geopotential coordinates	439
30.3.3	Cyclonic and anti-cyclonic orientation	439
30.3.4	Geostrophic relation in pressure coordinates	440
30.4	Planetary geostrophy and thermal wind	440
30.4.1	Stating the planetary geostrophic equations	440
30.4.2	Planetary geostrophic vorticity equation	441
30.4.3	Taylor-Proudman theorem and vertical stiffening	442
30.4.4	Vorticity balance	443
30.4.5	Thermal wind balance for the ocean	444
30.4.6	Thermal wind balance for the atmosphere	445
30.4.7	Further reading	446
30.4.8	Exercises	446

30.5	Ekman layers	447
30.5.1	Non-dimensionalization	447
30.5.2	Ekman number and Ekman layer thickness	448
30.5.3	Integral properties of the Ekman layer	449
30.5.4	Further reading	450
30.6	Natural coordinates	450
30.6.1	Natural coordinates	451
30.6.2	Material acceleration	451
30.6.3	Coriolis acceleration and pressure gradient acceleration	452
30.6.4	Horizontal momentum equation	452
30.6.5	Further reading	452

30.1 Primitive equations

Throughout this chapter, we make use of the hydrostatic primitive equations derived in Section ??

$$\rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla + f \hat{\mathbf{z}} \wedge \right] \mathbf{u} = -\nabla_z p \quad (30.1a)$$

$$\frac{\partial p}{\partial z} = -g \rho \quad (30.1b)$$

$$\frac{D\rho}{Dt} = \rho \nabla \cdot \mathbf{v}, \quad (30.1c)$$

where the velocity vector is written

$$\mathbf{v} = \mathbf{u} + \hat{\mathbf{z}} w = \hat{\mathbf{x}} u + \hat{\mathbf{y}} v + \hat{\mathbf{z}} w, \quad (30.2)$$

and the horizontal gradient operator is

$$\nabla_z = \hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y. \quad (30.3)$$

For some of the scaling analysis we assume an incompressible fluid as per the Boussinesq approximation (Section 29.1). In this case, the mass continuity equation (30.1c) becomes the non-divergent condition

$$\nabla \cdot \mathbf{v} = 0. \quad (30.4)$$

Furthermore, ρ in the horizontal momentum equation (30.1a) is converted to a constant reference density ρ_0 .

30.2 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. Indeed, one may wish to consider the alternative, whereby the atmosphere is stationary relative to the stars, in which case atmospheric motion relative to the ground would involve huge winds (see Exercise 30.1 for more on this point). Second, human scale horizontal length scales are generally far too small to take notice of the rotation of the planet. This point is quantified by considering the Rossby number, which includes a horizontal length scale, a velocity scale, and angular rotation speed.

30.2.1 Scaling for the Rossby number

The *Rossby number* measures the ratio of the Coriolis acceleration to the material acceleration. The material acceleration has two contributions; one from local time tendencies and one from advection. We expose typical scales for the acceleration by writing

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{u} \sim \frac{U}{T} + \frac{U^2}{L} + \frac{WU}{H}, \quad (30.5)$$

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 27.2). Likewise, the Coriolis acceleration scales as

$$f \hat{\mathbf{z}} \wedge \mathbf{u} \sim f_0 U, \quad (30.6)$$

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

$$\frac{W}{H} \sim \frac{U}{L} \Rightarrow W \sim U \frac{H}{L}. \quad (30.7)$$

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

$$\frac{H}{L} \ll 1, \quad (30.8)$$

as per the hydrostatic balance discussed in Section 27.2. Consequently, the vertical velocity scale is much less than the horizontal

$$W \ll U. \quad (30.9)$$

30.2.2 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

$$Ro = \frac{U^2/L}{f_0 U} = \frac{U}{f_0 L}. \quad (30.10)$$

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.

30.2.3 Ratio of local time tendency to Coriolis acceleration

A complementary way to understand the Rossby number is as a ratio of the local time tendency for the horizontal velocity, to the Coriolis acceleration

$$Ro = \frac{U/T}{U f_0} = \frac{1/T}{f_0} \quad (30.11)$$

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.

30.2.4 Rossby number for a kitchen sink

Consider flow in a kitchen sink. 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

$$Ro \approx 10^2 - 10^3. \quad (30.12)$$

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.

30.2.5 Rossby number for a Gulf Stream ring

For a Gulf Stream ring, 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

$$Ro \approx 10^{-2} - 10^{-1}, \quad (30.13)$$

thus indicating the importance of the Coriolis acceleration for dynamics of Gulf Stream rings.

30.2.6 Further reading

Section 6.6.6 of [Marshall and Plumb \(2008\)](#) discuss rotating physics in the context of laboratory tank experiments.

30.2.7 Exercise

EXERCISE 30.1: SMALL ROSSBY NUMBER AT HUMAN SCALES

Consider motion of a car at a speed $U \sim 10^5 \text{ m hour}^{-1}$ and a length scale of $L \sim 10 \text{ m}$.

- (a) What is the rotation period required to 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 13.2.4.
- (d) Discuss one or two astronomical objects that have very large rotational speeds.

30.3 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

$$\mathbf{u} \cdot (\hat{\mathbf{z}} \wedge \mathbf{u}) = 0. \quad (30.14)$$

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 30.1). In the southern hemisphere, where $f < 0$, it acts in the opposite direction.

30.3.1 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 10 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.

30.3.2 Geostrophic relation in geopotential coordinates

Mathematically, the geostrophic balance becomes important when the Rossby number is small. In this case, the leading order dynamical balance in the momentum equation (30.1a) is between the Coriolis acceleration and pressure gradient acceleration

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = -\rho^{-1} \nabla_z p, \quad (30.15)$$

or equivalently, we have the expression for the geostrophic velocity¹

$$\mathbf{u}_g = \frac{\hat{\mathbf{z}} \wedge \nabla p}{\rho f}. \quad (30.16)$$

Note that the equator is special since the Coriolis parameter, $f = 2\Omega \sin \phi$, vanishes, thus precluding the utility of geostrophy near the equator.

30.3.3 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 30.1). 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 south, geostrophic motion around a low pressure center is clockwise, which is a negatively oriented

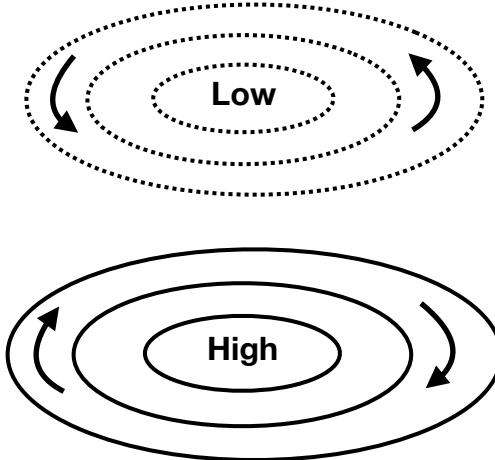


Figure 30.1: Geostrophic motion around low and high pressure centers in the northern hemisphere ($f = 2\Omega \sin \phi > 0$ in the north). The counter-clockwise motion around the low pressure center is in the same sense as the planetary rotation, and is thus termed cyclonic. Cyclonic motion in the Southern 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, and is thus termed anti-cyclonic.

¹We can write either ∇ or ∇_z in equation (30.16). The reason is that the $\hat{\mathbf{z}} \wedge$ operator selects only the horizontal portion of the gradient.

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 30.1).

30.3.4 Geostrophic relation in pressure coordinates

The hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho g \quad (30.17)$$

can be used to eliminate density from the geostrophic balance (30.15) to render

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = \frac{g \nabla_z p}{\partial p / \partial z}. \quad (30.18)$$

The right hand side is minus the gradient of the geopotential

$$\Phi = g z \quad (30.19)$$

along surfaces of constant pressure (see Appendix 8.12.2)

$$f \hat{\mathbf{z}} \wedge \mathbf{u}_g = -\nabla_p \Phi \Rightarrow f \mathbf{u}_g = \hat{\mathbf{z}} \wedge \nabla_p \Phi. \quad (30.20)$$

This is a useful expression of geostrophy for the compressible atmosphere.

30.4 Planetary geostrophy and thermal wind

We here introduce the planetary geostrophic (PG) balance, which is quite useful for theories of the large-scale ocean circulation. We only state the equations here, providing a more systematic justification for the shallow water system in Section 42.3 and a continuously stratified fluid in Section 43.2. Certain implications follow from considering the vorticity budget for this fluid.

30.4.1 Stating the planetary geostrophic equations

The governing equations for PG are based on the Boussinesq equations stated in Section 29.1.6, with the assumption of a steady state geostrophic balance for the horizontal momentum

$$\rho_o f (\hat{\mathbf{z}} \wedge \mathbf{u}) = -\nabla p - \rho g \hat{\mathbf{z}} + \frac{\partial \boldsymbol{\tau}}{\partial z} \quad (30.21a)$$

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (30.21b)$$

$$\frac{Db}{Dt} = \dot{b}. \quad (30.21c)$$

Note that the stress $\boldsymbol{\tau}$ is assumed to act just in the horizontal directions. Furthermore, the vertical component of the momentum equation (30.21a) is the hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho g. \quad (30.22)$$

Velocity is slaved to buoyancy

The only time derivative appearing in the PG equations appears in the buoyancy equation. All other equations are diagnostic. As the buoyancy evolves, the 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.

How pressure affects velocity shears

The horizontal momentum is directly affected by horizontal pressure gradient forces. Furthermore, the hydrostatic balance says that the vertical shear of the horizontal pressure gradient is determined by horizontal density gradients

$$\frac{\partial(\nabla_z p)}{\partial z} = -g \nabla_z \rho. \quad (30.23)$$

Consequently, in the presence of horizontal density gradients, the horizontal pressure gradient forces have a depth dependence. That depth dependence in turn imparts a vertical shear to the horizontal velocity. This connection between horizontal density gradients and vertical shears in the horizontal velocity forms the foundation for the thermal wind relation detailed in Section 30.4.5.

Vertical transfer of horizontal stress and subgrid scale parameterizations

We introduced a horizontal stress (dimensions of force per area) into the momentum equation

$$\boldsymbol{\tau} = (\tau^x, \tau^y, 0). \quad (30.24)$$

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. The development of SGS parameterizations remains an active area of theoretical physical oceanography.

30.4.2 Planetary geostrophic vorticity equation

The vertical component of relative vorticity is given by

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \quad (30.25)$$

As to be discussed in Chapter 37, we form the relative vorticity budget by taking the curl of the momentum equation.

Curl of the PG momentum equation

Taking the curl of the momentum equation (30.21a), and rearranging terms, leads to the planetary geostrophic vorticity equation

$$-\rho_o f \frac{\partial \mathbf{u}}{\partial z} + \hat{z} \rho_o \nabla \cdot (f \mathbf{u}) = -g \nabla \wedge (\hat{z} \rho) + \frac{\partial (\nabla \wedge \boldsymbol{\tau})}{\partial z}. \quad (30.26)$$

Note that $\nabla \cdot (f \mathbf{u}) = \nabla_z \cdot (f \mathbf{u})$ since \mathbf{u} is the horizontal velocity vector. Introducing buoyancy (Section 29.1.2)

$$b = -g \left[\frac{\rho - \rho_o}{\rho_o} \right] \quad (30.27)$$

leads to

$$-f \frac{\partial \mathbf{u}}{\partial z} + \hat{z} \nabla_z \cdot (f \mathbf{u}) = \nabla \wedge (\hat{z} b) + \frac{1}{\rho_o} \frac{\partial (\nabla \wedge \boldsymbol{\tau})}{\partial z}. \quad (30.28)$$

Pressure removed from vorticity equation

One of the key reasons to study vorticity is that its evolution is not explicitly affected by pressure gradients, since the curl of the pressure gradient vanishes. However, pressure gradients remain since they directly impact the velocity field through the momentum equation, and the vorticity is the curl of velocity.

Absence of relative vorticity in the PG vorticity equation

It is notable that there is no explicit appearance of the relative vorticity, ζ , in the planetary geostrophic vorticity equation (30.21a). The reason is that we dropped the material time derivative when forming the planetary geostrophic momentum equation (30.21a). By doing so, we drop all expressions of ζ in the vorticity equation. Planetary geostrophy is valid for those cases where

$$|\zeta| \ll |f|, \quad (30.29)$$

which means vorticity is dominated by the planetary vorticity. We encounter more complete versions of the vorticity equation in Chapter 37, 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 (30.28) 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 37).

30.4.3 Taylor-Proudman theorem and vertical stiffening

Consider the vorticity equation (30.28) on an f -plane with zero friction, in which the horizontal geostrophic motion is horizontally non-divergent

$$\nabla_z \cdot \mathbf{u} = 0 \quad f\text{-plane geostrophy.} \quad (30.30)$$

Use of continuity (equation (30.21b)) means there is no vertical stretching of a vertical material line element (Section 17.2.5)

$$\frac{\partial w}{\partial z} = 0. \quad (30.31)$$

As shown in Chapter 37, a vortex tube exhibits the same kinematics as a material line element (see Section 17.2). Hence, $\partial w / \partial z = 0$ means there is no vertical stretching of a vortex tube.

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 (30.28) implies that the horizontal velocity has zero vertical shear

$$\frac{\partial \mathbf{u}}{\partial z} = 0 \quad f\text{-plane and homogeneous density.} \quad (30.32)$$

This result is known as the Taylor-Proudman theorem. Laboratory experiments verify its utility by inserting a dye into a rapidly rotating tank of unstratified fluid. After a few rotation periods, the dye will form vertical sheets, or “Taylor curtains” whose center is along the axis of rotation. The fluid is said to have a “vertical stiffness” due to the effects of rotation.

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. Smaller vertical scales of variation are captured by an infinite hierarchy of *baroclinic* modes.

30.4.4 Vorticity balance

The vertical component to the vorticity balance (30.28) leads to

$$\nabla_z \cdot (f \mathbf{u}) = \frac{1}{\rho_o} \frac{\hat{\mathbf{z}} \cdot \partial (\nabla \wedge \boldsymbol{\tau})}{\partial z} \quad (30.33)$$

which can be written

$$\beta v = -f \nabla_z \cdot \mathbf{u} + \frac{1}{\rho_o} \frac{\hat{\mathbf{z}} \cdot \partial (\nabla \wedge \boldsymbol{\tau})}{\partial z} \quad (30.34)$$

where

$$\beta = \frac{\partial f}{\partial y} \quad (30.35)$$

is the gradient of planetary vorticity. The continuity equation (30.21b) can be used to remove the horizontal divergence, which yields the vorticity balance

$$\beta v = f \frac{\partial w}{\partial z} + \frac{1}{\rho_o} \frac{\hat{\mathbf{z}} \cdot \partial (\nabla \wedge \boldsymbol{\tau})}{\partial z}. \quad (30.36)$$

The left hand side represents the meridional advection of planetary vorticity. The first term on the right hand side represents the vortex stretching by planetary vorticity; i.e., planetary induction or the β -effect discussed in Section 37.9.3. 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 (30.36) in Section 43.2.1 where we consider its depth integrated form.

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 27.2.2 for a discussion in the context of the hydrostatic balance.

30.4.5 Thermal wind balance for the ocean

Horizontal components to the inviscid vorticity equation (equation (30.28) with τ set to zero) form the *thermal wind balance*

$$f \frac{\partial u}{\partial z} = -\nabla \wedge (\hat{z} b), \quad (30.37)$$

which takes on the component form

$$f \frac{\partial u}{\partial z} = -\frac{\partial b}{\partial y} \quad f \frac{\partial v}{\partial z} = \frac{\partial b}{\partial x}. \quad (30.38)$$

Note that these relations can also be derived directly from taking the vertical derivative of the horizontal momentum equation (30.21a) and then using the horizontal gradient of the hydrostatic balance (30.22). In either case, the thermal wind balance (30.37) 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.

Thermal wind and the atmospheric jet stream and 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 30.2 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

$$f \frac{\partial \bar{u}}{\partial z} = \frac{g}{\rho_0} \frac{\partial \bar{\rho}}{\partial y} = -\frac{\partial \bar{b}}{\partial y} > 0, \quad (30.39)$$

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 (30.37) between two constant depth surfaces leads to

$$\mathbf{u}(z) = \mathbf{u}(z_{\text{ref}}) - f^{-1} \nabla \wedge \hat{z} \int_{z_{\text{ref}}}^z b dz. \quad (30.40)$$

Hence, knowledge of the buoyancy field (e.g., through hydrographic measurements of temperature and salinity in

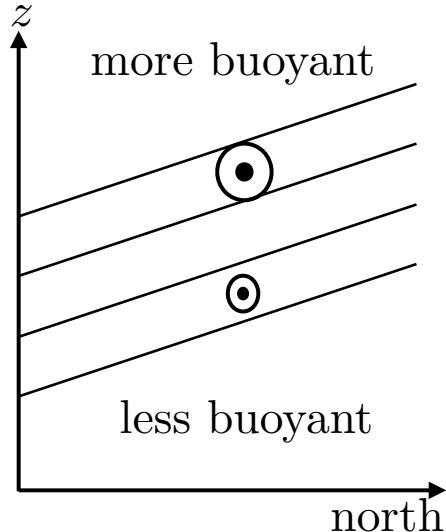


Figure 30.2: Schematic of thermal wind balance in the northern hemisphere with north to the right and east out of the page. Shown here are surfaces of constant density or buoyancy. Density increases with depth and when moving poleward ($\partial \bar{b} / \partial y < 0$). The zonal average removes zonal variations, thus leading to $\partial_x \bar{b} = 0$. The thermal wind velocity associated with this density field is eastward (towards the reader), and it increases when moving upward. The same eastward thermal wind velocity results in the southern hemisphere where $\partial_y \bar{b} > 0$ and $f < 0$.

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.

30.4.6 Thermal wind balance for the atmosphere

The large-scale atmosphere is compressible and predominantly in hydrostatic balance. The expression for geostrophic balance (30.20) 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 (30.20) to render

$$f \frac{\partial \mathbf{u}}{\partial p} = \hat{\mathbf{z}} \wedge \nabla_p \left[\frac{\partial \Phi}{\partial p} \right]. \quad (30.41)$$

The hydrostatic relation $\partial p/\partial z = -\rho g$ takes the form

$$\frac{\partial p}{\partial \Phi} = -\rho \Rightarrow \frac{\partial \Phi}{\partial p} = -\alpha \quad (30.42)$$

in which case

$$f \frac{\partial \mathbf{u}}{\partial p} = -\hat{\mathbf{z}} \wedge \nabla_p \alpha. \quad (30.43)$$

The specific volume takes the following form for an ideal gas atmosphere (see Section 19.5.1)

$$\alpha = \rho^{-1} = \frac{R^M T}{p}. \quad (30.44)$$

Since the horizontal derivative in the thermal wind relation (30.43) is along pressure surfaces, we have

$$f \frac{\partial \mathbf{u}}{\partial p} = -\frac{R^M}{p} [\hat{\mathbf{z}} \wedge \nabla_p T]. \quad (30.45)$$

This expression gives rise to the name “thermal wind”.

As for the ocean in equation (30.40), we vertically integrate the thermal wind expression (30.45), only now do so between two pressure levels

$$\mathbf{u}(p_a) - \mathbf{u}_g(p_b) = f^{-1} R^M \hat{\mathbf{z}} \wedge \nabla_p \left[\int_{p_a}^{p_b} \frac{T dp}{p} \right], \quad (30.46)$$

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 lower altitude (greater pressure)

$$\mathbf{u}_T = \mathbf{u}(p_a) - \mathbf{u}(p_b) \quad \text{with } p_a < p_b \quad (30.47)$$

so that

$$\mathbf{u}_T = \frac{R^M}{f} \hat{\mathbf{z}} \wedge \nabla_p \bar{T}^{\ln p}, \quad (30.48)$$

where we introduced the log-pressure weighted temperature between the two pressure surfaces

$$\bar{T}^{\ln p} = \int_{p_a}^{p_b} \frac{T dp}{p}. \quad (30.49)$$

The relation (30.48) 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.

30.4.7 Further reading

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.

30.4.8 Exercises

EXERCISE 30.2: THE BETA SPIRAL

Consider a steady state Boussinesq planetary geostrophic fluid in the absence of mixing. Write the geostrophic velocity as

$$u = |\mathbf{u}| \cos \Delta \quad v = |\mathbf{u}| \sin \Delta, \quad (30.50)$$

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 30.3: ALTERNATIVE FORM OF THERMAL WIND

Consider a fluid with density a function of pressure and potential temperature

$$\rho = \rho(p, \theta). \quad (30.51)$$

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

$$\frac{\partial \mathbf{u}}{\partial z} = \left[\frac{N^2}{f \rho g} \right] (\hat{\mathbf{z}} \wedge \nabla_\theta p), \quad (30.52)$$

where

$$N^2 = -\frac{g}{\rho} \frac{\partial \rho}{\partial \theta} \frac{\partial \theta}{\partial z} = g \beta_\theta \frac{\partial \theta}{\partial z} > 0 \quad (30.53)$$

is the squared buoyancy frequency, assumed positive so that the fluid is gravitationally stable in the vertical (see Section 20.3.3). The term β_θ is the thermal expansion coefficient written in terms of potential temperature (Section 20.2.5). Finally, the horizontal gradient projected onto constant θ surfaces is given by (see Section 8.12.2)

$$\nabla_\theta = \hat{\mathbf{x}} \left[\frac{\partial}{\partial x} \right]_{y,\theta} + \hat{\mathbf{y}} \left[\frac{\partial}{\partial y} \right]_{x,\theta} \quad (30.54a)$$

$$= \nabla_z - \left[\frac{\nabla_z \theta}{\partial \theta / \partial z} \right] \frac{\partial}{\partial z}. \quad (30.54b)$$

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 $\rho = \rho(\theta, p)$.

30.5 Ekman layers

We study a Boussinesq hydrostatic fluid within a turbulent boundary layer, where the balance of terms in the momentum equation is between friction, pressure, and Coriolis

$$\mathbf{f} \wedge \mathbf{u} = -\nabla_z \psi + \mathbf{F}. \quad (30.55)$$

We find it useful to separate velocity into a geostrophic contribution

$$\mathbf{f} \wedge \mathbf{u}_g = -\nabla_z \psi \quad (30.56)$$

and an ageostrophic or Ekman contribution

$$\mathbf{f} \wedge \mathbf{u}_E = \mathbf{F}. \quad (30.57)$$

That is, the Ekman velocity is a balance between Coriolis acceleration and frictional acceleration.

Details for the frictional acceleration are generally not well defined, since they depend on details of the turbulence. [Vallis \(2017\)](#) assumes the form of an eddy viscous stress (equation (2.274) of [Vallis \(2006\)](#))

$$\mathbf{F}^{\text{viscous}} = \frac{1}{\rho_o} \frac{\partial \boldsymbol{\tau}}{\partial z} = \frac{\partial}{\partial z} \left[A \frac{\partial \mathbf{u}}{\partial z} \right], \quad (30.58)$$

with A a vertical turbulent kinematic viscosity of dimensions L^2/T . In contrast, [Marshall and Plumbe \(2008\)](#) assume the form of a Rayleigh drag (equation (7.28) of [Marshall and Plumbe \(2008\)](#))

$$\mathbf{F}^{\text{drag}} = -\frac{k \mathbf{u}}{\delta}, \quad (30.59)$$

where δ is a depth scale and k is a drag coefficient with dimensions L/T .

30.5.1 Non-dimensionalization

Non-dimensionalize the hydrostatic Boussinesq equations by setting

$$(x, y) = L(\hat{x}, \hat{y}) \quad z = L_z \hat{z} \quad (u, v) = U(\hat{u}, \hat{v}) \quad f = f_0 \hat{f} \quad \psi = \Psi \hat{\psi} \quad (30.60)$$

where the hatted variables are non-dimensional, and we introduced typical scales for length, velocity, Coriolis, and pressure. We will make the vertical scale, L_z , precise in the following, by defining it to be the thickness over which friction is nontrivial. The pressure scale, Ψ , is assumed to follow geostrophic scaling

$$\Psi = f_0 U L. \quad (30.61)$$

Inserting the relations (30.60) into equation (30.55) leads to the non-dimensional frictional geostrophic equation

$$\hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla} \hat{\psi} + \frac{\mathbf{F}}{f_0 U}. \quad (30.62)$$

30.5.2 Ekman number and Ekman layer thickness

The ratio of scales for the frictional acceleration to the Coriolis acceleration defines the Ekman number

$$Ek = \frac{\text{frictional acceleration}}{\text{Coriolis acceleration}}. \quad (30.63)$$

For the viscous stress form of frictional acceleration

$$\mathbf{F}^{\text{viscous}} = \frac{AU}{L_z^2} \left[\frac{\partial^2 \hat{\mathbf{u}}}{\partial z^2} \right], \quad (30.64)$$

the Ekman number is

$$Ek = \frac{A}{f_0 L_z^2}. \quad (30.65)$$

If we take the vertical scale, L_z , equal to the depth scale over which interior flow processes occur, $L_z = H$, then the Ekman number will generally be very small, in which case the Laplacian friction term is negligible. Note, however, that since the Ekman number multiplies the highest derivative in equation (30.62), setting it to zero represents a singular limit. We therefore expect there to be a boundary layer inside of which the Ekman number is order unity, and the Laplacian viscous friction is nontrivial.

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.³ The viscous friction acceleration, (30.64), 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

$$Ek = 1 \Rightarrow \delta^{\text{viscous}} = \left[\frac{A}{f_0} \right]^{1/2}. \quad (30.66)$$

The eddy viscosity is not easy to deduce 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

$$A = f_0 (\delta^{\text{viscous}})^2. \quad (30.67)$$

In the atmosphere, the boundary layer thickness is order 1000 m, so that at mid-latitudes, with $f_0 = 10^{-4} \text{ s}^{-1}$, we expect

$$A^{\text{atmos}} \sim 10^2 \text{ m}^2 \text{ s}^{-1}. \quad (30.68)$$

In the ocean, the upper ocean boundary layer depth is roughly 50 m, in which case

$$A^{\text{ocean}} \sim 0.25 \text{ m}^2 \text{ s}^{-1}. \quad (30.69)$$

For the drag form of friction

$$\mathbf{F}^{\text{drag}} = - \left[\frac{k U}{\delta} \right] \hat{\mathbf{u}}, \quad (30.70)$$

the Ekman number takes the form

$$Ek = \frac{k}{\delta f_0}. \quad (30.71)$$

³When 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.

The Ekman number is unity, and hence friction important, if the thickness δ takes on the Ekman value

$$\delta^{\text{drag}} = \frac{k}{f_0}. \quad (30.72)$$

The two expressions for the Ekman layer thicknesses are the same if the drag coefficient and viscosity are related by

$$\delta^{\text{drag}} = \delta^{\text{viscous}} \iff k = \sqrt{A f_0}. \quad (30.73)$$

30.5.3 Integral properties of the Ekman layer

It is possible to establish some integral properties of the Ekman layer, even without knowing the precise details of the eddy viscous closure. We focus here on the upper ocean Ekman layer between the Ekman layer bottom and the ocean free surface

$$-\delta(x, y, t) \leq z \leq \eta(x, y, t). \quad (30.74)$$

30.5.3.1 Horizontal Ekman mass transport

Integrating the horizontal Ekman balance (30.57) over the depth of the Ekman layer leads to

$$\mathbf{f} \wedge \mathbf{M}_E = \int_{-\delta}^{\eta} \rho_o \mathbf{F} dz, \quad (30.75)$$

where

$$\mathbf{M}_E = \int_{-\delta}^{\eta} \rho_o \mathbf{u}_E dz \quad (30.76)$$

is the depth integrated ageostrophic mass transport within the Ekman boundary layer. Assume friction in the form of a viscous stress (equation (30.58)), so that

$$\mathbf{M}_E = -f^{-1} \hat{\mathbf{z}} \wedge [\boldsymbol{\tau}(\eta) - \boldsymbol{\tau}(-\delta)]. \quad (30.77)$$

Stress at the bottom of the Ekman layer, $\boldsymbol{\tau}(-\delta)$, matches to that in the ocean interior, which is much smaller than stress at the ocean surface, $\boldsymbol{\tau}(\eta)$. Hence, the ageostrophic Ekman layer mass transport takes the form

$$\mathbf{M}_E = -f^{-1} \hat{\mathbf{z}} \wedge \boldsymbol{\tau}(\eta). \quad (30.78)$$

Mass transport in the Ekman layer is thus at right angles to the surface stress; a consequence of the Coriolis force.

30.5.3.2 Vertical transport into the Ekman layer

The horizontal Ekman transport induces a vertical transport into or out of the Ekman layer. To obtain an expression for the vertical transport, vertically integrate the continuity equation $\nabla \cdot \mathbf{v} = 0$ over the depth of the Ekman layer between $-\delta(x, y, t) \leq z \leq \eta(x, y, t)$

$$\int_{-\delta}^{\eta} \nabla \cdot \mathbf{v} dz = \frac{\partial}{\partial x} \left[\int_{-\delta}^{\eta} u dz \right] + \frac{\partial}{\partial y} \left[\int_{-\delta}^{\eta} v dz \right] + [w(\eta) - \mathbf{u}(\eta) \cdot \nabla \eta] - [w(-\delta) + \mathbf{u}(-\delta) \cdot \nabla \delta]. \quad (30.79)$$

For small slopes of the free surface at $z = \eta$, the outward normal to the ocean surface is

$$\hat{\mathbf{n}}^{\text{surf}} \approx \hat{\mathbf{z}} - \nabla \eta \quad \text{at } z = \eta. \quad (30.80)$$

Likewise, the outward normal (directed into the ocean interior) at the base of a gently sloping Ekman layer base at $z = -\delta$ is

$$\hat{\mathbf{n}}^{\text{base}} \approx -(\hat{\mathbf{z}} + \nabla\delta) \quad \text{at } z = -\delta. \quad (30.81)$$

Hence, the integrated continuity equation can be written

$$\frac{\partial}{\partial x} \left[\int_{-\delta}^{\eta} u \, dz \right] + \frac{\partial}{\partial y} \left[\int_{-\delta}^{\eta} v \, dz \right] + \mathbf{v}(\eta) \cdot \hat{\mathbf{n}}^{\text{surf}} + \mathbf{v}(-\delta) \cdot \hat{\mathbf{n}}^{\text{base}} = 0. \quad (30.82)$$

The surface kinematic boundary condition says that a nonzero surface transport arises from precipitation, evaporation, river runoff, as well as time changes to the free surface height

$$\mathbf{v}(\eta) \cdot \hat{\mathbf{n}}^{\text{surf}} = q_w + \frac{\partial \eta}{\partial t}. \quad (30.83)$$

Mass transport crosses the base of the Ekman layer due to net accumulation or depletion of mass transported horizontally into the layer, and across the free surface. We are interested in diagnosing this transport, which we write

$$w_B = w(-\delta) + \mathbf{u}(-\delta) \cdot \nabla\delta, \quad (30.84)$$

and which is given by

$$w_B = \mathbf{v}(\eta) \cdot \hat{\mathbf{n}}^{\text{surf}} + \frac{\partial}{\partial x} \left[\int_{-\delta}^{\eta} u \, dz \right] + \frac{\partial}{\partial y} \left[\int_{-\delta}^{\eta} v \, dz \right]. \quad (30.85)$$

The horizontal velocity appearing in the vertical integrals arises from geostrophic and ageostrophic (Ekman) currents

$$\nabla \cdot \left[\int_{-\delta}^{\eta} \mathbf{u} \, dz \right] = \nabla \cdot \left[\int_{-\delta}^{\eta} (\mathbf{u}_g + \mathbf{u}_e) \, dz \right] = \rho_o^{-1} \nabla \cdot (\mathbf{M}_g + \mathbf{M}_e), \quad (30.86)$$

where we introduced the horizontal mass transport from geostrophic and ageostrophic effects. We computed the ageostrophic effects in equation (30.78), expressing this transport in terms of the surface stress. Consequently, we have the vertical transport crossing the Ekman layer base

$$w_B = \mathbf{v}(\eta) \cdot \hat{\mathbf{n}}^{\text{surf}} + \rho_o^{-1} [\nabla \cdot \mathbf{M}_g + \hat{\mathbf{z}} \cdot \nabla \wedge (\boldsymbol{\tau}(\eta)/f)]. \quad (30.87)$$

This equation corresponds to equation (2.296a) of [Vallis \(2006\)](#), where we here keep the additional term $\mathbf{v}(\eta) \cdot \hat{\mathbf{n}}^{\text{surf}}$.

30.5.4 Further reading

This section is a supplement to Section 2.12 of [Vallis \(2006\)](#) 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 a more complete discussion.

30.6 Natural coordinates

We here make use of some elementary differential geometry to decompose the forces acting on a fluid parcel undergoing horizontal motion. The decomposition is into forces acting parallel to the instantaneous velocity, and forces acting normal to the velocity.

30.6.1 Natural coordinates

Natural coordinates for horizontal motion are defined by a locally orthogonal set of unit vectors

$$\hat{z} = \text{vertical direction} \quad (30.88a)$$

$$\hat{t} = \text{tangent to velocity} \quad (30.88b)$$

$$\hat{n} = \hat{z} \wedge \hat{t}. \quad (30.88c)$$

In this equation, \hat{t} is tangent to the velocity vector (which is horizontal), so that

$$\mathbf{u} = |\mathbf{u}| \hat{t}. \quad (30.89)$$

The unit vector \hat{n} is normal to the velocity and points to the left of the trajectory.

30.6.2 Material acceleration

When writing the velocity according to equation (30.89), we in turn decompose the acceleration into the change in speed and change in direction

$$\frac{D\mathbf{u}}{Dt} = \frac{D|\mathbf{u}|}{Dt} \hat{t} + |\mathbf{u}| \frac{D\hat{t}}{Dt}. \quad (30.90)$$

Following our discussion of rotation in Section 10.5 (see Figure 10.3), the magnitude of the direction change can be written in terms of the infinitesimal angle swept out by the motion as the fluid moves along a trajectory

$$|\delta\hat{t}| = \delta\vartheta. \quad (30.91)$$

The infinitesimal angle swept out by the trajectory is related to the radius of curvature, R , and the arc-length, δs , traversed by the trajectory

$$\delta\vartheta = \frac{\delta s}{R}. \quad (30.92)$$

Finally, the infinitesimal change in tangent, $\delta\hat{t}$, is directed normal to the motion along the \hat{n} direction. We see this property by noting that

$$\hat{t} \cdot \hat{t} = 1 \Rightarrow \delta\hat{t} \cdot \hat{t} = 0. \quad (30.93)$$

That is, $\delta\hat{t}$ is orthogonal to \hat{t} , so that it points parallel to \hat{n} (see also Section 3.1.4).

Bringing these results together leads to the constraints for the increments following the motion

$$\delta\hat{t} = \hat{n} \frac{\delta s}{R}, \quad (30.94)$$

so that the material time change is

$$\frac{D\hat{t}}{Dt} = \frac{D\hat{t}}{Ds} \frac{Ds}{Dt} \quad (30.95a)$$

$$= \frac{\hat{n}}{R} \frac{Ds}{Dt} \quad (30.95b)$$

$$= \frac{\hat{n}}{R} |\mathbf{u}|, \quad (30.95c)$$

where the speed is given by the time change of the arc-length along the trajectory

$$|\mathbf{u}| = \frac{Ds}{Dt}. \quad (30.96)$$

Combining these results renders the acceleration

$$\frac{D\mathbf{u}}{Dt} = \frac{D|\mathbf{u}|}{Dt} \hat{\mathbf{t}} + \frac{|\mathbf{u}|^2}{R} \hat{\mathbf{n}}. \quad (30.97)$$

The acceleration has thus been decomposed into the change in speed of the parcel plus the centripetal acceleration due to the curvature of the trajectory.

30.6.3 Coriolis acceleration and pressure gradient acceleration

The Coriolis acceleration takes the following form in natural coordinates

$$-f \hat{\mathbf{z}} \wedge \mathbf{u} = -(\hat{\mathbf{z}} \wedge \hat{\mathbf{t}}) f |\mathbf{u}| \quad (30.98a)$$

$$= -\hat{\mathbf{n}} f |\mathbf{u}|, \quad (30.98b)$$

so that the Coriolis acceleration always points normal to the direction of the flow. In contrast, the pressure gradient has two components

$$\nabla p = \hat{\mathbf{t}} (\hat{\mathbf{t}} \cdot \nabla p) + \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \nabla p) \quad (30.99a)$$

$$= \hat{\mathbf{t}} \frac{\partial p}{\partial s} + \hat{\mathbf{n}} \frac{\partial p}{\partial n}. \quad (30.99b)$$

30.6.4 Horizontal momentum equation

Bringing the above results together leads to the horizontal momentum equation decomposed into natural coordinates

$$\frac{D|\mathbf{u}|}{Dt} = -\frac{\partial p}{\partial s} \quad (30.100a)$$

$$\frac{|\mathbf{u}|^2}{R} + f |\mathbf{u}| = -\frac{\partial p}{\partial n}. \quad (30.100b)$$

These equations decompose the forces into those acting parallel to and normal to the velocity. Motion parallel to the pressure contours experiences no pressure gradient ($\partial p / \partial s = 0$), so that the speed of this motion 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 pressure gradient and Coriolis.

30.6.5 Further reading

Section 3.2.1 of [Holton \(1992\)](#) details the use of natural coordinates, with a similar decomposition provided in Section 7.10 of [Gill \(1982\)](#).

31

Dynamics with generalized vertical coordinates[†]

In this chapter we derive the dynamical equations using generalized vertical coordinates. These equations provide the foundations for many current treatments found in ocean and atmosphere modeling. After deriving the general equations we specialize to a hydrostatic Boussinesq fluid and make use of isopycnal vertical coordinates. Isopycnal coordinates are synonymous with buoyancy coordinates and thus correspond to specific entropy coordinates in atmospheric modeling. Next we describe 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 8 and the corresponding kinematics in Chapter 18. We also make use of the dynamical equations derived in Chapter 26. For most purposes in this chapter we find Cartesian horizontal coordinates sufficient. However, we note some places where spherical coordinates warrant special consideration. Sections 31.3 and 31.4 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.

31.1	GVC equations of motion	454
31.1.1	Mass and tracer equations	454
31.1.2	Momentum equation	454
31.1.3	Eulerian flux-form horizontal momentum equation	455
31.1.4	Vector-invariant horizontal momentum equation	455
31.1.5	Hydrostatic balance with constant gravitational acceleration	456
31.1.6	Difficulties computing the horizontal pressure gradient	456
31.2	Isopycnal equations for hydrostatic Boussinesq fluids	457
31.2.1	Montgomery potential and the pressure force	457
31.2.2	Material time derivative	459
31.2.3	Layer thickness equation	459
31.2.4	Equations of motion	460
31.2.5	Vector-invariant horizontal momentum equation	460

31.2.6	Connection to the shallow water equations	460
31.3	A split-explicit algorithm for hydrostatic models	461
31.3.1	Two common vertical coordinates	461
31.3.2	Depth integrated mass and volume budgets	462
31.3.3	Fast and slow motions	463
31.3.4	The pressure gradient for Boussinesq fluids	464
31.3.5	The pressure gradient for non-Boussinesq fluids	465
31.3.6	Depth integrated momentum equation	465
31.3.7	Summary of the split-explicit algorithm	466
31.3.8	Comments	467
31.4	The vertical Lagrangian-remap algorithm	468
31.4.1	Ocean model equations	468
31.4.2	Outlining the algorithm	469
31.4.3	Two specific examples	471
31.4.4	Connection to operator splitting	472
31.4.5	Comments	473
31.5	Numerically diagnosing fluid particle trajectories	473
31.5.1	Basics of estimating fluid particle trajectories	473
31.5.2	High wave number power in the vertical velocity	473
31.5.3	Trajectories from the vertical Lagrangian-remap algorithm	474
31.5.4	Interpolation versus extrapolation	474

31.1 GVC 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 18.8 and 18.9.

31.1.1 Mass and tracer equations

We derived the GVC version of mass continuity in Section 18.8, and the tracer equation in Section 18.9, here written in their Eulerian flux-form

$$\frac{\partial(\rho h)}{\partial t} + \nabla_\sigma \cdot (\rho h \mathbf{u}) + \delta_\sigma(\rho w^{(\dot{\sigma})}) = 0 \quad (31.1a)$$

$$\frac{\partial(h \rho C)}{\partial t} + \nabla_\sigma \cdot (h \rho C \mathbf{u} + h \mathbf{J}^h) + \delta_\sigma(\rho C w^{(\dot{\sigma})} + z_\sigma \nabla \sigma \cdot \mathbf{J}) = 0. \quad (31.1b)$$

Compatibility is maintained between so long as the tracer equation reduces to the thickness 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.

31.1.2 Momentum equation

From Section 26.1.6, the horizontal and vertical components to the momentum equation are

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{u} = -\rho \nabla_z \Phi - \nabla_z p + \rho \mathbf{F}^h \quad \text{horizontal} \quad (31.2a)$$

$$\rho \frac{Dw}{Dt} = -\rho \frac{\partial \Phi}{\partial z} - \frac{\partial p}{\partial z} + \rho F^z \quad \text{vertical.} \quad (31.2b)$$

We transform the horizontal derivatives according to (see equation (8.64))

$$\nabla_z = \nabla_\sigma - (\nabla_\sigma z) \partial_z \quad (31.3)$$

thus leading to the horizontal momentum equation

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{u} = -\rho [\nabla_\sigma - (\nabla_\sigma z) \partial_z] \Phi - [\nabla_\sigma - (\nabla_\sigma z) \partial_z] p + \rho \mathbf{F}^h. \quad (31.4)$$

The corresponding vertical momentum equation is transformed rather trivially into

$$\rho \frac{Dw}{Dt} = -\frac{\partial \sigma}{\partial z} \left[\rho \frac{\partial \Phi}{\partial \sigma} - \frac{\partial p}{\partial \sigma} \right] + \rho F^z. \quad (31.5)$$

However, the vertical velocity component, w , is not directly relevant for GVCs. Rather, it is the dia-surface velocity, $w^{(\dot{\sigma})}$, from Section 18.3 that concerns us, with the two velocities related by (see equation (18.38))

$$w^{(\dot{\sigma})} = w - \left[\frac{\partial z}{\partial t} \right]_\sigma - \mathbf{u} \cdot \nabla_\sigma z. \quad (31.6)$$

For some choices of GVC, it can be more convenient to compute $w^{(\dot{\sigma})}$ according to its connection to material changes in σ

$$w^{(\dot{\sigma})} = \frac{\partial z}{\partial \sigma} \frac{D\sigma}{Dt}. \quad (31.7)$$

Furthermore, as described in Section 31.4, 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.

31.1.3 Eulerian flux-form horizontal momentum equation

As shown in Section 26.1.4, the spherical form of the horizontal momentum equation includes a contribution from the relative acceleration that has a form as for a tracer (Section 18.9)

$$h \rho \frac{Du}{Dt} = \frac{\partial(h \rho u)}{\partial t} + \nabla_\sigma \cdot (h \rho u \mathbf{u}) + \delta_\sigma(\rho u w^{(\dot{\sigma})}) \quad (31.8a)$$

$$h \rho \frac{Dv}{Dt} = \frac{\partial(h \rho v)}{\partial t} + \nabla_\sigma \cdot (h \rho v \mathbf{u}) + \delta_\sigma(\rho v w^{(\dot{\sigma})}). \quad (31.8b)$$

Also from Section 26.1.4, we see that the spherical form of the horizontal momentum equation includes a metric term containing the vertical velocity component, $w = Dz/Dt$. The appearance of w makes it awkward to transform the spherical form of the horizontal momentum equations into their generalized vertical coordinate form. In contrast, the vector-invariant form is more amenable.

31.1.4 Vector-invariant horizontal momentum equation

As noted in Section 26.1.5, the *vector-invariant* form of the momentum equation eliminates the metric terms that appear in the flux-form equations. The vector-invariant form is also suited for deriving the vorticity equation (see Section 37.7). Here, we start with the material time derivative in the GVC form (18.36c) so that the horizontal acceleration is given by

$$\frac{D\mathbf{u}}{Dt} = \left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + (\mathbf{u} \cdot \nabla_\sigma) \mathbf{u} + (w^{(\dot{\sigma})} \partial_z) \mathbf{u}. \quad (31.9)$$

Now make use of the identity (see Section 3.3.4)

$$(\mathbf{u} \cdot \nabla_\sigma) \mathbf{u} = \nabla_\sigma K + (\nabla_\sigma \wedge \mathbf{u}) \wedge \mathbf{u}, \quad (31.10)$$

where

$$K = \mathbf{u} \cdot \mathbf{u}/2 \quad (31.11)$$

is the kinetic energy per mass of the horizontal flow. Introducing the GVC version of the relative vorticity (see Section 39.5.1)

$$\tilde{\zeta} \equiv \hat{\mathbf{z}} \cdot (\nabla_\sigma \wedge \mathbf{u}) = \left[\frac{\partial v}{\partial x} \right]_\sigma - \left[\frac{\partial u}{\partial y} \right]_\sigma \quad (31.12)$$

renders

$$\frac{D\mathbf{u}}{Dt} = \left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + \nabla_\sigma K + \hat{\mathbf{z}} \tilde{\zeta} \wedge \mathbf{u} + (w^{(\dot{\sigma})} \partial_z) \mathbf{u}, \quad (31.13)$$

so that the horizontal momentum equation takes the form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + w^{(\dot{\sigma})} \frac{\partial \mathbf{u}}{\partial z} + (2\boldsymbol{\Omega} + \hat{\mathbf{z}} \tilde{\zeta}) \wedge \mathbf{u} = -\nabla_\sigma K - \nabla_z \Phi - (1/\rho) \nabla_z p + \mathbf{F}^h, \quad (31.14)$$

where again $\nabla_z = \nabla_\sigma - (\nabla_\sigma z) \partial_z$ as per equation (31.3). This equation is form-invariant regardless the horizontal coordinates, thus motivating its name.

31.1.5 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 11.1) with g a constant gravitational acceleration, then the horizontal momentum equation becomes

$$\rho \frac{D\mathbf{u}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{u} = -[\nabla_\sigma - (\nabla_\sigma z) \partial_z] p + \rho \mathbf{F}^h \quad \text{if } \Phi = g z. \quad (31.15)$$

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

$$\rho \frac{D\mathbf{u}}{Dt} + \rho f \hat{\mathbf{z}} \wedge \mathbf{u} = -[\nabla_\sigma p + \rho \nabla_\sigma \Phi] + \rho \mathbf{F}^h, \quad (31.16)$$

which also takes on the vector-invariant form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + w^{(\dot{\sigma})} \frac{\partial \mathbf{u}}{\partial z} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_\sigma K - (1/\rho) [\nabla_\sigma p + \rho \nabla_\sigma \Phi] + \mathbf{F}^h. \quad (31.17)$$

31.1.6 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 (31.3) can lead to numerical difficulties. For example, with a simple geopotential and a hydrostatic fluid, equation (31.16) shows that the horizontal pressure gradient takes the form

$$\nabla_z p = \nabla_\sigma p + \rho \nabla_\sigma \Phi = \nabla_\sigma p + g \rho (\nabla_\sigma z), \quad (31.18)$$

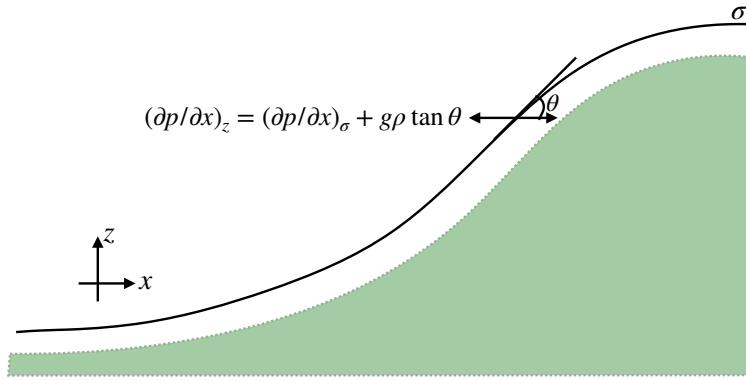


Figure 31.1: Illustrating how the horizontal pressure gradient is decomposed into two terms, one aligned with the surface of constant σ , and another associated with the slope of the σ -surface relative to the horizontal, $\tan \theta = (\partial z / \partial x)_\sigma$. We here consider the decomposition 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.

with this decomposition illustrated in Figure 31.1 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. \(2008\)](#).

31.2 Isopycnal equations for hydrostatic Boussinesq fluids

In this section we write 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 31.1, we find it pedagogically useful to start from the equations written using the geopotential vertical coordinate (see Section 29.1.6)

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z \psi + \mathbf{F} \quad \text{horizontal momentum} \quad (31.19a)$$

$$\frac{\partial \psi}{\partial z} = b \quad \text{hydrostatic} \quad (31.19b)$$

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad \text{continuity} \quad (31.19c)$$

$$\frac{Db}{Dt} = \dot{b} \quad \text{thermodynamics,} \quad (31.19d)$$

where $\mathbf{v} = (\mathbf{u}, w)$ is the full velocity field, \mathbf{u} is its horizontal component, ψ is the dynamic pressure, and b is the buoyancy.

31.2.1 Montgomery potential and the pressure force

We here consider the horizontal pressure force appearing in isopycnal models, in which we uncover the importance of the Montgomery potential.

Horizontal pressure gradient force

Following the discussion in Section 31.1.2, the horizontal pressure gradient transforms as

$$\nabla_z \psi = \nabla_b \psi - \frac{\partial \psi}{\partial z} \nabla_b z \quad (31.20a)$$

$$= \nabla_b \psi - b \nabla_b z \quad (31.20b)$$

$$= \nabla_b (\psi - bz) \quad (31.20c)$$

$$= \nabla_b M, \quad (31.20d)$$

where

$$M = \psi - bz \quad (31.21)$$

defines the Montgomery potential.

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 31.1).

Equation (31.20c) 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

$$\frac{\partial M}{\partial b} = \frac{\partial \psi}{\partial b} - b \frac{\partial z}{\partial b} - z \quad (31.22a)$$

$$= \frac{\partial \psi}{\partial z} \frac{\partial z}{\partial b} - b \frac{\partial z}{\partial b} - z \quad (31.22b)$$

$$= -z, \quad (31.22c)$$

where we made use of the hydrostatic balance $\partial \psi / \partial z = b$ (equation (31.19b)). This result means that M is the buoyancy coordinate version of pressure.

31.2.2 Material time derivative

As seen in Section 18.4, there are two equivalent forms for the material time derivative

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} \right]_z + \mathbf{u} \cdot \nabla_z + w \frac{\partial}{\partial z} \quad \text{geopotential form} \quad (31.23a)$$

$$= \left[\frac{\partial}{\partial t} \right]_b + \mathbf{u} \cdot \nabla_b + w^{(b)} \frac{\partial}{\partial z} \quad \text{isopycnal form}, \quad (31.23b)$$

where

$$w^{(b)} = \frac{\partial z}{\partial b} \frac{Db}{Dt} \quad (31.24)$$

is the diapycnal velocity component that measures the rate of flow crossing buoyancy surfaces (Section 18.3). Besides differences in the spatial operators, it is important to note that the time derivative operators in equations (31.23a) and (31.23b) 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

$$(u, v) = \frac{D(x, y)}{Dt}. \quad (31.25)$$

31.2.3 Layer 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 18.8.1, and thus quote the result here

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) + \delta_b w^{(b)} = 0, \quad (31.26)$$

where the isopycnal layer thickness (dimensions of length) is given by

$$h = \frac{\partial z}{\partial b} db = N^{-2} db. \quad (31.27)$$

We here introduced the squared buoyancy frequency

$$N^2 = \frac{\partial b}{\partial z} \quad (31.28)$$

and note that its inverse is the specific thickness

$$\frac{\partial z}{\partial b} = \text{specific thickness.} \quad (31.29)$$

Furthermore,

$$\delta_b = db \frac{\partial}{\partial b} \quad (31.30)$$

is the non-dimensional differential b -operator.

Notably, in the presence of a diabatic term, $\dot{b} \neq 0$, the three terms in the full thickness equation (31.26) are coupled. We discussed examples in Section 18.5 when examining the vertical velocity and dia-surface velocity. In contrast, for idealized studies we often consider the adiabatic case with $\dot{b} = 0$. In this case layer thickness is altered only through horizontal rearrangement of volume within a layer according to

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) = 0 \quad \text{adiabatic thickness equation.} \quad (31.31)$$

31.2.4 Equations of motion

Bringing the pieces together leads to the isopycnal version of the hydrostatic Boussinesq equations

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \mathbf{u} + (w^{(b)} \partial_z) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b M + \mathbf{F}^h \quad (31.32a)$$

$$\frac{\partial M}{\partial b} = -z \quad (31.32b)$$

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) + \delta_b w^{(b)} = 0. \quad (31.32c)$$

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.

31.2.5 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 31.1.4. Introducing the isopycnal version of the relative vorticity (see Section 39.5.1)

$$\hat{\mathbf{z}} \tilde{\zeta} \equiv \nabla_b \wedge \mathbf{u} \quad (31.33)$$

renders the vector-invariant horizontal momentum equation

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (w^{(b)} \partial_z) \mathbf{u} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b B + \mathbf{F}^h, \quad (31.34)$$

where

$$B = M + \mathbf{u} \cdot \mathbf{u}/2 = \psi - b z + \mathbf{u} \cdot \mathbf{u}/2 \quad (31.35)$$

is the Bernoulli potential for a hydrostatic Boussinesq fluid (see Section 26.3.2). Note that we can further introduce the isopycnal potential vorticity (Section 39.5.2)

$$\tilde{Q} = \frac{f + \tilde{\zeta}}{h} \quad (31.36)$$

to bring the momentum equation to the form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (w^{(b)} \partial_z) \mathbf{u} + \tilde{Q} h \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b B + \mathbf{F}^h. \quad (31.37)$$

This form is commonly used as the starting point for certain theoretical analyses, particularly when considering the adiabatic limit in which $w^{(b)} = 0$.

31.2.6 Connection to the shallow water equations

We can make use of the material time derivative operator (31.23b) to write the material form of the adiabatic and inviscid equations (31.32a)-(31.32c)

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_b M \quad (31.38a)$$

$$\frac{\partial M}{\partial b} = -z \quad (31.38b)$$

$$\frac{Dh}{Dt} + h \nabla_b \cdot \mathbf{u} = 0. \quad (31.38c)$$

These isopycnal momentum and thickness equations are isomorphic to those for a single layer of adiabatic shallow water fluid (see Section 33.1). This isomorphism allows us to derive the vorticity and potential vorticity equations in Section 39.5, making use of the shallow water manipulations from Section 36.3. 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.

31.3 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 single layer shallow water gravity waves of Chapter 35. Depending on the ocean depth, these *external* 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.

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 31.1.3

$$\frac{\partial(h\rho\mathbf{u})}{\partial t} + \nabla_\sigma \cdot [\mathbf{u}(h\rho\mathbf{u})] + \delta_\sigma(\rho\mathbf{u}w^{(\dot{\sigma})}) + f h\rho\hat{\mathbf{z}} \wedge \mathbf{u} = -h[\nabla_\sigma p + \rho\nabla_\sigma\Phi] + \rho h\mathbf{F}. \quad (31.39)$$

31.3.1 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 31.2 as part of a hybrid vertical coordinate algorithm. In Section 31.4 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

$$\sigma = p^* = p_b^0 \left[\frac{p - p_a}{p_b - p_a} \right] \implies 0 \leq p^* \leq p_b^0. \quad (31.40)$$

In this equation we made use of

$$p_b - p_a = g \int_{-H}^\eta \rho dz = g \sum_k h \rho, \quad (31.41)$$

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 (31.41) introduced the discrete vertical sum over the k grid cells in a numerical ocean model. Finally, the pressure

$$p_b^0 = \rho_0 g H, \quad (31.42)$$

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

$$\rho h = \rho \frac{\partial z}{\partial \sigma} d\sigma = \rho \frac{p_b - p_a}{p_b^0} (-\rho g)^{-1} d\sigma = -(1/g) \frac{p_b - p_a}{p_b^0} dp^*. \quad (31.43)$$

The pressure increment, dp^* , is negative since pressure decreases moving upward in the fluid column. Consequently, the layer thickness is a positive number

$$dp^* < 0 \implies h > 0, \quad (31.44)$$

and the depth integrated layer thickness times density yields the expected result

$$\sum_k h \rho = -(1/g) \frac{p_b - p_a}{p_b^0} \sum_k dp^* = (1/g) (p_b - p_a). \quad (31.45)$$

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

$$\sigma = z^* = H \left[\frac{z - \eta}{H + \eta} \right] \implies -H \leq z^* \leq 0, \quad (31.46)$$

so that the layer thickness is given by

$$h = \frac{\partial z}{\partial z^*} dz^* = (1 + \eta/H) dz^*. \quad (31.47)$$

Consequently, the depth integrated layer thickness is the expected result

$$\sum_k h = (1 + \eta/H) \sum_k dz^* = H + \eta. \quad (31.48)$$

31.3.2 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 15.3 we know that the time change for this column mass is determined by equation (15.109)

$$\frac{\partial}{\partial t} \left[\int_{-H}^{\eta} \rho dz \right] = -\nabla \cdot \mathbf{U}^\rho + Q_m, \quad (31.49)$$

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

$$g^{-1} \partial_t (p_b - p_a) = -\nabla \cdot \mathbf{U}^\rho + Q_m. \quad (31.50)$$

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. In a similar manner, Exercise 16.2 leads to the free surface equation, which expresses the conservation of volume for a column of Boussinesq fluid

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}, \quad (31.51)$$

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.

31.3.3 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

$$\mathbf{u} = \underbrace{\left[\mathbf{u} - \frac{\sum_k \mathbf{u} \rho h}{\sum_k \rho h} \right]}_{\text{slow}} + \underbrace{\left[\frac{\sum_k \mathbf{u} \rho h}{\sum_k \rho h} \right]}_{\text{fast}} \equiv \hat{\mathbf{u}} + \bar{\mathbf{u}}^z. \quad (31.52)$$

The depth-averaged or barotropic velocity is given by

$$\bar{\mathbf{u}}^z = \frac{\mathbf{U}^\rho}{\sum_k \rho h} \quad (31.53)$$

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,

$$\sum_k \rho h \hat{\mathbf{u}} = 0. \quad (31.54)$$

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

$$\hat{\mathbf{u}} = \mathbf{u}' - \frac{\sum_k \mathbf{u}' \rho h}{\sum_k \rho h}. \quad (31.55)$$

Completing the updates of \mathbf{u}' and \mathbf{U}^ρ allows for an update of the full horizontal velocity via

$$\mathbf{u} = \left[\mathbf{u}' - \frac{\sum_k \mathbf{u}' \rho h}{\sum_k \rho h} \right] + \frac{\mathbf{U}^\rho}{\sum_k \rho h}. \quad (31.56)$$

31.3.4 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 \(1982\)](#))

$$\rho_0 = 1035 \text{ kg/m}^3. \quad (31.57)$$

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

$$\frac{\partial p}{\partial z} = -g \rho = -g (\rho_o + \rho'), \quad (31.58)$$

which has an associated split in the hydrostatic pressure field

$$p = p_a(x, y, t) + p_o(x, y, t) + p'(x, y, z, t), \quad (31.59)$$

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

$$p_o(z = \eta) = 0 \quad \text{and} \quad p'(z = \eta) = 0, \quad (31.60)$$

which leads to

$$p_o = -g \rho_o (z - \eta) = -\rho_o (\Phi - g \eta), \quad (31.61)$$

and

$$p' = g \int_z^\eta \rho' dz, \quad (31.62)$$

so that

$$p = p_a + g \rho_o \eta - \rho_o \Phi + p'. \quad (31.63)$$

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

$$\nabla_z p = \nabla_\sigma p + \rho \nabla_\sigma \Phi \quad (31.64a)$$

$$= \nabla_\sigma (p_a + g \rho_o \eta - \rho_o \Phi + p') + \rho \nabla_\sigma \Phi \quad (31.64b)$$

$$= \underbrace{\nabla_\sigma (p_a + g \rho_o \eta)}_{\text{fast}} + \underbrace{\nabla_\sigma p' + \rho' \nabla_\sigma \Phi}_{\text{slow}}. \quad (31.64c)$$

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.

31.3.5 The pressure gradient for non-Boussinesq fluids

We now offer the complementary discussion for pressure based vertical coordinates. As seen in Section 31.3.2, 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

$$\Phi + g H = g \int_{-H}^z dz \quad (31.65a)$$

$$= g \int_{p_b}^p \frac{\partial z}{\partial p} dp \quad (31.65b)$$

$$= - \int_{p_b}^p \rho^{-1} dp \quad (31.65c)$$

$$= - \int_{p_b}^p (\rho_o^{-1} + \rho^{-1} - \rho_o^{-1}) dp \quad (31.65d)$$

$$= (p_b - p)/\rho_o + \rho_o^{-1} \int_{p_b}^p (\rho'/\rho) dp \quad (31.65e)$$

$$= (p_b - p)/\rho_o - (g/\rho_o) \int_{-H}^z \rho' dz. \quad (31.65f)$$

We are thus led to the identity

$$\rho_o \Phi = p_b - p + \rho_o (\Phi_b + \Phi'), \quad (31.66)$$

where

$$\rho_o \Phi' = -g \int_{-H}^z \rho' dz \quad (31.67)$$

is an anomalous geopotential similar to the anomalous hydrostatic pressure introduced in Section 31.3.4, and

$$\Phi_b = -g H \quad (31.68)$$

is the geopotential at the ocean bottom. The horizontal pressure gradient is therefore written

$$\nabla_\sigma p + \rho \nabla_\sigma \Phi = \nabla_\sigma p + (\rho/\rho_o) \nabla (p_b + \rho_o \Phi_b) - (\rho/\rho_o) \nabla_\sigma p + \rho \nabla_\sigma \Phi' \quad (31.69a)$$

$$= \underbrace{(\rho/\rho_o) \nabla (p_b + \rho_o \Phi_b)}_{\text{fast}} + \underbrace{\rho \nabla_\sigma \Phi'}_{\text{slow}} - \underbrace{(\rho'/\rho_o) \nabla_\sigma p}_{\text{slow}}. \quad (31.69b)$$

31.3.6 Depth integrated momentum equation

The thickness weighted horizontal momentum equation (31.39) 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

$$\begin{aligned} (\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho &= - \sum_k \nabla_\sigma \cdot [\mathbf{u} (h \rho \mathbf{u})] - \sum_k h [\nabla_\sigma p + \rho \nabla_\sigma \Phi] + \sum_k \rho h \mathbf{F} \\ &\quad + [\boldsymbol{\tau}^{\text{surf}} - \boldsymbol{\tau}^{\text{bott}} + Q_m \mathbf{u}_m]. \end{aligned} \quad (31.70)$$

On the left hand side we wrote

$$(\partial_t + f \hat{z} \wedge) \sum_k h \rho \mathbf{u} = (\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho. \quad (31.71)$$

On the right hand side of equation (31.70) 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 (31.69b) to write

$$\sum_k h [\nabla_\sigma p + \rho \nabla_\sigma \Phi] = \sum_k (\rho/\rho_0) h \nabla (p_b + \rho_o \Phi_b) + \sum_k h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p] \quad (31.72a)$$

$$= \frac{p_b - p_a}{g \rho_0} \nabla (p_b + \rho_o \Phi_b) + \sum_k h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p], \quad (31.72b)$$

so that the depth-integrated momentum equation (31.70) takes on the form

$$(\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho + \frac{p_b - p_a}{g \rho_0} \nabla (p_b + \rho_o \Phi_b) = - \sum_k h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p] \\ - \sum_k \nabla_\sigma \cdot [\mathbf{u} (h \rho \mathbf{u})] + \sum_k \rho h \mathbf{F} + [\boldsymbol{\tau}^{\text{surf}} - \boldsymbol{\tau}^{\text{bott}} + Q_m \mathbf{u}_m]. \quad (31.73)$$

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 $\sqrt{g H}$ (see Chapter 35). 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 (31.64c) of the pressure gradient. The resulting depth-integrated horizontal momentum equation is given by

$$(\partial_t + f \hat{z} \wedge) \mathbf{U}^\rho + (H + \eta) \nabla (p_a + g \rho_0 \eta) = - \sum_k h [\nabla_\sigma p' + \rho' \nabla_\sigma \Phi] \\ - \sum_k \nabla_\sigma \cdot [\mathbf{u} (h \rho_0 \mathbf{u})] + \sum_k \rho_0 h \mathbf{F} + [\boldsymbol{\tau}^{\text{surf}} - \boldsymbol{\tau}^{\text{bott}} + Q_m \mathbf{u}_m]. \quad (31.74)$$

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.

31.3.7 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

$$g^{-1} \partial_t(p_b - p_a) = -\nabla \cdot \mathbf{U}^\rho + Q_m \quad (31.75)$$

$$(\partial_t + f \hat{\mathbf{z}} \wedge) \mathbf{U}^\rho + \frac{p_b - p_a}{g \rho_0} \nabla (p_b + \rho_o \Phi_b) = \mathbf{G}, \quad (31.76)$$

where \mathbf{G} symbolizes those terms on the right hand side of equation (31.73). 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

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}, \quad (31.77)$$

$$(\partial_t + f \hat{\mathbf{z}} \wedge) \mathbf{U}^\rho + (H + \eta) \nabla (p_a + g \rho_0 \eta) = \mathbf{G}. \quad (31.78)$$

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 31.3.3 according to

$$\frac{\partial(h \rho \mathbf{u}')}{\partial t} + \nabla_\sigma \cdot [\mathbf{u} (h \rho \mathbf{u})] + \delta_\sigma(\rho \mathbf{u} w^{(\dot{\sigma})}) + f h \rho \hat{\mathbf{z}} \wedge \mathbf{u} = -h [\rho \nabla_\sigma \Phi' - (\rho'/\rho_o) \nabla_\sigma p] + \rho h \mathbf{F}. \quad (31.79)$$

Once \mathbf{u}' is updated we compute the baroclinic velocity according to equation (31.55)

$$\hat{\mathbf{u}} = \mathbf{u}' - \frac{\sum_k \mathbf{u}' \rho h}{\sum_k \rho h}. \quad (31.80)$$

The full velocity is then updated by

$$\mathbf{u} = \hat{\mathbf{u}} + \frac{\mathbf{U}^\rho}{\sum_k \rho h}. \quad (31.81)$$

31.3.8 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 itself 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).

31.4 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 31.3.1). 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](#)).

31.4.1 Ocean model equations

We consider the Boussinesq hydrostatic ocean primitive equations formulated in their generalized vertical coordinate form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_\sigma + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} + w^{(\dot{\sigma})} \frac{\partial \mathbf{u}}{\partial z} = -\nabla_\sigma (K + p/\rho_0) - (\rho/\rho_0) \nabla_\sigma \Phi + \mathbf{F}^h \quad \text{momentum (31.82)}$$

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad \text{hydrostatic (31.83)}$$

$$\left[\frac{\partial h}{\partial t} \right]_\sigma + \nabla_\sigma \cdot (h \mathbf{u}^\dagger) + \delta_\sigma w^{(\dot{\sigma})} = 0 \quad \text{thickness (31.84)}$$

$$\left[\frac{\partial (h C)}{\partial t} \right]_\sigma + \nabla_\sigma \cdot (h C \mathbf{u}^\dagger + h \mathbf{J}^h) + \delta_\sigma (C w^{(\dot{\sigma})} + z_\sigma \nabla \sigma \cdot \mathbf{J}) = 0 \quad \text{tracer equation (31.85)}$$

$$\rho = \rho(S, \Theta, -g\rho_0 z) \quad \text{eqn of state. (31.86)}$$

Here are some characteristics of these equations and how they are handled in light of the vertical Lagrangian-remap algorithm.

- The horizontal momentum equation (31.82) 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 24.2). 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 31.2 offers a schematic of these points.

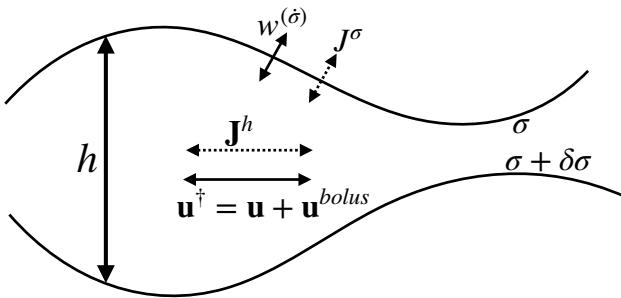


Figure 31.2: Illustrating how thickness and tracer concentration evolve in an isopycnal ocean model or more generally using a vertical Lagrangian-remap algorithm. Thickness evolves according to the layer convergence of the horizontal thickness flux, $-\nabla_\sigma \cdot (h \mathbf{u}^\dagger)$, as well 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, $-\nabla_\sigma \mathbf{J}^h - \delta_\sigma (z_\sigma \nabla_\sigma \cdot \mathbf{J})$. 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 $w^{(\dot{\sigma})} \neq 0$.

31.4.2 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 (31.82)-(31.86) are ideally setup for a vertical Lagrangian-remap algorithm. The key to this suitability is the dia-surface velocity component, $w^{(\dot{\sigma})}$, which measures transport of fluid crossing the σ surface (see Section 18.3). We here outline the vertical Lagrangian-remap algorithm, with Figure 31.3 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 18.3, 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 15.4). We thus consider a σ -surface in a real fluid that follows the fluid velocity (the barycentric velocity of Section 15.4) 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

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_{\sigma} + (f + \tilde{\zeta}) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_{\sigma} (K + p/\rho_0) - (\rho/\rho_0) \nabla_{\sigma} \Phi + \mathbf{F}^h \quad (31.87)$$

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad (31.88)$$

$$\left[\frac{\partial h}{\partial t} \right]_{\sigma} + \nabla_{\sigma} \cdot [h \mathbf{u}^{\dagger}] = 0 \quad (31.89)$$

$$\left[\frac{\partial (h C)}{\partial t} \right]_{\sigma} + \nabla_{\sigma} \cdot (h C \mathbf{u}^{\dagger} + h \mathbf{J}^h) + \delta_{\sigma}(z_{\sigma} \nabla \sigma \cdot \mathbf{J}) = 0 \quad (31.90)$$

$$\rho = \rho(S, \Theta, -g\rho_0 z). \quad (31.91)$$

The second panel of Figure 31.3 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 31.3. 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 31.4.4, 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 31.3 illustrates this final step in the vertical Lagrangian-remap algorithm.

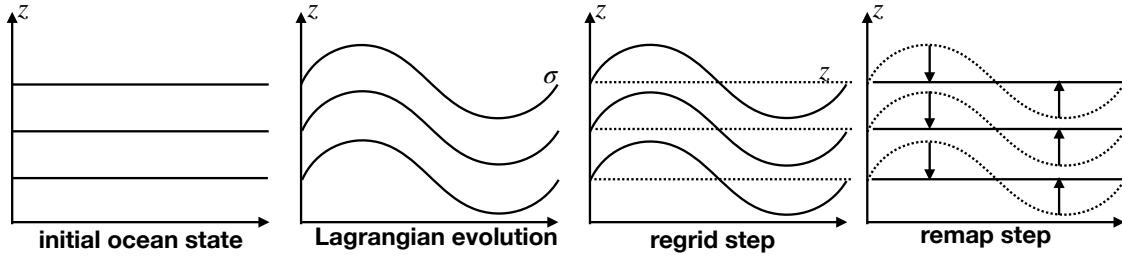


Figure 31.3: Illustrating the various steps for the vertical Lagrangian-remap algorithm used to time step the fluid state. The first panel shows an initial state where the σ -coordinate surfaces are horizontal. The second panel shows the vertical Lagrangian evolution of the σ -surfaces as determined 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.

31.4.3 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 31.3 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 (18.51) means that the vertical position of a Θ -surface evolves according to

$$\left[\frac{\partial z}{\partial t} \right]_{\Theta} = -w^{(\dot{\Theta})}. \quad (31.92)$$

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.

31.4.4 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 (31.82)-(31.86) 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

$$h^* = h^{(n)} - \Delta t \nabla_{\sigma} \cdot (h^{(n)} \mathbf{u}^\dagger) \quad \text{thickness} \quad (31.93a)$$

$$h^* C^* = h^{(n)} C^{(n)} - \Delta t \left[\nabla_{\sigma} \cdot (h C \mathbf{u}^\dagger + h \mathbf{J}^h) + \delta_{\sigma} (z_{\sigma} \nabla_{\sigma} \cdot \mathbf{J}) \right] \quad \text{tracer} \quad (31.93b)$$

$$h^{(n+1)} = h^* - \Delta t \delta_{\sigma} w^{(\dot{\sigma})} \quad \text{regrid} \quad (31.93c)$$

$$h^{(n+1)} C^{(n+1)} = h^* C^* - \Delta t \delta_{\sigma} (w^{(\dot{\sigma})} C^*) \quad \text{remap tracer.} \quad (31.93d)$$

Equation (31.93a) provides an intermediate thickness, h^* , resulting from the vertical Lagrangian update realized by dropping $w^{(\dot{\sigma})}$. Similarly, equation (31.93b) 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 (31.85) is dropped. Equation 31.93c 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 (31.93c). The diagnosed dia-surface transport is then used to remap the tracer concentration in equation (31.93d). 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 (31.93a) into equation (31.93c) leads to the time discretized form of the full thickness equation (31.84). Similarly, substituting equation (31.93b) into equation (31.93d) recovers the time discrete form of the tracer equation (31.85). This sanity check verifies that the operator split algorithm offers a consistent discretization of the thickness and tracer equations.

31.4.5 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.

31.5 Numerically diagnosing fluid particle trajectories

We introduced fluid particle trajectories in Section 14.3 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 31.4. This discussion serves to outline a method to compute trajectories and to emphasize key aspects of the vertical Lagrangian-remap algorithm.

31.5.1 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}

$$\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, t_0) + \int_{t_0}^t \mathbf{V}(\mathbf{a}, t') dt'. \quad (31.94)$$

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

$$\mathbf{X}(t) = \mathbf{X}(t_0) + \int_{t_0}^t [\hat{\mathbf{x}} U(t') + \hat{\mathbf{y}} V(t') + \hat{\mathbf{z}} W(t')] dt', \quad (31.95)$$

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 8.6.

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 (31.94). There are a variety of methods used for this interpolation, with [van Sebille et al. \(2018\)](#) offering a review.

31.5.2 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

$$\partial_z w = -\nabla_z \cdot \mathbf{u} \quad \text{incompressible.} \quad (31.96)$$

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 35.2 and 35.3). 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.

31.5.3 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 (31.95), 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 31.4 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 (31.95). 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 (31.95). Their vertical position is determined according to the vertical position of its σ -surface. The middle panel of Figure 31.4 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 31.4 provides a schematic for this step.

31.5.4 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

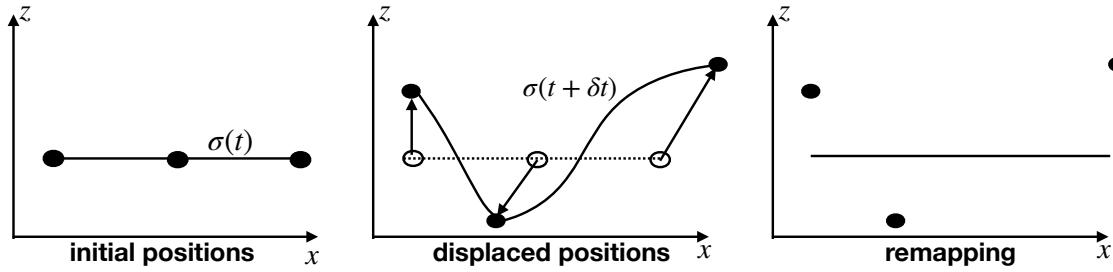


Figure 31.4: Illustrating particle trajectories with the vertical Lagrangian-remap algorithm. The left panel shows an initial horizontal line of fluid particles situated on a constant σ -surface, assumed here to be a geopotential. The middle panel shows the result of the Lagrangian step, with the displaced 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.¹ 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.

¹Time 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.

32

Space-time dependent gravity[†]

We here 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 26 as well as the gravitational and centrifugal accelerations from Section 11.1. We dispense with tensor notation in this chapter, with subscripts used here as descriptive labels rather than tensor indices.

32.1	Gravitational potential	478
32.1.1	Simple geopotential	478
32.1.2	General geopotential	478
32.1.3	Comments	479
32.2	Momentum equations	479
32.3	Primitive equations	479
32.4	Depth independent perturbed geopotential	480
32.5	Forces contributing to ocean tides	480
32.5.1	Tidal acceleration in a spherically symmetric gravity field	480
32.5.2	Heuristics of tidal acceleration on the surface of a sphere	481
32.5.3	Gravitational potential for an idealized earth-moon system	484
32.5.4	Comments on realistic tides	487

32.1 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.

32.1.1 Simple geopotential

As detailed in Section 11.1, 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 11.1.2),

$$\mathbf{g} = -\nabla \Phi, \quad (32.1)$$

with Φ the geopotential. In most applications of this book, the local vertical direction is denoted by

$$z = r - R, \quad (32.2)$$

with $z = 0$ the geopotential surface corresponding to a resting ocean. The geopotential in this case is given by

$$\Phi \approx \Phi_0 = g z, \quad (32.3)$$

with $g \approx 9.8 \text{ m s}^{-2}$ the typical value used for the acceleration due to gravity at the earth's surface.

32.1.2 General geopotential

Consider a generalized geopotential written in the form

$$\Phi = \Phi_0(r) + \Phi_1(r, \lambda, \phi, t), \quad (32.4)$$

where $\Phi_0(r)$ is the unperturbed geopotential given by equation (32.3), 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 (32.4), 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.

32.1.3 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 (32.6b).

32.2 Momentum equations

As detailed in Section 26.1.3, the inviscid momentum equation for a rotating fluid in a gravitational field is given by

$$\rho \frac{D\mathbf{v}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{v} = -(\nabla p + \rho \nabla \Phi). \quad (32.5)$$

In writing the momentum equation in the form (32.5), 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 .

32.3 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

$$\frac{\partial p}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad (32.6a)$$

$$= -\rho(g + \partial_z \Phi_1). \quad (32.6b)$$

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

$$\rho \frac{D\mathbf{u}}{Dt} + \hat{\mathbf{z}} f \wedge \rho \mathbf{u} = -(\rho \nabla_z \Phi_1 + \nabla_z p) \quad (32.7)$$

where ∇_z is the horizontal gradient taken on surfaces of constant z . In their oceanic Boussinesq form (Chapter 29), the inviscid horizontal momentum equation becomes

$$\frac{D\mathbf{u}}{Dt} + \hat{\mathbf{z}} f \wedge \mathbf{u} = -(1/\rho_0)(\rho_0 \nabla_z \Phi_1 + \nabla_z p) \quad (32.8)$$

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 .

32.4 Depth independent perturbed geopotential

A particularly simple form of Φ_1 occurs when it is depth independent,

$$\Phi_1 = \Phi_1(\lambda, \phi, t), \quad (32.9)$$

in which case the hydrostatic balance (32.6b) 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

$$\Phi_1 = -g h, \quad (32.10)$$

with $h = h(\lambda, \phi, t)$ the perturbed geopotential height field. The full geopotential is thus written

$$\Phi = g(z - h), \quad (32.11)$$

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

$$-\int_{-H}^{\eta} \nabla_z \Phi_1 dz = g \int_{-H}^{\eta} \nabla_z h dz = g(H + \eta) \nabla_z h. \quad (32.12)$$

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.

32.5 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.

32.5.1 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 32.1 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 11.1.1)

$$\mathbf{g} = -\frac{GM}{r^2} \hat{\mathbf{r}}, \quad (32.13)$$

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 32.1, 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

$$\mathbf{g}(r_B) - \mathbf{g}(r_A) = \mathbf{g}(r_0 - L/2) - \mathbf{g}(r_0 + L/2), \quad (32.14)$$

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

$$\mathbf{g}(r_B) - \mathbf{g}(r_A) \approx -L \frac{\partial \mathbf{g}}{\partial r} = -2L \frac{GM}{r_0^3} \hat{\mathbf{r}} = (2L/r_0) \mathbf{g}(r_0). \quad (32.15)$$

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 32.5.3 the gravitational acceleration generated from a remote body (e.g., the moon) acting on the surface of a sphere (e.g., the earth).

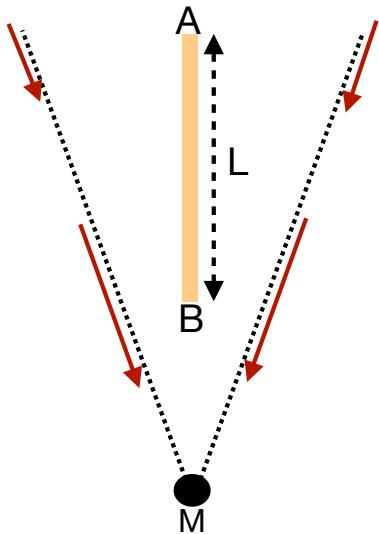


Figure 32.1: Tidal acceleration is the acceleration that acts on a finite sized object placed in a non-uniform gravitational field. The finite object is here depicted as a narrow rod of length L placed in the gravity field of a spherically symmetric body of mass M . That portion of the rod closer to 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.

32.5.2 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

32.2 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 32.5.1, 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 32.1, the tidal acceleration at point *B* relative to the center of the earth is given by

$$\mathbf{g}(r_B) - \mathbf{g}(R_{\text{em}}) = (2R_e/R_{\text{em}}) \mathbf{g}(R_{\text{em}}). \quad (32.16)$$

This acceleration points towards the moon. In contrast, the tidal acceleration at point *A* relative to the center of the earth is given by

$$\mathbf{g}(r_A) - \mathbf{g}(R_{\text{em}}) = -(2R_e/R_{\text{em}}) \mathbf{g}(R_{\text{em}}), \quad (32.17)$$

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 center. Hence, they 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. We can compute the gravitational acceleration at these intermediate points, but the trigonometry is somewhat complex. We prefer here to invoke 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 32.2. 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 32.5.3 by computing the gravitational potential for this idealized earth-moon system and then taking the gradient to compute the gravitational acceleration (see Figure 32.4).

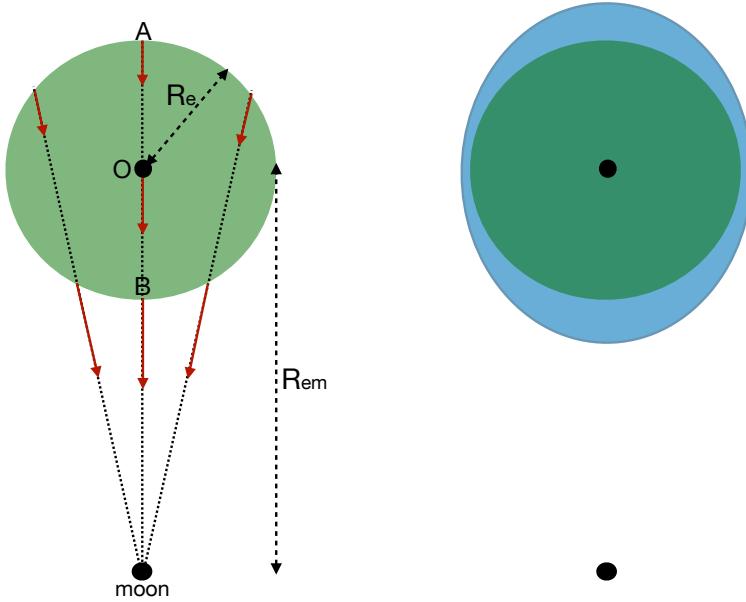


Figure 32.2: Illustrating the tidal force on the surface of a sphere. The sphere is an ideal depiction of the earth and the smaller massive object is the moon. The distance between the center of the earth and moon is R_{em} , and the radius of the earth is R_e . The left panel shows representative moon-generated 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 32.1. 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 have two bulges as shown in the right panel, resulting from the tidal accelerations acting along the earth's surface. As shown in Section 32.5.3, 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.

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

$$\mathbf{a}_{\text{orbital centrifugal}} = -\mathbf{g}(R_{\text{em}}). \quad (32.18)$$

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 11.1). 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 (32.16) and (32.17). We are thus led to the same result as before.

32.5.3 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 11.1.1, whereby the gravitational potential for a point at distance r from the center of a spherical earth is given by

$$\Phi_e(r) = -\frac{GM_e}{r}, \quad (32.19)$$

where M_e is the mass of the earth. The corresponding radial gravitational acceleration is given by

$$\mathbf{g}_e = -\nabla\Phi_e = -\frac{GM_e \hat{\mathbf{r}}}{r^2}. \quad (32.20)$$

The same considerations hold for the moon's gravitational potential. Hence, referring to Figure 32.3, the moon's gravitational potential evaluated at a distance L from the moon's center is given by

$$\Phi_m(L) = -\frac{GM_m}{L}. \quad (32.21)$$

Trigonometry leads to the law of cosines relation

$$L^2 = (R_{em} - r \cos \psi)^2 + (r \sin \psi)^2 = R_{em}^2 + r^2 - 2r R_{em} \cos \psi, \quad (32.22)$$

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 32.3).

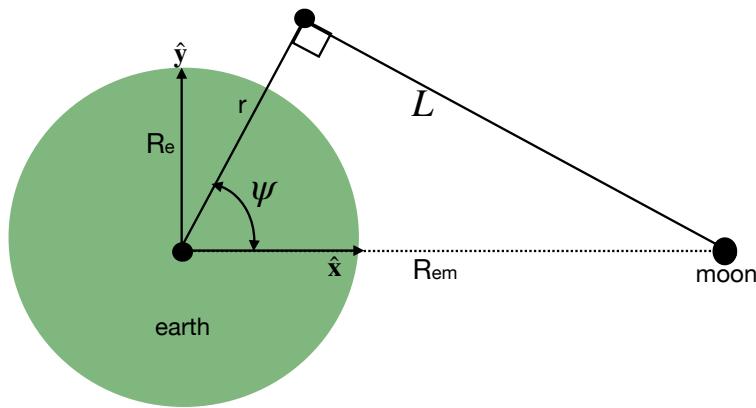


Figure 32.3: Geometry of an idealized earth-moon system. The center of the earth is a distance R_{em}^2 from the center of the moon; the moon has a mass M_m ; and the earth has a radius R_e . An arbitrary test point is shown a distance L from the center of the moon, r from the center of the earth, and with a polar angle ψ relative to the $\hat{\mathbf{x}}$ axis, where the $\hat{\mathbf{x}}$ axis points from the earth's 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 7.3 for details on relating polar and Cartesian coordinates.

Identifying the leading order contributions

Assuming the test point in Figure 32.3 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

$$\Phi_m(L) = -\frac{GM_m}{L} = -\frac{GM_m}{R_{\text{em}}} \left[1 + \frac{r \cos \psi}{R_{\text{em}}} + \frac{r^2}{2R_{\text{em}}^2} (3 \cos^2 \psi - 1) + \mathcal{O}(r/R_{\text{em}})^3 \right]. \quad (32.23)$$

We thus identify the leading three terms to the geopotential

$$\Phi_m^{(0)} = -\frac{GM_m}{R_{\text{em}}} \quad (32.24)$$

$$\Phi_m^{(1)} = -\frac{GM_m}{R_{\text{em}}^2} r \cos \psi \quad (32.25)$$

$$\Phi_m^{(2)} = -\frac{GM_m}{2R_{\text{em}}^3} r^2 (3 \cos^2 \psi - 1). \quad (32.26)$$

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.

Acceleration maintaining the orbiting earth-moon system

For the first order term, $\Phi_m^{(1)}$, we introduce the Cartesian coordinate as in Figure 32.3 to write

$$\Phi_m^{(1)} = -\frac{GM_m x}{R_{\text{em}}^2}, \quad (32.27)$$

where $x = r \cos \psi$ is the distance along \hat{x} . Hence, the gradient of $\Phi_m^{(1)}$ leads to the gravitational acceleration

$$\mathbf{g}_m^{(1)} = -\nabla \Phi_m^{(1)} = \hat{x} \frac{GM_m}{R_{\text{em}}^2}. \quad (32.28)$$

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 32.5.2, 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

$$\frac{M_m/R_{\text{em}}^2}{M_e/R_e^2} \approx 3.4 \times 10^{-6}, \quad (32.29)$$

where we set

$$M_e = 5.97 \times 10^{24} \text{ kg} \quad M_m = 7.35 \times 10^{22} \text{ kg} = (1/81.2) M_e \quad (32.30a)$$

$$R_e = 6378 \text{ km} \quad R_{\text{em}} = 384 \times 10^3 \text{ km} = 60.2 R_e. \quad (32.30b)$$

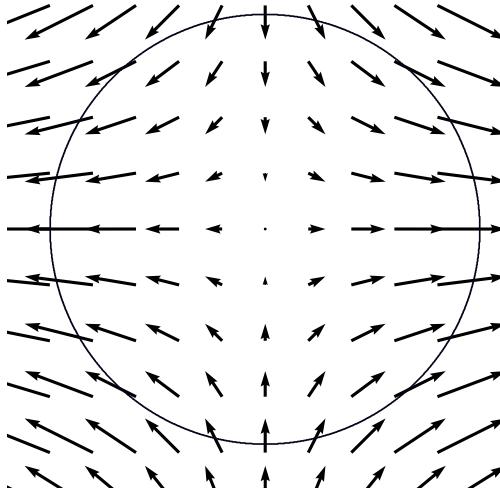


Figure 32.4: The tide producing gravitational acceleration $\mathbf{g}_m^{(2)}$ given by equation (32.34). The moon is assumed to be positioned in the equatorial plane of the earth.

Tide producing geopotential

The main tide producing acceleration results from $\Phi_m^{(2)}$. Introducing the second Cartesian coordinate, $y = r \sin \psi$, leads to

$$\Phi_m^{(2)} = -\frac{GM_m}{2R_{em}^3} r^2 (3 \cos^2 \psi - 1) = -\frac{GM_m}{2R_{em}^3} (2x^2 - y^2). \quad (32.31)$$

The corresponding perturbed geopotential height field (see equation (32.11)) is given by

$$h = -\frac{\Phi_m^{(2)}}{g} = \frac{R_e^2}{2R_{em}^3} \frac{M_m}{M_e} r^2 (3 \cos^2 \psi - 1). \quad (32.32)$$

Letting $r = R_e$ renders

$$h = \frac{R_e^4}{2R_{em}^3} \frac{M_m}{M_e} (3 \cos^2 \psi - 1) \approx 2.8 \times 10^{-8} R_e (3 \cos^2 \psi - 1). \quad (32.33)$$

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 32.2 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

$$\mathbf{g}_m^{(2)} = -\nabla \Phi_m^{(2)} = \frac{GM_m}{R_{em}^3} (2x \hat{\mathbf{x}} - y \hat{\mathbf{y}}). \quad (32.34)$$

We illustrate the vector field $\mathbf{g}_m^{(2)}$ in Figure 32.4. 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 7.3.2

$$\hat{\mathbf{r}} = \hat{\mathbf{x}} \cos \psi + \hat{\mathbf{y}} \sin \psi \quad (32.35a)$$

$$\hat{\psi} = -\hat{\mathbf{x}} \sin \psi + \hat{\mathbf{y}} \cos \psi \quad (32.35b)$$

thus rendering

$$\mathbf{g}_m^{(2)} = \frac{GM_m R_e}{R_{em}^3} \left[\hat{\mathbf{r}} (3 \cos^2 \psi - 1) - (3/2) \hat{\psi} \sin 2\psi \right], \quad (32.36)$$

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 32.5.2 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

$$\mathbf{g}_e^{(2)} = g_e \frac{M_m}{M_e} \frac{R_e^3}{R_{em}^3} \left[\hat{\mathbf{r}} (3 \cos^2 \psi - 1) - (3/2) \hat{\psi} \sin 2\psi \right], \quad (32.37)$$

where $g_e = GM_e/R_e^2$ 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 32.4, thus leading to tidal motion.

32.5.4 Comments on 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

$$\frac{\text{moon tidal acceleration}}{\text{sun tidal acceleration}} = \frac{M_m/R_{em}^3}{M_s/R_{es}^3} \approx 2.2 \quad (32.38)$$

where we set

$$M_s = 1.99 \times 10^{30} \text{ kg} \quad R_{es} = 23460 R_e. \quad (32.39)$$

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 then projecting forward in time.

Part VII

The shallow water system

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.

33

Formulation of 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.

33.1	A single shallow water layer	492
33.1.1	Pressure within the fluid layer	492
33.1.2	Momentum equation	493
33.1.3	Thickness equation	493
33.1.4	Eulerian flux-form momentum equation	495
33.1.5	Kinematic boundary conditions	496
33.1.6	Stretching and vertical velocity	498
33.1.7	Comments and further reading	498
33.2	Emphasizing the hydrostatic approximation	499
33.2.1	Hydrostatic balance with respect to background density	499
33.2.2	Density is a uniform constant	499
33.2.3	Fluid is incompressible	500
33.2.4	Comments	500
33.3	A shallow water fluid in a rotating tank	500
33.3.1	Equations of motion	500
33.3.2	Free surface shape in solid-body rotation	501
33.3.3	Further reading	501
33.4	Reduced gravity model for the upper ocean	501
33.4.1	Momentum and thickness equations for the active layer	502
33.4.2	Relating undulations of the top and bottom layer interfaces	502
33.4.3	Momentum equation with reduced gravity	502
33.4.4	Further reading	503
33.5	Stacked shallow water equations	503
33.5.1	Model formulation	504
33.5.2	Further reading	505
33.6	Exercises	505

33.1 A single shallow water layer

Consider a homogeneous layer of fluid in a uniform effective gravitational field (gravity plus 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 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.

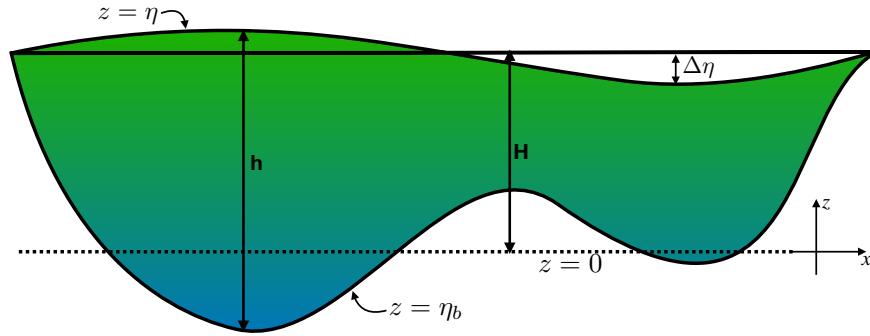


Figure 33.1: A single layer of shallow water fluid with h the thickness of the water column, H the averaged fluid depth (if the area integral of η_b is zero), $z = \eta$ the free surface height measured with respect to $z = 0$, and $z = \eta_b$ the height of the bottom boundary measured with respect 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. Finally, we assume the horizontal area occupied by the fluid to be constant, thus ignoring the case of water running up or down a beach, for example.

33.1.1 Pressure within the fluid layer

Figure 33.1 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

$$\frac{\partial p}{\partial z} = -\rho g. \quad (33.1)$$

Recall from Section 27.2 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.

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

$$p(x, y, z, t) = p_o(x, y, t) + g \rho \int_z^\eta dz \quad (33.2a)$$

$$= p_o(x, y, t) + g \rho [\eta(x, y, t) - z], \quad (33.2b)$$

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

$$\nabla_z p = \nabla_z p_o + g \rho \nabla_z \eta. \quad (33.3)$$

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. With $p_o = 0$, horizontal pressure forces within the fluid layer are determined solely by undulations in the free surface

$$\nabla_z p = g \rho \nabla \eta \quad \text{if } p_o = 0. \quad (33.4)$$

Pressure is larger in regions of high free surface (thick regions), and smaller in thin regions.

33.1.2 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

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta, \quad (33.5)$$

where

$$\mathbf{v} = (\mathbf{u}, w) \quad (33.6)$$

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

The Coriolis parameter, \mathbf{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

$$\frac{D\mathbf{u}}{Dt} = \left[\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \right] \mathbf{u} \quad (33.7)$$

so that the shallow water momentum equation (33.5) takes on the form

$$\left[\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + f \hat{\mathbf{z}} \wedge \right] \mathbf{u} = -g \nabla \eta. \quad (33.8)$$

33.1.3 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 33.2. Note that since the fluid density is constant,

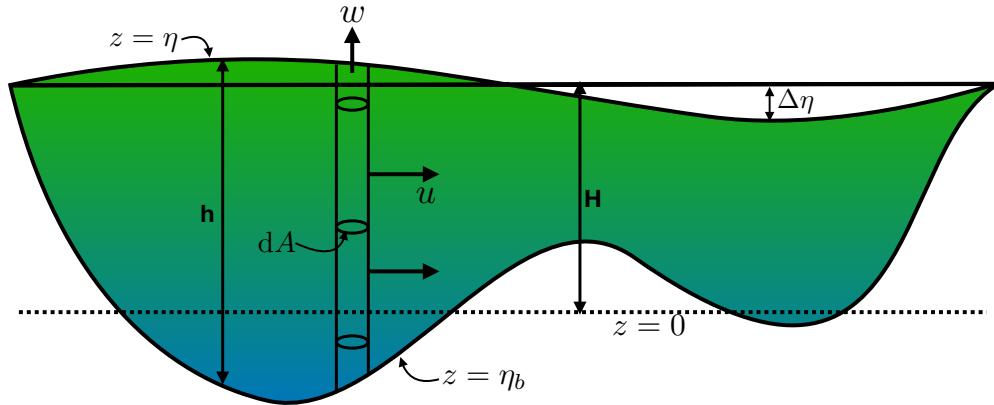


Figure 33.2: Mass budget for a column of shallow water fluid with cross-sectional area dA , constant density ρ , and thickness h . In the absence of boundary mass fluxes, the column mass is affected only by horizontal transport.

mass conservation is the same as volume conservation. Hence, the terms “mass conservation” and “volume conservation” are commonly used interchangeably.

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

$$h = \eta - \eta_b \quad (33.9)$$

(see Figure 33.2). The total mass of fluid in this column is given by

$$M = \int dA \int_{\eta_b}^{\eta} \rho dz = \rho \int (\eta - \eta_b) dA = \rho \int h dA. \quad (33.10)$$

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

$$\frac{dM}{dt} = \rho \int \frac{\partial h}{\partial t} dA, \quad (33.11)$$

where

$$\frac{\partial h}{\partial t} = \frac{\partial (\eta - \eta_b)}{\partial t} = \frac{\partial \eta}{\partial t}, \quad (33.12)$$

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

$$\text{mass per time entering column} = -\rho \int \mathbf{u} \cdot \hat{\mathbf{n}} dS, \quad (33.13)$$

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

$$\text{mass per time entering column} = -\rho \oint dl \int \mathbf{u} \cdot \hat{\mathbf{n}} dz, \quad (33.14)$$

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

$$-\rho \oint dl \int \mathbf{u} \cdot \hat{\mathbf{n}} dz = -\rho \oint h \mathbf{u} \cdot \hat{\mathbf{n}} dl \quad (33.15a)$$

$$= -\rho \int \nabla \cdot (h \mathbf{u}) dA, \quad (33.15b)$$

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 (33.11), and noting that the horizontal cross-sectional area is arbitrary, yields an equation for the layer thickness

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = 0. \quad (33.16)$$

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 (33.16) in the material form

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (33.17)$$

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 (33.15b) 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

$$\text{mass per time entering column} = -\rho \int [(u h)_{\text{east}} - (u h)_{\text{west}}] dy - \rho \int [(v h)_{\text{north}} - (v h)_{\text{south}}] dx. \quad (33.18)$$

Taking the limit as the column becomes infinitesimal leads to

$$\text{mass per time entering column} = -\rho \int \left[\frac{\partial(u h)}{\partial x} + \frac{\partial(v h)}{\partial y} \right] dx dy = -\rho \int \nabla \cdot (h \mathbf{u}) dA, \quad (33.19)$$

thus recovering the result (33.15b).

33.1.4 Eulerian flux-form momentum equation

The momentum and thickness equations are given in their Lagrangian form as

$$\frac{Du}{Dt} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (33.20)$$

To write the Eulerian flux-form momentum equation, we write the thickness weighted material acceleration as

$$h \frac{Du}{Dt} = h \frac{Du}{Dt} + \mathbf{u} \left[\frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} \right] = \frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot [\mathbf{u} (h \mathbf{u})], \quad (33.21)$$

so that the momentum equation takes the form

$$\frac{\partial(h \mathbf{u})}{\partial t} + \nabla \cdot [\mathbf{u} (h \mathbf{u})] + \mathbf{f} \wedge (h \mathbf{u}) = -g h \nabla \eta. \quad (33.22)$$

Writing $\eta = \eta_b + h$ allows us to more clearly expose the contributions from bottom topography

$$\frac{\partial(h\mathbf{u})}{\partial t} + \nabla \cdot [\mathbf{u}(h\mathbf{u})] + \mathbf{f} \wedge (h\mathbf{u}) = -g[(1/2)\nabla h^2 + h\nabla\eta_b]. \quad (33.23)$$

Expanding this vector equation into its two horizontal components renders

$$\frac{\partial(hu)}{\partial t} + \partial_x(hu^2 + gh^2/2) + \partial_y(huv) - vh f = -gh \partial_x \eta_b \quad (33.24a)$$

$$\frac{\partial(hv)}{\partial t} + \partial_x(huv) + \partial_y(hv^2 + gh^2/2) + uh f = -gh \partial_y \eta_b. \quad (33.24b)$$

We make use of this form when discussing force balances in a shallow water layer in Section 34.2.6.

33.1.5 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 15.7 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

$$z = \eta_b(x, y). \quad (33.25)$$

This location can equivalently be specified mathematically by the surface

$$s(x, y, z) = \eta_b(x, y) - z = 0. \quad (33.26)$$

The outward normal at this surface is thus given by

$$\hat{\mathbf{n}} = \frac{\nabla s}{|\nabla s|} = \frac{\nabla \eta_b - \hat{\mathbf{z}}}{\sqrt{1 + \nabla \eta_b \cdot \nabla \eta_b}}. \quad (33.27)$$

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

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta_b. \quad (33.28)$$

That is, fluid can move tangentially to the bottom, but not normal to the bottom. Making use of the bottom outward normal (33.27) leads to

$$w = \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \eta_b. \quad (33.29)$$

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 (33.29) 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

$$\frac{Ds}{Dt} = \frac{D(z - \eta_b)}{Dt} = 0 \quad \text{at } z = \eta_b. \quad (33.30)$$

Rearrangement of this result leads to the Eulerian constraint (33.29). Equivalently, we can write this boundary condition in the form

$$w = \frac{D\eta_b}{Dt} \quad \text{at } z = \eta_b. \quad (33.31)$$

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

$$s(x, y, z, t) = z - \eta(x, y, t) = 0. \quad (33.32)$$

The outward normal to the free surface is thus given by

$$\hat{\mathbf{n}} = \frac{\nabla s}{|\nabla s|} = \frac{\hat{\mathbf{z}} - \nabla \eta}{\sqrt{1 + \nabla \eta \cdot \nabla \eta}}. \quad (33.33)$$

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

$$(\mathbf{v} - \mathbf{v}^{(s)}) \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta(x, y, t), \quad (33.34)$$

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

$$\frac{\partial s}{\partial t} + \mathbf{v}^{(s)} \cdot \nabla s = 0. \quad (33.35)$$

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

$$\mathbf{v}^{(s)} \cdot \hat{\mathbf{n}} = -\frac{\partial_t s}{|\nabla s|} = \frac{\partial_t \eta}{\sqrt{1 + \nabla \eta \cdot \nabla \eta}}. \quad (33.36)$$

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 (33.36) in the no-normal flow constraint (33.34) then leads to the surface kinematic boundary condition

$$w - \mathbf{u} \cdot \nabla \eta = \frac{\partial \eta}{\partial t} \quad \text{at } z = \eta. \quad (33.37)$$

As for the bottom kinematic boundary condition written as (33.30), we can interpret this result materially, in which case

$$\frac{Ds}{Dt} = \frac{D(z - \eta)}{Dt} = 0 \quad \text{at } z = \eta. \quad (33.38)$$

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

$$w = \frac{D\eta}{Dt} \quad \text{at } z = \eta. \quad (33.39)$$

33.1.6 Stretching and vertical velocity

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

$$\nabla \cdot \mathbf{v} = 0 \Rightarrow \frac{\partial w}{\partial z} = -\nabla \cdot \mathbf{u}. \quad (33.40)$$

This result also follows since material parcels in the constant density shallow water layer maintain a constant volume (see Section 16.1). 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

$$w(z) = w(\eta_b) - (z - \eta_b) \nabla \cdot \mathbf{u}, \quad (33.41)$$

so that the vertical velocity is a linear function of depth. Applying this equation at the ocean surface yields

$$w(\eta) = w(\eta_b) - (\eta - \eta_b) \nabla \cdot \mathbf{u}. \quad (33.42)$$

Eliminating the horizontal convergence between equations (33.41) and (33.42) leads to

$$w(z) - w(\eta_b) = \left(\frac{z - \eta_b}{\eta - \eta_b} \right) [w(\eta) - w(\eta_b)]. \quad (33.43)$$

Making use of the surface kinematic boundary condition (33.39) and bottom kinematic boundary condition (33.31) renders the material form

$$\frac{1}{z - \eta_b} \left(\frac{D(z - \eta_b)}{Dt} \right) = \frac{1}{\eta - \eta_b} \left(\frac{D(\eta - \eta_b)}{Dt} \right). \quad (33.44)$$

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

$$\frac{D}{Dt} \left(\frac{z - \eta_b}{h} \right) = 0. \quad (33.45)$$

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 33.1). Consequently, equation (33.45) 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.

33.1.7 Comments and further reading

The shallow water model is ubiquitous in the geophysical fluid dynamics literature. An early application of the single shallow water layer was Laplace's studies of tides on a sphere (see exercise 33.4). Our presentation largely follows Chapter 3 of [Vallis \(2017\)](#). Another lucid discussion is given by [Salmon \(1998\)](#).

33.2 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

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \quad (33.46a)$$

$$\rho \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \rho \mathbf{v} = -\hat{\mathbf{z}} g \rho - \nabla p. \quad (33.46b)$$

In a manner analogous to our discussion of the Boussinesq approximation in Section 29.1, decompose the density and pressure into a depth dependent term and a deviation

$$\rho(x, y, z, t) = \rho_r(z) + \rho'(x, y, z, t) \quad (33.47a)$$

$$p(x, y, z, t) = p_r(z) + p'(x, y, z, t) \quad (33.47b)$$

where the reference pressure is in hydrostatic balance with the reference density

$$\frac{dp_r}{dz} = -\rho_r g. \quad (33.48)$$

Inserting this decomposition into the equation of motion leads to

$$\rho \frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \rho \mathbf{v} = -\hat{\mathbf{z}} g \rho' - \nabla p'. \quad (33.49)$$

33.2.1 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.

33.2.2 Density is a uniform constant

With $\rho = \rho_r(z)$, the mass continuity equation implies

$$w \frac{d\rho_r}{dz} = 0. \quad (33.50)$$

For a nonzero vertical velocity, this constraint means that the density is itself a constant in space and time

$$\rho = \text{constant}. \quad (33.51)$$

The perturbation hydrostatic pressure is thus given by

$$p'(x, y, z, t) = g \rho (\eta - z), \quad (33.52)$$

where $\eta = \eta(x, y, t)$ is the free surface height. The horizontal momentum equation thus takes the form

$$\frac{Du}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} = -g \nabla \eta. \quad (33.53)$$

An initial horizontal velocity that is depth independent will thus remain such.

33.2.3 Fluid is incompressible

With a uniform density, the mass continuity implies that the fluid is incompressible

$$\nabla \cdot \mathbf{v} = 0. \quad (33.54)$$

With \mathbf{u} depth independent, we can depth integrate $\nabla \cdot \mathbf{v} = 0$ as well as the kinematic boundary conditions in Section 33.1.5 to render the thickness equation

$$\frac{1}{h} \frac{Dh}{Dt} = -\nabla \cdot \mathbf{u}. \quad (33.55)$$

33.2.4 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 27.2). 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 29.1).

33.3 A shallow water fluid in a rotating tank

We introduced the geopotential in Section 11.1.2, which are surfaces where the effective gravitational force (sum of central gravity plus centrifugal) is constant. Correspondingly, we introduced geopotential vertical coordinates in Section 11.2.3. We here revisit that discussion in relation to a shallow water fluid undergoing constant rotation about the vertical ($f = 2\Omega\hat{\mathbf{z}}$). In this case, we see how the centrifugal acceleration leads to a parabolic shape for the surface of a rotating tank fluid in solid-body rotation.

33.3.1 Equations of motion

The equation of motion for a fluid in a rotating tank is given by (see also Exercise 27.4)

$$\frac{D\mathbf{u}}{Dt} + f\hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla(p/\rho + g_e z - \Omega^2 r^2/2). \quad (33.56)$$

Note the use of a gravitational and centrifugal potentials, which can be combined into a geopotential. However, we keep them separate here since we are interested in details of the motion as seen in a laboratory. We make use of polar coordinates as defined by

$$x = r \cos \theta \quad (33.57)$$

$$y = r \sin \theta, \quad (33.58)$$

with r the radial distance from the rotation axis and θ the angle made in a counter-clockwise direction from the positive x -axis.

33.3.2 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

$$p/\rho + g_e z - \Omega^2 r^2/2 = p_0/\rho, \quad (33.59)$$

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

$$h = \frac{(\Omega r)^2}{2 g_e} + \left(\frac{p_0 - p_a}{\rho g_e} \right). \quad (33.60)$$

For simplicity, assume the applied pressure is spatially constant. Hence, we specify p_0 according to the thickness at $r = 0$, so that

$$h(r) = \frac{(\Omega r)^2}{2 g_e} + h(0). \quad (33.61)$$

This column thickness is therefore parabolic in shape, with increasing thickness moving away from the rotation axis.

33.3.3 Further reading

The discussion in this section parallels that in Exercise 27.4. We consider the angular momentum for this system in Section 34.5. Refer to Section 6.6.4 of [Marshall and Plumb \(2008\)](#) for more discussion of rotating tank laboratory experiments.

33.4 Reduced gravity model for the upper ocean

The reduced gravity model describes an active layer of uniform density, ρ_1 , above a 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 ocean circulation whereby an active layer (e.g., the region above the pycnocline), sits above an inactive layer of zero motion. In this model, we introduce the *level of no motion*, below which (baroclinic) currents vanish.

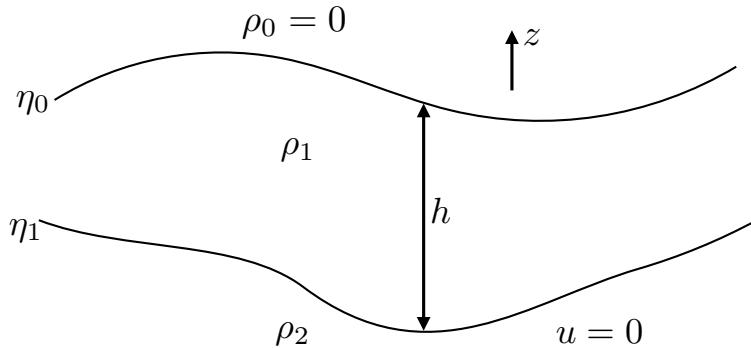


Figure 33.3: Reduced gravity model of shallow water fluid. The lower layer with density ρ_2 is dynamically inactive and thus has a zero velocity. The upper layer is dynamically active with thickness h and density ρ_1 . The dynamically active layer is bounded above by a zero density layer.

33.4.1 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 top layer is computed as (see Figure 33.3)

$$p_1(x, y, z, t) = p_0(x, y, t) + g \rho_1 (\eta_0 - z). \quad (33.62)$$

Since the fluid above the upper layer has zero density, we set

$$p_0 = 0. \quad (33.63)$$

We immediately see that the horizontal momentum equation for the top layer is written

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta_0. \quad (33.64)$$

The equations for the top layer are completed by use of the mass conservation to write the thickness equation

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (33.65)$$

33.4.2 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

$$p_2(x, y, z, t) = g \rho_1 (\eta_0 - \eta_1) + g \rho_2 (\eta_1 - z). \quad (33.66)$$

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

$$\nabla_z p_2 = g \rho_1 \nabla(\eta_0 - \eta_1) + g \rho_2 \nabla \eta_1 = 0. \quad (33.67)$$

This relation provides a constraint that links undulations of the top and bottom interfaces of the dynamically active layer

$$\eta_0 = -\eta_1 \left[\frac{\rho_2 - \rho_1}{\rho_1} \right] + \text{constant}. \quad (33.68)$$

The density ratio on the right hand side is positive but typically much smaller than unity. Hence, the relation (33.68) 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 33.4).

33.4.3 Momentum equation with reduced gravity

Relation (33.68) can be used to write the momentum equation (33.64) in the form

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = +g'_1 \nabla \eta_1, \quad (33.69)$$

where

$$g'_1 = g \left[\frac{\rho_2 - \rho_1}{\rho_1} \right] \ll g \quad (33.70)$$

defines the *reduced gravity*. It is more typical to make use of the momentum equation in the form (33.69), than the original form (33.64). The reason is that ocean hydrography measurements for decades¹ allow 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, $\nabla \eta_0$.

¹In oceanography, hydrography refers to measurements of temperature, salinity, and pressure; see [Talley et al. \(2011\)](#).

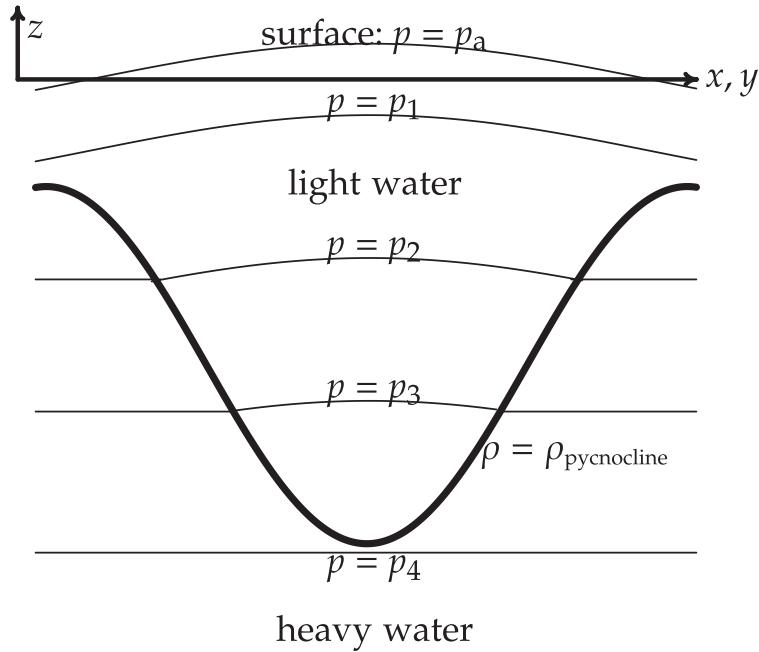


Figure 33.4: A vertical slice through a reduced gravity, or 1.5 layer, ocean in hydrostatic balance. Shown here is a plug of dynamically active light water, as may occur in a warm core eddy to the subtropical gyres, sitting on top of heavy water of zero motion. The free surface corresponds to η_0 in Figure 33.3, whereas the pycnocline (heavy black line) corresponds to the lower interface η_1 of Figure 33.3. The sea surface experiences an applied pressure $p = p_a$, assumed to be uniform for this idealized situation. Isolines of hydrostatic pressure are shown, with a slight upward bow to the isobars within the light water region, and flat isobars in the deeper region of zero motion. Note how sea level is a maximum above the pycnocline minimum, with this geometry reflected in equation (33.68). 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.

33.4.4 Further reading

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 33.4 is based on Figure 3.3 from [Tomczak and Godfrey \(1994\)](#) as well as Figure 37 from [Griffies et al. \(2014\)](#).

33.5 Stacked shallow water equations

We here consider two dynamically active shallow water layers as shown in Figure 33.5. This model offers the canonical tool for theoretical studies of baroclinic behaviour. The equations for more than two layers follows from the two-layer case by induction.

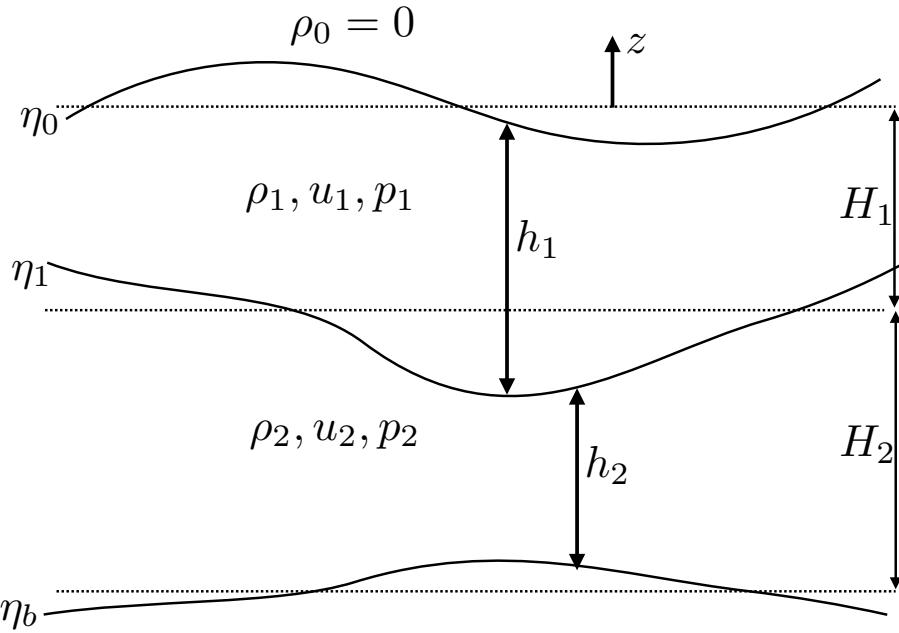


Figure 33.5: Two dynamically active layers of stacked shallow water fluid. The notation corresponds to that for the reduced gravity model of Figure 33.3, with two dynamically active layers shown in the present figure.

33.5.1 Model formulation

Each shallow water layer satisfies its own independent thickness equation, representing the conservation of mass for each layer

$$\frac{Dh_1}{Dt} = -h_1 \nabla \cdot \mathbf{u}_1 \quad (33.71)$$

$$\frac{Dh_2}{Dt} = -h_2 \nabla \cdot \mathbf{u}_2. \quad (33.72)$$

We next need the pressure in each layer to formulate the pressure forces for driving currents. As before, make use of the hydrostatic balance and integrate down from the surface (assuming zero mass layer above), which results in the pressure fields

$$p_1 = \rho_1 g (\eta_0 - z) \quad (33.73)$$

$$p_2 = \rho_1 g (\eta_0 - \eta_1) + \rho_2 g (\eta_1 - z). \quad (33.74)$$

It is convenient to write pressure in layer-two using the reduced gravity, which leads to

$$p_2 = \rho_1 g (\eta_0 - \eta_1) + \rho_2 g (\eta_1 - z) \quad (33.75a)$$

$$= g \eta_1 (\rho_2 - \rho_1) + g \rho_1 \eta_0 - g \rho_2 z \quad (33.75b)$$

$$= \rho_1 \left[g \eta_0 + g \eta_1 \frac{\rho_2 - \rho_1}{\rho_1} \right] - g \rho_2 z \quad (33.75c)$$

$$= \rho_1 (g \eta_0 + g'_1 \eta_1) - g \rho_2 z, \quad (33.75d)$$

where we introduced the reduced gravity between layers one and two

$$g'_1 = g \left[\frac{\rho_2 - \rho_1}{\rho_1} \right]. \quad (33.76)$$

The terms $g \rho_i z$ appearing in the layer pressures p_i have zero horizontal gradient. They hence play no dynamical role in determining the horizontal velocity and so will be dropped.

The horizontal momentum equations for the two layers take the general form

$$\rho_1 \left[\frac{D}{Dt} + \mathbf{f} \wedge \right] \mathbf{u}_1 = -\nabla p_1 \quad (33.77)$$

$$\rho_2 \left[\frac{D}{Dt} + \mathbf{f} \wedge \right] \mathbf{u}_2 = -\nabla p_2 \quad (33.78)$$

Making use of expressions (33.73) and (33.75d) for layer pressures leads to

$$\frac{D\mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -g \nabla \eta_0 \quad (33.79)$$

$$\frac{D\mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\frac{\rho_1}{\rho_2} (g \nabla \eta_0 + g'_1 \nabla \eta_1). \quad (33.80)$$

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 (33.71) and (33.72). We thus write

$$\eta_0 = \eta_b + h_1 + h_2 \quad (33.81)$$

$$\eta_1 = \eta_b + h_2, \quad (33.82)$$

so that

$$p_1 = \rho_1 g (\eta_b + h_1 + h_2) \quad (33.83)$$

$$p_2 = \rho_1 [g (\eta_b + h_1 + h_2) + g'_1 (\eta_b + h_2)], \quad (33.84)$$

thus resulting in the momentum equations

$$\frac{D\mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -g \nabla (\eta_b + h_1 + h_2) \quad (33.85)$$

$$\frac{D\mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\frac{\rho_1}{\rho_2} \nabla [g (\eta_b + h_1 + h_2) + g'_1 (\eta_b + h_2)]. \quad (33.86)$$

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 has a direct impact on pressure forces in the adjacent layer. Finally, when making the Boussinesq approximation, the density ratio ρ_1/ρ_2 is set to unity.

33.5.2 Further reading

The material in this section is based on Section 3.3 of [Vallis \(2017\)](#).

33.6 Exercises

EXERCISE 33.1: RELATIONS FOR VERTICAL VELOCITY (EXERCISE (3.2) OF [Vallis \(2006\)](#))

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

$$w = \left(\frac{z - \eta_b}{h} \right) \frac{Dh}{Dt} + \frac{D\eta_b}{Dt}. \quad (33.87)$$

Interpret the result, showing that it gives sensible answers at the top and bottom of the fluid layer.

EXERCISE 33.2: SHALLOW WATER LAYER WITH SURFACE VOLUME SOURCES

In Section 33.1 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 (33.38) and (33.39) 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 (33.45) 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 (36.24) 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 36.3.

EXERCISE 33.3: 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

$$\partial_x h [1 - Fr^2] = \partial_x \eta_b \quad (33.88)$$

where the Froude number is given by

$$Fr = \frac{u}{\sqrt{gh}}. \quad (33.89)$$

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 33.4: SHALLOW WATER EQUATIONS WITH TIDES

In Chapter 32 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 32.4, assume the perturbation geopotential is depth independent.

EXERCISE 33.5: 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 33.5.

34

Shallow water dynamics

We focus in this chapter on developing dynamical features for the shallow water system, including geostrophy, interfacial contact pressure forces, mechanical energy, and available potential energy.

READER'S GUIDE TO THIS CHAPTER

This chapter builds from the shallow water formulations in Chapter 33. We make use of the dynamical results here in many of the subsequent chapters.

34.1	Geostrophic balance and thermal wind	508
34.1.1	Geostrophy for a single shallow water layer	508
34.1.2	Geostrophy and thermal wind for two shallow water layers	508
34.1.3	Comments	509
34.2	Contact pressure forces and pressure form drag	510
34.2.1	Contact pressure force along vertical sides	510
34.2.2	Contact pressure force along the top and bottom interfaces	511
34.2.3	Form drag (or form stress)	512
34.2.4	Net integrated contact pressure force on the column	513
34.2.5	Contact pressure force on a single layer	513
34.2.6	Forces in a zonally re-entrant channel with topography	513
34.3	Mechanical energy budget for a shallow water layer	516
34.3.1	Gravitational potential energy	516
34.3.2	Kinetic energy per horizontal area	517
34.3.3	Mechanical energy	517
34.3.4	Further reading	518
34.4	Available potential energy of a shallow water layer	518
34.4.1	Further reading	519
34.5	Angular momentum for fluid in a rotating cylindrical tank	519
34.5.1	Angular momentum for a column of shallow water fluid	519
34.5.2	Material time evolution of the angular momentum	520
34.5.3	Materially invariant angular momentum	521
34.5.4	Comments	521
34.6	Exercises	521

34.1 Geostrophic balance and thermal wind

As described in Chapter 30, 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.

34.1.1 Geostrophy for a single shallow water layer

The geostrophic balance for a single shallow water fluid layer takes the form

$$\mathbf{f} \wedge \mathbf{u}_g = -g \nabla \eta, \quad (34.1)$$

or in component form

$$u_g = -\frac{g}{f} \frac{\partial \eta}{\partial y} \quad v_g = \frac{g}{f} \frac{\partial \eta}{\partial x}. \quad (34.2)$$

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 34.1 shows a schematic of the geostrophic balance for a single shallow water layer.

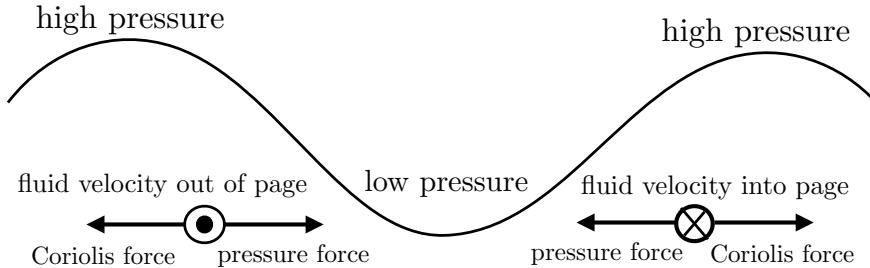


Figure 34.1: Side view of geostrophic balance for a single shallow water layer, here shown with two high pressure centers surrounding a low pressure center. The Coriolis force balances the pressure gradient force. In the northern hemisphere, where $f > 0$, geostrophic flow is counter-clockwise around a low pressure center and clockwise around a high pressure center.

34.1.2 Geostrophy and thermal wind for two shallow water layers

Now consider two shallow water layers. Recall the layer pressure equations (33.73) and (33.74), which leads to the pressure difference

$$p_1 - p_2 = g \eta_1 (\rho_1 - \rho_2) + g z (\rho_2 - \rho_1) \quad (34.3a)$$

$$= -g'_1 \rho_1 \eta_1 + g'_1 \rho_1 z, \quad (34.3b)$$

where the reduced gravity is given by

$$g'_1 = g \left[\frac{\rho_2 - \rho_1}{\rho_1} \right] \ll g. \quad (34.4)$$

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

$$\frac{D\mathbf{u}_1}{Dt} + \mathbf{f} \wedge \mathbf{u}_1 = -\rho_1^{-1} \nabla p_1 \quad (34.5a)$$

$$\frac{D\mathbf{u}_2}{Dt} + \mathbf{f} \wedge \mathbf{u}_2 = -\rho_1^{-1} \nabla p_2, \quad (34.5b)$$

where we used the top layer density as the reference density for the Boussinesq fluid. Making use of the pressure difference (34.3b) renders

$$\frac{D\Delta\mathbf{u}}{Dt} + \mathbf{f} \wedge \Delta\mathbf{u} = -\rho_1^{-1} \nabla(p_1 - p_2) \quad (34.6a)$$

$$= g'_1 \nabla \eta_1, \quad (34.6b)$$

where

$$\Delta\mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2 \quad (34.7)$$

is the vertical difference of the layer horizontal velocities. Hence, the difference in the geostrophic velocities for the two layers is proportional to the slope of the interface between the two layers

$$\mathbf{f} \wedge \Delta\mathbf{u}_g = g'_1 \nabla \eta_1, \quad (34.8)$$

which is equivalent to

$$f \Delta\mathbf{u}_g = -\hat{\mathbf{z}} \wedge g'_1 \nabla \eta_1, \quad (34.9)$$

or in component form

$$\Delta u_g = +\frac{g'_1}{f} \frac{\partial \eta_1}{\partial y} \quad \Delta v_g = -\frac{g'_1}{f} \frac{\partial \eta_1}{\partial x}. \quad (34.10)$$

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 is large. The Margules relation is a discretized version of the thermal wind relation discussed in Section 30.4.5.

34.1.3 Comments

An alternative definition of reduced gravity in equation (34.4) uses the average density,

$$\bar{\rho} = (\rho_1 + \rho_2)/2 \quad (34.11)$$

for the denominator, in which case the modified reduced gravity is

$$\bar{g}'_1 = g \left[\frac{\rho_2 - \rho_1}{\bar{\rho}} \right]. \quad (34.12)$$

Correspondingly, the pressure gradient on the right hand side of the momentum equations (34.5a) and (34.5b) have $1/\rho_1$ replaced by $1/\bar{\rho}$. Even so, the equation for the vertical shear evolution, (34.6b), remains unchanged.

34.2 Contact pressure forces and pressure form drag

Thus far, 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.

The connection between the two forms for the pressure force arise through the following form of Gauss's Law applied to scalar fields (see Section 3.7.2)

$$\mathbf{F}_{\mathcal{R}}^{\text{press}} = - \int_{\mathcal{R}} \nabla p \, dV = - \int_{\partial\mathcal{R}} p \hat{\mathbf{n}} \, dA. \quad (34.13)$$

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 26.1. Here, we pursue the contact force perspective as a means to understand the *form drag* or *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).

34.2.1 Contact pressure force along vertical sides

The pressure at a vertical position within a shallow water layer is given by (see Figure 34.2)

$$p(z) = \rho_i g (\eta_{i-1} - z) + p_{i-1}. \quad (34.14)$$

Integrating this pressure over the layer thickness yields

$$\int_{\eta_i}^{\eta_{i-1}} p(z) \, dz = \rho_i g [\eta_{i-1} (\eta_{i-1} - \eta_i) - (1/2) (\eta_{i-1}^2 - \eta_i^2)] + p_{i-1} h_i \quad (34.15a)$$

$$= \rho_i g h_i^2 / 2 + h_i p_{i-1}. \quad (34.15b)$$

For simplicity assume the column to be rectangular. The zonal pressure force, in the limit that the column becomes thin, takes the form

$$dy \int_{\eta_i}^{\eta_{i-1}} (p_L - p_R) \, dz = -dx dy \left((1/2) \rho_i g \frac{\partial h_i^2}{\partial x} + \frac{\partial(h_i p_{i-1})}{\partial x} \right) \quad (34.16a)$$

$$= -dx dy \frac{\partial}{\partial x} ((1/2) \rho_i g h_i^2 + h_i p_{i-1}), \quad (34.16b)$$

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

$$\mathbf{F}_{\text{sides}}^{\text{press}} = -dx dy \nabla ((1/2) \rho_i g h_i^2 + h_i p_{i-1}). \quad (34.17)$$

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 (34.13). However, it is useful as a means to see just how the integral theorem manifests within the shallow water system.

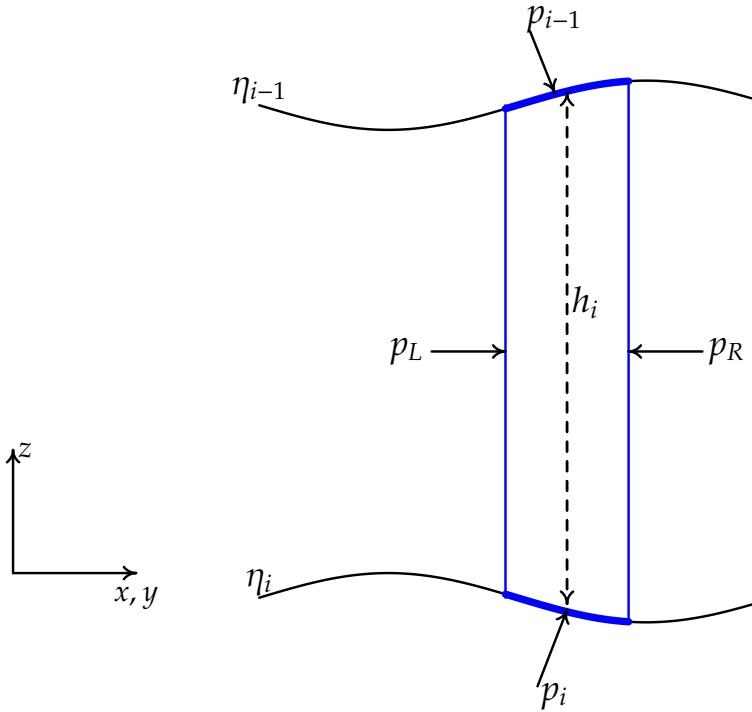


Figure 34.2: A schematic of the contact pressure force per area acting on the boundaries of a vertical column region within a shallow water layer of density ρ_i . The horizontal cross-sectional area of the column is depth independent. The interface at the lower boundary is at the vertical position $z = \eta_i$, 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.

34.2.2 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 acceleration a pressure form stress or drag. Form drag operates between two fluid layers with sloped interfaces. Form drag 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

$$\hat{\mathbf{n}}_{i-1} = \frac{\nabla(z - \eta_{i-1})}{|\nabla(z - \eta_{i-1})|} = \frac{\hat{z} - \nabla\eta_{i-1}}{\sqrt{1 + (\nabla\eta_{i-1})^2}}. \quad (34.18)$$

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

$$\hat{\mathbf{n}}_{i-1} = \frac{\hat{\mathbf{z}} - \hat{\mathbf{x}} \partial_x \eta_{i-1}}{\sqrt{1 + (\partial_x \eta_{i-1})^2}} \quad (34.19a)$$

$$= \frac{\hat{\mathbf{z}} - \hat{\mathbf{x}} \tan \phi_{i-1}}{\sqrt{1 + \tan^2 \phi_{i-1}}} \quad (34.19b)$$

$$= (\hat{\mathbf{z}} - \hat{\mathbf{x}} \tan \phi_{i-1}) \cos \phi_{i-1}, \quad (34.19c)$$

where we defined the interface slope as

$$\frac{\partial \eta_{i-1}}{\partial x} = \tan \phi_{i-1}, \quad (34.20)$$

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

$$dS_{i-1} = \frac{dx dy}{\cos \phi_{i-1}}. \quad (34.21)$$

Hence, the product of the area and the outward normal is given by

$$\hat{\mathbf{n}}_{i-1} dS_{i-1} = dx dy (\hat{\mathbf{z}} - \hat{\mathbf{x}} \partial_x \eta_{i-1}). \quad (34.22)$$

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

$$\hat{\mathbf{n}}_{i-1} dS_{i-1} = dx dy (\hat{\mathbf{z}} - \nabla \eta_{i-1}). \quad (34.23)$$

The contact pressure force at the top of the column is therefore given by

$$\mathbf{F}_{\text{top}}^{\text{press}} = -dx dy (\hat{\mathbf{z}} - \nabla \eta_{i-1}) p_{i-1}. \quad (34.24)$$

Analogous considerations lead to the contact pressure force at the bottom of the column

$$\mathbf{F}_{\text{bot}}^{\text{press}} = dx dy (\hat{\mathbf{z}} - \nabla \eta_i) p_i. \quad (34.25)$$

34.2.3 Form drag (or form stress)

As noted earlier, form drag is the horizontal projection of the contact pressure force acting on the sloped top or bottom interface of the fluid column

$$\mathbf{F}_{\text{top}}^{\text{form drag}} = dx dy (p_{i-1} \nabla \eta_{i-1}) \quad (34.26)$$

$$\mathbf{F}_{\text{bott}}^{\text{form drag}} = -dx dy (p_i \nabla \eta_i). \quad (34.27)$$

These forces are associated with the vertical exchange of horizontal momentum. This momentum exchange occurs without any exchange of matter. Rather, it occurs only through the mechanical imbalance of forces on the interfaces.

For a specific case, consider a top interface that slopes upward in the $\hat{\mathbf{x}}$ direction (e.g., see Figure 34.2). Form drag acts on this interface to accelerate the column in the $+\hat{\mathbf{x}}$ direction. For the bottom interface, a negatively sloped bottom interface is accelerated in the $+\hat{\mathbf{x}}$ direction (e.g., see Figure 34.2). Form drag at the ocean surface arises from the weight of the atmosphere above. Likewise, form drag at the ocean bottom arises from an exchange of momentum between the fluid and the solid-earth.

¹Equation (34.21) was also found in Section 15.7.3 when developing the kinematic boundary condition for a material interface.

34.2.4 Net integrated contact pressure force on the column

Summing the contact pressure forces (34.17), (34.24), and (34.25) leads to the net pressure force

$$\mathbf{F}_{\text{net}}^{\text{press}} = -dx dy [\nabla ((1/2) \rho_i g h_i^2 + h_i p_{i-1}) + (\hat{\mathbf{z}} - \nabla \eta_{i-1}) p_{i-1} - (\hat{\mathbf{z}} - \nabla \eta_i) p_i] \quad (34.28a)$$

$$= -M_i g \hat{\mathbf{z}} - dx dy [\nabla ((1/2) \rho_i g h_i^2 + h_i p_{i-1}) - p_{i-1} \nabla \eta_{i-1} + p_i \nabla \eta_i]. \quad (34.28b)$$

To reach this result, we used the hydrostatic relation for the vertical force

$$p_{i-1} - p_i = -\rho_i g h_i, \quad (34.29)$$

and introduced the column mass

$$M_i = \rho_i h_i dx dy. \quad (34.30)$$

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 (form drag) 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.

34.2.5 Contact pressure force on a single layer

As a check on our calculation of the contact pressure force (34.28b), consider a single shallow water layer under a massless atmosphere. In this case, the contact pressure force per mass is given by

$$\frac{\mathbf{F}_{\text{net}}^{\text{press}}}{M} = -g \hat{\mathbf{z}} - \frac{1}{\rho h} [\nabla ((1/2) \rho g h^2) + p_b \nabla \eta_b] \quad (34.31a)$$

$$= -g \hat{\mathbf{z}} - \frac{1}{\rho h} [\rho g h \nabla h + \rho g h \nabla \eta_b] \quad (34.31b)$$

$$= -g \hat{\mathbf{z}} - g \nabla \eta. \quad (34.31c)$$

The final equality made use of the identity

$$\eta = h + \eta_b \quad (34.32)$$

as shown in Figure 33.1. As expected, the horizontal component of this force equals to the pressure gradient body force per mass detailed in Section 33.1.1.

34.2.6 Forces in a zonally re-entrant channel with topography

We illustrate pressure form stress by considering flow 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 shallow water layer. What is the zonal and meridional force balance at steady state?

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

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = 0. \quad (34.33)$$

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

$$\int h v dx = 0 \quad (34.34)$$

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.

Meridional force balance

To study the meridional force balance we make use of the Eulerian flux-form momentum equation (33.24b)

$$\frac{\partial(h v)}{\partial t} + \partial_x(h u v) + \partial_y(h v^2 + g h^2/2) + u h f = -g h \partial_y \eta_b. \quad (34.35)$$

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

$$\frac{d}{dt} \int h v dA = - \int [g h \partial_y \eta + u h f] dA, \quad (34.36)$$

where we set $\eta = \eta_b + h$ according to Figure 33.1. In a steady state, the Coriolis force arising from zonal flow balances the meridional gradient in the surface height

$$\int h (u f + g \partial_y \eta) dA = 0. \quad (34.37)$$

Steady volume integrated zonal geostrophic transport is supported by the northern and southern walls that allow for a nonzero meridional gradient in surface height.

Zonal force balance

Now consider the zonal momentum equation written in its flux-form (see equation (33.24a))

$$\frac{\partial(h u)}{\partial t} + \partial_x(h u^2 + g h^2/2) + \partial_y(h u v) - v h f = -g h \partial_x \eta_b + \tau^x / \rho. \quad (34.38)$$

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 (34.34) eliminates the Coriolis force. We are thus left with the area integrated zonal momentum balance

$$\frac{d}{dt} \int \rho h u dA = \int [-g h \rho \partial_x \eta_b + \tau^x] dA. \quad (34.39)$$

We identify the pressure at the bottom of the layer due to the constant density fluid within the layer

$$p_{\text{bot}} = g h \rho, \quad (34.40)$$

so that the area integrated zonal momentum equation is

$$\frac{d}{dt} \int h \rho u dA = \int [-p_{\text{bot}} \partial_x \eta_b + \tau^x] dA. \quad (34.41)$$

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

$$\int p_{\text{bot}} \partial_x \eta_b dA = \int \tau^x dA. \quad (34.42)$$

Correlation between surface height and topography slope

Let us pursue the balance (34.42) 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 34.3.

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

$$p_{\text{bot}} \partial_x \eta_b = \rho g (\eta - \eta_b) \partial_x \eta_b = \rho g \eta \partial_x \eta_b - (g/2) \partial_x \eta_b^2. \quad (34.43)$$

The balance (34.42) thus becomes

$$\rho g \int \eta \partial_x \eta_b dA = \int \tau^x dA. \quad (34.44)$$

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.

Comments on stacked shallow water layers

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 30.5) 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.

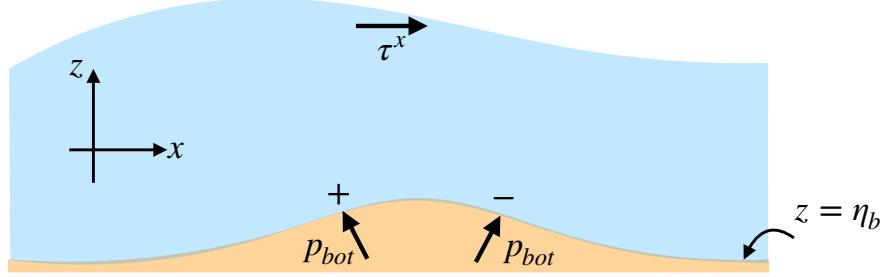


Figure 34.3: Side view of the wind stress and bottom pressure acting on a steady state layer of shallow water fluid flowing over a mountain. The eastward surface stress acts in the $+\hat{x}$ direction. As the shallow water layer acts on the solid-earth through its bottom pressure, the solid earth in 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 stresses 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.

34.3 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.

34.3.1 Gravitational potential energy

The gravitational potential energy per horizontal area of a shallow water fluid of constant density is given by

$$\mathcal{P} = g \rho \int_{\eta_b}^{\eta} z \, dz \quad (34.45a)$$

$$= \frac{g \rho}{2} (\eta^2 - \eta_b^2) \quad (34.45b)$$

$$= g \rho h (\eta - h/2). \quad (34.45c)$$

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 (34.45c) reduces to

$$\mathcal{P}_{\text{flat}} = g \rho h^2 / 2. \quad (34.46)$$

The material time tendency of the potential energy written in the form (34.45b) is

$$\frac{D\mathcal{P}}{Dt} = g \rho \left[\eta \frac{D\eta}{Dt} - \eta_b \frac{D\eta_b}{Dt} \right] \quad (34.47a)$$

$$= g \rho (\eta w_\eta - \eta_b w_b), \quad (34.47b)$$

where we used equations for the vertical velocity component from Section 33.1.6. Making further use of these equations renders

$$\frac{D\mathcal{P}}{Dt} = g \rho (\eta w_\eta - \eta_b w_b) \quad (34.48a)$$

$$= g \rho [\eta (w_b - h \nabla \cdot \mathbf{u}) - \eta_b w_b] \quad (34.48b)$$

$$= g \rho [w_b (\eta - \eta_b) - \eta h \nabla \cdot \mathbf{u}] \quad (34.48c)$$

$$= g \rho h (w_b - \eta \nabla \cdot \mathbf{u}), \quad (34.48d)$$

where we used the definition $h = \eta - \eta_b$.

34.3.2 Kinetic energy per horizontal area

As for the flat bottom case, the kinetic energy per horizontal area is

$$\mathcal{K} = \int_{\eta_b}^{\eta} \rho \mathbf{u}^2 / 2 dz = \rho h \mathbf{u}^2 / 2, \quad (34.49)$$

which has a material time derivative given by

$$\frac{D\mathcal{K}}{Dt} = -g \rho h \mathbf{u} \cdot \nabla \eta - \frac{\rho h \mathbf{u}^2}{2} \nabla \cdot \mathbf{u} \quad (34.50a)$$

$$= -g \rho h \mathbf{u} \cdot \nabla \eta - \mathcal{K} \nabla \cdot \mathbf{u}. \quad (34.50b)$$

34.3.3 Mechanical energy

Hence, the material time derivative of the mechanical energy is given by

$$\frac{D(\mathcal{K} + \mathcal{P})}{Dt} = -g h \rho \mathbf{u} \cdot \nabla \eta - \mathcal{K} \nabla \cdot \mathbf{u} + g h \rho (w_b - \eta \nabla \cdot \mathbf{u}). \quad (34.51)$$

Expanding the material time derivative into its Eulerian components leads to

$$\frac{\partial}{\partial t} (\mathcal{K} + \mathcal{P}) + \mathbf{u} \cdot \nabla (\mathcal{K} + \mathcal{P}) = -g h \rho \mathbf{u} \cdot \nabla \eta - \mathcal{K} \nabla \cdot \mathbf{u} + g h \rho (w_b - \eta \nabla \cdot \mathbf{u}), \quad (34.52)$$

with rearrangement rendering

$$\frac{\partial}{\partial t} (\mathcal{K} + \mathcal{P}) + \nabla \cdot [\mathbf{u} (\mathcal{K} + \mathcal{P})] = \mathcal{P} \nabla \cdot \mathbf{u} - g h \rho \mathbf{u} \cdot \nabla \eta + g h \rho (w_b - \eta \nabla \cdot \mathbf{u}) \quad (34.53a)$$

$$= (\mathcal{P} - g h \rho \eta) \nabla \cdot \mathbf{u} + g h \rho (w_b - \mathbf{u} \cdot \nabla \eta) \quad (34.53b)$$

$$= -(g \rho h^2 / 2) \nabla \cdot \mathbf{u} + g h \rho (w_b - \mathbf{u} \cdot \nabla \eta) \quad (34.53c)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2) + g \rho h \mathbf{u} \cdot \nabla h + g h \rho (w_b - \mathbf{u} \cdot \nabla \eta) \quad (34.53d)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2) + g \rho h (\mathbf{u} \cdot \nabla (h - \eta) + w_b) \quad (34.53e)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2) + g \rho h (-\mathbf{u} \cdot \nabla \eta_b + w_b) \quad (34.53f)$$

$$= -\nabla \cdot (\mathbf{u} g \rho h^2 / 2), \quad (34.53g)$$

where we used the identity

$$w_b = \frac{D\eta_b}{Dt} = \mathbf{u} \cdot \nabla \eta_b, \quad (34.54)$$

which follows since $\partial\eta_b/\partial t = 0$. We are thus left with the conservation law

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathbf{F} = 0, \quad (34.55)$$

which has the specific expression

$$\frac{\partial}{\partial t}(\mathcal{K} + \mathcal{P}) + \nabla \cdot [\mathbf{u}(\mathcal{K} + \mathcal{P} + g\rho h^2/2)] = 0, \quad (34.56)$$

where the total mechanical energy is

$$\mathcal{E} = \mathcal{K} + \mathcal{P} = \frac{1}{2}\rho h \mathbf{u}^2 + \frac{1}{2}\rho g (\eta^2 - \eta_b^2), \quad (34.57)$$

and the mechanical energy flux is

$$\mathbf{F} = \mathbf{u}(\mathcal{K} + \mathcal{P} + g\rho h^2/2) \quad (34.58a)$$

$$= \mathbf{u}(\mathcal{K} + \rho g \eta h). \quad (34.58b)$$

34.3.4 Further reading

This discussion is based on material in Section 3.7.2 of [Vallis \(2017\)](#).

34.4 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

$$P = g\rho \int dA \int_0^\eta z dz = \frac{g\rho}{2} \int \eta^2 dA, \quad (34.59)$$

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

$$P_{\text{ref}} = \frac{g\rho}{2} \int H^2 dA. \quad (34.60)$$

The available potential energy is the difference

$$APE = P - P_{\text{ref}} \quad (34.61a)$$

$$= \frac{g\rho}{2} \int (\eta^2 - H^2) dA \quad (34.61b)$$

$$= \frac{g\rho}{2} \int (\eta^2 - \bar{\eta}^2) dA \quad (34.61c)$$

$$= \frac{g\rho}{2} \int (\eta')^2 dA, \quad (34.61d)$$

where

$$\bar{\eta} = \frac{\int \eta dA}{A} \quad (34.62)$$

is the area averaged surface height, and

$$\eta' = \eta - \bar{\eta} \quad (34.63)$$

is the anomalous sea surface. Furthermore, we set

$$\bar{\eta} = H \quad (34.64)$$

since the total volume of the shallow water layer is constant. Equation (34.61d) 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 non-zero APE.

34.4.1 Further reading

This section is a summary of the material in Section 3.11 of [Vallis \(2017\)](#).

34.5 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 33.3, where we developed the horizontal equation of motion

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla(p/\rho + g_e z - \Omega^2 r^2/2), \quad (34.65)$$

where $r^2 = x^2 + y^2$ is the radial distance from the rotational axis,

$$\Omega = f/2 \quad (34.66)$$

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 Appendix 7.3) in the following, in which case

$$x = r \cos \theta \quad (34.67a)$$

$$y = r \sin \theta, \quad (34.67b)$$

with the polar angle θ measured counter-clockwise from the positive x -axis.

34.5.1 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 Section 12.6 and Exercises 26.1 and 27.1)

$$L^z = \delta M [\mathbf{x} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] \cdot \hat{\mathbf{z}}, \quad (34.68)$$

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

$$\mathbf{U}_{\text{solid}} = (f/2) \hat{\mathbf{z}} \wedge \mathbf{x} = r \Omega \hat{\boldsymbol{\theta}}, \quad (34.69)$$

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

$$\mathbf{x} \wedge \mathbf{u} = (x v - y u) \hat{\mathbf{z}} = r^2 \dot{\theta} \hat{\mathbf{z}}, \quad (34.70)$$

where $\dot{\theta} = D\theta/Dt$ is the angular velocity. Likewise, we have

$$\mathbf{x} \wedge \mathbf{U}_{\text{solid}} = r^2 \Omega \hat{\mathbf{z}}, \quad (34.71)$$

so that the angular momentum can be written

$$L^z = \delta M [\mathbf{x} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] \cdot \hat{\mathbf{z}} = \delta M r^2 (\dot{\theta} + \Omega). \quad (34.72)$$

34.5.2 Material time evolution of the angular momentum

The material time evolution for the angular momentum is given by

$$\frac{DL^z}{Dt} = \delta M [\mathbf{u} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] \cdot \hat{\mathbf{z}} + \delta M \left[\mathbf{x} \wedge \left(\frac{D\mathbf{u}}{Dt} + \frac{D\mathbf{U}_{\text{solid}}}{Dt} \right) \right] \cdot \hat{\mathbf{z}}. \quad (34.73)$$

Using the solid-body rotation velocity given by equation (34.69), and with a constant rotation rate, yields

$$\mathbf{u} \wedge \mathbf{U}_{\text{solid}} + \mathbf{x} \wedge \frac{D\mathbf{U}_{\text{solid}}}{Dt} = \mathbf{u} \wedge (\boldsymbol{\Omega} \wedge \mathbf{x}) + \mathbf{x} \wedge (\boldsymbol{\Omega} \wedge \mathbf{u}) \quad (34.74a)$$

$$= (\mathbf{x} \cdot \mathbf{u}) f \hat{\mathbf{z}}. \quad (34.74b)$$

Making use of the material evolution of the horizontal velocity given by equation (34.65) renders

$$\left(\mathbf{x} \wedge \frac{D\mathbf{u}}{Dt} \right) \cdot \hat{\mathbf{z}} = (\mathbf{x} \wedge [-f \hat{\mathbf{z}} \wedge \mathbf{u} - \nabla (p/\rho + g_e z - \Omega^2 r^2/2)]) \cdot \hat{\mathbf{z}} \quad (34.75a)$$

$$= -f (\mathbf{x} \cdot \mathbf{u}) - (\mathbf{x} \wedge g \nabla \eta) \cdot \hat{\mathbf{z}}. \quad (34.75b)$$

The centrifugal term dropped out since

$$\mathbf{x} \wedge \nabla r^2 = 2 \mathbf{x} \wedge r \hat{\mathbf{r}} = 2 \mathbf{x} \wedge \mathbf{x} = 0. \quad (34.76)$$

The gravitational term dropped out since

$$(\mathbf{x} \wedge \nabla z) \cdot \hat{\mathbf{z}} = (\mathbf{x} \wedge \hat{\mathbf{z}}) \cdot \hat{\mathbf{z}} = 0, \quad (34.77)$$

as does the vertical component to the pressure gradient. We are thus left with

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g (\mathbf{x} \wedge \nabla \eta) \cdot \hat{\mathbf{z}}. \quad (34.78)$$

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 (34.78). 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 (34.78) into a more transparent form by switching to polar coordinates

$$\boldsymbol{x} \wedge \nabla \eta = r \hat{\boldsymbol{r}} \wedge \left(\hat{\boldsymbol{r}} \frac{\partial \eta}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial \eta}{\partial \theta} \right) = \frac{\partial \eta}{\partial \theta} \hat{\boldsymbol{z}}, \quad (34.79)$$

so that

$$\frac{1}{\delta M} \frac{DL^z}{Dt} = -g \frac{\partial \eta}{\partial \theta}. \quad (34.80)$$

This result is directly analogous to the angular momentum evolution for a fluid moving around a sphere as derived in Exercise 26.1. 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.

34.5.3 Materially invariant angular momentum

The angular momentum for a fluid column is materially invariant (i.e., a constant on a material fluid parcel) if

$$\frac{DL^z}{Dt} = 0 \iff \frac{\partial \eta}{\partial \theta} = 0. \quad (34.81)$$

For a flat bottom, equation (33.61) 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.

34.5.4 Comments

The material evolution equation (34.78) 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 11.1.2). 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 Plumb \(2008\)](#) for more discussion of rotating tank experiments.

34.6 Exercises

EXERCISE 34.1: TROPOPAUSE AND THERMOCLINE SLOPES

We here make use of the two-layer thermal wind relations from Section 34.1.2, also known as Margules' relation, here used to estimate the slope of the atmospheric tropopause and oceanic thermocline. 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 gradient of 60K. You may wish to consult [Wallace and Hobbs \(2006\)](#) for physical scales.
- (b) Estimate a vertical interfacial displacement in the ocean thermocline associated with a temperature gradient 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.

EXERCISE 34.2: CIRCULAR STEADY GEOSTROPHIC FLOW

Consider a single layer of shallow water fluid in steady geostrophic balance on a f -plane so that

$$f \hat{z} \wedge \mathbf{u}_g = -g \nabla \eta, \quad (34.82)$$

where $f > 0$ (northern hemisphere). Assume the free surface has a circular Gaussian shape

$$\eta = \eta_0 e^{-r^2/(2\sigma^2)} \quad (34.83)$$

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 14.7.2.

EXERCISE 34.3: 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

$$\mathbf{P} = \int \rho \mathbf{u} dV = \int \rho h \mathbf{u} dA. \quad (34.84)$$

Show that

$$\frac{d\mathbf{P}}{dt} = 0 \quad (34.85)$$

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 34.4: 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 34.5: APE FOR A SINGLE SHALLOW WATER LAYER WITH BOTTOM TOPOGRAPHY

Generalize the APE discussion in Section 34.4 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 34.6: APE FOR TWO SHALLOW WATER LAYERS

Compute the APE for two shallow water layers as in Figure 33.5 with nontrivial bottom topography, $z = \eta_b(x, y)$. Show that the APE is non-negative. Assume the domain is simply connected.

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.

35.1	The linearized shallow water system	523
35.1.1	Linearizing the shallow water equations	524
35.1.2	Relative vorticity of linear shallow water fluctuations	524
35.1.3	Potential vorticity of linear shallow water fluctuations	525
35.2	Non-rotating shallow water gravity waves	525
35.2.1	Dispersionless waves	526
35.2.2	Vanishing relative vorticity	526
35.3	Inertia-gravity (Poincaré) waves	527
35.3.1	Non-dimensionalization	527
35.3.2	Dispersion relation	527
35.3.3	Shortwave limit	529
35.3.4	Longwave limit	529
35.4	Shallow water Kelvin waves	529
35.5	Geostrophic adjustment	531
35.5.1	Posing the initial value problem	531
35.5.2	Adjustment in the absence of rotation	531
35.5.3	Adjustment with rotation	532
35.5.4	Comments	533
35.6	Exercises	534

35.1 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 force.

35.1.1 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

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (35.1a)$$

$$\frac{\partial h}{\partial t} + h \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla h = 0. \quad (35.1b)$$

Since the bottom is assumed flat, the surface height equals to the thickness (see Figure 33.1)

$$\eta = h \quad \text{flat bottom.} \quad (35.2)$$

Now consider small fluctuations of the thickness and velocity relative to a state of rest

$$\eta(x, y, t) = H + \eta'(x, y, t) \quad (35.3a)$$

$$\mathbf{u}(x, y, t) = 0 + \mathbf{u}'(x, y, t). \quad (35.3b)$$

Substitution into the thickness equation (35.1b) leads to

$$\frac{\partial \eta'}{\partial t} + (H + \eta') \nabla \cdot \mathbf{u}' + \mathbf{u}' \cdot \nabla \eta' = 0. \quad (35.4)$$

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

$$\frac{\partial \eta'}{\partial t} + H \nabla \cdot \mathbf{u}' = 0 \quad \text{linearized surface height equation.} \quad (35.5)$$

Similarly, the linearized momentum equation takes the form (dropping the nonlinear advection term $(\mathbf{u}' \cdot \nabla) \mathbf{u}'$)

$$\frac{\partial \mathbf{u}'}{\partial t} + f \hat{\mathbf{z}} \wedge \mathbf{u}' = -g \nabla \eta' \quad \text{linearized momentum equation.} \quad (35.6)$$

35.1.2 Relative vorticity of linear shallow water fluctuations

We here consider how the vertical component of relative vorticity

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}) \quad (35.7)$$

evolves for the small amplitude shallow water fluctuations determined by equations (35.5) and (35.6). Taking the curl of the linearized momentum equation (35.6) leads to

$$\frac{\partial \zeta'}{\partial t} = -\nabla \cdot (f \mathbf{u}'). \quad (35.8)$$

On the f -plane and with the non-divergence condition $\nabla \cdot \mathbf{u}' + \partial_z w' = 0$, we have

$$\frac{\partial \zeta'}{\partial t} = f \frac{\partial w'}{\partial z}. \quad (35.9)$$

We can use the linearized thickness equation (35.5) to render an equivalent result

$$\frac{\partial \zeta'}{\partial t} = \frac{f}{H} \frac{\partial \eta'}{\partial t}. \quad (35.10)$$

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.

35.1.3 Potential vorticity of linear shallow water fluctuations

Rearranging the linearized vorticity equation (35.10) leads to the local (i.e., no advection) conservation law

$$\frac{\partial q}{\partial t} = 0, \quad (35.11)$$

where the linearized shallow water potential vorticity is given by²

$$q = \zeta' - \frac{f\eta'}{H}. \quad (35.12)$$

As mentioned in Section 35.1.2, 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 35.5).

35.2 Non-rotating shallow water gravity waves

For the non-rotating case ($f = 0$), the linear velocity and thickness equations are

$$\frac{\partial \eta'}{\partial t} = -H \nabla \cdot \mathbf{u}' \quad (35.13a)$$

$$\frac{\partial \mathbf{u}'}{\partial t} = -g \nabla \eta'. \quad (35.13b)$$

Equation (35.13a) 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

$$\frac{\partial^2 \eta'}{\partial t^2} = -H \frac{\partial(\nabla \cdot \mathbf{u}')}{\partial t} \quad (35.14a)$$

$$\frac{\partial(\nabla \cdot \mathbf{u}')}{\partial t} = -g \nabla^2 \eta'. \quad (35.14b)$$

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

$$\frac{\partial^2 \eta'}{\partial t^2} - g H \nabla^2 \eta' = 0. \quad (35.15)$$

The complement substitution leads to the slightly more general wave equation for the perturbation horizontal velocity

$$\frac{\partial^2 \mathbf{u}'}{\partial t^2} - g H \nabla(\nabla \cdot \mathbf{u}') = 0. \quad (35.16)$$

²The linearized potential vorticity (35.12) can be obtained by taking the limit of $|f| \gg |\zeta|$ and $H \gg \Delta\eta$ in the shallow water potential vorticity (see Section 36.3) $Q = (f + \zeta)/h$.

³In Appendix 1.4.1, we discuss the linear wave equation and show that it is the canonical hyperbolic partial differential equation.

35.2.1 Dispersionless waves

Assume a linear wave solution of the form

$$\eta' = \gamma \cos(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (35.17)$$

where γ is a constant amplitude,

$$\mathbf{k} = \hat{\mathbf{x}} k + \hat{\mathbf{y}} l \quad (35.18)$$

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 (35.13b), to the velocity fluctuation⁴

$$\mathbf{u}' = \mathbf{k} \frac{g \eta'}{\omega}, \quad (35.19)$$

which can be readily shown to satisfy the linearized velocity equation (35.16). Substitution of the surface height fluctuation (35.17) into the wave equation (35.16) leads to the relation between frequency and wave-number

$$\omega = \pm c |\mathbf{k}|, \quad (35.20)$$

where

$$c = \sqrt{g H} \quad (35.21)$$

is the shallow water gravity wave speed. The *dispersion relation* (35.20) 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.

35.2.2 Vanishing relative vorticity

Equation (35.14b) 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 (35.13b)), the associated vorticity has a zero time tendency

$$\frac{\partial(\nabla \wedge \mathbf{u}')}{\partial t} = 0. \quad (35.22)$$

This result follows from the discussion of vorticity for the shallow water waves in Section 35.1.2. Hence, if the initial flow configuration has zero vorticity, the linear gravity waves retain zero vorticity as they propagate. The velocity fluctuation (35.19) indeed has zero vorticity since

$$\hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}') = \frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y} \quad (35.23a)$$

$$= \frac{g}{\omega} (l \partial_x - k \partial_y) \eta' \quad (35.23b)$$

$$= -\frac{g \gamma}{\omega} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) (l k - k l) \quad (35.23c)$$

$$= 0. \quad (35.23d)$$

⁴Formally, we have \mathbf{u}' specified only up to an arbitrary function of space. We set that function to zero without loss of generality.

35.3 Inertia-gravity (Poincaré) waves

Now let the Coriolis parameter, f , be a nonzero constant, so that the linearized thickness equation (35.5) and momentum equation (35.6) take the form

$$\frac{\partial \eta'}{\partial t} = -H \nabla \cdot \mathbf{u}' \quad (35.24a)$$

$$\frac{\partial u'}{\partial t} - f v' = -g \frac{\partial \eta'}{\partial x} \quad (35.24b)$$

$$\frac{\partial v'}{\partial t} + f u' = -g \frac{\partial \eta'}{\partial y}. \quad (35.24c)$$

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 12.3, 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 35.2. That is, both f and g play a role as restoring forces for the waves.

The free surface equation (35.24a) remains the same as for the non-rotating case in equation (35.13a). 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.

35.3.1 Non-dimensionalization

It is convenient to non-dimensionalize the linear equations (35.24a)-(35.24c) by writing

$$\mathbf{x} = L \hat{\mathbf{x}}, \quad \mathbf{u}' = U \hat{\mathbf{u}}, \quad t = \frac{L \hat{t}}{U}, \quad f = \hat{f} T^{-1}, \quad \eta' = H \hat{\eta}, \quad (35.25)$$

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 (35.24a)-(35.24c) leads to the non-dimensional system

$$\frac{\partial \hat{\eta}}{\partial \hat{t}} + \frac{\partial \hat{u}}{\partial \hat{x}} + \frac{\partial \hat{v}}{\partial \hat{y}} = 0 \quad (35.26a)$$

$$\frac{\partial \hat{u}}{\partial \hat{t}} - \hat{f} \hat{v} = -\hat{c}^2 \frac{\partial \hat{\eta}}{\partial \hat{x}} \quad (35.26b)$$

$$\frac{\partial \hat{v}}{\partial \hat{t}} + \hat{f} \hat{u} = -\hat{c}^2 \frac{\partial \hat{\eta}}{\partial \hat{y}}, \quad (35.26c)$$

where

$$\hat{c} = \frac{\sqrt{g H}}{U} \quad (35.27)$$

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.

35.3.2 Dispersion relation

To obtain a dispersion relation we let

$$(\hat{u}, \hat{v}, \hat{\eta}) = (\tilde{u}, \tilde{v}, \tilde{\eta}) e^{i(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}} - \hat{\omega} \hat{t})}, \quad (35.28)$$

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 (35.28) given the horizontal symmetry of the linearized system (35.26a)-(35.26c). Substitution into equations (35.26a)-(35.26c) leads to the dispersion relation

$$\begin{pmatrix} -i\hat{\omega} & -\hat{f} & i\hat{c}^2\hat{k} \\ \hat{f} & -i\hat{\omega} & i\hat{c}^2\hat{l} \\ i\hat{k} & i\hat{l} & -i\hat{\omega} \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\eta} \end{pmatrix} = 0. \quad (35.29)$$

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

$$\hat{\omega} \left[\hat{\omega}^2 - \hat{f}^2 - \hat{c}^2 (\hat{k}^2 + \hat{l}^2) \right] = 0. \quad (35.30)$$

The $\hat{\omega} = 0$ solution corresponds to time-independent geostrophic motion. Reintroducing dimensions, the second solution satisfies the dispersion relation

$$\omega^2 = f^2 + c^2 (k^2 + l^2). \quad (35.31)$$

Figure 35.1 illustrates this relation. We discuss the shortwave and longwave limits in the following.

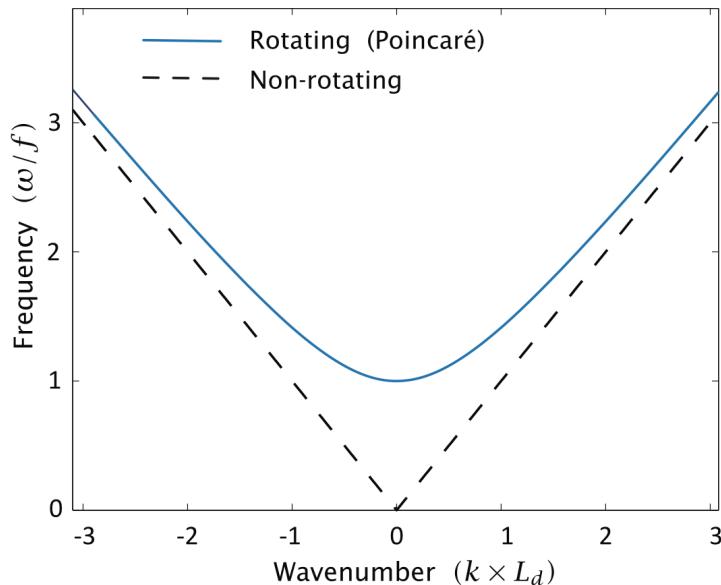


Figure 35.1: This is Figure 3.8 of [Vallis \(2017\)](#), illustrating the dispersion relation (35.31) for Poincaré waves. Frequency is scaled by the Coriolis frequency, f , and wavenumber by the inverse deformation radius \sqrt{gH}/f . For small wave number (large wave length relative to the deformation 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.

35.3.3 Shortwave limit

The short wave limit is in the regime where

$$k^2 + l^2 \gg \frac{f^2}{g H} = \frac{1}{L_d^2}, \quad (35.32)$$

where we introduced the shallow water deformation radius⁵

$$L_d = \frac{c}{f} = \frac{\sqrt{g H}}{f}. \quad (35.33)$$

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

$$\lambda \ll 2\pi L_d. \quad (35.34)$$

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 (35.31) reduces to the non-rotating dispersion relation (35.20)

$$\omega \approx \pm c |\mathbf{k}|. \quad (35.35)$$

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.

35.3.4 Longwave limit

The opposite limit occurs when

$$k^2 + l^2 \ll \frac{1}{L_d^2}, \quad (35.36)$$

so the waves are much longer than the deformation radius. In this limit the dispersion relation is

$$\omega = \pm f, \quad (35.37)$$

which are known as inertial oscillations. As discussed in Section 12.3, inertial oscillations are unaffected by pressure forces otherwise arising from surface height undulations. Instead, they are determined only by the Coriolis frequency.

35.4 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 35.5.3.

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

$$\frac{\partial \eta'}{\partial t} = -H \frac{\partial u'}{\partial x} \quad (35.38a)$$

$$\frac{\partial u'}{\partial t} = -g \frac{\partial \eta'}{\partial x} \quad (35.38b)$$

$$f u' = -g \frac{\partial \eta'}{\partial y}. \quad (35.38c)$$

Equations (35.38a) and (35.38c) lead to the one-dimensional wave equation for the zonal velocity fluctuation

$$\frac{\partial^2 u'}{\partial t^2} - c^2 \frac{\partial^2 u'}{\partial x^2} = 0, \quad (35.39)$$

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

$$u'(x, y, t) = F_1(x + ct, y) + F_2(x - ct, y), \quad (35.40)$$

with corresponding surface height displacement

$$\eta'(x, y, t) = \sqrt{H/g} [-F_1(x + ct, y) + F_2(x - ct, y)]. \quad (35.41)$$

We now substitute this form of the solution into equation (35.38c) to determine the y -dependence

$$\frac{\partial F_1}{\partial y} = \frac{f F_1}{c} \quad (35.42a)$$

$$\frac{\partial F_2}{\partial y} = -\frac{f F_2}{c} \quad (35.42b)$$

with solutions

$$F_1 = F(x + ct) e^{y/L_d} \quad (35.43a)$$

$$F_2 = G(x - ct) e^{-y/L_d} \quad (35.43b)$$

where $L_d = c/f$ is the shallow water deformation radius (equation (35.33)). To ensure boundedness in the region $y > 0$ where the fluid is assumed to exist, we drop the F_1 solution, thus leaving

$$u' = e^{-y/L_d} G(x - ct) \quad (35.44a)$$

$$v' = 0 \quad (35.44b)$$

$$\eta' = (H/g)^{1/2} e^{-y/L_d} G(x - ct). \quad (35.44c)$$

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 Appendix 1.4.1 for more on general solutions to the linear wave equation.

35.5 Geostrophic adjustment

The geostrophic balance presented in Sections 30.3 and 34.1 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 35.1. Consequently, the adjustment consists of linear inertia-gravity waves that maintain a locally static potential vorticity (Section 35.1.3). For brevity in notation, we here drop all primes on the linear fluctuating terms.

35.5.1 Posing the initial value problem

We solve for the $t > 0$ evolution of surface height and velocity by making use of the linearized equations

$$\frac{\partial \mathbf{u}}{\partial t} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (35.45)$$

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0 \quad (35.46)$$

$$\zeta - \frac{f\eta}{H} = q(x, y) \quad (35.47)$$

where we dropped the primes for brevity, and where $q(x, y)$ is the static potential vorticity determined by the initial conditions (Section 35.1.3). To illustrate the geostrophic adjustment in an analytically tractable manner, consider the following step initial conditions for the surface height

$$\eta(x, t = 0) = \begin{cases} +\eta_0 & x < 0 \\ -\eta_0 & x > 0, \end{cases} \quad (35.48)$$

which can be written

$$\eta(x, t = 0) = -\eta_0 \operatorname{sgn}(x). \quad (35.49)$$

The velocity is assumed to be zero initially

$$\mathbf{u}(x, y, t = 0) = 0. \quad (35.50)$$

Correspondingly, the initial relative vorticity vanishes so that the linearized potential vorticity is initialized as

$$q(x, y) = \frac{f\eta_0}{H} \operatorname{sgn}(x). \quad (35.51)$$

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.

35.5.2 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

$$\eta(x, t) = -\frac{\eta_0}{2} [\operatorname{sgn}(x + ct) + \operatorname{sgn}(x - ct)], \quad (35.52)$$

where $c = \sqrt{gH}$ is the speed for non-rotating gravity waves. The meridional velocity remains zero, whereas the zonal velocity

$$\frac{\partial u}{\partial t} = -g \frac{\partial \eta}{\partial x} \quad (35.53)$$

is given by

$$u(x, t) = \frac{g\eta_0}{2c} [\operatorname{sgn}(x + ct) - \operatorname{sgn}(x - ct)]. \quad (35.54)$$

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.

35.5.3 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

$$f\hat{z} \wedge \mathbf{u} = -g \nabla \eta \quad (35.55)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (35.56)$$

$$q = \frac{f\eta_0}{H} \operatorname{sgn}(x). \quad (35.57)$$

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

$$\psi = \frac{g\eta}{f}. \quad (35.58)$$

The equilibrium state is written in terms of the streamfunction according to

$$u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x} \quad (\nabla^2 - L_d^{-2}) \psi = q(x, y), \quad (35.59)$$

where we introduced the *deformation radius*

$$L_d = \frac{\sqrt{gH}}{f}. \quad (35.60)$$

We motivate the name “deformation radius” in the following.

The initial condition (35.49) 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

$$\frac{d^2\psi}{dx^2} - L_d^{-2} \psi = \frac{f\eta_0}{H} \operatorname{sgn}(x). \quad (35.61)$$

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

$$\frac{d^2\psi}{dx^2} - L_d^{-2} \psi = \frac{f\eta_0}{H}. \quad (35.62)$$

The particular solution is

$$\psi_p = -L_d^2 \frac{f\eta_0}{H} = -\frac{g\eta_0}{f} \quad (35.63)$$

and the homogeneous solution is

$$\psi_h = \frac{g\eta_0}{f} e^{-x/L_d} \quad (35.64)$$

so that

$$\psi = -\frac{g\eta_0}{f} \left[1 - e^{-x/L_d} \right]. \quad (35.65)$$

The $x < 0$ solution is found similarly, so that

$$\psi = \frac{g\eta_0}{f} \begin{cases} -\left(1 - e^{-x/L_d}\right) & x > 0 \\ \left(1 - e^{x/L_d}\right) & x < 0, \end{cases} \quad (35.66)$$

which means that the equilibrium surface height is

$$\eta = \eta_0 \begin{cases} -\left(1 - e^{-x/L_d}\right) & x > 0 \\ \left(1 - e^{x/L_d}\right) & x < 0, \end{cases} \quad (35.67)$$

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

$$u = 0 \quad v = -\frac{g\eta_0}{f L_d} e^{-|x|/L_d}. \quad (35.68)$$

The equilibrium velocity thus consists of a jet that is perpendicular to the surface height front.

35.5.4 Comments

As illustrated in Figure 35.2, 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.

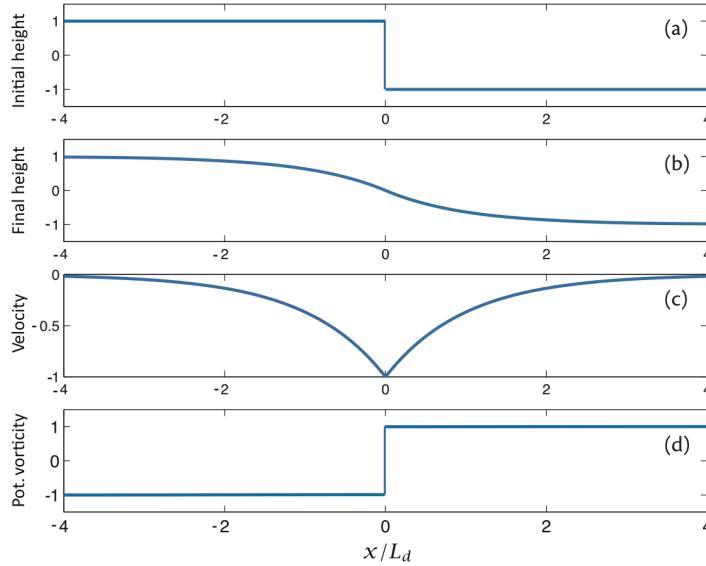


Figure 35.2: This is Figure 3.10 of [Vallis \(2017\)](#), illustrating solutions to the linear geostrophic adjustment of a rotating shallow water layer to a surface height perturbation. The top panel shows the initial surface height (35.49), and the second panel the equilibrium surface height (35.67). The third panel shows the equilibrium meridional velocity (35.68) comprised of a jet centered at $x = 0$. The final panel shows the static potential vorticity (35.57).

35.6 Exercises

EXERCISE 35.1: 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 35.2: 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

$$\frac{\partial \mathbf{u}}{\partial t} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (35.69a)$$

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0, \quad (35.69b)$$

where

$$h(x, y, t) = H + \eta(x, y, t) \quad (35.70)$$

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.

Part VIII

Vorticity and potential vorticity

Vorticity is a kinematic property of a fluid that locally measures the spin of a fluid element. Its kinematic and dynamic properties are fundamental to understanding fluid flow. Vorticity takes on an especially important role for geophysical fluid flows given the ubiquitous nature of vorticity induced by planetary rotation. Potential vorticity (PV) is a strategically chosen component of the vector vorticity, with PV of great use for understanding the rotating and stratified fluid found in geophysical flows.

We start this part of the book by introducing vorticity and potential vorticity for the shallow water model in Chapter 36. This material allows us to capture a flavor for vorticity and PV without the added mathematics needed for continuously stratified fluids. Chapter 37 then develops the fundamentals of vorticity and circulation, which are related through Stokes' Theorem. We here encounter the famous Kelvin Circulation Theorem, which identifies the materially conserved nature of circulation in perfect barotropic flows. Chapter 38 then presents the foundations of potential vorticity, with Appendix 39 then providing a suite of detailed examples.

This part of the book makes extensive use of the Cartesian tensor algebra and vector calculus detailed in Chapters 2 and 3.

36

Shallow water vorticity and potential vorticity

We here consider vorticity and potential vorticity within the shallow water system.

36.1	PV for a rotating deformable cylinder	539
36.1.1	Mass conservation	539
36.1.2	Angular momentum conservation	540
36.1.3	Potential vorticity conservation	540
36.1.4	Connecting angular momentum and vorticity	540
36.1.5	Comments and further reading	541
36.2	The vorticity equation for a shallow water layer	541
36.3	Potential vorticity for a shallow water layer	542
36.3.1	Algebraic derivation	542
36.3.2	Derivation based on Kelvin's circulation theorem	543
36.3.3	Material conservation of an arbitrary function of PV	544
36.3.4	Further reading	544
36.4	Exercises	544

36.1 PV for a rotating deformable cylinder

To introduce the notion of potential vorticity, consider a deformable cylinder of constant density ρ , radius R , and height H . Allow the cylinder to exhibit solid body rotation about the central axis of the cylinder. We might think of the cylinder as a rotating material region of a constant density fluid, in which case the time derivatives are material derivatives.

36.1.1 Mass conservation

Mass conservation is a kinematic property of the cylinder that relates the radius and thickness. With constant density, mass conservation for a material region is equivalent to volume conservation. A materially constant cylinder mass

$$M = \pi R^2 H \rho \quad (36.1)$$

implies

$$\frac{2}{R} \frac{DR}{Dt} = - \frac{1}{H} \frac{DH}{Dt}. \quad (36.2)$$

That is, mass conservation means that as twice the relative radius increases, the relative height decreases.

36.1.2 Angular momentum conservation

A second constraint arises from angular momentum conservation. Since the cylinder is moving 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

$$\boldsymbol{\Omega} = \Omega \hat{z}. \quad (36.3)$$

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)

$$I_{mn} = \delta_{mn} \frac{M R^2}{2}, \quad (36.4)$$

so that the angular momentum is

$$\mathbf{L} = \frac{M R^2}{2} \Omega \hat{z}. \quad (36.5)$$

The familiar “ice-skater” example occurs when the cylinder height is fixed and the radius changes (skater’s arms are brought in towards the body, or outward away from the body). To maintain constant angular momentum and constant mass, the angular velocity Ω increases (spins faster) when the radius decreases, and vice versa. Explicitly, we have $dL/dt = 0$ and $dM/dt = 0$ rendering

$$\frac{2}{R} \frac{dR}{dt} = -\frac{1}{\Omega} \frac{d\Omega}{dt}. \quad (36.6)$$

36.1.3 Potential vorticity conservation

Combining angular momentum conservation (36.6) with mass conservation (36.2) leads to

$$\frac{D}{Dt} \frac{\Omega}{H} = 0. \quad (36.7)$$

Equation (36.7) is a statement of potential vorticity conservation for the material fluid column, with potential vorticity given by

$$Q \equiv \frac{\Omega}{H}. \quad (36.8)$$

If the column thickens then the rotational velocity increases to maintain $Q = \Omega/H$ constant.

36.1.4 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 (36.8) is generalized to the absolute vorticity. We encounter this generalization in Section 36.2. Furthermore, as shown in Section 37.1.2, the vorticity for solid body motion equals to twice the rotation rate, Ω . Hence, the numerator for the potential vorticity of the solid-body rotating cylinder equals to one-half the vorticity. In Appendix 40, we connect angular momentum and vorticity for arbitrary fluid motion.

36.1.5 Comments and further reading

Much of the discussion in this section is motivated by a similar presentation given by [Salmon \(1998\)](#). The solid-body rotating cylinder illustrates a process recurring for more general fluids. Namely, it identifies the two mechanical properties contributing to the potential vorticity conservation law (36.7): 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.

36.2 The vorticity equation for a shallow water layer

The vorticity for the full flow and of the horizontal flow are written as

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad \boldsymbol{\omega}^* = \nabla \wedge \mathbf{u}. \quad (36.9)$$

The vector identity (see Section 3.3.4)

$$(\mathbf{u} \cdot \nabla) \mathbf{u} = (1/2) \nabla(\mathbf{u} \cdot \mathbf{u}) - \mathbf{u} \wedge (\nabla \wedge \mathbf{u}) \quad (36.10)$$

allows us to write the shallow water momentum equation (33.5)

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta \quad (36.11)$$

in the “vector-invariant” form

$$\frac{\partial \mathbf{u}}{\partial t} + \boldsymbol{\omega}_a^* \wedge \mathbf{u} = -\nabla(g \eta + \mathbf{u}^2/2) \quad (36.12)$$

where we introduced the absolute vorticity

$$\boldsymbol{\omega}_a^* = \boldsymbol{\omega}^* + f \hat{\mathbf{z}}. \quad (36.13)$$

We next make use of the vector identity (see again Section 3.3.4)

$$\nabla \wedge (\boldsymbol{\omega}_a^* \wedge \mathbf{u}) = \boldsymbol{\omega}_a^* (\nabla \cdot \mathbf{u}) - \mathbf{u} (\boldsymbol{\omega}_a^* \cdot \nabla) + (\mathbf{u} \cdot \nabla) \boldsymbol{\omega}_a^* - (\boldsymbol{\omega}_a^* \cdot \nabla) \mathbf{u} \quad (36.14a)$$

$$= \boldsymbol{\omega}_a^* (\nabla \cdot \mathbf{u}) + (\mathbf{u} \cdot \nabla) \boldsymbol{\omega}_a^*. \quad (36.14b)$$

For this result, we set

$$\nabla \cdot \boldsymbol{\omega}_a^* = \nabla \cdot \boldsymbol{\omega}^* + \nabla \cdot (f \hat{\mathbf{z}}) = 0 \quad (36.15)$$

since this expression involves the divergence of a curl (first right hand side term) and since f has no z dependence. We furthermore set

$$(\boldsymbol{\omega}_a^* \cdot \nabla) \mathbf{u} = 0 \quad (36.16)$$

since $\boldsymbol{\omega}_a^*$ is perpendicular to the surface in which \mathbf{u} lives. We now apply the operator $\hat{\mathbf{z}} \cdot (\nabla \wedge)$ onto the vector-invariant momentum equation (36.12) to yield a prognostic equation for the absolute vorticity

$$\frac{\partial \zeta_a}{\partial t} + \nabla \cdot (\mathbf{u} \zeta_a) = 0, \quad (36.17)$$

where

$$\zeta_a = \hat{z} \cdot (\boldsymbol{\omega}^* + \hat{z} f) = \hat{z} \cdot (\boldsymbol{\omega} + \hat{z} f) = \zeta + f \quad (36.18)$$

is the vertical component of the absolute vorticity, and

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad (36.19)$$

is the vertical component to the relative vorticity. We can write the Eulerian flux-form vorticity equation (36.17) in the material form

$$\frac{D\zeta_a}{Dt} = -\zeta_a \nabla \cdot \mathbf{u}, \quad (36.20)$$

where the material time derivative is determined by the horizontal flow

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla. \quad (36.21)$$

36.3 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 36.1. However, the derivation here follows more traditional fluid mechanical notions than used for the rotating cylinder. 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 37).

Figure 36.1 summarizes the key elements to PV conservation for a shallow water layer. Namely, shallow water PV conservation arises from combining mass conservation (material conservation of $h A$), and Kelvin's circulation theorem for a small aspect ratio fluid (material conservation of ζA). As noted in our discussion of the rotating cylinder in Section 36.1, PV conservation generally arises from the combination of a kinematic constraint (mass conservation) and a dynamical constraint.

36.3.1 Algebraic derivation

To derive the potential vorticity equation, we make use of the vorticity equation (36.20) and combine it with mass conservation. Mass conservation in the form of the material thickness equation (33.17) leads to the following expression for the divergence of the horizontal velocity

$$\nabla \cdot \mathbf{u} = -\frac{1}{h} \frac{Dh}{Dt}. \quad (36.22)$$

We make use of this result to eliminate the horizontal divergence in the vorticity equation (36.20), thus yielding

$$\frac{D\zeta_a}{Dt} = \frac{\zeta_a}{h} \frac{Dh}{Dt}. \quad (36.23)$$

This equation then leads to the material conservation law

$$\frac{DQ}{Dt} = 0, \quad (36.24)$$

where we introduced the shallow water potential vorticity

$$Q = \frac{\zeta_a}{h} = \frac{\zeta + f}{h}. \quad (36.25)$$

Potential vorticity (36.25) is the ratio of absolute vorticity to the thickness of the fluid layer. The material conservation law (36.24) says that this ratio remains constant for the shallow water layer.

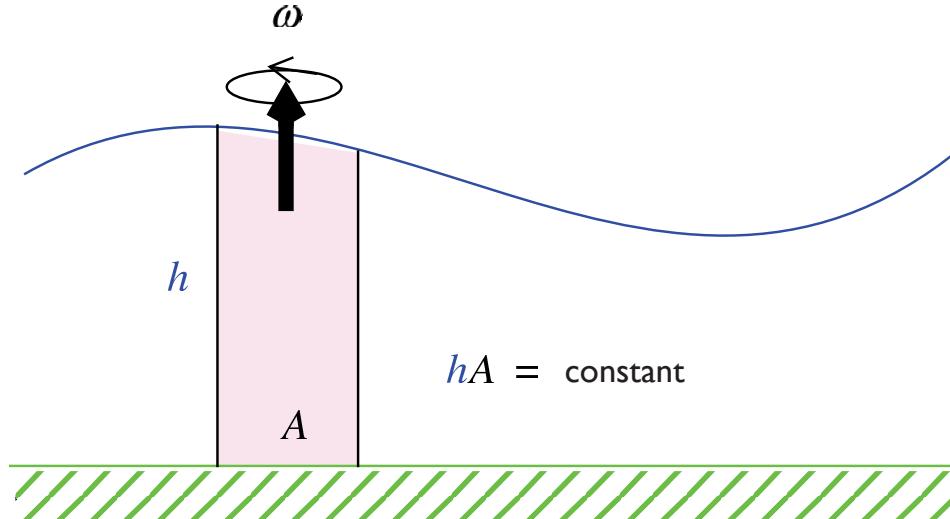


Fig. 4.10 The mass of a column of fluid, hA , is conserved in the shallow water system. Furthermore, the vorticity is tied to material lines so that ζA is also a material invariant, where $\zeta = \omega \cdot \mathbf{k}$ is the vertical component of the vorticity. From this, ζ/h must be materially conserved; that is, $D(\zeta/h)/Dt = 0$, which is the conservation of potential vorticity in a shallow water system. In a rotating system this generalizes to $D[(\zeta + f)/h]/Dt = 0$.

Figure 36.1: Illustrating the conservation of PV for a layer of shallow water fluid, as per Figure 4.10 of [Vallis \(2006\)](#). PV conservation is a merging of mass conservation, (material conservation of $h A$), and Kelvin's circulation theorem for a small aspect ratio fluid (material conservation of ζA).

36.3.2 Derivation based on Kelvin's circulation theorem

An ideal homogeneous (constant density) fluid layer satisfies Kelvin's circulation theorem (Section 37.6). Although we have yet to discuss Kelvin's Theorem (Section 37.6), we here make use of its result for the shallow water PV evolution. When applied to an infinitesimal circuit in the fluid, Kelvin's theorem says that

$$\frac{D(\omega_a \cdot \hat{\mathbf{n}} \delta S)}{Dt} = 0, \quad (36.26)$$

where ω_a is the absolute vorticity

$$\omega_a = \omega + f \hat{\mathbf{z}}, \quad (36.27)$$

$\hat{\mathbf{n}} \delta S$ is the infinitesimal surface area enclosed by the closed circuit, with $\hat{\mathbf{n}}$ the outward normal to the surface. For the shallow water system, we decompose absolute vorticity into

$$\omega_a = \hat{\mathbf{z}} (\zeta + f) + \omega_h, \quad (36.28)$$

where ζ is the vertical component to the relative vorticity (equation (36.19)), and

$$\omega_h = \hat{\mathbf{x}} \frac{\partial w}{\partial y} - \hat{\mathbf{y}} \frac{\partial w}{\partial x} \quad (36.29)$$

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 (36.28) into Kelvin's theorem (36.28) leads to

$$\frac{D}{Dt} [(\zeta + f) \delta A + \boldsymbol{\omega}_h \cdot \boldsymbol{\delta S}_h] = 0. \quad (36.30)$$

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 (36.30) is much smaller than the first, so that for the shallow water fluid we have

$$\frac{D}{Dt} \left[\left(\frac{\zeta + f}{h} \right) h \delta A \right] = 0, \quad (36.31)$$

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 conserved

$$\frac{D(h \delta A)}{Dt} = 0, \quad (36.32)$$

so that equation (36.31) yields the material conservation of shallow water PV

$$\frac{D}{Dt} \left[\frac{\zeta + f}{h} \right] = 0, \quad (36.33)$$

where $Q = (\zeta + f)/h$ is the shallow water potential vorticity (equation (36.25)).

36.3.3 Material conservation of an arbitrary function of PV

The material conservation law for PV, equation (36.24), means that any function, $F(Q)$ is also materially conserved. We see this property through the chain rule

$$\frac{DF}{Dt} = F'(Q) \left[\frac{DQ}{Dt} \right] = 0. \quad (36.34)$$

Since F is arbitrary, there are an infinite number of material invariants corresponding to distinct functions F .

36.3.4 Further reading

The material in this section is a summary of that in Chapter 3 of [Vallis \(2017\)](#).

36.4 Exercises

EXERCISE 36.1: 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

$$\frac{D}{Dt} \left[\frac{\zeta + f}{\eta - \eta_b} \right] = 0, \quad (36.35)$$

where η is the height of the free surface and η_b is the height of the bottom topography (see Figure 33.1). 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 36.2: 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 36.3: SHALLOW WATER PV WITH FRICTION

Consider the rotating shallow water equations with friction

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta + \mathbf{F} \quad \frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} = 0, \quad (36.36)$$

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 36.4: 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

$$h \mathbf{u} = \nabla \wedge (\hat{\mathbf{z}} \Psi). \quad (36.37)$$

- (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

$$Q = Q(\Psi). \quad (36.38)$$

- (d) Show that the Bernoulli function,

$$B = g \eta + \mathbf{u} \cdot \mathbf{u}/2 \quad (36.39)$$

is also a constant along the same streamlines; i.e.,

$$B = B(\Psi). \quad (36.40)$$

- (e) What is the functional relation between the Bernoulli function and the potential vorticity?

EXERCISE 36.5: 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 36.2.

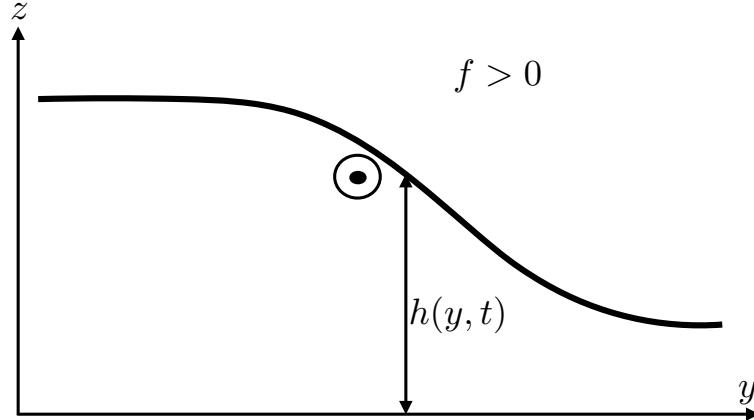


Figure 36.2: Schematic of a zonally symmetric front in a shallow water layer in the northern hemisphere ($f > 0$). The thickness decreases to the north. If the flow is in geostrophic balance, then the northward pressure gradient is in geostrophic balance with a southward Coriolis acceleration arising from an eastward (out of the page) geostrophic current (see also Figure 34.1).

- Write the potential vorticity, Q , assuming the fluid is in geostrophic balance. Write in terms of meridional derivatives of the layer thickness.
- 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.
- 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 ?
- 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.

EXERCISE 36.6: 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 30.3 or the 2d barotropic equation in Section 41.1). 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

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla(g\eta + \alpha\Gamma) \quad (36.41a)$$

$$\frac{Dh}{Dt} = -h\nabla \cdot \mathbf{u}. \quad (36.41b)$$

The parameter $\alpha > 0$ is a constant and the field Γ is given by the Laplacian of the horizontal flow divergence

$$\Gamma = \nabla^2 \mathcal{D}, \quad (36.42)$$

where

$$\mathcal{D} = \nabla \cdot \mathbf{u}. \quad (36.43)$$

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

$$\tilde{\eta} = \eta + \frac{\alpha\Gamma}{g}. \quad (36.44)$$

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 (36.20) remains unchanged in the presence of divergence-damping.
- (b) Show that the potential vorticity equation (36.24) remains unchanged in the presence of divergence-damping.
- (c) Show that the horizontal divergence evolves according to

$$\frac{\partial \mathcal{D}}{\partial t} = \left[\frac{\partial \mathcal{D}}{\partial t} \right]_{\alpha=0} - \alpha \nabla^2 \Gamma. \quad (36.45)$$

- (d) Show that the evolution of gravitational potential energy per horizontal area

$$\mathcal{P} = g\rho \int_{\eta_b}^{\eta} z dz \quad (36.46)$$

remains unchanged from that determined in Section 34.3.1.

- (e) Show that the kinetic energy per horizontal area evolves according to

$$\frac{\partial \mathcal{K}}{\partial t} + \nabla \cdot (\mathbf{u} \mathcal{K}) = -h\rho g \nabla \tilde{\eta}, \quad (36.47)$$

where

$$\mathcal{K} = \frac{1}{2} \int_{\eta_b}^{\eta} \rho \mathbf{u}^2 dz = \rho h \mathbf{u}^2 / 2, \quad (36.48)$$

is the horizontal kinetic energy per area (Section 34.3.2).

- (f) Determine the evolution equation for global integrated kinetic energy

$$\frac{\partial}{\partial t} \left[\int \mathcal{K} dA \right] = \frac{\partial}{\partial t} \left[\int \int_{\eta_b}^{\eta} (\rho \mathbf{u}^2 / 2) dz dA \right]. \quad (36.49)$$

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 34.5. Show that the material evolution of angular momentum relative to the vertical rotational axis is given by

$$\frac{1}{\delta M} \frac{D L^z}{Dt} = -g \frac{\partial \eta}{\partial \phi} + \mathcal{T}. \quad (36.50)$$

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

$$\frac{\partial}{\partial t} \int L^z = \alpha \rho \int \Gamma \frac{\partial \eta}{\partial \phi} dA. \quad (36.51)$$

where we assume the bottom topography is flat so that $h = \eta$.

- (i) The linearized thickness equation (see Section 35.3) for a flat bottom is given by

$$\frac{\partial \eta}{\partial t} + H \nabla \cdot \mathbf{u} = 0, \quad (36.52)$$

where H is the thickness of the resting fluid layer. Show that the time change for the global integrated angular momentum is given by

$$\frac{\partial}{\partial t} \int L^z = -\frac{\alpha \rho}{H} \int \left[\frac{\partial}{\partial t} \nabla^2 \eta \right] \frac{\partial \eta}{\partial \phi} dA. \quad (36.53)$$

37

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 notions of vorticity and circulation and develop results based on Kelvin's circulation theorem.

- Discuss Figure 14.1 from *Thorne and Blandford* (2017), who show three example flow fields and their vorticity; very illustrative.
- Lie derivative as per section 14.2.1 of *Thorne and Blandford* (2017).

37.1	Introducing vorticity	550
37.1.1	Irrotational flow	550
37.1.2	Vorticity for solid-body rotation	550
37.2	Circulation	551
37.2.1	Connecting circulation and vorticity	551
37.3	Vorticity and circulation for the free vortex	551
37.3.1	Motion of a fluid particle	552
37.3.2	Circulation	552
37.3.3	Further reading	552
37.4	Uniform translation and solid-body rotation	553
37.4.1	Rigid or solid-body motion	553
37.4.2	Example rigid body motion	554
37.4.3	Circulation for rigid body motion	554
37.5	Kinematics of vortex lines and vortex tubes	555
37.5.1	Defining vortex lines and vortex tubes	555
37.5.2	Kinematic properties of vortex lines and tubes	555
37.5.3	Helmholz's first theorem	556
37.5.4	Helmholz's second theorem	556
37.5.5	Helmholz's third theorem	556
37.5.6	Further reading	556
37.6	Kelvin's Circulation Theorem for non-rotating flow	557
37.7	Vorticity dynamics	558
37.7.1	Vector-invariant velocity equation	558
37.7.2	Basic form of the vorticity equation	558
37.7.3	Massaged form of the vorticity equation	559
37.8	Vorticity and circulation for a non-rotating barotropic fluid	560

37.8.1	Frozen-in nature of vorticity	560
37.8.2	Stretching and tilting of vortex tubes	560
37.9	Circulation and vorticity for rotating fluids	562
37.9.1	Circulation and vorticity	563
37.9.2	Material evolution of absolute circulation	563
37.9.3	Relative circulation induced by planetary circulation	564
37.9.4	A two-dimensional fluid example	566
37.10	Exercises	566

37.1 Introducing vorticity

Vorticity is the curl of the velocity field

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v}. \quad (37.1)$$

Vorticity measures the rotational or spin aspects 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, vorticity and angular momentum are generally not the same, instead offering complementary measures to characterize rotating fluid motion.

37.1.1 Irrotational flow

Irrotational fluid flow is characterized by a zero vorticity

$$\boldsymbol{\omega} = 0 = \text{irrotational flow}. \quad (37.2)$$

Since the curl of a gradient vanishes, irrotational flow has a velocity field given by the gradient of a velocity potential

$$\nabla \wedge \mathbf{v} = 0 \Rightarrow \mathbf{v} = \nabla \Psi. \quad (37.3)$$

Furthermore, if the flow is incompressible, as in a Boussinesq fluid (Section 29.1), then the velocity potential is a harmonic function

$$\nabla \cdot \mathbf{v} = 0 \Rightarrow \nabla^2 \Psi = 0. \quad (37.4)$$

Most geophysical flows have nonzero vorticity, though there are some examples, such as non-rotating linear gravity waves, with vanishing vorticity (Section 35.2.2).

37.1.2 Vorticity for solid-body rotation

Consider a fluid in solid-body rotational motion with angular velocity

$$\mathbf{v}_{\text{solid-body}} = \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (37.5)$$

where \mathbf{x} is the position vector of a fluid particle. As shown in Figure 3.4 and in Exercise 37.1, this solid-body motion has a vorticity given by

$$\nabla \wedge \mathbf{v}_{\text{solid-body}} = 2 \boldsymbol{\Omega}. \quad (37.6)$$

For planetary fluid dynamics, rotation of the planet imparts *planetary vorticity* to fluids. The total vorticity is the vector sum of the *relative vorticity*, $\boldsymbol{\omega}$, and the planetary vorticity

$$\boldsymbol{\omega}_{\text{total}} = \mathbf{v}_{\text{solid-body}} + \boldsymbol{\omega}, \quad (37.7)$$

where

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad (37.8)$$

is the vorticity of the fluid due to motion relative to the rotating sphere, with \mathbf{v} the velocity field with respect to the rotating sphere.

37.2 Circulation

Circulation is the oriented closed loop integral of velocity as projected onto the path of the loop

$$C \equiv \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r}. \quad (37.9)$$

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 3.4). The difference between two very close positions along the circuit defines the increment

$$d\mathbf{r} = \mathbf{r}(\varphi + \delta\varphi) - \mathbf{r}(\varphi). \quad (37.10)$$

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.

37.2.1 Connecting circulation and vorticity

Stokes' Theorem (Section 3.6) renders the identity

$$C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int_S (\nabla \wedge \mathbf{v}) \cdot \hat{\mathbf{n}} dS = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (37.11)$$

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 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.

37.3 Vorticity and circulation for the free vortex

Consider a two-dimensional rotating fluid in the $x - y$ plane with angular velocity given by

$$\boldsymbol{\Omega} = \frac{\mathbf{x} \wedge \mathbf{v}}{r^2} = \frac{K \hat{\mathbf{z}}}{r^2}. \quad (37.12)$$

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*.

37.3.1 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

$$\mathbf{v} \cdot \mathbf{x} = 0. \quad (37.13)$$

The velocity for this pure rotational flow is given by (see Section 10.5)

$$\mathbf{v} = \boldsymbol{\Omega} \wedge \mathbf{x} = \frac{K(-y \hat{\mathbf{x}} + x \hat{\mathbf{y}})}{r^2} = \frac{K \hat{\boldsymbol{\theta}}}{r}, \quad (37.14)$$

where $\hat{\boldsymbol{\theta}}$ is the polar angle unit vector oriented in the counter-clockwise direction (see Appendix 7.3). Each component of the velocity falls off as $1/r$ when moving away from the origin. Away from the origin, the vorticity vector vanishes

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} = 0, \quad (37.15)$$

whereas it is singular at the origin.

We illustrate this velocity field in Figure 37.1, where snapshots are taken of two fluid parcels moving around the center. The parcels are marked with a line to distinguish their orientations. One can imagine these marked parcels as two Saturn-like planets orbiting around the center, or as paddle-wheels placed in the flow. As the parcels move counter-clockwise, the marked lines on the parcels remain oriented at the same angle. That is, the parcels orbit but they do not spin. Likewise, paddle wheels placed in this flow will not spin. This example thus illustrates a flow with non-zero angular momentum but with zero vorticity, with the lack of vorticity seen due to the absence of spin on the “planets.”

Although vorticity is zero away from the origin, the angular momentum is nonzero, as expected since the fluid is rotating around the center. The angular momentum for this system arises just from the strain in the fluid (see Section 17.2.6). The presence of strain is apparent since the two sample parcels move relative to one another, thus manifesting the nonzero strain field.

37.3.2 Circulation

The 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

$$C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int_0^{2\pi} \mathbf{v} \cdot \hat{\boldsymbol{\theta}} r d\theta = 2\pi K. \quad (37.16)$$

To reach this result, we set the line element to

$$d\mathbf{r} = \hat{\boldsymbol{\theta}} r d\theta \quad (37.17)$$

and inserted the velocity (37.14) represented in cylindrical polar coordinates

$$\mathbf{v} \cdot \hat{\boldsymbol{\theta}} = \frac{K}{r}. \quad (37.18)$$

37.3.3 Further reading

This material is based on Section 4.1.2 of [Vallis \(2017\)](#).

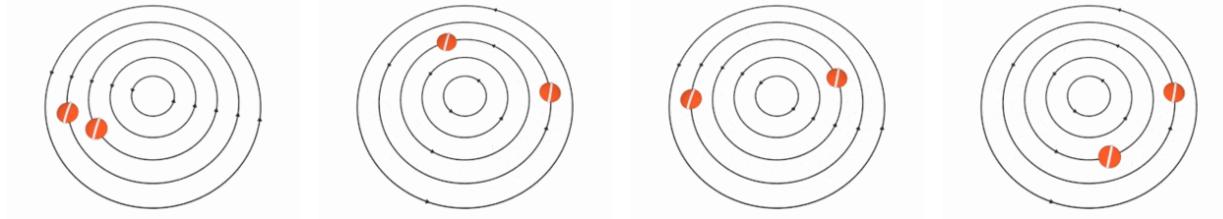


Figure 37.1: Irrotational flow in the presence of a free vortex. For all points except the origin, the vorticity vanishes. Note how the orientation remains fixed for the two illustrative parcels as they orbit around the center. Also note that the growing separation of the particles results from the non-zero strain in the fluid. These images were taken from the Wikipedia entry on “vortex”.

37.4 Uniform translation and solid-body rotation

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 resting in the rotating frame (see Exercise 37.1).

37.4.1 Rigid or solid-body motion

Rigid or solid-body fluid motion occurs when all fluid parcels are rigidly locked into their relative positions, as if in a rigid solid body. There are two kinds of rigid body motions: uniform translation and uniform rotation. The velocity field for this motion is given by

$$\mathbf{v} = \mathbf{U} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (37.19)$$

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 dependent. The rate of strain tensor vanishes for uniform translation or solid-body motion (see Exercise 17.1)

$$S_{mn} = \frac{1}{2}(\partial_m v_n + \partial_n v_m) = 0. \quad (37.20)$$

However, the vorticity field is nonzero (see Exercise 37.1)

$$\boldsymbol{\omega} = \nabla \wedge (\boldsymbol{\Omega} \wedge \mathbf{x}) = 2\boldsymbol{\Omega}. \quad (37.21)$$

The factor of two in this equation is geometric; it needs no physical explanation.

37.4.2 Example rigid body motion

Consider the circular rigid body rotation on a plane shown in Figure 37.2, in which the velocity field is given by

$$\mathbf{v} = \boldsymbol{\Omega} \wedge \mathbf{x} \quad (37.22a)$$

$$= |\boldsymbol{\Omega}| (-y \hat{\mathbf{x}} + x \hat{\mathbf{y}}) \quad (37.22b)$$

$$= |\boldsymbol{\Omega}| r \hat{\theta}. \quad (37.22c)$$

Assuming the center of mass to be at the circle center, the angular momentum for the flow is determined by that of a solid-body. Furthermore, the nonzero vorticity is illustrated by rotation of the lines marking the orientation of the parcels. The fluid parcels are rigidly locked into their relative position. But as they orbit the center, they spin on their respective axes, thus manifesting the nonzero vorticity.

37.4.3 Circulation for rigid body motion

For rigid body motion, the circulation around a circular path of radius R is given by

$$C = \oint \mathbf{v} \cdot d\mathbf{r} \quad (37.23)$$

$$= \oint (\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{r} \quad (37.24)$$

$$= R^2 |\boldsymbol{\Omega}| \oint d\theta \quad (37.25)$$

$$= 2\pi R^2 |\boldsymbol{\Omega}| \quad (37.26)$$

$$= 2A |\boldsymbol{\Omega}|, \quad (37.27)$$

where $A = \pi R^2$ is the area of the circle. Hence, the area averaged circulation is twice the angular rotation rate, which is the magnitude of the vorticity

$$C/A = |\boldsymbol{\omega}| = 2 |\boldsymbol{\Omega}|. \quad (37.28)$$

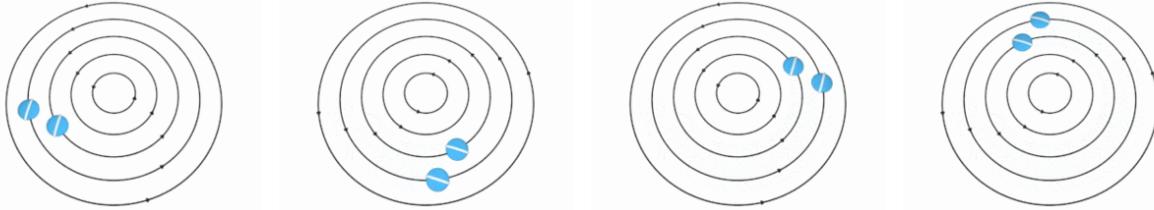


Figure 37.2: Rigid body rotation, where the vorticity is $\boldsymbol{\omega} = 2\boldsymbol{\Omega}$. Note how the two parcels maintain a constant relative distance, and that the marked lines on the parcels spin about the axis of a parcel. These images were taken from the Wikipedia entry on “vortex”.

37.5 Kinematics of vortex lines and vortex tubes

We now establish some kinematic results to help in understanding vorticity and its evolution, following from the kinematics of material line elements discussed in Section 17.2.3. We do not invoke dynamical principles here, instead focusing on kinematics.

37.5.1 Defining 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)

$$\mathbf{x}(\varphi) = x(\varphi) \hat{\mathbf{x}} + y(\varphi) \hat{\mathbf{y}} + z(\varphi) \hat{\mathbf{z}}. \quad (37.29)$$

The three coordinates of the line are constrained so that the line is tangent to vorticity at each point, which means

$$\frac{dx/d\varphi}{\omega_x} = \frac{dy/d\varphi}{\omega_y} = \frac{dz/d\varphi}{\omega_z}. \quad (37.30)$$

These equations are directly analogous to those used to compute velocity streamlines (Section 14.7.2)

$$\frac{dx/d\varsigma}{u} = \frac{dy/d\varsigma}{v} = \frac{dz/d\varsigma}{w}, \quad (37.31)$$

where ς is the parameter along the streamline. As 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 14.7). 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. 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 14.5.

37.5.2 Kinematic properties of vortex lines and tubes

Vorticity has zero divergence

$$\nabla \cdot \boldsymbol{\omega} = \nabla \cdot (\nabla \wedge \mathbf{v}) = 0, \quad (37.32)$$

which follows since the divergence of a curl vanishes. Integrating the non-divergence relation over an arbitrary closed volume within the fluid leads to

$$\int_V \nabla \cdot \boldsymbol{\omega} dV = \int_{\partial V} \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0, \quad (37.33)$$

where we made use of Gauss's Law to reach the surface integral expression, with $\hat{\mathbf{n}}$ the outward normal on the boundary ∂V . 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 (37.33) to be taken over a volume along a chosen vortex tube (sometimes called a vortex filament), bounded by a top and bottom cross-section

through the tube.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

$$\int_A \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS_A + \int_B \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS_B = 0. \quad (37.34)$$

Use of Stoke's Theorem transfers the above vorticity constraint to a constraint on the circulation around the circumference of the tube, so that

$$\oint_A \mathbf{v} \cdot d\mathbf{r} + \oint_B \mathbf{v} \cdot d\mathbf{r} = 0. \quad (37.35)$$

The constraints (37.34) and (37.35) are kinematic, so they hold for any vorticity field at all times. We now consider some consequences of this result.

37.5.3 Helmholtz's first theorem

Since the cross-sectional slices used to derive the circulation constraint (37.35) are arbitrary, the constraint holds throughout the full extent of the vortex tube. Hence, 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. This result is known as Helmholtz's first theorem.

As a corollary, we refer to the vorticity constraint (37.34) 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 must increase where the area decreases, which in turn means that the linear velocity increases as the area reduces. Think of a tornado, which is a natural expression of a vortex tube. Near the ground, the cross-sectional area of the tornado is small, with the linear velocity of a point within the tube relatively large. Near the tornado top, the cross-sectional area is large so the linear velocity is relatively small.

37.5.4 Helmholtz's second theorem

The vorticity constraint (37.34) 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), or intersect a boundary (as for a tornado).

37.5.5 Helmholtz's third theorem

Helmholtz's third theorem states that an unforced inviscid barotropic fluid (equation (37.42)) that has zero vorticity will remain irrotational forever. This theorem is a special case of Kelvin's Circulation Theorem, which is discussed in Section 37.6.

37.5.6 Further reading

A particularly insightful discussion of these ideas can be found in Chapter 5 of ?. Additionally, the [Physics Girl video](#) offers a vivid illustration of Helmholtz's Theorems in a swimming pool.

¹In Exercise 16.9, we developed a similar set of results for a streamtube in an incompressible fluid.

37.6 Kelvin's Circulation Theorem for non-rotating flow

Kelvin's Circulation Theorem is concerned with the evolution of circulation around a closed material circuit, or equivalently (through Stokes' Theorem) with the change in vorticity penetrating the enclosed area

$$\frac{DC}{Dt} = \frac{D}{Dt} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \frac{D}{Dt} \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS. \quad (37.36)$$

The material time derivative, D/Dt , is appropriate since we are asking how circulation evolves following the fluid flow. We here consider the case of non-rotating flow, with rotating flow considered in Section 37.9. The result found here for the non-rotating flow is equivalent to the absolute circulation (relative circulation plus circulation due to planetary rotation) found in the rotating case.

Formulation of the theorem

The material time derivative in equation (37.36) commutes with the material line integral, so that

$$\frac{DC}{Dt} = \frac{D}{Dt} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \oint_{\partial S} \frac{D(\mathbf{v} \cdot d\mathbf{r})}{Dt}. \quad (37.37)$$

The material evolution of \mathbf{v} is determined by Newton's Law of motion, which for a non-rotating flow is given by

$$\frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F}. \quad (37.38)$$

The material time derivative of the differential line element moving around the circuit equals to the differential of the velocity on the circuit

$$\frac{D(d\mathbf{r})}{Dt} = d\mathbf{v}. \quad (37.39)$$

Consequently, the material evolution of circulation becomes

$$\frac{DC}{Dt} = \frac{D}{Dt} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} \quad (37.40a)$$

$$= \oint_{\partial S} \left[\left(-\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F} \right) \cdot d\mathbf{r} + \frac{1}{2} d\mathbf{v}^2 \right] \quad (37.40b)$$

$$= \oint_{\partial S} \left[-\frac{1}{\rho} \nabla p + \mathbf{F} \right] \cdot d\mathbf{r} \quad (37.40c)$$

$$= \int_S \left[-\nabla \wedge \left(\frac{1}{\rho} \nabla p \right) + \nabla \wedge \mathbf{F} \right] \cdot \hat{\mathbf{n}} dS \quad (37.40d)$$

$$= \int_S (\mathbf{S}_o + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad (37.40e)$$

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

$$\mathbf{S}_o = \frac{\nabla \rho \wedge \nabla p}{\rho^2}. \quad (37.41)$$

Special case of barotropic flow

The solenoidal vector vanishes for a constant density fluid, so that $\nabla\rho = 0$ such as for a single layer of shallow water fluid. More generally, the solenoidal vector vanishes for barotropic flow, in which

$$p = p(\rho) \Rightarrow \text{barotropic flow.} \quad (37.42)$$

Kelvin's theorem then follows, which states that for inviscid barotropic flow, the circulation around any closed material circuit in the fluid remains constant

$$\frac{DC}{Dt} = \frac{D}{Dt} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \frac{D}{Dt} \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS = 0. \quad (37.43)$$

This is a rather remarkable result that greatly constrains the motion of a barotropic perfect fluid.

37.7 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 from (see Section 26.1.6)

$$\rho \left[\frac{D}{Dt} + 2\boldsymbol{\Omega} \wedge \right] \mathbf{v} = -\nabla p - \rho \nabla \Phi + \rho \mathbf{F}. \quad (37.44)$$

In this equation, Φ is the geopotential and/or the potential for any conservative force, and \mathbf{F} arises from non-conservative viscous stresses.

37.7.1 Vector-invariant velocity equation

We now convert the advective-form momentum equation to vector-invariant form by making use of the vector identity (see Section 3.3.4)

$$\boldsymbol{\omega} \wedge \mathbf{v} = -(1/2) \nabla(\mathbf{v} \cdot \mathbf{v}) + (\mathbf{v} \cdot \nabla) \mathbf{v}. \quad (37.45)$$

This identity allows us to eliminate velocity self-advection in favor of vorticity

$$\frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega}_a \wedge \mathbf{v} = -\frac{1}{\rho} \nabla p - \nabla \left[\frac{1}{2} \mathbf{v}^2 + \Phi \right] + \mathbf{F}, \quad (37.46)$$

where

$$\boldsymbol{\omega}_a = \boldsymbol{\omega} + 2\boldsymbol{\Omega} \quad (37.47)$$

is the absolute vorticity.

37.7.2 Basic form of the vorticity equation

Taking the curl of the vector-invariant momentum equation (37.46) removes the mechanical energy per mass, $\mathbf{v}^2/2 + \Phi$, thus leaving

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v}) = \frac{1}{\rho^2} (\nabla \rho \wedge \nabla p) + \nabla \wedge \mathbf{F}. \quad (37.48)$$

Since $\boldsymbol{\Omega}$ has zero time tendency, we have

$$\frac{\partial \boldsymbol{\omega}_a}{\partial t} = \frac{\partial (\boldsymbol{\omega} + 2\boldsymbol{\Omega})}{\partial t} = \frac{\partial \boldsymbol{\omega}}{\partial t}, \quad (37.49)$$

so that equation (37.48) can be written

$$\frac{\partial \boldsymbol{\omega}_a}{\partial t} + \nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v}) = \frac{1}{\rho^2}(\nabla \rho \wedge \nabla p) + \nabla \wedge \mathbf{F}. \quad (37.50)$$

37.7.3 Massaged form of the vorticity equation

Physical interpretation of the term $\nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v})$ appearing in the prognostic equation (37.50) can be made more transparent by using yet another vector identity

$$\nabla \wedge (\boldsymbol{\omega}_a \wedge \mathbf{v}) = (\mathbf{v} \cdot \nabla)\boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \nabla)\mathbf{v} + \boldsymbol{\omega}_a \nabla \cdot \mathbf{v} - \mathbf{v} \nabla \cdot \boldsymbol{\omega}_a \quad (37.51a)$$

$$= (\mathbf{v} \cdot \nabla)\boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \nabla)\mathbf{v} - \frac{\boldsymbol{\omega}_a}{\rho} \frac{D\rho}{Dt}. \quad (37.51b)$$

The second equality required the continuity equation

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}, \quad (37.52)$$

and the non-divergent nature of the absolute vorticity

$$\nabla \cdot \boldsymbol{\omega}_a = \nabla \cdot (\nabla \wedge \mathbf{v} + 2\boldsymbol{\Omega}) = 0. \quad (37.53)$$

Equation (37.50) thus takes the form

$$\frac{D\boldsymbol{\omega}_a}{Dt} - \frac{\boldsymbol{\omega}_a}{\rho} \frac{D\rho}{Dt} = (\boldsymbol{\omega}_a \cdot \nabla)\mathbf{v} + \frac{1}{\rho^2}(\nabla \rho \wedge \nabla p) + \nabla \wedge \mathbf{F}, \quad (37.54)$$

which can be written

$$\frac{D\tilde{\boldsymbol{\omega}}_a}{Dt} = (\tilde{\boldsymbol{\omega}}_a \cdot \nabla)\mathbf{v} + \frac{1}{\rho^3}(\nabla \rho \wedge \nabla p) + \frac{1}{\rho} \nabla \wedge \mathbf{F}, \quad (37.55)$$

where

$$\tilde{\boldsymbol{\omega}}_a = \frac{\boldsymbol{\omega}_a}{\rho}. \quad (37.56)$$

Equation (37.55) 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, $(\tilde{\boldsymbol{\omega}}_a \cdot \nabla)\mathbf{v}$, will be explored in Section 37.8 in the simplified context of a barotropic fluid. The second term is known as the baroclinicity, and is commonly written in terms of the solenoidal vector given by

$$\mathbf{S}_o = \frac{\nabla \rho \wedge \nabla p}{\rho^2} = -\nabla \alpha \wedge \nabla p, \quad (37.57)$$

with

$$\alpha = \rho^{-1} \quad (37.58)$$

the specific volume. A solenoid is a tube perpendicular to both $\nabla \alpha$ and ∇p . Solenoids introduce a torque at a point in the fluid that affects vorticity. Solenoids vanish in the special case of $p = p(\rho)$, which characterizes barotropic flow (see equation (37.42)).

37.8 Vorticity and circulation for a non-rotating barotropic fluid

To help develop intuition for the term $(\tilde{\omega}_a \cdot \nabla) \mathbf{v}$ appearing in the vorticity equation (37.55), consider the special case of an incompressible, inviscid, non-rotating barotropic fluid, in which case the vorticity equation (37.55) reduces to

$$\frac{D\boldsymbol{\omega}}{Dt} = (\boldsymbol{\omega} \cdot \nabla) \mathbf{v}. \quad (37.59)$$

37.8.1 Frozen-in nature of vorticity

Recall from Section 37.5.1 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

$$\delta\mathbf{x}(0) = A \boldsymbol{\omega}(\mathbf{x}, 0), \quad (37.60)$$

where A is a dimensionful scalar 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 (37.59) has precisely the same mathematical form as the material line element equation (17.19)

$$\frac{D(\delta\mathbf{x})}{Dt} = (\delta\mathbf{x} \cdot \nabla) \mathbf{v}. \quad (37.61)$$

Consequently, the relation (37.60) holds for all time with A 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 $\delta\mathbf{x} = A \boldsymbol{\omega}$. In this way, the vorticity is attached to the particular material line element, lending us to say that vorticity is a “frozen-in” property. Importantly, this result holds only for the restricted case of an incompressible, inviscid, barotropic fluid. Nonetheless, it offers great insight into the more general situation occurring in real fluids.

37.8.2 Stretching and tilting of vortex tubes

When vortex lines, which are also material lines, are stretched or bent, then vorticity responds. To help understand the response, consider again the barotropic incompressible vorticity equation (37.59)

$$\frac{D\boldsymbol{\omega}}{Dt} = (\boldsymbol{\omega} \cdot \nabla) \mathbf{v} \quad (37.62)$$

and focus just the vertical vorticity component

$$\frac{D\omega^z}{Dt} = \omega^x \left(\frac{\partial w}{\partial x} \right) + \omega^y \left(\frac{\partial w}{\partial y} \right) + \omega^z \left(\frac{\partial w}{\partial z} \right). \quad (37.63)$$

Note that since $\nabla \cdot \boldsymbol{\omega} = 0$, we can write this equation as

$$\frac{D\omega^z}{Dt} = \nabla \cdot (w \boldsymbol{\omega}), \quad (37.64)$$

though in the following we find it more useful to focus on the form given by equation (37.63). The following discussion emulates that given for a material line element in Section 17.2.5.

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

$$\frac{D\omega^z}{Dt} = \omega^z \frac{\partial w}{\partial z}. \quad (37.65)$$

Since the fluid is incompressible, the volume of an infinitesimal portion of the vortex tube is materially constant

$$\frac{D(\delta V)}{Dt} = 0, \quad (37.66)$$

which means that the vertical extent, δz , and cross-sectional area, δA , are constrained

$$\frac{1}{\delta z} \frac{D(\delta z)}{Dt} + \frac{1}{\delta A} \frac{D(\delta A)}{Dt} = 0. \quad (37.67)$$

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 (37.61)) allows us to write

$$\frac{1}{\delta z} \frac{D(\delta z)}{Dt} = \frac{\partial w}{\partial z}, \quad (37.68)$$

so that the vorticity equation (37.65) becomes

$$\frac{D\omega^z}{Dt} = \omega^z \frac{\partial w}{\partial z} \quad (37.69a)$$

$$= \omega^z \left[\frac{1}{\delta z} \frac{D(\delta z)}{Dt} \right] \quad (37.69b)$$

$$= -\omega^z \left[\frac{1}{\delta A} \frac{D(\delta A)}{Dt} \right]. \quad (37.69c)$$

Rearrangement leads to

$$\frac{D(\omega^z \delta A)}{Dt} = 0, \quad (37.70)$$

which is an expression of Kelvin's circulation theorem (equation (37.43)) 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 (37.63). Namely, as the vortex tube is stretched and its cross-sectional area is compressed, the vorticity increases so to maintain a constant circulation around the tube, as per Kelvin's theorem. Stretching a column increases vorticity in the direction of the stretching, and compressing a column reduces vorticity.

Tilting

Now consider an initially horizontal vortex tube, where $\omega^z = 0$. Focus on just one of the two horizontal directions, so that equation (37.63) for the vertical vorticity becomes

$$\frac{D\omega^z}{Dt} = \omega^x \frac{\partial w}{\partial x}. \quad (37.71)$$

²Be mindful to distinguish the symbols for the vertical component of vorticity, ω^z , and the vertical component of velocity, w .

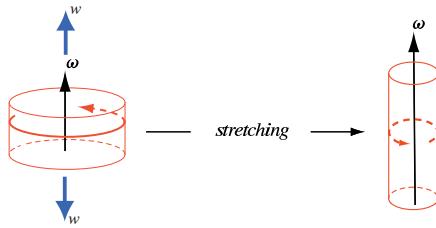


Figure 37.3: Illustrating the stretching process as per Figure 4.5 of [Vallis \(2017\)](#). As the cross-sectional area of the vortex tube shrinks, and the vertical extent of the tube stretches, the vorticity along the axis of the tube increases according to Kelvin's Circulation Theorem.

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 (37.61)) is given by

$$\frac{D(\delta\mathbf{x})}{Dt} = \delta x \frac{\partial \mathbf{v}}{\partial x}. \quad (37.72)$$

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.

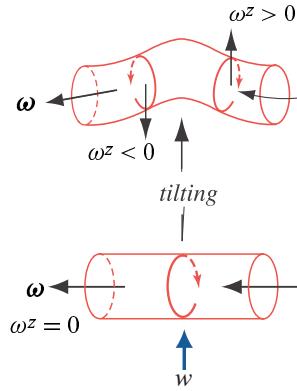


Figure 37.4: Illustrating the tilting process as per Figure 4.4 of [Vallis \(2017\)](#). The initial vortex tube is assumed to be aligned parallel to the x -axis, so that it has zero projection in the vertical direction. A horizontal shear of the vertical velocity (i.e., $\partial w / \partial x \neq 0$) will deform the vortex tube. Upon deforming, the tube picks up a nonzero projection in the vertical, which means that it now has a nonzero vertical component to vorticity.

37.9 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.

37.9.1 Circulation and vorticity

The inertial or absolute velocity (i.e., velocity measured in an inertial frame) is given by (Section 10.9.1)

$$\mathbf{v}_a = \mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (37.73)$$

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). We thus compute the absolute circulation by

$$C_a = \oint_{\partial S} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{x}) \cdot d\mathbf{r} \quad (37.74a)$$

$$= C + C_{\text{planet}}, \quad (37.74b)$$

where the circulation measured in the rotating frame is

$$C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} \quad (37.75)$$

and the circulation associated with the rotating planet is

$$C_{\text{planet}} = \oint_{\partial S} (\boldsymbol{\Omega} \wedge \mathbf{x}) \cdot d\mathbf{r}. \quad (37.76)$$

Again, $d\mathbf{r}$ is the differential line element moving around the circuit as mentioned in Section 37.2.

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

$$C = \oint_{\partial S} \mathbf{v} \cdot d\mathbf{r} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS \quad \text{relative circulation} \quad (37.77a)$$

$$C_{\text{planet}} = \oint_{\partial S} (\boldsymbol{\Omega} \wedge \mathbf{x}) \cdot d\mathbf{r} = \int_S \boldsymbol{\omega}_{\text{planet}} \cdot \hat{\mathbf{n}} dS \quad \text{planetary circulation} \quad (37.77b)$$

$$C_a = \oint_{\partial S} \mathbf{v}_a \cdot d\mathbf{r} = \int_S \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} dS \quad \text{absolute circulation,} \quad (37.77c)$$

where

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} \quad \text{relative vorticity} \quad (37.78a)$$

$$\boldsymbol{\omega}_{\text{planet}} = \nabla \wedge (\boldsymbol{\Omega} \wedge \mathbf{x}) = 2\boldsymbol{\Omega} \quad \text{planetary vorticity} \quad (37.78b)$$

$$\boldsymbol{\omega}_a = \boldsymbol{\omega} + \boldsymbol{\omega}_{\text{planet}} \quad \text{absolute vorticity.} \quad (37.78c)$$

37.9.2 Material evolution of absolute circulation

We are interested in how the absolute circulation evolves when moving with the fluid

$$\frac{DC_a}{Dt} = \frac{D}{Dt} \oint_{\partial S} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{r}, \quad (37.79)$$

where $\mathbf{r} = \mathbf{x}$ is the position of a fluid particle, 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 37.6 leads to

$$\frac{DC_a}{Dt} = \frac{D}{Dt} \oint_{\partial S} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{r} \quad (37.80a)$$

$$= \oint_{\partial S} \left[\frac{D\mathbf{v}}{Dt} + \boldsymbol{\Omega} \wedge \frac{D\mathbf{r}}{Dt} \right] \cdot d\mathbf{r} + \oint_{\partial S} (\mathbf{v} + \boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{v} \quad (37.80b)$$

$$= \oint_{\partial S} \left[\frac{D\mathbf{v}}{Dt} + \boldsymbol{\Omega} \wedge \mathbf{v} \right] \cdot d\mathbf{r} + \oint_{\partial S} (\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{v} \quad (37.80c)$$

$$= \oint_{\partial S} \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] \cdot d\mathbf{r}. \quad (37.80d)$$

To reach this result we set

$$\mathbf{v} = \frac{D\mathbf{r}}{Dt} \quad (37.81)$$

and used the identity

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{v} = \frac{1}{2} \oint_{\partial S} dv^2 = 0 \quad (37.82)$$

as well as

$$\oint_{\partial S} (\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot d\mathbf{v} = \oint_{\partial S} d[(\boldsymbol{\Omega} \wedge \mathbf{r}) \cdot \mathbf{v}] - \oint_{\partial S} (\boldsymbol{\Omega} \wedge d\mathbf{r}) \cdot \mathbf{v} \quad (37.83a)$$

$$= \oint_{\partial S} (\boldsymbol{\Omega} \wedge \mathbf{v}) \cdot d\mathbf{r}. \quad (37.83b)$$

Now insert the momentum equation into (37.80d) to yield

$$\frac{DC_a}{Dt} = \oint_{\partial S} \left[\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \wedge \mathbf{v} \right] \cdot d\mathbf{r}. \quad (37.84a)$$

$$= \oint_{\partial S} \left[-\frac{1}{\rho} \nabla p - \nabla \Phi + \mathbf{F} \right] \cdot d\mathbf{r}. \quad (37.84b)$$

$$= \oint_{\partial S} \left[-\frac{dp}{\rho} + \mathbf{F} \cdot d\mathbf{r} \right]. \quad (37.84c)$$

Making use of Stokes' Theorem leads to the material evolution of absolute circulation

$$\frac{DC_a}{Dt} = \oint_{\partial S} \left[-\frac{dp}{\rho} + \mathbf{F} \cdot d\mathbf{r} \right] = \int_S (\mathbf{S}_o + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS, \quad (37.85)$$

where $\mathbf{S}_o = \rho^{-2} \nabla \rho \wedge \nabla p$ is the solenoidal vector (equation (37.57)).

The circulation theorem (37.85) is the same as obtained for the non-rotating Kelvin's Circulation Theorem in Section 37.6 (see equation (37.40e)). This equivalence holds since circulation is an objective property of the fluid and so its evolution is unchanged when moving to a non-inertial rotating frame.

37.9.3 Relative circulation induced by planetary circulation

As given by equation (37.74b), the absolute circulation equals to the circulation of fluid measured in the rotating frame (relative circulation) plus circulation of the rotating frame itself (planetary circulation)

$$C_a = C + C_{\text{planet}} = C + 2 \int_S \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS. \quad (37.86)$$

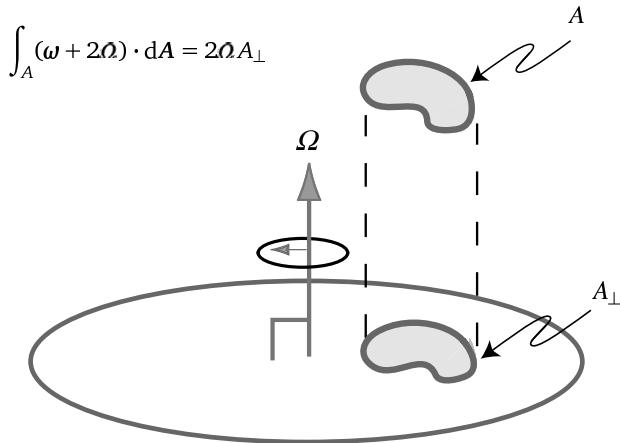


Fig. 4.7 The projection of a material circuit on to the equatorial plane. If a fluid element moves poleward, keeping its orientation to the local vertical fixed (e.g., it stays horizontal) then the area of its projection on to the equatorial plane increases. If its total (absolute) circulation is to be maintained, then the vertical component of the relative vorticity must diminish; that is, $\int_A (\omega + 2\Omega) \cdot dA = \int_A (\zeta + f) dA = \text{constant}$. Thus, the β term in $D(\zeta + f)/Dt = D\zeta/Dt + \beta v = 0$ ultimately arises from the *tilting* of a parcel relative to the axis of rotation as it moves meridionally.

Figure 37.5: Figure 4.7 from [Vallis \(2017\)](#), illustrating the beta effect.

The absolute circulation equation (37.85) allows us to write an evolution equation for the relative circulation

$$\frac{DC}{Dt} = -\frac{DC_{\text{planet}}}{Dt} + \frac{DC_a}{Dt} \quad (37.87a)$$

$$= -2 \frac{D}{Dt} \left[\int_S \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS \right] + \int_S (\mathbf{S}_o + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS. \quad (37.87b)$$

We assume that the planetary rotation is a constant³ $\boldsymbol{\Omega} = \Omega \hat{\mathbf{Z}}$, so that

$$2 \int_S \boldsymbol{\Omega} \cdot \hat{\mathbf{n}} dS = 2\Omega \int_S \hat{\mathbf{Z}} \cdot \hat{\mathbf{n}} dS = 2\Omega A_\perp, \quad (37.88)$$

where A_\perp is the projection onto the horizontal plane of the area enclosed by the circuit (see Figure 37.5). Hence, the circulation in the rotating frame materially changes according to

$$\frac{DC}{Dt} = \underbrace{-2\Omega \frac{DA_\perp}{Dt}}_{\text{planetary induction}} + \underbrace{\int_S (\mathbf{S}_o + \nabla \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} dS}_{\text{solenoids plus friction}}. \quad (37.89)$$

Latitudinal movement of a material circuit changes A_\perp , which in turn modifies the relative circulation. This process of relative circulation change via changes in planetary rotation is termed *planetary induction*, or the *beta effect*. [Holton \(1992\)](#) calls equation (37.89) the *Bjerknes Circulation Theorem* (see his equation (4.5)). In many cases, the solenoidal and friction terms are sub-dominant, such as for an ideal barotropic fluid where both terms vanish. In these cases, the evolution of relative circulation is dominated by the beta effect.

³We follow the notational conventions of Figure 10.4 with one exception. Here, the vertical Cartesian direction through the north pole is written $\hat{\mathbf{Z}}$ in order to avoid confusion with the local vertical direction $\hat{\mathbf{z}}$ determined by the geopotential.

37.9.4 A two-dimensional fluid example

To garner more insight into the beta effect, consider an ideal incompressible and two-dimensional flow (zero vertical velocity) on a rotating sphere. In the rotating frame, circulation around an infinitesimal closed loop is

$$C = A \zeta, \quad (37.90)$$

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 16.5). The material evolution of circulation is therefore given by

$$\frac{DC}{Dt} = A \frac{D\zeta}{Dt} = -2\Omega \frac{DA_{\perp}}{Dt}, \quad (37.91)$$

where the second equality follows from the circulation equation (37.89). Let the material circuit be at a latitude ϕ so that the projection of the loop area onto the equatorial plane is (see Figure 37.5)

$$A_{\perp} = A \sin \phi. \quad (37.92)$$

Hence, material evolution of the circulation is

$$\frac{DC}{Dt} = A \frac{D\zeta}{Dt} \quad (37.93a)$$

$$= -2\Omega \frac{DA_{\perp}}{Dt} \quad (37.93b)$$

$$= -2A\Omega \frac{D \sin \phi}{Dt} \quad (37.93c)$$

$$= -2A\Omega \cos \phi \frac{D\phi}{Dt} \quad (37.93d)$$

$$= -A \left[\frac{2\Omega \cos \phi}{R} \right] \left[R \frac{D\phi}{Dt} \right] \quad (37.93e)$$

$$= -A\beta v, \quad (37.93f)$$

where we introduced the meridional velocity component

$$v = R \frac{D\phi}{Dt} \quad (37.94)$$

and the meridional derivative of the planetary vorticity

$$\beta = \frac{df}{dy} = \frac{1}{R} \frac{d}{d\phi} (2\Omega \sin \phi) = \frac{2\Omega \cos \phi}{R}. \quad (37.95)$$

The result (37.93f) says that meridional motion on a rotating sphere induces relative circulation. It furthermore motivates the name *beta effect* for planetary induction of relative vorticity.

37.10 Exercises

EXERCISE 37.1: VORTICITY FOR SOLID-BODY MOTION

Show that a fluid in solid-body rotation with angular velocity

$$\mathbf{v}_{\text{solid-body}} = \boldsymbol{\Omega} \wedge \mathbf{x}, \quad (37.96)$$

has a vorticity given by

$$\nabla \wedge \mathbf{v}_{\text{solid-body}} = 2\boldsymbol{\Omega}. \quad (37.97)$$

EXERCISE 37.2: 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

$$\mathcal{V} = \int dA \int dz = \int h dA = \int (H + \Delta\eta - \eta_b) dA = H \mathcal{A}, \quad (37.98)$$

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 33.1). Additionally, recall that $z = 0$ is set according to

$$\int \eta_b dA = 0. \quad (37.99)$$

Volume conservation then ensures that

$$\int \Delta\eta dA = 0. \quad (37.100)$$

- (a) Determine the volume average of the vorticity $\hat{z} \cdot \boldsymbol{\omega}_{\text{solid}}$ arising from the solid-body rotation

$$\langle \hat{z} \cdot \boldsymbol{\omega}_{\text{solid}} \rangle = \mathcal{V}^{-1} \int \hat{z} \cdot \boldsymbol{\omega}_{\text{solid}} dV. \quad (37.101)$$

- (b) Determine the area average of the relative vorticity,

$$\bar{\zeta} = \mathcal{A}^{-1} \int \hat{z} \cdot \boldsymbol{\omega} dA, \quad (37.102)$$

in terms of the circulation around the boundary of the domain.

- (c) Determine the volume average of the relative vorticity

$$\langle \zeta \rangle = \mathcal{V}^{-1} \int \hat{z} \cdot \boldsymbol{\omega} dV. \quad (37.103)$$

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 37.3: VELOCITY POTENTIAL FOR THE FREE VORTEX

What is the velocity potential (37.3) for the free vortex whose velocity field is given by (37.14)?

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 7.3.

EXERCISE 37.4: STRAIN TENSOR FOR THE FREE VORTEX

Determine all components to the strain tensor

$$S_{pq} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \\ \frac{1}{2} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) & \frac{\partial v}{\partial y} \end{bmatrix} \quad (37.104)$$

for the free vortex as specified by the velocity field (37.14). Present the answer in the form of a 2×2 matrix.

EXERCISE 37.5: VANISHING VISCOS FRICITION FOR SOLID-BODY MOTION

As seen in Section 22.2.5, viscous effects from molecular viscosity in an incompressible fluid appear in the momentum equation as a Laplacian weighted by a constant molecular viscosity

$$\text{viscous force per mass} = \mu \nabla^2 \mathbf{v}, \quad (37.105)$$

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.

EXERCISE 37.6: FRICTION IN THE VORTICITY EQUATION

Add a viscous term of the form

$$\mathbf{F} = \mu \nabla^2 \mathbf{v}, \quad (37.106)$$

with μ a constant molecular viscosity. How is the vorticity equation modified?

EXERCISE 37.7: 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

$$\rho(x) = \rho_0 (1 - \gamma x) \quad (37.107)$$

where ρ_0 and γ are positive constants (with dimensions of density and inverse length, respectively).

- (a) Compute the density gradient $\nabla \rho$ and draw a schematic.
- (b) Compute the pressure gradient ∇p and draw a schematic at $x = 0$.
- (c) Compute the baroclinicity (or solenoidal) vector $\mathbf{S}_0 = \rho^{-2} (\nabla \rho \wedge \nabla p)$. Draw a schematic.
- (d) Describe the vorticity induced by the baroclinicity vector.

EXERCISE 37.8: CIRCULATION WITH ISLANDS

Our discussion of Stokes' Theorem has been thus far restricted to a simply connected domain, in which

$$\oint_C \mathbf{v} \cdot d\mathbf{r} = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS. \quad (37.108)$$

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, thus adding a level of complexity to the World Ocean that is absent on an AquaPlanet. Figure 37.6 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.

Derive the following expression for the circulation in multiply-connected regions

$$\oint_C \mathbf{v} \cdot d\mathbf{r} = \sum_{n=1}^N \left(\oint_{C_n} \mathbf{v} \cdot d\mathbf{r} \right) + \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} dS, \quad (37.109)$$

where N is the number of islands. In words, this result says that the circulation around a region equals to the circulation around the islands within the region, plus the normal component of the vorticity integrated over the area within the region. Removing the islands allows the island contours to be shrunk to zero size, in which case we recover the simply connected result (37.108).

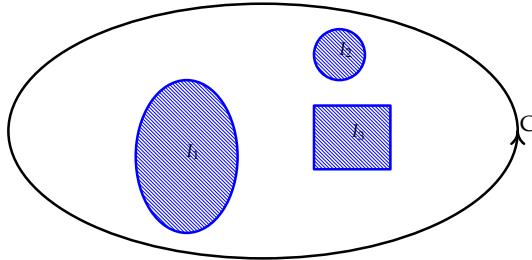


Figure 37.6: A region of the ocean consisting of three islands, I_1 , I_2 , and I_3 , with the contour, C , drawn around the three islands. The contour cannot be shrunk to a point without crossing over the islands, thus indicating that the domain is multiply connected.

EXERCISE 37.9: EVOLUTION OF CIRCULATION AROUND ISLANDS

The momentum equation for a homogeneous layer of inviscid shallow water on a tangent plane is given by

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -g \nabla \eta. \quad (37.110)$$

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

$$\mathbf{u} \cdot \nabla = u \partial_x + v \partial_y. \quad (37.111)$$

Use of a vector identity allows us to write

$$\frac{\partial \mathbf{u}}{\partial t} + (f + \zeta) \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla (\mathbf{u}^2/2 + g \eta), \quad (37.112)$$

where

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}) \quad (37.113)$$

is the vorticity of the shallow water fluid.

Consider an island, such as one shown in Figure 37.6. Each island is static and impenetrable to fluid flow, which means that

$$\mathbf{u} \cdot \hat{\mathbf{n}} = 0 \quad (37.114)$$

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

$$\mathbf{u} \wedge d\mathbf{r} = 0. \quad (37.115)$$

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

$$\frac{d}{dt} \oint_I \mathbf{u} \cdot d\mathbf{r} = 0. \quad (37.116)$$

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 37.10: HELICITY FOR AN IDEAL BAROTROPIC FLUID

Consider a closed material volume, \mathcal{R} , of isentropic non-rotating 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,

$$\hat{\mathbf{n}} \cdot \boldsymbol{\omega} = 0. \quad (37.117)$$

Such volumes are defined by closed vortex tubes, such as a smoke ring or linked smoke rings. The *helicity* of the fluid within the volume is defined as the integration of the helicity density, $\mathbf{v} \cdot \boldsymbol{\omega}$, over the closed volume

$$\mathbb{H} = \int_{\mathcal{R}} \mathbf{v} \cdot \boldsymbol{\omega} dV. \quad (37.118)$$

In Cartesian coordinates, the helicity density takes the form

$$\mathbf{v} \cdot \boldsymbol{\omega} = u(\partial_y w - \partial_z v) + v(\partial_z u - \partial_x w) + w(\partial_x v - \partial_y u). \quad (37.119)$$

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) For an inviscid barotropic fluid, show that helicity is materially conserved following the material volume

$$\frac{D\mathbb{H}}{Dt} = 0. \quad (37.120)$$

- (b) Discuss why helicity is not defined for a shallow water fluid.

Make use of the following hints.

- For a single-component isentropic fluid, the pressure can be related to the enthalpy per mass via (Section 19.3.3)

$$d\mathcal{H} = \rho^{-1} dp, \quad (37.121)$$

so that

$$\nabla \mathcal{H} = \rho^{-1} \nabla p. \quad (37.122)$$

- 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.

38

Potential vorticity

We here detail general properties of potential vorticity (PV). The simplest case is that of a barotropic fluid, which forms our starting point. However, realistic geophysical flows are baroclinic, which are fluids where “PV thinking” is most powerful. The general idea used for the derivation of PV conservation is to introduce a materially conserved scalar field used to orient the vorticity. Consequently, PV is a function of the chosen scalar.

For a barotropic fluid, the choice of scalar field is rather arbitrary, so long as it is materially conserved. However, for a baroclinic fluid we are much more constrained, since the scalar must orient vorticity in a direction to annihilate the solenoidal term. We can only annihilate the solenoidal term under certain restricted cases. Nonetheless, even when PV is only partially conserved, it remains a useful diagnostic field for use in both constraining and interpreting motions.

- Integrate ρQ within a buoyancy layer that intersects boundaries as in the ocean. Determine the time tendency for this quantity. Make use of the impermeability theorem to show that changes only occur through boundary input. Deal carefully with the boundary terms. Discuss possibility for buoyancy surfaces to intersect boundary as a normal to the boundary.

38.1	Kelvin’s circulation theorem for a perfect barotropic fluid	572
38.2	Kelvin’s circulation theorem for a perfect baroclinic fluid	573
38.3	Material evolution of PV with non-conservative processes	574
38.4	Eulerian flux-form PV budget	575
38.5	Integrated entropic PV	576
38.5.1	Integrated PV in a region bounded by an isentrope	576
38.5.2	Integrated PV in a region bounded by two isentropes	577
38.5.3	Integrated PV in a region bounded by land and an isentrope	577
38.6	Impermeability theorem for isentropic PV	578
38.6.1	Zero cross-isentrope flux of PV	578
38.6.2	Further reading	579
38.7	Exercises	580

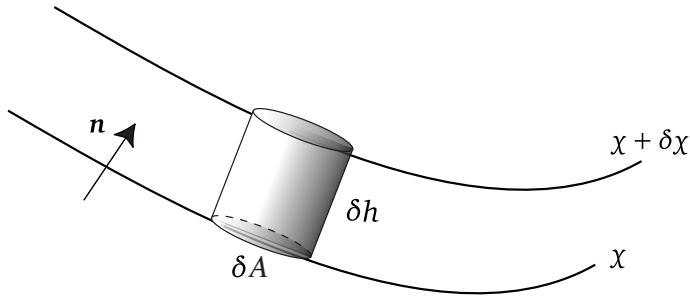


Fig. 4.8 An infinitesimal fluid element, bounded by two isosurfaces of the conserved tracer χ . As $D\chi/Dt = 0$, then $D\delta\chi/Dt = 0$.

Figure 38.1: Illustrating the geometry of a pillbox region of fluid between two iso-surfaces of a scalar field χ , as per Figure 4.8 of [Vallis \(2006\)](#).

38.1 Kelvin's circulation theorem for a perfect barotropic fluid

Consider a perfect barotropic fluid. As for the shallow water discussion in Section 36.3.2, we apply Kelvin's circulation theorem to an infinitesimal material circuit in the fluid

$$\frac{D}{Dt}(\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} \delta A) = 0. \quad (38.1)$$

The conservation of PV is built from specializing this result. For that purpose, introduce a materially conserved scalar field

$$\frac{D\chi}{Dt} = 0. \quad (38.2)$$

We make use of isosurfaces of χ to orient the material circuit. In particular, referring to Figure 38.1, let the circuit bound a tube¹ whose two ends sit on two isosurfaces, χ and $\chi + \delta\chi$. The tube volume is given by

$$\delta V = \delta A \delta h, \quad (38.3)$$

where δh is the distance between the χ and $\chi + \delta\chi$ isosurfaces. The unit normal direction orienting the area δA is given by (see Figure 38.1)

$$\hat{\mathbf{n}} = \frac{\nabla\chi}{|\nabla\chi|}. \quad (38.4)$$

The separation between the two isosurfaces is related to the normal derivative of χ through

$$\delta\chi = \nabla\chi \cdot \delta\mathbf{x} = |\nabla\chi| \hat{\mathbf{n}} \cdot \delta\mathbf{x} = |\nabla\chi| \delta h. \quad (38.5)$$

As an aside, note that this result is equivalent to (see [Vallis \(2017\)](#))

$$\delta\chi = |\nabla\chi| \delta h = (\hat{\mathbf{n}} \cdot \nabla\chi) \delta h = \frac{\delta\chi}{\delta n} \delta h. \quad (38.6)$$

Hence, the distance (or thickness) between the two isosurfaces is

$$\delta h = \frac{\delta\chi}{|\nabla\chi|}. \quad (38.7)$$

¹This tube is not generally a vortex tube. Rather, it is merely a tube bounded by isosurfaces of constant χ .

This equation has a straightforward geometric interpretation indicated in Figure 38.1. Namely, the geometric separation between two isosurfaces, δh , is reduced in regions of strong gradients and increased in regions of weak gradients.

We now have the necessary pieces in place to write

$$\omega_a \cdot \hat{\mathbf{n}} \delta A = \frac{\omega_a \cdot \nabla \chi}{|\nabla \chi|} \delta A \quad (38.8a)$$

$$= \frac{\omega_a \cdot \nabla \chi}{|\nabla \chi|} \frac{\delta V}{\delta h} \quad (38.8b)$$

$$= (\omega_a \cdot \nabla \chi) \frac{\delta V}{\delta \chi} \quad (38.8c)$$

$$= \frac{\omega_a \cdot \nabla \chi}{\rho} \frac{\rho \delta V}{\delta \chi}. \quad (38.8d)$$

Mass is materially constant, so that

$$\frac{D(\rho \delta V)}{Dt} = 0. \quad (38.9)$$

Likewise, by assumption χ is materially constant, so that the increment between two χ isosurfaces is materially constant

$$\frac{D(\delta \chi)}{Dt} = 0. \quad (38.10)$$

Bringing these elements into Kelvin's circulation theorem (38.1) leads to the material conservation law for potential vorticity, Q

$$\frac{DQ}{Dt} = 0 \quad \text{where } Q = \frac{\omega_a \cdot \nabla \chi}{\rho}. \quad (38.11)$$

38.2 Kelvin's circulation theorem for a perfect baroclinic fluid

Now consider the case of a perfect baroclinic fluid, in which Kelvin's circulation theorem for an infinitesimal circuit takes the form

$$\frac{D}{Dt}(\omega_a \cdot \hat{\mathbf{n}} \delta A) = \mathbf{S}_o \cdot \hat{\mathbf{n}} \delta A, \quad (38.12)$$

where the source on the right hand side is given by the solenoidal term

$$\mathbf{S}_o = \frac{\nabla \rho \wedge \nabla p}{\rho^2}. \quad (38.13)$$

Now assume there exists a materially constant scalar, θ , that annihilates the solenoidal term (see Figure 38.2), so that

$$\mathbf{S}_o \cdot \hat{\mathbf{n}} = \frac{\mathbf{S}_o \cdot \nabla \theta}{|\nabla \theta|} = 0. \quad (38.14)$$

In that case, the barotropic derivation detailed in Section 38.1 follows for the baroclinic case. Hence, the baroclinic PV remains materially constant

$$\frac{DQ}{Dt} = 0 \quad \text{where } Q = \frac{\omega_a \cdot \nabla \theta}{\rho}. \quad (38.15)$$

Existence of a materially constant PV for perfect baroclinic fluids depends on the existence of a materially constant scalar that annihilates the solenoidal term. The most common choices for this scalar in geophysical fluid applications are entropy or potential temperature in the atmosphere, and potential density in the ocean.

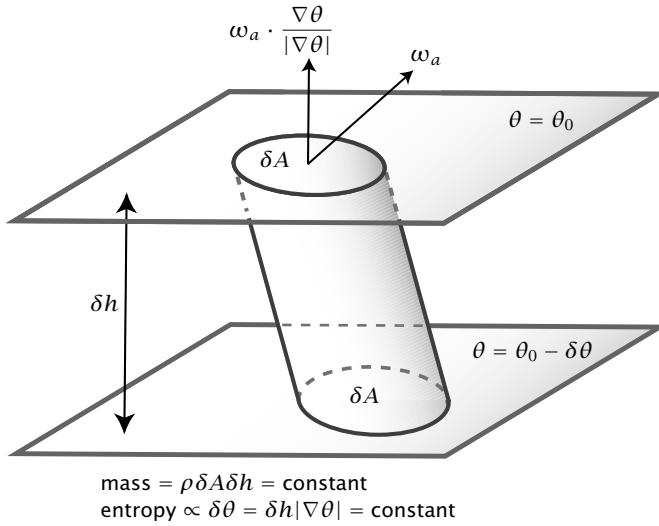


Fig. 4.9 Geometry of potential vorticity conservation. The circulation equation is $D[(\omega_a \cdot \mathbf{n})\delta A]/Dt = \mathbf{S} \cdot \mathbf{n} \delta A$, where $\mathbf{S} \propto \nabla\theta \times \nabla T$. We choose $\mathbf{n} = \nabla\theta/|\nabla\theta|$, where θ is materially conserved, to annihilate the solenoidal term on the right-hand side, and we note that $\delta A = \delta V/\delta h$, where δV is the volume of the cylinder, and the height of the column is $\delta h = \delta\theta/|\nabla\theta|$. The circulation is $C \equiv \omega_a \cdot \mathbf{n} \delta A = \omega_a \cdot (\nabla\theta/|\nabla\theta|)(\delta V/\delta h) = [\rho^{-1}\omega_a \cdot \nabla\theta](\delta M/\delta\theta)$, where $\delta M = \rho \delta V$ is the mass of the cylinder. As δM and $\delta\theta$ are materially conserved, so is the potential vorticity $\rho^{-1}\omega_a \cdot \nabla\theta$.

Figure 38.2: The geometry of a tube of fluid between two iso-surfaces of a scalar field θ , used to derive the PV conservation equation in a perfect baroclinic fluid, as per Figure 4.9 of Vallis (2006).

38.3 Material evolution of PV with non-conservative processes

We here consider the most general case of the PV equation in the presence of non-conservative processes. The derivation is algebraic and starts from the vorticity equation (37.55)

$$\frac{D\tilde{\omega}_a}{Dt} = (\tilde{\omega}_a \cdot \nabla) \mathbf{v} + \frac{1}{\rho} (S_o + \nabla \wedge \mathbf{F}), \quad (38.16)$$

where

$$\tilde{\omega}_a = \frac{\omega_a}{\rho} \quad (38.17)$$

is the absolute vorticity scaled by the density, \mathbf{F} is the friction vector, and

$$S_o = \frac{\nabla\rho \wedge \nabla p}{\rho^2} \quad (38.18)$$

is the solenoidal vector. We introduce a non-conservative scalar field

$$\frac{D\chi}{Dt} = \dot{\chi}, \quad (38.19)$$

with $\dot{\chi}$ arising from diffusion or other irreversible processes.

We next make use of the identity

$$(\tilde{\omega}_a \cdot \nabla) \frac{D\chi}{Dt} = \tilde{\omega}_a \cdot \frac{D(\nabla\chi)}{Dt} + [(\tilde{\omega}_a \cdot \nabla) \mathbf{v}] \cdot \nabla\chi, \quad (38.20)$$

with this identity readily proven by expanding terms assuming Cartesian coordinates. Rearrangement, and use of the scalar equation (38.19), leads to

$$\tilde{\omega}_a \cdot \frac{D(\nabla\chi)}{Dt} = (\tilde{\omega}_a \cdot \nabla) \dot{\chi} - [(\tilde{\omega}_a \cdot \nabla) \mathbf{v}] \cdot \nabla\chi. \quad (38.21)$$

Now project the vorticity equation (38.16) onto the direction normal to the χ isosurfaces

$$\nabla\chi \cdot \left(\frac{D\tilde{\omega}_a}{Dt} \right) = \nabla\chi \cdot [(\tilde{\omega}_a \cdot \nabla) \mathbf{v}] + \nabla\chi \cdot \left[\frac{1}{\rho} (\mathbf{S}_o + \nabla \wedge \mathbf{F}) \right]. \quad (38.22)$$

The sum of equations (38.21) and (38.22) leads to

$$\frac{D(\nabla\chi \cdot \tilde{\omega}_a)}{Dt} = (\tilde{\omega}_a \cdot \nabla) \dot{\chi} + \nabla\chi \cdot \left[\frac{1}{\rho} (\mathbf{S}_o + \nabla \wedge \mathbf{F}) \right]. \quad (38.23)$$

As in the discussion from Section 38.2, choose a special scalar so that it annihilates the solenoidal term

$$\nabla\theta \cdot \mathbf{S}_o = 0, \quad (38.24)$$

in which case we are led to the potential vorticity equation in the presence of irreversible terms

$$\rho \frac{D}{Dt} \left(\frac{\omega_a \cdot \nabla\theta}{\rho} \right) = (\omega_a \cdot \nabla) \dot{\theta} + \nabla\theta \cdot (\nabla \wedge \mathbf{F}). \quad (38.25)$$

We see that the material evolution of PV

$$Q = \frac{\omega_a \cdot \nabla\theta}{\rho} \quad (38.26)$$

is affected by diabatic processes and friction. In the absence of these irreversible processes, we are left with the same PV conservation statement (38.15) derived for the perfect fluid using Kelvin's circulation theorem. However, in their presence, PV can be either generated or destroyed. Such source/sink regions of potential vorticity are often localized to boundary regions, such as boundary layer regions in the atmosphere and ocean where strong mechanical and/or buoyancy processes are active.

38.4 Eulerian flux-form PV budget

Equation (38.25) provides an expression for the material evolution of potential vorticity in the presence of diabatic processes and friction. For many purposes, it is useful to transform the material evolution equation into a flux-form Eulerian equation. For that purpose we make use of the following identities

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad \text{material time derivative related to Eulerian} \quad (38.27a)$$

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v} \quad \text{mass conservation} \quad (38.27b)$$

$$(\omega_a \cdot \nabla) \dot{\theta} = \nabla \cdot (\omega_a \dot{\theta}) \quad \nabla \cdot \omega_a = 0 \quad (38.27c)$$

$$\nabla\theta \cdot (\nabla \wedge \mathbf{F}) = \nabla \cdot (\mathbf{F} \wedge \nabla\theta) \quad \text{divergence of curl vanishes.} \quad (38.27d)$$

These identities then lead to the flux-form Eulerian budget equation for PV

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot [\rho Q \mathbf{v} - \boldsymbol{\omega}_a \dot{\theta} - \mathbf{F} \wedge \nabla \theta] = 0. \quad (38.28)$$

The budget equation (38.28) says that the density-weighted potential vorticity

$$\rho Q = \boldsymbol{\omega}_a \cdot \nabla \theta. \quad (38.29)$$

has a local time tendency determined by the convergence of a flux

$$\mathbf{J} = \rho Q \mathbf{v} - \boldsymbol{\omega}_a \dot{\theta} - \mathbf{F} \wedge \nabla \theta. \quad (38.30)$$

The first term in the flux arises from advection; the second from diabatic processes; and the third from friction.

38.5 Integrated entropic PV

It is common to define potential vorticity using potential temperature in the atmosphere, θ , whereas for the ocean it is more common to use potential density (see Appendix ?? for details on ocean potential vorticity). We generically write this *entropic PV* in the form

$$Q = \frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{\rho} = \frac{\nabla \cdot (\boldsymbol{\omega}_a \theta)}{\rho}. \quad (38.31)$$

In this section we consider Q to be an intensive fluid property measuring the amount of PV substance per unit mass.² With this interpretation, the amount of PV substance is determined by the integral of ρQ over the volume,

$$\mathcal{I} = \int_{\mathcal{D}} Q \rho dV = \int_{\mathcal{D}} \nabla \cdot (\boldsymbol{\omega}_a \theta) dV = \int_{\partial \mathcal{D}} \boldsymbol{\omega}_a \theta \cdot \hat{\mathbf{n}} dA, \quad (38.32)$$

where the final equality used the divergence theorem. Hence, the volume integrated PV substance in a region is determined solely by values on the region boundary.

38.5.1 Integrated PV in a region bounded by an isentrope

Consider a volume of fluid bounded by a single potential temperature surface as shown in the left panel of Figure 38.3. Since the boundary is set by a constant θ surface, we can pull θ outside of the surface integral in equation (38.32) so that

$$\mathcal{I} = \int_{\partial \mathcal{D}} \boldsymbol{\omega}_a \theta \cdot \hat{\mathbf{n}} dS = \theta_2 \int_{\partial \mathcal{D}} \boldsymbol{\omega}_a \cdot \hat{\mathbf{n}} dS. \quad (38.33)$$

We can now use the divergence theorem to return to the volume integral, only now with θ remaining outside of the integrand, in which case

$$\mathcal{I} = \theta_2 \int_{\mathcal{D}} \nabla \cdot \boldsymbol{\omega}_a dV = 0, \quad (38.34)$$

where $\nabla \cdot \boldsymbol{\omega}_a = 0$ led to the final equality. This remarkable result follows without invoking any dynamical information. Hence, it is purely kinematic. It says that there is zero “PV substance” contained within any region bounded solely by an isentrope. It holds whether there are reversible or irreversible processes acting on the isentrope. Our only assumption is that the domain is fully enclosed by a θ isosurface.

²Recall our discussion of extensive and intensive fluid properties in Section 15.6.1.

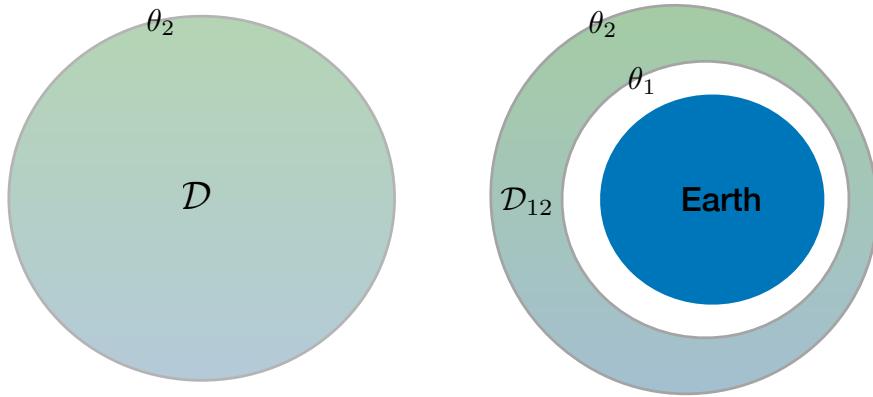


Figure 38.3: Integrating PV over regions bounded by potential temperature surfaces that do not intersect the ground. The left panel considers a single potential temperature surface, $\theta = \theta_2$, bounding the fluid region \mathcal{D} . The right panel considers an isentropic layer \mathcal{D}_{12} bounded by two potential temperatures $\theta_1 < \theta_2$ surrounding the earth, with neither temperature surface intersecting the ground. In both cases there is zero domain integrated PV in the regions.

38.5.2 Integrated PV in a region bounded by two isentropes

The identity (38.34) has a corollary, in which we consider a region bounded by two θ isosurfaces such as \mathcal{D}_{12} shown in the right panel of Figure 38.3. The above arguments hold for that region as well, since we can decompose the surface integral into two integrals separately over θ_1 and θ_2

$$\mathcal{I} = \int_{\mathcal{D}_{12}} \nabla \cdot (\boldsymbol{\omega}_a \theta) dV \quad (38.35a)$$

$$= \int_{\mathcal{D}_2} \nabla \cdot (\boldsymbol{\omega}_a \theta) dV - \int_{\mathcal{D}_1} \nabla \cdot (\boldsymbol{\omega}_a \theta) dV, \quad (38.35b)$$

where the domain \mathcal{D}_1 extends from the ground up to θ_1 and \mathcal{D}_2 extends from the ground up to θ_2 . Integration over the region below θ_1 cancels through the subtraction. Indeed, the region below θ_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{D} in Figure 38.3. Invoking the single isentrope result we see that both integrals separately vanish. We are thus led to a vanishing integral for the layer

$$\mathcal{I} = \int_{\mathcal{D}_{12}} \nabla \cdot (\boldsymbol{\omega}_a \theta) dV = 0. \quad (38.36)$$

Again, the key assumption is that no isentrope intersects land, in which case we are able to ignore the presence of land altogether and thus make use of the single isentrope result.

38.5.3 Integrated PV in a region bounded by land and an isentrope

Now consider a fluid domain bounded by an isentrope that intersects the ground (atmospheric example) or ocean surface (ocean example), as shown in Figure 38.3. The integrated PV is given

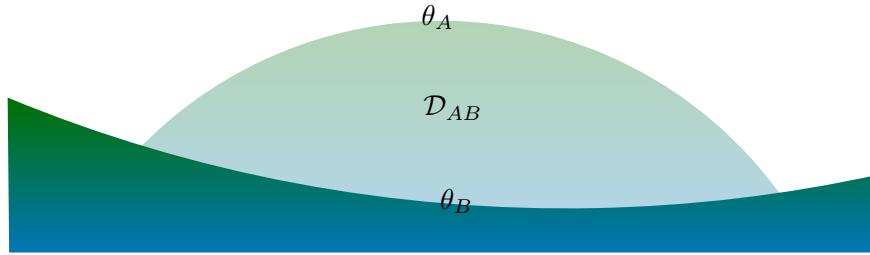


Figure 38.4: A fluid region, \mathcal{D}_{AB} , bounded by a potential temperature surface that intersects the ground. The potential temperature θ_A is an isosurface, whereas the boundary value θ_B is a space and time dependent function.

by

$$\mathcal{I} = \theta_A \int_{\theta=\theta_A} \omega_a \cdot \hat{n} dS + \int_{\text{ground}} \theta_B \omega_a \cdot \hat{n} dS \quad (38.37a)$$

$$= \theta_A \int_{\theta_A+\text{ground}} \omega_a \cdot \hat{n} dS + \int_{\text{ground}} (\theta_B - \theta_A) \omega_a \cdot \hat{n} dS \quad (38.37b)$$

$$= \int_{\text{ground}} (\theta_B - \theta_A) \omega_a \cdot \hat{n} dS, \quad (38.37c)$$

where we made use of the divergence theorem to reach the final equality. Consequently, the PV substance in a region enclosed by isentropes can change only when the isentropes intersect a boundary that has a non-constant entropy.

38.6 Impermeability theorem for isentropic PV

The above results made use of the property that ρQ is a total divergence, and as such they are purely kinematic. In this section we study the PV flux vector, \mathbf{J} , and prove that this flux never penetrates an isentrope. This result then provides another explanation for why ρQ remains constant within an isentrope unless the isentrope intersects a boundary.

38.6.1 Zero cross-isentrope flux of PV

Recall the PV evolution equation (38.28)

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot [\rho Q \mathbf{v} - \omega_a \dot{\theta} - \mathbf{F} \wedge \nabla \theta] = 0. \quad (38.38)$$

Now decompose the velocity into two components, one oriented parallel to isentropes and one oriented perpendicular

$$\mathbf{v}_{\parallel} = \mathbf{v} - \hat{n} (\mathbf{v} \cdot \hat{n}) \quad \mathbf{v}_{\perp} = -\frac{\hat{n} \partial \theta / \partial t}{|\nabla \theta|} \quad \mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} + \frac{\hat{n} \dot{\theta}}{|\nabla \theta|} \quad (38.39)$$

where

$$\hat{n} = \frac{\nabla \theta}{|\nabla \theta|} \quad (38.40)$$

is the normal vector on an isentrope. With this decomposition, the PV flux vector takes the form

$$\mathbf{J} = \rho \mathbf{v} Q - \dot{\theta} \boldsymbol{\omega}_a + \nabla \theta \wedge \mathbf{F} \quad (38.41a)$$

$$= \left[\mathbf{v}_{\parallel} + \mathbf{v}_{\perp} + \frac{\dot{\theta} \nabla \theta}{|\nabla \theta|^2} \right] \rho Q - \dot{\theta} \boldsymbol{\omega}_a + \nabla \theta \wedge \mathbf{F} \quad (38.41b)$$

$$= (\mathbf{v}_{b\parallel} + \mathbf{v}_{b\perp}) \rho Q - \dot{\theta} [\boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}}] + \nabla \theta \wedge \mathbf{F} \quad (38.41c)$$

$$= (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) \rho Q - \dot{\theta} (\boldsymbol{\omega}_a)_{\parallel} + \nabla \theta \wedge \mathbf{F} \quad (38.41d)$$

$$= \mathbf{v}_{\perp} \rho Q + \left(\rho Q \mathbf{v}_{\parallel} - \dot{\theta} (\boldsymbol{\omega}_a)_{\parallel} \right) + \nabla \theta \wedge \mathbf{F} \quad (38.41e)$$

$$\equiv \mathbf{J}_{\perp} + \mathbf{J}_{\parallel}, \quad (38.41f)$$

where

$$(\boldsymbol{\omega}_a)_{\parallel} = \boldsymbol{\omega}_a - (\boldsymbol{\omega}_a \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} \quad (38.42a)$$

$$= \boldsymbol{\omega}_a - \left[\frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{|\nabla \theta|^2} \right] \nabla \theta \quad (38.42b)$$

$$= \boldsymbol{\omega}_a - \frac{\rho Q}{|\nabla \theta|} \hat{\mathbf{n}}. \quad (38.42c)$$

The above results suggest that we write the PV equation in the form

$$\frac{\partial(\rho Q)}{\partial t} = -\nabla \cdot \mathbf{J} = -\nabla \cdot (\mathbf{v}_Q \rho Q). \quad (38.43)$$

The velocity

$$\mathbf{v}_Q = \frac{\mathbf{J}}{\rho Q} = \mathbf{v}_{\perp} + \mathbf{v}_{\parallel} - \frac{\dot{\theta} (\boldsymbol{\omega}_a)_{\parallel} + \mathbf{F} \wedge \nabla \theta}{\rho Q} \quad (38.44)$$

is an effective velocity that advects a notional ‘‘PV substance’’ through the fluid, and it has a normal component given by

$$\mathbf{v}_Q \cdot \hat{\mathbf{n}} = -\frac{\partial \theta / \partial t}{|\nabla \theta|} = \mathbf{v}_{\perp} \cdot \hat{\mathbf{n}}. \quad (38.45)$$

Recall from Section 15.7.2 that we discussed the kinematic boundary condition for a material surface. In that discussion we showed that the normal component to a velocity field that does not penetrate a moving surface is given precisely in the form given here for $\mathbf{v}_Q \cdot \hat{\mathbf{n}}$. We also provide a more general discussion of dia-surface transport in Chapter 18. From both of these discussions we are led to conclude that no PV substance crosses a constant entropy surface, even as that surface moves and even in the presence of irreversible processes. Instead, the normal projection of the PV flux, divided by ρQ , is identical to the projection of the velocity for a point on the entropy surface. We thus see that the entropy surface moves precisely in a way to account for changes in ρQ so that there is zero penetration of the isentrope by the PV flux. This is a rather remarkable result that has deep implications for PV budgets, some of which are pursued in Chapter 39.

38.6.2 Further reading

The impermeability theorem was introduced by [Haynes and McIntyre \(1987\)](#). This introduction to the literature was met by some confusion, which prompted clarification by [Haynes and McIntyre \(1990\)](#). It focuses attention to boundary processes for understanding how potential vorticity changes. For example, this emphasis on boundary processes is the focus of oceanographic studies of submesoscale instabilities ([Thomas et al., 2008](#)). We detail many of these points regarding boundary conditions in Section 39.2.

38.7 Exercises

EXERCISE 38.1: 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,

$$\rho = \rho_0 (1 - \alpha \theta), \quad (38.46)$$

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

$$b = -g \left[\frac{\rho - \rho_0}{\rho_0} \right] = \alpha g \theta. \quad (38.47)$$

The governing equations for a perfect fluid version of this system are given by

$$\frac{D\mathbf{v}}{Dt} + f(\hat{\mathbf{z}} \wedge \mathbf{v}) = -\nabla\phi + b\hat{\mathbf{z}} \quad (38.48)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (38.49)$$

$$\frac{D\theta}{Dt} = 0, \quad (38.50)$$

where

$$\phi = \frac{\delta p}{\rho_0} \quad (38.51)$$

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

$$Q_{\text{vert}}^{\text{bouss}} = (\zeta + f) N^2 \quad (38.52)$$

where ζ is the vertical component to the relative vorticity and N^2 is the squared buoyancy frequency (Section 20.3.3). 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

$$Q_{\text{horz}}^{\text{bouss}} = \boldsymbol{\omega} \cdot \nabla_z b \approx -f^{-1} |\nabla_z b|^2. \quad (38.53)$$

Hint: recall that for geostrophic flow, the vertical velocity is much smaller than horizontal.

EXERCISE 38.2: PV FOR NON-HYDROSTATIC BOUSSINESQ FLOW

Reconsider Exercise 38.1 in the presence of irreversible friction and buoyancy sources so that the governing equations are

$$\frac{D\mathbf{v}}{Dt} + f(\hat{\mathbf{z}} \wedge \mathbf{v}) = -\nabla\phi + b\hat{\mathbf{z}} + \mathbf{F} \quad (38.54)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (38.55)$$

$$\frac{D\theta}{Dt} = \dot{\theta} \quad (38.56)$$

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

$$\frac{\partial Q}{\partial t} = -\nabla \cdot \mathbf{J}. \quad (38.57)$$

What is the PV flux \mathbf{J} ? Note that your answer is unique up to the curl of an arbitrary vector (gauge symmetry).

- (c) A common diabatic process is written in the form of a damping source

$$\dot{b} = -\mu(b - b^*), \quad (38.58)$$

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

$$\mathbf{F} = -\lambda \mathbf{u}, \quad (38.59)$$

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, $Q^{pg} = f N^2$, where $N^2 = \partial b / \partial z$ is the squared buoyancy frequency.

- (e) Discuss the balance needed between forcing terms in \mathbf{J} 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.

EXERCISE 38.3: PV IN ISOPYCNAL COORDINATES WITH MIXING AND HEATING

In Section 39.5, we consider the PV equation for an adiabatic, inviscid, hydrostatic, Boussinesq fluid using isopycnal vertical coordinates. After carefully studying that section, here extend to the case of friction in the momentum equation and diabatic heating in the buoyancy equation.

- What is the material time evolution of PV with diabatic heating and friction?
- What is the flux-form Eulerian PV equation?

Present the answers in terms of isopycnal vertical coordinates rather than the traditional geopotential vertical coordinates.

1. HINT: There are two ways to proceed. One is to convert the non-hydrostatic PV equation in Exercise 38.2 to isopycnal coordinates, after making the hydrostatic approximation. Another is to start from the equations of motion in isopycnal coordinates and derive the vorticity equation and then the PV equation.

2. HINT: The equations of motion with diabatic heating and friction, written using isopycnal (or buoyancy) vertical coordinates, take the form

$$\left[\frac{\partial \mathbf{u}}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \mathbf{u} + \dot{b} \frac{\partial \mathbf{u}}{\partial b} + \mathbf{f} \wedge \mathbf{u} = -\nabla_b M + \mathbf{F} \quad (38.60a)$$

$$\frac{\partial M}{\partial b} = -z \quad (38.60b)$$

$$\left[\frac{\partial h}{\partial t} \right]_b + \nabla_b \cdot (h \mathbf{u}) = -\frac{\partial(h \dot{b})}{\partial b} \quad (38.60c)$$

$$\frac{D b}{D t} = \dot{b}. \quad (38.60d)$$

As seen by the thickness equation (38.60c), 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.

3. HINT: We can make use of the material time derivative operator (31.23b) to write the material form of the equations

$$\frac{D \mathbf{u}}{D t} + \mathbf{f} \wedge \mathbf{u} = -\nabla_b M + \mathbf{F} \quad (38.61a)$$

$$\frac{\partial M}{\partial b} = -z \quad (38.61b)$$

$$\frac{D h}{D t} + h \nabla_b \cdot \mathbf{u} = -h \frac{\partial \dot{b}}{\partial b}. \quad (38.61c)$$

Ocean potential vorticity[†]

We here present further details for potential vorticity and its budgets, with particular application to the ocean.

39.1	PV for the ocean	583
39.1.1	Solenoidal vector	584
39.1.2	PV based on potential density	584
39.1.3	An example equation of state with exact PV conservation	585
39.1.4	Further reading	585
39.2	PV budget for an isopycnal layer	586
39.2.1	Layer integrated budget	586
39.2.2	The land-sea boundary condition	587
39.2.3	The air-sea boundary condition	587
39.3	Hydrostatic Boussinesq fluid	587
39.3.1	Vorticity equation	588
39.3.2	Ertel potential vorticity	589
39.3.3	Potential vorticity flux vector	590
39.3.4	Potential vorticity flux vector and the Bernoulli potential	591
39.3.5	Impermeability theorem	594
39.4	Distinguishing buoyancy and density	596
39.4.1	Potential vorticity	596
39.4.2	Potential vorticity flux vector	597
39.4.3	Potential vorticity flux vector and the Bernoulli potential	597
39.4.4	Vector-invariant velocity equation	597
39.5	Vorticity and PV using isopycnal vertical coordinates	598
39.5.1	Derivation of the vorticity equation	598
39.5.2	Derivation of the potential vorticity equation	599
39.5.3	A note about specific thickness	599
39.5.4	Coordinate transforming vorticity and potential vorticity	600

39.1 PV for the ocean

The most relevant PV for atmospheric dynamics is based on entropy as the scalar field, with entropy directly proportional to potential temperature (see Section 19.8). However, the presence of moisture prevents there from being an exact PV conservation principle; only an approximate one holds. Likewise, in the ocean, a realistic equation of state along with non-constant temperature

and salinity preclude an exact PV conservation principle. The problem in practice is that there is no suitable scalar that can annihilate the solenoidal term in the vorticity equation. However, there are important approximate cases that allow for PV to still be of great use for oceanography.

39.1.1 Solenoidal vector

Recall the solenoidal vector is given by

$$\mathbf{S}_o = \frac{\nabla\rho \wedge \nabla p}{\rho^2}. \quad (39.1)$$

If we take the *in situ* density as the scalar field for defining PV, then $\mathbf{S}_o \cdot \hat{\mathbf{n}} = 0$. 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

$$\frac{D\rho}{Dt} = \frac{\partial\rho}{\partial S} \frac{DS}{Dt} + \frac{\partial\rho}{\partial\theta} \frac{D\theta}{Dt} + \frac{\partial\rho}{\partial p} \frac{Dp}{Dt}. \quad (39.2)$$

Even in the absence of sources or mixing of temperature and salinity, in which case $DS/Dt = 0$ and $D\theta/Dt = 0$, *in situ* density has a nonzero material time evolution due to mechanical effects $Dp/Dt \neq 0$. Hence, *in situ* density is not an appropriate scalar for use in developing a PV conservation principle.

39.1.2 PV based on potential density

Potential density is commonly used in oceanography (see Section 20.2.5), with potential density the *in situ* density referenced to a chosen pressure.¹ We write potential density as

$$\sigma(S, \theta) = \rho(S, \theta, p = p_R). \quad (39.3)$$

The material time derivative of potential density is

$$\frac{D\sigma}{Dt} = \frac{\partial\sigma}{\partial S} \frac{DS}{Dt} + \frac{\partial\sigma}{\partial\theta} \frac{D\theta}{Dt}, \quad (39.4)$$

which vanishes in the absence of irreversible salinity and temperature processes. When using potential density as the scalar field for PV, the solenoidal term takes the form

$$\rho^2 \mathbf{S}_o \cdot \nabla\sigma = (\nabla\rho \wedge \nabla p) \cdot \nabla\sigma \quad (39.5a)$$

$$= (\nabla\sigma \wedge \nabla\rho) \cdot \nabla p \quad (39.5b)$$

$$= [(\sigma_S \nabla S + \sigma_\theta \nabla\theta) \wedge (\rho_S \nabla S + \rho_\theta \nabla\theta + \rho_p \nabla p)] \cdot \nabla p \quad (39.5c)$$

$$= [(\sigma_S \nabla S + \sigma_\theta \nabla\theta) \wedge (\rho_S \nabla S + \rho_\theta \nabla\theta)] \cdot \nabla p \quad (39.5d)$$

$$= [\sigma_S \nabla S \wedge \rho_\theta \nabla\theta + \sigma_\theta \nabla\theta \wedge \rho_S \nabla S] \cdot \nabla p \quad (39.5e)$$

$$= (\sigma_S \rho_\theta - \sigma_\theta \rho_S) (\nabla S \wedge \nabla\theta) \cdot \nabla p, \quad (39.5f)$$

where we used the shorthand notation for partial derivatives

$$\rho_S = \frac{\partial\rho}{\partial S} \quad \sigma_S = \frac{\partial\sigma}{\partial S} \quad (39.6a)$$

$$\rho_\theta = \frac{\partial\rho}{\partial\theta} \quad \sigma_\theta = \frac{\partial\sigma}{\partial\theta}. \quad (39.6b)$$

¹Oceanographers typically choose pressure referenced to a standard atmospheric surface pressure.

Note that the triple product, $(\nabla S \wedge \nabla \theta) \cdot \nabla p$, also appears in the discussion of neutral helicity in Section 24.5 (see equation (24.106)). Equation (39.4) allows us to identify cases where the solenoidal vector is annihilated, $\mathbf{S}_o \cdot \nabla \sigma = 0$, thus yielding an exact PV conservation principle.

- **UNIFORM SALINITY OR UNIFORM POTENTIAL TEMPERATURE:** If salinity or potential temperature are spatially uniform, then $\mathbf{S}_o \cdot \nabla \sigma = 0$. The uniform salinity case was mentioned in Section 4.5.4 of [Vallis \(2006\)](#) as the only case where the ocean has an exact PV conservation principle. However, there are other cases, as described in the next bullet point.
- **ADDITIONAL 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 $\mathbf{S}_o \cdot \nabla \sigma \neq 0$. Nonetheless, the solenoidal term is annihilated if the *in situ* density has a pressure dependence that is additive, so that we can write

$$\rho(S, \theta, p) = \sigma(S, \theta) + F(p) - F(p_R) \Rightarrow \sigma_S \rho_\theta - \sigma_\theta \rho_S = 0, \quad (39.7)$$

which then leads to an exact PV conservation principle. Notably, we did not assume the equation of state to be linear; only that it has the special functional form in equation (39.7). 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.

39.1.3 An example equation of state with exact PV conservation

An explicit realization of the equation of state (39.7) 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

$$\rho(S, \theta, p) \approx \sigma(S, \theta) + (p - p_R) \underbrace{\left(\frac{\partial \rho}{\partial p} \right)_{S=S_R, \theta=\theta_R, p=p_R}}_{F(p) - F(p_R)} + H.O.T. \quad (39.8)$$

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

$$c_s^{-2} = \frac{\partial \rho}{\partial p} \approx \text{constant}. \quad (39.9)$$

In this case, the equation of state takes the form

$$\rho(S, \theta, p) \approx \sigma(S, \theta) + \frac{p - p_R}{c_s^2}, \quad (39.10)$$

where again $\sigma(S, \theta) = \rho(S, \theta, p_R)$ is the potential density referenced to $p = p_R$.

39.1.4 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 24.4.4).

39.2 PV budget for an isopycnal layer

We further illustrate the implications of the impermeability theorem of Section 38.6 by considering a buoyancy (isopycnal) layer within the ocean that intersects the land on one side and the atmosphere on the other (Figure 39.1). The 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 for the layer changes only through exchanges at the land-sea and air-sea boundaries. We here develop the mathematical form for these boundary conditions.

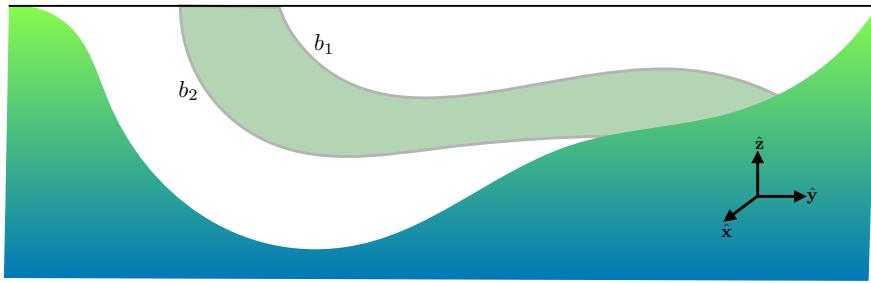


Figure 39.1: An isopycnal layer (gray region) bounded by buoyancy b_1 and $b_2 < b_1$, with the layer intersecting land at one side and the atmosphere at the other.

39.2.1 Layer integrated budget

In addition to waves, currents, entrainment, and detrainment affecting the layer interfaces, the layer side boundaries move and thus change both their vertical and horizontal extents. To formulate the PV budget within an isopycnal layer intersecting a boundary thus requires the Leibniz-Reynolds transport theorem derived in Section 15.6.4, which leads to

$$\frac{d}{dt} \left[\int_{\mathcal{D}} \rho Q dV \right] = \int_{\mathcal{D}} \left[\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot \left(\rho Q \frac{dx}{dt} \right) \right] dV, \quad (39.11)$$

where \mathcal{D} is the domain defined by the layer and dx/dt is the velocity for a point on the domain boundary. Making use of the PV equation, $\partial(\rho Q)/\partial t = -\nabla \cdot \mathbf{J}$, and the divergence theorem renders

$$\frac{d}{dt} \left[\int_{\mathcal{D}} \rho Q dV \right] = \int_{\mathcal{D}} \nabla \cdot \left[-\mathbf{J} + \left(\rho Q \frac{dx}{dt} \right) \right] dV \quad (39.12a)$$

$$= \int_{\partial\mathcal{D}} \left[-\mathbf{J} + \left(\rho Q \frac{dx}{dt} \right) \right] \cdot \hat{\mathbf{n}} dS. \quad (39.12b)$$

The velocity of a point along the isopycnal boundaries has a normal component that satisfies equation (38.45) (here applied to isopycnals rather than isotherms)

$$\frac{dx}{dt} \cdot \hat{\mathbf{n}} = -\frac{\partial b}{\partial t} / |\nabla b| \implies \left[-\mathbf{J} + \left(\rho Q \frac{dx}{dt} \right) \right] \cdot \hat{\mathbf{n}} = 0 \quad \text{isopycnal interfaces.} \quad (39.13)$$

As per the impermeability theorem, we conclude that the isopycnal layer integrated PV changes only via transfer across the land-sea and air-sea boundaries

$$\frac{d}{dt} \left[\int_{\mathcal{D}} \rho Q dV \right] = \int_{\text{boundaries}} \left[-\mathbf{J} + \left(\rho Q \frac{dx}{dt} \right) \right] \cdot \hat{\mathbf{n}} dS. \quad (39.14)$$

39.2.2 The land-sea boundary condition

Making use of equation (38.41a) at the land-sea boundary renders

$$-\mathbf{J} + \rho Q \dot{\mathbf{x}} = \rho Q (\dot{\mathbf{x}} - \mathbf{v}) + \dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}, \quad (39.15)$$

where we introduced the shorthand for the boundary velocity

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}. \quad (39.16)$$

At a solid boundary, the no-flux 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 land-sea boundary condition is thus given just by the irreversible terms

$$(-\mathbf{J} + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = (\dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} \quad \text{land-sea boundary.} \quad (39.17)$$

39.2.3 The air-sea boundary condition

We make use of the kinematic boundary condition derived in Section 15.7.3 for the permeable air-sea boundary, where the boundary condition (15.83) leads to

$$\rho \hat{\mathbf{n}} \cdot (\dot{\mathbf{x}} - \mathbf{v}) = \mathcal{Q}_m \quad \text{air-sea boundary} \quad (39.18)$$

with \mathcal{Q}_m the mass per time per surface area crossing the boundary. We are thus led to the PV boundary condition

$$(-\mathbf{J} + \rho Q \dot{\mathbf{x}}) \cdot \hat{\mathbf{n}} = Q \mathcal{Q}_m + (\dot{b} \boldsymbol{\omega}_a + \nabla b \wedge \mathbf{F}) \cdot \hat{\mathbf{n}} \quad \text{air-sea boundary.} \quad (39.19)$$

Besides the irreversible processes, PV is affected at the air-sea interface by the transfer of matter across the boundary.

39.3 Hydrostatic Boussinesq fluid

We here discuss the Ertel PV 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. The governing equations are

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{v} = -\nabla_z \phi + \mathbf{F} \quad (39.20a)$$

$$\frac{\partial \phi}{\partial z} = b \quad (39.20b)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (39.20c)$$

$$\frac{Db}{Dt} = \dot{b}. \quad (39.20d)$$

In this equation, the non-divergent velocity field is written

$$\mathbf{v} = (\mathbf{u}, w), \quad (39.21)$$

with $\mathbf{u} = (u, v, 0)$ the horizontal component. The perturbation pressure is

$$\phi = \frac{\delta p}{\rho_0} = \frac{p - p_0}{\rho_0}, \quad (39.22)$$

with $p_0 = p_0(z)$ in hydrostatic balance with the reference density

$$\frac{\partial p_0}{\partial z} = -g \rho_0 \quad (39.23)$$

and p the full hydrostatic pressure satisfying the hydrostatic balance

$$\frac{\partial p}{\partial z} = -g \rho. \quad (39.24)$$

The buoyancy is given by

$$b = -g \left[\frac{\rho - \rho_0}{\rho_0} \right], \quad (39.25)$$

with negative buoyancy for fluid more dense than the reference density ρ_0 , and vice versa. We assume the Coriolis parameter of the form $\mathbf{f} = f \hat{z}$, in which case

$$\mathbf{f} \wedge \mathbf{v} = \mathbf{f} \wedge \mathbf{u}. \quad (39.26)$$

Finally, the horizontal friction vector is given by

$$\mathbf{F} = (F^x, F^y, 0) \quad (39.27)$$

and the gradient operator is

$$\nabla = \nabla_z + \hat{z} \partial_z. \quad (39.28)$$

39.3.1 Vorticity equation

To derive the vorticity equation, it is useful to combine the horizontal momentum equation with the hydrostatic balance, in which case

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{v} = -\nabla \phi + b \hat{z} + \mathbf{F}. \quad (39.29)$$

As for the non-hydrostatic case (Section 37.7.1), 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

$$\boldsymbol{\omega}^{\text{hy}} = \nabla \wedge \mathbf{u} = -\hat{x} \partial_z v + \hat{y} \partial_z u + \hat{z} \zeta, \quad (39.30)$$

where

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad (39.31)$$

is the vertical component to the relative vorticity. It is then straightforward to show that

$$\boldsymbol{\omega}^{\text{hy}} \wedge \mathbf{v} = \hat{x} (w \partial_z u - v \partial_x u + v \partial_y u) + \hat{y} (w \partial_z v - u \partial_y u + u \partial_x v) - \hat{z} \partial_z (u^2 + v^2)/2 \quad (39.32a)$$

$$= w \partial_z \mathbf{u} + \zeta (\hat{x} v - \hat{u} u) - \hat{z} \partial_z (u^2 + v^2)/2, \quad (39.32b)$$

in which case

$$\nabla(\mathbf{u}^2/2) + \boldsymbol{\omega}^{\text{hy}} \wedge \mathbf{v} = \nabla(u^2 + v^2)/2 - \hat{z} \partial_z (u^2 + v^2)/2 + w \partial_z \mathbf{u} + \zeta (u \hat{y} - v \hat{x}) \quad (39.33a)$$

$$= (u \partial_x + v \partial_y + w \partial_z) \mathbf{u} \quad (39.33b)$$

$$= (\mathbf{v} \cdot \nabla) \mathbf{u}. \quad (39.33c)$$

The material time derivative of horizontal velocity can thus be written as

$$\frac{D\mathbf{u}}{Dt} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{u} = \frac{\partial \mathbf{u}}{\partial t} + \boldsymbol{\omega}^{\text{hy}} \wedge \mathbf{v} + \nabla(\mathbf{u}^2/2), \quad (39.34)$$

which then leads to the horizontal momentum equation

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{f} + \boldsymbol{\omega}^{\text{hy}}) \wedge \mathbf{v} = -\nabla(\phi + \mathbf{u}^2/2) + \hat{\mathbf{z}} b + \mathbf{F}. \quad (39.35)$$

Now take the curl, and make use of the identity

$$\nabla \wedge (\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \wedge \mathbf{v}) = (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_{\text{a}}^{\text{hy}} - (\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \cdot \nabla) \mathbf{v}, \quad (39.36)$$

where we introduced the absolute vorticity for the hydrostatic fluid

$$\boldsymbol{\omega}_{\text{a}}^{\text{hy}} = \mathbf{f} + \boldsymbol{\omega}^{\text{hy}}, \quad (39.37)$$

thus resulting in the vorticity equation

$$\frac{\partial \boldsymbol{\omega}_{\text{a}}^{\text{hy}}}{\partial t} + (\mathbf{v} \cdot \nabla) \boldsymbol{\omega}_{\text{a}}^{\text{hy}} = (\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \cdot \nabla) \mathbf{v} + \nabla \wedge \hat{\mathbf{z}} b + \nabla \wedge \mathbf{F}. \quad (39.38)$$

Since the Coriolis parameter is time independent, we can add it to the time derivative, thus leading to

$$\frac{D\boldsymbol{\omega}_{\text{a}}^{\text{hy}}}{Dt} = \underbrace{(\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \cdot \nabla) \mathbf{v}}_{\text{stretching + tilting}} + \underbrace{\nabla \wedge \hat{\mathbf{z}} b}_{\text{baroclinicity}} + \underbrace{\nabla \wedge \mathbf{F}}_{\text{friction}}. \quad (39.39)$$

Equation (39.39) 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 (37.55)). Both equations have a vorticity source due to stretching and tilting. Furthermore, baroclinicity vector for the hydrostatic Boussinesq fluid is given by $\nabla \wedge \hat{\mathbf{z}} b$, which is simpler than the form $(\nabla \rho \wedge \nabla p)/\rho^2$ (equation (37.57)) for the non-hydrostatic and non-Boussinesq fluid. Hence, we can diagnose the presence of baroclinicity for the hydrostatic Boussinesq fluid merely by noting whether there is a slope to the buoyancy surfaces relative to the horizontal. That is, a sloping buoyancy surface acts as a vorticity source for the hydrostatic Boussinesq fluid. A sloping buoyancy surface is therefore synonymous with a nontrivial baroclinic structure.

39.3.2 Ertel 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

$$\nabla b \cdot \frac{D\boldsymbol{\omega}_{\text{a}}^{\text{hy}}}{Dt} = \nabla b \cdot [(\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \cdot \nabla) \mathbf{v}] + \nabla b \cdot (\nabla \wedge \mathbf{F}), \quad (39.40)$$

where we used

$$\nabla b \cdot (\nabla \wedge \hat{\mathbf{z}} b) = 0. \quad (39.41)$$

We next make use of the identity

$$\frac{D(\partial b / \partial x^i)}{Dt} = \frac{\partial}{\partial x^i} \left[\frac{Db}{Dt} \right] - \nabla b \cdot \frac{\partial \mathbf{v}}{\partial x^i} \quad (39.42a)$$

$$= \frac{\partial \dot{b}}{\partial x^i} - \nabla b \cdot \frac{\partial \mathbf{v}}{\partial x^i}, \quad (39.42b)$$

so that

$$\boldsymbol{\omega}_a^{hy} \cdot \left[\frac{D \nabla b}{Dt} \right] = \boldsymbol{\omega}_a^{hy} \cdot \nabla \dot{b} - \nabla b \cdot [(\boldsymbol{\omega}_a^{hy} \cdot \nabla) \mathbf{v}]. \quad (39.43)$$

Making use of this result in equation (39.40) renders

$$\nabla b \cdot \frac{D \boldsymbol{\omega}_a^{hy}}{Dt} + \boldsymbol{\omega}_a^{hy} \cdot \frac{D \nabla b}{Dt} = \boldsymbol{\omega}_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \wedge \mathbf{F}), \quad (39.44)$$

which leads to

$$\frac{DQ}{Dt} = \boldsymbol{\omega}_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \wedge \mathbf{F}) \quad (39.45)$$

where

$$Q = \boldsymbol{\omega}_a^{hy} \cdot \nabla b = \boldsymbol{\omega}^{hy} \cdot \nabla b + f \frac{\partial b}{\partial z} \quad (39.46)$$

is the Ertel potential vorticity for a rotating hydrostatic Boussinesq fluid. Potential vorticity is materially conserved 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 (39.30). In this way, PV takes on the form

$$Q = \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + (\zeta + f) \frac{\partial b}{\partial z} = \hat{z} \cdot \left[\frac{\partial \mathbf{u}}{\partial z} \wedge \nabla b \right] + (\zeta + f) \frac{\partial b}{\partial z}. \quad (39.47)$$

39.3.3 Potential vorticity flux vector

The material form of the PV equation (39.45) is converted into its Eulerian flux-form via

$$\frac{\partial Q}{\partial t} + \nabla \cdot (\mathbf{v} Q) = \boldsymbol{\omega}_a^{hy} \cdot \nabla \dot{b} + \nabla b \cdot (\nabla \wedge \mathbf{F}) \quad (39.48a)$$

$$= \nabla \cdot [\dot{b} \boldsymbol{\omega}_a^{hy} + b (\nabla \wedge \mathbf{F})] \quad (39.48b)$$

$$= \nabla \cdot [\dot{b} \boldsymbol{\omega}_a^{hy} + \nabla \wedge (b \mathbf{F}) - \nabla b \wedge \mathbf{F}] \quad (39.48c)$$

$$= \nabla \cdot [\dot{b} \boldsymbol{\omega}_a^{hy} - \nabla b \wedge \mathbf{F}], \quad (39.48d)$$

where we used

$$\nabla \cdot \mathbf{v} = 0 \quad (39.49a)$$

$$\nabla \cdot \boldsymbol{\omega}_a^{hy} = 0 \quad (39.49b)$$

$$\nabla \cdot (\nabla \wedge \mathbf{F}) = 0 \quad (39.49c)$$

$$\nabla \cdot [\nabla \wedge (b \mathbf{F})] = 0. \quad (39.49d)$$

The conservation equation (39.48d) allows us to identify a potential vorticity flux vector

$$\mathbf{J} = \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A}, \quad (39.50)$$

so that the PV equation takes the form

$$\frac{\partial Q}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (39.51)$$

The potential vorticity flux (39.50) is comprised of an advective term

$$\mathbf{J}_{\text{advective}} = \mathbf{v} Q, \quad (39.52)$$

and non-advective terms arising from diabatic and frictional forcing,

$$\mathbf{J}_{\text{non-advective}} = -\dot{b}\boldsymbol{\omega}_{\text{a}}^{\text{hy}} + \nabla b \wedge \mathbf{F}, \quad (39.53)$$

as well as a gauge term,

$$\mathbf{J}_{\text{gauge}} = \nabla \wedge \mathbf{A}. \quad (39.54)$$

The gauge term remains arbitrary since the divergence of the curl vanishes, with the arbitrariness known as gauge freedom (see Section 16.7).

39.3.4 Potential vorticity flux vector and the Bernoulli potential

[Schär \(1993\)](#) provided a generalization of Bernoulli's Theorem, and [Vallis \(2006\)](#) offers a brief discussion in Section 4.8.1. [Marshall \(2000\)](#) and [Polton and Marshall \(2007\)](#) apply this theorem in a Boussinesq and hydrostatic ocean. There are some subtle points in the derivation that we wish to raise here, as noted by [Polton and Marshall \(2007\)](#).

Momentum equation

We start by expressing the momentum equation in terms of the Bernoulli potential. For this purpose, return to the horizontal momentum equation (39.35), and expand the expressions for the perturbation pressure and the buoyancy

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{f} + \boldsymbol{\omega}^{\text{hy}}) \wedge \mathbf{v} = -\nabla(\phi + \mathbf{u}^2/2) + \hat{\mathbf{z}} b + \mathbf{F} \quad (39.55\text{a})$$

$$= -\nabla(\mathbf{u}^2/2) - \frac{1}{\rho_0} \nabla(p - p_0) - \hat{\mathbf{z}} \left[\frac{g(\rho - \rho_0)}{\rho_0} \right] + \mathbf{F} \quad (39.55\text{b})$$

$$= -\nabla(\mathbf{u}^2/2) - \frac{1}{\rho_0} \nabla p - \hat{\mathbf{z}} \frac{g\rho}{\rho_0} + \mathbf{F} \quad (39.55\text{c})$$

$$= -\nabla \left[\frac{\mathbf{u}^2}{2} + \frac{p}{\rho_0} \right] - \hat{\mathbf{z}} \left[\frac{g\rho - g\rho_0 + g\rho_0}{\rho_0} \right] + \mathbf{F} \quad (39.55\text{d})$$

$$= -\nabla \left[\frac{\mathbf{u}^2}{2} + \frac{p}{\rho_0} + gz \right] - \hat{\mathbf{z}} \left[\frac{g(\rho - \rho_0)}{\rho_0} \right] + \mathbf{F} \quad (39.55\text{e})$$

$$= -\nabla B + \hat{\mathbf{z}} b + \mathbf{F}, \quad (39.55\text{f})$$

where we introduced the Bernoulli potential for a hydrostatic and Boussinesq fluid

$$B = \frac{\mathbf{u}^2}{2} + \frac{p}{\rho_0} + gz. \quad (39.56)$$

Potential vorticity flux

The flux-form potential vorticity conservation statement remains as given by equation (39.48d), and the PV flux is given by equation (39.50). However, we can make use of the gauge invariance of the PV flux to write the flux in a manner conducive to analyzing steady state conditions, with this approach motivated by the work of [Schär \(1993\)](#). For this purpose, operate with $\nabla b \wedge$ on the velocity equation (39.55f) to have

$$\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} + \nabla b \wedge (\boldsymbol{\omega}_{\text{a}}^{\text{hy}} \wedge \mathbf{v}) = -\nabla b \wedge \nabla B + \nabla b \wedge \hat{\mathbf{z}} b + \nabla b \wedge \mathbf{F}. \quad (39.57)$$

Now make use of the identity

$$\nabla b \wedge (\boldsymbol{\omega}_a^{hy} \wedge \mathbf{v}) = (\nabla b \cdot \mathbf{v}) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \mathbf{v} \quad (39.58)$$

in equation (39.57) to render

$$(\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a^{hy} - (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \mathbf{v} = -\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} - \nabla b \wedge \nabla B + \nabla b \wedge \hat{z} b + \nabla b \wedge \mathbf{F}. \quad (39.59)$$

Now write the PV flux given by equation (39.50) in the form

$$\mathbf{J} = \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (39.60a)$$

$$= \mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) - \left[\frac{\partial b}{\partial t} + \mathbf{v} \cdot \nabla b \right] \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (39.60b)$$

$$= [\mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla b) - (\mathbf{v} \cdot \nabla b) \boldsymbol{\omega}_a^{hy}] - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (39.60c)$$

$$= \left[\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} + \nabla b \wedge \nabla B - \nabla b \wedge \hat{z} b - \nabla b \wedge \mathbf{F} \right] - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \mathbf{A} \quad (39.60d)$$

$$= \nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \nabla B - \nabla b \wedge \hat{z} b + \nabla \wedge \mathbf{A}. \quad (39.60e)$$

$$= \nabla b \wedge \left[\frac{\partial \mathbf{u}}{\partial t} + \nabla B \right] - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla \wedge (\mathbf{A} - \hat{z} b^2/2). \quad (39.60f)$$

Choosing the gauge function according to

$$\mathbf{A} = \hat{z} (b^2/2) \quad (39.61)$$

renders the flux vector

$$\mathbf{J} = \nabla b \wedge \left(\frac{\partial \mathbf{u}}{\partial t} + \nabla B \right) - \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} \quad (39.62a)$$

$$= \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla b \wedge b \hat{z}. \quad (39.62b)$$

The second equality reintroduced the first form of the PV flux given by equation (39.50) and made use of the gauge choice (39.61).

In the steady state, the PV flux (39.62a) reduces to

$$\mathbf{J}^{ss} = \nabla b \wedge \nabla B. \quad (39.63)$$

Hence, the steady state potential vorticity flux is aligned with the intersection of surfaces of constant buoyancy and Bernoulli potential

$$\nabla b \cdot \mathbf{J}^{ss} = 0 \quad (39.64a)$$

$$\nabla B \cdot \mathbf{J}^{ss} = 0. \quad (39.64b)$$

This result is the Boussinesq/hydrostatic form of the more general result from *Schär* (1993).

Integral constraints for steady state

As noted by [Schär \(1993\)](#), [Marshall \(2000\)](#), and [Polton and Marshall \(2007\)](#), the steady state PV flux in the form (39.63) can be utilized to develop some useful integral constraints. Consider the integral of \mathbf{J}^{ss} over an arbitrary simply connected area, and make use of Stokes' Theorem

$$\int_{\mathcal{A}} \mathbf{J}^{\text{ss}} \cdot \hat{\mathbf{n}} \, dA = \int_{\mathcal{A}} (\nabla b \wedge \nabla B) \cdot \hat{\mathbf{n}} \, dA \quad (39.65a)$$

$$= - \int_{\mathcal{A}} [\nabla \wedge (b \nabla B)] \cdot \hat{\mathbf{n}} \, dA \quad (39.65b)$$

$$= - \oint_{\partial\mathcal{A}} b \nabla B \cdot d\mathbf{r} \quad (39.65c)$$

$$= - \oint_{\partial\mathcal{A}} b dB \quad (39.65d)$$

$$= \oint_{\partial\mathcal{A}} B db. \quad (39.65e)$$

The penultimate equality set

$$\nabla B \cdot d\mathbf{r} = dB, \quad (39.66)$$

and the final equality made use of

$$b dB = d(b B) - B db \quad (39.67)$$

and noted that

$$\oint_{\partial\mathcal{A}} d(B b) = 0, \quad (39.68)$$

since $d(B b)$ is a perfect differential. If we can find a closed contour where either the Bernoulli potential is a constant ($dB = 0$), or the buoyancy is a constant ($db = 0$), then we have the constraint

$$\int_{\mathcal{A}} \mathbf{J}^{\text{ss}} \cdot \hat{\mathbf{n}} \, dA = 0 \quad \text{area enclosed by contour with } B \text{ constant or } b \text{ constant.} \quad (39.69)$$

[Marshall \(2000\)](#) and [Polton and Marshall \(2007\)](#) make particular use of closed Bernoulli contours on constant depth surfaces, so that $\hat{\mathbf{n}} = \hat{\mathbf{z}}$. In regions where there are such closed contours, this constraint offers useful insight into the balances.

Integral constraints for non-steady state

Following [Polton and Marshall \(2007\)](#), we make use of the previous results to provide constraints on the non-steady state flow. For this purpose, again consider a closed contour with either B constant or b constant. Integrating over the enclosed area annihilates the $\nabla b \wedge \nabla B$ term as before, thus leaving

$$\int_{\mathcal{A}} \mathbf{J} \cdot \hat{\mathbf{n}} \, dA = \int_{\mathcal{A}} \left[\nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} - \frac{\partial b}{\partial t} \boldsymbol{\omega}_{\text{a}}^{\text{hy}} \right] \cdot \hat{\mathbf{n}} \, dA \quad (39.70a)$$

$$= \int_{\mathcal{A}} \left[\mathbf{v} Q - \dot{b} \boldsymbol{\omega}_{\text{a}}^{\text{hy}} + \nabla b \wedge \mathbf{F} + \nabla b \wedge b \hat{\mathbf{z}} \right] \cdot \hat{\mathbf{n}} \, dA. \quad (39.70b)$$

To reach the second equality, we made use of the two forms of the PV flux given by equations (39.62a) and (39.62b). Rerrangement thus leads to the balance

$$\int_{\mathcal{A}} \left[\mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla b \wedge b \hat{\mathbf{z}} - \nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} \right] \cdot \hat{\mathbf{n}} dA = 0. \quad (39.71)$$

Following [Polton and Marshall \(2007\)](#), we find closed Bernoulli contours on constant depth surfaces, in which case $(\nabla b \wedge b \hat{\mathbf{z}}) \cdot \hat{\mathbf{z}} = 0$ so that

$$\begin{aligned} & \int_{\mathcal{A}(B)} \left[\mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} - \nabla b \wedge \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} \right] \cdot \hat{\mathbf{z}} dA \\ &= \int_{\mathcal{A}(B)} \left[w Q + \boldsymbol{\omega}_a^{hy} \cdot \hat{\mathbf{z}} \left(\frac{\partial b}{\partial t} - \dot{b} \right) + \left[\left(\mathbf{F} - \frac{\partial \mathbf{u}}{\partial t} \right) \wedge \hat{\mathbf{z}} \right] \cdot \nabla b \right] dA \\ &= 0. \end{aligned} \quad (39.72)$$

39.3.5 Impermeability theorem

Equation (39.64a) indicates that the steady state form of the potential vorticity flux (39.63) is parallel to surfaces of constant buoyancy. That is, the steady state PV flux does not penetrate constant buoyancy surfaces. This is a very useful constraint placed on the PV flux, even in the presence of irreversible processes. Remarkably, [Haynes and McIntyre \(1990\)](#) showed that this result can be generalized to the time dependent case with a moving buoyancy surface. We here provide a proof for the hydrostatic and Boussinesq fluid, following the discussion given in Section 38.6 for non-hydrostatic and non-Boussinesq fluids.

Introduce the velocity components

$$\mathbf{v}_{b\parallel} = \mathbf{v} - \hat{\mathbf{n}} (\mathbf{v} \cdot \hat{\mathbf{n}}) \quad \mathbf{v}_{b\perp} = -\hat{\mathbf{n}} \frac{\partial b / \partial t}{|\nabla b|} \quad (39.73)$$

where

$$\hat{\mathbf{n}} = \frac{\nabla b}{|\nabla b|} \quad (39.74)$$

is the unit normal for a buoyancy surface and the parallel and perpendicular velocity components satisfy

$$\mathbf{v}_{b\parallel} \cdot \hat{\mathbf{n}} = 0 \quad (39.75a)$$

$$\mathbf{v}_{b\perp} \cdot \hat{\mathbf{n}} = -\frac{\partial b / \partial t}{|\nabla b|}. \quad (39.75b)$$

These identities then lead to the decomposition of the velocity field

$$\mathbf{v} = \mathbf{v}_{b\parallel} + \mathbf{v}_{b\perp} + \hat{\mathbf{n}} \frac{\dot{b}}{|\nabla b|}. \quad (39.76)$$

The velocity component $\mathbf{v}_{b\parallel}$ is oriented parallel to surfaces of constant buoyancy, whereas the velocity component $\mathbf{v}_{b\perp}$ is perpendicular. Furthermore, the relation (39.73) is satisfied by the velocity of a point on a constant buoyancy surface, which follows since

$$\frac{\partial b}{\partial t} + \mathbf{v}_{b\perp} \cdot \nabla b = 0. \quad (39.77)$$

When the buoyancy surfaces are material, so that $\dot{b} = 0$, then $\mathbf{v}_{b\perp} \cdot \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}}$, in which case there is no mass flux crossing the buoyancy surfaces. More generally, there is a non-zero mass flux crossing buoyancy surfaces due to diabatic processes.² However, as we now show, there remains, identically, zero potential vorticity flux crossing these surfaces.

Choosing the gauge (39.61) and using the velocity decomposition (39.76) renders the potential vorticity flux vector (39.50)

$$\mathbf{J} = \mathbf{v} Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + \nabla \wedge \hat{\mathbf{z}} b^2 / 2 \quad (39.78a)$$

$$= \left[\mathbf{v}_{b\parallel} + \mathbf{v}_{b\perp} + \frac{\dot{b} \nabla b}{|\nabla b|^2} \right] Q - \dot{b} \boldsymbol{\omega}_a^{hy} + \nabla b \wedge \mathbf{F} + b \nabla b \wedge \hat{\mathbf{z}} \quad (39.78b)$$

$$= (\mathbf{v}_{b\parallel} + \mathbf{v}_{b\perp}) Q - \dot{b} \left[\boldsymbol{\omega}_a^{hy} - \frac{(\boldsymbol{\omega}_a^{hy} \cdot \nabla b) \nabla b}{|\nabla b|^2} \right] + \nabla b \wedge (\mathbf{F} + b \hat{\mathbf{z}}) \quad (39.78c)$$

$$= (\mathbf{v}_{b\parallel} + \mathbf{v}_{b\perp}) Q - \dot{b} (\boldsymbol{\omega}_a^{hy})_{b\parallel} + \nabla b \wedge (\mathbf{F} + b \hat{\mathbf{z}}) \quad (39.78d)$$

$$= \mathbf{v}_{b\perp} Q + \left[\mathbf{v}_{b\parallel} Q - \dot{b} (\boldsymbol{\omega}_a^{hy})_{b\parallel} + \nabla b \wedge (\mathbf{F} + b \hat{\mathbf{z}}) \right] \quad (39.78e)$$

$$\equiv \mathbf{J}_{b\perp} + \mathbf{J}_{b\parallel} \quad (39.78f)$$

where

$$(\boldsymbol{\omega}_a^{hy})_{b\parallel} = \boldsymbol{\omega}_a^{hy} - \left[\frac{\boldsymbol{\omega}_a^{hy} \cdot \nabla b}{|\nabla b|^2} \right] \nabla b. \quad (39.79)$$

We therefore write the PV equation in the form

$$\frac{\partial Q}{\partial t} = -\nabla \cdot \mathbf{J} \quad (39.80a)$$

$$= -\nabla \cdot \left[\frac{\mathbf{J}}{Q} Q \right] \quad (39.80b)$$

$$= -\nabla \cdot (\mathbf{v}_Q Q). \quad (39.80c)$$

The velocity $\mathbf{v}_Q = \mathbf{J}/Q$ can be considered an effective velocity that advects a notional ‘‘PV substance’’ through the fluid. This velocity satisfies

$$\mathbf{v}_Q \cdot \hat{\mathbf{n}} = (\mathbf{J}/Q) \cdot \hat{\mathbf{n}} \quad (39.81a)$$

$$= -\frac{\partial b / \partial t}{|\nabla b|} \quad (39.81b)$$

$$= \mathbf{v}_{b\perp} \cdot \hat{\mathbf{n}}. \quad (39.81c)$$

That is,

$$\frac{\partial b}{\partial t} + \mathbf{v}_Q \cdot \nabla b = 0 \quad (39.82)$$

so that no PV substance penetrates a constant buoyancy surface, even as that surface moves and even in the presence of irreversible processes. Instead, the normal projection of the PV flux, divided by the potential vorticity, is identical to the projection of the velocity for a point on the buoyancy surface. This result represents the impermeability theorem of [Haynes and McIntyre \(1990\)](#) as realized for a hydrostatic and Boussinesq fluid.

²See Section 15.7.3 for a discussion of dia-surface mass fluxes in the context of kinematic boundary conditions, and Section 18.3 for a more general discussion of dia-surface transport.

We offer another derivation of the above result, here starting from the alternative form of the PV flux given by equation (39.62a). Projecting the flux (39.62a) onto the direction normal to the buoyancy surface leads to

$$\mathbf{J} \cdot \hat{\mathbf{n}} = -\frac{\partial b}{\partial t} \boldsymbol{\omega}_a^{hy} \cdot \hat{\mathbf{n}} = -Q \frac{\partial b / \partial t}{|\nabla b|}, \quad (39.83)$$

as in equation (39.81b).

39.4 Distinguishing buoyancy and density

The analysis in Section 39.3 ignored the distinction between *in situ* density and potential density. We here follow [Polton and Marshall \(2007\)](#) by developing the PV budget for the more general ocean system. For this purpose, start with the hydrostatic Boussinesq primitive equation in the form

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{v} = -\frac{1}{\rho_0} \nabla_z p + \mathbf{F} \quad (39.84a)$$

$$\frac{\partial p}{\partial z} = -\rho g \quad (39.84b)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (39.84c)$$

$$\frac{D\sigma}{Dt} = \dot{\sigma}, \quad (39.84d)$$

where σ is the potential density referenced to a chosen pressure. Following the manipulations in Section 39.3.4, we write the velocity equation as

$$\frac{\partial \mathbf{u}}{\partial t} + \boldsymbol{\omega}_a^{hy} \wedge \mathbf{v} = -\nabla B + \hat{\mathbf{z}} b + \mathbf{F}, \quad (39.85)$$

where

$$b = -g \left[\frac{\rho - \rho_0}{\rho_0} \right] \quad (39.86)$$

is the buoyancy defined using the *in situ* density. Following the manipulations from Section 39.3.1 leads to the material evolution of vorticity

$$\frac{D\boldsymbol{\omega}_a^{hy}}{Dt} = (\boldsymbol{\omega}_a^{hy} \cdot \nabla) \mathbf{v} + \nabla \wedge \hat{\mathbf{z}} b + \nabla \wedge \mathbf{F}, \quad (39.87)$$

which is the same as equation (39.39).

39.4.1 Potential vorticity

To develop a potential vorticity equation, we follow [Polton and Marshall \(2007\)](#) by taking the projection of the vorticity onto the direction normal to the potential density surface.

$$Q = \boldsymbol{\omega}_a^{hy} \cdot \nabla \sigma. \quad (39.88)$$

Notably, the evolution of this potential vorticity involves the baroclinicity vector $\nabla \wedge \hat{\mathbf{z}} b$. Following the details from Section 39.3.2, we are left with the potential vorticity equation

$$\frac{DQ}{Dt} = \nabla \sigma \cdot (\nabla \wedge \hat{\mathbf{z}} b) + \boldsymbol{\omega}_a^{hy} \cdot \nabla \dot{\sigma} + \nabla \sigma \cdot (\nabla \wedge \mathbf{F}). \quad (39.89)$$

This equation is the same as the PV equation (39.45), yet with the additional contribution from baroclinicity, $\nabla\sigma \cdot (\nabla \wedge \hat{\mathbf{z}} b)$. This term vanishes only when surfaces of constant potential density are parallel to those of constant *in situ* density (see Chapter ??), as for a linear equation of state. However, this term remains for the general case, which means that this PV is not materially conserved in the adiabatic and inviscid case. However, we choose to make use of potential density rather than *in situ* density, largely since ocean mesoscale eddy fluxes are generally aligned with potential density surfaces, or more accurately neutral directions (see Section 20.3).

39.4.2 Potential vorticity flux vector

We now write the material evolution equation (39.89) in a flux form, so that

$$\frac{\partial Q}{\partial t} = -\nabla \cdot (\mathbf{v} Q - \sigma \nabla \wedge \hat{\mathbf{z}} b - \dot{\sigma} \boldsymbol{\omega}_a^{hy} - \sigma \nabla \wedge \mathbf{F}) \quad (39.90a)$$

$$= -\nabla \cdot (\mathbf{v} Q + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) - \dot{\sigma} \boldsymbol{\omega}_a^{hy}). \quad (39.90b)$$

Hence, although the potential vorticity (39.88) is not materially conserved, even in adiabatic and inviscid flow, its local time density is determined by the convergence of a flux

$$\mathbf{J} = \mathbf{v} Q - \dot{\sigma} \boldsymbol{\omega}_a^{hy} + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) + \nabla \wedge \mathbf{A}. \quad (39.91)$$

This flux agrees with equation (39.50), again with the exception of the additional baroclinicity term $\nabla \sigma \wedge \hat{\mathbf{z}} b$.

39.4.3 Potential vorticity flux vector and the Bernoulli potential

Following the manipulations in Section 39.3.4, we can write the potential vorticity flux vector in the form

$$\mathbf{J} = \mathbf{v} Q - \dot{\sigma} \boldsymbol{\omega}_a^{hy} + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) + \nabla \wedge \mathbf{A}. \quad (39.92a)$$

$$= \mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla \sigma) - \left[\frac{\partial \sigma}{\partial t} + \mathbf{v} \cdot \nabla \sigma \right] \boldsymbol{\omega}_a^{hy} + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) + \nabla \wedge \mathbf{A} \quad (39.92b)$$

$$= [\mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla \sigma) - (\mathbf{v} \cdot \nabla \sigma) \boldsymbol{\omega}_a^{hy}] - \frac{\partial \sigma}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) + \nabla \wedge \mathbf{A}. \quad (39.92c)$$

Now make use of the identity (39.59) to write

$$\mathbf{J} = [\mathbf{v} (\boldsymbol{\omega}_a^{hy} \cdot \nabla \sigma) - (\mathbf{v} \cdot \nabla \sigma) \boldsymbol{\omega}_a^{hy}] - \frac{\partial \sigma}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) + \nabla \wedge \mathbf{A} \quad (39.93a)$$

$$= \left[\nabla \sigma \wedge \frac{\partial \mathbf{u}}{\partial t} + \nabla \sigma \wedge \nabla B - \nabla \sigma \wedge \hat{\mathbf{z}} b - \nabla \sigma \wedge \mathbf{F} \right] - \frac{\partial \sigma}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla \sigma \wedge (\hat{\mathbf{z}} b + \mathbf{F}) + \nabla \wedge \mathbf{A} \quad (39.93b)$$

$$= \nabla \sigma \wedge \left[\frac{\partial \mathbf{u}}{\partial t} + \nabla B \right] - \frac{\partial \sigma}{\partial t} \boldsymbol{\omega}_a^{hy} + \nabla \wedge \mathbf{A}. \quad (39.93c)$$

39.4.4 Vector-invariant velocity equation

We take this opportunity to express the velocity equation in a form commonly used to formulate the equations of numerical ocean models. It is also the vector-invariant form used to start in the

development of the vorticity equation (see Section 37.7.1 for the non-hydrostatic version). For this purpose, write the material time derivative in the form

$$\frac{D\mathbf{u}}{Dt} = \frac{\partial \mathbf{u}}{\partial t} + \zeta \hat{\mathbf{z}} \wedge \mathbf{u} + \nabla(\mathbf{u}^2/2) + [-\hat{\mathbf{x}} \partial_z v + \hat{\mathbf{y}} \partial_z u] \wedge \mathbf{v} \quad (39.94a)$$

$$= \frac{\partial \mathbf{u}}{\partial t} + \zeta \hat{\mathbf{z}} \wedge \mathbf{u} + \nabla(\mathbf{u}^2/2) + w \partial_z \mathbf{u} \quad (39.94b)$$

thus leading to the momentum equation

$$\left[\frac{\partial}{\partial t} + w \frac{\partial}{\partial z} + (f + \zeta) \hat{\mathbf{z}} \wedge \right] \mathbf{u} = -\nabla(\phi + \mathbf{u}^2/2) + \hat{\mathbf{z}} b + \mathbf{F}. \quad (39.95)$$

39.5 Vorticity and PV using isopycnal vertical coordinates

In Section 39.3, we show how vorticity in a Boussinesq hydrostatic fluid, 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 solenoidal 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 this desirable feature of buoyancy surfaces by construction. Hence, the derivation of the PV equation given in this section never requires us to eliminate the baroclinicity vector through projecting vorticity onto the direction defined by a chosen scalar field. 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 from the start, thus allowing the derivation to proceed as for the shallow water system in Section 36.3. For the shallow water fluid, there is no baroclinicity since density is uniform. For the continuously stratified fluid, an isopycnal description removes the baroclinicity *a priori* by working within layers of constant buoyancy. The one key assumption needed for this approach is that surfaces of constant buoyancy must remain stably stratified in the vertical.

39.5.1 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 (31.32a) leads to the isopycnal vorticity equation

$$\left[\frac{\partial \tilde{\zeta}_a}{\partial t} \right]_b + (\mathbf{u} \cdot \nabla_b) \tilde{\zeta}_a = -\tilde{\zeta}_a \nabla_b \cdot \mathbf{u} \quad (39.96)$$

where

$$\tilde{\zeta}_a = f + \hat{\mathbf{z}} \cdot (\nabla_b \wedge \mathbf{u}) = f + \tilde{\zeta} \quad (39.97)$$

is the absolute vorticity, written as the planetary vorticity plus the isopycnal relative vorticity. The left hand side of equation (39.96) is the material time derivative of absolute vorticity (see equation (31.23b)), so that we can write

$$\frac{D\tilde{\zeta}_a}{Dt} = -\tilde{\zeta}_a \nabla_b \cdot \mathbf{u}. \quad (39.98)$$

39.5.2 Derivation of the potential vorticity equation

We now make use of the thickness equation in the material form (31.38c) to eliminate the convergence $\nabla_b \cdot \mathbf{u}$ on the right hand side of equation (39.98), thus leading to

$$\frac{D\tilde{\zeta}_a}{Dt} - \frac{\tilde{\zeta}_a}{h} \frac{D\sigma}{Dt} = 0. \quad (39.99)$$

Introducing the isopycnal potential vorticity

$$\tilde{Q} = \frac{\tilde{\zeta}_a}{h} = \frac{f + \tilde{\zeta}}{h} \quad (39.100)$$

leads to

$$\frac{D\tilde{Q}}{Dt} = 0. \quad (39.101)$$

Expanding the material time derivative into its components according to equation (31.23b), and making use of the adiabatic form of the thickness equation (31.32c), leads to the Eulerian flux-form equation

$$\left[\frac{\partial(h\tilde{Q})}{\partial t} \right]_b + \nabla_b \cdot (h\tilde{Q}\mathbf{u}) = 0. \quad (39.102)$$

39.5.3 A note about specific thickness

Our use of the thickness, $h = (\partial z / \partial b) db$, for defining the potential vorticity (39.100) accords with the shallow water case as detailed in Chapter 36. A slightly modified form makes use of the specific thickness

$$h = \frac{\partial z}{\partial z}, \quad (39.103)$$

which, for an adiabatic fluid, satisfies the same equation as the thickness

$$\frac{Dh}{Dt} = -h \nabla_b \cdot \mathbf{u}. \quad (39.104)$$

Hence, PV defined according to the specific thickness

$$Q = \frac{f + \tilde{\zeta}}{h} \quad (39.105)$$

is also materially constant and thus satisfies the Eulerian flux-form equation

$$\left[\frac{\partial(hQ)}{\partial t} \right]_b + \nabla_b \cdot (hQ\mathbf{u}) = 0 \quad (39.106)$$

or equivalently using the specific thickness

$$\left[\frac{\partial(hQ)}{\partial t} \right]_b + \nabla_b \cdot (hQ\mathbf{u}) = 0. \quad (39.107)$$

39.5.4 Coordinate transforming vorticity and potential vorticity

As we just derived, 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

$$(\boldsymbol{\omega}^{\text{hy}} + f \hat{\mathbf{z}}) \cdot \nabla b = -\frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + f \right) \frac{\partial b}{\partial z} \quad (39.108a)$$

$$= \frac{\partial b}{\partial z} \left[f + \left(\frac{\partial v}{\partial x} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} \right) - \left(\frac{\partial u}{\partial y} - \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} \right) \right] \quad (39.108b)$$

$$= \frac{\partial b}{\partial z} \left[f + \left(\frac{\partial v}{\partial x} \right)_b - \left(\frac{\partial u}{\partial y} \right)_b \right] \quad (39.108c)$$

$$= \frac{f + (\partial v / \partial x)_b - (\partial u / \partial y)_b}{\partial z / \partial b} \quad (39.108d)$$

$$= \frac{f + \tilde{\zeta}}{h} \quad (39.108e)$$

$$= Q. \quad (39.108f)$$

In the above, we used the horizontal derivatives on constant buoyancy surfaces

$$\nabla_b = \hat{\mathbf{x}} \left[\frac{\partial}{\partial x} \right]_b + \hat{\mathbf{y}} \left[\frac{\partial}{\partial y} \right]_b \quad (39.109)$$

and defined the isopycnal relative vorticity

$$\tilde{\zeta} = \left[\frac{\partial v}{\partial x} \right]_b - \left[\frac{\partial u}{\partial y} \right]_b = \hat{\mathbf{z}} \cdot (\nabla_b \wedge \mathbf{u}). \quad (39.110)$$

40

Angular momentum, vorticity, and strain[†]

In Section 37.1.2, we showed that the vorticity for solid-body motion is 2Ω , twice the angular rotation rate. Since the angular momentum for solid body motion is also proportional to the angular rotation rate, there is a direct connection between angular momentum and vorticity for solid-body motion. The purpose of this appendix is to generalize that connection for arbitrary fluid motion. As shown here, the generalization requires us to also include strain.

40.1	A resume of point particle mechanics	601
40.1.1	Angular velocity and moment of inertia	602
40.1.2	Relating angular velocity to velocity	602
40.2	Linear momentum for material fluid regions	603
40.3	Angular momentum for material fluid regions	604
40.3.1	Taylor expanding the velocity	604
40.3.2	Strain and vorticity	605
40.3.3	Relating angular momentum to strain and vorticity	606
40.3.4	Comments	606

40.1 A resume of point particle mechanics

The linear momentum of a point particle is given by

$$\mathbf{P} = M \mathbf{V}, \quad (40.1)$$

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,

$$\mathbf{V} = \frac{d\mathbf{X}}{dt}. \quad (40.2)$$

The corresponding angular momentum is given by

$$\mathbf{L} = \mathbf{X} \wedge \mathbf{P} = M (\mathbf{X} \wedge \mathbf{V}). \quad (40.3)$$

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 12.6 that the component of angular momentum about the rotation axis is a constant of the motion.

40.1.1 Angular velocity and moment of inertia

Whereas linear momentum has physical dimensions of

$$[\mathbf{P}] \equiv \text{mass} \times \text{length} \times \text{time}^{-1}, \quad (40.4)$$

angular momentum has dimensions of

$$[\mathbf{L}] \equiv \text{mass} \times \text{length}^2 \times \text{time}^{-1}. \quad (40.5)$$

We can pursue the analog by introducing the angular velocity

$$\boldsymbol{\Omega} = \frac{\mathbf{X} \wedge \mathbf{V}}{|\mathbf{X}|^2}. \quad (40.6)$$

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

$$\boldsymbol{\Omega} \cdot \mathbf{X} = \boldsymbol{\Omega} \cdot \mathbf{V} = 0. \quad (40.7)$$

The angular velocity is not defined at the origin since $|\mathbf{X}| = 0$.

Inserting the definition of the angular velocity (40.6) into the angular momentum (40.3) renders

$$\mathbf{L} = M (\mathbf{X} \wedge \mathbf{V}) \quad (40.8a)$$

$$= M |\mathbf{X}|^2 \boldsymbol{\Omega} \quad (40.8b)$$

$$\equiv I \boldsymbol{\Omega}. \quad (40.8c)$$

In the final equality we introduced the moment of inertia for a point particle

$$I = M |\mathbf{X}|^2. \quad (40.9)$$

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 40.3.3).

40.1.2 Relating angular velocity to velocity

The cross product of the position vector with the angular velocity (40.6) is given by

$$\boldsymbol{\Omega} \wedge \mathbf{X} = \mathbf{V} - \mathbf{X} \frac{\mathbf{V} \cdot \mathbf{X}}{|\mathbf{X}|^2}. \quad (40.10)$$

For the special case of velocity \mathbf{V} orthogonal to the position vector, \mathbf{X} , we have

$$\mathbf{V} = \boldsymbol{\Omega} \wedge \mathbf{X} \quad \text{when } \mathbf{V} \cdot \mathbf{X} = 0. \quad (40.11)$$

In particular, for circular motion the velocity is orthogonal to the position.

40.2 Linear momentum for material fluid regions

We now consider the velocity and linear momentum of a connected material fluid region denoted by \mathcal{R} . Let an arbitrary fluid parcel within this region be marked with the material label \mathbf{a} . The position vector of the parcel is $\mathbf{X}(\mathbf{a}, t)$ and its velocity is¹

$$\mathbf{V}(\mathbf{a}, t) = \frac{D\mathbf{X}(\mathbf{a}, t)}{Dt}. \quad (40.12)$$

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

$$\mathbf{V}(\mathbf{a}, t) = \frac{D\mathbf{X}(\mathbf{a}, t)}{Dt} \quad (40.13a)$$

$$= \frac{D(\bar{\mathbf{X}} + \mathbf{X}')}{Dt} \quad (40.13b)$$

$$= \bar{\mathbf{V}}(t) + \mathbf{V}'(\mathbf{a}, t). \quad (40.13c)$$

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

$$\bar{\mathbf{V}} = \frac{D\bar{\mathbf{X}}}{Dt} \quad (40.14a)$$

$$= \frac{D}{Dt} \left[\frac{\int_{\mathcal{R}} \mathbf{x} \rho dV}{\int_{\mathcal{R}} \rho dV} \right] \quad (40.14b)$$

$$= \frac{1}{M} \int_{\mathcal{R}} \frac{D\mathbf{x}}{Dt} \rho dV \quad (40.14c)$$

$$= \frac{1}{M} \int_{\mathcal{R}} \mathbf{v} \rho dV. \quad (40.14d)$$

The identity (40.14c) follows since the material region maintains a constant mass,

$$M = \int_{\mathcal{R}} \rho dV, \quad (40.15)$$

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 15.3.4), the material derivative moves across the integral to act only on the position vector. The final equality, (40.14d), 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

$$\mathbf{v}(\mathbf{x}, t) = \frac{D\mathbf{x}}{Dt}. \quad (40.16)$$

It follows that the linear momentum for the material fluid region is given by

$$\mathbf{P} = \int_{\mathcal{R}} \mathbf{v} \rho dV = M \bar{\mathbf{V}}. \quad (40.17)$$

We conclude that the total linear momentum of an extended body equals to that of a point particle of mass $M \int_{\mathcal{R}} \rho dV$ moving with the center of mass velocity, $\bar{\mathbf{V}}$.

¹We choose to use the material time derivative notation, D/Dt , even though that notation is redundant when working with Lagrangian coordinates. It is of use here as a reminder of the Lagrangian nature of the formulation.

40.3 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

$$\mathbf{L} = \int_{\mathcal{R}} (\mathbf{x} \wedge \mathbf{v}) \rho dV. \quad (40.18)$$

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

$$\mathbf{L} = \int_{\mathcal{R}} (\mathbf{x} \wedge \mathbf{v}) \rho dV \quad (40.19a)$$

$$= \int_{\mathcal{R}} [(\bar{\mathbf{x}} + \mathbf{x}') \wedge \mathbf{v}] \rho dV \quad (40.19b)$$

$$= \bar{\mathbf{X}} \wedge \left(\int_{\mathcal{R}} \mathbf{v} \rho dV \right) + \int_{\mathcal{R}} (\mathbf{x}' \wedge \mathbf{v}) \rho dV \quad (40.19c)$$

$$= (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}} (\mathbf{x}' \wedge \mathbf{v}) \rho dV. \quad (40.19d)$$

The final equality introduced the linear momentum, (40.17), for the fluid region. The first term in equation (40.19d) 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, $\int_{\mathcal{R}} (\mathbf{x}' \wedge \mathbf{v}) \rho dV$, contributes to the angular momentum (40.18). 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 2, 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 2.4)

$$(\mathbf{A} \wedge \mathbf{B})_m = \epsilon_{mnp} A_n B_p. \quad (40.20)$$

40.3.1 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

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}(\bar{\mathbf{x}} + \mathbf{x}') \quad (40.21a)$$

$$\approx \mathbf{v}(\bar{\mathbf{x}}) + (\mathbf{x}' \cdot \nabla) \mathbf{v}|_{\mathbf{x}=\bar{\mathbf{x}}}. \quad (40.21b)$$

We are thus left with

$$\mathbf{L} = (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}} (\mathbf{x}' \wedge \mathbf{v}) \rho dV \quad (40.22a)$$

$$= (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}} [\mathbf{x}' \wedge \mathbf{v}(\bar{\mathbf{x}})] \rho dV + \int_{\mathcal{R}} [\mathbf{x}' \wedge (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})] \rho dV. \quad (40.22b)$$

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 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 (40.22b) vanishes

$$\int_{\mathcal{R}} [\mathbf{x}' \wedge \mathbf{v}(\bar{\mathbf{x}})] \rho dV = \left[\int_{\mathcal{R}} \mathbf{x}' \rho dV \right] \wedge \mathbf{v}(\bar{\mathbf{x}}) = 0, \quad (40.23)$$

where $\int_{\mathcal{R}} \mathbf{x}' \rho dV = 0$ by definition of the center of mass. The angular momentum is thus given by the two terms

$$\mathbf{L} = (\bar{\mathbf{X}} \wedge \mathbf{P}) + \int_{\mathcal{R}} [\mathbf{x}' \wedge (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})] \rho dV. \quad (40.24)$$

The m' th component of the second term can be written

$$\int_{\mathcal{R}} [\mathbf{x}' \wedge (\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})]_m \rho dV = \epsilon_{mnp} \int_{\mathcal{R}} x'_n [(\mathbf{x}' \cdot \nabla) \mathbf{v}(\bar{\mathbf{x}})]_p \rho dV \quad (40.25a)$$

$$= \epsilon_{mnp} \int_{\mathcal{R}} x'_n x'_q \partial_q v(\bar{\mathbf{x}})_p \rho dV \quad (40.25b)$$

$$= \epsilon_{mnp} \left[\int_{\mathcal{R}} x'_n x'_q \rho dV \right] \partial_q v(\bar{\mathbf{x}})_p. \quad (40.25c)$$

We removed the velocity derivatives

$$\partial_q v(\bar{\mathbf{x}})_p = \left[\frac{\partial v_p}{\partial x_q} \right]_{\mathbf{x}=\bar{\mathbf{x}}} \quad (40.26)$$

from the integral, since they are evaluated at the center of mass point and so do not participate in the integration.

40.3.2 Strain and vorticity

Following from the discussion in Section 17.2.4, we know that the velocity derivatives $\partial_q v_p$ appearing in equation (40.25c) form the components to a second order tensor. To expose the kinematics of this tensor, decompose it into its symmetric and anti-symmetric components

$$\partial_q v_p = \frac{1}{2}(\partial_q v_p + \partial_p v_q) + \frac{1}{2}(\partial_q v_p - \partial_p v_q) \quad (40.27a)$$

$$\equiv S_{qp} + A_{qp}. \quad (40.27b)$$

The symmetric tensor

$$S_{qp} = \frac{1}{2}(\partial_q v_p + \partial_p v_q) \quad (40.28)$$

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

$$2A_{qp} = \partial_q v_p - \partial_p v_q \quad (40.29a)$$

$$= (\delta_{qm} \delta_{pn} - \delta_{qn} \delta_{pm}) \partial_m v_n \quad (40.29b)$$

$$= \epsilon_{sqp} \epsilon_{smn} \partial_m v_n \quad (40.29c)$$

$$= \epsilon_{sqp} \omega_s, \quad (40.29d)$$

where $\omega_s = \epsilon_{smn} \partial_m v_n$ are components to the vorticity pseudo-vector

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v}. \quad (40.30)$$

40.3.3 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

$$L_m = (\bar{\mathbf{X}} \wedge \mathbf{P})_m + \epsilon_{mnp} \left[\int_{\mathcal{R}} x'_n x'_q \rho dV \right] S_{qp} + \epsilon_{mnp} \left[\int_{\mathcal{R}} x'_n x'_q \rho dV \right] A_{qp} \quad (40.31a)$$

$$= (\bar{\mathbf{X}} \wedge \mathbf{P})_m + \epsilon_{mnp} \left[\int_{\mathcal{R}} x'_n x'_q \rho dV \right] S_{qp} + \frac{1}{2} \epsilon_{mnp} \epsilon_{sqp} \left[\int_{\mathcal{R}} x'_n x'_q \rho dV \right] \omega_s \quad (40.31b)$$

$$= \underbrace{(\bar{\mathbf{X}} \wedge \mathbf{P})_m}_{\text{center of mass}} + \underbrace{\epsilon_{mnp} \left[\int_{\mathcal{R}} x'_n x'_q \rho dV \right] S_{qp}}_{\text{strain contribution}} + \underbrace{\frac{1}{2} \left[\int_{\mathcal{R}} (\mathbf{x}' \cdot \mathbf{x}' \delta_{ms} - x'_m x'_s) \rho dV \right] \omega_s}_{\text{vorticity contribution}}. \quad (40.31c)$$

As each point in the fluid can be considered the center of mass for an arbitrary material region, the decomposition (40.31c) is general.

- **CENTER OF MASS ANGULAR MOMENTUM:** The first term on the right hand side of equation (40.31c) 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 (40.8a)). 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 17.2.4). 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. 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 (40.31c) contains the vorticity as weighted by the moment of inertia tensor

$$I_{ms} \equiv \int_{\mathcal{R}} (\mathbf{x}' \cdot \mathbf{x}' \delta_{ms} - x'_m x'_s) \rho dV. \quad (40.32)$$

Since the material region is evolving and is not rigid, the moment of inertia tensor is time dependent. The contribution

$$L_m^{\text{vorticity}} \equiv \frac{1}{2} I_{ms} \omega_s \quad (40.33)$$

has the same form as angular momentum for a rigid body, with one-half the vorticity playing the role of angular velocity (see equation (40.8c) 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.

40.3.4 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 not the same when the fluid has non-zero straining

motion. It is only for the special case of a solid-body motion that the strain contribution to angular momentum vanishes.

The discussion in this appendix is supported by [Chatwin \(1973\)](#), as well as online notes “The Vorticity Equation and Conservation of Angular Momentum” from A.J. DeCaria.

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 30, 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 41. The nuts and bolts of the next two chapters involve methods of scaling analysis and asymptotic analysis via perturbation series. In Chapter 42, 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 Chapter 43. 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.

41

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 35.3). 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

41.1	Governing equations	612
41.2	Vorticity equation	612
41.2.1	Deriving the vorticity equation	613
41.2.2	Potential vorticity	613
41.2.3	Poisson equation for the streamfunction	613
41.3	Pressure	614
41.3.1	Lid pressure	614
41.3.2	Diagnostic relation for the pressure	614
41.4	Further reading	614
41.5	Exercises	615

41.1 Governing equations

The governing equations for 2d barotropic flow are the shallow water equations with the horizontal circulation assumed to be non-divergent

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z \phi \quad (41.1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (41.2)$$

where the pressure is normalized according to

$$\phi = p/\rho_0 \quad (41.3)$$

with ρ_0 a constant reference density, and where material evolution occurs with the two-dimensional non-divergent flow

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla. \quad (41.4)$$

With the horizontal flow $\mathbf{u} = (u, v, 0)$ non-divergent, we can introduce a streamfunction

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi \Rightarrow u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x}. \quad (41.5)$$

The advection operator can thus be written

$$\mathbf{u} \cdot \nabla \zeta = u \partial_x \zeta + v \partial_y \zeta \quad (41.6a)$$

$$= -\partial_y \psi \partial_x \zeta + \partial_x \psi \partial_y \zeta \quad (41.6b)$$

$$= \hat{\mathbf{z}} \cdot (\nabla \psi \wedge \nabla \zeta) \quad (41.6c)$$

$$\equiv J(\psi, \zeta), \quad (41.6d)$$

where J is the Jacobian operator

$$J(A, B) = \hat{\mathbf{z}} \cdot (\nabla A \wedge \nabla B) = \frac{\partial A}{\partial x} \frac{\partial B}{\partial y} - \frac{\partial A}{\partial y} \frac{\partial B}{\partial x}. \quad (41.7)$$

41.2 Vorticity equation

The vertical component of the relative vorticity is given by

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \wedge \mathbf{u}) = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla^2 \psi. \quad (41.8)$$

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

$$f = f_0 + \beta (y - y_0). \quad (41.9)$$

41.2.1 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 (41.1)) to arrive at

$$\frac{\partial}{\partial t} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) + \frac{\partial}{\partial x} [\nabla \cdot (\mathbf{u} v)] - \frac{\partial}{\partial y} [\nabla \cdot (\mathbf{u} u)] + f \nabla \cdot \mathbf{u} + \beta v = 0. \quad (41.10)$$

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

$$\frac{\partial}{\partial x} [\nabla \cdot (\mathbf{u} v)] - \frac{\partial}{\partial y} [\nabla \cdot (\mathbf{u} u)] = \mathbf{u} \cdot \nabla \zeta, \quad (41.11)$$

so that the 2d vorticity equation takes the form

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta + \beta v = 0, \quad (41.12)$$

which takes on the material evolution form

$$\frac{D\zeta}{Dt} = -\beta v. \quad (41.13)$$

41.2.2 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

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot (\mathbf{u} q). \quad (41.14)$$

Similarly, we have the potential vorticity equation

$$\frac{\partial q}{\partial t} = -\nabla \cdot (\mathbf{u} q), \quad (41.15)$$

or the material conservation form

$$\frac{Dq}{Dt} = 0. \quad (41.16)$$

41.2.3 Poisson equation for the streamfunction

Given initial and boundary conditions, the barotropic vorticity equation (41.14) allows us to determine the evolution of vorticity. We can in turn invert the Poisson equation (see Section 1.4.3 for discussion of the Poisson equation)

$$\nabla^2 \psi = \zeta \quad (41.17)$$

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.

41.3 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?

41.3.1 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.

41.3.2 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 (41.1) 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

$$-\nabla^2\phi = \partial_x[\nabla \cdot (\mathbf{u} u)] + \partial_y[\nabla \cdot (\mathbf{u} v)] - f\zeta + \beta u. \quad (41.18)$$

Given sufficient boundary conditions, this elliptic partial differential equation (Section 1.4.3) can be inverted to find the pressure field. In Exercise 29.2, 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.

41.4 Further reading

In exercise 41.2 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 \(2004\)](#)).

41.5 Exercises

EXERCISE 41.1: 2D BAROTROPIC SYSTEM AND GRAVITY WAVES

Are there gravity waves for the 2d barotropic system described in Section 41.1. Why? Hint: recall the discussion of gravity waves for the shallow water system in Section 35.3. A one-sentence answer is sufficient.

EXERCISE 41.2: 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

$$\mathcal{K} = \frac{1}{2} \int_{\mathcal{A}} \mathbf{u} \cdot \mathbf{u} \, dA, \quad (41.19)$$

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

$$\mathcal{C} = \int_{\mathcal{A}} \zeta \, dA. \quad (41.20)$$

- (c) Show that the domain integrated enstrophy is constant in time for f -plane motion ($\beta = 0$)

$$\mathcal{Z}^{(\zeta)} = \int_{\mathcal{A}} \zeta^2 \, dA. \quad (41.21)$$

- (d) Show that the domain integrated potential enstrophy is constant in time even with $\beta \neq 0$

$$\mathcal{Z}^{(q)} = \int_{\mathcal{A}} q^2 \, dA. \quad (41.22)$$

EXERCISE 41.3: 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

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta + \beta y) = -\nu \nabla^4 \zeta, \quad (41.23)$$

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

$$\frac{dC}{dt} = - \oint \left[\psi \frac{\partial q}{\partial s} + \nu \frac{\partial (\nabla^2 \zeta)}{\partial n} \right] ds, \quad (41.24)$$

where s is the arc-length along the boundary of the area, n is a coordinate normal to the boundary, and the integration is oriented counter-clockwise.

EXERCISE 41.4: 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 14.1. 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

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) = 0. \quad (41.25)$$

- (a) Show that the material evolution of the vorticity gradient is given by

$$\frac{D(\nabla \zeta)}{Dt} = -J(\nabla \psi, \zeta). \quad (41.26)$$

- (b) Show that the material evolution of the squared vorticity gradient is given by

$$\frac{D|\nabla \zeta|^2}{Dt} = 2 J(\zeta, \nabla \psi) \cdot \nabla \zeta. \quad (41.27)$$

EXERCISE 41.5: 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 34.5.

As in Section 34.5.1, the relative angular momentum for a region of fluid is given by

$$\mathbf{L} = \int dA \int (\mathbf{x} \wedge \mathbf{v}) \rho dz, \quad (41.28)$$

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

$$\mathbf{L} = \rho H \int_{\mathcal{A}} (\mathbf{x} \wedge \mathbf{u}) dA, \quad (41.29)$$

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

$$L^z = \rho H \int_{\mathcal{A}} \hat{\mathbf{z}} \cdot (\mathbf{x} \wedge \mathbf{u}) dA. \quad (41.30)$$

Show for a simply connected and bounded region, L^z can be written

$$L^z = 2 \rho H \int_{\mathcal{A}} (\psi_b - \psi) dA \quad (41.31)$$

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 16.3.2 that the streamfunction equals to a spatial constant when evaluated along the domain boundary.

42

Shallow water PG and QG

In this chapter we develop the mechanical equations for planetary geostrophy (PG) and quasi-geostrophy (QG) within the shallow water fluid system. This material generally follows that in Chapter 5 of [Vallis \(2017\)](#).

42.1	Scaling analysis and the Buckingham-II theorem	617
42.2	Shallow water equations	618
42.2.1	Dimensional scales	618
42.2.2	Physical dimensions	619
42.2.3	Number of non-dimensional parameters	619
42.2.4	Choosing the non-dimensional parameters	619
42.2.5	Assumed values for the non-dimensional parameters	621
42.2.6	Deformation radius and the free surface undulation scale	622
42.2.7	Non-dimensional shallow water equations	623
42.3	Shallow water planetary geostrophy	625
42.4	Shallow water quasi-geostrophy	625
42.4.1	Quasi-geostrophic scaling	625
42.4.2	Outlining the asymptotic method	626
42.4.3	Zeroth order asymptotic equations	627
42.4.4	First order asymptotic equations	628
42.4.5	Dimensional equations and quasi-geostrophic PV	629
42.5	Exercises	630

42.1 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

$$N_{\text{dimensionless}} = K - R. \quad (42.1)$$

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.

42.2 Shallow water equations

A single-layer of inviscid shallow water fluid is governed by the momentum and thickness equations (Chapter 33)

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad (42.2a)$$

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = 0, \quad (42.2b)$$

where (see Figure 33.1)

$$\eta = \eta_b + h = H + \Delta\eta. \quad (42.3)$$

42.2.1 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 33.1).
- ★ L = horizontal/lateral length scale of motions under consideration. Note that 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

$$c = \sqrt{g H}. \quad (42.4)$$

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 forces, 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.

42.2.2 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

$$M = \int \rho dV [\equiv] L^2 H \rho. \quad (42.5)$$

Relatedly, we are unconcerned with the fluid density, since it is uniform and does not explicitly appear in any of the governing equations.

42.2.3 Number of non-dimensional parameters

The Buckingham-II theorem then says we have

$$N_{\text{dimensionless}} = 9 - 2 = 7 \quad (42.6)$$

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.

42.2.4 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

$$\delta_{\text{vertical/horizontal}} = \frac{\text{vertical depth scale}}{\text{horizontal length scale}} = \frac{H}{L}. \quad (42.7)$$

2. RATIO OF HORIZONTAL SCALE TO PLANETARY SCALE: The ratio of the lateral length scale to the planet radius is

$$\delta_{\text{horizontal/planet}} = \frac{\text{lateral length scale}}{\text{planetary length scale}} = \frac{L}{R}. \quad (42.8)$$

3. RATIO OF FREE SURFACE UNDULATION TO DEPTH: The ratio of the free surface undulation scale to the vertical depth scale is

$$\delta_{\text{free surface/depth}} = \frac{\text{free surface undulation scale}}{\text{vertical depth scale}} = \frac{\mathcal{H}}{H}. \quad (42.9)$$

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

$$\delta_{\text{bottom/free surface}} = \frac{\text{bottom topography undulation scale}}{\text{free surface undulation scale}} = \frac{\mathcal{B}}{\mathcal{H}}. \quad (42.10)$$

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

$$Fr = \frac{U}{c} = \frac{U}{\sqrt{g H}}. \quad (42.11)$$

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

$$Ro = \frac{\text{parcel acceleration}}{\text{Coriolis acceleration}}. \quad (42.12)$$

The particle acceleration scale is determined by the local time tendency and advection

$$\text{particle acceleration} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \quad (42.13a)$$

$$\sim \frac{U}{T} + \frac{U^2}{L}. \quad (42.13b)$$

We assume that the time scale is determined by advection, so that

$$T \sim \frac{L}{U} \Rightarrow \frac{U^2}{L} = \frac{U}{T}, \quad (42.14)$$

in which case the Rossby number is given by

$$Ro = \frac{1}{f T} = \frac{U}{f L}. \quad (42.15)$$

Another interpretation for the Rossby number is the ratio of the relative vorticity to the planetary vorticity

$$Ro = \frac{\text{relative vorticity}}{\text{planetary vorticity}} \quad (42.16)$$

With the relative vorticity scaling as U/L and the planetary vorticity scaling as f , we recover the expression (42.15) for the Rossby number.

7. GEOSTROPHIC NUMBER: We define the geostrophic number as the ratio of the Coriolis acceleration to the pressure gradient acceleration¹

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}}. \quad (42.17)$$

The Coriolis acceleration scales as

$$\text{Coriolis acceleration} \sim f U \quad (42.18)$$

whereas the pressure gradient acceleration, $-g \nabla \eta$, scales as

$$\text{pressure gradient acceleration} \sim \frac{g \mathcal{H}}{L}, \quad (42.19)$$

so that

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}} = \frac{f U}{(g/L) \mathcal{H}}. \quad (42.20)$$

42.2.5 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

$$\delta_{\text{vertical/horizontal}} \ll 1. \quad (42.21)$$

This assumption was made when formulating the shallow water system, which is based on hydrostatic balance (see Section 33.1). 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

$$\delta_{\text{horizontal/planet}} \ll 1 \quad \text{quasi-geostrophy} \quad (42.22a)$$

$$\delta_{\text{horizontal/planet}} \sim 1 \quad \text{planetary geostrophy}. \quad (42.22b)$$

3. RATIO OF FREE SURFACE UNDULATION TO DEPTH: This ratio will be implied by other scaling assumptions along with the dynamical equations.

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.

4. SMALL 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

$$\delta_{\text{bottom/free surface}} \ll 1. \quad (42.23)$$

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

$$Ro = \frac{U}{f L} \ll 1, \quad (42.24)$$

which means that the Coriolis acceleration is a leading order term in the horizontal momentum equation (42.2a).

7. UNIT GEOSTROPHIC NUMBER: The geostrophic number is assumed to be order unity

$$Ge \sim 1. \quad (42.25)$$

This assumption means that the Coriolis acceleration and pressure gradient acceleration scale together

$$f U \sim (g/L) \mathcal{H}. \quad (42.26)$$

Making use of the momentum equation (42.2a), 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 42.2.6.

42.2.6 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, (42.26), to express the free surface undulation scale according to

$$\Delta\eta \sim \mathcal{H} = \frac{f U L}{g} = Ro \frac{f^2 L^2}{g} = Ro H \frac{f^2 L^2}{g H} = Ro H \left[\frac{L}{L_d} \right]^2. \quad (42.27)$$

In the final equality, we introduced the deformation radius

$$L_d = \frac{\sqrt{g H}}{f}. \quad (42.28)$$

We encountered the deformation radius when discussing geostrophic adjustment in Section 35.5. 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

$$Fr = \frac{U}{\sqrt{g H}} = \frac{U}{f L_d} = Ro \frac{L}{L_d}. \quad (42.29)$$

Furthermore, the squared ratio of the deformation radius to the lateral length scale is termed the Burger number

$$F^{-1} = Bu = \left[\frac{L_d}{L} \right]^2. \quad (42.30)$$

Use of the Burger number allows us to write the Froude number as

$$Fr = \frac{Ro}{\sqrt{Bu}} = Ro \sqrt{F} \quad (42.31)$$

and the free surface height undulation scale as

$$\mathcal{H} = H Ro \left[\frac{L}{L_d} \right]^2 = H \frac{Ro}{Bu} = H Ro F = H \frac{Fr^2}{Ro}. \quad (42.32)$$

Hence, the ratio of the free surface undulations to the depth scale is given by

$$\delta_{\text{free surface/depth}} = \frac{\mathcal{H}}{H} = Ro \left[\frac{L}{L_d} \right]^2 = \frac{Ro}{Bu} = Ro F = \frac{Fr^2}{Ro}. \quad (42.33)$$

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.

42.2.7 Non-dimensional shallow water equations

To non-dimensionalize the shallow water equations, introduce non-dimensional variables, denoted by a hat, according to

$$t = T \hat{t} \quad (x, y) = L(\hat{x}, \hat{y}) \quad \partial_t = T^{-1} \partial_{\hat{t}} \quad \nabla = L^{-1} \hat{\nabla} \quad (42.34a)$$

$$(u, v) = U(\hat{u}, \hat{v}) \quad f = f_o \hat{f} \quad \Delta\eta = \mathcal{H} \hat{\eta} \quad \eta = H + \Delta\eta = H + \mathcal{H} \hat{\eta}. \quad (42.34b)$$

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 (42.2a) renders

$$\frac{U}{T} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + \frac{U^2}{L} (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{\mathbf{u}} + f_o U (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = - \frac{g \mathcal{H}}{L} \hat{\nabla} \hat{\eta}. \quad (42.35)$$

As before, we assume the time scale is given by the advection time

$$T = \frac{L}{U} = \frac{1}{R_o f_0}, \quad (42.36)$$

so that dividing by $f_0 U$ leads to

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{\mathbf{u}} \right] + (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = - \left[\frac{g H}{f_o L U} \frac{Ro}{Bu} \right] \hat{\nabla} \hat{\eta}, \quad (42.37)$$

where we set $\mathcal{H} = H(Ro/Bu)$ according to equation (42.33). We reduce the factor on the right hand side according to

$$\frac{gH}{f_o LU} \frac{Ro}{Bu} = \frac{gH}{f_o LU} \frac{U}{f_o L} \frac{L^2}{L_d^2} = \frac{gH}{f_o LU} \frac{U}{f_o L} \frac{L^2 f_o^2}{gH} = 1. \quad (42.38)$$

Hence, the non-dimensional inviscid shallow water momentum equation takes on the rather elegant form

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{\mathbf{u}} \right] + \hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla} \hat{\eta}. \quad (42.39)$$

Introducing the non-dimensional material time derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}} \cdot \hat{\nabla} \quad (42.40)$$

brings the momentum equation to

$$Ro \frac{D\hat{\mathbf{u}}}{Dt} + \hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla} \hat{\eta}. \quad (42.41)$$

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 (42.2b) can be written as

$$\frac{\partial \Delta \eta}{\partial t} + (H + \Delta \eta - \eta_b) \nabla \cdot \mathbf{u} + (\mathbf{u} \cdot \nabla) (H + \Delta \eta - \eta_b) = 0, \quad (42.42)$$

where we wrote (see Figure 33.1)

$$h = H + \Delta \eta - \eta_b. \quad (42.43)$$

Our assumed scaling for the bottom topography, (42.23), allows us to drop η_b from this equation. We also note that H is a constant, and so has a zero gradient. Hence, the continuity equation takes the form

$$\frac{\partial \Delta \eta}{\partial t} + (H + \Delta \eta) \nabla \cdot \mathbf{u} + (\mathbf{u} \cdot \nabla) \Delta \eta = 0. \quad (42.44)$$

Introduction of the dimensionless variables leads to

$$H \frac{Ro}{Bu} \left[\frac{1}{T} \frac{\partial}{\partial \hat{t}} + \frac{U}{L} \hat{\mathbf{u}} \cdot \hat{\nabla} \right] \hat{\eta} + \frac{UH}{L} \left[1 + \frac{Ro}{Bu} \hat{\eta} \right] \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (42.45)$$

With the time scale set by advection, $T = L/U$, we have

$$\frac{Ro}{Bu} \frac{D\hat{\eta}}{Dt} + \left[1 + \frac{Ro}{Bu} \hat{\eta} \right] \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (42.46)$$

Equivalently, introducing the inverse Burger number, $F = Bu^{-1}$ renders

$$F Ro \frac{D\hat{\eta}}{Dt} + (1 + F Ro \hat{\eta}) \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (42.47)$$

42.3 Shallow water planetary geostrophy

We now make use of the non-dimensional equations derived in Section 42.2.7 to derive the dynamical equations for planetary geostrophy and quasi-geostrophy. For planetary geostrophy, we drop the parcel acceleration term from the momentum equation (42.41), given that it is order Rossby number smaller than the Coriolis and pressure gradient accelerations. This assumption leads to the geostrophic balance

$$\hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla} \hat{\eta}. \quad (42.48)$$

For the mass continuity equation (42.47), we must make an assumption about the ratio of the Rossby number to the Burger number. For planetary geostrophy, we assume

$$Ro \sim Bu. \quad (42.49)$$

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

$$L \gg L_d. \quad (42.50)$$

In dimensional form, the planetary geostrophic equations take the form

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad (42.51a)$$

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} \quad (42.51b)$$

$$\eta = \eta_b + h. \quad (42.51c)$$

These equations are equivalent to the material conservation equation for the shallow water planetary potential vorticity

$$\frac{DQ}{Dt} = 0, \quad (42.52)$$

where

$$Q = \frac{f}{h} \quad \text{planetary geostrophy.} \quad (42.53)$$

We prove this assertion in exercise 42.1.

42.4 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.

42.4.1 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 (42.33), 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.

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.

42.4.2 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 42.2.7, and make use of the assumed $Bu \sim 1$ scaling

$$Ro \frac{D\hat{\mathbf{u}}}{Dt} + (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = -\hat{\nabla}\hat{\eta} \quad (42.54a)$$

$$F Ro \frac{D\hat{\eta}}{Dt} + (1 + F Ro \hat{\eta}) \hat{\nabla} \cdot \hat{\mathbf{u}} = 0. \quad (42.54b)$$

We maintain the dimensionless parameter

$$F \equiv Bu^{-1} \quad (42.55)$$

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

$$\hat{u} = \hat{u}_0 + Ro \hat{u}_1 + Ro^2 \hat{u}_2 + \dots \quad (42.56a)$$

$$\hat{v} = \hat{v}_0 + Ro \hat{v}_1 + Ro^2 \hat{v}_2 + \dots \quad (42.56b)$$

$$\hat{\eta} = \hat{\eta}_0 + Ro \hat{\eta}_1 + Ro^2 \hat{\eta}_2 + \dots \quad (42.56c)$$

In addition to expanding the prognostic variables, we expand the non-dimensional Coriolis parameter in terms of the Rossby number

$$\hat{\mathbf{f}} = \frac{\mathbf{f}}{f_0} \quad (42.57a)$$

$$= \frac{(f_0 + \beta y) \hat{\mathbf{z}}}{f_0} \quad (42.57b)$$

$$\equiv (\hat{f}_0 + Ro \hat{\beta} \hat{y}) \hat{\mathbf{z}}, \quad (42.57c)$$

where $\hat{\mathbf{z}}$ is the unit vector in the vertical,

$$\hat{\beta} \hat{y} = \frac{\beta y}{Ro f_0} = T \beta y, \quad (42.58)$$

and

$$\hat{f}_0 = \frac{f_0}{f_0} = 1. \quad (42.59)$$

Although $\hat{f}_0 = 1$, it is useful to retain this term as a placeholder. The Coriolis expression (42.57c), in particular the assumed scaling (42.58), 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 27.3).

The practical goal of asymptotic analysis is to develop a closed set of prognostic equations for functions appearing in the asymptotic expansions (42.56a)-(42.57c). 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.

42.4.3 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 (42.56a)-(42.56c) into the non-dimensional partial differential equations (42.54a) and (42.54b). 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

$$Ro \frac{D\hat{\mathbf{u}}_0}{Dt} + (\hat{f}_0 + Ro \hat{\beta} \hat{y}) \hat{\mathbf{z}} \wedge (\hat{\mathbf{u}}_0 + Ro \hat{\mathbf{u}}_1) = -\hat{\nabla}(\hat{\eta}_0 + Ro \hat{\eta}_1) \quad (42.60a)$$

$$F Ro \frac{D\hat{\eta}_0}{Dt} + \hat{\nabla} \cdot \hat{\mathbf{u}}_0 + Ro \hat{\nabla} \cdot \hat{\mathbf{u}}_1 + F Ro \hat{\eta}_0 \hat{\nabla} \cdot \hat{\mathbf{u}}_0 = 0. \quad (42.60b)$$

Terms balancing at order Ro^0 are given by

$$\hat{f}_0 \hat{\mathbf{z}} \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla} \hat{\eta}_0 \quad (42.61a)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}}_0 = 0, \quad (42.61b)$$

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 (42.61a) leads to the non-divergence condition (42.61b). Given the non-divergence condition (42.61b), the zeroth order

velocity field can be written in terms of a streamfunction

$$\hat{u}_0 = -\frac{\partial \hat{\psi}_0}{\partial \hat{y}} \quad \hat{v}_0 = \frac{\partial \hat{\psi}_0}{\partial \hat{x}} \quad \hat{\zeta}_0 = \hat{\nabla}^2 \hat{\psi}_0, \quad (42.62)$$

where the zeroth order streamfunction is the ratio of the zeroth order surface height to zeroth order Coriolis parameter

$$\hat{\psi}_0 = \frac{\hat{\eta}_0}{\hat{f}_0}, \quad (42.63)$$

and we introduced the zeroth order vorticity, $\hat{\zeta}_0$, which will appear in subsequent steps.

42.4.4 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

$$\frac{D_0 \hat{u}_0}{Dt} + \hat{f}_0 \hat{z} \wedge \hat{u}_1 + \hat{\beta} \hat{y} \hat{z} \wedge \hat{u}_0 = -\hat{\nabla} \hat{\eta}_1 \quad (42.64a)$$

$$F \frac{D_0 \hat{\eta}_0}{Dt} + \hat{\nabla} \cdot \hat{u}_1 = 0. \quad (42.64b)$$

At this order, the material time derivative makes use of *only* the zeroth order geostrophic horizontal velocity

$$\frac{D_0}{Dt} = \frac{\partial}{\partial \hat{t}} + \hat{u}_0 \cdot \hat{\nabla}. \quad (42.65)$$

The set of first order equations (42.64a) and (42.64b) 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 (42.64a) eliminates the pressure gradient, $\hat{\nabla} \hat{\eta}_1$, thus producing the vorticity equation

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{u}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \hat{\nabla} \cdot \hat{u}_1. \quad (42.66)$$

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

$$\frac{D_0 (\hat{\zeta}_0 + \hat{\beta} \hat{y})}{Dt} = -\hat{f}_0 \hat{\nabla} \cdot \hat{u}_1. \quad (42.67)$$

As anticipated, we need one more step to close the system, since the evolution of zeroth order vorticity in equations (42.66) and (42.67) 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 (42.64b), thus leading to a prognostic equation involving just zeroth order terms

$$\frac{\partial (\hat{\zeta}_0 + \hat{\beta} \hat{y} - F \hat{f}_0 \hat{\eta}_0)}{\partial \hat{t}} + (\hat{u}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y} - F \hat{f}_0 \hat{\eta}_0) = 0, \quad (42.68)$$

which can be written in the material form

$$\frac{D_0}{Dt} \left[\hat{\zeta}_0 + \hat{\beta} \hat{y} - F \hat{f}_0 \hat{\eta}_0 \right] = 0. \quad (42.69)$$

Finally, we introduce the geostrophic streamfunction $\hat{\psi}_0 = \hat{\eta}_0 / \hat{f}_0$ (equation (42.63)) to render

$$\frac{D_0}{Dt} \left[\hat{\nabla}^2 \hat{\psi}_0 + \hat{\beta} \hat{y} - F \hat{f}_0^2 \hat{\psi}_0 \right] = 0. \quad (42.70)$$

Equations (43.103) and (42.70) 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 (42.65)).

42.4.5 Dimensional equations and quasi-geostrophic PV

The material conservation equation (42.70) 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 42.2.7, so that

$$\hat{t} = T^{-1} t \quad (\hat{x}, \hat{y}) = L^{-1} (x, y) \quad \partial_{\hat{t}} = T \partial_t \quad \hat{\nabla} = L \nabla \quad (42.71a)$$

$$(\hat{u}, \hat{v}) = U^{-1} (u, v) \quad \hat{f}_0 = \frac{f_0}{f_0} \quad \hat{\beta} \hat{y} = \frac{\beta y}{Ro f_0} = \frac{L}{U} \beta y \quad (42.71b)$$

$$\hat{\eta} = \mathcal{H}^{-1} \Delta \eta \quad \mathcal{H} = H F Ro \quad \hat{\zeta} = \frac{L}{U} \zeta = L^2 \nabla^2 \psi, \quad (42.71c)$$

where the quasi-geostrophic streamfunction is

$$\psi = \frac{g \Delta \eta}{f_0}. \quad (42.72)$$

Note that [Vallis \(2017\)](#) defines the streamfunction as

$$\psi_{\text{vallis}} = \frac{g \eta}{f_0} = \frac{g (H + \Delta \eta)}{f_0}, \quad (42.73)$$

which differs by the constant $g H / f_0$. There is no difference in the dynamics, since a streamfunction is defined only up to a constant. Also note that for QG, we take $F \sim 1$.

We now make use of these relations in the non-dimensional quasi-geostrophic potential vorticity (QGPV) equation (43.103)

$$\hat{q} = \hat{\zeta}_0 + \hat{\beta} \hat{y} - \hat{f}_0 \hat{\eta}_0 \quad (42.74a)$$

$$= \frac{L}{U} (\zeta + \beta y) - \frac{\Delta \eta}{\mathcal{H}} \quad (42.74b)$$

$$= \frac{L}{U} (\zeta + \beta y) - \frac{\Delta \eta}{H Ro} \quad (42.74c)$$

$$= \frac{L}{U} \left[\zeta + \beta y - \frac{f_0 \Delta \eta}{H} \right] \quad (42.74d)$$

$$= \frac{L}{U} \left[\zeta + \beta y - \frac{g \Delta \eta}{f_0} \frac{1}{L_d^2} \right] \quad (42.74e)$$

$$= \frac{L}{U} [\zeta + \beta y - L_d^{-2} \psi]. \quad (42.74f)$$

Multiplying both sides by f_0 leads to the dimensionful QGPV for the shallow water fluid layer

$$q = Ro f_0 \hat{q} = \zeta + \beta y - L_d^{-2} \psi. \quad (42.75)$$

42.5 Exercises

EXERCISE 42.1: PV CONSERVATION FOR PG

Show that the planetary geostrophic equations

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad (42.76a)$$

$$\frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u} \quad (42.76b)$$

$$\eta = \eta_b + h \quad (42.76c)$$

are equivalent to

$$\mathbf{f} \wedge \mathbf{u} = -g \nabla \eta \quad (42.77)$$

$$\frac{DQ}{Dt} = 0 \quad (42.78)$$

$$\eta = \eta_b + h \quad (42.79)$$

$$Q = \frac{f}{h}. \quad (42.80)$$

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.

EXERCISE 42.2: CONSTRAINTS ON STEADY STATE PLANETARY GEOSTROPHIC FLOW

Consider a shallow water fluid satisfying the planetary geostrophic equations developed in Section 42.3. 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 33.1 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 42.3: LINEARIZED SHALLOW WATER PV

The potential vorticity for a shallow water layer is given by

$$Q = \frac{\zeta + f}{h}. \quad (42.81)$$

It is materially constant when the flow is inviscid

$$\frac{DQ}{Dt} = 0. \quad (42.82)$$

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

$$\frac{D}{Dt} \left[\zeta + \beta y - \frac{f_0 \Delta\eta}{H} \right] = 0. \quad (42.83)$$

Hint: read Section 43.5.1.

- (b) Using f -plane geostrophic balance, obtain an expression for ζ in terms of η .

EXERCISE 42.4: CONSTRAINT ON f -PLANE GEOSTROPHIC FLOW FROM BOTTOM TOPOGRAPHY
Throughout the quasi-geostrophic scaling for the shallow water in Section 42.2, we assumed the bottom topography has a tiny amplitude relative to the undulations of the free surface height (see equation (42.23)). 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.

43

Continuously stratified PG and QG

In this chapter, we extend the shallow water discussions in Chapter 42 to continuously stratified fluids. We make use of stratified geophysical fluid dynamics from Chapters 26 and 30, as well as potential vorticity from Chapter 38. Material in this chapter is a supplement to Chapter 5 in [Vallis \(2017\)](#).

43.1	Continuously stratified Boussinesq fluid	634
43.1.1	Dimensional parameters	634
43.1.2	Physical dimensions and non-dimensional parameters	635
43.1.3	Choosing the non-dimensional parameters	635
43.1.4	Relating the buoyancy scale to the Coriolis acceleration scale	637
43.1.5	Richardson number	637
43.1.6	The Rossby deformation radius	638
43.1.7	Assumed values for the non-dimensional parameters	638
43.1.8	Non-dimensional Boussinesq equations	639
43.2	Planetary geostrophy for continuously stratified Boussinesq	640
43.2.1	Depth integrated vorticity budget	641
43.2.2	Depth integrated flow and f/H contours	643
43.3	Quasi-geostrophy for continuously stratified Boussinesq	645
43.3.1	Zeroth order asymptotic equations	646
43.3.2	First order asymptotic equations	647
43.3.3	Dimensional QG-PV equation	648
43.3.4	Properties of the steady state flow field	649
43.3.5	Constant background buoyancy frequency	650
43.3.6	Buoyancy advection at the boundaries	650
43.4	Dimensions of various forms for potential vorticity	651
43.5	Connecting QG-PV to Ertel PV	652
43.5.1	Shallow water layer	652
43.5.2	Continuously stratified hydrostatic Boussinesq fluid	652
43.6	Energetics of a continuously stratified QG fluid	654
43.6.1	Kinetic energy	654
43.6.2	Available potential energy	655
43.6.3	Energy conversion	656
43.6.4	Scaling APE and KE	656
43.7	Exercises	657

43.1 Continuously stratified Boussinesq fluid

Our starting point is the adiabatic stratified hydrostatic Boussinesq equations (Section 29.1)

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \wedge \mathbf{u} = -\nabla_z \phi \quad (43.1a)$$

$$\frac{\partial \phi}{\partial z} = b \quad (43.1b)$$

$$\frac{Db}{Dt} = 0 \quad (43.1c)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (43.1d)$$

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 (units 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

$$b = \tilde{b}(z) + b'(x, y, z, t), \quad (43.2)$$

and introduce the corresponding background squared buoyancy frequency

$$N^2 = \frac{\partial \tilde{b}}{\partial z}. \quad (43.3)$$

With this decomposition, the buoyancy equation (43.1c) takes the form

$$\frac{Db'}{Dt} + w N^2 = 0. \quad (43.4)$$

We also introduce an associated decomposition of the hydrostatic pressure

$$\phi = \tilde{\phi}(z) + \phi'(x, y, z, t) \quad (43.5)$$

where $\tilde{\phi}$ is hydrostatically balanced by \tilde{b}

$$\frac{\partial \tilde{\phi}}{\partial z} = \tilde{b}, \quad (43.6)$$

and the fluctuating pressure, ϕ' , is hydrostatically balanced by b'

$$\frac{\partial \phi'}{\partial z} = b'. \quad (43.7)$$

Hence, the scale for the fluctuating pressure, Φ , is related to the scale for the fluctuating buoyancy, B , and the depth scale, H

$$\Phi/H = B. \quad (43.8)$$

We return to this scaling relation, and others, in the following.

43.1.1 Dimensional parameters

Following the shallow water discussion in Section 42.2.1, 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' (units pressure divided by density = length scale \times acceleration).
- ★ B = scale of buoyancy fluctuations b' (units 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

Note that we dropped the wave speed for present purposes, since it does not affect the asymptotics. We also dropped the bottom topography scale, assuming it is small for present purposes.

43.1.2 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

$$N_{\text{dimensionless}} = 9 - 2 = 7 \quad (43.9)$$

non-dimensional parameters.

43.1.3 Choosing the non-dimensional parameters

Following the shallow water discussion in Section 42.2.4, 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

$$\delta_{\text{vertical/horizontal}} = \frac{\text{vertical depth scale}}{\text{horizontal length scale}} = \frac{H}{L}. \quad (43.10)$$

2. RATIO OF HORIZONTAL SCALE TO PLANETARY SCALE: The ratio of the lateral length scale to the planet radius is given by

$$\delta_{\text{horizontal/planet}} = \frac{\text{lateral length scale}}{\text{planetary length scale}} = \frac{L}{R}. \quad (43.11)$$

3. RATIO VERTICAL TO HORIZONTAL VELOCITY SCALE: The ratio of the vertical to horizontal velocity is given by

$$\frac{\text{vertical velocity scale}}{\text{horizontal velocity scale}} = \frac{W}{U}. \quad (43.12)$$

4. HYDROSTATIC NUMBER: The hydrostatic number is the ratio of the pressure gradient scale to the buoyancy scale. For the hydrostatic fluid fluctuations

$$\frac{\Phi}{H} = B, \quad (43.13)$$

where B is the scale for the buoyancy fluctuations. We encountered this relation earlier in equation (43.8).

5. ROSSBY NUMBER: The Rossby number is the ratio of the fluid parcel acceleration scale to the Coriolis acceleration

$$Ro = \frac{\text{parcel acceleration}}{\text{Coriolis acceleration}} = \frac{U}{fL}, \quad (43.14)$$

where we again assume time scales advectively

$$T \sim \frac{L}{U}. \quad (43.15)$$

6. GEOSTROPHIC NUMBER: The ratio of the Coriolis acceleration to the pressure gradient acceleration defines the geostrophic number

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}}. \quad (43.16)$$

The Coriolis acceleration scales as

$$\text{Coriolis acceleration} \sim fU \quad (43.17)$$

whereas the pressure gradient acceleration from the fluctuating pressure, ϕ' , scales as

$$\text{pressure gradient acceleration} \sim \frac{\Phi}{L}, \quad (43.18)$$

so that

$$Ge = \frac{\text{Coriolis acceleration}}{\text{pressure gradient acceleration}} = \frac{fU}{(\Phi/L)}. \quad (43.19)$$

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

$$\frac{\text{fluctuating buoyancy frequency}}{\text{background buoyancy frequency}} = \frac{B/H}{N^2}. \quad (43.20)$$

43.1.4 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 (43.13) we have

$$B = \frac{\Phi}{H}. \quad (43.21)$$

Additionally, assuming geostrophic scaling, equation (43.19) means that the fluctuating pressure has a scale related to the Coriolis acceleration scale according to

$$\Phi = f U L. \quad (43.22)$$

Hence, the scale for the fluctuating buoyancy is given by

$$B = \frac{f U L}{H}. \quad (43.23)$$

43.1.5 Richardson number

The *Richardson number* is the non-dimensional ratio of the squared buoyancy frequency to the squared vertical shear of the horizontal velocity

$$Ri = \frac{N^2}{|\partial_z \mathbf{u}|^2}. \quad (43.24)$$

In regions where $Ri \ll 1$, the vertical shear is strong and the flow tends to be unstable to *Kelvin-Helmholtz 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

$$Ri = \frac{N^2}{(U/H)^2}, \quad (43.25)$$

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

$$Bu = \left[\frac{L_d}{L} \right]^2 = \left[\frac{N H}{f L} \right]^2 = \frac{U^2 Ri}{U^2/(Ro)^2} = (Ro)^2 Ri. \quad (43.26)$$

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 (43.26) thus means that the Richardson number scales as

$$Ri \sim (Ro)^{-2} \quad \text{QG flow.} \quad (43.27)$$

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.

43.1.6 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

$$L_d = H \frac{N}{f}. \quad (43.28)$$

The ratio f/N appears frequently in rotating/stratified fluids, and is sometimes called the Prandtl ratio

$$\frac{f}{N} = \text{Prandtl ratio}. \quad (43.29)$$

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 \(2017\)](#)).

43.1.7 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 42.2.5.

1. SMALL VERTICAL TO HORIZONTAL ASPECT RATIO: The aspect ratio is generally small for large-scale atmospheric and oceanic fluid systems

$$\delta_{\text{vertical/horizontal}} \ll 1. \quad (43.30)$$

This assumption was made when making the hydrostatic approximation (Section 27.2).

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

$$\delta_{\text{horizontal/planet}} \ll 1 \quad \text{quasi-geostrophy} \quad (43.31a)$$

$$\delta_{\text{horizontal/planet}} \sim 1 \quad \text{planetary geostrophy}. \quad (43.31b)$$

3. SMALL RATIO VERTICAL TO HORIZONTAL VELOCITY SCALE: The continuity equation implies

$$\frac{W}{H} = \frac{U}{L}, \quad (43.32)$$

so that

$$W = U \frac{H}{L}. \quad (43.33)$$

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 (43.1b) means that the scales for a buoyancy fluctuation and pressure fluctuation are related by (see equation (43.8))

$$\Phi = H B. \quad (43.34)$$

5. SMALL ROSSBY NUMBER: The Rossby number is assumed small

$$Ro = \frac{U}{f L} = \frac{1}{f T} \ll 1, \quad (43.35)$$

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

$$Ge \sim 1, \quad (43.36)$$

which means that the Coriolis acceleration and pressure gradient acceleration scale together

$$f U \sim \frac{\Phi}{L}. \quad (43.37)$$

This scaling is consistent with the momentum equation (43.1a) 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

$$\frac{B/H}{N^2} = \frac{\Phi}{H^2 N^2} = \frac{f U L}{H^2 N^2} = \frac{U}{f L} \frac{L^2 f^2}{H^2 N^2} = Ro \frac{L^2}{L_d^2}, \quad (43.38)$$

where we introduced the deformation radius for the continuously stratified system

$$L_d = H \frac{N}{f}. \quad (43.39)$$

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 in Section 43.3.3.

43.1.8 Non-dimensional Boussinesq equations

Following the shallow water approach in Section 42.2.7, we introduce non-dimensional variables according to

$$t = T \hat{t} \quad (x, y) = L (\hat{x}, \hat{y}) \quad \partial_t = T^{-1} \partial_{\hat{t}} \quad \nabla_z = L^{-1} \hat{\nabla}_z \quad \partial_z = H^{-1} \partial_{\hat{z}} \quad (43.40a)$$

$$(u, v) = U (\hat{u}, \hat{v}) \quad w = W \hat{w} \quad f = f_0 \hat{f} \quad \phi' = f_0 U L \hat{\phi} \quad b' = B \hat{b} = \left[\frac{f_0 U L}{H} \right] \hat{b}. \quad (43.40b)$$

For the second equality in the buoyancy scale, we made use of the relation (43.23) to connect the buoyancy fluctuation scale to the Coriolis acceleration scale. We also make use of the following relations between scales

$$T = \frac{L}{U} \quad W = \frac{U H}{L} \quad Ro = \frac{U}{f_0 L} = \frac{1}{T f_0}. \quad (43.41)$$

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 (43.1a) renders

$$\frac{U}{T} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + \frac{U^2}{L} (\hat{\mathbf{u}} \cdot \hat{\nabla}_z) \hat{\mathbf{u}} + \frac{WU}{H} \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} + f_o U (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = -f_0 U \hat{\nabla}_z \hat{\phi}. \quad (43.42)$$

Hence, dividing by $f_o U$ leads to

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}_{\hat{z}}) \hat{\mathbf{u}} + \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} \right] + (\hat{\mathbf{f}} \wedge \hat{\mathbf{u}}) = -\hat{\nabla}_z \hat{\phi}. \quad (43.43)$$

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

$$\frac{\partial \hat{\phi}}{\partial \hat{z}} = \hat{b}, \quad (43.44)$$

and the non-dimensional continuity equation is

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0. \quad (43.45)$$

The buoyancy equation (43.4) requires a bit more work to non-dimensionalize. The material time derivative takes the form

$$\frac{Db'}{Dt} = \frac{B}{T} \frac{D\hat{b}}{D\hat{t}} = \frac{U}{L} \frac{f_0 U L}{H} \frac{D\hat{b}}{D\hat{t}} = \frac{f_0 U^2}{H} \frac{D\hat{b}}{D\hat{t}}, \quad (43.46)$$

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

$$N^2 w = N^2 W \hat{w} = N^2 U (H/L) \hat{w} = L_d^2 \frac{U f_0^2}{H L} \hat{w}, \quad (43.47)$$

where we introduced the deformation scale, $L_d = HN/f$, from equation (43.39). Bringing these two pieces together leads to

$$Ro F \frac{D\hat{b}}{D\hat{t}} + \hat{w} = 0, \quad (43.48)$$

where we introduced the Burger number

$$Bu = F^{-1} = \left[\frac{L_d}{L} \right]^2. \quad (43.49)$$

43.2 Planetary geostrophy for continuously stratified Boussinesq

Just like for the shallow water model in Section 42.3, 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

$$F Ro = Ro/Bu \sim 1, \quad (43.50)$$

or

$$L^2 \sim L_d^2 Ro^{-1}. \quad (43.51)$$

With this scaling, and with the Rossby number small, the momentum equation (43.43) reduces to geostrophic balance. The continuity and buoyancy equations do not reduce at all. Hence, in dimensional form, the adiabatic planetary geostrophic equations for a stratified Boussinesq fluid take the form

$$\frac{Db'}{Dt} + w N^2 = 0 \quad (43.52a)$$

$$\mathbf{f} \wedge \mathbf{u} = -\nabla \phi' \quad (43.52b)$$

$$\frac{\partial \phi'}{\partial z} = b' \quad (43.52c)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (43.52d)$$

We could just as well write these equations in terms of the full buoyancy $b = \tilde{b}(z) + b'$, in which

$$\frac{Db}{Dt} = 0 \quad (43.53a)$$

$$\mathbf{f} \wedge \mathbf{u} = -\nabla \phi \quad (43.53b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (43.53c)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (43.53d)$$

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 (43.52b). This situation contrasts with the QG approach, where it is only the horizontal advection that contributes to material time evolution at leading order (Section 43.3).

43.2.1 Depth integrated vorticity budget

As in Section 30.4, we here introduce a vertical stress divergence to the planetary geostrophic system, as well as the potential for diabatic processes, so that the equations of motion are

$$\rho_o f (\hat{\mathbf{z}} \wedge \mathbf{u}) = -\nabla p - \rho g \hat{\mathbf{z}} + \frac{\partial \boldsymbol{\tau}}{\partial z} \quad (43.54a)$$

$$\nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (43.54b)$$

$$\frac{Db}{Dt} = \dot{b}. \quad (43.54c)$$

In Section 30.4.2 we derived the vorticity equation for the planetary geostrophic system, whose vertical component is given by equation (30.36)

$$\rho_o \beta v = \frac{\partial}{\partial z} [\rho_o f w + \hat{\mathbf{z}} \cdot (\nabla \wedge \boldsymbol{\tau})]. \quad (43.55)$$

Vertical integration from the ocean bottom at $z = -H(x, y)$ to free surface at $z = \eta(x, y, t)$ leads to

$$\rho_o \beta V = \rho_o f [w(\eta) - w(-H)] + \hat{\mathbf{z}} \cdot (\nabla \wedge \Delta \boldsymbol{\tau}), \quad (43.56)$$

where

$$\Delta\tau = \tau(\eta) - \tau(-H) \quad (43.57)$$

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.

Bottom kinematics and dynamics

The bottom kinematic boundary condition (Section 15.7.1) leads to

$$w(-H) = -\mathbf{u}(-H) \cdot \nabla H. \quad (43.58)$$

Evaluating the horizontal momentum equation (30.21a) at the ocean bottom yields

$$\rho_o f \hat{\mathbf{z}} \wedge \mathbf{u}(-H) = -\nabla p_b + \frac{\partial \tau(-H)}{\partial z}, \quad (43.59)$$

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

$$\rho_o f \hat{\mathbf{z}} \wedge \mathbf{u}(-H) = -\nabla p_b. \quad (43.60)$$

Note that all velocity components vanish at the bottom when imposing a no-slip bottom boundary condition. We thus consider $\mathbf{u}(-H)$ as the horizontal geostrophic velocity within the bottom boundary layer arising from the bottom pressure gradient.

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 16.2) yields

$$w(\eta) = -\frac{Q_m}{\rho_0} + \frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta, \quad (43.61)$$

in which the steady state balance is¹

$$w(\eta) = -\frac{Q_m}{\rho_0} + \mathbf{u} \cdot \nabla \eta. \quad (43.62)$$

Evaluating the horizontal momentum equation (30.21a) at the ocean surface renders

$$\rho_o f \hat{\mathbf{z}} \wedge \mathbf{u}(\eta) = -\nabla p_a + \frac{\partial \tau(\eta)}{\partial z}, \quad (43.63)$$

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

$$\rho_o f \hat{\mathbf{z}} \wedge \mathbf{u}(\eta) = -\nabla p_a. \quad (43.64)$$

¹Even for transient solutions, the balance in equation (43.61) is largely that found in equation (43.62) 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.

²We 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 Section 4.10 of [Kundu et al. \(2012\)](#)), in which case $\partial \tau(\eta) / \partial z = 0$, thus leading to equation (43.64).

Meridional transport balances

Plugging expressions (43.64) and (43.60) into the depth integrated balance (43.56) renders the depth integrated meridional transport

$$\rho_o \beta V = \nabla p_a \wedge \nabla \eta + \nabla p_b \wedge \nabla H + \hat{z} \cdot (\nabla \wedge \Delta \boldsymbol{\tau}) - f Q_m. \quad (43.65)$$

Note that 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

$$p_b = \rho_0 g H + p'_b \quad (43.66)$$

so that

$$\rho_o \beta V = \nabla p_a \wedge \nabla \eta + \nabla p'_b \wedge \nabla H + \hat{z} \cdot (\nabla \wedge \Delta \boldsymbol{\tau}) - f Q_m. \quad (43.67)$$

This equation shows that the steady state depth integrated meridional transport is balanced by four terms: gradients in the atmospheric pressure and sea level; gradients in the anomalous bottom pressure and bottom topography; differences between the turbulent stresses at the ocean top and bottom; and boundary mass fluxes. There are interesting cases where only some or just one of the right hand side terms dominates. For example, it is common to assume a rigid lid for the upper ocean boundary condition, in which case we drop the surface pressure and surface mass flux. In this case a meridional transport with a flat bottom is driven solely by differences between the surface and bottom stresses, whereas topographic variations can produce sizable meridional flows that are balanced by the bottom term $\nabla p'_b \wedge \nabla H$.

43.2.2 Depth integrated flow and f/H contours

As shown in Section 42.3, 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. We discover much about potential vorticity later in this book. Here, we note that when the flow is steady and 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 f/h PV. The quasi-stationary approximation is a bit more general than the rigid lid introduced in Section 43.2.1, 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).

Massaging the depth integrated momentum budget

The depth integrated horizontal momentum equation (43.54a) is given by

$$\rho_o f \hat{z} \wedge \mathbf{U} = - \int_{-H}^{\eta} \nabla_z p \, dz + \Delta \boldsymbol{\tau} \quad (43.68)$$

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

$$\int_{-H}^{\eta} p \, dz = \int_{-H}^{\eta} [d(pz) - z \, dp] = p_a \eta + p_b H + \int_{-H}^{\eta} g \rho z \, dz, \quad (43.69)$$

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

$$\mathcal{P} = \int_{-H}^{\eta} g \rho z dz. \quad (43.70)$$

These results then lead to the depth integrated horizontal pressure gradient³

$$\int_{-H}^{\eta} \nabla_z p dz = \nabla \int_{-H}^{\eta} p dz - p_a \nabla \eta - p_b \nabla H \quad (43.71a)$$

$$= \nabla [p_a \eta + p_b H + \mathcal{P}] - p_a \nabla \eta - p_b \nabla H \quad (43.71b)$$

$$= \eta \nabla p_a + H \nabla p_b + \nabla \mathcal{P}, \quad (43.71c)$$

which in turn renders the depth integrated horizontal momentum balance

$$\rho_o f \hat{z} \wedge \mathbf{U} = -\eta \nabla p_a - H \nabla p_b - \nabla \mathcal{P} + \Delta \boldsymbol{\tau}. \quad (43.72)$$

The momentum balance is here written in terms of pressure form drags on the top and bottom interfaces of the fluid column (see Section 34.2 for more discussion of pressure form drag), the gradient of the potential energy per area, and the difference in stress at the top and bottom.

Flow relative to f/H

Dividing the depth integrated momentum equation (43.72) by the depth H and taking the curl leads to

$$\nabla \cdot (\mathbf{U} f/H) = \hat{z} \cdot [\nabla(\eta/H) \wedge \nabla p_a + \nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)]. \quad (43.73)$$

Performing the chain rule on the left hand side leads to

$$\mathbf{U} \cdot \nabla(f/H) + (f/H) \nabla \cdot \mathbf{U} = \hat{z} \cdot [\nabla(\eta/H) \wedge \nabla p_a + \nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)]. \quad (43.74)$$

Exercise 16.2 indicates that the column volume budget for an incompressible fluid leads to the free surface time tendency

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho_0} - \nabla \cdot \mathbf{U}. \quad (43.75)$$

Assuming a steady state then leads to boundary mass fluxes balancing a divergence in the depth integrated flow, in which case

$$\mathbf{U} \cdot \nabla(f/H) = -\frac{f}{H} \frac{Q_m}{\rho_0} + \hat{z} \cdot [\nabla(\eta/H) \wedge \nabla p_a + \nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)]. \quad (43.76)$$

This balance indicates that there are many sources for the depth integrated flow to deviate from contours of constant f/H .

³We write $\nabla_z p$ inside the integral of equation (43.71a), since p is a function of depth. However, we use the more concise ∇ in equations (43.71b) and (43.71c) since all terms are only spatial functions of the horizontal coordinates x, y meaning that ∇ reduces to ∇_z .

Special case of flow respecting 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

$$\mathbf{U} = \hat{\mathbf{z}} \wedge \nabla \Psi. \quad (43.77)$$

The balance (43.76) in turn takes the form

$$[\nabla \Psi \wedge \nabla(f/H)] \cdot \hat{\mathbf{z}} = [\nabla(1/H) \wedge \nabla \mathcal{P} - \nabla \wedge (\Delta \boldsymbol{\tau}/H)] \cdot \hat{\mathbf{z}}, \quad (43.78)$$

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 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 $\nabla(1/H) \wedge \nabla \mathcal{P} = 0$.

43.3 Quasi-geostrophy for continuously stratified Boussinesq

We proceed much like for the single layer of shallow water fluid in Section 42.4. In particular, quasi-geostrophic scaling from Section 42.4.1 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 43.1.8

$$Ro \left[\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}_{\hat{z}}) \hat{\mathbf{u}} + \hat{w} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{z}} \right] + \hat{\mathbf{f}} \wedge \hat{\mathbf{u}} = -\hat{\nabla}_{\hat{z}} \hat{\phi} \quad (43.79a)$$

$$\frac{\partial \hat{\phi}}{\partial \hat{z}} = \hat{b} \quad (43.79b)$$

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0 \quad (43.79c)$$

$$Ro F \frac{D\hat{b}}{Dt} + \hat{w} = 0. \quad (43.79d)$$

We expand the prognostic variables in an asymptotic series in Rossby number

$$\hat{u} = \hat{u}_0 + Ro \hat{u}_1 + Ro^2 \hat{u}_2 + \dots \quad (43.80a)$$

$$\hat{v} = \hat{v}_0 + Ro \hat{v}_1 + Ro^2 \hat{v}_2 + \dots \quad (43.80b)$$

$$\hat{w} = \hat{w}_0 + Ro \hat{w}_1 + Ro^2 \hat{w}_2 + \dots \quad (43.80c)$$

$$\hat{b} = \hat{b}_0 + Ro \hat{b}_1 + Ro^2 \hat{b}_2 + \dots \quad (43.80d)$$

along with the expansion (42.57c) for the Coriolis parameter

$$\hat{\mathbf{f}} = (\hat{f}_0 + Ro \hat{\beta} \hat{y}) \hat{\mathbf{k}}, \quad (43.81)$$

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 (42.58))

$$\hat{\beta} \hat{y} = \frac{\beta y}{Ro f_0} = T \beta y. \quad (43.82)$$

Furthermore, under quasi-geostrophic scaling it is important to retain the depth dependence of the Burger number through its dependence on the background stratification $N^2(z)$

$$F(z) = \left[\frac{L}{L_d} \right]^2 = \frac{1}{[N(z)]^2} \left[\frac{L f_0}{H} \right]^2. \quad (43.83)$$

Additionally, the Burger number is order unity since the horizontal length scales are on the order of the deformation radius

$$Bu = F^{-1} \sim 1 \Rightarrow L \sim L_d. \quad (43.84)$$

43.3.1 Zeroth order asymptotic equations

The zeroth order asymptotic equations take the form

$$\hat{\mathbf{f}}_0 \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla}_z \hat{\phi}_0 \quad (43.85a)$$

$$\frac{\partial \hat{\phi}_0}{\partial \hat{z}} = \hat{b}_0 \quad (43.85b)$$

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_0 + \frac{\partial \hat{w}_0}{\partial \hat{z}} = 0 \quad (43.85c)$$

$$\hat{w}_0 = 0. \quad (43.85d)$$

The first equation represents f -plane geostrophy, which means that the horizontal velocity has zero divergence

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_0 = 0. \quad (43.86)$$

Equation (43.85b) means the zeroth order buoyancy determines the zeroth order hydrostatic pressure. Since the horizontal velocity has zero divergence, the continuity equation (43.85c) means that the vertical velocity is depth independent

$$\frac{\partial \hat{w}_0}{\partial \hat{z}} = 0. \quad (43.87)$$

If it vanishes somewhere, such as a solid boundary, then it vanishes everywhere. This is a manifestation of the Taylor-Proudman theorem (see Section 30.4.3). Indeed, a vanishing \hat{w}_0 is required by the zeroth-order buoyancy equation (43.85d). Hence, the non-dimensional velocity has a nonzero contribution only at order Ro^1

$$\hat{w} = Ro \hat{w}_1 + Ro^2 \hat{w}_2 + \dots, \quad (43.88)$$

thus manifesting the vertical stiffening of fluid columns found in rotating fluids. Hence, the dimensionful vertical velocity has the asymptotic expansion

$$w = W \hat{w} = WRo (\hat{w}_1 + Ro \hat{w}_2 + \dots), \quad (43.89)$$

so that to leading to order Ro^1

$$\hat{w}_1 = \frac{w}{WRo}. \quad (43.90)$$

Since the zeroth-order velocity is non-divergent, we can introduce a geostrophic streamfunction

$$\hat{u}_0 = -\frac{\partial \hat{\psi}_0}{\partial \hat{y}} \quad \hat{v}_0 = \frac{\partial \hat{\psi}_0}{\partial \hat{x}} \quad \hat{\zeta}_0 = \hat{\nabla}^2 \hat{\psi}_0, \quad (43.91)$$

where the zeroth-order streamfunction is the ratio of the zeroth order pressure to zeroth order Coriolis parameter

$$\hat{\psi}_0 = \frac{\hat{\phi}_0}{\hat{f}_0}. \quad (43.92)$$

Note also that the zeroth-order system satisfies the thermal wind balance

$$\hat{f}_0 \hat{z} \wedge \frac{\partial \hat{\mathbf{u}}_0}{\partial \hat{z}} = -\hat{\nabla}_z \hat{b}_0. \quad (43.93)$$

Finally, note that the zeroth order buoyancy is related to the streamfunction through the hydrostatic balance

$$\hat{b}_0 = \frac{\partial \hat{\phi}_0}{\partial \hat{z}} = \hat{f}_0 \frac{\partial \hat{\psi}_0}{\partial \hat{z}}. \quad (43.94)$$

43.3.2 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

$$\frac{D_0 \hat{\mathbf{u}}_0}{D\hat{t}} + \hat{f}_0 \hat{\mathbf{k}} \wedge \hat{\mathbf{u}}_1 + \hat{\beta} \hat{y} \hat{\mathbf{k}} \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla}_z \hat{\phi}_1 \quad (43.95a)$$

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_1 + \frac{\partial \hat{w}_1}{\partial \hat{z}} = 0 \quad (43.95b)$$

$$F \frac{D_0 \hat{b}_0}{D\hat{t}} + \hat{w}_1 = 0. \quad (43.95c)$$

At this order, the material time derivative makes use *only* of the zeroth order geostrophic horizontal velocity

$$\frac{D_0}{D\hat{t}} = \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla}. \quad (43.96)$$

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 (43.95a) eliminates the pressure gradient, $\hat{\nabla} \hat{\phi}_1$, thus producing the vorticity equation

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \hat{\nabla}_z \cdot \hat{\mathbf{u}}_1. \quad (43.97)$$

We make use of the continuity equation (43.95b) to eliminate the horizontal convergence

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = \hat{f}_0 \frac{\partial \hat{w}_1}{\partial \hat{z}}. \quad (43.98)$$

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 (43.95c). When doing so, it is important to keep the $F(z)$ depth dependence according to equation (43.83), with this depth dependence arising from the background stratification. The resulting vorticity equation is

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = -\hat{f}_0 \frac{\partial}{\partial \hat{z}} \left[F \frac{D_0 \hat{b}_0}{D\hat{t}} \right]. \quad (43.99)$$

We now use the identity

$$\frac{\partial}{\partial \hat{z}} \left[F(\hat{z}) \frac{D_0 \hat{b}_0}{D\hat{t}} \right] = \frac{\partial}{\partial \hat{z}} \left[F \left(\frac{\partial}{\partial \hat{t}} + \hat{\mathbf{u}}_0 \cdot \hat{\nabla} \right) \hat{b}_0 \right] \quad (43.100a)$$

$$= \frac{D_0}{D\hat{t}} \left[\frac{\partial}{\partial \hat{z}} (F \hat{b}_0) \right] + F \frac{\partial \hat{\mathbf{u}}_0}{\partial \hat{z}} \cdot \hat{\nabla}_z \hat{b}_0 \quad (43.100b)$$

$$= \frac{D_0}{D\hat{t}} \left[\frac{\partial}{\partial \hat{z}} (F \hat{b}_0) \right], \quad (43.100c)$$

where we set

$$\frac{\partial \hat{\mathbf{u}}_0}{\partial \hat{z}} \cdot \hat{\nabla}_z \hat{b}_0 = 0 \quad (43.101)$$

since the zeroth-order velocity maintains thermal wind balance (43.93). Bringing terms together then leads to the material conservation equation for quasi-geostrophic potential vorticity

$$\frac{D_0}{D\hat{t}} \left[\hat{\zeta}_0 + \hat{\beta} \hat{y} + \hat{f}_0 \frac{\partial}{\partial \hat{z}} (F \hat{b}_0) \right] = 0. \quad (43.102)$$

43.3.3 Dimensional QG-PV equation

From equation (43.102), we identify the non-dimensional QG-PV

$$\hat{q} = \hat{\zeta} + \hat{\beta} \hat{y} + \hat{f}_0 \frac{\partial (F \hat{b}_0)}{\partial \hat{z}}, \quad (43.103)$$

where we dropped the 0 asymptotic subscript for brevity. Introducing dimensional quantities to the right hand side yields (recall $\hat{f}_0 = 1$ and $\hat{\beta} \hat{y} = T \beta y = (L/U) \beta y$)

$$\hat{q} = \frac{L}{U} (\zeta + \beta y) + \frac{\partial}{\partial z} \left(\frac{H F b'}{B} \right). \quad (43.104)$$

The scale for the fluctuating buoyancy is given by equation (43.23)

$$B = \frac{f_0 U L}{H}, \quad (43.105)$$

and the inverse Burger number is given by equation (43.83)

$$F(z) = \left[\frac{L}{L_d} \right]^2 = \frac{1}{[N(z)]^2} \left[\frac{L f_0}{H} \right]^2. \quad (43.106)$$

These terms then yield for the non-dimensional PV

$$\hat{q} = \frac{L}{U} (\zeta + \beta y) + \frac{\partial}{\partial z} \left(\frac{H F b'}{B} \right) \quad (43.107a)$$

$$= \frac{L}{U} (\zeta + \beta y) + \frac{1}{Ro} \left[\frac{\partial}{\partial z} \left(\frac{b'}{N^2} \right) \right] \quad (43.107b)$$

$$= \frac{L}{U} (\nabla_z^2 \psi + \beta y) + \frac{f_0}{Ro} \left[\frac{\partial}{\partial z} \left(\frac{1}{N^2} \frac{\partial \psi}{\partial z} \right) \right], \quad (43.107c)$$

where we introduced the QG streamfunction

$$\psi = \frac{\phi'}{f_0} \quad (43.108)$$

for the final equality, and made use of the identities

$$u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x} \quad \zeta = \nabla_z^2 \psi \quad b' = f_0 \left(\frac{\partial \psi}{\partial z} \right). \quad (43.109)$$

We thus identify the dimensional QG-PV

$$q = Ro f_0 \hat{q} \quad (43.110a)$$

$$= \zeta + \beta y + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] \quad (43.110b)$$

$$= \nabla_z^2 \psi + \beta y + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right]. \quad (43.110c)$$

We can add a constant to q without changing the dynamics. Consequently, some authors like to add f_0 in which case

$$q = \underbrace{\nabla_z^2 \psi}_{\text{relative vorticity}} + \underbrace{f_0 + \beta y}_{\text{planetary vorticity}} + \underbrace{\frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right]}_{\text{stretching by } f}. \quad (43.111)$$

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 (43.98) to see why the third term represents stretching).

43.3.4 Properties of the steady state flow field

The geostrophic velocity takes the following form in terms of the QG streamfunction

$$\mathbf{u} = \hat{\mathbf{z}} \wedge \nabla \psi. \quad (43.112)$$

This equality then allows us to write the following equivalent forms for the material time derivative of QG-PV

$$\frac{Dq}{Dt} = \frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q \quad (43.113a)$$

$$= \frac{\partial q}{\partial t} + (\hat{\mathbf{z}} \wedge \nabla \psi) \cdot \nabla q \quad (43.113b)$$

$$= \frac{\partial q}{\partial t} + (\nabla \psi \wedge \nabla q) \cdot \hat{\mathbf{z}} \quad (43.113c)$$

$$= \frac{\partial q}{\partial t} + J(\psi, q), \quad (43.113d)$$

where the final equality introduced the Jacobian operator

$$J(\psi, q) = \frac{\partial \psi}{\partial x} \frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial q}{\partial x}. \quad (43.114)$$

For a perfect fluid, in which $Dq/Dt = 0$, a steady state (zero Eulerian time derivative) is realized when

$$\mathbf{u} \cdot \nabla q = (\nabla\psi \wedge \nabla q) \cdot \hat{\mathbf{z}} = J(\psi, q) = 0. \quad (43.115)$$

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

$$\psi = \psi(q) \Rightarrow J(\psi, q) = 0. \quad (43.116)$$

43.3.5 Constant background buoyancy frequency

Consider the QG-PV for the special case of a constant background buoyancy frequency,

$$N^2 = \text{constant}. \quad (43.117)$$

For this case the QG-PV in equation (43.110c) can be written

$$q = \beta y + \nabla_z^2 \psi + \frac{\partial}{\partial z} \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] \quad (43.118a)$$

$$= \beta y + \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{f_0^2}{N^2} \frac{\partial^2 \psi}{\partial z^2} \quad (43.118b)$$

$$= \beta y + \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial \tilde{z}^2}. \quad (43.118c)$$

For the final equality we introduced the vertical coordinate

$$\tilde{z} = \frac{N}{f} z. \quad (43.119)$$

Since $N/f \gg 1$ for the stably stratified flows considered in QG, we term \tilde{z} a *stretched* vertical coordinate.

43.3.6 Buoyancy advection at the boundaries

We need boundary conditions in order to invert the elliptic QG-PV equation (43.111) 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 tangential flow to zero for the inviscid case. The top and bottom boundaries are less trivial and require some care.

Returning to the Boussinesq equations (43.1a)-(43.1c), we focus on the buoyancy equation, written as in equation (43.4)

$$\frac{Db'}{Dt} + w N^2 = 0, \quad (43.120)$$

where advection is via the horizontal geostrophic currents. Inserting the geostrophic streamfunction

$$b' = f_0 \frac{\partial \psi}{\partial z} \quad (43.121)$$

leads to

$$\frac{\partial}{\partial t} \left(\frac{\partial \psi}{\partial z} \right) + \mathbf{u} \cdot \nabla \left(\frac{\partial \psi}{\partial z} \right) + w N^2 = 0. \quad (43.122)$$

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

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \Rightarrow w = \mathbf{u} \cdot \nabla \eta_b, \quad (43.123)$$

where $\hat{\mathbf{n}}$ is the outward normal to the boundary (Section 15.7.1). The second contribution arises from Ekman pumping or suction (Section 30.5). 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 \(2006\)](#))

$$w_{\text{Ekman}} = r \nabla_z^2 \psi, \quad (43.124)$$

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

$$w = \mathbf{u} \cdot \nabla \eta_b + r \nabla_z^2 \psi. \quad (43.125)$$

Using this expression in the buoyancy equation (43.122) leads to the boundary evolution of buoyancy

$$\frac{\partial}{\partial t} \left(\frac{\partial \psi}{\partial z} \right) + \mathbf{u} \cdot \nabla \left(\frac{\partial \psi}{\partial z} + N^2 \eta_b \right) + N^2 r \nabla_z^2 \psi = 0. \quad (43.126)$$

We say that the QG fluid system is characterized by horizontal advection of PV in the interior, and advection of buoyancy on the boundaries.

43.4 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:

$$\frac{\zeta + f}{h} [\equiv] (\text{time} \times \text{length})^{-1}. \quad (43.127)$$

- Entropic Ertel PV based on potential temperature,

$$\frac{\omega_a \cdot \nabla \theta}{\rho} [\equiv] \frac{\text{length}^2 \times \text{temperature}}{\text{time} \times \text{mass}}, \quad (43.128)$$

where θ is the potential temperature.

- Ertel PV with buoyancy in a Boussinesq fluid

$$\omega_a \cdot \nabla b [\equiv] \text{time}^{-3}. \quad (43.129)$$

- Quasi-geostrophic PV

$$q = \nabla_z^2 \psi + \beta y + \frac{\partial}{\partial z} \left(\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right) [\equiv] \text{time}^{-1}. \quad (43.130)$$

43.5 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 38.

43.5.1 Shallow water layer

The Ertel PV for a single layer of shallow water fluid is (Section 36.3)

$$Q = \frac{f + \zeta}{h}, \quad (43.131)$$

where

$$h = H + \Delta\eta - \eta_b \quad (43.132)$$

is the thickness of the layer (see Figure 33.1). 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

$$Q_{\text{PG}} = \frac{f}{h}. \quad (43.133)$$

The QG limit requires a bit more algebra, whereby

$$Q = \frac{\zeta + f}{h} \quad (43.134a)$$

$$= \frac{\zeta + f}{H(1 + \Delta\eta/H)} \quad (43.134b)$$

$$\approx \frac{1}{H} (\zeta + f) \left[1 - \frac{\Delta\eta}{H} \right] \quad (43.134c)$$

$$\approx \frac{1}{H} \left[\zeta + f_0 + \beta y - f_0 \frac{\Delta\eta}{H} \right]. \quad (43.134d)$$

43.5.2 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 38.1, for which

$$Q = \boldsymbol{\omega}_a \cdot \nabla b = \frac{\partial u}{\partial z} \frac{\partial b}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b}{\partial x} + (\zeta + f) \frac{\partial b}{\partial z}. \quad (43.135)$$

Now split the buoyancy into a depth dependent background and a deviation from the background

$$b = \tilde{b}(z) + b'(x, y, z, t), \quad (43.136)$$

and write its vertical derivative as

$$\frac{\partial b}{\partial z} = N^2 + \frac{\partial b'}{\partial z}, \quad (43.137)$$

where

$$N^2 = \frac{\partial \tilde{b}}{\partial z} \quad (43.138)$$

is the squared buoyancy frequency for the background buoyancy field. Rather than introducing non-dimensional variables⁴, we work more briefly by arranging the terms in the PV according to their QG scaling in terms of the Rossby number

$$Q = [f_0 N^2] + \left[(\beta y + \zeta) N^2 + f_0 \frac{\partial b'}{\partial z} \right] + \left[(\beta y + \zeta) \frac{\partial b'}{\partial z} + \frac{\partial u}{\partial z} \frac{\partial b'}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial b'}{\partial x} \right]. \quad (43.139)$$

We drop the third bracket term, as it is order Ro^2 , and write

$$Q = \tilde{Q} + N^2 q_*, \quad (43.140)$$

where

$$\tilde{Q} = f_0 N^2 \quad (43.141)$$

is the f-plane planetary geostrophic PV, and

$$q_* = \beta y + \zeta + \frac{f_0}{N^2} \frac{\partial b'}{\partial z}. \quad (43.142)$$

The material conservation of PV now takes the form

$$N^2 \frac{Dq_*}{Dt} + w \left[1 + \frac{q_*}{f_0} \right] \frac{\partial \tilde{Q}}{\partial z} = 0 \quad (43.143)$$

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

$$\frac{Dq_*}{Dt} + \frac{w}{N^2} \frac{\partial \tilde{Q}}{\partial z} = 0. \quad (43.144)$$

To eliminate the vertical velocity component, we introduce the buoyancy equation

$$\frac{Db'}{Dt} + w N^2 = 0, \quad (43.145)$$

so that

$$\frac{Dq_*}{Dt} - \frac{f_0}{N^4} \frac{\partial N^2}{\partial z} \frac{Db'}{Dt} = 0. \quad (43.146)$$

Writing

$$\frac{\partial}{\partial z} \left[\frac{1}{N^2} \right] = -\frac{1}{N^4} \frac{\partial N^2}{\partial z} \quad (43.147)$$

leads to

$$\frac{Dq_*}{Dt} + f_0 \left[\frac{\partial N^{-2}}{\partial z} \right] \frac{Db'}{Dt} = 0. \quad (43.148)$$

⁴SMG: Should introduce non-dimensional variables in future versions of the notes to enhance the analysis from *Vallis (2006)* Section 5.5.1.

Since the material time derivative operator only involves horizontal advection, we can merge these two terms to render

$$\frac{Dq_*}{Dt} + f_0 \left(\frac{\partial N^{-2}}{\partial z} \right) \frac{Db'}{Dt} = \frac{D}{Dt} \left[q_* + f_0 b' \left(\frac{\partial N^{-2}}{\partial z} \right) \right] \quad (43.149a)$$

$$= \frac{D}{Dt} \left[\beta y + \zeta + \frac{f_0}{N^2} \left(\frac{\partial b'}{\partial z} \right) + f_0 b' \left(\frac{\partial N^{-2}}{\partial z} \right) \right] \quad (43.149b)$$

$$= \frac{D}{Dt} \left[\beta y + \zeta + f_0 \frac{\partial}{\partial z} \left(\frac{b'}{N^2} \right) \right] \quad (43.149c)$$

$$= 0. \quad (43.149d)$$

The term inside the bracket is the QG-PV given by equation (43.111).

43.6 Energetics of a continuously stratified QG fluid

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

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta = f_0 \frac{\partial w}{\partial z} \quad (43.150a)$$

$$\frac{\partial b}{\partial t} + \mathbf{u} \cdot \nabla b = -w N^2, \quad (43.150b)$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \quad b' = f_0 \frac{\partial \psi}{\partial z} \quad u = -\frac{\partial \psi}{\partial y} \quad v = \frac{\partial \psi}{\partial x} \quad \zeta = \nabla_z^2 \psi. \quad (43.151)$$

43.6.1 Kinetic energy

The kinetic energy per mass for the total fluid domain is given by the integral

$$\mathcal{K} = \frac{1}{2} \int (\mathbf{u} \cdot \mathbf{u}) dV = \frac{1}{2} \int (\nabla \psi \cdot \nabla \psi) dV, \quad (43.152)$$

and its time derivative is

$$\frac{d\mathcal{K}}{dt} = \int \nabla \psi \cdot \nabla \left[\frac{\partial \psi}{\partial t} \right] dV. \quad (43.153)$$

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

$$\frac{d\mathcal{K}}{dt} = \int \nabla \psi \cdot \nabla \left[\frac{\partial \psi}{\partial t} \right] dV \quad (43.154a)$$

$$= \int [\nabla \cdot [\psi \nabla (\partial \psi / \partial t)] - \psi \partial (\nabla_z^2 \psi) / \partial t] dV \quad (43.154b)$$

$$= - \int \psi \frac{\partial \zeta}{\partial t} dV, \quad (43.154c)$$

where we dropped the lateral boundary term and introduced relative vorticity. Use of the vorticity equation (43.150a) yields

$$\frac{d\mathcal{K}}{dt} = - \int \psi \frac{\partial \zeta}{\partial t} dV \quad (43.155a)$$

$$= \int \psi (\mathbf{u} \cdot \nabla \zeta - f_0 \partial_z w) dV. \quad (43.155b)$$

The first term vanishes, since

$$\int \psi (\mathbf{u} \cdot \nabla \zeta) dV = \int \psi \nabla \cdot (\mathbf{u} \zeta) dV \quad (43.156a)$$

$$= \int [\nabla \cdot (\psi \mathbf{u} \zeta) - \nabla \psi \cdot \mathbf{u} \zeta] dV \quad (43.156b)$$

$$= 0, \quad (43.156c)$$

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

$$\frac{d\mathcal{K}}{dt} = - \int \psi f_0 \frac{\partial w}{\partial z} dV. \quad (43.157)$$

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

$$\frac{d\mathcal{K}}{dt} = - \int \psi f_0 \frac{\partial w}{\partial z} dV \quad (43.158a)$$

$$= - \int f_0 \left[\frac{\partial(w\psi)}{\partial z} - w \frac{\partial\psi}{\partial z} \right] dV \quad (43.158b)$$

$$= \int f_0 w \frac{\partial\psi}{\partial z} dV, \quad (43.158c)$$

where we dropped the boundary term given the rigid boundaries. Making use of

$$b' = f_0 \partial\psi/\partial z \quad (43.159)$$

leads to

$$\frac{d\mathcal{K}}{dt} = \int w b' dV. \quad (43.160)$$

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 26.2. It is reassuring to see the same behavior in the quasi-geostrophic system.

43.6.2 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

$$\mathcal{A} = \frac{1}{2} \int \left[\frac{f_0}{N} \frac{\partial\psi}{\partial z} \right]^2 dV. \quad (43.161)$$

Taking a time derivative leads to

$$\frac{d\mathcal{A}}{dt} = \int \left[\frac{f_0}{N} \right]^2 \frac{\partial \psi}{\partial z} \frac{\partial}{\partial t} \frac{\partial \psi}{\partial z} dV \quad (43.162a)$$

$$= \int \frac{f_0}{N^2} \frac{\partial \psi}{\partial z} [-w N^2 - \nabla \cdot (\mathbf{u} b)] dV, \quad (43.162b)$$

where we used the buoyancy equation (43.150b) 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

$$\int \frac{f_0}{N^2} \frac{\partial \psi}{\partial z} [\nabla \cdot (\mathbf{u} b)] dV = \int \left[\frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} \right] \mathbf{u} \cdot \nabla \left[\frac{\partial \psi}{\partial z} \right] dV \quad (43.163a)$$

$$= \frac{1}{2} \int \nabla \cdot \left[\mathbf{u} \left(\frac{f_0}{N} \frac{\partial \psi}{\partial z} \right)^2 \right] dV \quad (43.163b)$$

$$= 0. \quad (43.163c)$$

Consequently, the APE changes according to

$$\frac{d\mathcal{A}}{dt} = - \int w f_0 \frac{\partial \psi}{\partial z} dV = - \int w b' dV. \quad (43.164)$$

43.6.3 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

$$\frac{d(\mathcal{K} + \mathcal{A})}{dt} = 0. \quad (43.165)$$

The buoyancy work conversion term is given by

$$\text{buoyancy work} = \int w f_0 \frac{\partial \psi}{\partial z} dV = \int w b' dV, \quad (43.166)$$

which has the same form as that encountered for the conversion between potential and kinetic energy in the unapproximated equations (Section 26.2)

43.6.4 Scaling APE and KE

The scale for the kinetic energy is given by

$$\mathcal{K} = \frac{1}{2} \int (\nabla \psi \cdot \nabla \psi) dV \sim L^{-2} \Psi V \quad (43.167)$$

and the APE scale is

$$\mathcal{A} = \frac{1}{2} \int \left[\frac{f_0}{N} \frac{\partial \psi}{\partial z} \right]^2 dV \sim L_d^{-2} \Psi V, \quad (43.168)$$

where we wrote Ψ for the streamfunction scale, and V for the domain volume. Taking the ratio yields

$$\frac{\mathcal{K}}{\mathcal{A}} \sim \left[\frac{L_d}{L} \right]^2 = \left[\frac{H}{L} \right]^2 \left[\frac{N}{f_0} \right]^2 = Bu. \quad (43.169)$$

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.

43.7 Exercises

EXERCISE 43.1: RIGID LID AND f/H CONTOURS

Derive equation (43.78) for rigid lid flow.

EXERCISE 43.2: 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 30.5, 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.

- 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.
- Assume friction arises just from vertical shears in the horizontal velocity, so that

$$\mathbf{F} = \frac{\partial}{\partial z} \left[\nu \frac{\partial \mathbf{u}}{\partial z} \right], \quad (43.170)$$

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

$$\frac{Dq^{\text{approx}}}{Dt} = A \nabla_z^2 q^{\text{approx}}, \quad (43.171)$$

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 43.3: TRADITIONAL FORM OF THE QUASI-GEOSTROPHIC ω -EQUATION

As discussed in Section 43.3, 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 [43.3.2](#), the non-dimensional velocity, continuity, buoyancy and vorticity equations for β -plane flow valid at order Ro^1 are given by

$$\frac{D_0 \hat{\mathbf{u}}_0}{Dt} + \hat{f}_0 \hat{\mathbf{k}} \wedge \hat{\mathbf{u}}_1 + \hat{\beta} \hat{y} \hat{\mathbf{k}} \wedge \hat{\mathbf{u}}_0 = -\hat{\nabla}_z \hat{\phi}_1 \quad (43.172a)$$

$$\hat{\nabla}_z \cdot \hat{\mathbf{u}}_1 + \frac{\partial \hat{w}_1}{\partial \hat{z}} = 0 \quad (43.172b)$$

$$F \frac{D_0 \hat{b}_0}{Dt} + \hat{w}_1 = 0 \quad (43.172c)$$

$$\frac{\partial \hat{\zeta}_0}{\partial \hat{t}} + (\hat{\mathbf{u}}_0 \cdot \hat{\nabla}) (\hat{\zeta}_0 + \hat{\beta} \hat{y}) = \hat{f}_0 \frac{\partial \hat{w}_1}{\partial \hat{z}}. \quad (43.172d)$$

Write the dimensional form of these four equations, showing all steps when moving from the dimensionless to dimensional equations. For the horizontal velocity, write

$$\mathbf{u}_g = U \hat{\mathbf{u}}_0 \quad (43.173)$$

for that portion of the geostrophic flow that is horizontally non-divergent, and

$$\mathbf{u}_{ag} = U(\hat{\mathbf{u}} - \hat{\mathbf{u}}_0) \approx Ro U \hat{\mathbf{u}}_1 \quad (43.174a)$$

$$\phi_{ag} = \Phi(\hat{\phi} - \hat{\phi}_0) \approx Ro \Phi \hat{\phi}_1. \quad (43.174b)$$

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

$$\mathcal{L}w = \sigma, \quad (43.175)$$

where

$$\mathcal{L} = N^2 \nabla_z^2 + f_0^2 \frac{\partial^2}{\partial z^2} \quad (43.176)$$

is a linear partial differential operator and

$$\sigma = f_0 J(\psi, \zeta + \beta y) - \nabla^2 J(\psi, b) \quad (43.177)$$

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 [??](#). What about when $N^2 < 0$? What new physical phenomena do you expect when $N^2 < 0$?

EXERCISE 43.4: HOSKINS' FORM OF THE QUASI-GEOSTROPHIC ω -EQUATION

We here rederive the ω -equation from Exercise 43.3 using methods introduced by *Hoskins et al.* (1978). It is not necessary to have solved Exercise 43.3 to solve the present exercise.

Hoskins' approach reveals an insightful form for the source function contributing to vertical motion. As in Exercise 43.3, we work with the adiabatic and hydrostatic Boussinesq system (see Section 43.1)

$$\frac{Du}{Dt} - fv = -\frac{\partial \phi}{\partial x} \quad (43.178a)$$

$$\frac{Dv}{Dt} + fu = -\frac{\partial \phi}{\partial y} \quad (43.178b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (43.178c)$$

$$\frac{Db}{Dt} = 0 \quad (43.178d)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0 \quad (43.178e)$$

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (43.178f)$$

In Exercise 43.3, 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

$$\mathbf{u} = \mathbf{u}_g + \mathbf{u}_{ag}, \quad (43.179)$$

with the geostrophic velocity balancing the horizontal gradient of the full pressure field

$$f\mathbf{u}_g = \hat{\mathbf{z}} \wedge \nabla \phi. \quad (43.180)$$

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 43.3.1 or Exercise 43.3). 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

$$\frac{D\mathbf{u}}{Dt} + f\hat{\mathbf{z}} \wedge \mathbf{u}_{ag} = 0, \quad (43.181)$$

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

$$\frac{\partial \mathbf{u}_g}{\partial t} + (\mathbf{u}_g \cdot \nabla_z) \mathbf{u}_g + f\hat{\mathbf{z}} \wedge \mathbf{u}_{ag} = 0 \quad (43.182)$$

and the buoyancy equation equal to

$$\frac{\partial b}{\partial t} + \mathbf{u}_g \cdot \nabla_z b + N^2 w = 0, \quad (43.183)$$

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 43.3.2). However, Hoskins' momentum equation (43.182) has no pressure gradient on the right hand side, whereas an asymptotic approach has contributions from higher order pressure gradients (Section 43.3.2). 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

$$\left[\frac{\partial}{\partial t} + \mathbf{u}_g \cdot \nabla_z \right] |\nabla_z b|^2 = \mathbf{Q} \cdot \nabla_z b. \quad (43.184)$$

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 14.1.

- (b) Show that the quasi-geostrophic ω -equation on an f -plane can be written

$$N^2 \nabla_z^2 w + f_0^2 \frac{\partial^2 w}{\partial z^2} = 2 \nabla_z \cdot \mathbf{Q}. \quad (43.185)$$

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 (43.185) and show each of the relevant steps.

- (c) The equation (43.185) for the vertical velocity takes the form

$$\mathcal{L}w = \sigma, \quad (43.186)$$

where

$$\mathcal{L} = N^2 \nabla_z^2 + f_0^2 \frac{\partial^2}{\partial z^2} \quad (43.187)$$

is a linear partial differential operator and

$$\sigma = 2 \nabla_z \cdot \mathbf{Q} \quad (43.188)$$

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 ???. What about when $N^2 < 0$? What new physical phenomena do you expect when $N^2 < 0$?

44

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.

44.1	Three types of instabilities	662
44.1.1	Elements of the instabilities	662
44.1.2	The nature of the base state and the perturbations	663
44.1.3	Comments	663
44.2	Centrifugal instability of cyclostrophically balanced flow	664
44.2.1	Equations of motion	664
44.2.2	Solid-body motion	665
44.2.3	Stability analysis based on energetic arguments	665
44.2.4	Stability analysis based on parcel arguments	667
44.2.5	Comments	668
44.3	Potential momentum	668
44.3.1	Linear momentum and potential momentum	669
44.3.2	Zonal potential momentum on a β -plane	670
44.3.3	Further reading	670
44.4	Horizontal inertial instability of a geostrophically balanced front	670
44.4.1	Equations of motion and equilibrium state	671
44.4.2	Stability analysis based on energetic arguments	672
44.4.3	Stability analysis based on parcel arguments	673
44.4.4	Comments	674
44.5	Symmetric (isentropic inertial) instability of a baroclinic front	675
44.5.1	Equations of motion and equilibrium state	675
44.5.2	Stability analysis based on parcel arguments	676
44.5.3	Comments	678
44.6	Secondary circulation along baroclinic fronts	678
44.6.1	Hydrostatic and Boussinesq fluid on an f -plane	679

44.6.2	Geostrophic momentum approximation	680
44.6.3	Secondary ageostrophic circulation	681
44.6.4	Ageostrophic overturning circulation for a symmetric front	682
44.6.5	Connection to potential vorticity and symmetric instability	683

44.1 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

44.1.1 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).

44.1.2 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 44.5 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.

44.1.3 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.

44.2 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 34.5 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

44.2.1 Equations of motion

We studied the angular momentum of this rotating shallow water system in Section 34.5, 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 33.1. Making use of those earlier results allows us to write down the horizontal components to the equations of motion

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta}{\partial r} + r^{-3} (l^z)^2 \quad (44.1a)$$

$$\frac{Dl^z}{Dt} = -g \frac{\partial \eta}{\partial \theta}. \quad (44.1b)$$

In these equations, η is the free surface height for the shallow water layer (see the schematic in Figure ??),

$$l^z = \hat{z} \cdot [\mathbf{r} \wedge (\mathbf{u} + \mathbf{U}_{\text{solid}})] = r (v^\theta + r \Omega) \quad (44.2)$$

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

$$v^r = \frac{Dr}{Dt} \quad r^{-1} v^\theta = \frac{D\theta}{Dt}. \quad (44.3)$$

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 44.5.

The material evolution of the radial velocity (equation (44.1a)) 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

$$\frac{Dv^r}{Dt} = 0 \implies g \frac{\partial \eta}{\partial r} = \frac{(v^\theta + r \Omega)^2}{r} \quad \text{cyclostrophic balance.} \quad (44.4)$$

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

$$\frac{\partial \eta}{\partial \theta} = 0 \implies \frac{Dl^z}{Dt} = 0. \quad (44.5)$$

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, $v^\theta = v^\theta(r)$.

44.2.2 Solid-body motion

Consider the rotating fluid undergoing solid-body rotation. In this case, the angular momentum per mass is given by

$$l_{\text{solid-body}}^z = r^2 \Omega \quad (44.6)$$

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

$$\frac{d[l^z(r)]^2}{dr} = 4r^3 \Omega^2 > 0. \quad (44.7)$$

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.

44.2.3 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

$$E(r_1) = (\delta m/2) [v^\theta(r_1) + r_1 \Omega]^2 = (\delta m/2) [l^z(r_1)/r_1]^2. \quad (44.8)$$

The initial kinetic energy for the two rings is thus given by the sum

$$E_{\text{init}} = (\delta m/2) ([l^z(r_1)/r_1]^2 + [l^z(r_2)/r_2]^2). \quad (44.9)$$

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

$$E_{\text{swap}} = (\delta m/2) ([l^z(r_1)/r_2]^2 + [l^z(r_2)/r_1]^2). \quad (44.10)$$

The difference in energy is given by

$$E_{\text{swap}} - E_{\text{init}} = (\delta m/2) ([l^z(r_2)]^2 - [l^z(r_1)]^2)(r_1^{-2} - r_2^{-2}). \quad (44.11)$$

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

$$\frac{d[l^z(r)]^2}{dr} < 0 \implies \frac{d}{dr} \left[r^4 (\dot{\theta} + \Omega)^2 \right] = \frac{d}{dr} \left[r^2 (v^\theta + r \Omega)^2 \right] < 0. \quad (44.12)$$

Rearrangement leads to the instability condition

$$(\dot{\theta} + \Omega)(\zeta + f) < 0, \quad (44.13)$$

where

$$\zeta = \frac{1}{r} \frac{d(r v^\theta)}{dr} \quad (44.14)$$

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

$$(\dot{\theta} + \Omega)(\zeta + f) > 0, \quad (44.15)$$

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.

44.2.4 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 (44.1a) leads to

$$\frac{Dv^r}{Dt} = 0 = -g \frac{\partial \bar{\eta}}{\partial r} + r^{-3} (\bar{l}^z)^2 \quad (44.16)$$

Subtracting this equilibrium state from the full momentum equation (44.1a) leads to an equation for radial acceleration of perturbations about the equilibrium state

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta'}{\partial r} + r^{-3} [(l^z)^2 - (\bar{l}^z)^2], \quad (44.17)$$

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

$$(l^z)^2(r + \Delta r) = (l^z)^2(r) = (\bar{l}^z)^2(r), \quad (44.18)$$

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

$$(l^z)^2(r + \Delta r) - (\bar{l}^z)^2(r + \Delta r) = (\bar{l}^z)^2(r) - (\bar{l}^z)^2(r + \Delta r) \approx -\Delta r \frac{d(\bar{l}^z)^2(r)}{dr}, \quad (44.19)$$

so that the radial momentum equation at $r + \Delta r$ takes on the form

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta'}{\partial r} - \frac{\Delta r}{(r + \Delta r)^3} \frac{d(\bar{l}^z)^2(r)}{dr}. \quad (44.20)$$

We thus see that if the squared angular momentum decreases upon moving the ring to a larger radius, then the second term in equation (44.20) 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 (44.12) derived using energetic arguments.

Centrifugal oscillations

Ignoring the perturbation pressure gradient, and introducing a centrifugal frequency

$$\sigma^2(r) \equiv \frac{1}{(r + \Delta r)^3} \frac{d(\bar{l}^z)^2(r)}{dr} \quad (44.21)$$

leads to the oscillator equation for the deviation from the equilibrium radial position

$$\frac{D^2 \Delta r}{Dt^2} + \sigma^2 \Delta r = 0, \quad (44.22)$$

where $u^r = D\Delta r/Dt$. For stable cases with $\sigma^2 > 0$, the parcel oscillates around the equilibrium position with period

$$T_{\text{centrifugal}} = \frac{2\pi}{\sigma}. \quad (44.23)$$

In contrast, for the unstable case with $\sigma^2 < 0$, then Δr grows exponentially.

44.2.5 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 (44.20). 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.

44.3 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 44.1 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.

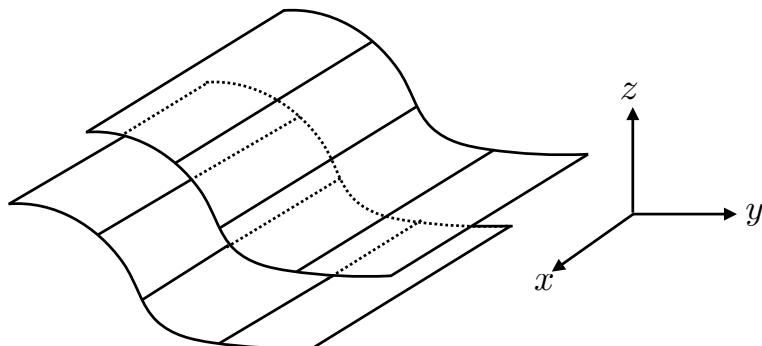


Figure 44.1: Example of a zonally symmetric baroclinic frontal region, showing two iso-buoyancy surfaces with $b = b(y, z, t)$. With $\partial b / \partial y > 0$ as drawn, the corresponding zonal thermal wind shear, $f \partial u / \partial z = -\partial b / \partial y < 0$ is westward; i.e., stronger westward flow with increasing height.

44.3.1 Linear momentum and potential momentum

The horizontal linear momentum per mass in a Boussinesq perfect fluid on an f -plane evolves according to

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \wedge \mathbf{u} = -\nabla_z p. \quad (44.24)$$

Since f is a constant, this equation can be written

$$\frac{DM}{Dt} = -\nabla_z p, \quad (44.25)$$

where we introduced the *potential momentum* per mass

$$\mathbf{M} \equiv \mathbf{u} + f \hat{\mathbf{z}} \wedge \mathbf{x} = (u - f y) \hat{\mathbf{x}} + (v + f x) \hat{\mathbf{y}}, \quad (44.26)$$

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 12.2. 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 44.1, we assume that all fields are independent of the zonal direction so that

$$\frac{\partial p}{\partial x} = 0. \quad (44.27)$$

A vanishing zonal pressure gradient means that the *zonal potential momentum* per mass is a Lagrangian invariant

$$\frac{DM^x}{Dt} = 0 \quad \text{where } M^x \equiv u - f y. \quad (44.28)$$

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

$$u_0 - f y_0 = y_1 - f y_1, \quad (44.29)$$

so that the zonal velocity at the new latitude is given by

$$u_1 = u_0 - f(y_0 - y_1). \quad (44.30)$$

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 (44.30) we see that

$$u_0 = M^x(y_1) + f y_0. \quad (44.31)$$

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 44.4, 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

$$\frac{Du}{Dt} = fv - \frac{1}{\rho_0} \frac{\partial p}{\partial x} \quad (44.32a)$$

$$\frac{Dv}{Dt} = -fu. \quad (44.32b)$$

In this case the meridional potential momentum is materially invariant

$$\frac{DM^y}{Dt} = 0 \quad \text{where } M^y \equiv v + fx. \quad (44.33)$$

44.3.2 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 (44.25), 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

$$\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - f v = 0, \quad (44.34)$$

where we assumed zonal symmetry ($\partial_x = 0$). Now write the Coriolis parameter in the form

$$\Gamma = f_0 y + \beta y^2 / 2 \implies f = \frac{d\Gamma}{dy}, \quad (44.35)$$

so that the zonal momentum equation takes the form

$$\frac{\partial(u - \Gamma)}{\partial t} + v \frac{\partial(u - \Gamma)}{\partial y} + w \frac{\partial(u - \Gamma)}{\partial z} = 0. \quad (44.36)$$

We thus see that that $M = u - \Gamma$ is materially invariant for this zonally symmetric front.

44.3.3 Further reading

See Section 12.3 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 \(2010\)](#).

44.4 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 44.2 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 44.2, 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 44.5 maintains the adiabatic nature of displacements when probing for instabilities.

44.4.1 Equations of motion and equilibrium state

The horizontal momentum equation for an inviscid Boussinesq fluid on an f -plane is given by

$$\frac{D\mathbf{u}}{Dt} + f \hat{z} \wedge \mathbf{u} = -\frac{1}{\rho_0} \nabla_z p. \quad (44.37)$$

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

$$u_g = -\frac{1}{f \rho_0} \frac{\partial p}{\partial y} \quad (44.38a)$$

$$v = 0. \quad (44.38b)$$

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 (44.37) leads to

$$\frac{Du}{Dt} = fv \quad (44.39a)$$

$$\frac{Dv}{Dt} = f(u_g - u). \quad (44.39b)$$

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 12.2, we write the along-front momentum equation as the material time derivative of the along-front potential momentum per mass³

$$m = u - fy, \quad (44.40)$$

bringing the suite of perturbation equations to

$$\frac{Dm}{Dt} = 0 \quad (44.41a)$$

$$\frac{Dv}{Dt} = f(u_g - u). \quad (44.41b)$$

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 44.5.

Material invariance of the along-front potential momentum plays a fundamental role in the stability analysis.

44.4.2 Stability analysis based on energetic arguments

We follow the energetic arguments given in Section 44.2.3 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

$$E = (1/2) u_g^2 = (1/2) (m_g + fy)^2, \quad (44.42)$$

where we replaced the geostrophic velocity with the geostrophic potential momentum through equation (44.5.1). 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

$$E_{\text{init}} = (1/2) (m_g(y_1) + fy_1)^2 + (1/2) (m_g(y_2) + fy_2)^2. \quad (44.43)$$

Swapping the parcels while leaving the potential momentum unchanged leads to the kinetic energy in the swapped state

$$E_{\text{swap}} = (1/2) (m_g(y_1) + fy_2)^2 + (1/2) (m_g(y_2) + fy_1)^2. \quad (44.44)$$

A bit of algebra leads to the difference in kinetic energy for the two states

$$E_{\text{swap}} - E_{\text{init}} = -f \Delta y \Delta m_g, \quad (44.45)$$

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

$$E_{\text{swap}} - E_{\text{init}} = -f (\Delta y)^2 \frac{dm_g}{dy} \quad (44.46a)$$

$$= -f (\Delta y)^2 \left(\frac{\partial u_g}{\partial y} - f \right) \quad (44.46b)$$

$$= f (\Delta y)^2 (\zeta_g + f), \quad (44.46c)$$

where

$$\zeta_g = -\frac{\partial u_g}{\partial y} \quad (44.47)$$

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

$$-f \frac{\partial m_g}{\partial y} = f (\zeta_g + f) < 0 \quad \text{inertial instability.} \quad (44.48)$$

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

$$-f(\partial m/\partial y) = f^2 > 0 \quad \text{zero along-front flow.} \quad (44.49)$$

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 (44.48) 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 44.2.3.

Details of the instability condition

The instability condition (44.48) 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 $\hat{x} \wedge \hat{y} = \hat{z}$.

$$\text{northern hemisphere } (f > 0): \quad \frac{\partial m_g}{\partial y} > 0 \quad \zeta_g < -|f| \quad \frac{\partial u_g}{\partial y} > +|f| \quad (44.50a)$$

$$\text{southern hemisphere } (f < 0): \quad \frac{\partial m_g}{\partial y} < 0 \quad \zeta_g > +|f| \quad \frac{\partial u_g}{\partial y} < -|f|. \quad (44.50b)$$

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.

44.4.3 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 (44.48) 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

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)], \quad (44.51)$$

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

$$m(y) = m_g(y) = u_g(y) - fy \quad (44.52)$$

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 $u_g(y)$

$$m(y + \Delta y) = u(y + \Delta y) - f(y + \Delta y) \quad (44.53a)$$

$$= m(y) \quad (44.53b)$$

$$= u_g(y) - fy, \quad (44.53c)$$

which leads to

$$u(y + \Delta y) = u_g(y) + f\Delta y. \quad (44.54)$$

The cross-front acceleration (44.51) thus takes the form

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)] \quad (44.55a)$$

$$= f[u_g(y + \Delta y) - u_g(y) - f\Delta y] \quad (44.55b)$$

$$\approx f\Delta y \left(\frac{\partial u_g}{\partial y} - f \right) \quad (44.55c)$$

$$= -\Delta y f (\zeta_g + f). \quad (44.55d)$$

As with the energetic arguments in Section 44.4.2, we are left with the instability condition

$$f (\zeta_g + f) < 0 \implies \text{inertially unstable}. \quad (44.56)$$

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 44.2 for centrifugal instability, as well as for a vertical column of fluid in Section 20.3 for gravitational stability.

44.4.4 Comments

As stated in Section 44.1.3, 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.

44.5 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 44.4.3, 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.

44.5.1 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 31.2. The horizontal momentum equation is thus given by

$$\frac{Du}{Dt} + f \hat{z} \wedge \mathbf{u} = -\nabla_b M \quad (44.57)$$

where ∇_b is a horizontal gradient computed along constant buoyancy surfaces. Furthermore,

$$M = p/\rho_0 - b z \quad (44.58)$$

is the Montgomery potential that contributes the acceleration

$$-\nabla_b M = -\nabla_b(p/\rho_0) + b \nabla_b z. \quad (44.59)$$

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

$$u_g = -\frac{1}{f \rho_0} \left[\frac{\partial M}{\partial y} \right]_b \quad (44.60a)$$

$$v = 0. \quad (44.60b)$$

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 (44.37) leads to

$$\frac{Du}{Dt} = fv \quad (44.61a)$$

$$\frac{Dv}{Dt} = f(u_g - u), \quad (44.61b)$$

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 12.2 and 44.4.1, we write the along-front momentum equation as the material time derivative of the along-front potential momentum per mass $m = u - fy$, bringing the perturbation equations to

$$\frac{Dm}{Dt} = 0 \quad (44.62a)$$

$$\frac{Dv}{Dt} = f(u_g - u). \quad (44.62b)$$

44.5.2 Stability analysis based on parcel arguments

We follow the parcel analysis for inertial instability in Section 44.4.3, 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

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)], \quad (44.63)$$

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

$$m(y) = m_g(y) = u_g(y) - fy \quad (44.64)$$

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 $u_g(y)$

$$m(y + \Delta y) = u(y + \Delta y) - f(y + \Delta y) \quad (44.65a)$$

$$= m(y) \quad (44.65b)$$

$$= u_g(y) - fy, \quad (44.65c)$$

which leads to

$$u(y + \Delta y) = u_g(y) + f\Delta y. \quad (44.66)$$

The cross-front acceleration (44.63) thus takes the form

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)] \quad (44.67a)$$

$$= f[u_g(y + \Delta y) - u_g(y) - f\Delta y] \quad (44.67b)$$

$$\approx f\Delta y \left[\left(\frac{\partial u_g}{\partial y} \right)_b - f \right] \quad (44.67c)$$

$$= -\Delta y f (\zeta_g^b + f). \quad (44.67d)$$

In the final equality we introduced the relative geostrophic vorticity as written in buoyancy coordinates (see Section 31.2)

$$\zeta_g^b = \left[\frac{\partial v_g}{\partial x} \right]_b - \left[\frac{\partial u_g}{\partial y} \right]_b = - \left[\frac{\partial u_g}{\partial y} \right]_b. \quad (44.68)$$

Furthermore, we can introduce the Boussinesq Ertel potential vorticity in the form

$$Q = (\zeta^b + f) N^2, \quad (44.69)$$

with

$$N^2 = \frac{\partial b}{\partial z} \quad (44.70)$$

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

$$f(\zeta_g^b + f) < 0 \implies \text{symmetrically unstable}. \quad (44.71a)$$

$$f Q_g < 0 \implies \text{symmetrically unstable}. \quad (44.71b)$$

The instability condition (44.71a) is a direct translation of the inertial instability condition (44.56), translating from horizontal displacements to isopycnal displacements.

Relative slope of buoyancy and potential momentum surfaces: Part I

The instability condition (44.71a) 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

$$\left[\frac{\partial m}{\partial y} \right]_b = \left[\frac{\partial u}{\partial y} \right]_b - f = -(\zeta^b + f). \quad (44.72)$$

Consequently, the instability condition (44.71a) is equivalent to

$$f \left[\frac{\partial m_g}{\partial y} \right]_b > 0 \implies \text{symmetrically unstable}. \quad (44.73)$$

This condition is directly analogous to the inertial instability condition (44.48), 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 44.2.

Relative slope of buoyancy and potential momentum surfaces: Part II

As another way to write the instability condition (44.73), we make use of the expression (8.68b) 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

$$\left[\frac{\partial m_g}{\partial y} \right]_b = \left[\frac{\partial m_g}{\partial y} \right]_z - \frac{\partial b / \partial y}{\partial b / \partial z} \frac{\partial m_g}{\partial z} \quad (44.74a)$$

$$= N^{-2} \hat{x} \cdot (\nabla m_g \wedge \nabla b) \quad (44.74b)$$

$$= -N^{-2} \hat{x} \cdot (\nabla b \wedge \nabla m_g) \quad (44.74c)$$

$$= - \left[\frac{\partial b}{\partial y} \right]_{m_g}. \quad (44.74d)$$

We are thus led to the additional expression for a symmetrically unstable base state

$$f \left[\frac{\partial b}{\partial y} \right]_{m_g} < 0 \implies \text{symmetrically unstable.} \quad (44.75)$$

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 44.2.

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

$$\omega + f\hat{z} = \hat{x} \wedge \nabla m, \quad (44.76)$$

so that

$$Q = (\omega + f\hat{z}) \cdot \nabla b \quad (44.77a)$$

$$= (\nabla b \wedge \nabla m) \cdot \hat{x} \quad (44.77b)$$

$$= N^2 (\zeta^b + f). \quad (44.77c)$$

The condition for symmetric instability thus can be written

$$f Q_g = -N^2 f \left[\frac{\partial m_g}{\partial y} \right]_b = N^2 f \left[\frac{\partial b}{\partial y} \right]_{m_g} < 0. \quad (44.78)$$

44.5.3 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.

44.6 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 44.5. However, the focus here is not on the stability per se, but instead on deriving an expression for the overturning circulation.

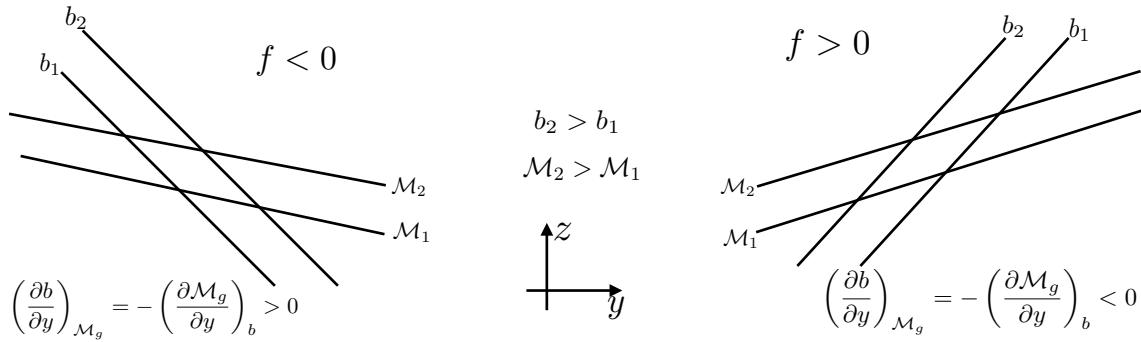


Figure 44.2: Example flow configurations that are symmetrically unstable; i.e., inertially unstable to a symmetric perturbation along a constant buoyancy surface. We show example buoyancy surfaces and potential momentum surfaces for the southern hemisphere (left) and northern hemisphere (right). The instability conditions (44.73) and (44.75) 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).

44.6.1 Hydrostatic and Boussinesq fluid on an f -plane

We frame our analysis within the adiabatic and hydrostatic Boussinesq equations (see Section 43.1) on an f -plane

$$\frac{Du}{Dt} - fv = -\frac{\partial \phi}{\partial x} \quad (44.79a)$$

$$\frac{Dv}{Dt} + fu = -\frac{\partial \phi}{\partial y} \quad (44.79b)$$

$$\frac{\partial \phi}{\partial z} = b \quad (44.79c)$$

$$\frac{Db}{Dt} = 0 \quad (44.79d)$$

$$\nabla \cdot \mathbf{v} = \nabla_z \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0. \quad (44.79e)$$

The geostrophic velocity is given in terms of the pressure field by

$$\mathbf{u}_g = f^{-1} \hat{z} \wedge \nabla_z \phi \implies u_g = -\frac{1}{f} \frac{\partial \phi}{\partial y} \quad v_g = \frac{1}{f} \frac{\partial \phi}{\partial x}. \quad (44.80)$$

The horizontal momentum equations (44.79a)-(44.79b) can thus be written

$$\mathbf{u} = \mathbf{u}_g + \hat{\mathbf{z}} \wedge \mathcal{D}\mathbf{u} \implies u = u_g - \mathcal{D}v \quad v = v_g + \mathcal{D}u, \quad (44.81)$$

where we introduced the dimensionless material time operator

$$\mathcal{D} = \frac{1}{f} \frac{D}{Dt}. \quad (44.82)$$

One step of iteration then leads to

$$u = u_g - \mathcal{D}(v_g + \mathcal{D}u) \quad (44.83a)$$

$$v = v_g + \mathcal{D}(u_g - \mathcal{D}v). \quad (44.83b)$$

Recall that on an f -plane, the geostrophic velocity \mathbf{u}_g is horizontally non-divergent.

44.6.2 Geostrophic momentum approximation

The *geostrophic momentum approximation* assumes the horizontal velocity takes the form

$$\mathbf{u} \approx \mathbf{u}_{gm} = \mathbf{u}_g + \hat{\mathbf{z}} \wedge \mathcal{D}\mathbf{u}_g. \quad (44.84)$$

For simplicity in notation, we drop the “gm” subscripts in the following. The geostrophic momentum approximation holds so long as

$$|u| \gg |\mathcal{D}^2 u| \quad (44.85a)$$

$$|v| \gg |\mathcal{D}^2 v|. \quad (44.85b)$$

Rearranging the geostrophic momentum approximation (44.84) leads to

$$\mathcal{D}\mathbf{u}_g + \hat{\mathbf{z}} \wedge (\mathbf{u} - \mathbf{u}_g) = 0. \quad (44.86)$$

Reintroducing the material time derivative yields

$$\frac{D\mathbf{u}_g}{Dt} + f\hat{\mathbf{z}} \wedge \mathbf{u}_{ag} = 0 \quad (44.87)$$

where \mathbf{u}_{ag} is the ageostrophic velocity based on the geostrophic momentum velocity

$$\mathbf{u}_{ag} = \mathbf{u} - \mathbf{u}_g = \hat{\mathbf{z}} \wedge \mathcal{D}\mathbf{u}_g. \quad (44.88)$$

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

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla = \frac{\partial}{\partial t} + (\mathbf{u}_g + \mathbf{u}_{ag}) \cdot \nabla_z + w \frac{\partial}{\partial z} \quad (44.89)$$

with

$$\mathbf{v} = (\mathbf{u}_g + \mathbf{u}_{ag}) + \hat{\mathbf{z}} w. \quad (44.90)$$

44.6.3 Secondary ageostrophic circulation

In Sections 44.4 and 44.5 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 44.6.4. For that purpose, start from the zonal momentum equation, buoyancy equation, and continuity equation within the semi-geostrophic system

$$\frac{\partial u_g}{\partial t} + (\mathbf{u}_g \cdot \nabla_z) u_g + (\mathbf{u}_{ag} \cdot \nabla_z) u_g + w \frac{\partial u_g}{\partial z} - f v_{ag} = 0 \quad (44.91a)$$

$$\frac{\partial b}{\partial t} + \mathbf{u}_g \cdot \nabla_z b + \mathbf{u}_{ag} \cdot \nabla_z b + w N^2(z) = 0 \quad (44.91b)$$

$$\frac{\partial u_{ag}}{\partial x} + \frac{\partial v_{ag}}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (44.91c)$$

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

$$\frac{\partial}{\partial t} \frac{\partial u_g}{\partial z} + \left(\frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial u_g}{\partial z} + \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial u_g}{\partial z} + \frac{\partial}{\partial z} (w \frac{\partial u_g}{\partial z}) - f \frac{\partial v_{ag}}{\partial z} = 0 \quad (44.92a)$$

$$\frac{\partial}{\partial t} \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial b}{\partial y} + \frac{\partial w}{\partial y} N^2 = 0. \quad (44.92b)$$

We now make use of thermal wind for the geostrophic velocity

$$f \frac{\partial \mathbf{u}_g}{\partial z} = \hat{\mathbf{z}} \wedge \nabla_z b \implies \frac{\partial u_g}{\partial z} = -\frac{1}{f} \frac{\partial b}{\partial y} \quad \frac{\partial v_g}{\partial z} = \frac{1}{f} \frac{\partial b}{\partial x} \quad (44.93)$$

thus rendering

$$\frac{\partial}{\partial t} \frac{\partial b}{\partial y} - f \left(\frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial b}{\partial y} - f \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial b}{\partial y} + \frac{\partial}{\partial z} (w \frac{\partial b}{\partial y}) + f^2 \frac{\partial v_{ag}}{\partial z} = 0 \quad (44.94a)$$

$$\frac{\partial}{\partial t} \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_g \cdot \nabla_z) \frac{\partial b}{\partial y} + \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b + (\mathbf{u}_{ag} \cdot \nabla_z) \frac{\partial b}{\partial y} + \frac{\partial w}{\partial y} N^2 = 0. \quad (44.94b)$$

Subtracting equation (44.94b) from equation (44.94a) eliminates the time derivative thus revealing the diagnostic relation

$$-f \left(\frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z \right) u_g - f \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + \frac{\partial}{\partial z} (w \frac{\partial b}{\partial y}) + f^2 \frac{\partial v_{ag}}{\partial z} - \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z \right) b - \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b - \frac{\partial w}{\partial y} N^2 = 0. \quad (44.95)$$

We now make use of thermal wind and horizontal non-divergence for the geostrophic velocity to write (as in the solution to Exercise 43.4)

$$f \frac{\partial \mathbf{u}_g}{\partial z} \cdot \nabla_z u_g = \frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z b = -Q^{(y)}, \quad (44.96)$$

where $Q^{(y)}$ is the meridional component of the geostrophic \mathbf{Q} -vector (see Exercise 43.4)

$$\mathbf{Q} = - \left(\frac{\partial \mathbf{u}_g}{\partial x} \cdot \nabla_z b \right) \hat{\mathbf{x}} - \left(\frac{\partial \mathbf{u}_g}{\partial y} \cdot \nabla_z b \right) \hat{\mathbf{y}}. \quad (44.97)$$

Introduction of $Q^{(y)}$ into equation (44.95) yields

$$-f \left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z \right) u_g + \frac{\partial w}{\partial z} \frac{\partial b}{\partial y} + w \frac{\partial^2 b}{\partial z \partial y} + f^2 \frac{\partial v_{ag}}{\partial z} - \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z \right) b - \frac{\partial w}{\partial y} N^2 = -2Q^{(y)}. \quad (44.98)$$

Again making use of thermal wind and $\nabla_z \cdot \mathbf{u}_g = 0$ allows us to write

$$-f\left(\frac{\partial \mathbf{u}_{ag}}{\partial z} \cdot \nabla_z\right)u_g - \left(\frac{\partial \mathbf{u}_{ag}}{\partial y} \cdot \nabla_z\right)b = f\frac{\partial u_{ag}}{\partial z}\frac{\partial v_g}{\partial y} - f\frac{\partial v_{ag}}{\partial z}\frac{\partial u_g}{\partial y} + f\frac{\partial u_{ag}}{\partial y}\frac{\partial v_g}{\partial z} + f\frac{\partial v_{ag}}{\partial y}\frac{\partial u_g}{\partial z}. \quad (44.99)$$

The mixed partial derivative of the buoyancy vanishes

$$\frac{\partial^2 b}{\partial z \partial y} = \frac{\partial}{\partial y} \frac{\partial b}{\partial z} = \frac{\partial N^2(z)}{\partial y} = 0, \quad (44.100)$$

which follows since we are assuming a background vertical stratification that is independent of horizontal direction. Bringing these results together into equation (44.98) leads to

$$f\frac{\partial u_{ag}}{\partial z}\frac{\partial v_g}{\partial y} + f\frac{\partial u_{ag}}{\partial y}\frac{\partial v_g}{\partial z} + f\frac{\partial v_{ag}}{\partial y}\frac{\partial u_g}{\partial z} + \frac{\partial w}{\partial z}\frac{\partial b}{\partial y} + f\frac{\partial v_{ag}}{\partial z}\left(f - \frac{\partial u_g}{\partial y}\right) - \frac{\partial w}{\partial y}N^2 = -2Q^{(y)}. \quad (44.101)$$

Another use of thermal wind brings this result to the form

$$f\frac{\partial u_{ag}}{\partial z}\frac{\partial v_g}{\partial y} + f\frac{\partial u_{ag}}{\partial y}\frac{\partial v_g}{\partial z} + f\frac{\partial u_g}{\partial z}\left(\frac{\partial v_{ag}}{\partial y} - \frac{\partial w}{\partial z}\right) + f\frac{\partial v_{ag}}{\partial z}\left(f - \frac{\partial u_g}{\partial y}\right) - \frac{\partial w}{\partial y}N^2 = -2Q^{(y)}. \quad (44.102)$$

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 (44.102) becomes a diagnostic equation for the ageostrophic overturning streamfunction.

44.6.4 Ageostrophic overturning circulation for a symmetric front

The general result (44.102) is now specialized by assuming the zonal velocity is purely geostrophic (as in a zonal geostrophic front) so that

$$u_{ag} = 0. \quad (44.103)$$

For this flow, the ageostrophic flow in the cross-flow/depth plane is non-divergent

$$\frac{\partial v_{ag}}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (44.104)$$

The diagnostic equation (44.102) now takes on the specialized form for a symmetric front

$$-2f\frac{\partial u_g}{\partial z}\frac{\partial w}{\partial z} + f(f + \zeta_g)\frac{\partial v_{ag}}{\partial z} - \frac{\partial w}{\partial y}N^2 = -2Q^{(y)}, \quad (44.105)$$

where

$$\zeta_g = -\frac{\partial u_g}{\partial y} \quad (44.106)$$

is the vertical component of the geostrophic relative vorticity. Introducing an overturning streamfunction for the cross-flow/vertical ageostrophic circulation

$$v_{ag} = \frac{\partial \psi}{\partial z} \quad w = -\frac{\partial \psi}{\partial y}, \quad (44.107)$$

and using thermal wind brings equation (44.105) into the form

$$(N^2 \partial_{yy} - 2 \partial_y b \partial_{yz} + f(f + \zeta_g) \partial_{zz})\psi = -2Q^{(y)}. \quad (44.108)$$

Equation (44.108) 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 $Q^{(y)}$.

44.6.5 Connection to potential vorticity and symmetric instability

The partial differential equation (44.108) can be written

$$\mathcal{K}\psi = -2Q^{(y)}, \quad (44.109)$$

where

$$\mathcal{K} = N^2 \partial_{yy} - 2 \partial_y b \partial_{zy} + f(f + \zeta_g) \partial_{zz} \quad (44.110)$$

is a linear partial differential operator that is a function of the geostrophic flow and the buoyancy. Following the considerations in Section 1.4.3, we know that this operator is elliptic if the following inequality holds

$$\left(\frac{\partial b}{\partial y}\right)^2 - N^2 f(f + \zeta_g) < 0. \quad (44.111)$$

We can relate the ellipticity condition (44.111) to the Ertel potential vorticity for the Boussinesq geostrophic flow. For this purpose, write the geostrophic vorticity as

$$\boldsymbol{\omega}_g = \nabla \wedge \mathbf{u}_g \quad (44.112a)$$

$$= -\hat{\mathbf{x}} \frac{\partial v_g}{\partial z} + \hat{\mathbf{y}} \frac{\partial u_g}{\partial z} + \hat{\mathbf{z}} \left(\frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \right) \quad (44.112b)$$

$$= -\frac{1}{f} \left(\hat{\mathbf{x}} \frac{\partial b}{\partial x} + \hat{\mathbf{y}} \frac{\partial b}{\partial y} \right) + \hat{\mathbf{z}} \left(\frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \right). \quad (44.112c)$$

If we assume the front is zonally symmetric, then the geostrophic vorticity takes the form

$$\boldsymbol{\omega}_g^{2d} = -\frac{1}{f} \frac{\partial b}{\partial y} \hat{\mathbf{y}} - \hat{\mathbf{z}} \frac{\partial u_g}{\partial y}, \quad (44.113)$$

in which case the Ertel potential vorticity (for the geostrophic and Boussinesq flow) takes the form

$$q_g^{2d} = \nabla b \cdot (\boldsymbol{\omega}_g + f\hat{\mathbf{z}}) \quad (44.114a)$$

$$= -\frac{1}{f} \left(\frac{\partial b}{\partial y} \right)^2 + N^2 \left(f - \frac{\partial u_g}{\partial y} \right). \quad (44.114b)$$

$$= -\frac{1}{f} \left(\frac{\partial b}{\partial y} \right)^2 + N^2 (f + \zeta_g). \quad (44.114c)$$

Ellipticity of the PDE (44.108) is thus assured so long as

$$fq_g^{2d} = -\left(\frac{\partial b}{\partial y}\right)^2 + N^2 f(f + \zeta_g) > 0. \quad (44.115)$$

The PDE (44.5) transitions to a hyperbolic system when $fq_g^{2d} < 0$, which is the condition for symmetric instability detailed in Section 44.5.

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.
- Adcroft, A., and J.-M. Campin, Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models, *Ocean Modelling*, 7(3), 269–284, doi:10.1016/j.ocemod.2003.09.003, 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, 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, S. Legg, B. G. Reichl, 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, in prep*, 2019.
- Aiki, H., T. Jacobson, and T. Yamagata, Parameterizing ocean eddy transports from surface to bottom, *Journal of Geophysical Research*, 31, L19 302, 2004.
- Andrews, D., and M. McIntyre, An exact theory of nonlinear waves on a Lagrangian-mean flow, *J. Fluid Mech.*, 89, 609–646, 1978a.
- Andrews, D. G., and M. E. McIntyre, On wave action and its relatives, *J. Fluid Mech.*, 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.
- 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.
- 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.
- Cushmin-Roisin, B., Subduction, in *Dynamics of the oceanic surface mixed-layer*, pp. 181–196, Hawaii Institute of Geophysical Special Publications, 1987.
- 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.
- DeSzeoeke, R. A., and A. F. Bennett, Microstructure fluxes across density surfaces, *Journal of Physical Oceanography*, 23, 2254–2264, 1993.
- Drazin, P., and W. Reid, *Hydrodynamic stability*, Cambridge University Press, Cambridge, UK, 527 pp, 1981.
- 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, *Phys. 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, GFDDs 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.
- 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, 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.
- 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.
- 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. 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.
- 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.
- 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.
- Killworth, P. D., On the parameterization of eddy transfer Part I: Theory, *J. Mar. Res.*, 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.
- Lin, S. J., A finite volume integration method for computing pressure gradient force in general vertical coordinates, *Quart. J. Roy. Meteorol. Soc.*, 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. 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 Res.*, 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, *Proc. Nat. Acad. Sci. USA*, 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., 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 Res.*, 45, 1977–2010, 1998.
- Munk, W. H., Abyssal recipes, *Deep-Sea Res.*, 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.
- Nurser, A. G., and S. M. Griffies, Relating diffusive surface salinity fluxes to boundary freshwater and salt fluxes, *Journal of Physical Oceanography*, *in prep*, 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, *J. Fluid Mech.*, 122, 347–367, 1982.
- 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.
- Schmitt, R. W., Double diffusion in oceanography, *Annual Review of Fluid Mechanics*, 26, 255–285, 1994.
- 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, *Mon. Weather Rev.*, 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.
- 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.
- 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 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.
- 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.
- 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:10.1029/2008GL036138, 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.