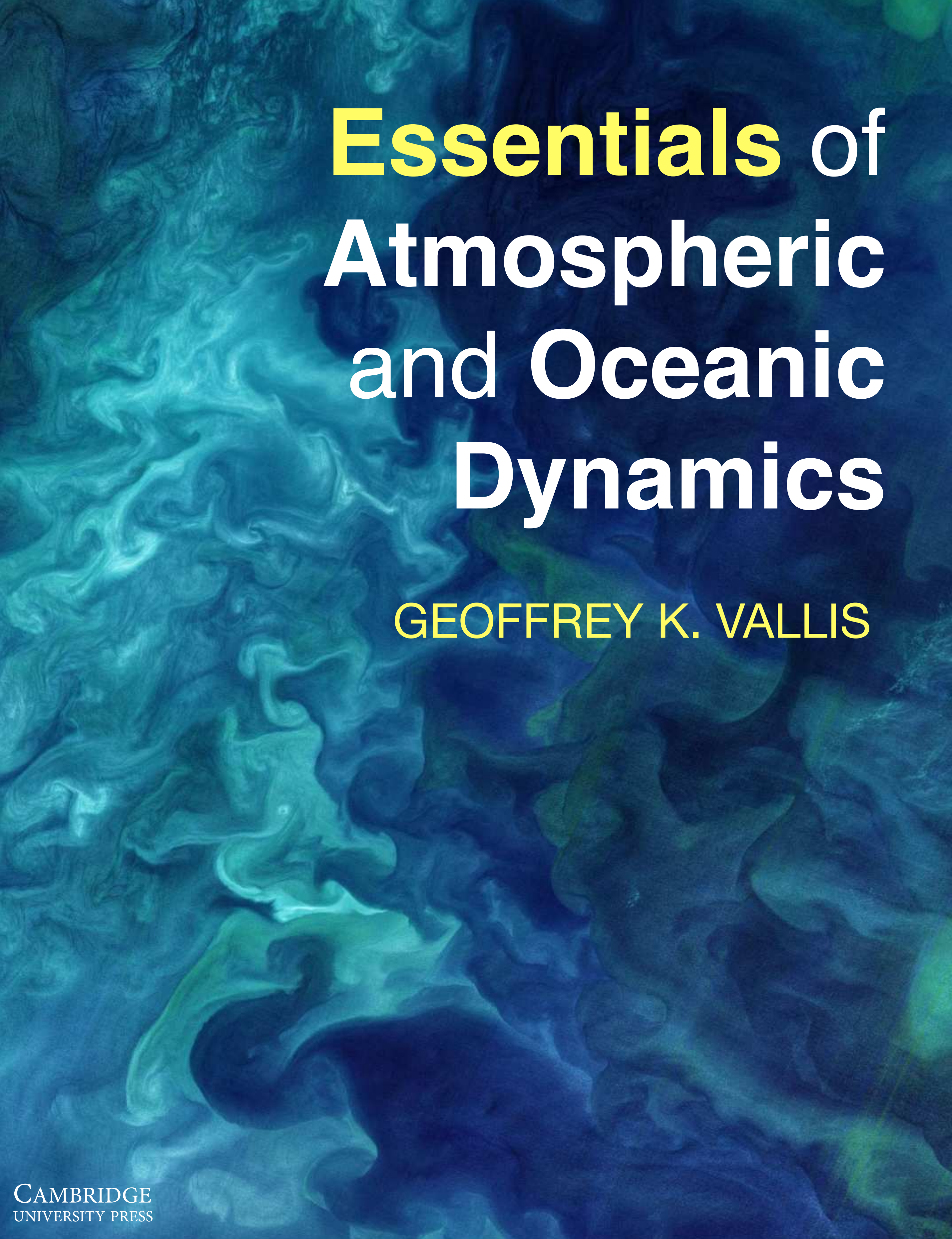

Essentials of Atmospheric and Oceanic Dynamics

GEOFFREY K. VALLIS

CAMBRIDGE
UNIVERSITY PRESS

The "big book" [AOFD] by Vallis is a treasure, but I suspect that this new *Essentials* is destined to be used much more widely in classrooms. Vallis does a superb job of communicating the peculiar tensions between deductive reasoning and physical intuition that underlie this science. The new book is more approachable but no less rigorous. I especially appreciate how the various equation sets are derived in succinct but meaningful ways in the first few chapters, and then used as tools to explore the dynamics in the chapters that follow. It's almost the perfect introductory textbook on this subject, and I plan to use it in my own courses.

Brian E. J. Rose,
University at Albany

He's done it again. In *Essentials*, Geoff Vallis has produced a text that is useful to the student and the experienced scientist alike. While the content is simplified and shortened compared to its parent text, Vallis now provides even more descriptive explanations to support readers in their quest to navigate the physics of fluid flows. These explanations pair well with the theory, serving as an accessible introduction to students while also supporting the more experienced scientist as they put all of the pieces together. This will certainly be a future favourite for reading groups. Even readers with dog-eared versions of the parent book will want a copy of *Essentials*, for in it Vallis has added an entirely new chapter on planetary atmospheres, allowing the interested reader to venture into outer space to apply their newly honed GFD expertise.

Elizabeth A. Barnes,
Colorado State University

For the past decade, Geoff Vallis' book *Atmospheric and Oceanic Fluid Dynamics* has been the "go to" encyclopaedic resource, but it is too lengthy and comprehensive to use as a course textbook. With this superb new shorter volume, Geoff Vallis provides us with the definitive graduate-level textbook, with just the right balance of essential topics alongside glimpses of more advanced topics at the cutting edge of research. The extensive use of margin notes, diamonds to indicate advanced topics, and a comprehensive set of problems will ensure that *Essentials of Atmospheric and Oceanic Dynamics* has much to offer students and researchers at all levels. The book opens with the quote: "Seek simplicity, accept complexity. Exploit simplification, avoid complication." On all counts, this book succeeds magnificently!

David Marshall,
University of Oxford

Vallis' insights into the fundamentals and applications go a long way towards making otherwise complex topics readily grasped by those willing to study. He does not shy away from mathematics where needed, nor does he smother the reader with mathematics where pedagogically unnecessary. Those making it through this book will be ready to tackle a huge suite of research questions related to atmosphere and ocean fluid mechanics. Hence, this book serves an incredibly important role to the academic community. In a nutshell, we need more smart researchers who are adept at atmosphere and ocean dynamics to help understand how those dynamics are increasingly being affected by humanity's choices.

Essentials of Atmospheric and Oceanic Dynamics (EAOD) fills an important niche by offering an articulate and authoritative textbook to be worked through by advanced undergraduates and/or entering graduate students taking courses. The inclusion of exercises in EAOD is incredibly valuable for both students and teachers clamouring for more problem sets to test understanding. Whereas Vallis' previous book, *Atmospheric and Oceanic Fluid Dynamics* (AOFD) is the mother reference, EAOD offers a pedagogical entrée for those wishing to test the waters, including some deep waters. I will happily keep both books on my shelf and make use of them for personal study and to support the teaching of geophysical fluid dynamics.

Vallis has a clear writing style that brings the reader into the subject in an authoritative and friendly manner. He is a wise guru and gentle tutor. The subject of ocean and atmosphere fluid mechanics has matured greatly through his efforts at writing AOFD. EAOD furthers that maturation by allowing for a broader readership to tap into his brain. Well done Geoff!

Stephen M. Griffies,
GFDL, Princeton University.

As its parent book became the bible of the field, but also grew in size and the number of topics it covered in its latest edition, this new book provides a perfect balance and introduction to the essential topics, giving a quick reference without going into all the details. In the Vallis tradition, it is presented clearly, perfectly packaged, and is well organized for both atmospheric and oceanic fluid dynamics. Its simplicity will make it majestically appealing both for people outside the discipline looking for an accessible, yet complete, introduction, and for students within the field at all levels. The inclusion of planetary atmospheres broadens the scope and makes it appealing to a wider and growing audience. Anyone with a background in physics can get the essentials using this book.

Yohai Kaspi,
Weizmann Institute of Science

Contents

Note: In the text itself more advanced sections are marked with a diamond, ✦, and may be omitted on a first reading. If a section is so marked then the marking applies to all the subsections within it.

Preface

Seek simplicity, accept complexity.
Exploit simplification, avoid complication.

This is an introductory book on the dynamics of atmospheres and oceans, with a healthy dose of geophysical fluid dynamics. It is written roughly at the level of advanced or upper-division undergraduates and beginning graduate students, but parts of it will be accessible to first- or second-year undergraduates and I hope that practising scientists will also find it useful. The book is designed for students and scientists who want an introduction to the subject but who may not want all the detail, at least not yet, and its prerequisites are just familiarity with some vector calculus and basic classical physics. Thus, it is meant to be accessible to non-specialists and students who will not necessarily go on to become professional dynamicists. However, as well as very basic material the book does include some elementary introductions to a few 'advanced' topics, such as the residual circulation and turbulence theory, as well as material on the general circulation of the atmosphere and ocean. The more advanced parts could easily be omitted for a first course and, like difficult ski slopes, are marked with a diamond, ✦. Readers may explore these topics more in the references provided, or in this book's parent, *Atmospheric and Oceanic Fluid Dynamics.* Nearly all the topics in this book, except those in the chapter on planetary atmospheres, are dealt with in greater detail there.

What is in the book

The book is divided into three Parts. The first, and longest, provides the foundation for the study of the dynamics of the atmosphere and ocean. It does not assume any prior knowledge of fluid dynamics or thermodynamics, although readers who have such knowledge may be able to skim Chapter 1. The rest of Part I provides an introduction to 'geophysical fluid dynamics', the subject that remains at the heart of atmospheric and oceanic dynamics and without which the subject would be largely qualitative and/or computational. Here we discuss the effects of rotation and stratification, leading into shallow water theory and the quasi-geostrophic and planetary-geostrophic equations. Rossby waves, gravity waves, baroclinic instability and elementary treatments of wave–mean-flow interaction and turbulence round out Part I.

Parts II and III focus on the large scale dynamics and circulation of the atmosphere and ocean, respectively. Our main focus in both Parts is what is sometimes called 'the general circulation', meaning the large-scale quasi-steady and/or time-averaged circulation, but this circulation depends on the effects of time-dependent eddies — the atmosphere's Ferrel Cell may be considered to be 'driven' by the effects of baroclinic instability and Rossby waves. And the El Niño phenomenon, described in the final chapter, is explicitly time dependent. One feature of this book that is not in the parent book is a chapter discussing some of the general principles of planetary atmospheres, a topic of increasing interest because of the new, sometimes quite spectacular, observations of the planets in our Solar System and beyond.

How to use the book

The contents of the book are about enough for a two-term course in atmosphere–ocean dynamics. A term-long, first course in geophysical fluid dynamics could, for example, be based on Part I, omitting some of the earlier or later chapters depending on the students' backgrounds and interests. A term-long course in atmospheric and/or oceanic circulation could be based on Part II and/or Part III, supplementing the material with review articles or research papers as needed, perhaps using data sets to look at the real world (and other planets, if Chapter 13 is to be studied). Alternatively, one could combine aspects of Parts I and II, or Parts I and III, to construct an 'Atmospheric Dynamics' or 'Oceanic Dynamics' course.

If the book is to be used for self-study it could simply be read from beginning to end, although many other pathways are possible and may be preferable. Parts II and III depend on the material in Part I, but the material is reasonably self-contained, and readers who already have some knowledge of geophysical fluid dynamics should feel free to start at a later chapter, or with Part II or Part III. A few problems are collected at the end of some chapters; these are designed to test understanding as well as to fill in gaps and extend the material in the book itself. Many other problems at varying levels of difficulty can be found on the web site of this book, which can easily be found with a search engine. The reader will also see a number of margin notes throughout the book, rather like the ones to the left. The book itself was typeset using LaTeX with Crimson fonts for text, Cronos Pro for sans serif and Minion Math for equations.

Margin notes that are set in a roman (i.e., upright) font emphasize or expand on something that is in the main text.

Margin notes set in italics are asides or historical anecdotes.

I would like to thank Matt Lloyd, Zoë Pruce and Richard Smith at Cambridge University Press for their expert guidance through the writing and production process, as well as many colleagues and students — too many to list, but they know who they are — for their many comments, corrections and criticisms. If you, the reader, have other comments, major or minor, do please contact me.

Part I

Geophysical Fluids

CHAPTER

1 Fluid Fundamentals

FLUIDS, LIKE SOLIDS, move if they are pushed and they warm if they are heated. But, unlike solids, they flow and deform. In this chapter we establish the governing equations of motion for a fluid, with particular attention to air and seawater — the fluids of the atmosphere and ocean, respectively. Readers who already have knowledge of fluid dynamics may skim this chapter and begin reading more seriously at Chapter 2, where we begin to look at the effects of rotation and stratification.

The fluid dynamical equations of motion determine the evolution of a fluid. The equations are based on Newton's laws of motion and the laws of thermodynamics, and embody the principles of conservation of momentum, energy and mass. Initial conditions and boundary conditions are needed to solve the equations.

1.1 Time Derivatives for Fluids

1.1.1 Field and Material Viewpoints

In solid-body mechanics one is normally concerned with the position and momentum of an identifiable object, such as a football or a planet, as it moves through space. In principle we could treat fluids the same way and try to follow the properties of individual fluid parcels as they flow along, perhaps getting hotter or colder as they move. This perspective is known as the *material* or *Lagrangian* viewpoint. However, in fluid dynamical problems we generally would like to know what the values of velocity, density and so on are at *fixed points* in space as time passes. A weather forecast we care about tells us how warm it will be where we live and, if we are given that, we may not care where a particular fluid parcel comes from or where it subsequently goes. Since the fluid is a continuum, this knowledge is equivalent to knowing how the fields of the dynamical variables evolve in space and time. This viewpoint is known as the *field* or *Eulerian* viewpoint.

Although the field viewpoint will often turn out to be the most practically useful, the material description is invaluable both in deriving the equations and in the subsequent insight it frequently provides. This is because the important quantities from a fundamental point of view are often those which are associated with a given fluid element: it is these which directly enter Newton's laws of motion and the thermodynamic equations. It is thus important to have a relationship between the rate of change of quantities associated with a given fluid element and the local rate of change of a field. The material derivative (also called the advective derivative or Lagrangian derivative) provides this relationship.

The Lagrangian viewpoint is named for the Franco-Italian J. L. Lagrange (1736–1813), one of the most renowned mathematicians of his time. The Eulerian point of view is named for Leonhard Euler (1707–1783), the great Swiss mathematician. In fact, Euler is also largely responsible for the Lagrangian view, but the attribution became tangled over time.

1.1.2 The Material Derivative of a Fluid Property

A *fluid element* is an infinitesimal, indivisible, piece of fluid — effectively a very small fluid parcel of fixed mass. The *material derivative,* or the *Lagrangian derivative,* is the rate of change of a property (such as temperature or momentum) of a particular fluid element or finite mass of fluid; that is, it is the total time derivative of a property of a piece of fluid.

Let us suppose that a fluid is characterized by a given velocity field $\boldsymbol{v}(\boldsymbol{x}, t)$, which determines its velocity throughout. Let us also suppose that the fluid has another property φ, and let us seek an expression for the rate of change of φ of a fluid element. Since φ is changing in time and in space we use the chain rule,

$$\delta\varphi = \frac{\partial\varphi}{\partial t}\delta t + \frac{\partial\varphi}{\partial x}\delta x + \frac{\partial\varphi}{\partial y}\delta y + \frac{\partial\varphi}{\partial z}\delta z = \frac{\partial\varphi}{\partial t}\delta t + \delta\boldsymbol{x}\cdot\nabla\varphi. \tag{1.1}$$

This is true in general for any δt, δx, etc. The total time derivative is then

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{\partial\varphi}{\partial t} + \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}\cdot\nabla\varphi. \tag{1.2}$$

If this equation is to provide a material derivative we must identify the time derivative in the second term on the right-hand side with the rate of change of position of a fluid element, namely its velocity. Hence, the material derivative of the property φ is

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{\partial\varphi}{\partial t} + \boldsymbol{v}\cdot\nabla\varphi. \tag{1.3}$$

The right-hand side expresses the material derivative in terms of the local rate of change of φ plus a contribution arising from the spatial variation of φ, experienced only as the fluid parcel moves. Because the material derivative is so common, and to distinguish it from other derivatives, we denote it by the operator $\mathrm{D}/\mathrm{D}t$. Thus, the material derivative of the field φ is

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \frac{\partial\varphi}{\partial t} + (\boldsymbol{v}\cdot\nabla)\varphi. \tag{1.4}$$

The brackets in the last term of this equation are helpful in reminding us that $(\boldsymbol{v}\cdot\nabla)$ is an operator acting on φ. The operator $\partial/\partial t + (\boldsymbol{v}\cdot\nabla)$ is the *Eulerian representation of the Lagrangian derivative as applied to a field.*

Material derivative of vector field

The material derivative may act on a vector field $\boldsymbol{b}$, in which case

$$\frac{\mathrm{D}\boldsymbol{b}}{\mathrm{D}t} = \frac{\partial \boldsymbol{b}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{b}. \tag{1.5}$$

In Cartesian coordinates this is

$$\frac{\mathrm{D}\boldsymbol{b}}{\mathrm{D}t} = \frac{\partial \boldsymbol{b}}{\partial t} + u\frac{\partial \boldsymbol{b}}{\partial x} + v\frac{\partial \boldsymbol{b}}{\partial y} + w\frac{\partial \boldsymbol{b}}{\partial z}, \tag{1.6}$$

and for a particular component of $\boldsymbol{b}$, b^x say,

$$\frac{\mathrm{D}b^x}{\mathrm{D}t} = \frac{\partial b^x}{\partial t} + u\frac{\partial b^x}{\partial x} + v\frac{\partial b^x}{\partial y} + w\frac{\partial b^x}{\partial z}, \tag{1.7}$$

and similarly for b^y and b^z. In coordinate systems other than Cartesian the advective derivative of a vector is not simply the sum of the advective derivatives of its components, because the coordinate vectors themselves change direction with position; this will be important when we deal with spherical coordinates.

The Eulerian derivative is the rate of change of a property at a fixed location in space. The material derivative is the rate of change of a property of a given piece of fluid, which may be moving and so changing its position.

1.1.3 Material Derivative of a Volume

The volume that a given, unchanging, mass of fluid occupies is deformed and advected by the fluid motion, and there is no reason why it should remain constant. Rather, the volume will change as a result of the movement of each element of its bounding material surface, and in particular it will change if there is a non-zero normal component of the velocity at the fluid surface. That is, if the volume of some fluid is $\int \mathrm{d}V$, then

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \mathrm{d}V = \int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}, \tag{1.8}$$

where the subscript V indicates that the integral is a definite integral over some finite volume V, and the limits of the integral are functions of time since the volume is changing. The integral on the right-hand side is over the closed surface, S, bounding the volume. Although intuitively apparent (to some), this expression may be derived more formally using Leibniz's formula for the rate of change of an integral whose limits are changing. Using the divergence theorem on the right-hand side, (1.8) becomes

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \mathrm{d}V = \int_V \nabla \cdot \boldsymbol{v}\,\mathrm{d}V. \tag{1.9}$$

The rate of change of the volume of an infinitesimal fluid element of volume ΔV is obtained by taking the limit of this expression as the volume tends to zero, giving

$$\lim_{\Delta V \to 0} \frac{1}{\Delta V}\frac{\mathrm{D}\Delta V}{\mathrm{D}t} = \nabla \cdot \boldsymbol{v}. \tag{1.10}$$

We will often write such expressions informally as

$$\frac{D\Delta V}{Dt} = \Delta V \nabla \cdot \boldsymbol{v}, \tag{1.11}$$

with the limit implied.

Consider now the material derivative of some fluid property, ξ say, multiplied by the volume of a fluid element, ΔV. Such a derivative arises when ξ is the amount per unit volume of ξ-substance — the mass density or the amount of a dye per unit volume, for example. Then we have

$$\frac{D}{Dt}(\xi \Delta V) = \xi \frac{D\Delta V}{Dt} + \Delta V \frac{D\xi}{Dt}. \tag{1.12}$$

Using (1.11) this becomes

$$\frac{D}{Dt}(\xi \Delta V) = \Delta V \left(\xi \nabla \cdot \boldsymbol{v} + \frac{D\xi}{Dt} \right), \tag{1.13}$$

and the analogous result for a finite fluid volume is just

$$\frac{D}{Dt} \int_V \xi \, dV = \int_V \left(\xi \nabla \cdot \boldsymbol{v} + \frac{D\xi}{Dt} \right) dV. \tag{1.14}$$

This expression is to be contrasted with the Eulerian derivative for which the volume, and so the limits of integration, are fixed and we have

$$\frac{d}{dt} \int_V \xi \, dV = \int_V \frac{\partial \xi}{\partial t} \, dV. \tag{1.15}$$

Now consider the material derivative of a fluid property φ multiplied by the mass of a fluid element, $\rho \Delta V$, where ρ is the fluid density. Such a derivative arises when φ is the amount of φ-substance per unit mass (note, for example, that the momentum of a fluid element is $\rho \boldsymbol{v} \Delta V$). The material derivative of $\varphi \rho \Delta V$ is given by

$$\frac{D}{Dt}(\varphi \rho \Delta V) = \rho \Delta V \frac{D\varphi}{Dt} + \varphi \frac{D}{Dt}(\rho \Delta V). \tag{1.16}$$

But $\rho \Delta V$ is just the mass of the fluid element, and that is constant — that is how a fluid element is defined. Thus the second term on the right-hand side vanishes and

$$\frac{D}{Dt}(\varphi \rho \Delta V) = \rho \Delta V \frac{D\varphi}{Dt} \qquad \text{and} \qquad \frac{D}{Dt} \int_V \varphi \rho \, dV = \int_V \rho \frac{D\varphi}{Dt} \, dV, \tag{1.17a,b}$$

where (1.17b) applies to a finite volume. That expression may also be derived more formally using Leibniz's formula for the material derivative of an integral, and the result also holds when φ is a vector. The result is quite different from the corresponding Eulerian derivative, in which the volume is kept fixed; in that case we have:

$$\frac{d}{dt} \int_V \varphi \rho \, dV = \int_V \frac{\partial}{\partial t}(\varphi \rho) \, dV. \tag{1.18}$$

Various material and Eulerian derivatives are summarized in the shaded box on the facing page.

Material and Eulerian Derivatives

The material derivatives of a scalar (φ) and a vector ($\boldsymbol{b}$) field are given by:

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \frac{\partial \varphi}{\partial t} + \boldsymbol{v} \cdot \nabla \varphi, \qquad \frac{\mathrm{D}\boldsymbol{b}}{\mathrm{D}t} = \frac{\partial \boldsymbol{b}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{b}. \tag{D.1}$$

Various material derivatives of integrals are:

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \varphi \,\mathrm{d}V = \int_V \left(\frac{\mathrm{D}\varphi}{\mathrm{D}t} + \varphi \nabla \cdot \boldsymbol{v}\right) \mathrm{d}V = \int_V \left(\frac{\partial \varphi}{\partial t} + \nabla \cdot (\varphi \boldsymbol{v})\right) \mathrm{d}V, \tag{D.2}$$

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \mathrm{d}V = \int_V \nabla \cdot \boldsymbol{v} \,\mathrm{d}V, \tag{D.3}$$

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \rho\varphi \,\mathrm{d}V = \int_V \rho \frac{\mathrm{D}\varphi}{\mathrm{D}t} \,\mathrm{d}V. \tag{D.4}$$

These formulae also hold if φ is a vector. The Eulerian derivative of an integral is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \varphi \,\mathrm{d}V = \int_V \frac{\partial \varphi}{\partial t} \,\mathrm{d}V, \tag{D.5}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \mathrm{d}V = 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\varphi \,\mathrm{d}V = \int_V \frac{\partial \rho\varphi}{\partial t} \,\mathrm{d}V. \tag{D.6}$$

1.2 The Mass Continuity Equation

In classical mechanics mass is absolutely conserved and in solid-body mechanics we normally do not need an explicit equation of mass conservation. However, in fluid mechanics a fluid may flow into and away from a particular location, and fluid density may change, and we need an equation to describe that change.

1.2.1 An Eulerian Derivation

We first derive the mass conservation equation from an Eulerian point of view; that is, our reference frame is fixed in space and the fluid flows through it. Consider an infinitesimal, rectangular cuboid, control volume, $\Delta V = \Delta x \Delta y \Delta z$ that is fixed in space, as in Fig. 1.1. Fluid moves into or out of the volume through its surface, including through its faces in the y–z plane of area $\Delta A = \Delta y \Delta z$ at coordinates x and $x + \Delta x$. The accumulation of fluid within the control volume due to motion in the x-direction is

CHAPTER

2

Equations for a Rotating Planet

PLANETS ARE ALMOST SPHERES. They also rotate. Here we consider how the equations of motion are affected by these facts, first by looking at how rotation affects the dynamics and then by expressing the equations in spherical coordinates.

2.1 EQUATIONS IN A ROTATING FRAME OF REFERENCE

Newton's second law of motion, that the rate of change of momentum of a body is proportional to the imposed force, applies in so-called inertial frames of reference that are either stationary or moving only with a constant rectilinear velocity relative to the distant galaxies. Now Earth spins around its axis once a day, so the surface of the Earth is *not* an inertial frame. Nevertheless, it is very convenient to describe the motion of the atmosphere or ocean relative to Earth's surface rather than in some inertial frame. How we do that is the subject of this section.

2.1.1 Rate of Change of a Vector

Consider first a vector $\boldsymbol{C}$ of constant length rotating relative to an inertial frame at a constant angular velocity $\boldsymbol{\Omega}$. Then, in a frame rotating with that same angular velocity it appears stationary and constant. If in a small interval of time δt the vector $\boldsymbol{C}$ rotates through a small angle $\delta\lambda$ then the change in $\boldsymbol{C}$, as perceived in the inertial frame, is given by (see Fig. 2.1)

$$\delta \boldsymbol{C} = |\boldsymbol{C}| \cos\vartheta \, \delta\lambda \, \boldsymbol{m}, \tag{2.1}$$

where the vector $\boldsymbol{m}$ is the unit vector in the direction of change of $\boldsymbol{C}$, which is perpendicular to both $\boldsymbol{C}$ and $\boldsymbol{\Omega}$. But the rate of change of the angle λ is just, by definition, the angular velocity so that $\delta\lambda = |\boldsymbol{\Omega}|\delta t$ and

$$\delta \boldsymbol{C} = |\boldsymbol{C}||\boldsymbol{\Omega}| \sin\widehat{\vartheta}\, \boldsymbol{m}\, \delta t = \boldsymbol{\Omega} \times \boldsymbol{C}\, \delta t, \tag{2.2}$$

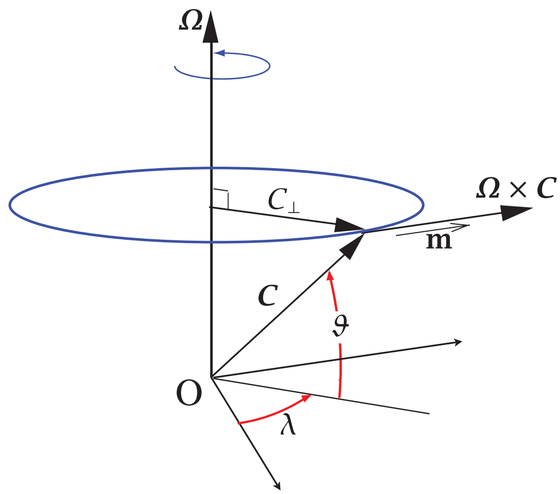

Fig. 2.1: A vector C rotating at an angular velocity Ω. It appears to be a constant vector in the rotating frame, whereas in the inertial frame it evolves according to $(\mathrm{d}C/\mathrm{d}t)_I = \Omega \times C$.

using the definition of the vector cross-product, where $\widehat{\vartheta} = (\pi/2 - \vartheta)$ is the angle between $\boldsymbol{\Omega}$ and $\boldsymbol{C}$. Thus

$$\left(\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}\right)_I = \boldsymbol{\Omega} \times \boldsymbol{C}, \tag{2.3}$$

where the left-hand side is the rate of change of $\boldsymbol{C}$ as perceived in the inertial frame.

Now consider a vector $\boldsymbol{B}$ that changes in the inertial frame. In a small time δt the change in $\boldsymbol{B}$ as seen in the rotating frame is related to the change seen in the inertial frame by

$$(\delta \boldsymbol{B})_I = (\delta \boldsymbol{B})_R + (\delta \boldsymbol{B})_{\mathrm{rot}}, \tag{2.4}$$

where the terms are, respectively, the change seen in the inertial frame, the change due to the vector itself changing as measured in the rotating frame, and the change due to the rotation. Using (2.2) $(\delta \boldsymbol{B})_{\mathrm{rot}} = \boldsymbol{\Omega} \times \boldsymbol{B}\,\delta t$, and so the rates of change of the vector $\boldsymbol{B}$ in the inertial and rotating frames are related by

$$\boxed{\left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{B}.} \tag{2.5}$$

This relation applies to a vector $\boldsymbol{B}$ that, as measured at any one time, is the same in both inertial and rotating frames.

2.1.2 Velocity and Acceleration in a Rotating Frame

The velocity of a body is not measured to be the same in the inertial and rotating frames, so care must be taken when applying (2.5) to velocity. First apply (2.5) to $\boldsymbol{r}$, the position of a particle, to obtain

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{r} \tag{2.6}$$

or

$$\boldsymbol{v}_I = \boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}. \tag{2.7}$$

We refer to $\boldsymbol{v}_R$ and $\boldsymbol{v}_I$ as the relative and inertial velocity, respectively, and (2.7) relates the two. Apply (2.5) again, this time to the velocity $\boldsymbol{v}_R$ to give

$$\left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{v}_R, \tag{2.8}$$

or, using (2.7)

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}_I - \boldsymbol{\Omega} \times \boldsymbol{r})\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{v}_R, \tag{2.9}$$

or

$$\left(\frac{\mathrm{d}\boldsymbol{v}_I}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{v}_R + \frac{\mathrm{d}\boldsymbol{\Omega}}{\mathrm{d}t} \times \boldsymbol{r} + \boldsymbol{\Omega} \times \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I. \tag{2.10}$$

Then, noting that

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{r} = (\boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}), \tag{2.11}$$

and assuming that the rate of rotation is constant, (2.10) becomes

$$\left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R = \left(\frac{\mathrm{d}\boldsymbol{v}_I}{\mathrm{d}t}\right)_I - 2\boldsymbol{\Omega} \times \boldsymbol{v}_R - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}). \tag{2.12}$$

This equation may be interpreted as follows. The term on the left-hand side is the rate of change of the relative velocity as measured in the rotating frame. The first term on the right-hand side is the rate of change of the inertial velocity as measured in the inertial frame (the inertial acceleration, which is, by Newton's second law, equal to the force on a fluid parcel divided by its mass). The second and third terms on the right-hand side (including the minus signs) are the *Coriolis force* and the *centrifugal force* per unit mass. Neither of these are usually regarded as true forces — they may be thought of as quasi-forces (i.e., 'as if' forces); that is, when a body is observed from a rotating frame it behaves as if unseen forces are present that affect its motion.

Centrifugal force

If $\boldsymbol{r}_\perp$ is the perpendicular distance from the axis of rotation (see Fig. 2.1 and substitute $\boldsymbol{r}$ for $\boldsymbol{C}$), then, because $\boldsymbol{\Omega}$ is perpendicular to $\boldsymbol{r}_\perp$, $\boldsymbol{\Omega} \times \boldsymbol{r} = \boldsymbol{\Omega} \times \boldsymbol{r}_\perp$. Then, using the vector identity $\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}_\perp) = (\boldsymbol{\Omega} \cdot \boldsymbol{r}_\perp)\boldsymbol{\Omega} - (\boldsymbol{\Omega} \cdot \boldsymbol{\Omega})\boldsymbol{r}_\perp$ and noting that the first term is zero, we see that the centrifugal force per unit mass is just given by

$$\boldsymbol{F}_{\mathrm{ce}} = -\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}) = \Omega^2 \boldsymbol{r}_\perp. \tag{2.13}$$

This may usefully be written as the gradient of a scalar potential,

$$\boldsymbol{F}_{\mathrm{ce}} = -\nabla \Phi_{\mathrm{ce}}. \tag{2.14}$$

where $\Phi_{\mathrm{ce}} = -(\Omega^2 r_\perp^2)/2 = -(\boldsymbol{\Omega} \times \boldsymbol{r}_\perp)^2/2$.

Coriolis force

The Coriolis force per unit mass is given by

$$F_{\text{Co}} = -2\boldsymbol{\Omega} \times \boldsymbol{v}_R. \tag{2.15}$$

We consider the effects of the Coriolis force extensively, but for now we just note three basic properties:

(i) There is no Coriolis force on bodies that are stationary in the rotating frame.

(ii) The Coriolis force acts to deflect moving bodies at right angles to their direction of travel.

(iii) The Coriolis force does no work on a body because it is perpendicular to the velocity, and so $\boldsymbol{v}_R \cdot (\boldsymbol{\Omega} \times \boldsymbol{v}_R) = 0$.

The Coriolis force is named for Gaspard-Gustave de Coriolis (1792–1843) who discussed the effect in an engineering context in 1835, although the basic effect may have been first recognized (as were so many things) by Leonhard Euler (1707–1783). The (now-called) Coriolis term is also contained in Laplace's tidal equations, formulated in 1776, published in English in Laplace (1832). William Ferrel (1817–1891) was perhaps the first to appreciate the effect of the force on Earth's circulation, identifying and discussing the relevant term ($2\Omega v \sin \vartheta$) in Laplace's equations.

2.1.3 Equations of Motion in a Rotating Frame

Momentum equation

Since (2.12) simply relates the accelerations of a particle in the inertial and rotating frames, then in the rotating frame of reference the three-dimensional momentum equation may be written

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v} = -\frac{1}{\rho}\nabla p - \nabla \Phi_{\text{ce}} + \boldsymbol{g}, \tag{2.16}$$

where all velocities and accelerations are measured with respect to the inertial frame. Since the centrifugal term does not vary with the fluid motion we can incorporate it into gravitational force, $\boldsymbol{g}$, so giving an 'effective gravity' that varies slightly with position over Earth's surface.

Mass continuity and the thermodynamic equation

The mass conservation equation and the thermodynamic equation are unchanged in a rotating frame. To see this consider the material derivative of some variable, φ, such as temperature or density. The material derivative is just the rate of change of φ of an identifiable fluid parcel and that clearly does not depend on the reference frame. Thus, without further ado, we can write

$$\left(\frac{\mathrm{D}\varphi}{\mathrm{D}t}\right)_R = \left(\frac{\mathrm{D}\varphi}{\mathrm{D}t}\right)_I, \tag{2.17}$$

where the material derivatives are $(\mathrm{D}\varphi/\mathrm{D}t)_R = (\partial\varphi/\partial t)_R + \boldsymbol{v}_R \cdot \nabla\varphi$ and $(\mathrm{D}\varphi/\mathrm{D}t)_I = (\partial\varphi/\partial t)_I + \boldsymbol{v}_I \cdot \nabla\varphi$. The individual terms differ in the two frames; that is $(\partial\varphi/\partial t)_R \neq (\partial\varphi/\partial t)_I$, but the material derivatives are equal.

Further, the divergence operator is the same in the inertial and rotating frame. Using (2.7), we have that

$$\nabla \cdot \boldsymbol{v}_I = \nabla \cdot (\boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}) = \nabla \cdot \boldsymbol{v}_R \tag{2.18}$$

since $\nabla \cdot (\boldsymbol{\Omega} \times \boldsymbol{r}) = 0$. Thus, using (2.17) and (2.18), the mass conservation equation (1.27b) may be written

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla \cdot \boldsymbol{v}_R = 0, \tag{2.19}$$

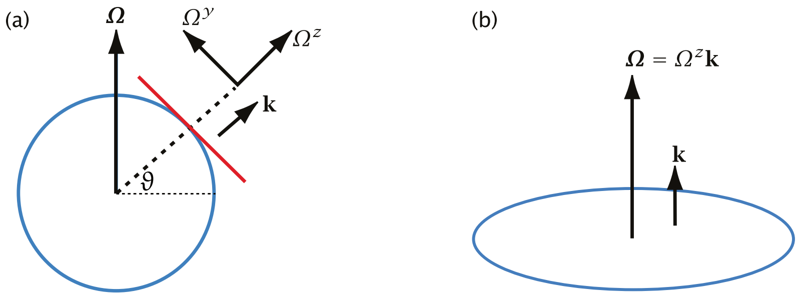

Fig. 2.3: (a) On the sphere the rotation vector $\boldsymbol{\Omega}$ can be decomposed into two components, one in the local vertical and one in the local horizontal, pointing toward the pole. That is, $\boldsymbol{\Omega} = \Omega_y\hat{\mathbf{j}} + \Omega_z\hat{\mathbf{k}}$ where $\Omega_y = \Omega\cos\vartheta$ and $\Omega_z = \Omega\sin\vartheta$. In geophysical fluid dynamics, the rotation vector in the local vertical is often the more important component in the horizontal momentum equations. On a rotating disk, (b), the rotation vector $\boldsymbol{\Omega}$ is parallel to the local vertical $\hat{\mathbf{k}}$.

The thermodynamic equation, in potential temperature form, is just an advection equation so that using (2.20), its (adiabatic) spherical coordinate form is

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \frac{\partial\theta}{\partial t} + \frac{u}{r\cos\vartheta}\frac{\partial\theta}{\partial\lambda} + \frac{v}{r}\frac{\partial\theta}{\partial\vartheta} + w\frac{\partial\theta}{\partial r} = 0, \tag{2.35}$$

and similarly for tracers such as water vapour or salt.

Momentum equation

Recall that the inviscid momentum equation is:

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega}\times\boldsymbol{v} = -\frac{1}{\rho}\nabla p - \nabla\Phi, \tag{2.36}$$

where Φ is the geopotential. In spherical coordinates the directions of the coordinate axes change with position and so the component expansion of (2.36) is

$$\begin{aligned}\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} &= \frac{\mathrm{D}u}{\mathrm{D}t}\hat{\mathbf{i}} + \frac{\mathrm{D}v}{\mathrm{D}t}\hat{\mathbf{j}} + \frac{\mathrm{D}w}{\mathrm{D}t}\hat{\mathbf{k}} + u\frac{\mathrm{D}\hat{\mathbf{i}}}{\mathrm{D}t} + v\frac{\mathrm{D}\hat{\mathbf{j}}}{\mathrm{D}t} + w\frac{\mathrm{D}\hat{\mathbf{k}}}{\mathrm{D}t} &\text{(2.37a)}\\ &= \frac{\mathrm{D}u}{\mathrm{D}t}\hat{\mathbf{i}} + \frac{\mathrm{D}v}{\mathrm{D}t}\hat{\mathbf{j}} + \frac{\mathrm{D}w}{\mathrm{D}t}\hat{\mathbf{k}} + \boldsymbol{\Omega}_{\text{flow}}\times\boldsymbol{v}, &\text{(2.37b)}\end{aligned}$$

using (2.32). Using either (2.37a) and the expressions for the rates of change of the unit vectors given in (2.32), or (2.37b) and the expression for $\boldsymbol{\Omega}_{\text{flow}}$ given in (2.31), (2.37) becomes

$$\begin{aligned}\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} = \hat{\mathbf{i}}\left(\frac{\mathrm{D}u}{\mathrm{D}t} - \frac{uv\tan\vartheta}{r} + \frac{uw}{r}\right) + \hat{\mathbf{j}}\left(\frac{\mathrm{D}v}{\mathrm{D}t} + \frac{u^2\tan\vartheta}{r} + \frac{vw}{r}\right)\\ + \hat{\mathbf{k}}\left(\frac{\mathrm{D}w}{\mathrm{D}t} - \frac{u^2+v^2}{r}\right).\end{aligned} \tag{2.38}$$

Using the definition of a vector cross-product, the Coriolis term is:

$$\begin{aligned}2\boldsymbol{\Omega}\times\boldsymbol{v} &= \begin{vmatrix}\hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}}\\ 0 & 2\Omega\cos\vartheta & 2\Omega\sin\vartheta\\ u & v & w\end{vmatrix}\\ &= \hat{\mathbf{i}}\,(2\Omega w\cos\vartheta - 2\Omega v\sin\vartheta) + \hat{\mathbf{j}}\,2\Omega u\sin\vartheta - \hat{\mathbf{k}}\,2\Omega u\cos\vartheta.\end{aligned} \tag{2.39}$$

Chapter

3

Dynamics on a Rotating Planet

We now put the equations of motion to use, and in so doing start our journey into the *dynamics* of fluid motion on a rotating planet. We begin rather gently by way of an introduction to scaling, which is the basis of the art of making sensible approximations.

3.1 A Gentle Introduction to Scaling

The units we use to measure length, velocity and so on are irrelevant to the dynamics, and SI units may not be the most appropriate ones for a given problem. Rather, it is useful to express the equations of motion in terms of 'nondimensional' variables, by which we mean expressing every variable as the ratio of its value to some reference value. We try to choose the reference value as a natural one for a given flow, in order that, where possible, the nondimensional variables are order-unity quantities, and doing this is called *scaling the equations*. Much of the art of fluid dynamics lies in choosing sensible scaling factors for the problem at hand for then the sizes of the various terms become clear, and we here we give a simple, non-rotating, example.

3.1.1 The Reynolds Number

Consider the constant-density momentum equation in Cartesian coordinates. If a typical velocity is U, a typical length is L, a typical time scale is T, and a typical value of the pressure deviation is Φ, then the approximate sizes of the various terms in the momentum equation are given by

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\nabla \phi + \nu \nabla^2 \boldsymbol{v}, \tag{3.1a}$$

$$\frac{U}{T} \qquad \frac{U^2}{L} \sim \frac{\Phi}{L} \qquad \nu \frac{U}{L^2}. \tag{3.1b}$$

Osborne Reynolds (1842–1912) was an Irish born (Belfast) physicist who was professor of engineering at Manchester University from 1868–1905. He was also one of the first scientists to think about the concept of group velocity.

The ratio of the inertial (i.e., the advective) terms to the viscous terms is $(U^2/L)/(\nu U/L^2) = UL/\nu$, and this is the *Reynolds number*. More formally, we can nondimensionalize the momentum equation by writing

$$\widehat{\boldsymbol{v}} = \frac{\boldsymbol{v}}{U}, \qquad \widehat{\boldsymbol{x}} = \frac{\boldsymbol{x}}{L}, \qquad \hat{t} = \frac{t}{T}, \qquad \widehat{\phi} = \frac{\phi}{\Phi}, \tag{3.2}$$

where the terms with hats on are *nondimensional* values of the variables and the capitalized quantities are known as *scaling values*, and these are the approximate magnitudes of the variables. We now choose the scaling values so that the nondimensional variables are of order unity, or $\widehat{u} = \mathcal{O}(1)$. Thus, for example, we choose U so that $u = \mathcal{O}(U)$, where the notation should be taken to mean that the magnitude of the variable u is approximately U, or that $u \sim U$, and we say that 'u scales like U'.

In this problem, we have no way to scale pressure and time except with the velocity and length scales we have chosen, and the only dimensionally correct choices are then

$$T = \frac{L}{U}, \qquad \Phi = U^2. \tag{3.3}$$

Substituting (3.2) and (3.3) into the momentum equation gives

$$\frac{U^2}{L}\left[\frac{\partial \widehat{\boldsymbol{v}}}{\partial \hat{t}} + (\widehat{\boldsymbol{v}} \cdot \nabla)\widehat{\boldsymbol{v}}\right] = -\frac{U^2}{L}\nabla\widehat{\phi} + \frac{\nu U}{L^2}\nabla^2\widehat{\boldsymbol{v}}, \tag{3.4}$$

where we use the convention that when ∇ operates on a nondimensional variable it is a nondimensional operator. Equation (3.4) simplifies to

$$\frac{\partial \widehat{\boldsymbol{v}}}{\partial \hat{t}} + (\widehat{\boldsymbol{v}} \cdot \nabla)\widehat{\boldsymbol{v}} = -\nabla\widehat{\phi} + \frac{1}{Re}\nabla^2\widehat{\boldsymbol{v}}, \tag{3.5}$$

where

$$\boxed{Re \equiv \frac{UL}{\nu}} \tag{3.6}$$

is, again, the Reynolds number. If we have chosen our length and velocity scales sensibly — that is, if we have scaled them properly — each variable in (3.5) is order unity, with the viscous term being multiplied by the parameter $1/Re$. There are two important conclusions:

(i) The ratio of the importance of the inertial terms to the viscous terms is given by the *Reynolds number,* defined by (3.6). In the absence of other forces, such as those due to gravity and rotation, the Reynolds number is the only nondimensional parameter explicitly appearing in the momentum equation. Hence its value, along with the boundary conditions and geometry, controls the behaviour of the system.

(ii) More generally, by scaling the equations of motion appropriately the parameters determining the behaviour of the system become explicit. *Scaling the equations is intelligent nondimensionalization.*

Nondimensionalizing the equations does not, however, absolve the investigator from the responsibility of producing dimensionally correct equations. One should regard nondimensional equations as dimensional equations in units appropriate for the problem at hand.

3.2 Hydrostatic Balance

Life is too short to solve every complex problem in detail, and the atmospheric and oceanic sciences abound with complex problems. In their usual form the fluid dynamical equations alone are a set of six nonlinear partial differential equations (three momentum equations, a thermodynamic equation, a mass continuity equation and an equation of state) describing velocity, pressure, temperature and density. To solve real-world problems we need to add water vapour or salinity, as well as the equations of radiative transfer. All this makes for a complex system, and to make progress we need to simplify where possible and eliminate unimportant effects. We have already seen how we might do that for fluids of nearly constant density in making the Boussinesq approximation, and we now look at the effects of gravity and rotation and see how these give rise hydrostatic balance and geostrophic balance, the dominant balances in the vertical and horizontal directions, respectively. The corresponding states, hydrostasy and geostrophy, are not exactly realized, but their approximate satisfaction has profound consequences on the behaviour of atmospheres and oceans.

We begin with hydrostatic balance. We first encountered it in Section 1.3.3 but now we take a closer look. We start by scaling the equations, just as we did in the previous section.

3.2.1 Scaling Estimates

Consider the relative sizes of terms in the vertical momentum equation, (2.42c):

$$\frac{W}{T} + \frac{UW}{L} + \frac{W^2}{H} + \Omega U \sim \left| \frac{1}{\rho}\frac{\partial p}{\partial z} \right| + g. \tag{3.7}$$

For most large-scale motion in the atmosphere and ocean the terms on the right-hand side are orders of magnitude larger than those on the left, and therefore must be approximately equal. Explicitly, suppose $W \sim 1\,\mathrm{cm\,s^{-1}}$, $L \sim 10^5\,\mathrm{m}$, $H \sim 10^3\,\mathrm{m}$, $U \sim 10\,\mathrm{m\,s^{-1}}$, $T = L/U$. Then by substituting into (3.7) it seems that the pressure term is the only one which could balance the gravitational term, and we are led to approximate (2.42c) by,

$$\frac{\partial p}{\partial z} = -\rho g. \tag{3.8}$$

This equation, which is a vertical momentum equation, is known as *hydrostatic balance*.

However, (3.8) is not always a useful equation! Let us suppose that the density is a constant, ρ_0 . We can then write the pressure as

$$p(x, y, z, t) = p_0(z) + p'(x, y, z, t), \qquad \text{where} \qquad \frac{\partial p_0}{\partial z} \equiv -\rho_0 g. \tag{3.9}$$

That is, p_0 and ρ_0 are in hydrostatic balance. On the f-plane, the inviscid vertical momentum equation becomes, without approximation,

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho_0}\frac{\partial p'}{\partial z}. \tag{3.10}$$

Thus, *for constant density fluids the gravitational term has no dynamical effect:* there is no buoyancy force, and the pressure term in the horizontal momentum equations can be replaced by p'. Hydrostatic balance, and in particular (3.9), is not a useful vertical momentum equation in this case. If the fluid is stratified, we should therefore subtract off the hydrostatic pressure associated with the mean density before we can determine whether hydrostasy is a useful *dynamical* approximation, accurate enough to determine the horizontal pressure gradients. This is automatic in the Boussinesq equations, where the vertical momentum equation is

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial \phi}{\partial z} + b, \tag{3.11}$$

and the hydrostatic balance of the basic state is already subtracted out. In the more general equation,

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{3.12}$$

we need to compare the advective term on the left-hand side with the pressure variations arising from horizontal flow in order to determine whether hydrostasy is an appropriate vertical momentum equation.

3.2.2 Hydrostatic Balance and the Aspect Ratio

In a Boussinesq fluid we write the horizontal and vertical momentum equations as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla_z \phi, \qquad \frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial \phi}{\partial z} + b. \tag{3.13a,b}$$

With $\boldsymbol{f} = 0$, (3.13a) implies the scaling

$$\phi \sim U^2. \tag{3.14}$$

If we then use mass conservation, $\nabla_z \cdot \boldsymbol{u} + \partial w/\partial z = 0$, to scale vertical velocity we find

$$w \sim W = \frac{H}{L}U = \alpha U, \tag{3.15}$$

where $\alpha = H/L$ is the aspect ratio. The advective terms in the vertical momentum equation all scale as

$$\frac{\mathrm{D}w}{\mathrm{D}t} \sim \frac{UW}{L} = \frac{U^2 H}{L^2}. \tag{3.16}$$

Using (3.14) and (3.16) the ratio of the advective term to the pressure gradient term in the vertical momentum equations then scales as

$$\frac{|\mathrm{D}w/\mathrm{D}t|}{|\partial\phi/\partial z|} \sim \frac{U^2 H/L^2}{U^2/H} \sim \left(\frac{H}{L}\right)^2. \tag{3.17}$$

Thus, the condition for hydrostasy, that $|\mathrm{D}w/\mathrm{D}t|/|\partial\phi/\partial z| \ll 1$, is

$$\alpha^2 \equiv \left(\frac{H}{L}\right)^2 \ll 1. \tag{3.18}$$

The advective term in the vertical momentum may then be neglected. Thus, *hydrostatic balance arises from a small aspect ratio approximation.*

We can obtain the same result more formally by nondimensionalizing the momentum equations. Using uppercase symbols to denote scaling values we write

$$(x,y) = L(\widehat{x},\widehat{y}), \qquad z = H\widehat{z}, \qquad \boldsymbol{u} = U\widehat{\boldsymbol{u}}, \qquad w = W\widehat{w} = \frac{HU}{L}\widehat{w},$$
$$t = T\hat{t} = \frac{L}{U}\hat{t}, \qquad \phi = \Phi\widehat{\phi} = U^2\widehat{\phi}, \qquad b = B\hat{b} = \frac{U^2}{H}\hat{b}, \tag{3.19}$$

where the hatted variables are nondimensional and the scaling for w is suggested by the mass conservation equation, $\nabla_z \cdot \boldsymbol{u} + \partial w/\partial z = 0$. Substituting (3.19) into (3.13) (with $\boldsymbol{f} = 0$) gives us the nondimensional equations

$$\frac{\mathrm{D}\widehat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} = -\nabla\widehat{\phi}, \qquad \alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}\hat{t}} = -\frac{\partial\widehat{\phi}}{\partial\widehat{z}} + \hat{b}, \tag{3.20a,b}$$

where $\mathrm{D}/\mathrm{D}\hat{t} = \partial/\partial\hat{t} + \widehat{u}\partial/\partial\widehat{x} + \widehat{v}\partial/\partial\widehat{y} + \widehat{w}\partial/\partial\widehat{z}$ and we use the convention that when ∇ operates on nondimensional quantities the operator itself is nondimensional. From (3.20b) it is clear that hydrostatic balance obtains when $\alpha^2 \ll 1$, that is when the aspect ratio is small.

3.3 Geostrophic and Thermal Wind Balance

We now consider the dominant dynamical balance in the horizontal components of the momentum equation. In the horizontal plane (meaning along geopotential surfaces) we find that the Coriolis term is much larger than the advective terms and the dominant balance is between it and the horizontal pressure force. This balance is called *geostrophic balance,* and it occurs when the Rossby number is small, as we now investigate.

3.3.1 The Rossby Number

The *Rossby number* characterizes the importance of rotation in a fluid. It is, essentially, the ratio of the magnitude of the relative acceleration to the Coriolis acceleration, and it is of fundamental importance in geophysical fluid dynamics. It arises from a simple scaling of two of the terms horizontal momentum equation, namely

$$\frac{\partial\boldsymbol{u}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{u} + \boldsymbol{f}\times\boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{3.21a}$$
$$\frac{U^2}{L} \qquad\qquad fU \tag{3.21b}$$

The Rossby number, U/fL, is named for C.-G. Rossby (1898–1957), a Swedish scientist who worked for many years in the United States and who was one of the great pioneers of dynamical meteorology in the mid-twentieth century. The Russian meteorologist I. Kibel introduced a similar number in 1940 and the number is sometimes called the Kibel or Rossby–Kibel number.

axis of rotation, and the flow is effectively two dimensional. This result is known as the *Taylor–Proudman effect,* namely that for constant density flow in geostrophic and hydrostatic balance the vertical derivatives of the horizontal and the vertical velocities are zero. At zero Rossby number, if the vertical velocity is zero somewhere in the flow, it is zero everywhere in that vertical column; furthermore, the horizontal flow has no vertical shear, and the fluid moves like a slab. The effects of rotation have provided a *stiffening* of the fluid in the vertical.

The Taylor–Proudman effect is named for G. I. Taylor and I. Proudman who wrote papers developing the result in 1921 and 1916, respectively. The effect is sometimes called the Taylor–Proudman 'theorem', but it is more usefully thought of as a physical effect, with manifestations even when the conditions for its satisfaction are not precisely met — which they never are.

In neither the atmosphere nor the ocean do we observe precisely such vertically coherent flow, mainly because of the effects of stratification. However, it is typical of geophysical fluid dynamics that the assumptions underlying a derivation are not fully satisfied, yet there are manifestations of it in real flow. For example, one might have naïvely expected, because $\partial w/\partial z = -\nabla_z \cdot \boldsymbol{u}$, that the scales of the various variables would be related by $W/H \sim U/L$. However, if the flow is rapidly rotating we expect that the horizontal flow will be in near geostrophic balance and therefore nearly divergence free; thus $\nabla_z \cdot \boldsymbol{u} \ll U/L$, and $W \ll HU/L$.

3.3.4 Thermal Wind Balance

Thermal wind balance arises by combining the geostrophic and hydrostatic approximations, and this is most easily done in the context of the Boussinesq equations, or in pressure coordinates. Beginning with the Boussinesq equations, geostrophic balance may be written

$$-fv_g = -\frac{\partial \phi}{\partial x}, \qquad f u_g = -\frac{\partial \phi}{\partial y}. \tag{3.37a,b}$$

Combining these relations with hydrostatic balance, $\partial\phi/\partial z = b$, gives

$$\boxed{f\frac{\partial v_g}{\partial z} = \frac{\partial b}{\partial x}, \qquad f\frac{\partial u_g}{\partial z} = -\frac{\partial b}{\partial y}.} \tag{3.38a,b}$$

These equations represent *thermal wind balance,* and the vertical derivative of the geostrophic wind is the 'thermal wind'.

If the density or buoyancy is constant then the right-hand sides of (3.38) are zero and there is no shear, recovering the Taylor–Proudman result. But suppose that the temperature falls in the poleward direction. Then thermal wind balance implies that the (eastward) wind will increase with height — just as is observed in the atmosphere! In general, a vertical shear of the horizontal wind is associated with a horizontal temperature gradient, and this is one of the most simple and far-reaching effects in geophysical fluid dynamics. The underlying physical mechanism is illustrated in Fig. 3.2.

Geostrophic and thermal wind balance in pressure coordinates

In pressure coordinates geostrophic balance is just

$$\boldsymbol{f} \times \boldsymbol{u}_g = -\nabla_p \Phi, \tag{3.39}$$

Chapter

4

Shallow Water Equations

THE *SHALLOW WATER EQUATIONS* are a set of equations that describe, not surprisingly, a shallow layer of fluid, and in particular one that is in hydrostatic balance and has constant density. The equations are useful for two reasons:

(i) They are a simpler set of equations than the full three-dimensional ones, and so allow for a much more straightforward analysis of sometimes complex problems.

(ii) In spite of their simplicity, the equations provide a reasonably realistic representation of a variety of phenomena in atmospheric and oceanic dynamics.

Put simply, the shallow water equations are a very useful *model* for geophysical fluid dynamics. Let's dive head first into the equations and see what they can do for us.

4.1 Shallow Water Equations of Motion

The shallow water equations apply, by definition, to a fluid layer of constant density in which the horizontal scale of the flow is much greater than the layer depth, and which have a free surface at the top (or sometimes at the bottom). Because the fluid is of constant density the fluid motion is fully determined by the momentum and mass continuity equations, and because of the assumed small aspect ratio the hydrostatic approximation is well satisfied, as we discussed in Section 3.2.2. Thus, consider a fluid above which is another fluid of negligible density, as illustrated in Fig. 4.1. Our notation is that $\boldsymbol{v} = u\hat{\mathbf{i}} + v\hat{\mathbf{j}} + w\hat{\mathbf{k}}$ is the three-dimensional velocity and $\boldsymbol{u} = u\hat{\mathbf{i}} + v\hat{\mathbf{j}}$ is the horizontal velocity, $h(x, y)$ is the thickness of the liquid column, H is its mean height, and η is the height of the free surface. In a flat-bottomed container $\eta = h$, whereas in general $h = \eta - \eta_B$, where η_B is the height of the floor of the container.

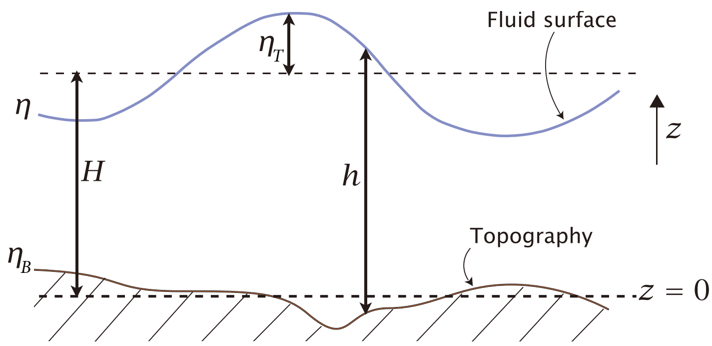

Fig. 4.1: A shallow water system where h is the thickness of a water column, H its mean thickness, η the height of the free surface and η_B is the height of the lower, rigid, surface above some arbitrary origin, typically chosen such that the average of η_B is zero. The quantity η_B is the deviation free surface height so we have $\eta = \eta_B + h = H + \eta_T$.

4.1.1 Momentum Equations

The vertical momentum equation is just the hydrostatic equation,

$$\frac{\partial p}{\partial z} = -\rho_0 g, \tag{4.1}$$

and, because density is assumed constant, we may integrate this to

$$p(x, y, z, t) = -\rho_0 g z + p_o. \tag{4.2}$$

At the top of the fluid, $z = \eta$, the pressure is determined by the weight of the overlying fluid and this is negligible. Thus, $p = 0$ at $z = \eta$, giving

$$p(x, y, z, t) = \rho_0 g(\eta(x, y, t) - z). \tag{4.3}$$

The key assumption underlying the shallow water equations is that of a small aspect ratio, so that $H/L \ll 1$, where H is the fluid depth and L the horizontal scale of motion. This gives rise to the hydrostatic approximation, and this in turn leads to the z-independence of the velocity field and the 'sloshing' nature of the flow.

The consequence of this is that *the horizontal gradient of pressure is independent of height*. That is

$$\nabla_z p = \rho_0 g \nabla_z \eta, \qquad \text{where} \qquad \nabla_z = \hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y}. \tag{4.4}$$

(In the rest of this chapter we drop the subscript z unless that causes ambiguity; the three-dimensional gradient operator is denoted by ∇_3. We also mostly use Cartesian coordinates, but the shallow water equations may certainly be applied over a spherical planet.) The horizontal momentum equations therefore become

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\frac{1}{\rho_0}\nabla p = -g\nabla\eta. \tag{4.5}$$

The right-hand side of this equation is independent of the vertical coordinate z. Thus, if the flow is initially independent of z, it must stay so. (This z-independence is unrelated to that arising from the rapid rotation necessary for the Taylor–Proudman effect.) The velocities u and v are functions of x, y and t only, and the horizontal momentum equation is therefore

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = \frac{\partial \boldsymbol{u}}{\partial t} + u\frac{\partial \boldsymbol{u}}{\partial x} + v\frac{\partial \boldsymbol{u}}{\partial y} = -g\nabla\eta. \tag{4.6}$$

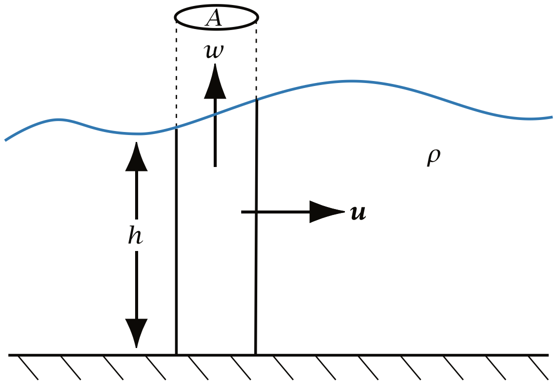

Fig. 4.2: The mass budget for a column of area A in a flat-bottomed shallow water system. The fluid leaving the column is $\oint \rho h \boldsymbol{u} \cdot \boldsymbol{n}\, \mathrm{d}l$ where $\boldsymbol{n}$ is the unit vector normal to the boundary of the fluid column. There is a non-zero vertical velocity at the top of the column if the mass convergence into the column is non-zero.

In the presence of rotation, (4.6) easily generalizes to

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \tag{4.7}$$

where $\boldsymbol{f} = f\hat{\mathbf{k}}$. Just as with the fully three-dimensional equations, f may be constant or may vary with latitude, so that on a spherical planet $f = 2\Omega \sin\vartheta$ and on the β-plane $f = f_0 + \beta y$.

4.1.2 Mass Continuity Equation

The mass contained in a fluid column of height h and cross-sectional area A is given by $\int_A \rho_0 h\, \mathrm{d}\boldsymbol{A}$ (see Fig. 4.2). If there is a net flux of fluid across the column boundary (by advection) then this must be balanced by a net increase in the mass in A, and therefore a net increase in the height of the water column. The mass convergence into the column is given by

$$F_m = \text{mass flux in} = -\int_S \rho_0 \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{S}, \tag{4.8}$$

where S is the area of the vertical boundary of the column. The surface area of the column is composed of elements of area $h\boldsymbol{n}\,\delta l$, where δl is a line element circumscribing the column and $\boldsymbol{n}$ is a unit vector perpendicular to the boundary, pointing outwards. Thus (4.8) becomes

$$F_m = -\oint \rho_0 h \boldsymbol{u} \cdot \boldsymbol{n}\, \mathrm{d}l. \tag{4.9}$$

Using the divergence theorem in two dimensions, (4.9) simplifies to

$$F_m = -\int_A \nabla \cdot (\rho_0 \boldsymbol{u} h)\, \mathrm{d}A, \tag{4.10}$$

where the integral is over the cross-sectional area of the fluid column (looking down from above). This is balanced by the local increase in height of the water column, given by

$$F_m = \frac{\mathrm{d}}{\mathrm{d}t}\int \rho_0\, \mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t}\int_A \rho_0 h\, \mathrm{d}A = \int_A \rho_0 \frac{\partial h}{\partial t}\, \mathrm{d}A. \tag{4.11}$$

The Shallow Water Equations

For a single-layer fluid, and including the Coriolis term, the inviscid shallow water equations are:

$$\text{momentum:} \qquad \frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \tag{SW.1}$$

$$\text{mass continuity:} \qquad \frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla\cdot\boldsymbol{u} = 0, \tag{SW.2}$$

$$\text{or} \qquad \frac{\partial h}{\partial t} + \nabla\cdot(h\boldsymbol{u}) = 0, \tag{SW.3}$$

where $\boldsymbol{u}$ is the horizontal velocity, h is the total fluid thickness, η is the height of the upper free surface, and h and η are related by

$$h(x, y, t) = \eta(x, y, t) - \eta_B(x, y), \tag{SW.4}$$

where η_B is the height of the lower surface (the bottom topography). The material derivative is

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \boldsymbol{u}\cdot\nabla = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}, \tag{SW.5}$$

with the rightmost expression holding in Cartesian coordinates.

Because ρ_0 is constant, the balance between (4.10) and (4.11) leads to

$$\int_A \left[\frac{\partial h}{\partial t} + \nabla\cdot(\boldsymbol{u}h)\right] \mathrm{d}A = 0, \tag{4.12}$$

and because the area is arbitrary the integrand itself must vanish, whence,

$$\boxed{\frac{\partial h}{\partial t} + \nabla\cdot(\boldsymbol{u}h) = 0 \qquad \text{or} \qquad \frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla\cdot\boldsymbol{u} = 0.} \tag{4.13a,b}$$

This derivation holds whether or not the lower surface is flat. If it is, then $h = \eta$, and if not $h = \eta - \eta_B$. Equations (4.7) and (4.13) form a complete set, summarized in the shaded box above.

4.1.3 Reduced Gravity Equations

Consider now a single shallow moving layer of fluid *on top* of a deep, quiescent fluid layer (Fig. 4.3), and beneath a fluid of negligible inertia. This configuration is often used as a model of the upper ocean: the upper layer represents flow in perhaps the upper few hundred metres of the ocean, the lower layer being the near-stagnant abyss. If we turn the model upside-down we have a perhaps slightly less realistic model of the atmosphere: the lower layer represents motion in the troposphere above which lies an

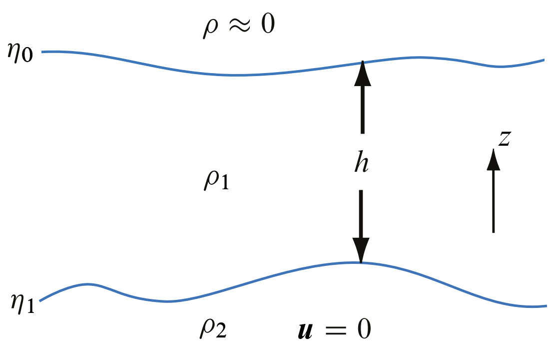

Fig. 4.3: The reduced gravity shallow water system. An active layer lies over a deep, denser, quiescent layer. In a common variation the upper surface is held flat by a rigid lid, and $\eta_0 = 0$.

inactive stratosphere. The equations of motion are virtually the same in both cases, but for definiteness we'll think about the oceanic case.

The pressure in the upper layer is given by integrating the hydrostatic equation down from the upper surface. Thus, at a height z in the upper layer

$$p_1(z) = g\rho_1(\eta_0 - z), \tag{4.14}$$

where η_0 is the height of the upper surface. Hence, everywhere in the upper layer,

$$\frac{1}{\rho_1}\nabla p_1 = g\nabla\eta_0, \tag{4.15}$$

and the momentum equation is

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta_0. \tag{4.16}$$

In the lower layer the pressure is also given by the weight of the fluid above it. Thus, at some level z in the lower layer,

$$p_2(z) = \rho_1 g(\eta_0 - \eta_1) + \rho_2 g(\eta_1 - z). \tag{4.17}$$

But if this layer is motionless the horizontal pressure gradient in it is zero and therefore

$$\rho_1 g\eta_0 = -\rho_1 g'\eta_1 + \text{constant}, \tag{4.18}$$

where $g' = g(\rho_2 - \rho_1)/\rho_1$ is the *reduced gravity*, and in the ocean $\rho_2 - \rho_1)/\rho \ll 1$ and $g' \ll g$. The momentum equation becomes

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = g'\nabla\eta_1. \tag{4.19}$$

The equations are completed by the usual mass conservation equation,

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla\cdot\boldsymbol{u} = 0, \tag{4.20}$$

where $h = \eta_0 - \eta_1$. Because $g \gg g'$, (4.18) shows that surface displacements are *much smaller* than the displacements at the interior interface. We see this in the real ocean where the mean interior isopycnal displacements may be several tens of metres but variations in the mean height of the ocean surface are of the order of centimetres.

The mass conservation equation, (4.13b) may be written as

$$-(\zeta+f)\nabla\cdot\boldsymbol{u}=\frac{\zeta+f}{h}\frac{\mathrm{D}h}{\mathrm{D}t}, \tag{4.34}$$

and using this equation and (4.32) we obtain

$$\frac{\mathrm{D}}{\mathrm{D}t}(\zeta+f)=\frac{\zeta+f}{h}\frac{\mathrm{D}h}{\mathrm{D}t}, \tag{4.35}$$

which is equivalent to

$$\boxed{\frac{\mathrm{D}Q}{\mathrm{D}t}=0 \qquad \text{where} \qquad Q=\left(\frac{\zeta+f}{h}\right).} \tag{4.36}$$

The important quantity Q is known as the *potential vorticity*, and (4.36) is the potential vorticity equation.

4.3 Shallow Water Waves

Let us now look at the gravity waves that occur in shallow water. To isolate the essence we consider waves in a single fluid layer, with a flat bottom and a free upper surface, in which gravity provides the sole restoring force.

4.3.1 Non-Rotating Shallow Water Waves

Given a flat bottom the fluid thickness is equal to the free surface displacement (Fig. 4.1), and taking the basic state of the fluid to be at rest we let

$$h(x,y,t)=H+h'(x,y,t)=H+\eta'(x,y,t), \tag{4.37a}$$

$$\boldsymbol{u}(x,y,t)=\boldsymbol{u}'(x,y,t). \tag{4.37b}$$

The mass conservation equation, (4.13b), then becomes

$$\frac{\partial\eta'}{\partial t}+(H+\eta')\nabla\cdot\boldsymbol{u}'+\boldsymbol{u}'\cdot\nabla\eta'=0, \tag{4.38}$$

and neglecting squares of small quantities this yields the linear equation

$$\frac{\partial\eta'}{\partial t}+H\nabla\cdot\boldsymbol{u}'=0. \tag{4.39}$$

Similarly, linearizing the momentum equation, (4.7) with $\boldsymbol{f}=0$, yields

$$\frac{\partial\boldsymbol{u}'}{\partial t}=-g\nabla\eta'. \tag{4.40}$$

Eliminating velocity by differentiating (4.39) with respect to time and taking the divergence of (4.40) leads to

$$\frac{\partial^2\eta'}{\partial t^2}-gH\nabla^2\eta'=0, \tag{4.41}$$

CHAPTER

5

Geostrophic Theory

GEOSTROPHIC AND HYDROSTATIC BALANCE are the two dominant balances in meteorology and oceanography and in this chapter we exploit these balances to derive various simplified sets of equation. The 'problem' with the full equations is that they are *too* complete, and they contain motions that we don't always care about — sound waves and gravity waves for example. If we can eliminate these modes from the outset then our path toward understanding is not littered with obstacles.

Our specific goal is to derive various sets of 'geostrophic equations', in particular the planetary-geostrophic and quasi-geostrophic equations, by making use of the fact that geostrophic and hydrostatic balance are closely satisfied. We do this first for the shallow water equations and then for the stratified, three-dimensional equations. We will use the Boussinesq equations, but a treatment in pressure coordinates would be very similar. The bottom topography, η_B, can be an unneeded complication in the derivations below and readers may wish to simplify by setting $\eta_B = 0$.

5.1 Scaling the Shallow Water Equations

In order to simplify the equations of motion we first *scale* them — we choose the scales we wish to describe, and then determine the approximate sizes of the terms in the equations. We then eliminate the small terms and derive a set of equations that is simpler than the original set but that consistently describes motion of the chosen scale. With the odd exception, we will denote the scales of variables by capital letters; thus, if L is a typical length scale of the motion we wish to describe, and U is a typical velocity scale, then

$$\begin{aligned} (x, y) &\sim L \qquad \text{or} \qquad (x, y) = \mathcal{O}(L), \\ (u, v) &\sim U \qquad \text{or} \qquad (u, v) = \mathcal{O}(U), \end{aligned} \tag{5.1}$$

and similarly for the other variables in the equations.

We then write the equations of motion in a nondimensional form by writing the variables as

$$(x, y) = L(\widehat{x}, \widehat{y}), \qquad (u, v) = U(\widehat{u}, \widehat{v}), \tag{5.2}$$

where the hatted variables are nondimensional and, by supposition, are $\mathcal{O}(1)$. The various terms in the momentum equation then scale as:

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \tag{5.3a}$$

$$\frac{U}{T} \qquad \frac{U^2}{L} \qquad fU \sim g\frac{\mathcal{H}}{L}, \tag{5.3b}$$

where the ∇ operator acts in the x–y plane and $\mathcal{H}$ is the amplitude of the variations in the surface displacement. We choose an 'advective scale' for time, meaning that $T = L/U$ and $t = \hat{t}L/U$, and the time derivative then scales the same way as the advection. The ratio of the advective term to the rotational term in the momentum equation (5.3) is $(U^2/L)/(fU) = U/fL$; this is the Rossby number that we previously encountered.

The geostrophic theory of this chapter applies when the Rossby number is small. On Earth the theory is generally appropriate for large-scale flow in the mid- and high latitude atmosphere and ocean. On other planets the applicability of geostrophic theory depends on how rapidly the planet rotates and how big it is. Venus has a rotation rate some 200 times slower than Earth and the Rossby number of the large-scale circulation is quite large. Jupiter rotates much faster than Earth (a Jupiter day is about 10 hours), and the planet is also much bigger, and the Rossby number remains small even close to the equator.

We are interested in flows for which the Rossby number is small, in which case the Coriolis term is largely balanced by the pressure gradient. From (5.3b), variations in η scale according to

$$\mathcal{H} = \frac{fUL}{g} = Ro\frac{f^2L^2}{g} = RoH\frac{L^2}{L_d^2}, \tag{5.4}$$

where $L_d = \sqrt{gH}/f$ is the deformation radius and H is the mean depth of the fluid. The ratio of variations in fluid height to the total fluid height thus scales as

$$\frac{\mathcal{H}}{H} \sim Ro\frac{L^2}{L_d^2}. \tag{5.5}$$

Now, the thickness of the fluid, h, may be written as the sum of its mean and a deviation, h_D

$$h = H + h_D = H + (\eta_T - \eta_B), \tag{5.6}$$

where, referring to Fig. 4.1, η_B is the height of the bottom topography and η_T is the height of the fluid above its mean value. Given the scalings above, the deviation height of the fluid may be written as

$$\eta_T = Ro\frac{L^2}{L_d^2}H\widehat{\eta}_T \quad \text{and} \quad \eta = H + \eta_T = H\left(1 + Ro\frac{L^2}{L_d^2}\widehat{\eta}_T\right), \tag{5.7}$$

where $\widehat{\eta}_T$ is the $\mathcal{O}(1)$ nondimensional value of the surface height deviation. We apply the same scalings to h itself and, if $h_D = h - H = \eta_T - \eta_B$ is the deviation of the thickness from its mean value, then

$$h = H + h_D = H\left(1 + Ro\frac{L^2}{L_d^2}\widehat{h}_D\right), \tag{5.8}$$

where $\widehat{h}_D$ is the nondimensional deviation thickness of the fluid layer.

Nondimensional momentum equation

If we use (5.7) and (5.8) to scale height variations, (5.2) to scale lengths and velocities, and an advective scaling for time, then, and since $\nabla\widehat{\eta} = \nabla\widehat{\eta}_T$, the momentum equation (5.3) becomes

$$Ro\left[\frac{\partial\widehat{\boldsymbol{u}}}{\partial\widehat{t}} + (\widehat{\boldsymbol{u}}\cdot\nabla)\widehat{\boldsymbol{u}}\right] + \widehat{\boldsymbol{f}}\times\widehat{\boldsymbol{u}} = -\nabla\widehat{\eta}, \tag{5.9}$$

where $\widehat{\boldsymbol{f}} = \hat{\mathbf{k}}\widehat{f} = \hat{\mathbf{k}}f/f_0$, where f_0 is a representative value of the Coriolis parameter. (If f is a constant, then $\widehat{f} = 1$, but it is informative to explicitly write $\widehat{f}$ in the equations. Also, where the operator ∇ operates on a nondimensional variable, the differentials are taken with respect to the nondimensional variables $\widehat{x}$, $\widehat{y}$.) All the variables in (5.9) will now be assumed to be of order unity, and the Rossby number multiplying the local time derivative and the advective terms indicates the smallness of those terms. By construction, the dominant balance in (5.9) is the geostrophic balance between the last two terms.

Nondimensional mass continuity (height) equation

The (dimensional) mass continuity equation can be written as

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla\cdot\boldsymbol{u} = 0 \quad \text{or} \quad \frac{1}{H}\frac{\mathrm{D}h_D}{\mathrm{D}t} + \left(1 + \frac{h_D}{H}\right)\nabla\cdot\boldsymbol{u} = 0, \tag{5.10}$$

since $\mathrm{D}h/\mathrm{D}t = \mathrm{D}h_D/\mathrm{D}t$. Using (5.2) and (5.8) the above equation may be written

$$Ro\left(\frac{L}{L_d}\right)^2\frac{\mathrm{D}\widehat{h}_D}{\mathrm{D}\widehat{t}} + \left[1 + Ro\left(\frac{L}{L_d}\right)^2\widehat{h}_D\right]\nabla\cdot\widehat{\boldsymbol{u}} = 0. \tag{5.11}$$

Equations (5.9) and (5.11) are the nondimensional versions of the full shallow water equations of motion. Since the Rossby number is small we might expect that some terms in this equation can be eliminated with little loss of accuracy, depending on the size of the second nondimensional parameter, $(L/L_d)^2$, as we now explore.

5.2 Geostrophic Shallow Water Equations

5.2.1 Planetary-Geostrophic Equations

We now derive simplified equation sets that are appropriate in particular parameter regimes, beginning with an equation set appropriate for the very largest scales. Specifically, we take

$$Ro \ll 1, \qquad \frac{L}{L_d} \gg 1 \quad \text{such that} \quad Ro\left(\frac{L}{L_d}\right)^2 = \mathcal{O}(1). \tag{5.12}$$

The first inequality implies we are considering flows in geostrophic balance. The second inequality means we are considering flows much larger than the deformation radius. The ratio of the deformation radius to scale of motion of the fluid is called the Burger number; that is, $Bu \equiv L_d/L$, so here we are considering small Burger-number flows.

The smallness of the Rossby number means that we can neglect the material derivative in the momentum equation, (5.9), leaving geostrophic balance. Thus, in dimensional form, the momentum equation may be written, in vectorial or component forms, as

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\eta, \quad \text{or} \quad fv = g\frac{\partial \eta}{\partial x}, \qquad fu = -g\frac{\partial \eta}{\partial y}. \tag{5.13}$$

The *planetary-geostrophic* equations are appropriate for geostrophically balanced flow at very large scales. In the shallow water version, they consist of the full mass conservation equation along with geostrophic balance.

Looking now at the mass continuity equation, (5.11), we see that there are no small terms that can be eliminated. Thus, we have simply the full mass conservation equation,

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\boldsymbol{u}) = 0, \tag{5.14}$$

where h and η are related by $\eta = \eta_B + h$, where η_B is the height of the bottom topography. Equations (5.13) and (5.14) form the *planetary geostrophic shallow water equations*. There is *only one time derivative* in the equations, so there can be no gravity waves. The system is evolved purely through the mass continuity equation, and the flow field is *diagnosed* from the height field.

Planetary-geostrophic potential vorticity

In the (full) shallow water equations potential vorticity is conserved, meaning that

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{\zeta + f}{h}\right) = 0. \tag{5.15}$$

In the planetary-geostrophic equations we can use (5.13) and (5.14) to show that this conservation law becomes

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{f}{h}\right) = 0, \tag{5.16}$$

as might be expected since ζ is smaller than f by a factor of the Rossby number. An alternate derivation of the planetary-geostrophic equations is to go directly from (5.15) to (5.16), by virtue of the smallness of the Rossby number, and then simply use (5.16) instead of (5.14) as the evolution equation.

5.2.2 Quasi-Geostrophic Equations

The *quasi-geostrophic equations* are appropriate for scales of the same order as the deformation radius, and so for

$$Ro \ll 1, \qquad \frac{L}{L_d} = \mathcal{O}(1) \quad \text{so that} \quad Ro\left(\frac{L}{L_d}\right)^2 \ll 1. \tag{5.17}$$

Since the Rossby number is small the momentum equations again reduce to geostrophic balance, namely (5.13). In the mass continuity equation, we now eliminate all terms involving Rossby number to give

$$\nabla \cdot \boldsymbol{u} = 0. \tag{5.18}$$

The quasi-geostrophic equations are appropriate for geostrophically balanced flow at so-called synoptic scales, or weather scales. This scale is mainly determined by the Rossby radius of deformation which is about 1000 km in the atmosphere and 100 km (and less in high latitudes) in the ocean.

Neither geostrophic balance nor (5.18) are prognostic equations, and it appears we have derived an uninteresting, static, set of equations. In fact we haven't gone far enough, since nothing in our derivation says that these quantities do not evolve. To see this, let us suppose that the Coriolis parameter is nearly constant, which is physically consistent with the idea that scales of motion are comparable to the deformation scale. Geostrophic balance with a constant Coriolis parameter gives

$$f_0 u = -g\frac{\partial \eta}{\partial y}, \quad -f_0 v = -g\frac{\partial \eta}{\partial z}, \quad \text{giving} \quad \nabla \cdot \boldsymbol{u} = 0. \tag{5.19}$$

That it to say, the geostrophic flow is divergence-free and we therefore should not suppose that $\nabla \cdot \boldsymbol{u} = 0$ is the dominant term in the height equation.

However, with a little more care we can in fact derive a set of equations that evolves under these conditions, and that furthermore is extraordinarily useful, for it describes the flow on the scales of motion corresponding to weather. We make three explicit assumptions:

(i) The Rossby number is small and the flow is in near geostrophic balance.

(ii) The scales of motion are similar to the deformation scales, so that $L \sim L_d$ and $Ro(L/L_d)^2 \ll 1$.

(iii) Variations of the Coriolis parameter are small, so that $f = f_0 + \beta y$ where $\beta y \ll f_0$.

The velocity is then equal to a geostrophic component, $\boldsymbol{u}_g$ plus an ageostrophic component, $\boldsymbol{u}_a$ where $|\boldsymbol{u}_g| \gg |\boldsymbol{u}_a|$ and the geostrophic velocity satisfies

$$f_0 \times \boldsymbol{u}_g = -g\nabla\eta, \tag{5.20}$$

which, because of the use of a constant Coriolis parameter (assumption *(iii)*), implies $\nabla \cdot \boldsymbol{u}_g = 0$.

We proceed from the shallow water vorticity equation which, as in (4.32), is

$$\frac{\partial \zeta}{\partial t} + (\boldsymbol{u} \cdot \nabla)(\zeta + f) = -(\zeta + f)\nabla \cdot \boldsymbol{u}. \tag{5.21}$$

The thermodynamic equation then becomes

$$\frac{\mathrm{D}b'}{\mathrm{D}t} + N^2 w = 0, \tag{5.33}$$

where $N^2 = \partial\tilde{b}/\partial z$ and the advective derivative is still three-dimensional. We then let $\phi = \tilde{\phi}(z) + \phi'$, where $\tilde{\phi}$ is hydrostatically balanced by $\tilde{b}$, and the hydrostatic equation becomes

$$\frac{\partial \phi'}{\partial z} = b'. \tag{5.34}$$

Equations (5.33) and (5.34) replace (5.31c) and (5.31b), and ϕ' is used in (5.31a).

5.3.1 Scaling the Equations

We scale the basic variables by supposing that

$$(x, y) \sim L, \quad (u, v) \sim U, \quad t \sim \frac{L}{U}, \quad z \sim H, \quad f \sim f_0, \quad N \sim N_0, \tag{5.35}$$

where the scaling variables (capitalized, except for f_0) are chosen to be such that the nondimensional variables have magnitudes of the order of unity, and the constant N_0 is a representative value of the stratification. We presume that the scales chosen are such that the Rossby number is small; that is $Ro = U/(f_0 L) \ll 1$. In the momentum equation the pressure term then balances the Coriolis force,

$$|\boldsymbol{f} \times \boldsymbol{u}| \sim |\nabla \phi'|, \tag{5.36}$$

and so the pressure scales as

$$\phi' \sim \Phi = f_o U L. \tag{5.37}$$

Using the hydrostatic relation, (5.37) implies that the buoyancy scales as

$$b' \sim B = \frac{f_0 U L}{H}, \tag{5.38}$$

and from this we obtain

$$\frac{(\partial b'/\partial z)}{N^2} \sim Ro \frac{L^2}{L_d^2}, \tag{5.39}$$

where

$$L_d = \frac{N_0 H}{f_0} \tag{5.40}$$

is the deformation radius in the continuously-stratified fluid, analogous to the quantity $\sqrt{gH}/f_0$ in the shallow water system, and we use the same symbol for both. In the continuously-stratified system, *if the scale of motion is the same as or smaller than the deformation radius, and the Rossby number is small, then the variations in stratification are small.* The choice of scale is the key difference between the planetary-geostrophic and quasi-geostrophic equations.

Finally, at least for now, we nondimensionalize the vertical velocity by using the mass conservation equation,

$$\frac{\partial w}{\partial z} = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \tag{5.41}$$

with the scaling

$$w \sim W = \frac{UH}{L}. \tag{5.42}$$

This scaling will not necessarily be correct if the flow is geostrophically balanced. In this case we can then estimate w by cross-differentiating geostrophic balance (with $\tilde{\rho}$ constant for simplicity) to obtain the linear geostrophic vorticity equation and corresponding scaling:

$$\beta v \approx f\frac{\partial w}{\partial z}, \qquad w \sim W = \frac{\beta UH}{f_0}. \tag{5.43a,b}$$

If variations in the Coriolis parameter are large and $\beta \sim f_0/L$, then (5.43b) is the same as (5.42), but if f is nearly constant then $W \ll UH/L$.

Given the scalings above (using (5.42) for w) we nondimensionalize by setting

$$\begin{gathered}(\hat{x}, \hat{y}) = L^{-1}(x, y), \qquad \hat{z} = H^{-1}z, \qquad (\hat{u}, \hat{v}) = U^{-1}(u, v), \qquad \hat{t} = \frac{U}{L}t, \\ \hat{w} = \frac{L}{UH}w, \qquad \hat{f} = f_0^{-1}f, \qquad \hat{\phi} = \frac{\phi'}{f_0UL}, \qquad \hat{b} = \frac{H}{f_0UL}b',\end{gathered} \tag{5.44}$$

where the hatted variables are nondimensional. The horizontal momentum and hydrostatic equations then become

$$Ro\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} + \hat{\boldsymbol{f}} \times \hat{\boldsymbol{u}} = -\nabla\hat{\phi}, \tag{5.45}$$

and

$$\frac{\partial\hat{\phi}}{\partial\hat{z}} = \hat{b}. \tag{5.46}$$

The nondimensional mass conservation equation is simply

$$\nabla \cdot \hat{\boldsymbol{v}} = \left(\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}} + \frac{\partial\hat{w}}{\partial\hat{z}}\right) = 0, \tag{5.47}$$

and the nondimensional thermodynamic equation is

$$\frac{f_0UL}{H}\frac{U}{L}\frac{\mathrm{D}\hat{b}}{\mathrm{D}\hat{t}} + \widehat{N}^2N_0^2\frac{HU}{L}\hat{w} = 0, \tag{5.48}$$

or, re-arranging,

$$Ro\frac{\mathrm{D}\hat{b}}{\mathrm{D}\hat{t}} + \left(\frac{L_d}{L}\right)^2\widehat{N}^2\hat{w} = 0. \tag{5.49}$$

The nondimensional equations are summarized in the box on the following page.

Chapter

6 Rossby Waves

WAVES ARE FAMILIAR TO ALMOST EVERYONE. Gravity waves cover the ocean surface, sound waves allow us to talk and light waves enable us to see. This chapter provides an introduction to their properties, paying particular attention to a wave that is especially important to the large scale flow in both ocean and atmosphere — the Rossby wave. We start with an elementary introduction to wave kinematics, discussing such concepts as phase speed and group velocity. Then, beginning with Section 6.3, we discuss the dynamics of Rossby waves, and this part may be considered to be the natural follow-on from the geostrophic theory of the previous chapter. Rossby waves then reappear frequently in later chapters.

6.1 Fundamentals and Formalities

6.1.1 Definitions and Kinematics

A wave is more easily recognized than defined. Loosely speaking, a wave is a propagating disturbance that has a characteristic relationship between its frequency and size, called a *dispersion relation.* To see what all this means, and what a dispersion relation is, suppose that a disturbance, $\psi(\boldsymbol{x}, t)$ (where ψ might be velocity, streamfunction, pressure, etc.), satisfies the equation

$$L(\psi) = 0, \tag{6.1}$$

where L is a linear operator, typically a polynomial in time and space derivatives; one example is $L(\psi) = \partial\nabla^2\psi/\partial t + \beta\partial\psi/\partial x$. If (6.1) has constant coefficients (if β is constant in this example) then harmonic solutions may often be found that are a superposition of *plane waves,* each of which satisfy

$$\psi = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}\theta(\boldsymbol{x},t)} = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{6.2}$$

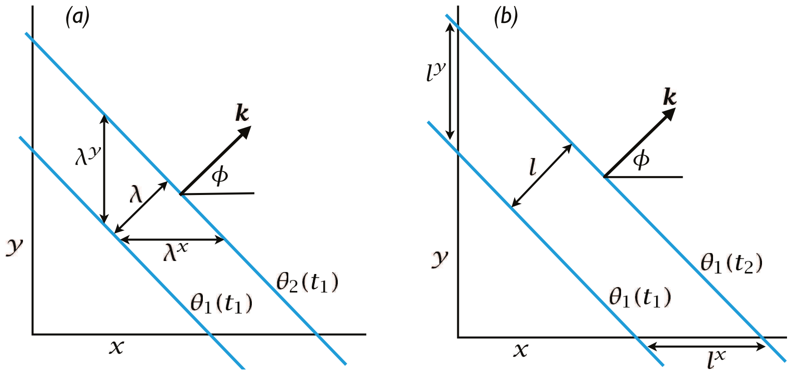

Fig. 6.1: The propagation of a two-dimensional wave. (a) Two lines of constant phase (e.g., two wavecrests) at a time t_1. The wave is propagating in the direction $\boldsymbol{k}$ with wavelength λ. (b) A line of constant phase at two successive times. The phase speed is the speed of advancement of the wavecrest in the direction of travel, and so $c_p = l/(t_2 - t_1)$. The phase speed in the x-direction is the speed of propagation of the wavecrest along the x-axis, and $c_p^x = l^x/(t_2 - t_1) = c_p/\cos\phi$.

where $\widetilde{\psi}$ is a complex constant, θ is the phase, ω is the wave frequency and $\boldsymbol{k}$ is the vector wavenumber (k, l, m) (also written as (k^x, k^y, k^z) or, in subscript notation, k_i). The prefix Re denotes the real part of the expression, but we will drop it if there is no ambiguity.

Waves are characterized by having a particular relationship between the frequency and wavevector known as the *dispersion relation*. This is an equation of the form

$$\omega = \Omega(\boldsymbol{k}), \tag{6.3}$$

where $\Omega(\boldsymbol{k})$, or $\Omega(k_i)$, and meaning $\Omega(k, l, m)$, is some function determined by the form of L in (6.1) and which thus depends on the particular type of wave — the function is different for sound waves, light waves and the Rossby waves and gravity waves we will encounter in this book. Unless it is necessary to explicitly distinguish the function Ω from the frequency ω, we often write $\omega = \omega(\boldsymbol{k})$.

6.1.2 Wave Propagation and Phase Speed

A common property of waves is that they propagate through space with some velocity, which in special cases might be zero. Waves in fluids may carry energy and momentum but do not necessarily transport fluid parcels themselves. Further, it turns out that the speed at which properties like energy are transported (the group speed) may be different from the speed at which the wave crests themselves move (the phase speed). Let's try to understand this statement, beginning with the phase speed. A summary of key results is given on page 107.

Phase speed

Consider the propagation of monochromatic plane waves, for that is all that is needed to introduce the phase speed. Given (6.2) a wave will propagate in the direction of $\boldsymbol{k}$ (Fig. 6.1). At a given instant and location we can align our coordinate axis along this direction, and we write $\boldsymbol{k} \cdot \boldsymbol{x} = Kx^*$, where x^* increases in the direction of $\boldsymbol{k}$ and $K^2 = |\boldsymbol{k}|^2$ is the magnitude of the wavenumber. With this, we can write (6.2) as

$$\psi = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(Kx^* - \omega t)} = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}K(x^* - ct)}, \tag{6.4}$$

where $c = \omega/K$. From this equation it is evident that the phase of the wave propagates at the speed c in the direction of $\boldsymbol{k}$, and we define the *phase speed* by

$$c_p \equiv \frac{\omega}{K}. \tag{6.5}$$

The wavelength of the wave, λ, is the distance between two wavecrests — that is, the distance between two locations along the line of travel whose phase differs by 2π — and evidently this is given by

$$\lambda = \frac{2\pi}{K}. \tag{6.6}$$

In (for simplicity) a two-dimensional wave, and referring to Fig. 6.1, the wavelength and wave vectors in the x- and y-directions are given by,

$$\lambda^x = \frac{\lambda}{\cos\phi}, \quad \lambda^y = \frac{\lambda}{\sin\phi}, \qquad k^x = K\cos\phi, \quad k^y = K\sin\phi. \tag{6.7}$$

In general, lines of constant phase intersect both the coordinate axes and propagate along them. The speed of propagation along these axes is given by

$$c_p^x = c_p\frac{l^x}{l} = \frac{c_p}{\cos\phi} = c_p\frac{K}{k^x} = \frac{\omega}{k^x}, \qquad c_p^y = c_p\frac{l^y}{l} = \frac{c_p}{\sin\phi} = c_p\frac{K}{k^y} = \frac{\omega}{k^y}, \tag{6.8}$$

using (6.5) and (6.7), and again referring to Fig. 6.1 for notation. The speed of phase propagation along any one of the axes is in general *larger* than the phase speed in the primary direction of the wave. The phase speeds are clearly *not* components of a vector: for example, $c_p^x \neq c_p\cos\phi$. Analogously, the wavevector $\boldsymbol{k}$ is a true vector, whereas the wavelength λ is not.

To summarize, the phase speed and its components are given by

$$\boxed{c_p = \frac{\omega}{K}, \qquad c_p^x = \frac{\omega}{k^x}, \quad c_p^y = \frac{\omega}{k^y}.} \tag{6.9}$$

6.1.3 The Dispersion Relation

The above description is kinematic, in that it applies to almost any disturbance that has a wavevector and a frequency. The particular *dynamics* of a wave are determined by the relationship between the wavevector and the frequency; that is, by the *dispersion relation*. Once the dispersion relation is known a great many of the properties of the wave follow in a more-or-less straightforward manner. Picking up from (6.3), the dispersion relation is a functional relationship between the frequency and the wavevector of the general form

$$\omega = \Omega(\boldsymbol{k}). \tag{6.10}$$

Perhaps the simplest example of a linear operator that gives rise to waves is the one-dimensional equation

$$\frac{\partial\psi}{\partial t} + c\frac{\partial\psi}{\partial x} = 0. \tag{6.11}$$

Wave Fundamentals

- A wave is a propagating disturbance that has a characteristic relationship between its frequency and size, known as the dispersion relation. Waves typically arise as solutions to a linear problem of the form $L(\psi) = 0$, where L is, commonly, a linear operator in space and time. Two examples are

$$\frac{\partial^2 \psi}{\partial t^2} - c^2 \nabla^2 \psi = 0 \qquad \text{and} \qquad \frac{\partial}{\partial t}\nabla^2 \psi + \beta \frac{\partial \psi}{\partial x} = 0, \tag{WF.1}$$

  where the second example gives rise to Rossby waves.
- Solutions to the governing equation are often sought in the form of plane waves that have the form

$$\psi = \mathrm{Re}\, A \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{WF.2}$$

  where A is the wave amplitude, $\boldsymbol{k} = (k, l, m)$ is the wavevector, and ω is the frequency.
- The dispersion relation connects the frequency and wavevector through an equation of the form $\omega = \Omega(\boldsymbol{k})$ where Ω is some function. The relation is normally derived by substituting a trial solution like (WF.2) into the governing equation. For the examples of (WF.1) we obtain $\omega = c^2 K^2$ and $\omega = -\beta k / K^2$ where $K^2 = k^2 + l^2 + m^2$ or, in two dimensions, $K^2 = k^2 + l^2$.
- The phase speed is the speed at which the wave crests move. In the direction of propagation and in the x, y and z directions the phase speeds are given by, respectively,

$$c_p = \frac{\omega}{K}, \qquad c_p^x = \frac{\omega}{k}, \quad c_p^y = \frac{\omega}{l}, \quad c_p^z = \frac{\omega}{m}, \tag{WF.3}$$

  where $K = 2\pi/\lambda$ and λ is the wavelength. The wave crests have both a speed (c_p) and a direction of propagation (the direction of $\boldsymbol{k}$), like a vector, but the components defined in (WF.3) are not the components of that vector.
- The group velocity is the velocity at which a wave packet or wave group moves. It is a vector and is given by

$$\boldsymbol{c}_g = \frac{\partial \omega}{\partial \boldsymbol{k}} \quad \text{with components} \quad c_g^x = \frac{\partial \omega}{\partial k}, \quad c_g^y = \frac{\partial \omega}{\partial l}, \quad c_g^z = \frac{\partial \omega}{\partial m}. \tag{WF.4}$$

  Most physical quantities of interest are transported at the group velocity.

Substituting a trial solution of the form $\psi = \mathrm{Re}\, A\mathrm{e}^{\mathrm{i}(kx-\omega t)}$ into (6.11) we obtain $(-\mathrm{i}\omega + c\mathrm{i}k)A = 0$, giving the dispersion relation

$$\omega = ck. \tag{6.12}$$

The phase speed of this wave is $c_p = \omega/k = c$. A couple of other examples of governing equations, dispersion relations and phase speeds are:

$$\frac{\partial^2 \psi}{\partial t^2} - c^2 \nabla^2 \psi = 0, \quad \omega^2 = c^2 K^2, \quad c_p = \pm c, \quad c_p^x = \pm \frac{cK}{k}, \quad c_p^y = \pm \frac{cK}{l}, \tag{6.13a}$$

$$\frac{\partial}{\partial t}\nabla^2 \psi + \beta \frac{\partial \psi}{\partial x} = 0, \quad \omega = \frac{-\beta k}{K^2}, \quad c_p = \frac{\omega}{K}, \quad c_p^x = -\frac{\beta}{K^2}, \quad c_p^y = -\frac{\beta k/l}{K^2}, \tag{6.13b}$$

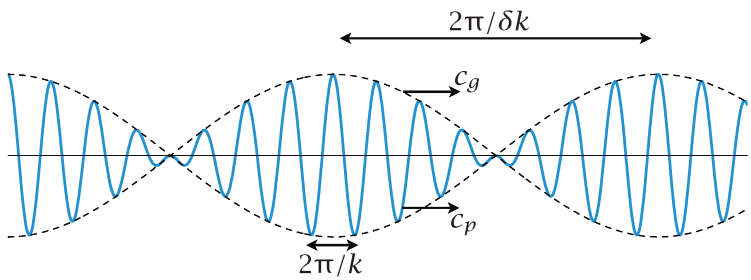

Fig. 6.2: Superposition of two sinusoidal waves with wavenumbers k and $k + \delta k$, producing a wave (solid line) that is modulated by a slowly varying wave envelope or packet (dashed). The envelope moves at the group velocity, $c_g = \partial\omega/\partial k$, and the phase moves at the group speed, $c_p = \omega/k$.

where $K^2 = k^2 + l^2$ and the examples are two-dimensional, with variation in x and y only.

A wave is said to be *nondispersive* if the phase speed is independent of the wavelength. This condition is satisfied for the simple example (6.11) but is manifestly not satisfied for (6.13b), and these waves (Rossby waves, in fact) are *dispersive*. Waves of different wavelengths then travel at different speeds so that a group of waves will spread out — disperse — even if the medium is homogeneous. When a wave is dispersive there is another characteristic speed at which the waves propagate, the group velocity, and we come to this shortly.

Most media are inhomogeneous, but if the medium varies sufficiently slowly in space and time — and in particular if the variations are slow compared to the wavelength and period — we may still have a *local* dispersion relation between frequency and wavevector,

$$\omega = \Omega(\boldsymbol{k}; \boldsymbol{x}, t), \tag{6.14}$$

where $\boldsymbol{x}$ and t are slowly varying parameters. We resume our discussion of this topic in Section 6.5, but before that we introduce the group velocity.

6.2 Group Velocity

Group velocity seems to have been first articulated in about 1841 by the Irish mathematician and physicist William Rowan Hamilton (1806–1865), who is also remembered for his formulation of 'Hamiltonian mechanics.' Hamilton was largely motivated by optics, and it was George Stokes, Osborne Reynolds and John Strutt (also known as Lord Rayleigh) who further developed and generalized the idea in fluid dynamics in the nineteenth and early twentieth centuries.

Information and energy do not, in general, propagate at the phase speed. Rather, most quantities of interest propagate at the *group velocity*, a quantity of enormous importance in wave theory. Roughly speaking, group velocity is the velocity at which a packet or a group of waves will travel, whereas the individual wave crests travel at the phase speed. To introduce the idea we will consider the superposition of plane waves, noting that a truly monochromatic plane wave already fills all space uniformly so that there can be no propagation of energy from place to place.

6.2.1 Superposition of Two Waves

Consider the linear superposition of two waves. Limiting attention to the one-dimensional case, consider a disturbance that is the sum of two waves,

$$\psi = \mathrm{Re}\, \widetilde{\psi}(\mathrm{e}^{\mathrm{i}(k_1 x - \omega_1 t)} + \mathrm{e}^{\mathrm{i}(k_2 x - \omega_2 t)}). \tag{6.15}$$

We linearize about a constant zonal flow, U, by writing

$$\frac{\partial}{\partial t}\nabla^2\psi' + U\frac{\partial\nabla^2\psi'}{\partial x} + \beta\frac{\partial\psi'}{\partial x} = 0. \tag{6.24}$$

This equation is just a single-layer version of (6.22), with $\partial\overline{q}/\partial y = \beta$, $q' = \nabla^2\psi'$ and $v' = \partial\psi'/\partial x$.

The coefficients in (6.24) are not functions of y or z; this is not a requirement for wave motion to exist but it does enable solutions to be found more easily. Let us seek solutions in the form of a plane wave, namely

$$\psi' = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \tag{6.25}$$

where $\widetilde{\psi}$ is a complex constant. Solutions of this form are valid in a domain with doubly-periodic boundary conditions; solutions in a channel can be obtained using a meridional variation of $\sin ly$, with no essential changes to the dynamics. The amplitude of the oscillation is given by $\widetilde{\psi}$ and the phase by $kx+ly-\omega t$, where k and l are the x- and y-wavenumbers and ω is the frequency of the oscillation.

Substituting (6.25) into (6.24) yields

$$[(-\omega + Uk)(-K^2) + \beta k]\widetilde{\psi} = 0, \tag{6.26}$$

where $K^2 = k^2 + l^2$. For non-trivial solutions the above equation implies

$$\boxed{\omega = Uk - \frac{\beta k}{K^2}\,,} \tag{6.27}$$

and this is the *dispersion relation* for barotropic Rossby waves. Evidently the velocity U Doppler shifts the frequency by the amount Uk. The components of the phase speed and group velocity are given by, respectively,

$$c_p^x \equiv \frac{\omega}{k} = U - \frac{\beta}{K^2}, \qquad c_p^y \equiv \frac{\omega}{l} = U\frac{k}{l} - \frac{\beta k}{K^2 l}, \tag{6.28a,b}$$

and

$$c_g^x \equiv \frac{\partial\omega}{\partial k} = U + \frac{\beta(k^2-l^2)}{(k^2+l^2)^2}, \qquad c_g^y \equiv \frac{\partial\omega}{\partial l} = \frac{2\beta kl}{(k^2+l^2)^2}. \tag{6.29a,b}$$

The phase speed in the absence of a mean flow is *westward,* with waves of longer wavelengths travelling more quickly, and the eastward current speed required to hold the waves of a particular wavenumber stationary (i.e., $c_p^x = 0$) is $U = \beta/K^2$. The background flow U evidently just provides a uniform shift to the phase speed, and (in this case) can be transformed away by a change of coordinate. The x-component of the group velocity may also be written as the sum of the phase speed plus a positive quantity, namely

$$c_g^x = c_p^x + \frac{2\beta k^2}{(k^2+l^2)^2}. \tag{6.30}$$

This means that the zonal group velocity for Rossby wave packets moves eastward relative to its zonal phase speed. A stationary wave ($c_p^x = 0$) has

Chapter

7

Gravity Waves

Waves arise when a system is perturbed and a restoring force tries to bring the system back to equilibrium; the system then overshoots and oscillations ensue. Gravity waves are waves in a fluid in which gravity provides the restoring force. For gravity to have an effect the fluid density must vary, and thus the waves must either exist at a fluid interface or in a stratified fluid — and a fluid interface is just an abrupt form of stratification. It is thus common to think of gravity waves as being either internal waves or surface waves: the former being in the interior of a fluid where the density changes may be continuous and the latter at a fluid interface, and naturally enough the two waves have many similarities. We considered surface waves in the hydrostatic, shallow water case in Chapter 4; now we consider internal waves in the continuously-stratified equations.

Gravity waves are those waves that exist in a fluid for which gravity provides the restoring force. Gravitational waves are a disturbance in the fabric of spacetime caused by accelerating massive bodies, as predicted by the general theory of relativity.

7.1 Internal Waves in a Continuously-Stratified Fluid

Internal gravity waves are waves that are internal to a stratified fluid and that owe their existence to the restoring force of gravity. In this section we will consider the simplest and most fundamental case, that of internal waves in a Boussinesq fluid with constant stratification and no background rotation. To this end, consider a fluid, initially at rest, in which the background buoyancy varies only with height and so the buoyancy frequency, N, is a function only of z. The system satisfies the Boussinesq equations (Section 2.5) and linearizing those equations of motion about this basic state gives the linear momentum equations,

$$\frac{\partial \boldsymbol{u}'}{\partial t} = -\nabla\phi', \qquad \frac{\partial w'}{\partial t} = -\frac{\partial \phi'}{\partial z} + b', \tag{7.1a,b}$$

and the mass continuity and thermodynamic equations,

$$\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} = 0, \qquad \frac{\partial b'}{\partial t} + w' N^2 = 0. \tag{7.1c,d}$$

Our notation is such that $\boldsymbol{u} \equiv u\hat{\mathbf{i}} + v\hat{\mathbf{j}}$, $\boldsymbol{v} \equiv u\hat{\mathbf{i}} + v\hat{\mathbf{j}} + w\hat{\mathbf{k}}$, where $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ are the unit vectors in the x, y and z directions, and the gradient operator is horizontal unless noted. Thus, $\nabla \equiv \hat{\mathbf{i}}\,\partial_x + \hat{\mathbf{j}}\,\partial_y$ and $\nabla_3 \equiv \hat{\mathbf{i}}\,\partial_x + \hat{\mathbf{j}}\,\partial_y + \hat{\mathbf{k}}\,\partial_z$.

A little algebra gives a single equation for w',

$$\left[\frac{\partial^2}{\partial t^2}\left(\nabla^2 + \frac{\partial^2}{\partial z^2}\right) + N^2\nabla^2\right] w' = 0. \tag{7.2}$$

This equation is evidently *not* isotropic. If N^2 is a constant — that is, if the background buoyancy varies linearly with z — then the coefficients of each term are constant, and we may then seek solutions of the form

$$w' = \mathrm{Re}\ \widetilde{w} \mathrm{e}^{\mathrm{i}(kx+ly+mz-\omega t)}, \tag{7.3}$$

where Re denotes the real part, a denotation that will frequently be dropped unless ambiguity arises, and other variables oscillate in a similar fashion. Using (7.3) in (7.2) yields the dispersion relation:

$$\omega^2 = \frac{(k^2 + l^2)N^2}{k^2 + l^2 + m^2} = \frac{K^2 N^2}{K_3^2}, \tag{7.4}$$

where $K^2 = k^2 + l^2$ and $K_3^2 = k^2 + l^2 + m^2$. The frequency (see Fig. 7.1) is thus always less than N, approaching N for small horizontal scales, $K^2 \gg m^2$. If we neglect pressure perturbations, as in the parcel argument of Section 3.4, then the two equations,

$$\frac{\partial w'}{\partial t} = b', \qquad \frac{\partial b'}{\partial t} + w' N^2 = 0, \tag{7.5}$$

form a closed set, and give $\omega^2 = N^2$.

If the basic state density increases with height then $N^2 < 0$ and the basic state is unstable. The disturbance grows exponentially according to $\exp(\sigma t)$ where

$$\sigma = \mathrm{i}\omega = \pm\frac{K\widetilde{N}}{K_3}, \tag{7.6}$$

where $\widetilde{N}^2 \equiv -N^2$ and $K_3 = \sqrt{K_3^2}$. Most convective activity in the ocean and atmosphere is, ultimately, related to an instability of this form, although of course there are many complicating issues — water vapour in the atmosphere, salt in the ocean, the effects of rotation and so forth.

7.1.1 Hydrostatic Internal Waves

Let us now suppose that the fluid satisfies the hydrostatic Boussinesq equations. The linearized two-dimensional equations of motion become

$$\frac{\partial \boldsymbol{u}'}{\partial t} = -\nabla\phi', \qquad 0 = -\frac{\partial \phi'}{\partial z} + b', \tag{7.7a}$$

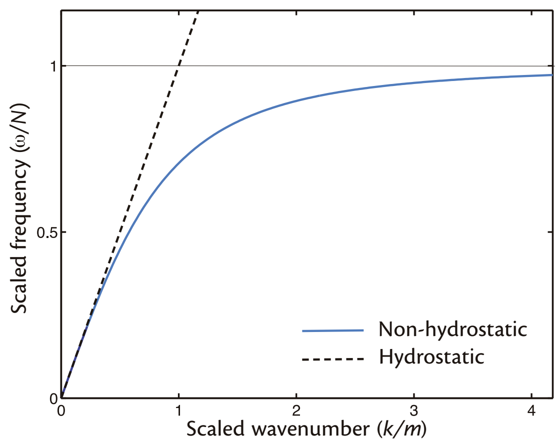

Fig. 7.1: Scaled frequency, ω/N, plotted as a function of scaled horizontal wavenumber, k/m, using the full dispersion relation of (7.4) with $l = 0$ (solid line, asymptoting to unit value for large k/m), and with the hydrostatic dispersion relation (7.8) (dashed line, tending to ∞ for large k/m).

$$\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} = 0, \qquad \frac{\partial b'}{\partial t} + w'N^2 = 0, \tag{7.7b}$$

where these are the horizontal and vertical momentum equations, the mass continuity equation and the thermodynamic equation respectively. A little algebra gives the dispersion relation,

$$\omega^2 = \frac{(k^2 + l^2)N^2}{m^2}. \tag{7.8}$$

The frequency and, if N^2 is negative, the growth rate, are unbounded as $K^2/m^2 \to \infty$, and the hydrostatic approximation thus has quite unphysical behaviour for small horizontal scales. Many numerical models of the large-scale circulation in the atmosphere and ocean do make the hydrostatic approximation. In these models convection must be *parameterized;* otherwise, it would simply occur at the smallest scale available, namely the size of the numerical grid, and this type of unphysical behaviour should be avoided. In nonhydrostatic models convection must also be parameterized if the horizontal resolution of the model is too coarse to properly resolve the convective scales.

7.2 Properties of Internal Waves

Internal waves have a number of interesting and counter-intuitive properties — let's discuss them.

7.2.1 The Dispersion Relation

We can write the dispersion relation, (7.4), as

$$\omega = \pm N \cos\vartheta, \tag{7.9}$$

where $\cos^2\vartheta = K^2/(K^2+m^2)$ so that ϑ is the angle between the three-dimensional wave-vector, $\boldsymbol{k} = k\hat{\mathbf{i}} + l\hat{\mathbf{j}} + m\hat{\mathbf{k}}$, and the horizontal. The frequency is evidently a function only of N and ϑ, and, if this is given, the frequency is not a function of wavelength. This has some interesting consequences for wave reflection, as we see below.

We can also write the dispersion relation, (7.4), as

$$\frac{\omega^2}{N^2-\omega^2} = \frac{K^2}{m^2}. \tag{7.10}$$

Thus, and consistently with our first point, given the wave frequency the ratio of the vertical to the horizontal wavenumber is fixed.

7.2.2 Polarization Relations

The oscillations of pressure, velocity and buoyancy are, naturally, connected, and we can obtain the relations between them with some simple manipulations. If the pressure field is oscillating like $\phi' = \widetilde{\phi}\exp[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)] = \widetilde{\phi}\exp[\mathrm{i}(kx+ly+mz-\omega t)]$ then, using (7.1a), the horizontal velocity components satisfy

$$(\widetilde{u},\widetilde{v}) = (k,l)\,\omega^{-1}\widetilde{\phi}. \tag{7.11}$$

Evidently, since the frequency is real, the velocities are in phase with the pressure. We can obtain similar relations for the other variables and, since all the fields are real, it is convenient to express the relations in terms of sines and cosines. If we choose pressure to vary as a cosine then after some algebra we obtain

$$\phi = \Phi_0\cos(kx+ly+mz-\omega t), \tag{7.12a}$$

$$(u,v) = (k,l)\frac{\Phi_0}{\omega}\cos(kx+ly+mz-\omega t), \tag{7.12b}$$

$$w = \frac{-K^2}{m\omega}\Phi_0\cos(kx+ly+mz-\omega t), \tag{7.12c}$$

$$b = \frac{N^2K^2}{m\omega^2}\Phi_0\sin(kx+ly+mz-\omega t), \tag{7.12d}$$

where Φ_0 is a constant. The vertical velocity is thus in phase with the pressure perturbation, and for regions of positive m (and so with upward phase propagation) regions of high relative pressure are associated with downward fluid motion. The above relations between pressure, buoyancy and velocity are known as *polarization relations.*

7.2.3 Relation between Wave Vector and Velocity

On multiplying (7.12b) and (7.12c) by (k,l) and m, respectively, we see that

$$\boldsymbol{k}\cdot\widetilde{\boldsymbol{v}} = 0, \tag{7.13}$$

where $\boldsymbol{k}$ and $\widetilde{\boldsymbol{v}}$ are three-dimensional vectors. This means that, at any instant, the wave vector is perpendicular to the velocity vector, and the velocity is therefore aligned *along* the direction of the troughs and crests,

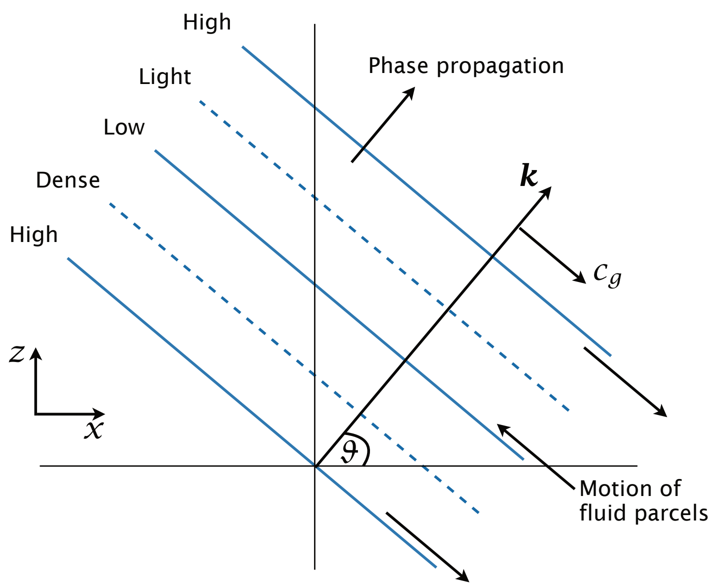

Fig. 7.2: An internal wave propagating in the direction $\boldsymbol{k}$. Both k and m are positive for the wave shown. The solid lines show crests and troughs of constant pressure, and the dashed lines the corresponding crests and troughs of buoyancy (or density). The motion of the fluid parcels is along the lines of constant phase, as shown, and is parallel to the group velocity and perpendicular to the phase speed.

along which there is no pressure gradient. If the wave vector is purely horizontal (i.e., $m = 0$), then the motion is purely vertical and $\omega = N$.

The vertical and horizontal velocities are related to the wavenumbers. If (for simplicity, and with no loss of generality) the motion is in the x–y plane with $v = l = 0$, then it is a corollary of (7.13) that

$$\frac{\tilde{u}}{\tilde{w}} = -\frac{m}{k}. \tag{7.14}$$

Furthermore, from (7.3) with $l = 0$, at any given instant all of the perturbation quantities in the wave are constant along the lines $kx+mz$ = constant. Thus, *all fluid parcel motions are parallel to the wave fronts.* Now, since the wave frequency is related to the background buoyancy frequency by $\omega = \pm N \cos\vartheta$, it follows that the fluid parcels oscillate along lines that are at an angle $\vartheta = \cos^{-1}(\omega/N)$ to the vertical. The polarization relations and the group and phase velocities are illustrated in Fig. 7.2. Let us now discuss the wave properties in a little more detail.

7.2.4 A Parcel Argument and Physical Interpretation

Let us consider first the dispersion relation itself and try to derive it more physically, or at least heuristically. Let us suppose there is a wave propagating in the (x, z) plane at some angle ϑ to the horizontal, with fluid parcels moving parallel to the troughs and crests, as in Fig. 7.2. In general the restoring force on a parcel is due to both the pressure gradient and gravity, but along the crests there is no pressure gradient. Referring to Fig. 7.3, for a total displacement Δs the restoring force, F_{res}, in the direc-

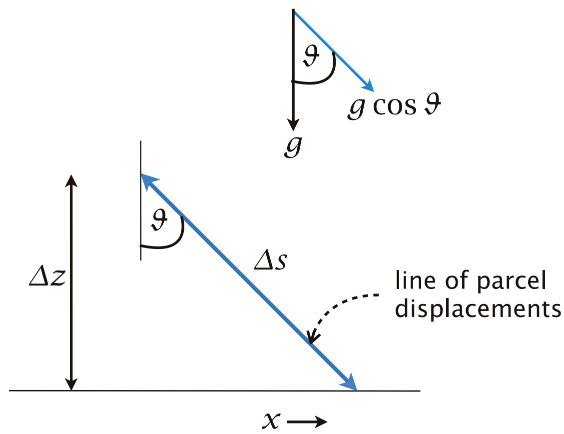

Fig. 7.3: Parcel displacements and associated forces in an internal gravity wave in which the parcel displacements are occurring at an angle ϑ to the vertical, as in Fig. 7.2.

tion of the particle displacement is

$$F_{\text{res}} = g\cos\vartheta\,\Delta\rho = g\cos\vartheta\,\frac{\partial\rho}{\partial z}\Delta z = g\cos\vartheta\,\frac{\partial\rho}{\partial z}\Delta s\cos\vartheta = \rho_0\frac{\partial b}{\partial z}\cos^2\vartheta\,\Delta s, \tag{7.15}$$

noting that $\Delta z = \cos\vartheta\,\Delta s$. The equation of motion of a parcel moving along a trough or crest is therefore

$$\rho_0\frac{\mathrm{d}^2\Delta s}{\mathrm{d}t^2} = -\rho_0 N^2\cos^2\vartheta\,\Delta s, \tag{7.16}$$

which implies a frequency $\omega = N\cos\vartheta$, as in (7.9). One of the $\cos\vartheta$ factors in (7.16) comes from the fact that the parcel displacement is at an angle to the direction of gravity, and the other comes from the fact that the restoring force that a parcel experiences is proportional to $N\cos\vartheta$. (The reader may also wish to refer ahead to Fig. 7.6 and Section 7.3.1 for a similar argument.)

Now consider the wave illustrated in Fig. 7.2. For this wave both k and m are positive, and the frequency is assumed positive by convention to avoid duplicative solutions. The slanting solid and dashed lines are lines of constant phase, and from (7.12) the buoyancy and pressure are 1/4 of a wavelength out of phase. When k and m are both positive the extrema in the buoyancy field lag the extrema in the vertical velocity by $\pi/2$, as illustrated. The perturbation velocities are zero along the lines of extreme buoyancy. This follows because the velocities are in phase with the pressure, which as we noted is out of phase with the buoyancy.

Given the direction of the fluid parcel displacement in Fig. 7.2, the direction of the phase propagation c_p up and to the right may be deduced from the following argument. Buoyancy perturbations arise because of vertical advection of the background stratification, $w'\partial b_0/\partial z = w'N^2$. A local maximum in rising motion, and therefore a tendency to increase the fluid density, is present along the 'Low' line 1/4 wavelength upward and to the right of the 'Dense' phase line. Thus, the density of fluid along the 'Low' phase line increases and the 'Dense' phase line moves upward and to

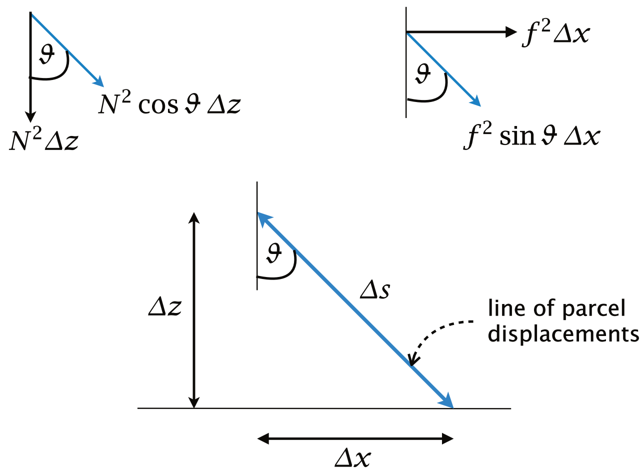

Fig. 7.6: Parcel displacements and associated forces in an inertia-gravity wave in which the parcel displacements are occurring at an angle ϑ to the vertical. Coriolis and buoyancy forces are present, and $\Delta s = \Delta z/\cos\vartheta = \Delta x/\sin\vartheta$.

7.3 ♦ Internal Waves in a Rotating Frame of Reference

In the presence of both a Coriolis force and stratification a displaced fluid will feel two restoring forces — one due to gravity and the other to rotation. The first gives rise to gravity waves, as we have discussed, and the second to inertial waves. When the two forces both occur the resulting waves are called *inertia-gravity waves.* The algebra describing them can be complicated so we begin with a simple parcel argument to lay bare the basic dynamics; refer to Section 7.2.4 as needed.

7.3.1 A Parcel Argument

Consider a parcel that is displaced along a slantwise path in the x–z plane, as shown in Fig. 7.6, with a horizontal displacement of Δx and a vertical displacement of Δz. Let us suppose that the fluid is Boussinesq and that there is a stable and uniform stratification given by $N^2 = -g\rho_0^{-1}\partial\rho_0/\partial z = \partial b/\partial z$. Referring to (7.15) as needed, the component of the restoring buoyancy force, F_b say, in the direction of the parcel oscillation is given by (7.15),

$$F_b = -N^2\cos\vartheta\,\Delta z = -N^2\cos^2\vartheta\,\Delta s. \tag{7.28}$$

The parcel also experiences a restoring Coriolis force, F_C, and the component of this in the direction of the parcel displacement is

$$F_C = -f^2\sin\vartheta\,\Delta x = -f^2\sin^2\vartheta\,\Delta s. \tag{7.29}$$

Here, and for the rest of the chapter, we denote the Coriolis parameter by f. It should be regarded as a constant in any given problem (so there are no Rossby waves), but its value varies with latitude. Using (7.28) and (7.29) the (Lagrangian) equation of motion for a displaced parcel is

$$\frac{\mathrm{d}^2\Delta s}{\mathrm{d}t^2} = -(N^2\cos^2\vartheta + f^2\sin^2\vartheta)\Delta s, \tag{7.30}$$

CHAPTER

8

Instability

WHAT HYDRODYNAMIC STATES OCCUR IN NATURE? Any flow must of course be a solution of the equations of motion, subject to the relevant initial and boundary conditions. There are many steady solutions to the equations of motion — certain purely zonal flows, for example — but the flows we experience are unsteady, time-dependent solutions, not steady solutions. Why should this be? It is because for any steady flow to persist it must be stable to those small perturbations that inevitably arise, but all the steady solutions that are known for the large-scale flow in the Earth's atmosphere and ocean have been found to be unstable.

Our focus in this chapter is on barotropic and baroclinic instability. *Baroclinic instability* (and we will define the term more precisely later on) is an instability that arises in rotating, stratified fluids that are subject to a horizontal temperature gradient. It is the instability that gives rise to the large- and mesoscale motion in the atmosphere and ocean — it produces atmospheric weather systems, for example — and so is, perhaps, the form of hydrodynamic instability that most affects the human condition. *Barotropic instability* is an instability that arises because of the shear in a flow, and may occur in fluids of constant density. It is important to us for two reasons: first, in its own right as an instability mechanism for jets and vortices and as an important process in turbulence; second, many problems in barotropic and baroclinic instability are very similar, so that the solutions and insight we obtain in the often simpler problems in barotropic instability may be useful in the baroclinic problem.

8.1 Kelvin–Helmholtz Instability

We first consider what is perhaps the simplest physically interesting instance of a fluid-dynamical instability — that of a constant-density flow with a shear perpendicular to the fluid's mean velocity, this being an example of a *Kelvin–Helmholtz instability*. Specifically, we consider two fluid masses of equal density, with an interface at $y = 0$, moving with velocities $-U$ and $+U$ in the x-direction, respectively, as in Fig. 8.1. There is no variation in the basic flow in the z-direction (normal to the page), and we will assume this is also true for the instability. This flow is clearly a solution of the Euler equations.

What happens if the flow is perturbed slightly? If the perturbation is initially small then even if it grows we can, for small times after the onset of instability, neglect the nonlinear interactions in the governing equations because these are the squares of small quantities. The equations determining the evolution of the initial perturbation are then the Euler equations linearized about the steady solution. Thus, denoting perturbation quantities with a prime and basic state variables with capital letters, for $y > 0$ the perturbation satisfies

$$\frac{\partial \boldsymbol{u}'}{\partial t} + U\frac{\partial \boldsymbol{u}'}{\partial x} = -\nabla p', \qquad \nabla\cdot\boldsymbol{u}' = 0, \tag{8.1a,b}$$

and a similar equation holds for $y < 0$, but with U replaced by $-U$. Given periodic boundary conditions in the x-direction, we may seek solutions of the form

$$\phi'(x, y, t) = \mathrm{Re}\sum_k \widetilde{\phi}_k(y)\exp[\mathrm{i}k(x - ct)], \tag{8.2}$$

where ϕ is any field variable (e.g., pressure or velocity), and Re denotes that only the real part should be taken. (Typically we use tildes over variables to denote Fourier-like modes, and we often omit the marker 'Re'.) Because (8.1a) is linear, the Fourier modes do not interact and we may confine attention to just one. Taking the divergence of (8.1a), the left-hand side vanishes and the pressure satisfies Laplace's equation

$$\nabla^2 p' = 0. \tag{8.3}$$

This has solutions in the form

$$p' = \begin{cases} \widetilde{p}_1 \mathrm{e}^{\mathrm{i}kx-ky}\mathrm{e}^{\sigma t} & y > 0, \\ \widetilde{p}_2 \mathrm{e}^{\mathrm{i}kx+ky}\mathrm{e}^{\sigma t} & y < 0, \end{cases} \tag{8.4}$$

where, anticipating the possibility of growing solutions, we have written the time variation in terms of a growth rate, $\sigma = -\mathrm{i}kc$. In general σ is complex: if it has a positive real component, the amplitude of the perturbation will grow and there is an instability; if σ has a non-zero imaginary component, then there will be oscillatory motion, and there may be both oscillatory motion *and* an instability. To obtain the dispersion relationship, we consider the y-component of (8.1a), namely (for $y > 0$)

$$\frac{\partial v_1'}{\partial t} + U\frac{\partial v_1'}{\partial x} = -\frac{\partial p_1'}{\partial y}. \tag{8.5}$$

Substituting a solution of the form $v_1' = \widetilde{v}_1\exp(\mathrm{i}kx + \sigma t)$ yields, with (8.4),

$$(\sigma + \mathrm{i}kU)\widetilde{v}_1 = k\widetilde{p}_1. \tag{8.6}$$

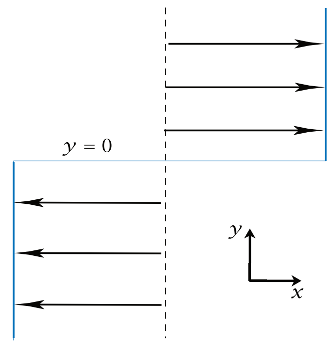

Fig. 8.1: A simple basic state giving rise to shear-flow instability. The velocity profile is discontinuous and the density is uniform. In more general problems in Kelvin–Helmholtz instability the density may also vary.

But the velocity normal to the interface is, at the interface, nothing but the rate of change of the position of the interface itself; that is, at $y = +0$

$$v_1 = \frac{\partial \eta'}{\partial t} + U \frac{\partial \eta'}{\partial x}, \tag{8.7}$$

or

$$\widetilde{v}_1 = (\sigma + \mathrm{i}kU)\widetilde{\eta}, \tag{8.8}$$

where η' is the displacement of the interface from its equilibrium position. Using this in (8.6) gives

$$(\sigma + \mathrm{i}kU)^2 \widetilde{\eta} = k\widetilde{p}_1. \tag{8.9}$$

The above few equations pertain to motion on the $y > 0$ side of the interface. Similar reasoning on the other side gives (at $y = -0$)

$$(\sigma - \mathrm{i}kU)^2 \widetilde{\eta} = -k\widetilde{p}_2. \tag{8.10}$$

But at the interface $p_1 = p_2$, because pressure must be continuous. The dispersion relationship then emerges from (8.9) and (8.10), giving

$$\sigma^2 = k^2 U^2. \tag{8.11}$$

This equation has two roots, one of which is positive. Thus, the amplitude of the perturbation grows exponentially, like $\mathrm{e}^{\sigma t}$, and the flow is *unstable*. The instability itself can be seen in the natural world when billow clouds appear wrapped up into spirals: the clouds are acting as tracers of fluid flow, and are a manifestation of the instability at finite amplitude, as seen later in Fig. 8.2.

8.2 Instability of Parallel Shear Flow

We now consider a little more systematically the instability of parallel shear flows. This is a classic problem in hydrodynamic stability theory, and there are two particular reasons for our own interest:

(i) The instability is an example of *barotropic instability,* which abounds in the ocean and atmosphere. Roughly speaking, barotropic instability arises when a flow is unstable by virtue of its horizontal shear, with gravitational and buoyancy effects being secondary.

(ii) The instability is in many ways analogous to *baroclinic instability,* which is the main instability giving rise to weather systems in the atmosphere and similar phenomena in the ocean.

We restrict attention to two-dimensional, incompressible flow; this illustrates the physical mechanisms in the most transparent way, in part because it allows for the introduction of a streamfunction and the automatic satisfaction of the mass continuity equation. In fact, for parallel two-dimensional shear flows the most unstable disturbances are two-dimensional ones.

The vorticity equation for incompressible two-dimensional flow is just

$$\frac{D\zeta}{Dt} = 0. \tag{8.12}$$

We suppose the basic state to be a parallel flow in the x-direction that may vary in the y-direction. That is

$$\overline{\boldsymbol{u}} = U(y)\hat{\mathbf{i}}. \tag{8.13}$$

The linearized vorticity equation is then

$$\frac{\partial \zeta'}{\partial t} + U\frac{\partial \zeta'}{\partial x} + v'\frac{\partial Z}{\partial y} = 0, \tag{8.14}$$

where $Z = -\partial_y U$. Because the mass continuity equation has the simple form $\partial u'/\partial x + \partial v'/\partial y = 0$, we may introduce a streamfunction ψ such that $u' = -\partial\psi'/\partial y$, $v' = \partial\psi'/\partial x$ and $\zeta' = \nabla^2\psi'$. The linear vorticity equation becomes

$$\frac{\partial \nabla^2\psi'}{\partial t} + U\frac{\partial \nabla^2\psi'}{\partial x} + \frac{\partial Z}{\partial y}\frac{\partial \psi'}{\partial x} = 0. \tag{8.15}$$

The coefficients of the x-derivatives are not themselves functions of x; thus, we may seek solutions that are harmonic functions (sines and cosines) in the x-direction, but the y dependence must remain arbitrary at this stage and we write

$$\psi' = \mathrm{Re}\,\widetilde{\psi}(y)\mathrm{e}^{\mathrm{i}k(x-ct)}. \tag{8.16}$$

The full solution is a superposition of all wavenumbers, but since the problem is linear the waves do not interact and it suffices to consider them separately. If c is purely real then c is the phase speed of the wave; if c has a positive imaginary component then the wave will grow exponentially and is thus *unstable*.

From (8.16) we have

$$u' = \widetilde{u}(y)\mathrm{e}^{\mathrm{i}k(x-ct)} = -\widetilde{\psi}_y\mathrm{e}^{\mathrm{i}k(x-ct)}, \tag{8.17a}$$

$$v' = \widetilde{v}(y)\mathrm{e}^{\mathrm{i}k(x-ct)} = \mathrm{i}k\widetilde{\psi}\mathrm{e}^{\mathrm{i}k(x-ct)}, \tag{8.17b}$$

$$\zeta' = \widetilde{\zeta}(y)\mathrm{e}^{\mathrm{i}k(x-ct)} = (-k^2\widetilde{\psi} + \widetilde{\psi}_{yy})\mathrm{e}^{\mathrm{i}k(x-ct)}, \tag{8.17c}$$

where the y subscript denotes a derivative. Using (8.17) in (8.14) gives

$$(U-c)(\widetilde{\psi}_{yy} - k^2\widetilde{\psi}) - U_{yy}\widetilde{\psi} = 0, \tag{8.18}$$

which is known as *Rayleigh's equation*. It is the linear vorticity equation for disturbances to parallel shear flow, and in the presence of a β-effect it generalizes slightly to

$$(U-c)(\widetilde{\psi}_{yy} - k^2\widetilde{\psi}) + (\beta - U_{yy})\widetilde{\psi} = 0, \tag{8.19}$$

which is known as the Rayleigh–Kuo equation.

John Strutt (1842–1919) became 3rd Baron Rayleigh on the death of his father in 1873 and, in a testament to the enduring British class system, is almost universally known as Lord Rayleigh. He made major contributions in many areas of physics, among them fluid mechanics (including the theory of sound and instability theory), the analysis of the composition of gases (leading to the discovery of argon), and scattering theory.

8.2.1 Piecewise Linear Flows

Although Rayleigh's equation is linear and has a simple form, it is nevertheless quite difficult to analytically solve for an arbitrary smoothly varying profile. It is simpler to consider *piecewise linear* flows, in which U_y is constant over some interval, with U or U_y changing abruptly to another value at a line of discontinuity, as for example in Fig. 8.1. The curvature, U_{yy} is accounted for through the satisfaction of matching conditions, analogous to boundary conditions, at the lines of discontinuity (as in Section 8.1), and solutions in each interval are then exponential functions.

Jump or matching conditions

The idea, then, is to solve the linearized vorticity equation separately in the continuous intervals in which vorticity is constant, matching the solution with that in the adjacent regions. The matching conditions arise from two physical conditions:

(i) That normal stress should be continuous across the interface. For an inviscid fluid this implies that pressure be continuous.

(ii) That the normal velocity of the fluid on either side of the interface should be consistent with the motion of the interface itself.

Let us consider the implications of these two conditions.

(i) Continuity of pressure
The linearized momentum equation in the direction along the interface is:

$$\frac{\partial u'}{\partial t} + U\frac{\partial u'}{\partial x} + v'\frac{\partial U}{\partial y} = -\frac{\partial p'}{\partial x}. \tag{8.20}$$

For normal modes, $u' = -\widetilde{\psi}_y \mathrm{e}^{\mathrm{i}k(x-ct)}$, $v' = \mathrm{i}k\widetilde{\psi}\mathrm{e}^{\mathrm{i}k(x-ct)}$ and $p' = \widetilde{p}\mathrm{e}^{\mathrm{i}k(x-ct)}$, and (8.20) becomes

$$\mathrm{i}k(U-c)\widetilde{\psi}_y - \mathrm{i}k\widetilde{\psi}U_y = -\mathrm{i}k\widetilde{p}. \tag{8.21}$$

Because pressure is continuous across the interface we have the first *matching* or *jump condition*,

$$\Delta[(U-c)\widetilde{\psi}_y - \widetilde{\psi}U_y] = 0, \tag{8.22}$$

where the operator Δ denotes the difference in the values of the argument (in square brackets) across the interface. That is, the quantity $(U-c)\widetilde{\psi}_y - \widetilde{\psi}U_y$ is continuous.
We can obtain this condition directly from Rayleigh's equation, (8.19), written in the form

$$[(U-c)\widetilde{\psi}_y - U_y\widetilde{\psi}]_y + [\beta - k^2(U-c)]\widetilde{\psi} = 0. \tag{8.23}$$

Integrating across the interface gives (8.22).

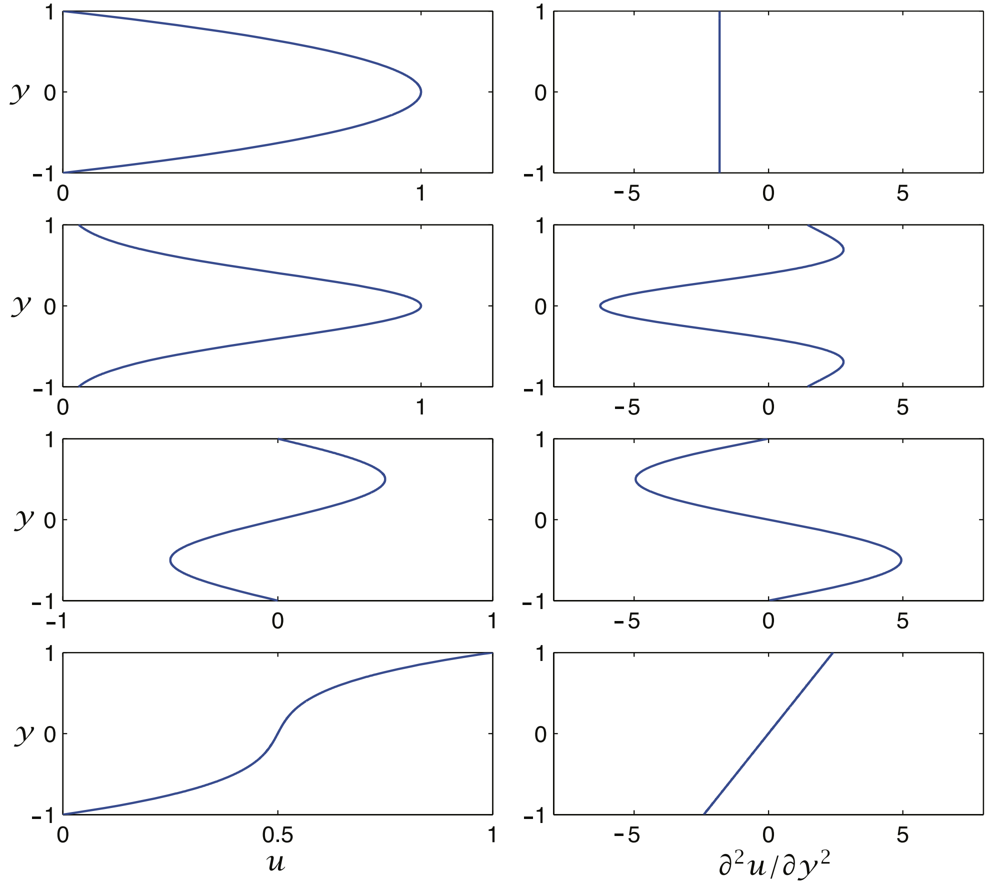

Fig. 8.5: Example parallel velocity profiles (left column) and their second derivatives (right column). From the top: Poiseuille flow ($u = 1 - y^2$); a Gaussian jet; a sinusoidal profile; a polynomial profile. By Rayleigh's criterion, the top profile is stable, whereas the lower three are potentially unstable. The bottom profile is in fact stable (although we do not demonstrate that here). If the β-effect were present and large enough it would stabilize the middle two profiles.

8.4.1 A Physical Picture

We first draw a picture of baroclinic instability as a form of 'sloping convection' in which the fluid, although statically stable, is able to release available potential energy when parcels move along a sloping path. To this end, let us first ask: what is the basic state that is baroclinically unstable? In a stably stratified fluid potential density decreases with height; we can also easily imagine a state in which the basic state temperature decreases, and the potential density increases, polewards. (We couch most of our discussion in terms of the Boussinesq equations and drop the qualifier 'potential' from density.) Can we construct a steady solution from these two conditions? The answer is yes, provided the fluid is also rotating; rotation is necessary because the meridional temperature gradient generally implies a meridional pressure gradient; there is nothing to balance this in the absence of rotation, and a fluid parcel would therefore accelerate. In a rotating fluid this pressure gradient can be balanced by the Coriolis force and a steady solution can be maintained even in the absence of viscosity. Consider a stably stratified Boussinesq fluid in geostrophic and hydrostatic balance on an f-plane, with buoyancy decreasing uniformly

CHAPTER

9

Waves and Mean-Flows

LINEAR DYNAMICS is mainly concerned with waves and instabilities that live on a pre-determined background flow. But the real world isn't quite like that. Rather, the mean state is the *result* of the combined effects of thermal and mechanical forcing (by radiation from the sun and, for the ocean, the winds) *plus the action of the waves and instabilities themselves*. In this chapter we explore the geophysical fluid dynamics underlying such wave–mean-flow interactions. We try to keep our discussion as elementary as possible by staying within the comfortable bounds of the quasi-geostrophic approximation and considering only zonal averages. Nevertheless the subject is often regarded as an advanced one, and all the sections in this chapter may be considered to be implicitly marked with a diamond, ✦.

9.1 Quasi-Geostrophic Wave–Mean-Flow Interaction

9.1.1 Preliminaries

The Jacobian operator for any two quantities a and b is given by $J(a, b) = \partial_x a\, \partial_y b - \partial_y a\, \partial_x b$. When a is the streamfunction the Jacobian gives the horizontal advection of b, namely $\boldsymbol{u} \cdot \nabla b$.

To fix our dynamical system and notation, we write down the Boussinesq quasi-geostrophic potential vorticity equation

$$\frac{\partial q}{\partial t} + J(\psi, q) = D, \tag{9.1}$$

where $J(\psi, q) = \partial\psi/\partial x\, \partial q/\partial y - \partial\psi/\partial y\, \partial q/\partial x$ and D represents any non-conservative terms. The potential vorticity in a Boussinesq system is

$$q = \beta y + \zeta + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2} b\right), \tag{9.2}$$

where ζ is the relative vorticity and b is the buoyancy perturbation from a background state characterized by N^2. (Nearly all the derivations in this chapter could be done in pressure coordinates with minor modifications.)

We refer to lines of constant b as isentropes. In terms of the streamfunction, the variables are

$$\zeta = \nabla^2\psi, \qquad b = f_0\frac{\partial\psi}{\partial z}, \qquad q = \beta y + \left[\nabla^2 + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial}{\partial z}\right)\right]\psi, \tag{9.3}$$

where $\nabla^2 \equiv (\partial_x^2 + \partial_y^2)$. The potential vorticity equation holds in the fluid interior; the boundary conditions on (9.3) are provided by the thermodynamic equation

$$\frac{\partial b}{\partial t} + J(\psi, b) + wN^2 = S, \tag{9.4}$$

where S represents heating terms. The vertical velocity at the boundary, w, is zero in the absence of topography and Ekman friction so that the boundary condition is just

$$\frac{\partial b}{\partial t} + J(\psi, b) = S. \tag{9.5}$$

Equations (9.1) and (9.5) are the evolution equations for the system, and if both D and S are zero they conserve both the total energy, $\widehat{E}$ and the total enstrophy, $\widehat{Z}$:

$$\begin{aligned} \frac{d\widehat{E}}{dt} &= 0, \qquad \widehat{E} = \frac{1}{2}\int_V (\nabla\psi)^2 + \frac{f_0^2}{N^2}\left(\frac{\partial\psi}{\partial z}\right)^2 dV, \\ \frac{d\widehat{Z}}{dt} &= 0, \qquad \widehat{Z} = \frac{1}{2}\int_V q^2\, dV, \end{aligned} \tag{9.6}$$

where V is a volume bounded by surfaces at which the normal velocity is zero, or that has periodic boundary conditions. The enstrophy is also conserved layerwise; that is, the horizontal integral of q^2 is conserved at every level.

9.1.2 Potential Vorticity Flux in the Linear Equations

Let us decompose the fields into a mean (to be denoted with an overbar) plus a perturbation (denoted with a prime), and let us suppose the perturbation fields are of small amplitude. (In linear problems, such as those considered in Chapter 8, we decomposed the flow into a 'basic state' plus a perturbation, with the basic state fixed in time. Our approach here is similar, but soon we will allow the mean state to evolve.) The linearized quasi-geostrophic potential vorticity equation is then

$$\frac{\partial q'}{\partial t} + \overline{u}\frac{\partial q'}{\partial x} + u'\frac{\partial \overline{q}}{\partial x} + \overline{v}\frac{\partial q'}{\partial y} + v'\frac{\partial \overline{q}}{\partial y} = D', \tag{9.7}$$

where D' represents eddy forcing and dissipation and, in terms of streamfunction,

$$\left(u'(x,y,z,t), v'(x,y,z,t)\right) = \left(-\frac{\partial\psi'}{\partial y}, \frac{\partial\psi'}{\partial x}\right), \tag{9.8a}$$

$$q'(x, y, z, t) = \nabla^2\psi' + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi'}{\partial z}\right). \tag{9.8b}$$

If the mean is a zonal mean then $\partial\overline{q}/\partial x = 0$ and $\overline{v} = 0$ (because v is purely geostrophic) and (9.7) simplifies to

$$\frac{\partial q'}{\partial t} + \overline{u}\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = D', \tag{9.9}$$

where

$$\overline{q} = \beta y - \frac{\partial \overline{u}}{\partial y} + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}\overline{b}\right), \qquad \frac{\partial \overline{q}}{\partial y} = \beta - \frac{\partial^2 \overline{u}}{\partial y^2} - \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \overline{u}}{\partial z}\right), \tag{9.10a,b}$$

having used the thermal wind relation,

$$f_0\frac{\partial \overline{u}}{\partial z} = -\frac{\partial \overline{b}}{\partial y}. \tag{9.11}$$

Multiplying (9.9) by q' and zonally averaging gives the enstrophy equation:

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{q'^2} = -\overline{v'q'}\frac{\partial \overline{q}}{\partial y} + \overline{D'q'}. \tag{9.12}$$

The quantity $\overline{v'q'}$ is the meridional flux of potential vorticity; this is downgradient (by definition) when the first term on the right-hand side is positive (i.e., $\overline{v'q'}\partial\overline{q}/\partial y < 0$), and it then acts to increase the variance of the perturbation. (This occurs, for example, when the flux is diffusive so that $\overline{v'q'} = -\kappa\partial\overline{q}/\partial y$, where κ may vary but is everywhere positive.) This argument may be inverted: for unforced, inviscid flow ($D = 0$), if the waves are growing, as for example in the canonical models of baroclinic instability discussed in Chapter 8, then *the potential vorticity flux is downgradient*.

If the second term on the right-hand side of (9.12) is negative, as it will be if D' is a dissipative process (e.g., if $D' = A\nabla^2 q'$ or if $D' = -rq'$, where A and r are positive) then a statistical balance can be achieved between enstrophy production via downgradient transport, and dissipation. If the waves are steady (by which we mean statistically steady, neither growing nor decaying in amplitude) and conservative (i.e., $D' = 0$) then we must have

$$\overline{v'q'} = 0. \tag{9.13}$$

Similar results follow for the buoyancy at the boundary; we start by linearizing the thermodynamic equation (9.5) to give

$$\frac{\partial b'}{\partial t} + \overline{u}\frac{\partial b'}{\partial x} + v'\frac{\partial \overline{b}}{\partial y} = S', \tag{9.14}$$

where S' is a diabatic source term. Multiplying (9.14) by b' and averaging gives

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{b'^2} = -\overline{v'b'}\frac{\partial \overline{b}}{\partial y} + \overline{S'b'}. \tag{9.15}$$

If the flow is adiabatic ($S' = 0$) then growing waves have a downgradient flux of buoyancy at the boundary. In the Eady problem there is no interior gradient of basic-state potential vorticity and all the terms in (9.12) are zero, but the perturbation grows at the boundary. If the waves are steady and adiabatic then, analogously to (9.13), $\overline{v'b'} = 0$. In models with discrete vertical layers or a finite number of levels it is common practice to absorb the boundary conditions into the definition of potential vorticity at top and bottom, as in the two level model of Section 5.6.

9.1.3 Wave–Mean-Flow Interaction

In linear problems we usually suppose that the mean-flow is fixed and that the zonal mean terms, $\overline{u}$ and $\overline{q}$ in (9.9), are functions only of y and z. However, in reality we might expect that the mean-flow would change because of momentum and heat flux convergences arising from the eddy–eddy interactions. To calculate these changes we begin with the potential vorticity equation (9.1) and, in the usual way, express the variables as a zonal mean plus an eddy term and obtain

$$\frac{\partial \overline{q}}{\partial t} + \overline{v}\frac{\partial \overline{q}}{\partial y} + \frac{\partial}{\partial y}(\overline{v'q'}) = \overline{D}. \tag{9.16}$$

Now, $\overline{v} = 0$ (since the flow is geostrophic) and the mean-flow thus evolves according to

$$\frac{\partial \overline{q}}{\partial t} + \frac{\partial}{\partial y}\overline{v'q'} = \overline{D}. \tag{9.17}$$

Similarly, at the boundary the mean buoyancy evolution equation is

$$\frac{\partial \overline{b}}{\partial t} + \frac{\partial}{\partial y}\overline{v'b'} = \overline{S}. \tag{9.18}$$

To obtain $\overline{u}$ from $\overline{q}$ and $\overline{b}$ we use thermal wind balance, (9.11), to define a streamfunction Ψ. That is, since $f_0 \partial\overline{u}/\partial z = -\partial\overline{b}/\partial y$, then

$$\left(\overline{u}, \frac{1}{f_0}\overline{b}\right) = \left(-\frac{\partial \Psi}{\partial y}, \frac{\partial \Psi}{\partial z}\right), \tag{9.19}$$

whence, using (9.10a), the zonal mean potential vorticity is

$$\overline{q}(y,z,t) - \beta y = \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \Psi}{\partial z}\right) + \frac{\partial^2 \Psi}{\partial y^2}. \tag{9.20}$$

If $\overline{q}$ is known in the interior from (9.18), and $\overline{b}$ (i.e., $f_0\partial\Psi/\partial z$) is known at the boundaries, then $\overline{u}$ and $\overline{b}$ in the interior may be obtained using (9.20) and (9.19b). The equations are also summarized in the shaded box on page 176.

To close the system we suppose that the eddy terms themselves evolve according to (9.9) and (9.14). If in those equations we were to include the eddy–eddy interaction terms we would simply recover the full system, so in neglecting those terms we have constructed an eddy–mean-flow system, commonly called a *wave–mean-flow* system because by eliminating the nonlinear terms in the perturbation equation the eddies will often be wavelike. It is important to realise that such systems do differ from linear ones in which we regard the mean flow as fixed; we have gone one step further and allowed the mean flow to evolve because of the effects of eddies, but we do not allow the eddies to interact with themselves.

We now consider some more properties of the waves themselves — how they propagate and what they conserve — beginning with a discussion of the potential vorticity flux and its relative, the so-called Eliassen–Palm flux.

9.2 Potential Vorticity Flux

The eddy flux of potential vorticity may be expressed in terms of vorticity and buoyancy fluxes as

$$v'q' = v'\zeta' + f_0 v' \frac{\partial}{\partial z}\left(\frac{b'}{N^2}\right). \tag{9.21}$$

The second term on the right-hand side can be written as

$$\begin{aligned} f_0 v' \frac{\partial}{\partial z}\left(\frac{b'}{N^2}\right) &= f_0 \frac{\partial}{\partial z}\left(\frac{v'b'}{N^2}\right) - f_0 \frac{\partial v'}{\partial z}\frac{b'}{N^2} \\ &= f_0 \frac{\partial}{\partial z}\left(\frac{v'b'}{N^2}\right) - f_0 \frac{\partial}{\partial x}\left(\frac{\partial \psi'}{\partial z}\right)\frac{b'}{N^2} \\ &= f_0 \frac{\partial}{\partial z}\left(\frac{v'b'}{N^2}\right) - \frac{f_0^2}{2N^2}\frac{\partial}{\partial x}\left(\frac{\partial \psi'}{\partial z}\right)^2, \end{aligned} \tag{9.22}$$

using $b' = f_0 \partial\psi'/\partial z$.

Similarly, the flux of relative vorticity can be written

$$v'\zeta' = -\frac{\partial}{\partial y}(u'v') + \frac{1}{2}\frac{\partial}{\partial x}(v'^2 - u'^2), \tag{9.23}$$

and using (9.22) and (9.23), (9.21) becomes

$$v'q' = -\frac{\partial}{\partial y}(u'v') + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}v'b'\right) + \frac{1}{2}\frac{\partial}{\partial x}\left((v'^2 - u'^2) - \frac{b'^2}{N^2}\right). \tag{9.24}$$

Thus the meridional potential vorticity flux, in the quasi-geostrophic approximation, can be written as the divergence of a vector: $v'q' = \nabla \cdot \boldsymbol{\mathcal{E}}$ where

$$\boldsymbol{\mathcal{E}} \equiv \frac{1}{2}\left((v'^2 - u'^2) - \frac{b'^2}{N^2}\right)\hat{\mathbf{i}} - (u'v')\,\hat{\mathbf{j}} + \left(\frac{f_0}{N^2}v'b'\right)\hat{\mathbf{k}}. \tag{9.25}$$

A particularly useful form of this arises after zonally averaging, for then (9.24) becomes

$$\overline{v'q'} = -\frac{\partial}{\partial y}\overline{u'v'} + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}\overline{v'b'}\right). \tag{9.26}$$

The vector defined by

$$\boldsymbol{\mathcal{F}} \equiv -\overline{u'v'}\,\hat{\mathbf{j}} + \frac{f_0}{N^2}\overline{v'b'}\,\hat{\mathbf{k}} \tag{9.27}$$

is the wave activity flux, often called the (quasi-geostrophic) Eliassen–Palm (EP) flux, and its divergence, given by (9.26), gives the poleward flux of potential vorticity:

$$\overline{v'q'} = \nabla_x \cdot \boldsymbol{\mathcal{F}}, \tag{9.28}$$

where $\nabla_x \cdot \equiv (\partial/\partial y, \partial/\partial z)\cdot$ is the divergence in the meridional-vertical plane, at constant x. Unless the meaning is unclear, the subscript x will be dropped.

9.2.1 The Eliassen–Palm Relation

The Eliassen–Palm flux and Eliassen–Palm relation are named for quantities that appeared in Eliassen & Palm (1961). That paper was mainly concerned with the transfer of energy in mountain waves rather than with matters related to potential vorticity flux and large-scale flow. 'Wave activity flux' is a descriptive alternative name for the EP flux.

On dividing by $\partial\overline{q}/\partial y$ and using (9.28), the enstrophy equation (9.12) becomes

$$\frac{\partial \mathcal{P}}{\partial t} + \nabla \cdot \boldsymbol{\mathcal{F}} = \mathcal{D}, \tag{9.29}$$

where

$$\mathcal{P} = \frac{\overline{q'^2}}{2\partial\overline{q}/\partial y}, \qquad \mathcal{D} = \frac{\overline{D'q'}}{\partial\overline{q}/\partial y}, \tag{9.30a,b}$$

and $\boldsymbol{\mathcal{F}}$ is given by (9.27). Equation (9.29) is known as the *Eliassen–Palm relation,* and it is a conservation law (when $\mathcal{D} = 0$) for the *pseudomomentum* $\mathcal{P}$. The conservation law is exact (in the linear approximation) if the mean-flow is constant in time; it is a good approximation if $\partial\overline{q}/\partial y$ varies slowly compared to the variation of $\overline{q'^2}$.

If we integrate (9.29) over a meridional area A bounded by walls where the eddy activity vanishes, and if $\mathcal{D} = 0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_A \mathcal{P}\,\mathrm{d}A = 0. \tag{9.31}$$

The integral is a 'wave activity' — a quantity that is quadratic in the amplitude of the perturbation and that is conserved in the absence of forcing and dissipation. The quantity $\mathcal{P}$ is an example of a 'wave activity density', generically denoted $\mathcal{A}$; other kinds of wave activity density exist — the pseudoenergy for example, but we do not consider them here. If there is no ambiguity we drop the word density and also refer to $\mathcal{A}$ and $\mathcal{P}$ as wave activities. Note that neither the perturbation energy nor the perturbation enstrophy are wave activities of the linearized equations, because there can be an exchange of energy or enstrophy between mean and perturbation — indeed, this is how a perturbation grows in baroclinic or barotropic instability! This is already evident from (9.12), or in general take (9.7) with $D' = 0$ and multiply by q' to give the enstrophy equation,

$$\frac{1}{2}\frac{\partial q'^2}{\partial t} + \frac{1}{2}\overline{\boldsymbol{u}} \cdot \nabla q'^2 + \boldsymbol{u}'q' \cdot \nabla\overline{q} = 0, \tag{9.32}$$

The value of this approach becomes more apparent when we consider multiple layers of fluid, or equivalently if we express the continuous system in isentropic coordinates, in which the thickness of isentropic layer of fluid is used as one of the state variables. The residual velocity in the TEM approach is the same as that which arises from a thickness-weighted average, and this velocity represents the actual average flow of fluid parcels more truthfully than does a conventional Eulerian average at a fixed height. The interested reader may pursue this topic in the references given at the end of the chapter.

9.4 The Non-Acceleration Result

In this section we derive an important result in wave–mean-flow dynamics, the so-called non-acceleration condition. Under certain conditions, to be made precise below, we can show that waves have no net effect on the mean-flow, an important and somewhat counter-intuitive result.

9.4.1 A Derivation from the Potential Vorticity Equation

Consider how the potential vorticity fluxes affect the mean fields. The unforced and inviscid zonally-averaged potential vorticity equation is

$$\frac{\partial \overline{q}}{\partial t} + \frac{\partial \overline{v'q'}}{\partial y} = 0. \tag{9.62}$$

Now, in quasi-geostrophic theory the geostrophically balanced velocity and buoyancy can be determined from the potential vorticity via an elliptic equation, as in (9.20), namely

$$\overline{q} - \beta y = \frac{\partial^2 \overline{\psi}}{\partial y^2} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \overline{\psi}}{\partial z}\right), \tag{9.63}$$

where $\overline{\psi}$ is such that $(\overline{u}, \overline{b}/f_0) = (-\partial\overline{\psi}/\partial y, \partial\overline{\psi}/\partial z)$. Differentiating (9.62) with respect to y we obtain

$$\left[\frac{\partial^2}{\partial y^2} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial}{\partial z}\right)\right]\frac{\partial \overline{u}}{\partial t} = (\nabla\cdot\boldsymbol{\mathcal{F}})_{yy}, \tag{9.64}$$

where $\nabla\cdot\boldsymbol{\mathcal{F}} = \overline{v'q'}$ is the divergence of the EP flux (in the y–z plane, i.e., $\nabla_x\cdot\boldsymbol{\mathcal{F}}$). This is determined using the wave activity equation for pseudomomentum which, from (9.29), is

$$\frac{\partial \mathcal{P}}{\partial t} + \nabla\cdot\boldsymbol{\mathcal{F}} = \mathcal{D}, \tag{9.65}$$

where $\mathcal{P}$ is the pseudomomentum. If the waves are statistically steady (i.e., $\partial\mathcal{P}/\partial t = 0$) and have no dissipation ($\mathcal{D} = 0$) then evidently $\nabla\cdot\boldsymbol{\mathcal{F}} = 0$. If there is no acceleration at the boundaries then the solution of (9.64) is

$$\frac{\partial \overline{u}}{\partial t} = 0. \tag{9.66}$$

Chapter

10 Turbulence

An apocryphal story that has been attributed to both Horace Lamb and Werner Heisenberg goes as follows. 'When I die and go to Heaven,' they are each said to have predicted, 'I would like to ask my Maker to explain two things, namely turbulence and quantum electrodynamics. About the latter I am hopeful of getting an answer.' Aside from the confidence of these two men as to where they were headed, the story speaks to the inherent difficulty of turbulence. But they may have been more likely to get an answer had their lives been more dissolute, for it is said that turbulence is the invention of the Devil, put on Earth to torment us.

Horace Lamb (1849–1934) was a British applied mathematician who published the classic text Hydrodynamics *in 1895, probably the oldest book on fluid mechanics still in print. Werner Heisenberg (1901–1976) was a German physicist most famous for the matrix formulation of quantum mechanics and the uncertainty principle. His doctoral thesis of 1923 was, however, concerned with turbulence, and he returned to the subject in the late 1940s.*

Putting aside these metaphysical issues, in this chapter we will give a introduction to three concrete aspects of turbulence: (i) turbulent diffusion; (ii) the classical spectral scaling theory of turbulence in two and three dimensions; and (iii) the theory of geostrophic turbulence. Before all that we'll describe what the 'problem of turbulence' actually is.

10.1 The Problem of Turbulence

What is turbulence? Roughly speaking, turbulence is high Reynolds number fluid flow with both spatial and temporal disorder, and a couple of beautiful sketches of what Leonardo da Vinci called *turbolenza* are shown in Fig. 10.1.

Traditionally, turbulent flow has often been thought of as occurring at small scales but in fact a turbulent flow has, as a consequence of being so disordered, a *range* of scales from large to small. The weather itself is an example of a turbulent flow — the great storms sweeping across the midlatitudes contain many scales of motion within them and, as we know from experience, are very hard to predict. Still, turbulence in general and weather in particular *do* have predictable aspects — we know that next winter will be colder than this summer, and that any given month in

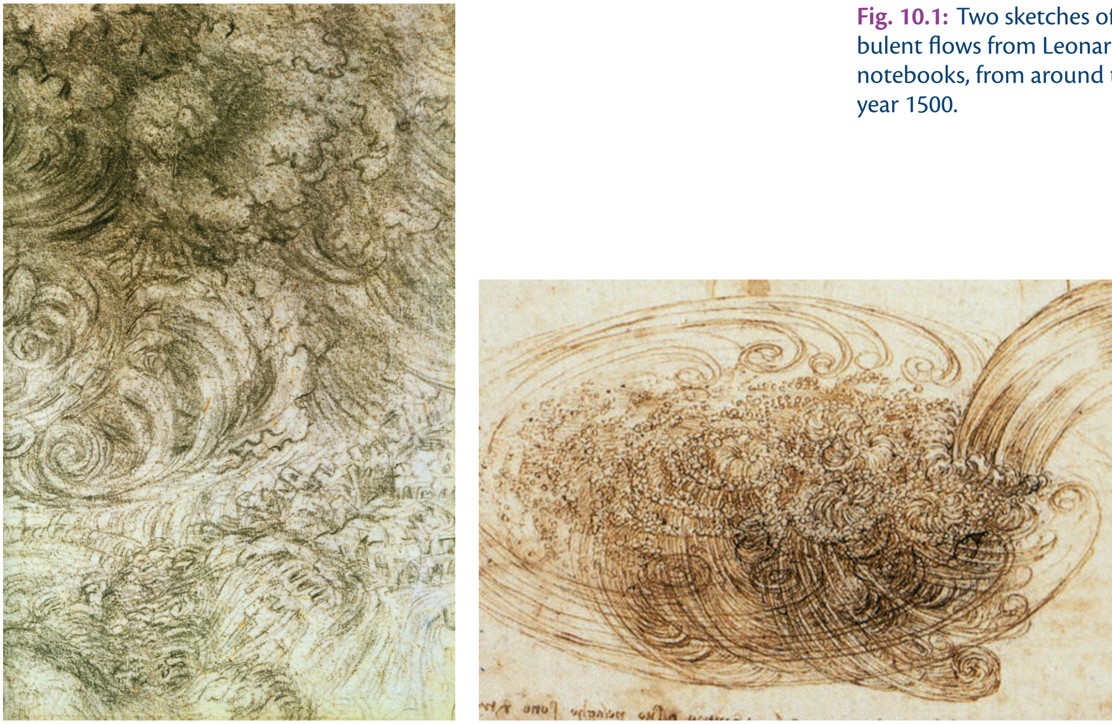

Fig. 10.1: Two sketches of turbulent flows from Leonardo's notebooks, from around the year 1500.

Spain will be warmer than the same month in the UK, and we know that if a storm is approaching from the west it is likely to be windier and rainier than normal, even if we do not know exactly when or where. We might like to be able to predict the *average* flow over a wide area or over a period of a time without necessarily predicting every detail. However, the details may be important — if not in themselves but because they have an effect on the large scale by transporting and mixing properties of the fluid; thus, a turbulent fluid will become well-mixed much more quickly than a laminar fluid. If we drop some ink into a glass of water then we can speed up the mixing by stirring the water, creating turbulent eddies that mix the ink into the water. However, to go beyond this picture and to properly understand the effect of the small scales on the large ones is very difficult because of the 'closure problem', as we now see.

Leonardo da Vinci (1452–1519) is now most famous for his paintings, in particular Mona Lisa *and* The Last Supper*, but he was also a sculptor, architect, inventor and scientist — the embodiment of a 'renaissance man'. He made a number of drawings of eddying flow and coined the word* turbolenza *to describe it.*

10.1.1 The Closure Problem

Let us suppose that a flow has a mean component and a fluctuating component, so that the velocity is given by

$$\boldsymbol{v} = \overline{\boldsymbol{v}} + \boldsymbol{v}'. \tag{10.1}$$

Here $\overline{\boldsymbol{v}}$ is the mean velocity field, and $\boldsymbol{v}'$ is the deviation from that mean. The mean might be a time average, in which case $\overline{\boldsymbol{v}}$ is a function only of space and not of time, or it might be a time mean over a finite period (e.g., a season if we are dealing with the weather), or it might be some form of ensemble mean. The average of the deviation is, by definition, zero; that is $\overline{\boldsymbol{v}'} = 0$. We would like to predict the evolution of the mean flow, $\overline{\boldsymbol{v}}$, without predicting the evolution of the eddying flow and to do this we might substitute (10.1) into the momentum equation and try to obtain a closed equation for the mean quantity $\overline{\boldsymbol{v}}$. To keep the algebra simple, and to avoid dealing with the full Navier–Stokes equations, let us carry out this program for a model nonlinear system that obeys

$$\frac{\mathrm{d}u}{\mathrm{d}t} + uu + ru = 0, \tag{10.2}$$

where r is a constant. The average of this equation is:

$$\frac{\mathrm{d}\overline{u}}{\mathrm{d}t} + \overline{uu} + r\overline{u} = 0. \tag{10.3}$$

The value of the term $\overline{uu}$ (i.e., $\overline{u^2}$) is not deducible simply by knowing $\overline{u}$, since it involves correlations between eddy quantities, namely $\overline{u'u'}$. That is, $\overline{uu} = \overline{u}\,\overline{u} + \overline{u'u'} \neq \overline{u}\,\overline{u}$. We can go to the next order to try (vainly!) to obtain an equation for $\overline{u}\,\overline{u}$. First multiply (10.2) by u to obtain an equation for u^2, and then average it to yield

$$\frac{1}{2}\frac{\mathrm{d}\overline{u^2}}{\mathrm{d}t} + \overline{uuu} + r\overline{u^2} = 0. \tag{10.4}$$

This equation contains the undetermined cubic term $\overline{uuu}$. An equation determining this would contain a quartic term, and so on in an unclosed hierarchy. Many methods of closing the hierarchy make assumptions about the relationship of $(n+1)$th order terms to nth order terms, for example by supposing that

$$\overline{uuuu} = \alpha\overline{uu}\,\overline{uu} + \beta\overline{uuu}, \tag{10.5}$$

where α and β are parameters to be determined, one may hope, by a theory. If we know that the variables are distributed normally then such closures can be made exact, but this is not generally true in turbulence and all closures that have been proposed so far are, at best, approximations.

This same closure problem arises in the Navier–Stokes equations. If density is constant (as we shall assume in this chapter) the x-momentum equation for an averaged flow is

$$\frac{\partial\overline{u}}{\partial t} + (\overline{\boldsymbol{v}}\cdot\nabla)\overline{u} = -\frac{\partial\overline{\phi}}{\partial x} - \nabla\cdot\overline{\boldsymbol{v}'u'}. \tag{10.6}$$

Written out in full in Cartesian coordinates, the last term is

$$\nabla\cdot\overline{\boldsymbol{v}'u'} = \frac{\partial}{\partial x}\overline{u'u'} + \frac{\partial}{\partial y}\overline{u'v'} + \frac{\partial}{\partial z}\overline{u'w'}. \tag{10.7}$$

These terms, and the similar ones in the y- and z-momentum equations, represent the effects of eddies on the mean flow and are known as *Reynolds stress* terms. The 'closure problem' of turbulence may be thought of as finding a representation of the Reynolds stresses in terms of mean flow quantities. Nobody has been able to usefully close the system without introducing physical assumptions not directly deducible from the equations of motion themselves. Indeed it is not clear that in general a useful closed-form solution even exists.

10.2 Turbulent Diffusion

The most widely used recipe to address the closure problem is by way of turbulent diffusion, or eddy diffusion. The idea comes by way of an analogy with molecular diffusion, and is roughly as follows. Suppose that a fluid carries with it a tracer, φ, that satisfies an equation like

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \kappa\nabla^2\varphi, \tag{10.8}$$

where κ is a molecular diffusivity. For simplicity we suppose the flow is two-dimensional and incompressible, and that the flow and the tracer have both a mean and a fluctuating component. The mean component of (10.8) may then be written as

$$\frac{\partial\overline{\varphi}}{\partial t} + \frac{\partial\overline{u}\overline{\varphi}}{\partial x} + \frac{\partial\overline{v}\overline{\varphi}}{\partial y} = -\frac{\partial\overline{u'\varphi'}}{\partial x} - \frac{\partial\overline{v'\varphi'}}{\partial y} + \kappa\nabla^2\overline{\varphi}, \tag{10.9}$$

Now, consider a fluctuating parcel of fluid that, on average, carries its value of φ with it a certain distance ℓ, a 'mixing length', before mixing with its surroundings. If there is a mean gradient of φ in the direction of movement (the y-direction, say) then the value of φ' is given by

$$\varphi' = -\ell\frac{\partial\varphi}{\partial y}. \tag{10.10}$$

If the dominant eddies have a typical speed v' then the eddy transport is given by

$$\overline{v'\varphi'} = -K\frac{\partial\overline{\varphi}}{\partial y}, \qquad \text{where} \qquad K = \overline{v'\ell}. \tag{10.11}$$

In this expression, K is an *eddy diffusivity*, the product of the velocity and length scale of the dominant eddies in the system. If we assume that a similar process occurs in the x-direction then (10.1) becomes

$$\frac{\partial\overline{\varphi}}{\partial t} + \frac{\partial\overline{u}\overline{\varphi}}{\partial x} + \frac{\partial\overline{v}\overline{\varphi}}{\partial y} = K\nabla^2\overline{\varphi} + \kappa\nabla^2\overline{\varphi}. \tag{10.12}$$

In most turbulent flows the eddy diffusivity is much larger than the molecular diffusivity κ because the mixing length is orders of magnitude larger than the corresponding molecular mixing length, which is the average distance that a molecule goes before interacting with another molecule. Thus, we commonly neglect the last term on the right-hand side in (10.12). However, the presence of a molecular viscosity is important in that it allows mixing to take place in the first instance; the turbulence amplifies the molecular mixing enormously, but that mixing must be present.

To suppose that turbulence acts like an enhanced diffusion is probably the most widely used parameterization of turbulence in practical situations. It is easily implemented and has a rational basis. On the other hand, there are many circumstances when it does not work, and momentum in large-scale atmospheric flows is typically not diffused downgradient, or diffused at all.

Equation (10.12) is a practical recipe, and no more, for treating the enhanced transport associated with a turbulent flow. It says that if we are unable to explicitly model the small scales of a turbulent flow, perhaps because we don't know what is happening at those scales, then we might be able to approximately simulate the effects of the small scales using a turbulent diffusion. The idea is rather ad hoc, because we don't have a good theory for the magnitude and structure of the eddy diffusion coefficient K, but it is often better than doing nothing.

The theory of turbulent diffusion stems from work by G. I. Taylor and L. Prandtl, two great figures in fluid dynamics in the early twentieth century.

10.2.1 Homogenization and Lack of Extrema

Now consider a tracer that is advected and diffused. The diffusion might be molecular or, if the effects of turbulence on a tracer are indeed diffusive, there might be an eddy diffusion. An important consequence of this is that, in the absence of additional forcing, there can be no extreme values of the tracer in the interior of the fluid and the diffusion acts to *homogenize* values of the tracer in broad regions.

Consider a tracer that obeys the equation

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \nabla \cdot (\kappa \nabla \varphi), \tag{10.13}$$

where $\kappa > 0$ and the advecting velocity is divergence-free. Given an extremum, there will then be a surrounding surface (in three dimensions), or a surrounding contour (in two), connecting constant values of φ. For definiteness consider three-dimensional incompressible flow which in the steady state flow satisfies

$$\nabla \cdot (\boldsymbol{v}\varphi) = \nabla \cdot (\kappa \nabla \varphi). \tag{10.14}$$

Integrating the left-hand side over the volume, V, enclosed by an isosurface, A, of φ, and applying the divergence theorem, gives

$$\iiint_V \nabla \cdot (\boldsymbol{v}\varphi)\,\mathrm{d}V = \iint_A (\boldsymbol{v}\varphi)\cdot \boldsymbol{n}\,\mathrm{d}A = \varphi \iint_A \boldsymbol{v}\cdot\boldsymbol{n}\,\mathrm{d}A = \varphi \iiint_V \nabla\cdot\boldsymbol{v}\,\mathrm{d}V = 0, \tag{10.15}$$

where $\boldsymbol{n}$ is a unit vector normal to the bounding surface. But the integral of the right-hand side of (10.14) over the same area is non-zero; that is

$$\iiint_V \nabla \cdot (\kappa \nabla \varphi)\,\mathrm{d}V = \iint_A \kappa \nabla\varphi \cdot \boldsymbol{n}\,\mathrm{d}A \neq 0, \tag{10.16}$$

if the integral surrounds an extremum. This is a contradiction for steady flow. Hence, there can be no isolated extrema of a conserved quantity in the interior of a fluid, if there is any diffusion at all. The result is kinematic, in that φ can be any tracer at all, active or passive.

Interpretation and consequences

The physical essence of the result is that the integrated effects of diffusion are non-zero surrounding an extremum, and cannot be balanced by where E is the energy density per unit mass, V is the volume of the domain, and the last equality serves to define the discrete energy spectrum $\mathcal{E}_{\boldsymbol{k}}$. We now assume that the turbulence is isotropic, and that the domain is sufficiently large that the sums in the above equations may be replaced by integrals. We may then write

$$\overline{E} = \frac{1}{V}\widehat{E} = \frac{1}{2V}\int_V \boldsymbol{v}^2 \,\mathrm{d}V = \int \mathcal{E}(k)\,\mathrm{d}k, \tag{10.20}$$

where $\overline{E}$ is the average energy, $\widehat{E}$ is the total energy and $\mathcal{E}(k)$ is the energy spectral density, or the energy spectrum, so that $\mathcal{E}(k)\,\delta k$ is the energy in the small wavenumber interval δk. Because of the assumed isotropy, the energy is a function only of the scalar wavenumber k, where $k^2 = \boldsymbol{k}\cdot\boldsymbol{k} = k^{x2} + k^{y2} + k^{z2}$. The units of $\mathcal{E}(k)$ are L^3/T^2 and the units of $\overline{E}$ are L^2/T^2.

10.3.2 Inertial-Range Theory

We now suppose that the fluid is stirred at large scales and that this energy is transferred to small scales where it is dissipated by viscosity. The key assumption is to suppose that, if the forcing scale is sufficiently larger than the dissipation scale, there exists a range of scales that is intermediate between the large scale and the dissipation scale and where neither forcing nor dissipation are explicitly important to the dynamics. This assumption, known as the *locality hypothesis,* depends on the nonlinear transfer of energy being sufficiently local (in spectral space). This intermediate range is known as the *inertial range,* because the inertial terms and not forcing or dissipation dominate in the momentum balance. If the rate of energy input per unit volume by stirring is equal to ε, then if we are in a steady state there must be a flux of energy from large to small scales that is also equal to ε, and an energy dissipation rate, also ε.

Now, we have no general theory for the energy spectrum of a turbulent fluid, but we might suppose it takes the general form

The theory described in this section is not an exact theory of turbulence. It relies on assumptions of spectral locality and the near constancy of the energy transfer across scales, and these assumptions are not exactly satisfied. Nevertheless, the theory has been enormously useful and is one of the enduring foundations of the field.

$$\mathcal{E}(k) = f(\varepsilon, k, k_0, k_\nu), \tag{10.21}$$

where the right-hand side denotes a function of the spectral energy flux or cascade rate ε, the wavenumber k, the forcing wavenumber k_0 and the wavenumber at which dissipation acts, k_ν (and $k_\nu \sim L_\nu^{-1}$). In general, the function f depends on the particular nature of the forcing. Now, the locality hypothesis essentially says that at some scale within the inertial range the flux of energy to smaller scales depends only on processes occurring at or near that scale. That is to say, the energy flux is only a function of $\mathcal{E}$ and k, or equivalently that the energy spectrum can be a function *only* of the energy flux ε and the wavenumber itself. From a physical point of view, as energy cascades to smaller scales the details of the forcing are forgotten but the effects of viscosity are not yet apparent, and the energy spectrum takes the form,

$$\mathcal{E}(k) = g(\varepsilon, k). \tag{10.22}$$

The function g is assumed to be *universal,* the same for every turbulent flow.

Part II

Atmospheres

Chapter

11

The Tropical Atmosphere

IN THIS CHAPTER AND THE TWO FOLLOWING we discuss the structure and circulation of planetary atmospheres. This chapter and the next focus mainly on Earth, first on the tropical circulation and then on the midlatitudes and the stratosphere. Then, in Chapter 13, we look at planetary atmospheres a little more generally. We begin with a brief observational overview of Earth's atmosphere as a whole.

11.1 An Observational Overview

Many of the main zonally- and/or time-averaged features of Earth's atmosphere can be seen in Figs. 11.1–11.3. The most prominent features are:

(i) The temperature falls monotonically with height to an altitude of about 16 km (in the tropics) or 8 km at high latitudes, before increasing with height. The lower region is called the *troposphere*, above which lies the *stratosphere*, and the boundary between the two is the *tropopause*.

(ii) The temperature also falls monotonically from equator to pole, at the surface falling from about 300 K to 240 K at the pole. The tropopause temperature is more uniform, varying from about 230 K to 210 K.

(iii) The surface winds are easterly in low latitudes (from about 30° S to 30° N), westerly in mid- and high latitudes, with weak polar easterlies in some seasons.

(iv) The winds increase in height, especially in midlatitudes, with pronounced westerly jets in both hemispheres centred at about 40° latitude in both hemispheres.

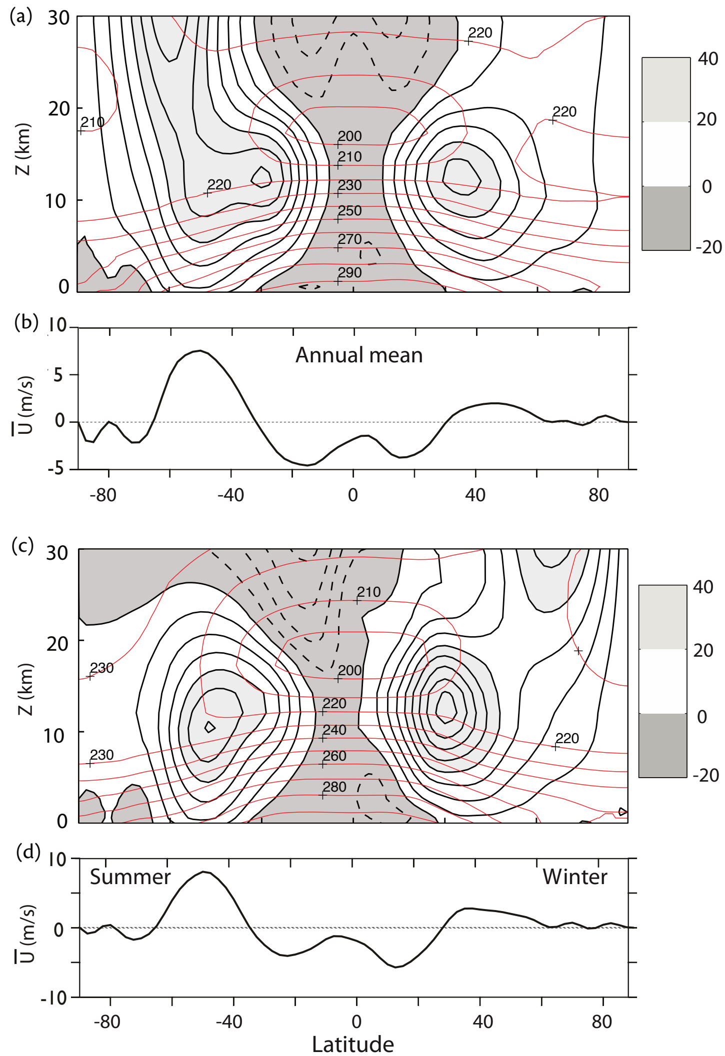

Fig. 11.1: (a) Annual mean, zonally-averaged zonal wind (heavy contours and shading) and the zonally-averaged temperature (red, thinner contours). (b) Annual mean, zonally averaged zonal winds at the surface. (c) and (d) Same as (a) and (b), except for northern hemisphere winter (December–January–February, or DJF).

The wind contours are at intervals of 5 m s^{-1} with dark shading for eastward winds above 20 m s^{-1} and light shading for all westward winds, and the temperature contours are labelled.

(*v*) In the meridional (y–z) plane the circulation in each hemisphere (Fig. 11.3) is characterized by:

(*i*) A 'direct' *Hadley Cell*, with warm air rising near the equator and sinking in the tropics. The winter Hadley Cell is much stronger and has greater latitudinal extent than the summer one, with the warm air rising at low latitudes in the summer hemisphere, sinking in the tropics in the winter hemisphere.

(*ii*) An 'indirect' *Ferrel Cell*, with cool air apparently rising around 60° and sinking in the tropics.

A direct cell is one that is thermally driven, with warm, buoyant fluid rising and cold fluid descending. An indirect cell may be mechanically driven, as we discuss later.

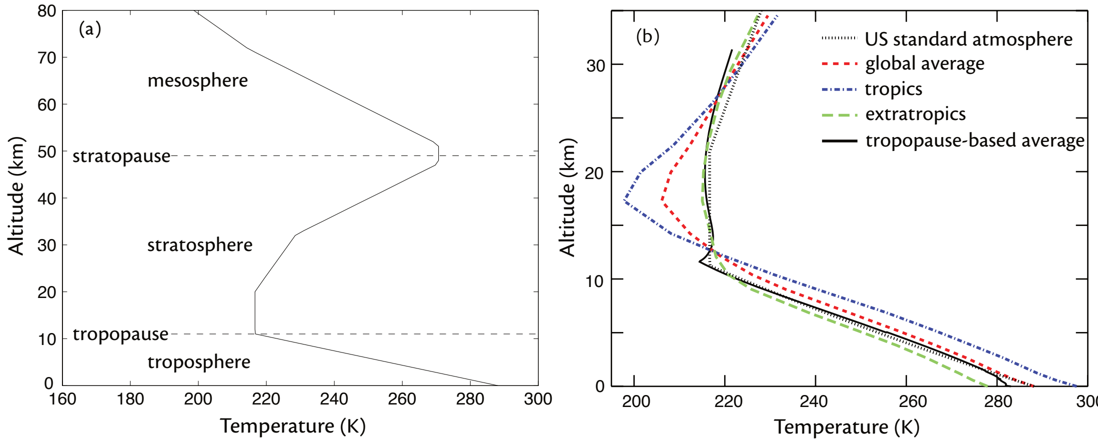

Fig. 11.2: (a) The temperature profile of the so-called US standard atmosphere. (b) Observed, annually averaged profiles of temperature in the atmosphere, where the ordinate is log-pressure. Tropics here is the average from 30°S to 30°N, and extratropics is the average over the rest of the globe. (The 'tropopause-based average' uses the tropopause itself as the origin of the height scale, set to be 11 km, but we shall not discuss it further.)

11.2 An Ideal Hadley Circulation

To makes matters as simple as possible we imagine the Earth to be a sphere with a uniform surface (so no mountains and no oceans), and that there are no seasons (unlike the case in Fig. 11.3. Because of the differential solar heating the air is warmer at low latitudes than at high, so we may reasonably imagine that the warm air rises and moves polewards before cooling and sinking at some high latitude, perhaps near the pole, and returning near the ground. Indeed such a concept was envisioned by George Hadley over 300 hundred years ago. However, observations tell us that the air does not go all the way to the pole; rather, it sinks in the subtropics at about 25–30°. There are two reasons why it must sink, one related to thermodynamic constraints and the other to hydrodynamic instabilities. Both are related to the properties of the air as it moves, conserving its angular momentum.

11.2.1 Zonally-Symmetric Equations of Motion

In Chapter 2 we wrote down the equations of motion on a sphere. If the flow is zonally symmetric (no longitudinal variation) then, with a little manipulation, the zonal momentum equation may be written in the form

$$\frac{\partial u}{\partial t} - (f + \zeta)v + w\frac{\partial u}{\partial z} = 0. \tag{11.1}$$

The variables in this equation are functions of latitude and height (ϑ and z) only, and not longitude, λ, and $\zeta = -(a\cos\vartheta)^{-1}\partial_\vartheta(u\cos\vartheta)$. If the vertical advection is small then a steady solution obeys

$$(f + \zeta)v = 0. \tag{11.2}$$

Presuming that the meridional flow v is non-zero then $f + \zeta = 0$, or equivalently, on the sphere,

$$2\Omega\sin\vartheta = \frac{1}{a}\frac{\partial u}{\partial\vartheta} - \frac{u\tan\vartheta}{a}. \tag{11.3}$$

The Hadley Cell is named for George Hadley (1685–1768), a British meteorologist who put forward perhaps the first scientific model of Earth's overturning circulation, in which the air rose near the equator and sank near the pole.

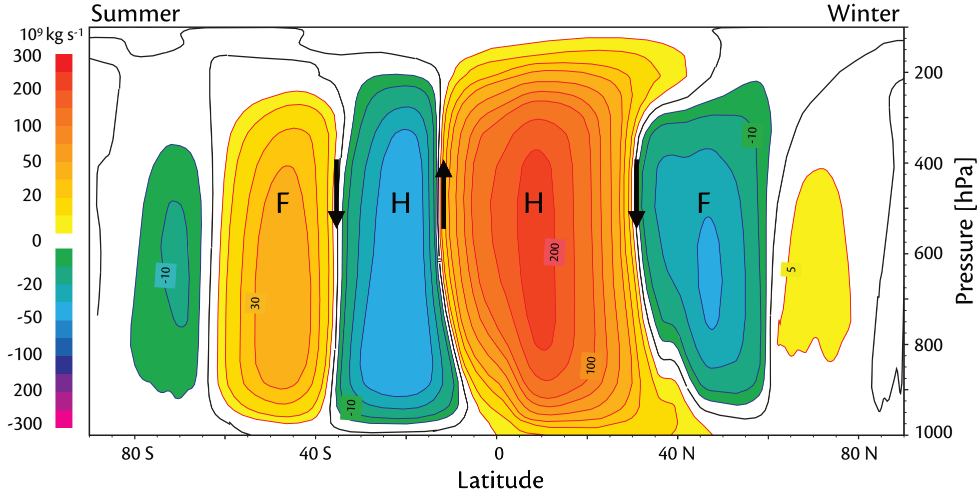

Fig. 11.3: The observed meridional overturning circulation (MOC) of the atmosphere ($\mathrm{kg\,s^{-1}}$) averaged over December–January– February. Note the direct *Hadley Cells* (H) particularly strong in winter with rising motion near the equator, descending motion in the subtropics, and the weaker, indirect, *Ferrel Cells* (F) in midlatitudes. (From an ECMWF reanalysis.)

At the equator we may assume that $u = 0$, because here parcels have risen from the surface where the flow is weak. Equation (11.3) then has a solution of

$$u = \Omega a \frac{\sin^2 \vartheta}{\cos \vartheta}. \tag{11.4}$$

This gives the zonal velocity of the poleward moving air in the upper branch of the (model) Hadley Cell, above the frictional boundary layer. Evidently the zonal velocity increases rapidly with latitude: at 20° and 40° the values of u are about 59 and 256 m s^{-1}respectively (look ahead to Fig. 11.6 or Fig. 11.7), becoming far larger than the observed values at midlatitudes.

Angular momentum conservation

An instructive interpretation of (11.4) comes by way of the conservation of angular momentum, m, of a ring of air at a latitude ϑ, as in Fig. 11.4. The angular momentum per unit pass of a parcel of air with zonal velocity u is

$$m = (u + \Omega a \cos \vartheta) a \cos \vartheta, \tag{11.5}$$

and if $u = 0$ at $\vartheta = 0$ and if m is conserved on a poleward moving parcel, then (11.5) leads directly to (11.4).

The air returning to the equator close to the surface has a small zonal velocity, meaning that there is a large thermal wind, which we may calculate using the thermal wind expression

$$2\Omega \sin \vartheta \frac{\partial u}{\partial z} = -\frac{1}{a}\frac{\partial b}{\partial \vartheta}, \tag{11.6}$$

where $b = g\,\delta\theta/\theta_0$ and $\delta\theta$ is the deviation of potential temperature from a constant reference value θ_0. (Be reminded that θ is potential temperature, This value of the divergence is used in (11.27b) and (11.27c) which, retaining all terms since none are obviously small, become

$$\frac{\partial \zeta}{\partial t} + \boldsymbol{u} \cdot \nabla(\zeta + f_0) + (\zeta + f)\frac{Q}{H} = -r\zeta, \tag{11.33b}$$

$$g\nabla^2 h = \hat{\mathbf{k}} \cdot \nabla \times [\boldsymbol{u}(\zeta + f_0)] - \frac{1}{H}\frac{\partial Q}{\partial t} - r\delta - \nabla^2 \frac{\boldsymbol{u}^2}{2}. \tag{11.33c}$$

The equation set (11.33) has but one prognostic equation, namely (11.33b), and so is truly balanced and may be thought of as a generalization of (11.25) to the case with non-zero heating. The divergence equation is a nonlinear balance equation except now with a diabatic term on the right-hand side. The divergent flow itself is computed using the height equation, by an assumed balance between adiabatic cooling and diabatic heating. The relationship between velocity and geopotential (or pressure) is the same as in the adiabatic case, because this arises through the momentum equation. Thus, even in the presence of a heating, gradients of geopotential and temperature remain relatively weak, a result that ultimately arises from the smallness of the Coriolis parameter. The importance of the result lies in what it implies about the response of the atmosphere to a localized heating: the equations provide a scaling for the response of the velocity, and suggest that the response may become spread out over a sufficient area to keep the temperature gradients small. This, taken with the scaling arguments of Section 5.3.1, is called the 'weak temperature gradient approximation'.

11.5 EFFECTS OF MOISTURE

In the rest of this chapter we talk about the two things that one immediately notices about the deep tropics — moisture and convection.

Moisture affects nearly every facet of tropical dynamics, and without it the tropics would be a very different place — no towering cumulonimbus clouds for example. However, convection *would* still take place since the basic state set up by the radiative forcing would still be convectively unstable.

11.5.1 Measures of Moisture

There are various measures of the amount of moisture in the atmosphere, so let us summarize them. The *absolute humidity* is the amount of water vapour per unit volume, with units of kg m^{-3}, or informally g m^{-3}. The *mixing ratio, w*, is the ratio of the mass of water vapour, m^v, to that of dry air, m^d, in some volume of air and is thus

$$w \equiv \frac{m^v}{m^d} = \frac{\rho^v}{\rho^d}. \tag{11.34}$$

It is a nondimensional measure but it is often expressed in terms of grams per kilogram. In the atmosphere values range from close to zero to about 2×10^{-2} (20 g kg^{-1}) in the tropics on a humid day.

The *specific humidity, q*, is the ratio of the mass of water vapour to the total mass of air — dry air plus water vapour — and so is

$$q \equiv \frac{m^v}{m^d + m^v} = \frac{w}{1+w} \quad \text{and} \quad w = \frac{q}{1-q}. \tag{11.35a,b}$$

CHAPTER

12

Midlatitudes and the Stratosphere

WHY DO THE DYNAMICS OF THE MIDLATITUDES differ from their tropical counterparts? Is the difference common to most planets or special to Earth? One difference, at least on Earth, is that the midlatitudes are baroclinically unstable, producing the weather. Even if the Hadley Cell were to terminate of its own accord before becoming baroclinically unstable, the radiative equilibrium temperature in midlatitudes has a meridional gradient that would be unstable. On other, more slowly rotating planets, the Hadley Cell might extend nearly all the way to the pole, in which case the planet may be thought of as entirely tropical! Venus and Titan (a moon of Saturn) are examples of such all-tropical planets, and others likely abound outside our Solar System.

Given this rather general point, in this chapter we will discuss two more specific properties of the Earth's midlatitudes: (i) The predominantly eastward surface winds and the strong eastward winds extending up to the tropopause. (ii) The meridional overturning circulation, or Ferrel Cell. Both features can be seen in Fig. 11.1 and Fig. 11.3, and both are consequences of the general phenomena of baroclinic instability and geostrophic turbulence, moulded by Earth's atmosphere, and they become intertwined in our discussion. None of the dynamics that we discuss in this chapter involves density variation in a truly essential way and readers may simplify the discussion by regarding density as constant.

12.1 JET FORMATION AND SURFACE WINDS

The atmosphere above the surface has a generally eastward flow, with a broad maximum about 10 km above the surface at around 40° in either hemisphere. But if we look a little more at the zonally average wind in Fig. 11.1(a), especially in the Southern Hemisphere, we see hints of there being two jets — one (the subtropical jet) at around 30°, and another somewhat poleward of this. The subtropical jet is associated with a strong meridional temperature gradient at the edge of the Hadley Cell and it is quite baroclinic — that is, there is a noticeable shear in the zonal wind. On the other hand, the midlatitude jet is more barotropic (it has little vertical structure, with less shear than the subtropical jet) and lies above an eastward surface flow. This flow feels the effect of surface friction and so there must be a momentum *convergence* into this region, as is seen in Fig. 12.4. This jet is known as the *eddy-driven jet.*

We encountered eddy-driven jets in our discussion of barotropic turbulence in Section 10.6. However, that case was homogeneous, with no preferred latitude for a particular jet, whereas in the atmosphere there appears to be but one midlatitude jet with a preferred average location, and in the sections that follow we discuss how this jet is maintained.

12.1.1 The Mechanism of Jet Production

For reference later on we establish a useful form of the zonal momentum equation. For two-dimensional, horizontally non-divergent flow we have

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} - fv = -\frac{\partial \phi}{\partial x} - D_u, \tag{12.1}$$

where D_u represents the effects of dissipation. We write the variables as the sum of a zonal mean plus a deviation so that $u = \overline{u} + u'$ and $v = \overline{v} + v'$, and for incompressible two dimensional flow $\overline{v} = 0$. The zonal average of (12.1) is then

$$\frac{\partial \overline{u}}{\partial t} + \frac{\partial \overline{u}\,\overline{v}}{\partial y} + \frac{\partial \overline{u'v'}}{\partial y} - f\overline{v} = -D_u, \tag{12.2}$$

but since $\overline{v} = 0$ we have

$$\frac{\partial \overline{u}}{\partial t} + \frac{\partial \overline{u'v'}}{\partial y} = -r\overline{u}, \tag{12.3}$$

where we also represent dissipation as a linear drag, with r being a constant.

We can write the momentum flux in terms of the vorticity flux since, for non-divergent two-dimensional flow,

$$v\zeta = \frac{1}{2}\frac{\partial}{\partial x}\left(v^2 - u^2\right) - \frac{\partial}{\partial y}(uv), \tag{12.4}$$

where $\zeta = \partial v/\partial x - \partial u/\partial y$ is the vorticity. After zonal averaging (12.4) gives

$$\overline{v'\zeta'} = -\frac{\partial \overline{u'v'}}{\partial y}, \tag{12.5}$$

and (12.3) becomes

$$\frac{\partial \overline{u}}{\partial t} = \overline{v'\zeta'} - r\overline{u}. \tag{12.6}$$

If we integrate the vorticity flux between two quiescent latitudes then, from (12.5), the integral vanishes. Thus, from (12.6), the mean wind, $\overline{u}$, must also vanish after integration over latitude and time.

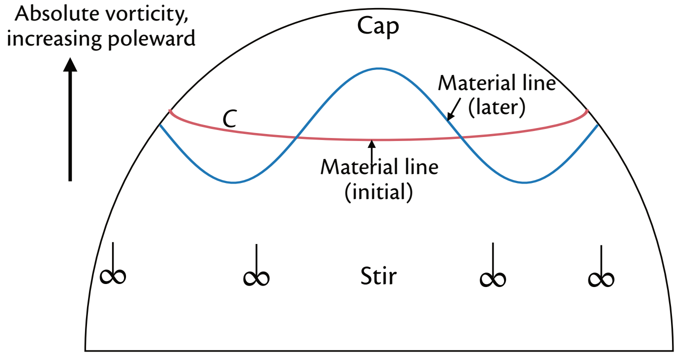

Fig. 12.1: Sketch of the effects of a midlatitude disturbance on the circulation around the latitude line C. If initially the absolute vorticity increases monotonically poleward, then the disturbance will bring fluid with lower absolute vorticity into the cap region. Then, using Stokes theorem, the velocity around the latitude line C will become more westward.

I. The vorticity budget

The argument we first present does not use the momentum equation directly; rather, it uses Kelvin's circulation theorem, and we use spherical coordinates. Suppose that the absolute vorticity normal to the surface , $\zeta + f$, where $f = 2\Omega \sin\vartheta$, increases monotonically poleward. (A sufficient condition for this is that the fluid is at rest.) By Stokes' theorem, the initial circulation, I_i, around a line of latitude circumscribing the polar cap is equal to the integral of the absolute vorticity over the cap. That is,

$$I_i = \int_{\text{cap}} \boldsymbol{\omega}_{ia} \cdot \mathrm{d}\boldsymbol{A} = \oint_C u_{ia}\,\mathrm{d}l = \oint_C (u_i + \Omega a \cos\vartheta)\,\mathrm{d}l, \tag{12.7}$$

where $\boldsymbol{\omega}_{ia}$ and u_{ia} are the initial absolute vorticity and absolute velocity, respectively, u_i is the initial zonal velocity in the Earth's frame of reference, and the line integrals are around the line of latitude. Let us take $u_i = 0$ and suppose there is a disturbance equatorward of the polar cap, and that this results in a distortion of the material line around the latitude circle C (Fig. 12.1).

Since the source of the disturbance is distant from the latitude of interest, if we neglect viscosity the circulation along the material line is conserved, by Kelvin's circulation theorem. Thus, vorticity with a lower value is brought into the region of the polar cap — that is, the region poleward of the latitude line C. Using Stokes' theorem again the circulation around the latitude circle C must therefore fall; that is, denoting later values with a subscript f,

$$I_f = \int_{\text{cap}} \boldsymbol{\omega}_{fa} \cdot \mathrm{d}\boldsymbol{A} < I_i\,, \tag{12.8}$$

so that

$$\oint_C (u_f + \Omega a \cos\vartheta)\,\mathrm{d}l < \oint_C (u_i + \Omega a \cos\vartheta)\,\mathrm{d}l, \tag{12.9}$$

and thus

$$\overline{u}_f < \overline{u}_i, \tag{12.10}$$

with the overbar indicating a zonal average. Thus, there is a tendency to produce *westward* flow poleward of the disturbance. By a similar argument westward flow is also produced equatorward of the disturbance —

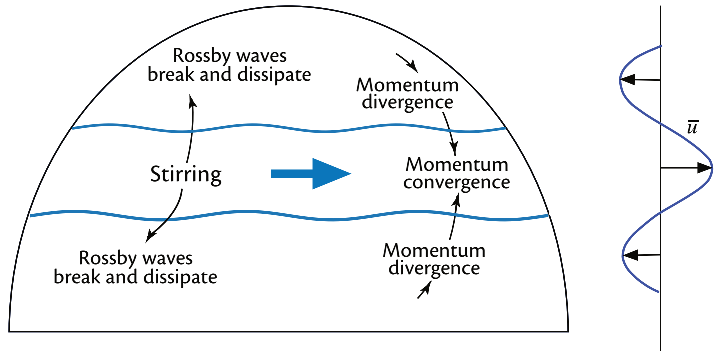

Fig. 12.2: Generation of zonal flow on a rotating sphere. Stirring in midlatitudes (by baroclinic eddies) generates Rossby waves that propagate away. Momentum converges in the region of stirring, producing eastward flow there and weaker westward flow on its flanks.

to see this one may (with care) apply Kelvin's theorem over all of the globe south of the source of the disturbance. Finally, note that the overall situation is the same in the Southern Hemisphere. Thus, on the surface of a rotating sphere, external stirring will produce westward flow *away* from the region of the stirring.

If the disturbance imparts no net angular momentum to the fluid then the integral of $\overline{u}\cos\vartheta$ over the entire hemisphere must be unaltered. But the fluid is accelerating westward away from the disturbance. Therefore, the fluid in the region of the disturbance must accelerate *eastward*, and this is the essence of the production of midlatitude westerlies on Earth, where the stirring is maintained by baroclinic instability.

II. Rossby waves and momentum flux

We have seen that a mean gradient of vorticity is an essential ingredient in the mechanism whereby a mean flow is generated by stirring. Given that, we expect Rossby waves to be excited, and we now show how those waves are intimately related to the momentum flux maintaining the mean flow.

If a stirring is present in midlatitudes then Rossby waves will be generated there before propagating away where they dissipate. To the extent that the waves are quasi-linear, then just away from the source region each wave has the form

$$\psi = \mathrm{Re}\, C e^{\mathrm{i}(kx+ly-\omega t)} = \mathrm{Re}\, C e^{\mathrm{i}(k(x-ct)+ly)}, \qquad (12.11)$$

where C is a constant, with dispersion relation (now back to the β-plane)

$$\omega = ck = \overline{u}k - \frac{\beta k}{k^2 + l^2}, \qquad (12.12)$$

provided that there is no meridional shear in the zonal flow. The meridional component of the group velocity is given by

$$c_g^y = \frac{\partial \omega}{\partial l} = \frac{2\beta k l}{(k^2 + l^2)^2}\,. \qquad (12.13)$$

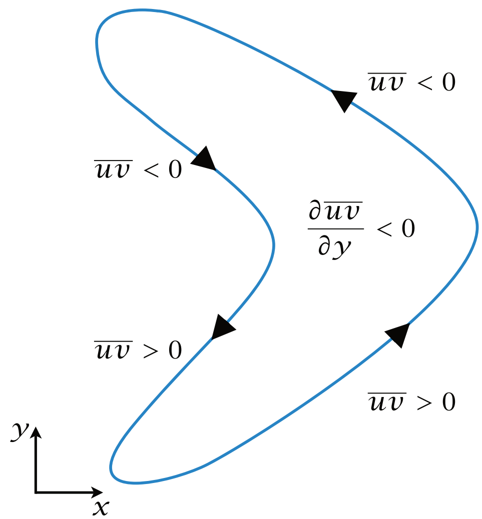

Fig. 12.3: The momentum transport in physical space, caused by the propagation of Rossby waves away from a source in midlatitudes. The ensuing bow-shaped eddies are responsible for a convergence of momentum in the centre of the bow, which in turn accelerates the mean flow eastward. If the arrows were reversed the momentum transport would still have the same sign.

Now, the direction of the group velocity must be away from the source region, because Rossby waves transport energy away from the disturbance. Thus, northward of the source kl is positive and southward of the source kl is negative. That the product kl can be positive or negative arises because for each k there are two possible values of l that satisfy the dispersion relation (12.12), namely

$$l = \pm\left(\frac{\beta}{\overline{u} - c} - k^2\right)^{1/2}, \tag{12.14}$$

assuming that the quantity in parentheses is positive.

The velocity variations associated with the Rossby waves are

$$u' = -\mathrm{Re}\, C\, il e^{i(kx+ly-\omega t)}, \qquad v' = \mathrm{Re}\, C\, ik e^{i(kx+ly-\omega t)}, \tag{12.15a,b}$$

and the associated momentum flux is

$$\overline{u'v'} = -\frac{1}{2}C^2 kl. \tag{12.16}$$

Thus, given that the sign of kl is determined by the group velocity, northward of the source the momentum flux associated with the Rossby waves is southward (i.e., $\overline{u'v'}$ is negative), and southward of the source the momentum flux is northward (i.e., $\overline{u'v'}$ is positive). That is, the momentum flux associated with the Rossby waves is *toward* the source region. Momentum thus converges in the region of the stirring, producing net eastward flow there and westward flow to either side (see Fig. 12.2).

If we think of this effect in physical space, then if kl is positive lines of constant phase ($kx + ly$ = constant) are tilted north-west/south-east, and the momentum flux associated with such a disturbance is negative (that is, $\overline{u'v'} < 0$). Similarly, if kl is negative then the constant-phase lines are tilted north-east/south-west and the associated momentum flux is positive ($\overline{u'v'} > 0$). The net result is a convergence of momentum flux into say. Here the eddy balance is between the Coriolis term and the frictional term, and integrating over this layer and taking the density there to be constant gives

$$-fV \approx -r\overline{u}_s, \tag{12.30}$$

where $V = \int_0^d \overline{v}\,\mathrm{d}z$ is the meridional transport in the boundary layer of height d, above which the stress vanishes. The surface return flow is poleward (i.e., $V > 0$ in the Northern Hemisphere) producing an eastward Coriolis force and an eastward surface flow. In this picture, then, the midlatitude eastward zonal flow at the surface is a consequence of the poleward flowing surface branch of the Ferrel Cell, this poleward flow being required by mass continuity given the equatorward flow in the upper branch of the cell. Seen this way, the Ferrel Cell is responsible for bringing the midlatitude eddy momentum flux convergence to the surface where it may be balanced by friction, as in Fig. 12.7.

A direct way to see that the surface flow must be eastward, given the eddy momentum flux convergence, is to vertically integrate (12.29) from the surface to the top of the atmosphere. By mass conservation, the Coriolis term vanishes (i.e., $\int_0^\infty f\rho\overline{v}\,\mathrm{d}z = 0$) and we obtain

$$\int_0^\infty \frac{\partial}{\partial y}(\overline{u'v'})\rho\,\mathrm{d}z = [\tau]_0^\infty = -r\rho_s\overline{u}_s. \tag{12.31}$$

That is, the surface wind is proportional to the vertically integrated eddy momentum flux convergence. Because there *is* a momentum flux convergence, the left-hand side is negative and the surface winds are positive, or eastward.

12.2.1 The Eulerian Meridional Overturning Circulation

We can obtain an explicit equation for the overturning circulation by combining the momentum equation and the thermodynamic equation using thermal wind balance. Neglecting all but the largest terms, the zonally-averaged zonal momentum equation may be written

$$\frac{\partial\overline{u}}{\partial t} - f\overline{v} = M, \tag{12.32a}$$

where $M = -\partial_y(\overline{u'v'}) + \rho^{-1}\partial\tau/\partial z$ contains the main eddy flux and frictional terms. At a similar level of approximation let us write the thermodynamic equation as

$$\frac{\partial\overline{b}}{\partial t} + N^2\overline{w} = J, \tag{12.32b}$$

where $J = Q_b - \partial_y(\overline{v'b'})$ is the sum of the heating, Q_b, and eddy forcing. We are assuming, as in quasi-geostrophic theory, that the mean stratification, N^2, is fixed and $\overline{b}$ represents only the (zonally averaged) deviations from this. Finally, we use the mass continuity equation to define a meridional streamfunction Ψ; that is

$$\frac{\partial\overline{v}}{\partial y} + \frac{\partial\overline{w}}{\partial z} = 0 \qquad \text{allows} \qquad \overline{w} = \frac{\partial\Psi}{\partial y}, \quad \overline{v} = -\frac{\partial\Psi}{\partial z}. \tag{12.33a,b}$$

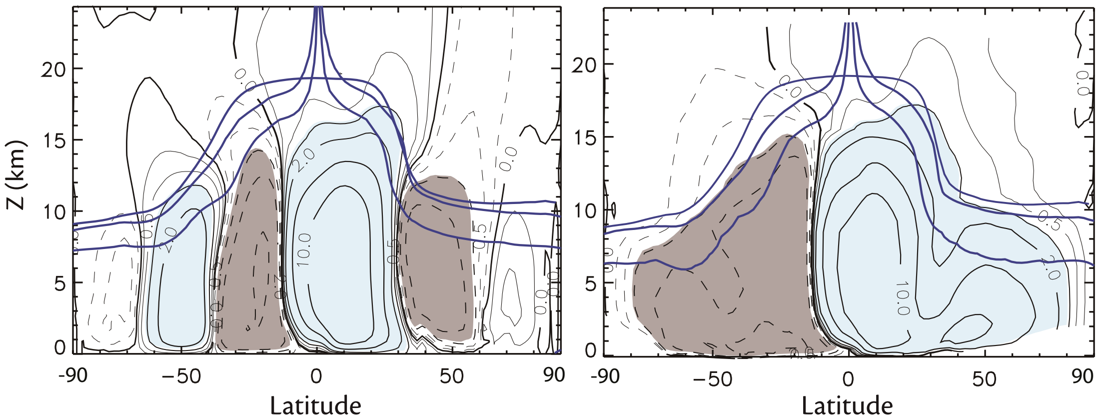

Fig. 12.8: Left: The observed zonally-averaged, Eulerian-mean, streamfunction in Northern Hemisphere winter (DJF, 1994–1997). Negative contours are dashed, and values greater or less than 10^{10} kg s^{-1} (10 Sv) are shaded, darker for negative values. The circulation is clockwise around the lighter shading. The three thick solid lines indicate various measures of the tropopause. Right: The residual meridional mass streamfunction. (Adapted from Juckes (2001).)

In a steady state we have, from (12.32b),

$$w = \frac{1}{N^2}\left[Q_b - \frac{\partial(\overline{v'b'})}{\partial y}\right]. \tag{12.36a}$$

Similarly, from the momentum equation the horizontal velocity and eddy momentum fluxes are related by, in a steady state,

$$-f\overline{v} = -\frac{\partial(\overline{u'v'})}{\partial y} + \frac{1}{\rho}\frac{\partial \tau}{\partial z}. \tag{12.36b}$$

Figure 12.7 (and Fig. 12.4) shows that both eddy heat and momentum fluxes produce an overturning circulation in the same sense as the observed Ferrel Cell. However, these fluxes are not independent of each other: it is a combination of them, and in particular the potential vorticity flux, that is really responsible for the overturning circulation, as we now see.

12.3 ♦ The Residual Ferrel Cell

A revealing way to describe the meridional overturning is by way of the residual circulation, as discussed in Section 9.3. Written in residual form on the f-plane, as in (9.50), the zonal momentum and buoyancy equations are

$$\frac{\partial \overline{u}}{\partial t} - f_0\overline{v}^* = \overline{v'q'} + F_u, \tag{12.37a}$$

$$\frac{\partial \overline{b}}{\partial t} + N^2\overline{w}^* = Q_b. \tag{12.37b}$$

In these equations F_u is a frictional term, Q_b is the heating term and $\overline{v'q'}$ is the eddy potential vorticity flux. The residual velocities, $\overline{v}^*$ and $\overline{w}^*$, are related to their Eulerian counterparts by

$$\overline{v}^* = \overline{v} - \frac{\partial}{\partial z}\left(\frac{1}{N^2}\overline{v'b'}\right), \qquad \overline{w}^* = \overline{w} + \frac{\partial}{\partial y}\left(\frac{1}{N^2}\overline{v'b'}\right). \tag{12.38}$$

Chapter

13 Planetary Atmospheres

EARTH IS BUT ONE PLANET. It is, at least for us humans, the most important and most interesting one, but there are very many others. There are at least seven other planets in the Solar System, all of them with atmospheres, and some of those planets have moons with atmospheres — Titan (orbiting Saturn) and Io (orbiting Jupiter) are two. At the time of writing we have observed about 4000 planets, and over 500 multiplanet solar systems, outside the Solar System but within our galaxy. With a little extrapolation we can estimate there are billions (yes, billions) of planets in our galaxy alone, and there are billions of other galaxies in the Universe. Many of these planets will have atmospheres, and some undoubtedly have oceans. And some, almost certainly, have life.

In this chapter our task is to apply geophysical fluid dynamical principles to these planetary atmospheres and thereby try to understand their circulation. The task is a hard one because the variety of planets is enormous — they differ from each other in their mass and composition, their emitting temperature, their size and rotation, whether they are terrestrial or gas giants (terms we define later) or something else entirely, and in a host of other parameters. There is much greater variety in planetary atmospheres than in the stars they orbit, and there can be no single theory of their circulation, no planetary equivalent of astronomy's main sequence of stars that shows the relation between stellar luminosity and effective temperature. On the other hand, the basic principles we have learned in earlier chapters apply to all planetary atmospheres, so we should not be engaged in describing planets one by one (although Earth is a special case). Indeed, because there are so many planets, we *must* look for general principles where we can, else we are hardly doing science at all.

In the sections that follow we aim to give a coherent but introductory treatment of these atmospheres, to put them into context and see how and where Earth's atmosphere might fit into the set of all planetary atmospheres. We begin with a descriptive taxonomy.

13.1 A Taxonomy of Planets

The formation and evolution of planetary atmospheres is a subject unto itself which we won't delve into, and here we give just a brief descriptive overview of some of the more common types. These types are not all orthogonal and a given planet may belong in two or more categories, and the definitions themselves are of disputed authority and subject to debate, and may well evolve over the years ahead. Readers of this book 20 years hence may well read this section with a knowing smile.

Capitalization by example
The Solar System is the system of bodies — Earth, the other planets, minor planets and so on — orbiting the Sun, as well as the Sun itself. Many planets have earth at their surface, but there is only one Earth and the Moon orbits around it. Earth has only one moon whereas Mars has two, Phobos and Deimos. Other solar systems exist in our galaxy, the Milky Way, and planets in these solar systems orbit around their own suns. There are many galaxies in the Universe and there may even be many universes.

Planets. A planet is defined to be a body that orbits its host star directly and is massive enough to be in hydrostatic equilibrium (effectively meaning it has formed under its own gravitational force and has a spheroidal shape) and to dominate its own orbit, clearing it of other bodies. In so far as this definition is official (i.e., as stated by the International Astronomical Union (IAU) in a statement in 2006) it does not apply to bodies in other solar systems, but the definition may usefully be taken to apply more generally. (Many scientists believe the requirement of clearing the orbit is also too restrictive, in which case dwarf planets, defined below, are also planets.) The planets in the Solar System are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune.

Dwarf planets are bodies that are large enough to form under their own gravity but are not able to clear their orbit of other bodies. Pluto is the most famous example, and Eris (in an orbit beyond Neptune) and Ceres (in the asteroid belt) are two others. More generally, *minor planets* are objects that orbit around a star that need not have formed under their own gravity, including dwarf planets and asteroids but excluding true planets and comets. There are hundreds of thousands of minor planets in the Solar System. The IAU also defines *small Solar System bodies* to be objects, including asteroids and comets, that directly orbit the Sun that are too small to be planets or dwarf planets.

Planetary bodies. This is a general term for objects that have formed under their own gravity, including objects that are not big enough to be true planets (under the definition above) and objects that are not in direct orbit around their host star, but excluding stars themselves. The category thus includes all the Solar System planets, the dwarf planets such as Pluto, large natural satellites such as Titan (in orbit around Saturn and with a thick methane atmosphere), Triton (in orbit around Neptune and with a thin, nitrogen atmosphere), and exoplanets.

Exoplanets are planets in other solar systems. They must be big enough to form under gravity but not so big as to form stars, and so are generally under the gravitational influence of a host star. The definition of exoplanets is generally taken to be looser than that for planets in our own solar system, and might include dwarf planets (if any were to be discovered) but would normally exclude comets and asteroids. A *rogue planet* is a planetary body that orbits the galactic centre and not a particular star.

Terrestrial planets. These are planetary bodies that, like Earth, have an atmosphere with a distinct lower boundary, often a rocky surface but also possibly an ocean or other distinct change of character. Other examples include Mars, Venus and Mercury, although Mercury's atmosphere is extremely thin. Some sources restrict the definition to planets of similar size to Earth, in which case much larger or smaller planets that are otherwise similar might be called *quasi-terrestrial*. On the other hand, the term terrestrial planet is often applied to objects that are not, by the IAU definition, planets, such as Pluto and Titan.

When discussing other planets a 'day', without a qualifying adjective, still refers to 86,400 seconds, and a 'year' refers to 365 days. One may of course refer to such things as 'Venusian day' or a 'Jovian year', in which case the terms apply to the planet in question.

Giant planets. A giant planet may be defined as any planet at least ten times more massive than Earth (although other definitions may differ slightly), including ice giants, gas giants (both defined below) and massive terrestrial planets.

Gas giants. These are giant planets, like Jupiter and Saturn, that are composed mainly of hydrogen and helium and that do not have a sharp interface between atmosphere and solid planet. Jupiter, for example, most likely has an outer layer of molecular hydrogen, an inner layer of metallic hydrogen and a molten rocky core, and is 300 times more massive than Earth. The outer layer contains water and other heavier compounds but, in general, gas giants are more than 90% hydrogen and helium, although not all of it is gaseous: much of the hydrogen may be in liquid form.

Ice giants. Giant planets that are composed of elements heavier than hydrogen and helium are called ice giants, although the name is something of a misnomer. Uranus and Neptune (which are about 15 times more massive than Earth) are both ice giants and have less than 20% hydrogen and helium, the rest being such elements as oxygen, nitrogen and carbonic compounds such as ammonia and methane.

Super-Earths. A super-Earth is a planet with a mass between that of the Earth and that of a giant planet. A super-Earth might in principle be a terrestrial or a gaseous planet, and occasionally they are called 'mini-Neptunes'. There are no super-Earths in the Solar System — the outer planets, Jupiter, Saturn, Uranus and Neptune, are all 'giants'.

Hot Jupiters. These are gas-giant exoplanets that are in close proximity to their host stars and that may be tidally locked (with one side permanently facing their sun), and so can be expected to have very high surface temperatures on one side, low on the other. Their orbital period may be of order tens of Earth days (which is very short compared to Earth), but this may also be their rotation period around their own axis (which would be very long compared to Earth). *Hot Neptunes* is the analogous name for somewhat less massive planets, of Neptune size, that may also be in close orbit around their star.

Brown dwarfs. These are hybrids between small stars and gas-giant planets, and may at some stage have undergone nuclear fusion and have elements much heavier than hydrogen and helium. We won't consider them further in this chapter.

Others. Various other planet types exist or are hypothesized to exist. For example, the unpronounceable *chthonian planets* are gas giants that are losing or have lost their outer layers and, still more exotically, *lava planets* are planetary bodies with a surface covered by molten lava. And so on.

We cannot hope to provide theories for all of these types of planets or the atmospheres they may contain; rather we will first focus on planets with a shallow, ideal-gas atmosphere and consider the effects of a few key parameters. In this context 'shallow' means that the depth of the atmosphere is a small fraction of the planetary radius in which there may be a 'weather layer' similar to the atmosphere of a terrestrial planet. Later in the chapter we will consider the dynamics of the deeper atmosphere immediately beneath the weather layers that might give rise to such things as the jets on Jupiter and Saturn.

13.2 Dimensional and Nondimensional Parameters

Consider a terrestrial planet with an atmosphere that obeys the primitive equations. It is forced by incoming solar ('shortwave') radiation which is balanced by outgoing infra-red ('longwave') radiation. There is (if the obliquity is low) more incoming radiation near the equator, and in many atmospheres much of the shortwave radiation is absorbed at the surface, and we represent this thermal forcing by a relaxation to a specified temperature that decreases with latitude and height. There is no external forcing in the momentum equation, aside from the effects of gravity in the vertical, but momentum is dissipated by the effects of friction near the surface, and we may represent this by the effects of a linear drag. If we use pressure coordinates then the equations of motion may be written,

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \omega\frac{\partial \boldsymbol{u}}{\partial p} + \boldsymbol{f}\times\boldsymbol{u} = \nabla\phi - r\boldsymbol{u}, \tag{13.1a}$$

$$\frac{\partial \phi}{\partial p} = -\frac{RT}{p} = -\frac{R\theta}{p_R}\left(\frac{p_R}{p}\right)^{c_v/c_p}, \tag{13.1b}$$

$$\nabla\cdot\boldsymbol{u} + \frac{\partial \omega}{\partial p} = 0, \tag{13.1c}$$

$$\frac{\partial \theta}{\partial t} + \boldsymbol{u}\cdot\nabla\theta + \omega\frac{\partial \theta}{\partial p} = Q_{[\theta]}, \tag{13.1d}$$

where $\boldsymbol{u}$ is the horizontal velocity and the ∇ operator is taken to be horizontal, meaning at constant pressure and we omit viscous and diffusion terms. The term $-r\boldsymbol{u}$ parameterizes surface drag and r is large only very close to the surface, and is negligible in the free atmosphere. The term $Q_{[\theta]}$ represents heating due to radiative forcing, and θ is the potential temperature, $\theta = T(p_R/p)^\kappa$ where p_R is the reference pressure, which we take to be the mean surface pressure, p_s, and $\kappa = R/c_p$.

For the purposes of nondimensionalization the radiative forcing is taken to be given by a relaxation to a radiative-equilibrium temperature field, θ^*, such that

$$Q_{[\theta]} = \frac{\theta^*(\vartheta, p) - \theta(\vartheta, p)}{\tau}, \tag{13.2}$$

where θ^* is a function of both latitude and height and τ is a relaxation timescale. On Earth, τ is about 10 days, and θ^* might vary by about 60 K from equator to pole and about 50 K from surface to the tropopause, and we denote these values θ_H and θ_V respectively. It is these variations that produce the circulation. For example (although the exact form is not of particular concern here) a possible recipe for θ^* is

$$\theta^* = \overline{\theta}\left[1 + \frac{\Delta_H}{3}(1 - 3\sin^2\vartheta) + \Delta_V Z\right]. \tag{13.3}$$

Here, $\overline{\theta}$ is the average surface temperature, Δ_H is a nondimensional parameter that determines the equator to pole temperature difference, Δ_V is a similar parameter for the vertical, and $Z = -\log(p/p_s)$. We then have $\theta_H = \overline{\theta}\,\Delta_H$ and $\theta_V = \overline{\theta}\,\Delta_V$.

We nondimensionalize (13.1) by writing

$$(\widehat{u}, \widehat{v}) = \frac{(u, v)}{U}, \quad (\widehat{x}, \widehat{y}) = \frac{(x, y)}{a}, \quad \widehat{\omega} = \frac{\omega a}{U p_s}, \quad \widehat{p} = \frac{p}{p_s}$$
$$\widehat{\theta} = \frac{\theta}{\overline{\theta}}, \quad \hat{t} = \frac{t}{T} \quad \widehat{\phi} = \frac{\phi}{\Phi}. \tag{13.4}$$

Here, as usual, the hats denote nondimensional quantities and T, U and Φ denote scaling values for time, horizontal velocity, and pressure, and we scale temperature and potential temperature with $\overline{\theta}$. To scale time we use the planetary rotation rate, $T = 1/\Omega$, and to scale velocity we use thermal wind balance, based on the radiative equilibrium temperature difference between equator and pole, as follows. The thermal wind relation for the zonal wind is

$$f\frac{\partial u}{\partial p} = -\frac{R}{pa}\frac{\partial T}{\partial \vartheta}, \tag{13.5}$$

where R is the gas constant. This suggests the scaling

$$U = \frac{R\theta_H}{\Omega a}. \tag{13.6}$$

The usual Rossby number is defined as $Ro = U/fL$. If we use (13.6) we can by analogy define the *external Rossby number, Ro_E,* also called the *thermal Rossby number,* by

$$Ro_E = \frac{U}{\Omega a} = \frac{R\theta_H}{\Omega^2 a^2}. \tag{13.7}$$

Unlike the Rossby number itself, this is an external parameter of the system and not an emergent property of the flow itself. Finally, for the nondimensionalization of the geopotential we use

$$\Phi = (\Omega a) \times U = R\theta_H, \tag{13.8}$$

which is analogous to the geostrophic scaling estimate $\phi \sim fUL$ encountered in Section 5.3.1.

The external Rossby number is one of the most important nondimensional numbers affecting the behaviour of a planet's atmosphere. Similar to the conventional Rossby number, U/fL, it is a measure of the importance of rotation, but now we take the velocity to be determined by a thermal wind determined by the radiative forcing of the planet, and the length scale to be the radius of the planet itself. The conventional Rossby number may be called the 'internal Rossby number' in this context.

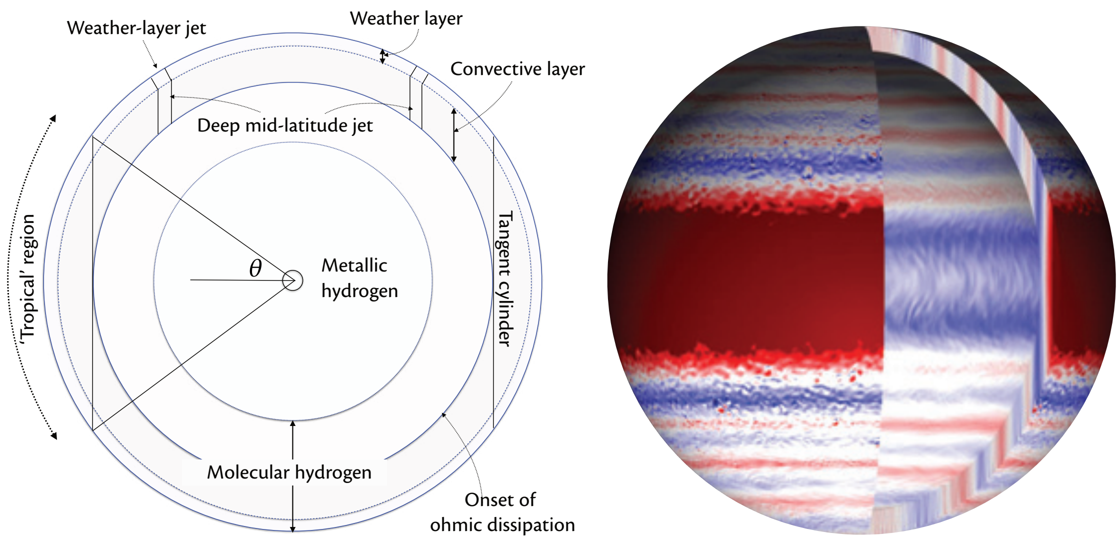

Fig. 13.12: Left: Sketch of the potential structure of Jupiter's atmosphere (not to scale) with the jets in a convective layer between a layer of ohmic dissipation and the very thin weather layer. Right: The zonal velocity on Jupiter obtained from a numerical simulation (from Heimpel *et al.* 2016) of a spherical shell with inner radius equal to 0.9 of the planetary radius, giving $\theta = 23°$. Red colours denote eastward flow and blue colours westward.

there might be about 10 jets between equator and pole. This is a little larger than the number observed (about 6) but the agreement is as good as can be expected given the nature of the scaling argument.

It is, however, quite possible that the jets extend a few thousand kilometres, and perhaps considerably more, into Jupiter's interior, as shown in Fig. 13.12. In this picture, a significant source of energy is the heat emanating from the planetary interior, and the ensuing convection creates a neutrally stratified region extending up to the weather layer, where solar absorption — and possibly water vapour and baroclinic instability — stabilize the fluid to dry convection. How deep the jets go is not known with any certainty. It was once thought that they should go down as far as the metallic core, but other ideas posit that 'ohmic dissipation', which arises because of the finite electrical conductivity of the molecular hydrogen, acts much closer to the surface than the metallic layer and may prevent the jets extending much deeper than a few thousand kilometres, but the exact depth at which this dissipation becomes significant remains uncertain. Assuming this deeper layer (the 'convective layer') in Fig. 13.9 does exist then jets will form within it, as we now discuss.

Jets in a deep atmosphere

If the jets on Jupiter do descend just a few thousand kilometres into the interior then they may be regarded as deep from the point of view of a meteorologist interested in the weather layer, but shallow from the point of a scientist studying planetary interiors!

Consider the schematic in the left panel in Fig. 13.12. We will suppose that the planetary radius is a and that there is a convective layer between some inner shell at a depth d and the outer radius (where the very thin weather layer resides). This convective layer may be naturally divided into three regions, one equatorward of the intersection of the tangent cylinder with the outer radius (and so with latitude less than θ) and denoted the 'tropical' region in Fig. 13.12, and two regions poleward of that, one in either hemisphere. Simple geometry indicates that the angle θ is given by $\theta = \cos^{-1}(a-d)/a$, with $\theta = 15°$ corresponding to $d = 2,400\,\mathrm{km}$. If on the other hand we were to take the inner radius to that of the transition to metallic hydrogen, and so with $d \approx 15,000$ km, then $\theta \approx 40°$. Let us make a simple model of these convective regions to illustrate how jets form within them, with a super-rotating jet in the tropical region.

Perhaps to a greater extent than elsewhere in the book, here we are making a somewhat speculative model of the phenomenon, not a quantitative or exact theory.

We will model the convective region as a layer of shallow water, obeying the potential vorticity equation

$$\frac{DQ}{Dt} = 0, \qquad Q = \left(\frac{\zeta + 2\Omega}{h}\right). \tag{13.33}$$

Here Q is the potential vorticity, Ω is the rotation rate of the planet (a constant) and ζ is the vorticity aligned with the planetary rotation (and not with the radial direction which would be conventional in a shallow atmosphere). The quantity h is the thickness of the convecting layer and we write this as $h = H + h'$, where H is the mean shell thickness and h' are small, time-dependent, deviations of that due to fluid motion, and $H \gg h'$. From Fig. 13.13 (and Fig. 13.12) we see that H varies in the y direction, decreasing toward the pole in the region poleward of the intersection with the tangent cylinder, but decreasing toward the equator in the region equatorward of the intersection with the tangent cylinder. It is this variation with mean thickness, and hence the variation of the background potential vorticity, that gives rise to a 'topographic beta effect' and hence to zonal jets. To see this explicitly, we make two more assumptions:

(i) The small Rossby number assumption, that $|2\Omega| \gg |\zeta|$.

(ii) The variations in mean height occur on a larger scale than the variations in vorticity.

The potential vorticity is then given by

$$Q = \left(\frac{\zeta + 2\Omega}{h}\right) \approx \left(\frac{\zeta + 2\Omega}{H}\right), \tag{13.34}$$

and, using the assumptions above, its evolution is given by

$$\frac{DQ}{Dt} \approx \frac{1}{H}\frac{D\zeta}{Dt} + 2\Omega \frac{D}{Dt}\left(\frac{1}{H}\right) = \frac{1}{H}\frac{D\zeta}{Dt} - \frac{2\Omega}{H^2}\boldsymbol{v}\cdot\nabla H, \tag{13.35}$$

and (13.33) becomes

$$\frac{D\zeta}{Dt} + \beta^* v = 0 \qquad \text{where} \qquad \beta^* = -\frac{2\Omega}{H}\frac{\partial H}{\partial y}, \tag{13.36a,b}$$

where v is the velocity in the y-direction. We see from Fig. 13.13 that β^* is positive in the region insider the tangent cylinder (the extra-tropics) and negative outside the tangent cylinder, in the tropics.

The planetary rotation organizes the convection below the weather layer into alternating zonal jets because of the topographic beta effect that arises due to the variations in thickness of the convecting layer. The convecting layer effectively forms a lower boundary condition for the weather layer, but the interaction of the two layers is far from understood and the jets in the weather layer might not even have a one-to-one correspondence with those in the convective layer.

Equation (13.36) is very similar to the familiar barotropic vorticity equation on the β-plane — compare it with (6.23) or (10.63). Thus, if the flow is turbulent, we may expect alternating zonal jets to form because of the interaction of Rossby waves with the eddying flow. The intensity of these jets, and the spacing between them, depends on the size of the turbulent flow produced by the convection, and that in turn depends on the heat flux coming up from the planetary interior and the viscosity. The value of β^* is actually similar to the value of β itself, because both are a consequence of the sphericity of the planet.

Part III

Oceans

CHAPTER

14

Wind-Driven Gyres

WE NOW START OUR VOYAGE into that other great fluid covering the Earth, the ocean, and we divide the voyage into three legs. In this first one we look at the essentially horizontal circulation that gives rise to the great gyres in the mid- and high latitudes. In the next chapter we look at the processes giving rise to the vertical structure of the ocean and the meridional overturning circulation, and in the third chapter we look at equatorial circulation and El Niño. Let us first take a brief look at the observations to see what we have to understand.

14.1 AN OBSERVATIONAL OVERVIEW

The aspect of the ocean that most affects the climate is the sea-surface temperature (SST), illustrated in Fig. 14.1. Aside from the to-be-expected latitudinal variation there is significant zonal variation — the western tropical Pacific is particularly warm, and the western Atlantic is warmer than the corresponding latitude in the east. These variations owe their existence to ocean currents, and the vertically averaged currents of the North Atlantic are illustrated in Fig. 14.2. The most striking features are the two main gyres — the clockwise, and anticyclonic, subtropical gyre between about 25°N and 50°N, and the anti-clockwise, and cyclonic, subpolar gyre north of that. We can see that these gyres are intensified in the west; the intensification is most obvious in the subtropical gyre, where the intense northward flowing current is known as the Gulf Stream, but is also present in the subpolar gyre.

The same features are present in *all* of the main basins of the world's ocean, as we see in Fig. 14.3, in both Northern and Southern Hemispheres. The western boundary current of the great subtropical gyre in the North Pacific, flowing northward off the coast of Japan, is known as the Kuroshio, and similar currents flow southward along the west coast of Australia and the west coast of Brazil and Argentina in the Southern

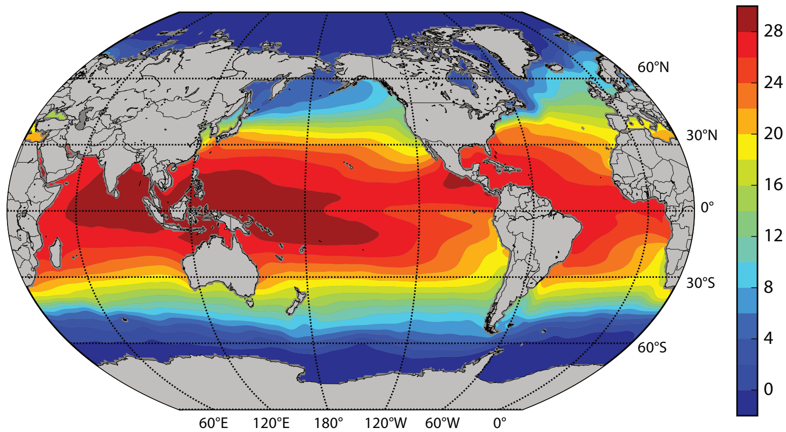

Fig. 14.1: The sea-surface temperature (SST, °C) of the world's ocean, as determined from a great many observations, combined in the World Ocean Circulation Experiment (WOCE).

Hemisphere. The existence of the great gyres, and that they are strongest in the west, has been known for centuries, and our main task in this chapter is to explain that. It turns out to be a much easier task to explain the vertically-integrated flow than the vertical structure of the flow and that is our focus.

14.2 SVERDRUP BALANCE

Harald Sverdrup (1888–1957) was a Norwegian meteorologist/oceanographer who is most famous for the balance that now bears his name, but he also played a leadership role in scientific policy and was the director of Scripps Institution of Oceanography from 1936–1948.

Let us begin by considering an ocean, forced by a wind stress, $\boldsymbol{\tau}_0 = \tau_0^x \hat{\mathbf{i}} + \tau_0^y \hat{\mathbf{j}}$, at the top, that satisfies the equations

$$-fv = -\frac{\partial \phi}{\partial x} + \frac{1}{\rho_0}\frac{\partial \tau^x}{\partial z}, \qquad fu = -\frac{\partial \phi}{\partial y} + \frac{1}{\rho_0}\frac{\partial \tau^y}{\partial z}, \qquad (14.1\text{a,b})$$

and $\tau_x \hat{\mathbf{i}} + \tau_y \hat{\mathbf{j}} = \boldsymbol{\tau}$ is the stress acting on the fluid. Since the ocean's density is very nearly constant we absorb the quantity $1/\rho_0$ into the definition of stress (the quantities $(\tau_x, \tau_y)/\rho_0$ are the 'kinematic stress' but are commonly, if a little loosely, just referred to as the stress). With this new definition of stress we rewrite (14.1) as

$$f(v_g - v) = \frac{\partial \tau^x}{\partial z}, \qquad f(u - u_g) = \frac{\partial \tau^y}{\partial z}, \qquad (14.2)$$

where (u_g, v_g) are the geostrophic velocities given by $f(u_g, v_g) = (-\partial\phi/\partial y, \partial\phi/\partial x)$. The left-hand side is just the ageostrophic velocity, and if we integrate vertically from the top of the ocean to the base of the Ekman layer, where the stress is by definition zero, we obtain

$$fV_a = -\tau_0^x, \qquad fU_a = \tau_0^y, \qquad \text{or} \qquad f\boldsymbol{U}_a = \hat{\mathbf{k}} \times \boldsymbol{\tau}_0, \qquad (14.3)$$

where $U_a = \int_{\mathrm{Ek}} (u - u_g)\,\mathrm{d}z$ is the integral of the ageostrophic velocity over the Ekman layer, and similarly for V_a, and $\boldsymbol{U}_a = U_a \hat{\mathbf{i}} + V_a \hat{\mathbf{j}}$. Evidently the ageostrophic Ekman transport is at right angles to the surface stress, and in the ocean the Ekman layer is of order tens of metres thick.

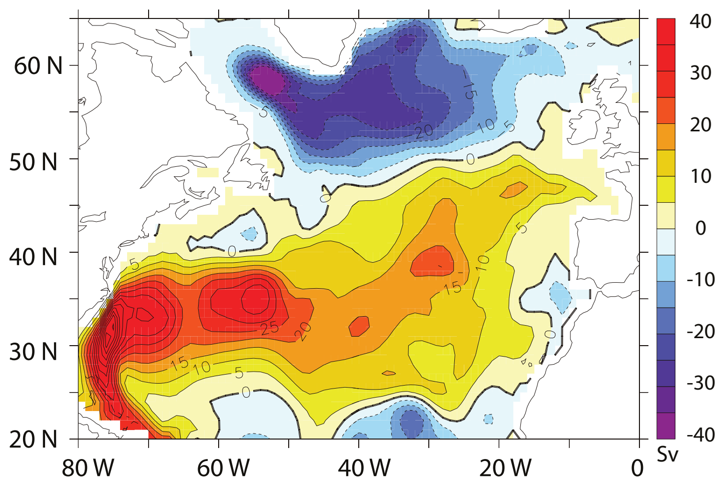

Fig. 14.2: The streamfunction of the vertically averaged flow in the North Atlantic, obtained by constraining a numerical model to observations so giving a 'state estimate'. Red shading indicates clockwise flow, and blue shading anticlockwise. (Courtesy of Rong Zhang using a GFDL model and climatologcal data.)

There is one particularly useful result we can obtain from (14.1). If we cross differentiate and use the mass conservation equation, $\partial u/\partial x + \partial v/\partial y + \partial w/\partial z = 0$, we obtain

$$f\frac{\partial w}{\partial z} + \beta v = \frac{\partial \tau^y}{\partial x} - \frac{\partial \tau^x}{\partial y}. \tag{14.4}$$

Now integrate from the top of the ocean (where $w = 0$) down to some level, z, below the base of the Ekman layer where the stress is zero, to obtain

$$w(z) + \int_z^0 \frac{\beta}{f} v \, \mathrm{d}z' = \frac{1}{f}\left[\frac{\partial \tau_0^y}{\partial x} - \frac{\partial \tau_0^x}{\partial y}\right]. \tag{14.5}$$

If we let the integral go over the entire depth of the ocean, and assume that the vertical velocity and the stress are zero at the ocean bottom, we obtain

$$\int \beta v \, \mathrm{d}z = \frac{\partial \tau_0^y}{\partial x} - \frac{\partial \tau_0^x}{\partial y}. \tag{14.6}$$

This expression is known as the *Sverdrup relation*. It is remarkable because it tells us that, at any location in the ocean, *the vertically integrated meridional velocity is given by the curl of the wind stress at the surface.* Although there are a number of caveats to this statement (as our assumptions are not exactly satisfied), the Sverdrup relation is one of the enduring foundations of physical oceanography.

14.3 Ocean Gyres

The equations of motion that govern the three-dimensional, large-scale flow in the oceans are the planetary-geostrophic equations, discussed in Chapter 5, namely

$$\frac{\mathrm{D}b}{\mathrm{D}t} = \dot{b}, \qquad \nabla_3 \cdot \boldsymbol{v} = 0, \tag{14.7a,b}$$

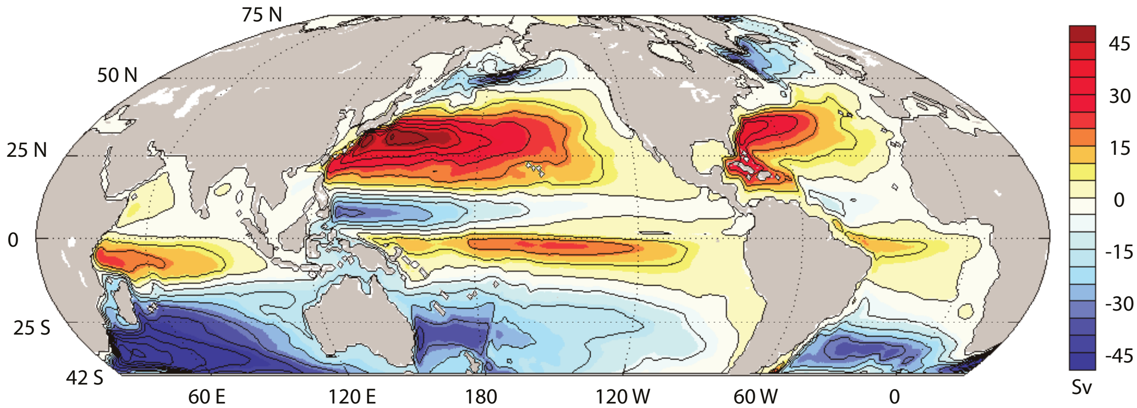

Fig. 14.3: A state estimate of the streamfunction of the vertically integrated flow for the near global ocean. Red shading indicates clockwise flow, and blue shading anticlockwise, but in both hemispheres the subtropical gyres are anticyclonic and the subpolar ones are cyclonic. (Courtesy of Patrick Heimbach, using the ECCO system.)

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi + \frac{\partial \boldsymbol{\tau}}{\partial z}, \qquad \frac{\partial \phi}{\partial z} = b. \tag{14.8a,b}$$

These equations are, respectively, the thermodynamic equation (14.7a), the mass continuity equation (14.7b), the horizontal momentum equation (14.8a), (i.e., geostrophic balance, plus a stress term), and the vertical momentum equation (14.8b) — that is, hydrostatic balance. The gradient and divergence operators are two dimensional, in the x–y plane, unless noted with a subscript 3. Simple as they may be compared to the full Navier–Stokes equations, the equations are still quite daunting: a prognostic equation for buoyancy is coupled to the advecting velocity via hydrostatic and geostrophic balance, and the resulting problem is quite nonlinear. However, it turns out that thermodynamic effects can effectively be eliminated by the simple device of vertical integration; the resulting equations are linear, and the only external forcing is that due to the wind stress. This device enables us to construct a rather simple but very revealing model of the ocean circulation, as follows.

The model of ocean gyres described here is due to Henry Stommel (1920–1992), one of the most creative physical oceanographers of the twentieth century. He spent most of his career at the Woods Hole Oceanographic Institution (WHOI) and played a major role in making physical oceanography a quantitative scientific endeavour. His forté was the construction and use of models that were both simple and relevant, and the observational and experimental testing of such models.

14.3.1 The Stommel Model

Take the curl of (14.8a) (that is, cross-differentiate its x and y components) and integrate over the depth of the ocean to give

$$\int \boldsymbol{f}\nabla\cdot\boldsymbol{u}\,\mathrm{d}z + \frac{\partial f}{\partial y}\int v\,\mathrm{d}z = \mathrm{curl}_z(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B), \tag{14.9}$$

where the operator curl_z is defined by $\mathrm{curl}_z\boldsymbol{A} \equiv \partial A^y/\partial x - \partial A^x/\partial y = \hat{\mathbf{k}}\cdot\nabla\times\boldsymbol{A}$, and the subscripts T and B are for top and bottom; the stress at the bottom, although small, must be retained to find a solution, as we will discover. Equation (14.9) then becomes

$$\beta V = \mathrm{curl}_z(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B), \tag{14.10}$$

where V is the vertical integral of v over the entire depth of the ocean (and similarly for U later on). Evidently, the thermodynamic fields do not affect the vertically integrated flow.

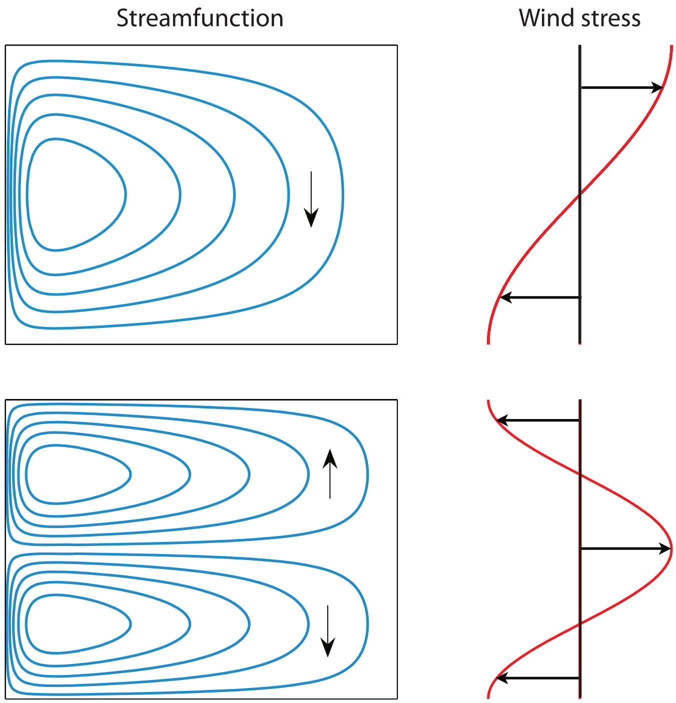

Fig. 14.5: Two solutions of the Stommel model. Upper panel shows the streamfunction of a single-gyre solution, with a wind stress proportional to $-\cos(\pi y/a)$ (in a domain of side a), and the lower panel shows a two-gyre solution, with wind stress proportional to $\cos(2\pi y/a)$. In both cases $\epsilon_S = 0.04$.

14.3.3 The Munk Problem: Using Viscosity Instead of Drag

A natural variation on the Stommel problem is to use a harmonic viscosity, $\nu\nabla^2\zeta$, in place of the drag term $-r\zeta$ in the vorticity equation, the argument being that the wind-driven circulation does not reach all the way to the ocean bottom so that an Ekman drag is not appropriate. This variation is called the 'Munk problem' or 'Munk model'. The problem is to find and understand the solution to the (dimensional) equation

$$\beta\frac{\partial\psi}{\partial x} = \mathrm{curl}_z\boldsymbol{\tau}_T + \nu\nabla^2\zeta = \mathrm{curl}_z\boldsymbol{\tau}_T + \nu\nabla^4\psi \tag{14.36}$$

in a given domain, for example a square of side a. The nondimensional version of this is

$$-\epsilon_M\nabla^4\widehat{\psi} + \frac{\partial\widehat{\psi}}{\partial\widehat{x}} = \mathrm{curl}_z\widehat{\boldsymbol{\tau}}_T, \tag{14.37}$$

where $\epsilon_M = (\nu/\beta a^3)$.

Walter Munk (1917–) is a Viennese-born American physical oceanographer who spent most of his career at Scrippts Institution of Oceanography. He has made important contributions to a host of problems in oceanography, especially in the areas of waves and tides.

Because the equation is of higher order we need two boundary conditions at each wall to solve the problem uniquely, and as before for one of them we choose $\psi = 0$ to satisfy the no-normal-flow condition. For the other condition it is common to use a no-slip condition; that is $\psi_n = 0$ where the subscript denotes the normal derivative of the streamfunction, so that, for example, at $x = 0$ and $x = a$ we have $v = 0$. As with the Stommel problem the solution may be found by boundary-layer methods, and

Chapter

15

The Overturning Circulation and Thermocline

In the previous chapter we studied the horizontal, vertically integrated, flow of the world's oceans. In this chapter we look at the vertical structure of the oceans and the *meridional overturning circulation* (MOC), which is the circulation in the vertical–meridional plane.

15.1 The Observations

Our main goals in this chapter are to explain two important phenomena:

(i) The structure of the temperature and density of the ocean in the vertical–meridional plane;

(ii) The circulation of the ocean in that same plane.

As one might expect it is much harder to observe the interior of the ocean than the surface ocean, or the atmosphere. Because water is almost opaque to electromagnetic radiation we actually have to drop instruments into the ocean to measure its deep properties. These days measurements come from a combination of moored instruments, hydrographic surveys, floats, gliders and satellites (which mostly measure surface properties). The various measurements are combined in some fashion (often in combination with a numerical model) to give a 'state estimate' of the ocean, and we now have a decent coarse-grained view of the density structure and circulation of the sub-surface ocean, although with far less detail than our view of the atmosphere.

15.1.1 The Thermocline

The density structure of the Atlantic Ocean (and the Pacific is similar) is illustrated in Fig. 15.1. Here we see that the main gradients of density are concentrated in the upper one kilometre or so of the ocean, in the *main thermocline,* which serves to connect the relatively warm surface waters with the much colder abyssal waters. (The main thermocline exists year

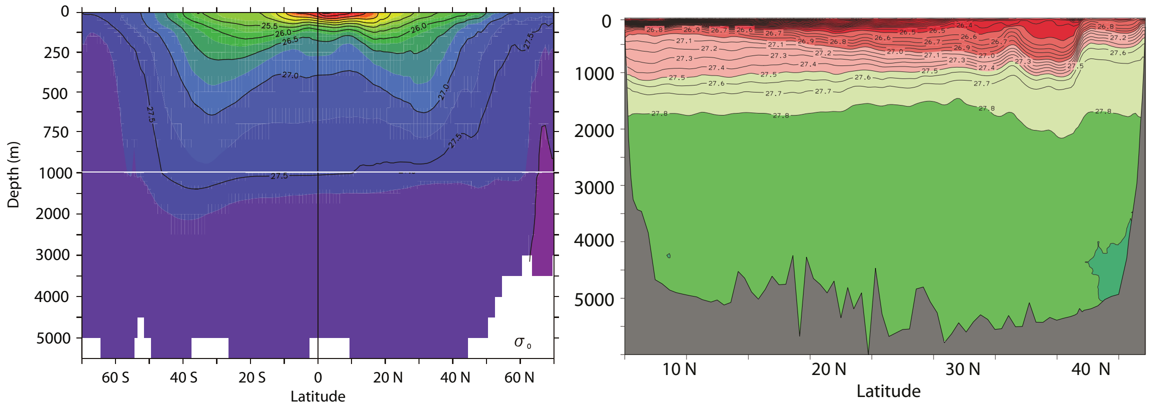

Fig. 15.1: The potential density in the Atlantic ocean. On the left is the climatological zonally-averaged field, plotted with a break in the vertical scale at 1000 m. On the right is a section at 53°W. Both plots show a region of rapid change of density (and temperature) concentrated in the upper kilometre, in the *main thermocline*, below which the density is much more uniform.

round; the seasonal thermocline, which is not visible in these plots, is a much shallower region near the surface over which the temperature gradient varies seasonally.) The thermocline is much weaker at high latitudes, since the near-surface waters are already cold, and it is shallower at low latitudes, as we see in Fig. 15.2. The abyssal temperature at all latitudes is about 2°C, which is similar to the surface temperature at high latitudes, and this is consistent with water at high latitudes sinking, spreading equatorward and filling the abyss.

15.1.2 The Meridional Overturning Circulation

Closely associated with the density structure of the ocean is the meridional overturning circulation, or MOC, and this is illustrated in Fig. 15.3. Focusing on the red, northern cell we see water sinking at high latitudes, spreading south at depth, and upwelling largely in the Southern Ocean; the water in this cell is called *North Atlantic Deep Water*, or NADW. The blue cell shows water sinking at high southern latitudes and spreading north underneath the NADW before rising to mid-depth and returning; this cell contains *Antarctic Bottom Water*, or AABW. The Pacific Ocean has

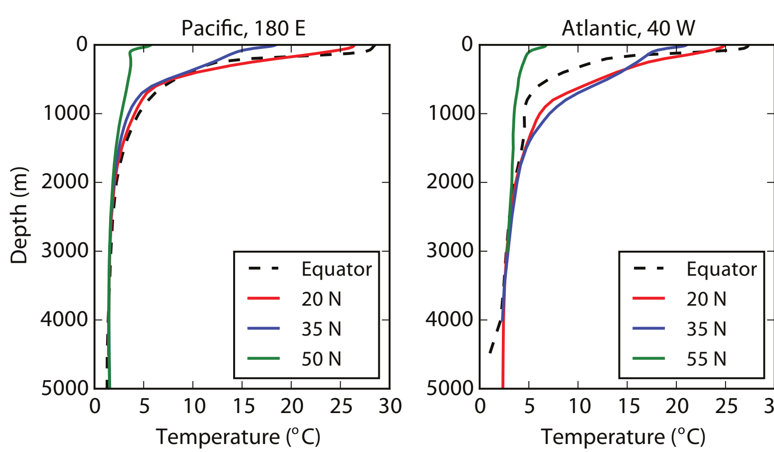

Fig. 15.2: Profiles of observed mean temperature in the North Pacific and North Atlantic at the longitudes and latitudes indicated. Note the shallowness of the equatorial thermoclines (especially in the Atlantic), and the weakness of the subpolar thermoclines.

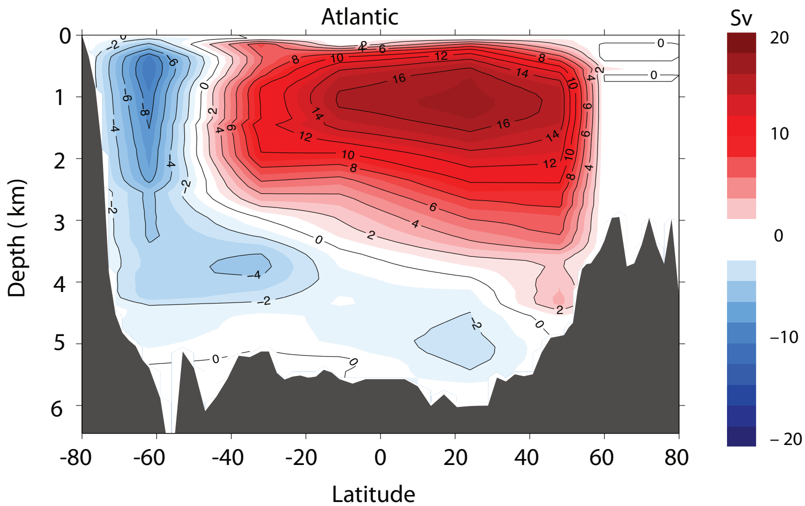

Fig. 15.3: The overturning circulation in the Atlantic Ocean as determined from observations in combination with a simple ('inverse') model. Red shading indicates a clockwise circulation, with water sinking in the North Atlantic and rising in the South, and this cell is predominantly *North Atlantic Deep Water*, or NADW. The blue, anticlockwise deep cell contains *Antarctic Bottom Water* that originates in the Southern Ocean and spreads northwards underneath the NADW. (Figure kindly provided by Loic Juillion using data and a methodology similar to that of Lumpkin & Speer 2007.)

a much weaker overturning circulation to such an extent that the globally-averaged overturning circulation largely reflects that of the Atlantic. The overturning circulation and the thermocline are, as one might expect, intimately linked and to explain one we must explain the other. Let us begin with a phenomenological discussion of the overturning circulation.

15.2 A Mixing-Driven Overturning Circulation

To begin with the simplest case let us consider the circulation in a closed, single hemispheric basin, and suppose that there is a net surface heating at low latitudes and a net cooling at high latitudes that maintains a meridional temperature gradient at the surface. It seems reasonable to imagine that there is a single overturning cell, with water sinking at high latitudes rising at low latitudes before returning to polar regions in the upper ocean, as illustrated schematically in Fig. 15.4 and Fig. 15.5. Is this a reasonable expectation? Can we explain why the water circulates at all?

15.2.1 Why the Water Circulates

Let us suppose that initially all the interior water is at some intermediate temperature, and we will also suppose that the flow in the interior is adiabatic, meaning that to a good approximation the subsurface water conserves its potential temperature as it moves around. Now, given a warm interior, cold surface water at high latitudes will be convectively unstable and will therefore sink, so that very quickly the dense water extends all the way to the ocean floor. By hydrostasy the pressure in the deep ocean is then higher at high latitudes than at low, where the water is warmer, and a pressure gradient then causes water to move equatorward, filling the abyss. Eventually, the *entire* ocean becomes filled with cold dense water of polar origin, except for a very thin layer at the surface, since the ocean surface at lower latitudes is kept at a higher temperature. Once the abyss

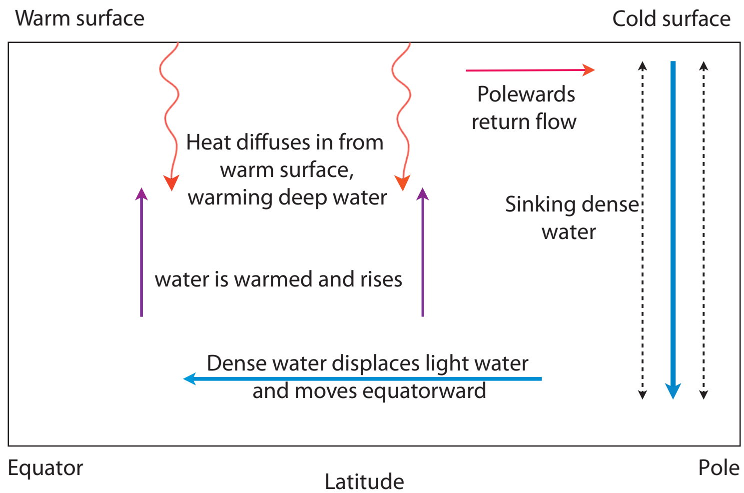

Fig. 15.4: Schematic of a single-celled meridional overturning circulation. Sinking is concentrated at high latitudes and upwelling spread out over lower latitudes.

is filled with dense water the surface polar waters *will no longer be convectively unstable.* The convection will thus cease and the circulation will halt! However, we know from observations that the deep ocean continues to circulate, albeit slowly, with the deep ocean completely overturning and the water being replaced on timescales of a few hundred years. There are two causes of the continued circulation, one being that the ocean mixes and the other being that the wind forcing at the top drives a deep circulation; we consider the effects of mixing first and come back to the wind-driving later in the chapter.

Mixing — either molecular mixing or in reality turbulent mixing, as discussed in Chapter 10 — will cause the higher surface temperatures in lower latitudes to diffuse down into the ocean interior. That is, the interior is slowly *warmed* by heat diffusion from above. This diffusion keeps the deep ocean slightly warmer than the cold polar surface waters, enabling the high-latitude convection and so the circulation itself to persist. The diffusion also extends the vertical temperature gradient down into the interior and we see in Fig. 15.2 how the vertical temperature profile varies with latitude. Except at the highest latitudes where the water is sinking and so almost uniform all the way to the bottom, we see that the temperature gradient is concentrated in the upper kilometre of the ocean, and this region is called the *main thermocline*. Why should the vertical temperature gradient be concentrated in the upper ocean? The upper ocean is the region of the gyres, which certainly creates a temperature gradient, but the underlying reason that the vertical temperature gradient is strongest there is more basic, as we now explore.

15.2.2 A Simple Kinematic Model of the Thermocline

In mid- and low latitudes cold water with polar origins upwells into a region of warmer water where high temperatures are diffusing down, and a simple model of this is the one-dimensional advective–diffusive balance,

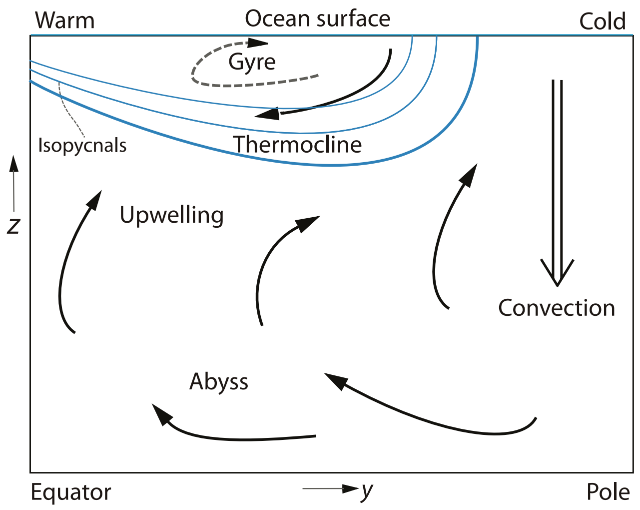

Fig. 15.5: Cartoon of a single-celled meridional overturning circulation, with a wall at the equator. Sinking is concentrated at high latitudes and upwelling spread out over lower latitudes. The thermocline is the boundary between the cold abyssal waters, with polar origins, and the warmer near-surface subtropical water. Wind forcing in the subtropics pushes the warm surface water into the fluid interior, deepening the thermocline as well as circulating as a gyre.

namely

$$w\frac{\partial T}{\partial z} = \kappa\frac{\partial^2 T}{\partial z^2}, \tag{15.1}$$

where w is the vertical velocity (which is positive), κ is a diffusivity and T is temperature. The equation represents a balance between the upwelling of cold water and the downward diffusion of heat. If w and κ are given and are constant, and if T is specified at the top ($T = T_T$ at $z = 0$) and if $T = T_B$ at great depth ($z = -\infty$) then the temperature falls exponentially away from the surface according to

$$T = (T_T - T_B)\mathrm{e}^{wz/\kappa} + T_B. \tag{15.2}$$

The scale at which temperature decays away from its surface value is given by

$$\delta = \frac{\kappa}{w}, \tag{15.3}$$

and this is an estimate of the thermocline thickness. It is not a useful a-priori estimate, because the magnitude of w depends on κ. However, it is reasonable to see if the observed ocean is broadly consistent with this expression. The diffusivity κ (which is an eddy diffusivity, maintained by small-scale turbulence) can be measured and is found to have values that range between $10^{-5}\,\mathrm{m^2\,s^{-1}}$ and $10^{-4}\,\mathrm{m^2\,s^{-1}}$ over much of the ocean, with higher values locally in some abyssal and shelf regions.

The vertical velocity is too small to be measured directly, but various estimates based on deep water production suggest a value of about $10^{-7}\,\mathrm{m\,s^{-1}}$. Using this and the smaller value of κ in (15.2) gives an e-folding vertical scale, κ/w, of order a hundred metres, beneath which the stratification is predicted to be very small (i.e., a nearly uniform density). Using the larger value of κ increases the vertical scale to 1000 m, similar to the observed value. Quantitative uncertainties aside, the model has a very robust result, that *the temperature gradient is concentrated in the upper ocean.*

as seen in Fig. 15.7. Here, C_0 is the strength of the convective source, which we take as given, $T_I(y)$ is the polewards flow in the interior, in the lower layer, across the latitude line at y, $T_W(y)$ is the equatorial flow in the deep western boundary current at y, and $U(y)$ is the total upwelling polewards of y. The terms on the left-hand side are mass sources to this region and the terms on the right-hand side are losses, and all are in units of $\mathrm{m^3\,s^{-1}}$ (since density is constant, mass balance and volume balance are synonymous). Over the entirety of the domain the source term must balance the upwelling, so that $C_0 = U(0)$, and we assume the upwelling is uniform.

The poleward transport in the interior is given using (15.13),

$$T_I(y) = \int vh \, \mathrm{d}x = \int \frac{fS}{\beta} \, \mathrm{d}x. \tag{15.15}$$

Now, since the upwelling S is uniform, and $\int S \, \mathrm{d}x \, \mathrm{d}y = SL_x L_y = U(0) = C_0$, we have

$$T_I(y) = \frac{fC_0}{\beta L_y} = \frac{C_0 y}{L_y}, \tag{15.16}$$

using $f = \beta y$. It is important to realise that this result is obtained using the potential vorticity equation and not the mass continuity equation.

The upwelling north of latitude y is given by

$$U(y) = SL_x(L_y - y) = C_0\left(1 - \frac{y}{L_y}\right). \tag{15.17}$$

Using (15.16) and (15.17) in (15.14) gives

$$\boxed{T_W(y) = \frac{2\,C_0 y}{L_y}.} \tag{15.18}$$

This is a remarkable result, for it tells us that the strength of the western boundary current near the source region is *twice* the strength of the source itself! The result arises because some of the flow in the deep layer is *recirculating,* going round and round without upwelling or coming from the source itself. The calculation itself is very approximate, but the fact that there is a deep western boundary current, and that the flow recirculates, transcend its limitations and these are robust predictions.

A final point to note is that we have taken the convective source to have a given magnitude. In reality, the strength of the source must match the strength of the upwelling, this being the strength of the overturning circulation itself. This is a function of the diapycnal diffusivity and the meridional temperature gradient, as described in Sections 15.2 and 15.3.

15.4 An Interhemispheric Overturning Circulation

As attractive as it may be, the theory of the overturning circulation and thermocline described in the preceding sections is only part of the picture.

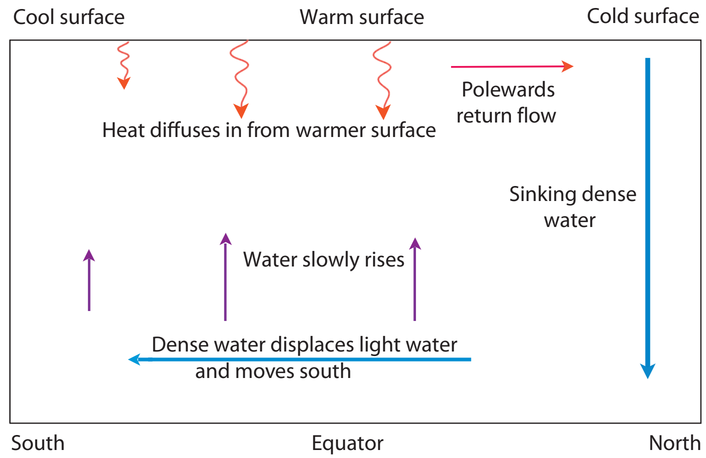

Fig. 15.8: An idealized interhemispheric overturning circulation. Water from the north sinks, because it is the densest water in the system. It displaces any lighter water, filling up the entire basin (i.e., both hemispheres) except for a thermocline near the surface. A circulation is maintained if heat diffuses in from the surface, warming the deep water and enabling it to rise. The strength of the circulation depends on the diffusivity, and if it is zero the circulation eventually halts.

In fact, much of the deep circulation is *interhemispheric:* we can see in Fig. 15.3 that much of the water that sinks in the North Atlantic upwells around 40°S or even further south in the Southern Ocean (although this only became truly apparent at the beginning of twenty-first century). In the rest of the chapter we try to understand why that should be.

15.4.1 A Basic Mechanism

An interhemispheric circulation of itself is of no particular surprise. For simplicity consider a 'shoebox' ocean consisting of a single basin stretching from high northern latitudes to high southern latitudes, and let us suppose that the surface at high latitudes in one hemisphere, say the North, is particularly cold and dense, as in Fig. 15.8. The physical situation then actually differs little from the situation described in Section 15.2. The densest water in the system sinks, and spreads equatorward. However, there is no reason that it should all upwell before it reaches the equator, although if the equatorial regions are warm the upwelling may be strong there because the downward diffusion of heat warms the deep water. Nonetheless, if the diffusion is small the densest water in the system displaces any lighter water and fills up both hemispheres of the basin, except for a thermocline near the surface. The flow away from the convective region occurs, as in the single-hemisphere model, in deep western boundary currents, with upwelling and return flow in the basin interior.

A non-zero circulation depends, as with the mixing-driven circulation, on there being a non-zero diffusivity to warm the deep water and allow it to rise. If the diffusivity were zero, then the entire basin would simply fill with the densest available water (with the exception of an infinitesimally thin layer at the surface) and the circulation would then halt.

15.4.2 A Wind-Driven Interhemispheric Circulation

The mixing-driven circulation described above is not the only mechanism, and is not in fact the main mechanism, whereby deep water ac-

Chapter

16

Equatorial Oceans and El Niño

Equatorial dynamics differs from its midlatitude counterpart because the Coriolis parameter is relatively small and the Rossby number large, and balanced and unbalanced dynamics then become intertwined. Yet if we move more than a few degrees away from the equator the Rossby number again becomes quite small, suggesting that familiar ways of investigating the dynamics — Sverdrup balance for example — might yet play a role. Not surprisingly, the equatorial ocean is the home to a multitude of interesting phenomena and in this chapter we discuss just two of the most striking, namely the equatorial undercurrent and El Niño. Let us first see what the observations tell us.

16.1 Observations of the Equatorial Ocean

The most distinctive features of equatorial oceans are illustrated in Fig. 16.1 and the top panel of Fig. 16.2, namely:

(i) A shallow westward flowing surface current, typically confined to the upper 50 m or less, strongest within a few degrees of the equator, although not always symmetric about the equator. Its speed is typically a few tens of centimetres per second.

(ii) A strong coherent eastward undercurrent extending to about 200 m depth, confined to within a few degrees of the equator. Its speed is up to a metre per second, and it is this current that dominates the vertically integrated transport at the equator. Beneath the undercurrent the flow is relatively weak.

(iii) Westward flow on either side of the undercurrent, with eastward countercurrents poleward of this. The Pacific countercurrent is strongest in the Northern Hemisphere, where it reaches the surface.

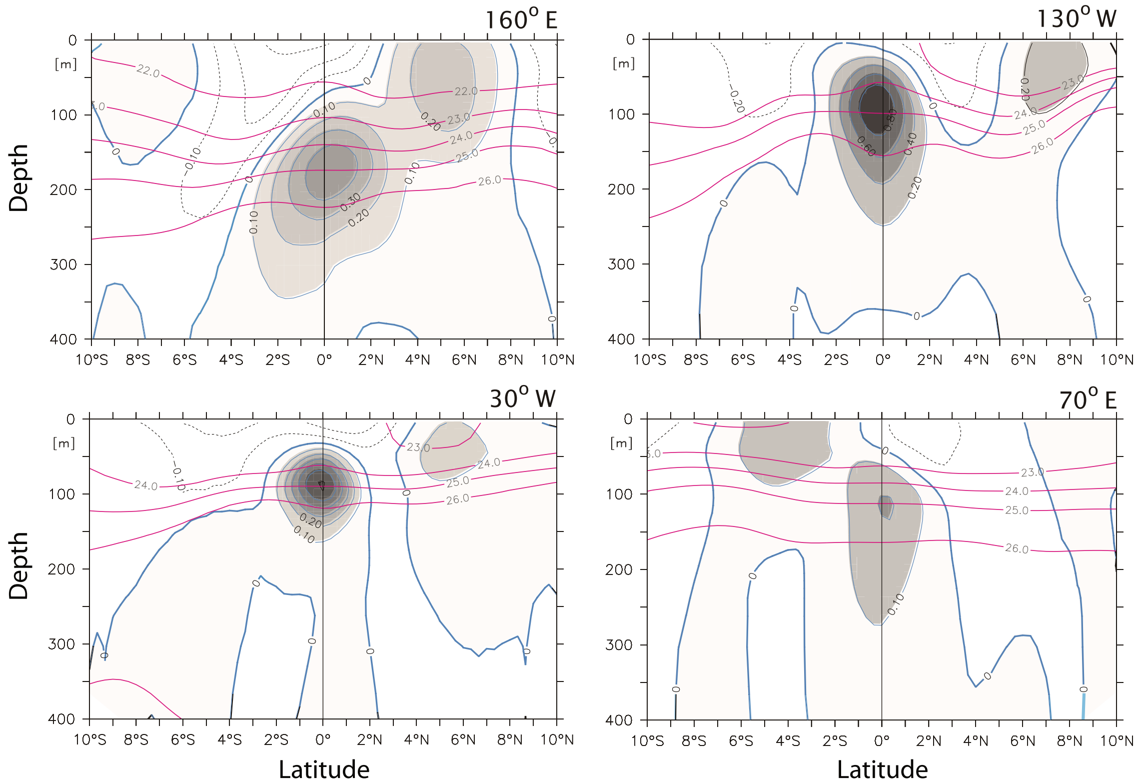

Fig. 16.1: Sections of the mean zonal current (shading and associated contours) at two longitudes in the Pacific (upper panels), in the Atlantic (lower left) and in the Indian Ocean (lower right). The contours are every 20 cm s^{-1} in the upper two panels and every 10 cm s^{-1} in the lower panels. Note the well-defined eastward undercurrent at the equator in all panels, and a weaker eastward counter-current at about 6°N and/or 6°S. The red, more horizontal, lines are isolines of potential density.

These features are largely common to both the Atlantic and Pacific Oceans and to a somewhat lesser extent in the Indian Ocean. We start our dynamical explorations with the vertically integrated flow.

16.2 Vertically Integrated Flow and Sverdrup Balance

In midlatitudes the large scale currents system may be understood using the planetary geostrophic equations of motion, with Sverdrup balance (Section 14.2) providing a solid foundation on which to build. As we approach lower latitudes the Coriolis parameter, f, decreases and the Rossby number increases and one might expect that dynamics based on geostrophic balance will ultimately fail. However, it is only very close to the equator that the Rossby number exceeds unity: if we take a velocity of 0.5 m s^{-1}and a length scale of 500 km then the Rossby number at 5° latitude is 0.08, at 2° it is 0.2 and at 1° it is 0.4. These numbers suggest that until we are virtually at the equator we can use some of the familiar tools from the midlatitude dynamics. Let us first see the extent to which the familiar Sverdrup balance can explain the vertically integrated flow. The horizontal momentum may be written

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla \phi + \frac{1}{\rho_0} \frac{\partial \boldsymbol{\tau}}{\partial z}, \qquad (16.1)$$

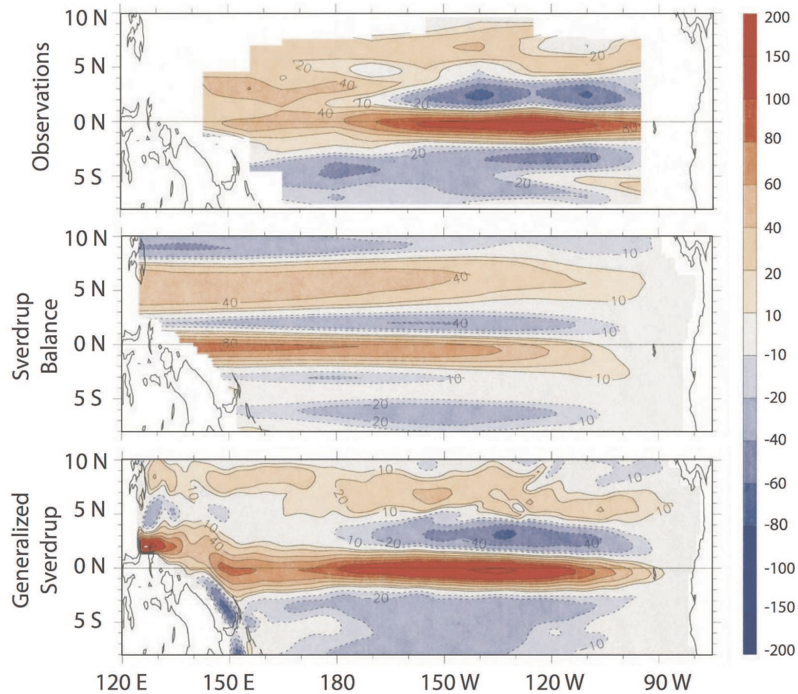

Fig. 16.2: Vertically integrated zonal transport in the Pacific. Red colours indicate eastward flow, blue colours westward. The top panel shows the observed flow, the middle panel shows the flow calculated using Sverdrup balance with the observed wind, and the bottom panel shows the flow calculated with a 'generalized' Sverdrup balance that includes the nonlinear terms in a diagnostic way.

where $\boldsymbol{\tau}$ is the stress on the fluid. As in earlier chapters, we will absorb the constant density, ρ_0, into the stress, so that $\boldsymbol{\tau}/\rho \to \boldsymbol{\tau}$. The mass conservation equation is

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{16.2}$$

which, on vertical integration over the depth of the ocean, gives

$$\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0, \tag{16.3}$$

where U and V are the vertically integrated zonal and meridional velocities (e.g., $U = \int u \, dz$) and we assume the ocean has a flat bottom and a rigid lid at the top. If we assume the flow is steady and integrate (16.1) vertically, then take the curl and use (16.3), we obtain

$$\beta V = \mathrm{curl}_z(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B) + \mathrm{curl}_z \boldsymbol{N}, \tag{16.4}$$

where the subscripts T and B denote top and bottom, $\boldsymbol{N}$ represents all the nonlinear terms and curl_z is defined by $\mathrm{curl}_z \boldsymbol{A} \equiv \partial A^y/\partial x - \partial A^x/\partial y = \hat{\mathbf{k}} \cdot \nabla_3 \times \boldsymbol{A}$. Equations (16.4) and (16.3) are closed equations for the vertically averaged flow.

If we neglect the nonlinear terms and the stress at the bottom (we'll come back to these terms later) then (16.4) becomes

$$\beta V = \mathrm{curl}_z \boldsymbol{\tau}_T. \tag{16.5}$$

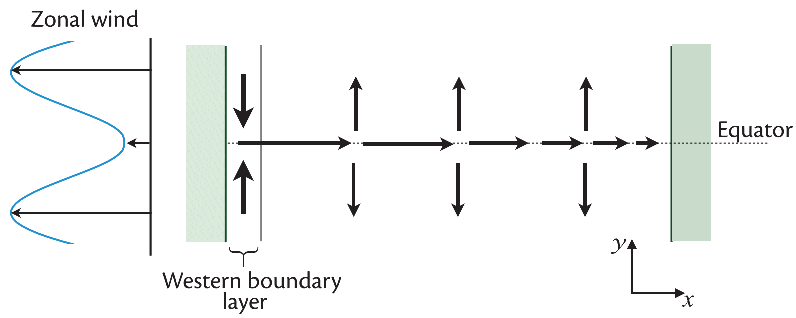

Fig. 16.3: Schema of Sverdrup flow at the equator between two meridional boundaries. The mean winds are all westward, but with a minimum in magnitude at the equator. By Sverdrup balance, (16.5), the wind stress produces the divergent meridional flow shown, which in turn induces an eastward equatorial zonal flow, strongest in the western part of the basin.

This is just Sverdrup balance, familiar from Chapter 14. The zonal transport is obtained by differentiating (16.5) with respect to y, using (16.3) to replace $\partial_y V$ with $\partial_x U$, and then integrating from the eastern boundary (x_E). This procedure gives

$$U = -\frac{1}{\beta}\int_{x_E}^{x} \frac{\partial}{\partial y} \mathrm{curl}_z \boldsymbol{\tau}_T \,\mathrm{d}x' + U(x_E, y). \tag{16.6}$$

We don't integrate from the western boundary because a boundary layer can be expected there, whereas the value of U at the eastern boundary, namely U_E, will be small.

The wind stress is known from observations and we can then use (16.6) to calculate U, which is found to be generally positive (eastward) at the equator. The solution is plotted in the middle panel of Fig. 16.2. There is a good but not perfect agreement with the observations, shown in the top panel. In the western equatorial Pacific the observed eastward flow is quite broad whereas the eastward Sverdrup flow is narrow, flanked on either side by westward flow, and much of this discrepancy can be attributed to the role of the nonlinear and frictional terms, as illustrated in the bottom panel of Fig. 16.2. To obtain the results shown, the the nonlinear terms (which have the form $\mathrm{curl}_z(\int \boldsymbol{u}\cdot\nabla\boldsymbol{u}\,\mathrm{d}z)$ are included in a diagnostic fashion. That is to say, the term $\mathrm{curl}_z\boldsymbol{N}$ is evaluated from observations and included on the right-hand side of (16.4) in order to calculate a 'generalized Sverdrup' flow, which (as one might expect) is in better agreement with the observations. Perhaps the most interesting point is that, even quite close to the equator and even without the nonlinear terms, Sverdrup balance provides a qualitatively correct picture of the vertically averaged flow, with the longitudinal structure of the flow sketched in Fig. 16.3.

16.2.1 Sensitivity of the Sverdrup Flow

Although the calculations of Sverdrup flow do show good agreement with observations, the calculation — and, most likely, the observed flow — is rather sensitive to the precise form of the winds. To illustrate this, suppose that $U(x_E, y) = 0$ and the stress is zonal and uniform, then (16.6) becomes

$$U(x, y) = \frac{1}{\beta}(x - x_E)\frac{\partial^2 \tau_T^x}{\partial y^2}. \tag{16.7}$$

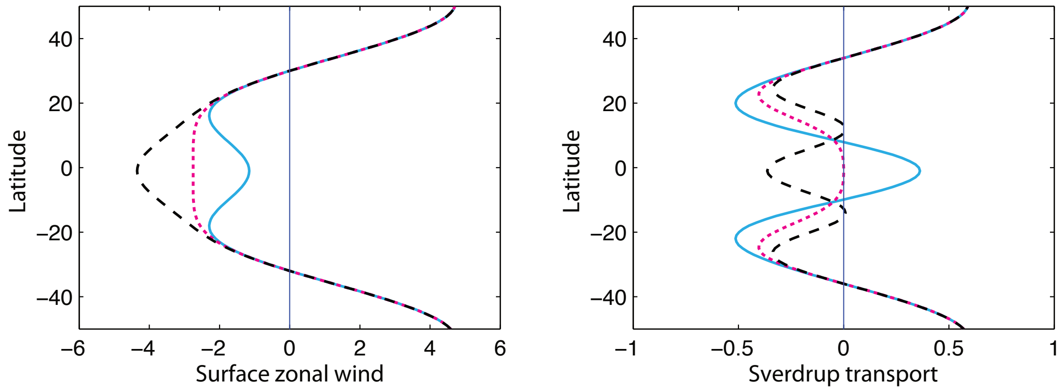

Fig. 16.4: The left panel shows three putative surface zonal (atmospheric) winds, u, all with westward winds in the tropics and with the solid line being the most realistic. The right panel shows the corresponding negative of the second derivative, $-\partial^2 u/\partial y^2$, proportional to the (oceanic) Sverdrup transport, in arbitrary units. The wind represented by solid (blue) line gives an eastward transport at the equator, as is observed, with the others differing markedly.

That is, the depth integrated flow is proportional to the second derivative of the zonal wind stress, and because $x < x_E$ we have $U \propto -\partial^2 \tau_T^x/\partial y^2$. Now, although the zonal wind is generally westward in the tropics there is a minimum in the magnitude of that wind near the equator (that is, there is a local maximum as sketched in the left panel Fig. 16.3) so that $\partial^2 \tau_T^x/\partial y^2$ is negative. Without this local maximum the Sverdrup flow would be westward at the equator.

This sensitivity of the Sverdrup flow to the wind pattern is illustrated in Fig. 16.4. The figure shows three surface zonal wind distributions, with the 'w' shaped solid line having a minimum in the westward flow (i.e., a minimum in the trade winds) at the equator and so being the most realistic. The right-hand panel shows the negative of the second derivative of the winds which is proportional to the zonal Sverdrup flow. Only in the one case (the blue line) does the wind produce an eastward Sverdrup flow. In fact, in the case illustrated with the dashed lines, the small changes in the meridional gradient of the wind between 15° and 20° produce large variations in the Sverdrup transport. Given this sensitivity, the small difference in the latitudinal variation of the Sverdrup flow and the observed flow, illustrated in the top and middle panels of Fig. 16.2, is not surprising and cannot be considered a major failure of the theory. However, the difference in the longitudinal structure of the two fields is indicative of the importance of other terms in the vorticity balance.

Although the Sverdrup flow is rather sensitive to the horizontal derivatives of wind pattern, the undercurrent itself is not, and let us turn our attention to that.

16.3 Dynamics of the Equatorial Undercurrent

The equatorial undercurrent is perhaps the single most conspicuous feature of the ocean current system at low latitudes and we now describe a model for it. Our model will be a local one, meaning that it is the direct effect of the winds in the equatorial region that drive the current, and although it does provide a simple, compelling explanation for the undercurrent it is an incomplete picture: it does not account for the remote effects of winds in building up a head of pressure that can produce an un-

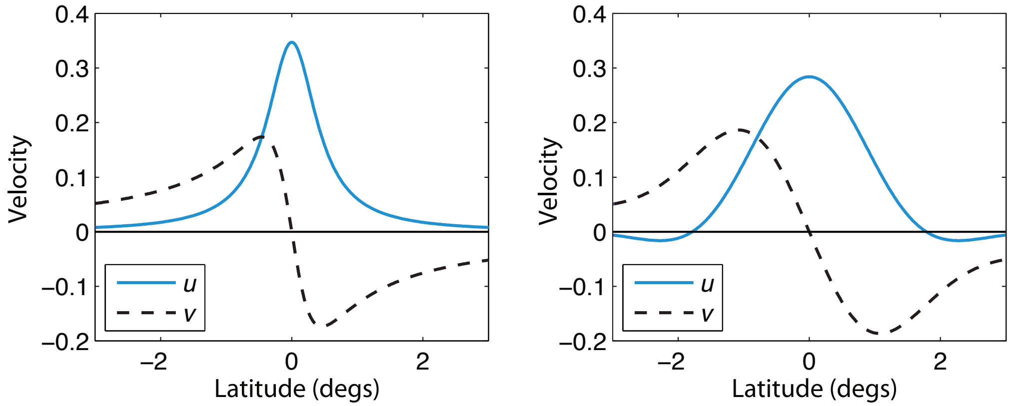

Fig. 16.6: Horizontal profiles of the undercurrent with friction represented by a linear drag (left, as given by (16.20)) and by a harmonic viscosity (right, see text), in dimensional units (m/s for velocity and degrees for latitude).

(i) The undercurrent is concentrated at the equator, decaying quite rapidly with latitude.

(ii) The deep meridional flow is zero at the equator, where $f = 0$, but is toward the equator in both hemispheres and therefore induces equatorial upwelling.

The latitudinal width of the undercurrent is determined by the ratio of β to r. Thus, with $\tau^y = 0$ (16.20a) becomes

$$u_2 = \frac{-\tau^x r}{H_1(r^2 + \beta^2 y^2)}, \tag{16.21}$$

and the width of the undercurrent scales as r/β — more friction gives a broader undercurrent.

Viscosity instead of drag

The frictional parameter r is a little arbitrary and unrealistic — friction does not act as a simple drag in the real ocean. To remedy this we can carry through a similar calculation with a viscosity instead of a drag, and we can also allow a continuous variation in the vertical instead of restricting ourselves to two layers. The equations of motion are similar except that terms like ru are replaced by $\nu\nabla^2 u$. The algebra to obtain a solution is now considerably more complicated, but the underlying mechanism producing the undercurrent is exactly the same and the solution itself is quite similar, as illustrated in the right-hand panel of Fig. 16.6. But we will not purse this topic further in this book; rather, let us turn our attention to that other great equatorial phenomenon, El Niño.

16.4 El Niño and the Southern Oscillation

Without capitalization, el niño is Spanish for male infant, whereas el Niño refers to the Christ Child and El Niño to the oceanic phenomenon.

El Niño! One of the most famous phenomena in the climate sciences, and certainly one with an enormous impact on humankind. El Niño is an anomalous warming of the surface waters in the eastern equatorial Pacific, peaking around Christmas-time, and its appealing name belies

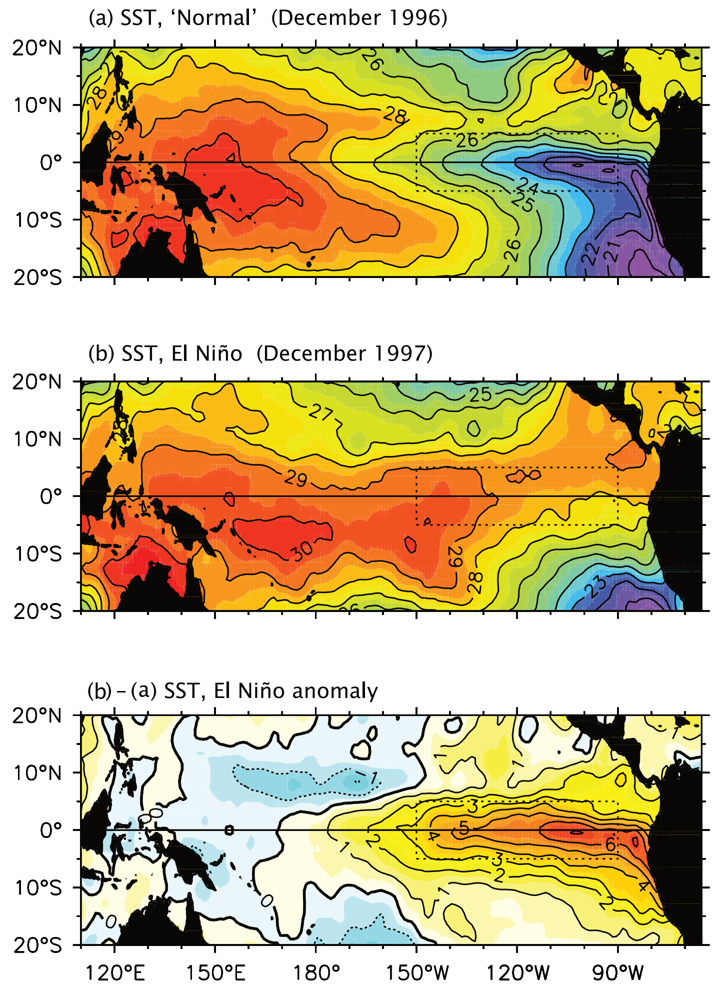

Fig. 16.7: The sea-surface temperature in December of a non-El Niño year (December 1996, top panel), a strong El Niño year (December 1997, middle panel) and their difference (bottom panel). An El Niño year is typically characterized by an anomalously warm tongue of water in the eastern tropical Pacific. The El Niño 3 region is the rectangular region demarcated by thin dotted lines in the eastern equatorial region. (Figure courtesy of A. Wittenberg, using data from NOAA.

its enormous power and global effects, bringing heavy rains to California and Northern Argentina and anomalously dry weather to South East Asia and Northern and Eastern Australia; it also raises the global average surface temperature by about half a degree Celsius. Taken with the associated changes in the atmosphere, in which case the whole phenomenon is known as the El Niño–Southern Oscillation (ENSO), it is the largest and most important source of global climate variability on interannual timescales.

16.4.1 A Descriptive Overview

Every few years the temperature of the surface waters in the eastern tropical Pacific rises quite significantly. The strongest warming takes place between about 5°S to 5°N, and from the west coast of Peru (a longitude of about 80°W) almost to the dateline, at 180°W, as illustrated in Fig. 16.7. The warming is large, with a difference in temperature up to 6°C from an

Bibliography

Acheson, D. J., 1990. *Elementary Fluid Dynamics*. Oxford University Press, 406 pp.

Ambaum, M. H. P., 2010. *Thermal Physics of the Atmosphere*. Wiley, 239 pp.

Andrews, D. G. & McIntyre, M. E., 1976. Planetary waves in horizontal and vertical shear: the generalized Eliassen–Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–2048.

Batchelor, G. K., 1969. Computation of the energy spectrum in homogeneous two-dimensional turbulence. *Phys. Fluids Suppl.*, **12**, II–233–239.

Bengtsson, L., Bonnet, R.-M., Grinspoon, D., Koumoutsaris, S. *et al.*, Eds., 2013. *Towards Understanding the Climate of Venus: Applications of Terrestrial Models to Our Sister Planet*. Springer Science & Business Media, 185 pp.

Bohren, C. F. & Albrecht, B. A., 1998. *Atmospheric Thermodynamics*. Oxford University Press, 416 pp.

Boussinesq, J., 1903. Théorie analytique de la chaleur (Analytic theory of heat). *Tome, Paris, Gauthier-Villars*, **II**, 170–172.

Boyd, J. P., 1976. The noninteraction of waves with the zonally averaged flow on a spherical Earth and the interrelationships of eddy fluxes of energy, heat and momentum. *J. Atmos. Sci.*, **33**, 2285–2291.

Brewer, A. W., 1949. Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere. *Quart. J. Roy. Meteor. Soc.*, **75**, 251–363.

Brillouin, L., 1926. La mécanique ondulatoire de Schrödinger; une méthode générale de resolution par approximations successives (The wave mechanics of Schrödinger: a general method of solution by successive approximation). *Comptes Rendus*, **183**, 24–26.

Bühler, O., 2009. *Waves and Mean Flows*. Cambridge University Press, 370 pp.

Carillo, C. N., 1892. Desertacion sobre las corrientes y estudios de la corriente Peruana de Humboldt. (Dissertation on currents and studies of the Peruvian Humboldt current). *Bol. Soc. Geogr. Lima*, **11**, 72–110.

Charney, J. G., 1947. The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, **4**, 135–162.

Charney, J. G., 1948. On the scale of atmospheric motion. *Geofys. Publ. Oslo*, **17** (2), 1–17.

Charney, J. G., 1963. A note on large-scale motions in the tropics. *J. Atmos. Sci.*, **20**, 607–609.

Charney, J. G., 1971. Geostrophic turbulence. *J. Atmos. Sci.*, **28**, 1087–1095.

Charney, J. G. & Drazin, P. G., 1961. Propagation of planetary scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83–109.

Charney, J. G. & Stern, M. E., 1962. On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172.

Clarke, A. J., 2008. *An Introduction to the Dynamics of El Niño and the Southern Oscillation*. Elsevier, 308 pp.

Cushman-Roisin, B. & Beckers, J.-M., 2011. *Introduction to Geophysical Fluid Dynamics*. 2nd edn. Academic Press, 828 pp.

Davidson, P., 2015. *Turbulence: An Introduction for Scientists and Engineers*. Oxford University Press, 688 pp.

Dickinson, R. E., 1969. Theory of planetary wave–zonal flow interaction. *J. Atmos. Sci.*, **26**, 73–81.

Dobson, G. M. B., 1956. Origin and distribution of the polyatomic molecules in the atmosphere. *Proc. Roy. Soc. Lond. A*, **236**, 187–193.

Döös, K. & Coward, A., 1997. The Southern Ocean as the major upwelling zone of the North Atlantic. *Int. WOCE Newsletter*, **27**, 3–4.

Drazin, P. G. & Reid, W. H., 1981. *Hydrodynamic Stability*. Cambridge University Press, 527 pp.

Durran, D. R., 1993. Is the Coriolis force really responsible for the inertial oscillations? *Bull. Am. Meteor. Soc.*, **74**, 2179–2184.

Durran, D. R., 2015. Lee waves and mountain waves. *Encycl. Atmos. Sci.*, 2nd edn, **4**, 95–102.

Durst, C. S. & Sutcliffe, R. C., 1938. The effect of vertical motion on the "geostrophic departure" of the wind. *Quart. J. Roy. Meteor. Soc.*, **64**, 240.

Eady, E. T., 1949. Long waves and cyclone waves. *Tellus*, **1**, 33–52.

Eady, E. T., 1950. The cause of the general circulation of the atmosphere. In *Cent. Proc. Roy. Meteor. Soc.* (1950), pp. 156–172.

Egger, J., 1999. Inertial oscillations revisited. *J. Atmos. Sci.*, **56**, 2951–2954.

Ekman, V. W., 1905. On the influence of the Earth's rotation on ocean currents. *Arch. Math. Astron. Phys.*, **2**, 1–52.

Eliassen, A. & Palm, E., 1961. On the transfer of energy in stationary mountain waves. *Geofys. Publ.*, **22**, 1–23.

Emanuel, K. A., 1994. *Atmospheric Convection*. Oxford University Press, 580 pp.

Fofonoff, N. P., 1954. Steady flow in a frictionless homogeneous ocean. *J. Mar. Res.*, **13**, 254–262.

Gill, A. E., 1982. *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.

Gnanadesikan, A., 1999. A simple predictive model for the structure of the oceanic pycnocline. *Science*, **283**, 2077–2079.

Haynes, P., 2005. Stratospheric dynamics. *Ann. Rev. Fluid Mech.*, **37**, 263–293.

Heimpel, M., Gastine, T. & Wicht, J., 2016. Simulation of deep-seated zonal jets and shallow vortices in gas giant atmospheres. *Nature Geos.*, **9**, 19–23.

Held, I. M., 2000. The general circulation of the atmosphere. In *Woods Hole Program in Geophysical Fluid Dynamics* (2000), pp. 66.

Held, I. M. & Hou, A. Y., 1980. Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.*, **37**, 515–533.

Holmes, M. H., 2013. *Introduction to Perturbation Methods*. 2nd edn. Springer, 436 pp.

Index