ELSEVIER
ACADEMIC
PRESS

An Introduction to DYNAMIC METEOROLOGY

FOURTH EDITION

James R. Holton

An Introduction to Dynamic Meteorology

FOURTH EDITION

This is Volume 88 in the
INTERNATIONAL GEOPHYSICS SERIES
A series of monographs and textbooks
Edited by RENATA DMOWSKA, JAMES R. HOLTON and H. THOMAS ROSSBY
A complete list of books in this series appears at the end of this volume.

AN INTRODUCTION TO DYNAMIC METEOROLOGY

Fourth Edition

JAMES R. HOLTON
Department of Atmospheric Sciences
University of Washington
Seattle, Washington

ELSEVIER
ACADEMIC
PRESS

Amsterdam Boston Heidelberg London New York Oxford Paris
San Diego San Francisco Singapore Sydney Tokyo

Senior Editor, Earth Sciences	Frank Cynar
Editorial Coordinator	Jennifer Helé
Senior Marketing Manager	Linda Beattie
Cover Design	Eric DeCicco
Composition	Integra Software Services
Printer/Binder	The Maple-Vail Manufacturing Group
Cover printer	Phoenix Color Corp

Elsevier Academic Press
200 Wheeler Road, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper. ♾

Library of Congress Cataloging-in-Publication Data
Application submitted

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 0-12-354015-1
CD-Rom ISBN: 0-12-354016-X

For all information on all Academic Press publications
visit our Web site at www.academicpressbooks.com

PRINTED IN THE UNITED STATES OF AMERICA
04 05 06 07 08 9 8 7 6 5 4 3 2 1

CONTENTS

Chapter 7 **Atmospheric Oscillations: Linear Perturbation Theory**

Chapter 8 **Synoptic-Scale Motions II: Baroclinic Instability**

Chapter 9 **Mesoscale Circulations**

Chapter 10 **The General Circulation**

PREFACE

In this fourth edition of *An Introduction to Dynamic Meteorology* I have retained the basic structure of all chapters of the previous edition. A number of minor corrections, pedagogical improvements, and updates of material are included throughout. The major departure from previous editions, however, is inclusion of a variety of computer-based exercises and demonstrations utilizing the MATLAB® programming language at the end of each chapter (MATLAB® is a registered trademark of The MathWorks, Inc.). These will, I hope, provide students with an opportunity to visualize and experiment with various aspects of dynamics not readily accessible through analytic problem solving.

I have chosen MATLAB because it is a high-level language with excellent graphing capabilities and is readily available to most university students. The ability within MATLAB to animate wave fields is particularly valuable as a learning aid in dynamic meteorology. It is not necessary to have much experience with MATLAB to solve most of the problems provided. In most cases MATLAB scripts (M-files) are provided on the accompanying CD, and the student need only run the scripts for various parameter choices or make minor revisions. Through studying the various examples, students should gradually be able to gain the confidence to program their own MATLAB scripts.

Much of the material included in this text is based on a two-term course sequence for seniors majoring in atmospheric sciences at the University of Washington. It would also be suitable for first-year graduate students with little previous background in meteorology. As in the previous editions the emphasis in the text is on physical principles rather than mathematical elegance. It is assumed that the reader has mastered the fundamentals of classical physics and has a thorough knowledge of elementary calculus. Some use is made of vector calculus. In most cases, however, the vector operations are elementary in nature so that the reader with little background in vector operations should not experience undue difficulties.

The fundamentals of fluid dynamics necessary for understanding large-scale atmospheric motions are presented in Chapters 1–5. These have undergone only minor revisions from the previous edition. The development of the Coriolis force in Section 1.5 has been substantially improved from previous editions. The discussion of the barotropic vorticity equation in Section 4.5 now introduces the streamfunction. As in previous editions, Chapter 6 is devoted to quasi-geostrophic theory, which is still fundamental to the understanding of large-scale extratropical motions. This chapter has been revised to provide increased emphasis on the role of potential vorticity and potential vorticity inversion. The presentation of the omega equation and the **Q** vector has been revised and improved.

In Chapter 9, the discussions of fronts, symmetric instability, and hurricanes have all been expanded and improved. Chapter 10 now includes a discussion of annular modes of variability, and the discussion of general circulation models has been rewritten. Chapter 11 has an improved discussion of El Niño and of steady equatorial circulations. Chapter 12 presents a revised discussion of the general circulation of the stratosphere, including discussions of the residual circulation and trace constituent transport. Finally, Chapter 13 has been updated to briefly summarize modern data assimilation techniques and ensemble forecasting.

Acknowledgments: I am indebted to a large number of colleagues and students for their continuing interest, suggestions, and help with various figures. I am particularly grateful to Drs. Dale Durran, Greg Hakim, Todd Mitchell, Adrian Simmons, David Thompson, and John Wallace for various suggestions and figures.

CHAPTER 1

Introduction

1.1 THE ATMOSPHERIC CONTINUUM

Dynamic meteorology is the study of those motions of the atmosphere that are associated with weather and climate. For all such motions the discrete molecular nature of the atmosphere can be ignored, and the atmosphere can be regarded as a continuous fluid medium, or *continuum*. A "point" in the continuum is regarded as a volume element that is very small compared with the volume of atmosphere under consideration, but still contains a large number of molecules. The expressions *air parcel* and *air particle* are both commonly used to refer to such a point. The various physical quantities that characterize the state of the atmosphere (e.g., pressure, density, temperature) are assumed to have unique values at each point in the atmospheric continuum. Moreover, these *field variables* and their derivatives are assumed to be continuous functions of space and time. The fundamental laws of fluid mechanics and thermodynamics, which govern the motions of the atmosphere, may then be expressed in terms of partial differential equations involving the field variables as dependent variables and space and time as independent variables.

The general set of partial differential equations governing the motions of the atmosphere is extremely complex; no general solutions are known to exist. To acquire an understanding of the physical role of atmospheric motions in determining the observed weather and climate, it is necessary to develop models based on systematic simplification of the fundamental governing equations. As shown in later chapters, the development of models appropriate to particular atmospheric motion systems requires careful consideration of the scales of motion involved.

1.2 PHYSICAL DIMENSIONS AND UNITS

The fundamental laws that govern the motions of the atmosphere satisfy the principle of *dimensional homogeneity*. That is, all terms in the equations expressing these laws must have the same physical dimensions. These dimensions can be expressed in terms of multiples and ratios of four dimensionally independent properties: length, time, mass, and thermodynamic temperature. To measure and compare the scales of terms in the laws of motion, a set of units of measure must be defined for these four fundamental properties.

In this text the international system of units (SI) will be used almost exclusively. The four fundamental properties are measured in terms of the SI *base units* shown in Table 1.1. All other properties are measured in terms of SI *derived units*, which are units formed from products or ratios of the base units. For example, velocity has the derived units of meter per second (m s^{-1}). A number of important derived units have special names and symbols. Those that are commonly used in dynamic meteorology are indicated in Table 1.2. In addition, the supplementary unit designating a plane angle, the radian (rad), is required for expressing angular velocity (rad s^{-1}) in the SI system.[1]

In order to keep numerical values within convenient limits, it is conventional to use decimal multiples and submultiples of SI units. Prefixes used to indicate such multiples and submultiples are given in Table 1.3. The prefixes of Table 1.3 may be affixed to any of the basic or derived SI units except the kilogram. Because the

Table 1.1 SI Base Units

Property	Name	Symbol
Length	Meter (meter)	m
Mass	Kilogram	kg
Time	Second	s
Temperature	Kelvin	K

[1] Note that *Hertz* measures frequency in *cycles* per second, not in radians per second.

Table 1.2 SI Derived Units with Special Names

Property	Name	Symbol
Frequency	Hertz	Hz (s^{-1})
Force	Newton	N (kg m s^{-2})
Pressure	Pascal	Pa (N m^{-2})
Energy	Joule	J (N m)
Power	Watt	W (J s^{-1})

kilogram already is a prefixed unit, decimal multiples and submultiples of mass are formed by prefixing the gram (g), not the kilogram (kg).

Although the use of non-SI units will generally be avoided in this text, there are a few exceptions worth mentioning:

1. In some contexts, the time units minute (min), hour (h), and day (d) may be used in preference to the second in order to express quantities in convenient numerical values.
2. The hectopascal (hPa) is the preferred SI unit for pressure. Many meteorologists, however, are still accustomed to using the millibar (mb), which is numerically equivalent to 1 hPa. For conformity with current best practice, pressures in this text will generally be expressed in hectopascals (e.g., standard surface pressure is 1013.25 hPa).
3. Observed temperatures will generally be expressed using the Celsius temperature scale, which is related to the thermodynamic temperature scale as follows:

$$T_C = T - T_0$$

where T_C is expressed in degrees Celsius (°C), T is the thermodynamic temperature in Kelvins (K), and $T_0 = 273.15$ K is the freezing point of water on the Kelvin scale. From this relationship it is clear that one Kelvin unit equals one degree Celsius.

Table 1.3 Prefixes for Decimal Multiples and Submultiples of SI Units

Multiple	Prefix	Symbol
10^6	Mega	M
10^3	Kilo	k
10^2	Hecto	h
10^1	Deka	da
10^{-1}	Deci	d
10^{-2}	Centi	c
10^{-3}	Milli	m
10^{-6}	Micro	μ

1.3 SCALE ANALYSIS

Scale analysis, or scaling, is a convenient technique for estimating the magnitudes of various terms in the governing equations for a particular type of motion. In scaling, typical expected values of the following quantities are specified: (1) magnitudes of the field variables; (2) amplitudes of fluctuations in the field variables; and (3) the characteristic length, depth, and time scales on which these fluctuations occur. These typical values are then used to compare the magnitudes of various terms in the governing equations. For example, in a typical midlatitude synoptic[2] cyclone the surface pressure might fluctuate by 10 hPa over a horizontal distance of 1000 km. Designating the amplitude of the horizontal pressure fluctuation by δp, the horizontal coordinates by x and y, and the horizontal scale by L, the magnitude of the horizontal pressure gradient may be estimated by dividing δp by the length L to get

$$\left(\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}\right) \sim \frac{\delta p}{L} = 10 \text{ hpa}/10^3 \text{ km } \left(10^{-3} \text{ Pa m}^{-1}\right)$$

Pressure fluctuations of similar magnitudes occur in other motion systems of vastly different scale such as tornadoes, squall lines, and hurricanes. Thus, the horizontal pressure gradient can range over several orders of magnitude for systems of meteorological interest. Similar considerations are also valid for derivative terms involving other field variables. Therefore, the nature of the dominant terms in the governing equations is crucially dependent on the horizontal scale of the motions. In particular, motions with horizontal scales of a few kilometers or less tend to have short time scales so that terms involving the rotation of the earth are negligible, while for larger scales they become very important. Because the character of atmospheric motions depends so strongly on the horizontal scale, this scale provides a convenient method for the classification of motion systems. Table 1.4 classifies examples of various types of motions by horizontal scale for the spectral region from 10^{-7} to 10^7 m. In the following chapters, scaling arguments are used extensively in developing simplifications of the governing equations for use in modeling various types of motion systems.

1.4 FUNDAMENTAL FORCES

The motions of the atmosphere are governed by the fundamental physical laws of conservation of mass, momentum, and energy. In Chapter 2, these principles are applied to a small volume element of the atmosphere in order to obtain the

[2] The term *synoptic* designates the branch of meteorology that deals with the analysis of observations taken over a wide area at or near the same time. This term is commonly used (as here) to designate the characteristic scale of the disturbances that are depicted on weather maps.

Table 1.4 Scales of Atmospheric Motions

Type of motion	Horizontal scale (m)
Molecular mean free path	10^{-7}
Minute turbulent eddies	$10^{-2} - 10^{-1}$
Small eddies	$10^{-1} - 1$
Dust devils	$1 - 10$
Gusts	$10 - 10^{2}$
Tornadoes	10^{2}
Cumulonimbus clouds	10^{3}
Fronts, squall lines	$10^{4} - 10^{5}$
Hurricanes	10^{5}
Synoptic cyclones	10^{6}
Planetary waves	10^{7}

governing equations. However, before deriving the complete momentum equation it is useful to discuss the nature of the forces that influence atmospheric motions.

These forces can be classified as either *body forces* or *surface forces*. Body forces act on the center of mass of a fluid parcel; they have magnitudes proportional to the mass of the parcel. Gravity is an example of a body force. Surface forces act across the boundary surface separating a fluid parcel from its surroundings; their magnitudes are independent of the mass of the parcel. The pressure force is an example.

Newton's second law of motion states that the rate of change of momentum (i.e., the acceleration) of an object, as measured relative to coordinates fixed in space, equals the sum of all the forces acting. For atmospheric motions of meteorological interest, the forces that are of primary concern are the pressure gradient force, the gravitational force, and friction. These *fundamental* forces are the subject of the present section. If, as is the usual case, the motion is referred to a coordinate system rotating with the earth, Newton's second law may still be applied provided that certain *apparent* forces, the centrifugal force and the Coriolis force, are included among the forces acting. The nature of these apparent forces is discussed in Section 1.5.

1.4.1 Pressure Gradient Force

We consider an infinitesimal volume element of air, $\delta V = \delta x \delta y \delta z$, centered at the point x_0, y_0, z_0 as illustrated in Fig. 1.1. Due to random molecular motions, momentum is continually imparted to the walls of the volume element by the surrounding air. This momentum transfer per unit time per unit area is just the pressure exerted on the walls of the volume element by the surrounding air. If the pressure at the center of the volume element is designated by p_0, then the pressure on the wall labeled A in Fig. 1.1 can be expressed in a Taylor series expansion as

$$p_0 + \frac{\partial p}{\partial x}\frac{\delta x}{2} + \text{higher order terms}$$

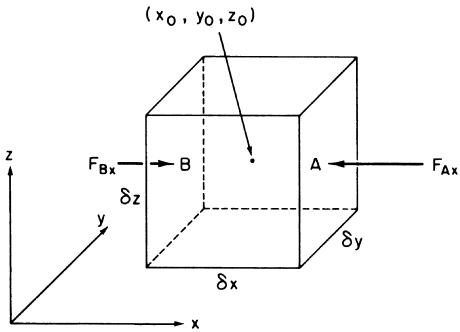


Fig. 1.1 The x component of the pressure gradient force acting on a fluid element.

Neglecting the higher order terms in this expansion, the pressure force acting on the volume element at wall A is

$$F_{Ax} = -\left(p_0 + \frac{\partial p}{\partial x}\frac{\delta x}{2}\right)\delta y\,\delta z$$

where $\delta y \delta z$ is the area of wall A. Similarly, the pressure force acting on the volume element at wall B is just

$$F_{Bx} = +\left(p_0 - \frac{\partial p}{\partial x}\frac{\delta x}{2}\right)\delta y\,\delta z$$

Therefore, the net x component of this force acting on the volume is

$$F_x = F_{Ax} + F_{Bx} = -\frac{\partial p}{\partial x}\,\delta x\,\delta y\,\delta z$$

Because the net force is proportional to the derivative of pressure in the direction of the force, it is referred to as the *pressure gradient force.*The mass m of the differential volume element is simply the density ρ times the volume: $m = \rho\delta x\delta y\delta z$. Thus, the x component of the pressure gradient force per unit mass is

$$\frac{F_x}{m} = -\frac{1}{\rho}\frac{\partial p}{\partial x}$$

Similarly, it can easily be shown that the y and z components of the pressure gradient force per unit mass are

$$\frac{F_y}{m} = -\frac{1}{\rho}\frac{\partial p}{\partial y} \qquad \text{and} \qquad \frac{F_z}{m} = -\frac{1}{\rho}\frac{\partial p}{\partial z}$$

so that the total pressure gradient force per unit mass is

$$\frac{\mathbf{F}}{m} = -\frac{1}{\rho}\,\nabla p \tag{1.1}$$

It is important to note that this force is proportional to the *gradient* of the pressure field, not to the pressure itself.

1.4.2 Gravitational Force

Newton's law of universal gravitation states that any two elements of mass in the universe attract each other with a force proportional to their masses and inversely proportional to the square of the distance separating them. Thus, if two mass elements M and m are separated by a distance $r \equiv |\mathbf{r}|$ (with the vector $\mathbf{r}$ directed toward m as shown in Fig. 1.2), then the force exerted by mass M on mass m due to gravitation is

$$\mathbf{F}_g = -\frac{GMm}{r^2}\left(\frac{\mathbf{r}}{r}\right) \tag{1.2}$$

where G is a universal constant called the gravitational constant. The law of gravitation as expressed in (1.2) actually applies only to hypothetical "point" masses since for objects of finite extent $\mathbf{r}$ will vary from one part of the object to another. However, for finite bodies, (1.2) may still be applied if $|\mathbf{r}|$ is interpreted as the distance between the centers of mass of the bodies. Thus, if the earth is designated as mass M and m is a mass element of the atmosphere, then the force per unit mass exerted on the atmosphere by the gravitational attraction of the earth is

$$\frac{\mathbf{F}_g}{m} \equiv \mathbf{g}^* = -\frac{GM}{r^2}\left(\frac{\mathbf{r}}{r}\right) \tag{1.3}$$

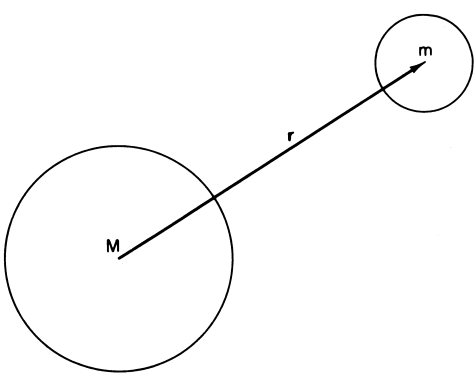


Fig. 1.2 Two spherical masses whose centers are separated by a distance r.

In dynamic meteorology it is customary to use the height above mean sea level as a vertical coordinate. If the mean radius of the earth is designated by a and the distance above mean sea level is designated by z, then neglecting the small departure of the shape of the earth from sphericity, $r = a + z$. Therefore, (1.3) can be rewritten as

$$\mathbf{g}^* = \frac{\mathbf{g}_0^*}{(1 + z/a)^2} \tag{1.4}$$

where $\mathbf{g}_0^* = -(GM/a^2)(\mathbf{r}/r)$ is the gravitational force at mean sea level. For meteorological applications, $z \ll a$ so that with negligible error we can let $\mathbf{g}^* = \mathbf{g}_0^*$ and simply treat the gravitational force as a constant.

1.4.3 Viscous Force

Any real fluid is subject to internal friction (viscosity), which causes it to resist the tendency to flow. Although a complete discussion of the resulting *viscous force* would be rather complicated, the basic physical concept can be illustrated by a simple experiment. A layer of incompressible fluid is confined between two horizontal plates separated by a distance l as shown in Fig. 1.3. The lower plate is fixed and the upper plate is placed into motion in the x direction at a speed u_0. Viscosity forces the fluid particles in the layer in contact with the plate to move at the velocity of the plate. Thus, at $z = l$ the fluid moves at speed $u(l) = u_0$, and at $z = 0$ the fluid is motionless. The force tangential to the upper plate required to keep it in uniform motion turns out to be proportional to the area of the plate, the velocity, and the inverse of the distance separating the plates. Thus, we may write $F = \mu A u_0 / l$ where μ is a constant of proportionality, the *dynamic viscosity coefficient*.

This force must just equal the force exerted by the upper plate on the fluid immediately below it. For a state of uniform motion, every horizontal layer of fluid of depth δz must exert the same force F on the fluid below. This may be

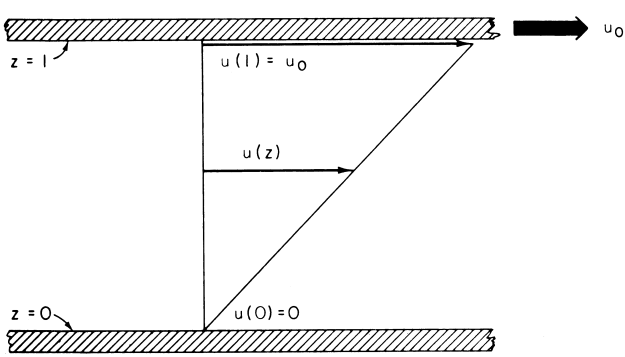


Fig. 1.3 One-dimensional steady-state viscous shear flow.

expressed in the form $F = \mu A \delta u / \delta z$ where $\delta u = u_0 \delta z / l$ is the velocity shear across the layer δz. The viscous force per unit area, or *shearing stress*, can then be defined as

$$\tau_{zx} = \lim_{\delta z \to 0} \mu \frac{\delta u}{\delta z} = \mu \frac{\partial u}{\partial z}$$

where subscripts indicate that τ_{zx} is the component of the shearing stress in the x direction due to vertical shear of the x velocity component.

From the molecular viewpoint, this shearing stress results from a net downward transport of momentum by the random motion of the molecules. Because the mean x momentum increases with height, molecules passing downward through a horizontal plane at any instant carry more momentum than those passing upward through the plane. Thus, there is a net downward transport of x momentum. This downward momentum transport per unit time per unit area is simply the shearing stress.

In a similar fashion, random molecular motions will transport heat down a mean temperature gradient and trace constituents down mean mixing ratio gradients. In these cases the transport is referred to as *molecular diffusion*. Molecular diffusion always acts to reduce irregularities in the field being diffused.

In the simple two-dimensional steady-state motion example given above there is no net viscous force acting on the elements of fluid, as the shearing stress acting across the top boundary of each fluid element is just equal and opposite to that acting across the lower boundary. For the more general case of nonsteady two-dimensional shear flow in an incompressible fluid, we may calculate the viscous force by again considering a differential volume element of fluid centered at (x, y, z) with sides $\delta x \delta y \delta z$ as shown in Fig. 1.4. If the shearing stress in the x direction acting through the center of the element is designated τ_{zx}, then the stress acting across the upper boundary on the fluid *below* may be written approximately as

$$\tau_{zx} + \frac{\partial \tau_{zx}}{\partial z} \frac{\delta z}{2}$$

while the stress acting across the lower boundary on the fluid *above* is

$$-\left[\tau_{zx} - \frac{\partial \tau_{zx}}{\partial z} \frac{\delta z}{2}\right]$$

(This is just equal and opposite to the stress acting across the lower boundary on the fluid *below*.) The net viscous force on the volume element acting in the x direction is then given by the sum of the stresses acting across the upper boundary on the fluid below and across the lower boundary on the fluid above:

$$\left(\tau_{zx} + \frac{\partial \tau_{zx}}{\partial z} \frac{\delta z}{2}\right) \delta x \, \delta y - \left(\tau_{zx} - \frac{\partial \tau_{zx}}{\partial z} \frac{\delta z}{2}\right) \delta x \, \delta y$$

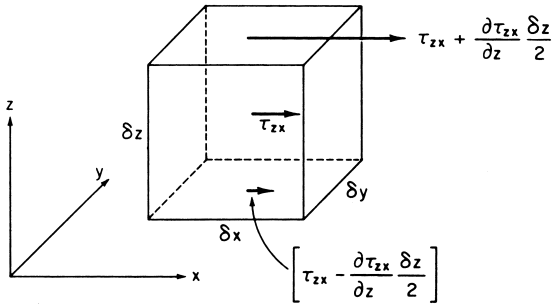


Fig. 1.4 The x component of the vertical shearing stress on a fluid element.

Dividing this expression by the mass $\rho\delta x\delta y\delta z$, we find that the viscous force per unit mass due to vertical shear of the component of motion in the x direction is

$$\frac{1}{\rho}\frac{\partial \tau_{zx}}{\partial z} = \frac{1}{\rho}\frac{\partial}{\partial z}\left(\mu\frac{\partial u}{\partial z}\right)$$

For constant μ, the right-hand side just given above may be simplified to $\nu\partial^2 u/\partial z^2$, where $\nu = \mu/\rho$ is the *kinematic viscosity coefficient*. For standard atmosphere conditions at sea level, $\nu = 1.46 \times 10^{-5}\,\mathrm{m}^2\,s^{-1}$. Derivations analogous to that shown in Fig. 1.4 can be carried out for viscous stresses acting in other directions. The resulting frictional force components per unit mass in the three Cartesian coordinate directions are

$$\begin{aligned} F_{rx} &= \nu\left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right] \\ F_{ry} &= \nu\left[\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right] \\ F_{rz} &= \nu\left[\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right] \end{aligned} \tag{1.5}$$

For the atmosphere below 100 km, ν is so small that molecular viscosity is negligible except in a thin layer within a few centimeters of the earth's surface where the vertical shear is very large. Away from this surface molecular boundary layer, momentum is transferred primarily by turbulent eddy motions. These are discussed in Chapter 5.

1.5 NONINERTIAL REFERENCE FRAMES AND "APPARENT" FORCES

In formulating the laws of atmospheric dynamics it is natural to use a *geocentric* reference frame, that is, a frame of reference at rest with respect to the rotating earth. Newton's first law of motion states that a mass in uniform motion relative to

a coordinate system fixed in space will remain in uniform motion in the absence of any forces. Such motion is referred to as *inertial motion*; and the fixed reference frame is an inertial, or absolute, frame of reference. It is clear, however, that an object at rest or in uniform motion with respect to the rotating earth is not at rest or in uniform motion relative to a coordinate system fixed in space. Therefore, motion that appears to be inertial motion to an observer in a geocentric reference frame is really accelerated motion. Hence, a geocentric reference frame is a *noninertial* reference frame. Newton's laws of motion can only be applied in such a frame if the acceleration of the coordinates is taken into account. The most satisfactory way of including the effects of coordinate acceleration is to introduce "apparent" forces in the statement of Newton's second law. These apparent forces are the inertial reaction terms that arise because of the coordinate acceleration. For a coordinate system in uniform rotation, two such apparent forces are required: the centrifugal force and the Coriolis force.

1.5.1 Centripetal Acceleration and Centrifugal Force

A ball of mass m is attached to a string and whirled through a circle of radius r at a constant angular velocity ω. From the point of view of an observer in inertial space the speed of the ball is constant, but its direction of travel is continuously changing so that its velocity is not constant. To compute the acceleration we consider the change in velocity $\delta\mathbf{V}$ that occurs for a time increment δt during which the ball rotates through an angle $\delta\theta$ as shown in Fig. 1.5. Because $\delta\theta$ is also the angle between the vectors $\mathbf{V}$ and $\mathbf{V} + \delta\mathbf{V}$, the magnitude of $\delta\mathbf{V}$ is just $|\delta\mathbf{V}| = |\mathbf{V}|\,\delta\theta$. If we divide by δt and note that in the limit $\delta t \to 0$, $\delta\mathbf{V}$ is directed toward the axis of rotation, we obtain

$$\frac{D\mathbf{V}}{Dt} = |\mathbf{V}|\frac{D\theta}{Dt}\left(-\frac{\mathbf{r}}{r}\right)$$

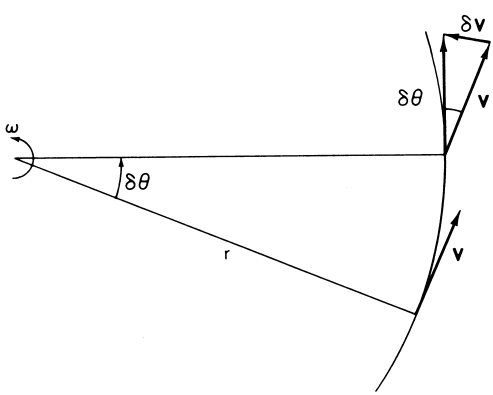


Fig. 1.5 Centripetal acceleration is given by the rate of change of the direction of the velocity vector, which is directed toward the axis of rotation, as illustrated here by $\delta\mathbf{V}$.

However, $|\mathbf{V}| = \omega r$ and $D\theta/Dt = \omega$, so that

$$\frac{D\mathbf{V}}{Dt} = -\omega^2 r \tag{1.6}$$

Therefore, viewed from fixed coordinates the motion is one of uniform acceleration directed toward the axis of rotation and equal to the square of the angular velocity times the distance from the axis of rotation. This acceleration is called *centripetal acceleration*. It is caused by the force of the string pulling the ball.

Now suppose that we observe the motion in a coordinate system rotating with the ball. In this rotating system the ball is stationary, but there is still a force acting on the ball, namely the pull of the string. Therefore, in order to apply Newton's second law to describe the motion relative to this rotating coordinate system, we must include an additional apparent force, the *centrifugal force*, which just balances the force of the string on the ball. Thus, the centrifugal force is equivalent to the inertial reaction of the ball on the string and just equal and opposite to the centripetal acceleration.

To summarize, observed from a fixed system the rotating ball undergoes a uniform centripetal acceleration in response to the force exerted by the string. Observed from a system rotating along with it, the ball is stationary and the force exerted by the string is balanced by a centrifugal force.

1.5.2 Gravity Force

An object at rest on the surface of the earth is not at rest or in uniform motion relative to an inertial reference frame except at the poles. Rather, an object of unit mass at rest on the surface of the earth is subject to a centripetal acceleration directed toward the axis of rotation of the earth given by $-\Omega^2\mathbf{R}$, where $\mathbf{R}$ is the position vector from the axis of rotation to the object and $\Omega = 7.292 \times 10^{-5}\,\mathrm{rad\,s^{-1}}$ is the angular speed of rotation of the earth.[3] Since except at the equator and poles the centripetal acceleration has a component directed poleward along the horizontal surface of the earth (i.e., along a surface of constant *geopotential*), there must be a net horizontal force directed poleward along the horizontal to sustain the horizontal component of the centripetal acceleration. This force arises because the rotating earth is not a sphere, but has assumed the shape of an oblate spheroid in which there is a poleward component of gravitation along a constant geopotential surface just sufficient to account for the poleward component of the centripetal acceleration at each latitude for an object at rest on the surface of the earth. In other words, from the point of view of an observer in an inertial reference frame, geopotential

[3] The earth revolves around its axis once every *sidereal* day, which is equal to 23 h 56 min 4 s (86,164 s). Thus, $\Omega = 2\pi/(86,164\,s) = 7.292 \times 10^{-5}\,\mathrm{rad\,s^{-1}}$.

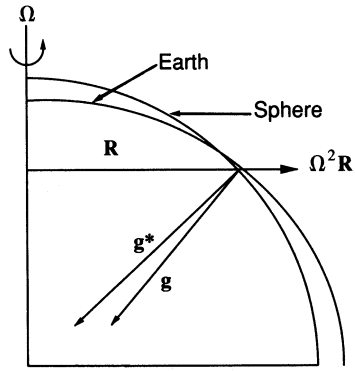


Fig. 1.6 Relationship between the true gravitation vector **g*** and gravity **g**. For an idealized homogeneous spherical earth, **g*** would be directed toward the center of the earth. In reality, **g*** does not point exactly to the center except at the equator and the poles. Gravity, **g**, is the vector sum of **g*** and the centrifugal force and is perpendicular to the level surface of the earth, which approximates an oblate spheroid.

surfaces slope upward toward the equator (see Fig. 1.6). As a consequence, the equatorial radius of the earth is about 21 km larger than the polar radius.

Viewed from a frame of reference rotating with the earth, however, a geopotential surface is everywhere normal to the sum of the true force of gravity, $\mathbf{g}^*$, and the centrifugal force $\Omega^2\mathbf{R}$ (which is just the reaction force of the centripetal acceleration). A geopotential surface is thus experienced as a level surface by an object at rest on the rotating earth. Except at the poles, the weight of an object of mass m at rest on such a surface, which is just the reaction force of the earth on the object, will be slightly less than the gravitational force $m\mathbf{g}^*$ because, as illustrated in Fig. 1.6, the centrifugal force partly balances the gravitational force. It is, therefore, convenient to combine the effects of the gravitational force and centrifugal force by defining *gravity* **g** such that

$$\mathbf{g} \equiv -g\mathbf{k} \equiv \mathbf{g}^* + \Omega^2\mathbf{R} \tag{1.7}$$

where **k** designates a unit vector parallel to the local vertical. Gravity, g, sometimes referred to as "apparent gravity," will here be taken as a constant ($g = 9.81\ m\ s^{-2}$). Except at the poles and the equator, **g** is not directed toward the center of the earth, but is perpendicular to a geopotential surface as indicated by Fig. 1.6. True gravity $\mathbf{g}^*$, however, is not perpendicular to a geopotential surface, but has a horizontal component just large enough to balance the horizontal component of $\Omega^2\mathbf{R}$.

Gravity can be represented in terms of the gradient of a potential function Φ, which is just the geopotential referred to above:

$$\nabla\Phi = -\mathbf{g}$$

However, because $\mathbf{g} = -g\mathbf{k}$ where $g \equiv |\mathbf{g}|$, it is clear that $\Phi = \Phi(z)$ and $d\Phi/dz = g$. Thus horizontal surfaces on the earth are surfaces of constant geopotential. If the value of geopotential is set to zero at mean sea level, the geopotential

$\Phi(z)$ at height z is just the work required to raise a unit mass to height z from mean sea level:

$$\Phi = \int_0^z g dz \tag{1.8}$$

Despite the fact that the surface of the earth bulges toward the equator, an object at rest on the surface of the rotating earth does not slide "downhill" toward the pole because, as indicated above, the poleward component of gravitation is balanced by the equatorward component of the centrifugal force. However, if the object is put into motion relative to the earth, this balance will be disrupted. Consider a frictionless object located initially at the North pole. Such an object has zero angular momentum about the axis of the earth. If it is displaced away from the pole in the absence of a zonal torque, it will not acquire rotation and hence will feel a restoring force due to the horizontal component of true gravity, which, as indicated above is equal and opposite to the horizontal component of the centrifugal force for an object at rest on the surface of the earth. Letting R be the distance from the pole, the horizontal restoring force for a small displacement is thus $-\Omega^2 R$, and the object's acceleration viewed in the inertial coordinate system satisfies the equation for a simple harmonic oscillator:

$$\frac{d^2 R}{dt^2} + \Omega^2 R = 0 \tag{1.9}$$

The object will undergo an oscillation of period $2\pi/\Omega$ along a path that will appear as a straight line passing through the pole to an observer in a fixed coordinate system, but will appear as a closed circle traversed in 1/2 day to an observer rotating with the earth (Fig. 1.7). From the point of view of an earthbound observer, there is an apparent deflection force that causes the object to deviate to the right of its direction of motion at a fixed rate.

1.5.3 The Coriolis Force and the Curvature Effect

Newton's second law of motion expressed in coordinates rotating with the earth can be used to describe the force balance for an object at rest on the surface of the earth, provided that an apparent force, the centrifugal force, is included among the forces acting on the object. If, however, the object is in motion along the surface of the earth, additional apparent forces are required in the statement of Newton's second law.

Suppose that an object of unit mass, initially at latitude ϕ moving zonally at speed u, relative to the surface of the earth, is displaced in latitude or in altitude by an impulsive force. As the object is displaced it will conserve its angular momentum in the absence of a torque in the east–west direction. Because the distance R to the axis of rotation changes for a displacement in latitude or altitude, the absolute angular velocity, $\Omega + u/R$, must change if the object is to conserve its absolute

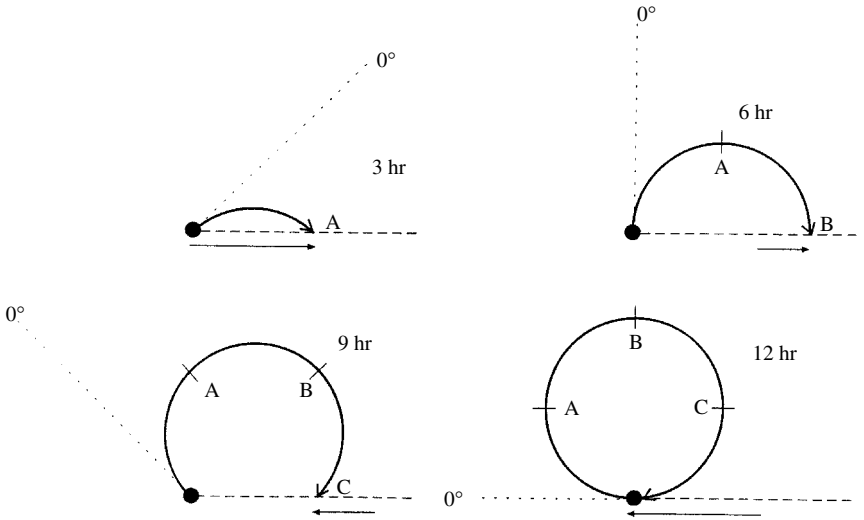


Fig. 1.7 Motion of a frictionless object launched from the north pole along the 0° longitude meridian at $t = 0$, as viewed in fixed and rotating reference frames at 3, 6, 9, and 12 h after launch. The horizontal dashed line marks the position that the 0° longitude meridian had at $t = 0$, and short dashed lines show its position in the fixed reference frame at subsequent 3 h intervals. Horizontal arrows show 3 h displacement vectors as seen by an observer in the fixed reference frame. Heavy curved arrows show the trajectory of the object as viewed by an observer in the rotating system. Labels A, B and C show the position of the object relative to the rotating coordinates at 3 h intervals. In the fixed coordinate frame the object oscillates back and forth along a straight line under the influence of the restoring force provided by the horizontal component of gravitation. The period for a complete oscillation is 24 h (only 1/2 period is shown) . To an observer in rotating coordinates, however, the motion appears to be at constant speed and describes a complete circle in a clockwise direction in 12 h.

angular momentum. Because Ω is constant, the relative zonal velocity must change. Thus, the object behaves as though a zonally directed deflection force were acting on it.

The form of the zonal deflection force can be obtained by equating the total angular momentum at the initial distance R to the total angular momentum at the displaced distance $R + \delta R$:

$$\left(\Omega + \frac{u}{R}\right) R^2 = \left(\Omega + \frac{u + \delta u}{R + \delta R}\right) (R + \delta R)^2$$

where δu is the change in eastward relative velocity after displacement. Expanding the right-hand side, neglecting second-order differentials, and solving for δu gives

$$\delta u = -2\Omega \delta R - \frac{u}{R} \delta R$$

Noting that $R = a\cos\phi$, where a is the radius of the earth and ϕ is latitude, dividing through by the time increment δt and taking the limit as $\delta t \to 0$, gives in the case of a meridional displacement in which $\delta R = -\sin\phi\delta y$ (see Fig. 1.8):

$$\left(\frac{Du}{Dt}\right) = \left(2\Omega\sin\phi + \frac{u}{a}\tan\phi\right)\frac{Dy}{Dt} = 2\Omega v\sin\phi + \frac{uv}{a}\tan\phi \tag{1.10a}$$

and for a vertical displacement in which $\delta R = +\cos\phi\delta z$:

$$\left(\frac{Du}{Dt}\right) = -\left(2\Omega\cos\phi + \frac{u}{a}\right)\frac{Dz}{Dt} = -2\Omega w\cos\phi - \frac{uw}{a} \tag{1.10b}$$

where $v = Dy/Ddt$ and $w = Dz/Dt$ are the northward and upward velocity components, respectively. The first terms on the right in (1.10a) and (1.10b) are the zonal components of the *Coriolis force* for meridional and vertical motions, respectively. The second terms on the right are referred to as *metric terms* or *curvature effects*. These arise from the curvature of the earth's surface.

A similar argument can be used to obtain the meridional component of the Coriolis force. Suppose now that the object is set in motion in the eastward direction by an impulsive force. Because the object is now rotating faster than the earth, the centrifugal force on the object will be increased. Letting **R** be the position vector

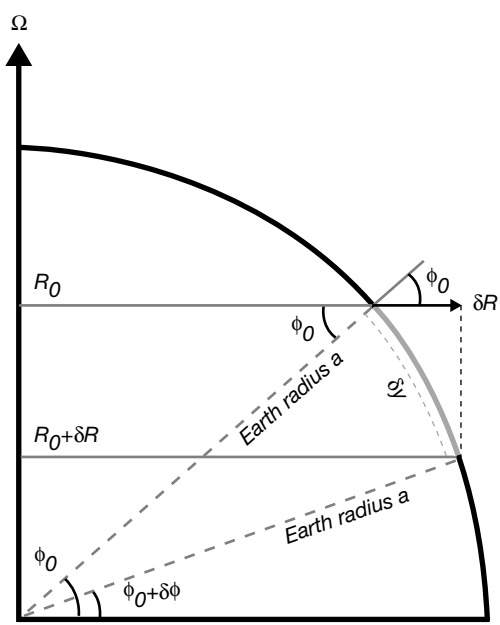


Fig. 1.8 Relationship of δR and $\delta y = a\delta\phi$ for an equatorward displacement.

from the axis of rotation to the object, the excess of the centrifugal force over that for an object at rest is

$$\left(\Omega + \frac{u}{R}\right)^2 \mathbf{R} - \Omega^2 \mathbf{R} = \frac{2\Omega u \mathbf{R}}{R} + \frac{u^2 \mathbf{R}}{R^2}$$

The terms on the right represent *deflecting* forces, which act outward along the vector **R** (i.e., perpendicular to the axis of rotation). The meridional and vertical components of these forces are obtained by taking meridional and vertical components of **R** as shown in Fig. 1.9 to yield

$$\left(\frac{Dv}{Dt}\right) = -2\Omega u \sin\phi - \frac{u^2}{a}\tan\phi \tag{1.11a}$$

$$\left(\frac{Dw}{Dt}\right) = 2\Omega u \cos\phi + \frac{u^2}{a} \tag{1.11b}$$

The first terms on the right are the meridional and vertical components, respectively, of the Coriolis forces for zonal motion; the second terms on the right are again the curvature effects.

For synoptic scale motions $|u| \ll \Omega R$, the last terms in (1.10a) and (1.11a) can be neglected in a first approximation. Therefore, relative horizontal motion produces a horizontal acceleration perpendicular to the direction of motion given by

$$\left(\frac{Du}{Dt}\right)_{Co} = 2\Omega v \sin\phi = fv \tag{1.12a}$$

$$\left(\frac{Dv}{Dt}\right)_{Co} = -2\Omega u \sin\phi = -fu \tag{1.12b}$$

where $f \equiv 2\Omega \sin\phi$ is the *Coriolis parameter*.

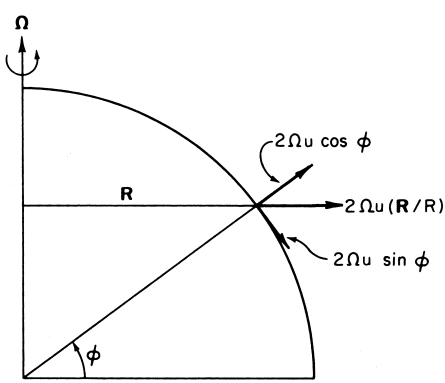


Fig. 1.9 Components of the Coriolis force due to relative motion along a latitude circle.

The subscript Co indicates that the acceleration is the part of the total acceleration due only to the Coriolis force. Thus, for example, an object moving eastward in the horizontal is deflected equatorward by the Coriolis force, whereas a westward moving object is deflected poleward. In either case the deflection is to the right of the direction of motion in the Northern Hemisphere and to the left in the Southern Hemisphere. The vertical component of the Coriolis force in (1.11b) is ordinarily much smaller than the gravitational force so that its only effect is to cause a very minor change in the apparent weight of an object depending on whether the object is moving eastward or westward.

The Coriolis force is negligible for motions with time scales that are very short compared to the period of the earth's rotation (a point that is illustrated by several problems at the end of the chapter). Thus, the Coriolis force is not important for the dynamics of individual cumulus clouds, but is essential to the understanding of longer time scale phenomena such as synoptic scale systems. The Coriolis force must also be taken into account when computing long-range missile or artillery trajectories.

As an example, suppose that a ballistic missile is fired due eastward at 43°N latitude ($f = 10^{-4}\,s^{-1}$ at 43°N). If the missile travels 1000 km at a horizontal speed $u_0 = 1000\,m\,s^{-1}$, by how much is the missile deflected from its eastward path by the Coriolis force? Integrating (1.12b) with respect to time we find that

$$v = -fu_0 t \tag{1.13}$$

where it is assumed that the deflection is sufficiently small so that we may let f and u_0 be constants. To find the total displacement we must integrate (1.13) with respect to time:

$$\int_0^t v dt = \int_{y_0}^{y_0+\delta y} dy = -fu_0 \int_0^t t dt$$

Thus, the total displacement is

$$\delta y = -fu_0 t^2/2 = -50\,\text{km}$$

Therefore, the missile is deflected southward by 50 km due to the Coriolis effect. Further examples of the deflection of objects by the Coriolis force are given in some of the problems at the end of the chapter.

The x and y components given in (1.12a) and (1.12b) can be combined in vector form as

$$\left(\frac{D\mathbf{V}}{Dt}\right)_{Co} = -f\mathbf{k} \times \mathbf{V} \tag{1.14}$$

where $\mathbf{V} \equiv (u, v)$ is the horizontal velocity, $\mathbf{k}$ is a vertical unit vector, and the subscript Co indicates that the acceleration is due solely to the Coriolis force. Since $-\mathbf{k} \times \mathbf{V}$ is a vector rotated 90° to the right of $\mathbf{V}$, (1.14) clearly shows the

deflection character of the Coriolis force. The Coriolis force can only change the direction of motion, not the speed of motion.

1.5.4 Constant Angular Momentum Oscillations

Suppose an object initially at rest on the earth at the point (x_0, y_0) is impulsively propelled along the x axis with a speed V at time $t = 0$. Then from (1.12a) and (1.12b), the time evolution of the velocity is given by $u = V \cos ft$ and $v = -V \sin ft$. However, because $u = Dx/Dt$ and $v = Dy/Dt$, integration with respect to time gives the position of the object at time t as

$$x - x_0 = \frac{V}{f} \sin ft \quad \text{and} \quad y - y_0 = \frac{V}{f} (\cos ft - 1) \tag{1.15a,b}$$

where the variation of f with latitude is here neglected. Equations (1.15a) and (1.15b) show that in the Northern Hemisphere, where f is positive, the object orbits clockwise (anticyclonically) in a circle of radius $R = V/f$ about the point $(x_0, y_0 - V/f)$ with a period given by

$$\tau = 2\pi R/V = 2\pi/f = \pi/(\Omega \sin \phi) \tag{1.16}$$

Thus, an object displaced horizontally from its equilibrium position on the surface of the earth under the influence of the force of gravity will oscillate about its equilibrium position with a period that depends on latitude and is equal to one sidereal day at 30° latitude and 1/2 sidereal day at the pole. Constant angular momentum oscillations (often referred to misleadingly as "inertial oscillations") are commonly observed in the oceans, but are apparently not of importance in the atmosphere.

1.6 STRUCTURE OF THE STATIC ATMOSPHERE

The thermodynamic state of the atmosphere at any point is determined by the values of pressure, temperature, and density (or specific volume) at that point. These field variables are related to each other by the equation of state for an ideal gas. Letting p, T ρ, and $\alpha (\equiv \rho^{-1})$ denote pressure, temperature, density, and specific volume, respectively, we can express the equation of state for dry air as

$$p\alpha = RT \quad or \quad p = \rho RT \tag{1.17}$$

where R is the gas constant for dry air ($R = 287\,\mathrm{J\,kg^{-1}\,K^{-1}}$).

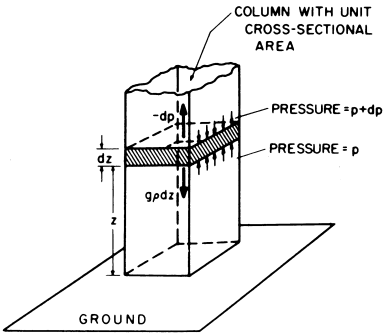


Fig. 1.10 Balance of forces for hydrostatic equilibrium. Small arrows show the upward and downward forces exerted by air pressure on the air mass in the shaded block. The downward force exerted by gravity on the air in the block is given by $\rho g dz$, whereas the net pressure force given by the difference between the upward force across the lower surface and the downward force across the upper surface is $-dp$. Note that dp is negative, as pressure decreases with height. (After Wallace and Hobbs, 1977.)

1.6.1 The Hydrostatic Equation

In the absence of atmospheric motions the gravity force must be exactly balanced by the vertical component of the pressure gradient force. Thus, as illustrated in Fig. 1.10,

$$dp/dz = -\rho g \tag{1.18}$$

This condition of *hydrostatic balance* provides an excellent approximation for the vertical dependence of the pressure field in the real atmosphere. Only for intense small-scale systems such as squall lines and tornadoes is it necessary to consider departures from hydrostatic balance. Integrating (1.18) from a height z to the top of the atmosphere we find that

$$p(z) = \int_z^\infty \rho g dz \tag{1.19}$$

so that the pressure at any point is simply equal to the weight of the unit cross section column of air overlying the point. Thus, mean sea level pressure $p(0) =$ 1013.25 hPa is simply the average weight per square meter of the total atmospheric column.[4] It is often useful to express the hydrostatic equation in terms of the geopotential rather than the geometric height. Noting from (1.8) that $d\Phi = g\,dz$ and from, (1.17) that $\alpha = RT/p$, we can express the hydrostatic equation in the form

$$gdz = d\Phi = -(RT/p)dp = -RTd\ln p \tag{1.20}$$

Thus, the variation of geopotential with respect to pressure depends only on temperature. Integration of (1.20) in the vertical yields a form of the *hypsometric equation*:

$$\Phi(z_2) - \Phi(z_1) = g_0(Z_2 - Z_1) = R\int_{p_2}^{p_1} Td\ln p \tag{1.21}$$

[4] For computational convenience, the mean surface pressure is often assumed to equal 1000 hPa.

Here $Z \equiv \Phi(z)/g_0$, is the *geopotential height*, where $g_0 \equiv 9.80665\,\mathrm{m\,s^{-2}}$ is the global average of gravity at mean sea level. Thus in the troposphere and lower stratosphere, Z is numerically almost identical to the geometric height z. In terms of Z the hypsometric equation becomes

$$Z_T \equiv Z_2 - Z_1 = \frac{R}{g_0}\int_{p_2}^{p_1} T\,d\ln p \tag{1.22}$$

where Z_T is the *thickness* of the atmospheric layer between the pressure surfaces p_2 and p_1. Defining a layer mean temperature

$$\langle T\rangle = \int_{p_2}^{p_1} T\,d\ln p\left[\int_{p_2}^{p_1} d\ln p\right]^{-1}$$

and a layer mean scale height $H \equiv R\langle T\rangle/g_0$ we have from (1.22)

$$Z_T = H\ln(p_1/p_2) \tag{1.23}$$

Thus the thickness of a layer bounded by isobaric surfaces is proportional to the mean temperature of the layer. Pressure decreases more rapidly with height in a cold layer than in a warm layer. It also follows immediately from (1.23) that in an isothermal atmosphere of temperature T, the geopotential height is proportional to the natural logarithm of pressure normalized by the surface pressure,

$$Z = -H\ln(p/p_0) \tag{1.24}$$

where p_0 is the pressure at $Z = 0$. Thus, in an isothermal atmosphere the pressure decreases exponentially with geopotential height by a factor of e^{-1} per scale height,

$$p(Z) = p(0)e^{-Z/H}$$

1.6.2 Pressure as a Vertical Coordinate

From the hydrostatic equation (1.18), it is clear that a single valued monotonic relationship exists between pressure and height in each vertical column of the atmosphere. Thus we may use pressure as the independent vertical coordinate and height (or geopotential) as a dependent variable. The thermodynamic state of the atmosphere is then specified by the fields of $\Phi(x, y, p, t)$ and $T(x, y, p, t)$.

Now the horizontal components of the pressure gradient force given by (1.1) are evaluated by partial differentiation holding z constant. However, when pressure is used as the vertical coordinate, horizontal partial derivatives must be

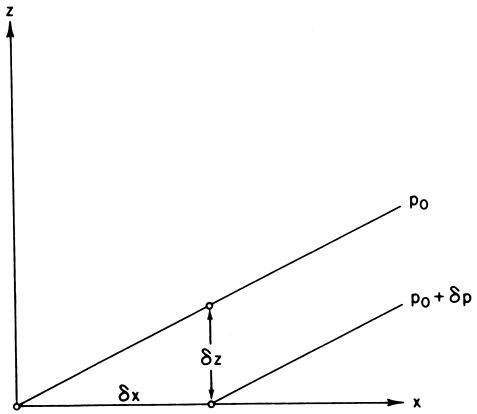


Fig. 1.11 Slope of pressure surfaces in the x, z plane.

evaluated holding p constant. Transformation of the horizontal pressure gradient force from height to pressure coordinates may be carried out with the aid of Fig. 1.11. Considering only the x, z plane, we see from Fig. 1.11 that

$$\left[\frac{(p_0 + \delta p) - p_0}{\delta x}\right]_z = \left[\frac{(p_0 + \delta p) - p_0}{\delta z}\right]_x \left(\frac{\delta z}{\delta x}\right)_p$$

where subscripts indicate variables that remain constant in evaluating the differentials. Thus, for example, in the limit $\delta z \to 0$

$$\left[\frac{(p_0 + \delta p) - p_0}{\delta z}\right]_x \to \left(-\frac{\partial p}{\partial z}\right)_x$$

where the minus sign is included because $\delta z < 0$ for $\delta p > 0$.

Taking the limits $\delta x, \delta z \to 0$ we obtain[5]

$$\left(\frac{\partial p}{\partial x}\right)_z = -\left(\frac{\partial p}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_p$$

which after substitution from the hydrostatic equation (1.18) yields

$$-\frac{1}{\rho}\left(\frac{\partial p}{\partial x}\right)_z = -g\left(\frac{\partial z}{\partial x}\right)_p = -\left(\frac{\partial \Phi}{\partial x}\right)_p \tag{1.25}$$

Similarly, it is easy to show that

$$-\frac{1}{\rho}\left(\frac{\partial p}{\partial y}\right)_z = -\left(\frac{\partial \Phi}{\partial y}\right)_p \tag{1.26}$$

[5] It is important to note the minus sign on the right in this expression!

Thus in the *isobaric* coordinate system the horizontal pressure gradient force is measured by the gradient of geopotential at constant pressure. Density no longer appears explicitly in the pressure gradient force; this is a distinct advantage of the isobaric system.

1.6.3 A Generalized Vertical Coordinate

Any single-valued monotonic function of pressure or height may be used as the independent vertical coordinate. For example, in many numerical weather prediction models, pressure normalized by the pressure at the ground $[\sigma \equiv p(x, y, z, t)/p_s(x, y, t)]$ is used as a vertical coordinate. This choice guarantees that the ground is a coordinate surface ($\sigma \equiv 1$) even in the presence of spatial and temporal surface pressure variations. Thus, this so-called σ coordinate system is particularly useful in regions of strong topographic variations.

We now obtain a general expression for the horizontal pressure gradient, which is applicable to any vertical coordinate $s = s(x, y, z, t)$ that is a single-valued monotonic function of height. Referring to Fig. 1.12 we see that for a horizontal distance δx the pressure difference evaluated along a surface of constant s is related to that evaluated at constant z by the relationship

$$\frac{p_C - p_A}{\delta x} = \frac{p_C - p_B}{\delta z}\,\frac{\delta z}{\delta x} + \frac{p_B - p_A}{\delta x}$$

Taking the limits as $\delta x, \delta z \to 0$ we obtain

$$\left(\frac{\partial p}{\partial x}\right)_s = \frac{\partial p}{\partial z}\left(\frac{\partial z}{\partial x}\right)_s + \left(\frac{\partial p}{\partial x}\right)_z \tag{1.27}$$

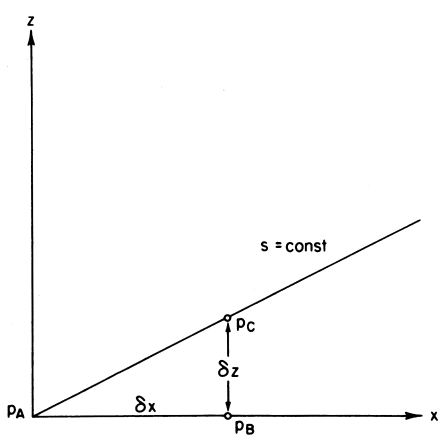


Fig. 1.12 Transformation of the pressure gradient force to s coordinates.

Using the identity $\partial p/\partial z = (\partial s/\partial z)(\partial p/\partial s)$, we can express (1.27) in the alternate form

$$\left(\frac{\partial p}{\partial x}\right)_s = \left(\frac{\partial p}{\partial x}\right)_z + \frac{\partial s}{\partial z}\left(\frac{\partial z}{\partial x}\right)_s \left(\frac{\partial p}{\partial s}\right) \tag{1.28}$$

In later chapters we will apply (1.27) or (1.28) and similar expressions for other fields to transform the dynamical equations to several different vertical coordinate systems.

PROBLEMS

1.1. Neglecting the latitudinal variation in the radius of the earth, calculate the angle between the gravitational force and gravity vectors at the surface of the earth as a function of latitude. What is the maximum value of this angle?

1.2. Calculate the altitude at which an artificial satellite orbiting in the equatorial plane can be a synchronous satellite (i.e., can remain above the same spot on the surface of the earth).

1.3. An artificial satellite is placed into a natural synchronous orbit above the equator and is attached to the earth below by a wire. A second satellite is attached to the first by a wire of the same length and is placed in orbit directly above the first at the same angular velocity. Assuming that the wires have zero mass, calculate the tension in the wires per unit mass of satellite. Could this tension be used to lift objects into orbit with no additional expenditure of energy?

1.4. A train is running smoothly along a curved track at the rate of 50 m s^{-1}. A passenger standing on a set of scales observes that his weight is 10% greater than when the train is at rest. The track is banked so that the force acting on the passenger is normal to the floor of the train. What is the radius of curvature of the track?

1.5. If a baseball player throws a ball a horizontal distance of 100 m at 30° latitude in 4 s, by how much is it deflected laterally as a result of the rotation of the earth?

1.6. Two balls 4 cm in diameter are placed 100 m apart on a frictionless horizontal plane at 43°N. If the balls are impulsively propelled directly at each other with equal speeds, at what speed must they travel so that they just miss each other?

1.7. A locomotive of 2×10^5 kg mass travels 50 m s^{-1} along a straight horizontal track at 43°N. What lateral force is exerted on the rails? Compare the magnitudes of the upward reaction force exerted by the rails for cases where the locomotive is traveling eastward and westward, respectively.

1.8. Find the horizontal displacement of a body dropped from a fixed platform at a height h at the equator neglecting the effects of air resistance. What is the numerical value of the displacement for $h = 5$ km?

1.9. A bullet is fired directly upward with initial speed w_0, at latitude ϕ. Neglecting air resistance, by what distance will it be displaced horizontally when it returns to the ground? (Neglect $2\Omega u \cos\phi$ compared to g in the vertical momentum equation.)

1.10. A block of mass $M = 1$ kg is suspended from the end of a weightless string. The other end of the string is passed through a small hole in a horizontal platform and a ball of mass $m = 10$ kg is attached. At what angular velocity must the ball rotate on the horizontal platform to balance the weight of the block if the horizontal distance of the ball from the hole is 1 m? While the ball is rotating, the block is pulled down 10 cm. What is the new angular velocity of the ball? How much work is done in pulling down the block?

1.11. A particle is free to slide on a horizontal frictionless plane located at a latitude ϕ on the earth. Find the equation governing the path of the particle if it is given an impulsive northward velocity $v = V_0$ at $t = 0$. Give the solution for the position of the particle as a function of time. (Assume that the latitudinal excursion is sufficiently small that f is constant.)

1.12. Calculate the 1000- to 500-hPa thickness for isothermal conditions with temperatures of 273- and 250 K, respectively.

1.13. Isolines of 1000- to 500-hPa thickness are drawn on a weather map using a contour interval of 60 m. What is the corresponding layer mean temperature interval?

1.14. Show that a homogeneous atmosphere (density independent of height) has a finite height that depends only on the temperature at the lower boundary. Compute the height of a homogeneous atmosphere with surface temperature $T_0 = 273$K and surface pressure 1000 hPa. (Use the ideal gas law and hydrostatic balance.)

1.15. For the conditions of Problem 1.14, compute the variation of the temperature with respect to height.

1.16. Show that in an atmosphere with uniform *lapse rate* γ (where $\gamma \equiv -dT/dz$) the geopotential height at pressure level p_1 is given by

$$Z = \frac{T_0}{\gamma}\left[1 - \left(\frac{p_0}{p_1}\right)^{-R\gamma/g}\right]$$

where T_0 and p_0 are the sea level temperature and pressure, respectively.

1.17. Calculate the 1000- to 500-hPa thickness for a constant lapse rate atmosphere with $\gamma = 6.5\text{K km}^{-1}$ and $T_0 = 273\text{K}$. Compare your results with the results in Problem 1.12.

1.18. Derive an expression for the variation of density with respect to height in a constant lapse rate atmosphere.

1.19. Derive an expression for the altitude variation of the pressure change δp that occurs when an atmosphere with a constant lapse rate is subjected to a height-independent temperature change δT while the surface pressure remains constant. At what height is the magnitude of the pressure change a maximum if the lapse rate is 6.5 K km^{-1}, $T_0 = 300$, and $\delta T = 2\text{K}$?

MATLAB EXERCISES

M1.1. This exercise investigates the role of the curvature terms for high-latitude constant angular momentum trajectories.

(a) Run the **coriolis.m** script with the following initial conditions: initial latitude 60°, initial velocity $u = 0$, $v = 40$ m s^{-1}, run time = 5 days. Compare the appearance of the trajectories for the case with the curvature terms included and the case with curvature terms neglected. Qualitatively explain the difference that you observe. Why is the trajectory not a closed circle as described in Eq. (1.15) of the text? [*Hint*: consider the separate effects of the term proportional to tan ϕ and of the spherical geometry.]

(b) Run **coriolis.m** with latitude 60°, $u = 0$, $v = 80$ m/s. What is different from case (a)? By varying the run time, see if you can determine how long it takes for the particle to make a full circuit in each case and compare this to the time given in Eq. (1.16) for $\phi = 60°$.

M1.2. Using the MATLAB script from Problem M1.1, compare the magnitudes of the lateral deflection for ballistic missiles fired eastward and westward at 43° latitude. Each missile is launched at a velocity of 1000 m s^{-1}and travels 1000 km. Explain your results. Can the curvature term be neglected in these cases?

M1.3. This exercise examines the strange behavior of constant angular momentum trajectories near the equator. Run the **coriolis.m** script for the following contrasting cases: a) latitude 0.5°, $u = 20$ m s^{-1}, $v = 0$, run time = 20 days and b) latitude 0.5°, $u = -20$ m s^{-1}, $v = 0$, run time = 20 days. Obviously, eastward and westward motion near the equator leads to very different behavior. Briefly explain why the trajectories are so different in

these two cases. By running different time intervals, determine the approximate period of oscillation in each case (i.e., the time to return to the original latitude.)

M1.4. More strange behavior near the equator. Run the script **const_ang_mom_traj1.m** by specifying initial conditions of latitude $= 0$, $u = 0$, $v = 50$ m s^{-1}, and a time of about 5 or 10 days. Notice that the motion is symmetric about the equator and that there is a net eastward drift. Why does providing a parcel with an initial poleward velocity at the equator lead to an eastward average displacement? By trying different initial meridional velocities in the range of 50 to 250 m s^{-1}, determine the approximate dependence of the maximum latitude reached by the ball on the initial meridional velocity. Also determine how the net eastward displacement depends on the initial meridional velocity. Show your results in a table or plot them using MATLAB.

Suggested References

Complete reference information is provided in the Bibliography at the end of the book.

Wallace and Hobbs, *Atmospheric Science: An Introductory Survey*, discusses much of the material in this chapter at an elementary level.

Curry and Webster, *Thermodynamics of Atmospheres and Oceans*, contains a more advanced discussion of atmospheric statistics.

Durran (1993) discusses the constant angular momentum oscillation in detail.

CHAPTER 2

Basic Conservation Laws

Atmospheric motions are governed by three fundamental physical principles: conservation of mass, conservation of momentum, and conservation of energy. The mathematical relations that express these laws may be derived by considering the budgets of mass, momentum, and energy for an infinitesimal *control volume* in the fluid. Two types of control volume are commonly used in fluid dynamics. In the *Eulerian* frame of reference the control volume consists of a parallelepiped of sides δx, δy, δz, whose position is fixed relative to the coordinate axes. Mass, momentum, and energy budgets will depend on fluxes caused by the flow of fluid through the boundaries of the control volume. (This type of control volume was used in Section 1.4.1.) In the *Lagrangian* frame, however, the control volume consists of an infinitesimal mass of "tagged" fluid particles; thus, the control volume moves about following the motion of the fluid, always containing the same fluid particles.

The Lagrangian frame is particularly useful for deriving conservation laws, as such laws may be stated most simply in terms of a particular mass element of the fluid. The Eulerian system is, however, more convenient for solving most problems because in that system the field variables are related by a set of partial differential equations in which the independent variables are the coordinates x, y, z, and t.

In the Lagrangian system, however, it is necessary to follow the time evolution of the fields for various individual fluid parcels. Thus the independent variables are x_0, y_0, z_0, and t, where x_0, y_0, and z_0 designate the position that a particular parcel passed through at a reference time t_0.

2.1 TOTAL DIFFERENTIATION

The conservation laws to be derived in this chapter contain expressions for the rates of change of density, momentum, and thermodynamic energy following the motion of particular fluid parcels. In order to apply these laws in the Eulerian frame it is necessary to derive a relationship between the rate of change of a field variable following the motion and its rate of change at a fixed point. The former is called the *substantial*, the *total*, or the *material* derivative (it will be denoted by D/Dt). The latter is called the *local* derivative (it is merely the partial derivative with respect to time). To derive a relationship between the total derivative and the local derivative, it is convenient to refer to a particular field variable (temperature, for example). For a given air parcel the location (x, y, z) is a function of t so that $x = x(t)$, $y = y(t)$, $z = z(t)$. Following the parcel, T may then be considered as truly a function only of time, and its rate of change is just the total derivative DT/Dt. In order to relate the total derivative to the local rate of change at a fixed point, we consider the temperature measured on a balloon that moves with the wind. Suppose that this temperature is T_0 at the point x_0, y_0, z_0 and time t_0. If the balloon moves to the point $x_0 + \delta x$, $y_0 + \delta y$, $z_0 + \delta z$ in a time increment δt, then the temperature change recorded on the balloon, δT, can be expressed in a Taylor series expansion as

$$\delta T = \left(\frac{\partial T}{\partial t}\right)\delta t + \left(\frac{\partial T}{\partial x}\right)\delta x + \left(\frac{\partial T}{\partial y}\right)\delta y + \left(\frac{\partial T}{\partial z}\right)\delta z + (\text{ higher order terms})$$

Dividing through by δt and noting that δT is the change in temperature following the motion so that

$$\frac{DT}{Dt} \equiv \lim_{\delta t \to 0} \frac{\delta T}{\delta t}$$

we find that in the limit $\delta t \to 0$

$$\frac{DT}{Dt} = \frac{\partial T}{\partial t} + \left(\frac{\partial T}{\partial x}\right)\frac{Dx}{Dt} + \left(\frac{\partial T}{\partial y}\right)\frac{Dy}{Dt} + \left(\frac{\partial T}{\partial z}\right)\frac{Dz}{Dt}$$

is the rate of change of T following the motion.

If we now let

$$\frac{Dx}{Dt} \equiv u, \quad \frac{Dy}{Dt} \equiv v, \quad \frac{Dz}{Dt} \equiv w$$

then u, v, w are the velocity components in the x, y, z directions, respectively, and

$$\frac{DT}{Dt} = \frac{\partial T}{\partial t} + \left(u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} + w\frac{\partial T}{\partial z} \right) \tag{2.1}$$

Using vector notation this expression may be rewritten as

$$\frac{\partial T}{\partial t} = \frac{DT}{Dt} - \mathbf{U} \bullet \boldsymbol{\nabla} T$$

where $\mathbf{U} = \mathbf{i}u + \mathbf{j}v + \mathbf{k}w$ is the velocity vector. The term $-\mathbf{U} \bullet \boldsymbol{\nabla} T$ is called the temperature *advection*. It gives the contribution to the local temperature change due to air motion. For example, if the wind is blowing from a cold region toward a warm region $-\mathbf{U} \bullet \boldsymbol{\nabla} T$ will be negative (cold advection) and the advection term will contribute negatively to the local temperature change. Thus, the local rate of change of temperature equals the rate of change of temperature following the motion (i.e., the heating or cooling of individual air parcels) plus the advective rate of change of temperature.

The relationship between the total derivative and local derivative given for temperature in (2.1) holds for any of the field variables. Furthermore, the total derivative can be defined following a motion field other than the actual wind field. For example, we may wish to relate the pressure change measured by a barometer on a moving ship to the local pressure change.

Example. The surface pressure decreases by 3 hPa per 180 km in the eastward direction. A ship steaming eastward at 10 km/h measures a pressure fall of 1 hPa per 3 h. What is the pressure change on an island that the ship is passing? If we take the x axis oriented eastward, then the local rate of change of pressure on the island is

$$\frac{\partial p}{\partial t} = \frac{Dp}{Dt} - u\frac{\partial p}{\partial x}$$

where Dp/Dt is the pressure change observed by the ship and u is the velocity of the ship. Thus,

$$\frac{\partial p}{\partial t} = \frac{-1 \text{ hPa}}{3 \text{ h}} - \left(10\frac{\text{km}}{\text{h}}\right)\left(\frac{-3 \text{ hPa}}{180 \text{ km}}\right) = -\frac{1 \text{ hPa}}{6 \text{ h}}$$

so that the rate of pressure fall on the island is only half the rate measured on the moving ship.

If the total derivative of a field variable is zero, then that variable is a conservative quantity following the motion. The local change is then entirely due to advection. As shown later, field variables that are approximately conserved following the motion play an important role in dynamic meteorology.

2.1.1 Total Differentiation of a Vector in a Rotating System

The conservation law for momentum (Newton's second law of motion) relates the rate of change of the absolute momentum following the motion in an inertial reference frame to the sum of the forces acting on the fluid. For most applications in meteorology it is desirable to refer the motion to a reference frame rotating with the earth. Transformation of the momentum equation to a rotating coordinate system requires a relationship between the total derivative of a vector in an inertial reference frame and the corresponding total derivative in a rotating system.

To derive this relationship, we let $\mathbf{A}$ be an arbitrary vector whose Cartesian components in an inertial frame are given by

$$\mathbf{A} = \mathbf{i}' A'_x + \mathbf{j}' A'_y + \mathbf{k}' A'_z$$

and whose components in a frame rotating with an angular velocity $\mathbf{\Omega}$ are

$$\mathbf{A} = \mathbf{i} A_x + \mathbf{j} A_y + \mathbf{k} A_z$$

Letting $D_a\mathbf{A}/Dt$ be the total derivative of $\mathbf{A}$ in the inertial frame, we can write

$$\begin{aligned}\frac{D_a\mathbf{A}}{Dt} &= \mathbf{i}'\frac{DA'_x}{Dt} + \mathbf{j}'\frac{DA'_y}{Dt} + \mathbf{k}'\frac{DA'_z}{Dt} \\ &= \mathbf{i}\frac{DA_x}{Dt} + \mathbf{j}\frac{DA_y}{Dt} + \mathbf{k}\frac{DA_z}{Dt} + \frac{D_a\mathbf{i}}{Dt}A_x + \frac{D_a\mathbf{j}}{Dt}A_y + \frac{D_a\mathbf{k}}{Dt}A_z\end{aligned}$$

The first three terms on the line above can be combined to give

$$\frac{DA}{Dt} \equiv \mathbf{i}\frac{DA_x}{Dt} + \mathbf{j}\frac{DA_y}{Dt} + \mathbf{k}\frac{DA_z}{Dt}$$

which is just the total derivative of $\mathbf{A}$ as viewed in the rotating coordinates (i.e., the rate of change of $\mathbf{A}$ following the relative motion).

The last three terms arise because the directions of the unit vectors $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ change their orientation in space as the earth rotates. These terms have a simple form for a rotating coordinate system. For example, considering the eastward directed unit vector:

$$\delta\mathbf{i} = \frac{\partial\mathbf{i}}{\partial\lambda}\delta\lambda + \frac{\partial\mathbf{i}}{\partial\phi}\delta\phi + \frac{\partial\mathbf{i}}{\partial z}\delta z$$

For solid body rotation $\delta\lambda = \Omega\delta t, \delta\phi = 0, \delta z = 0$, so that $\delta\mathbf{i}/\delta t = (\partial\mathbf{i}/\partial\lambda)\,(\delta\lambda/\delta t)$ and taking the limit $\delta t \to 0$,

$$\frac{D_a\mathbf{i}}{Dt} = \Omega\frac{\partial\mathbf{i}}{\partial\lambda}$$

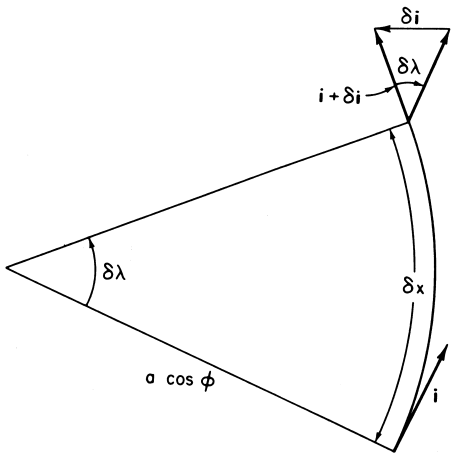


Fig. 2.1 Longitudinal dependence of the unit vector **i**.

But from Figs. 2.1 and 2.2, the longitudinal derivative of **i** can be expressed as

$$\frac{\partial \mathbf{i}}{\partial \lambda} = \mathbf{j}\sin\phi - \mathbf{k}\cos\phi$$

However, $\mathbf{\Omega} = (0, \Omega\sin\phi, \Omega\cos\phi)$ so that

$$\frac{D_a\mathbf{i}}{Dt} = \Omega(\mathbf{j}\sin\phi - \mathbf{k}\cos\phi) = \mathbf{\Omega}\times\mathbf{i}$$

In a similar fashion, it can be shown that $D_a\mathbf{j}/Dt = \mathbf{\Omega}\times\mathbf{j}$ and $D_a\mathbf{k}/Dt = \mathbf{\Omega}\times\mathbf{k}$. Therefore, the total derivative for a vector in an inertial frame is related to that in a rotating frame by the expression

$$\frac{D_a\mathbf{A}}{Dt} = \frac{D\mathbf{A}}{Dt} + \mathbf{\Omega}\times\mathbf{A} \tag{2.2}$$

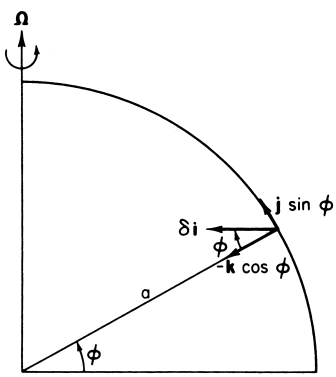


Fig. 2.2 Resolution of δi in Fig. 2.1 into northward and vertical components.

2.2 THE VECTORIAL FORM OF THE MOMENTUM EQUATION IN ROTATING COORDINATES

In an inertial reference frame, Newton's second law of motion may be written symbolically as

$$\frac{D_a \mathbf{U}_a}{Dt} = \sum \mathbf{F} \tag{2.3}$$

The left-hand side represents the rate of change of the absolute velocity $\mathbf{U}_a$, following the motion as viewed in an inertial system. The right-hand side represents the sum of the real forces acting per unit mass. In Section 1.5 we found through simple physical reasoning that when the motion is viewed in a rotating coordinate system certain additional apparent forces must be included if Newton's second law is to be valid. The same result may be obtained by a formal transformation of coordinates in (2.3).

In order to transform this expression to rotating coordinates, we must first find a relationship between $\mathbf{U}_a$ and the velocity relative to the rotating system, which we will designate by $\mathbf{U}$. This relationship is obtained by applying (2.2) to the position vector $\mathbf{r}$ for an air parcel on the rotating earth:

$$\frac{D_a \mathbf{r}}{Dt} = \frac{D\mathbf{r}}{Dt} + \mathbf{\Omega} \times \mathbf{r} \tag{2.4}$$

but $D_a\mathbf{r}/Dt \equiv \mathbf{U}_a$ and $D\mathbf{r}/Dt \equiv \mathbf{U}$; therefore (2.4) may be written as

$$\mathbf{U}_a = \mathbf{U} + \mathbf{\Omega} \times \mathbf{r} \tag{2.5}$$

which states simply that the absolute velocity of an object on the rotating earth is equal to its velocity relative to the earth plus the velocity due to the rotation of the earth.

Now we apply (2.2) to the velocity vector $\mathbf{U}_a$ and obtain

$$\frac{D_a \mathbf{U}_a}{Dt} = \frac{D\mathbf{U}_a}{Dt} + \mathbf{\Omega} \times \mathbf{U}_a \tag{2.6}$$

Substituting from (2.5) into the right-hand side of (2.6) gives

$$\begin{aligned} \frac{D_a \mathbf{U}_a}{Dt} &= \frac{D}{Dt}(\mathbf{U} + \mathbf{\Omega} \times \mathbf{r}) + \mathbf{\Omega} \times (\mathbf{U} + \mathbf{\Omega} \times \mathbf{r}) \\ &= \frac{D\mathbf{U}}{Dt} + 2\mathbf{\Omega} \times \mathbf{U} - \Omega^2 \mathbf{R} \end{aligned} \tag{2.7}$$

where $\mathbf{\Omega}$ is assumed to be constant. Here $\mathbf{R}$ is a vector perpendicular to the axis of rotation, with magnitude equal to the distance to the axis of rotation, so that with the aid of a vector identity,

$$\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}) = \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{R}) = -\Omega^2 \mathbf{R}$$

Equation (2.7) states that the acceleration following the motion in an inertial system equals the rate of change of relative velocity following the relative motion in the rotating frame plus the Coriolis acceleration due to relative motion in the rotating frame plus the centripetal acceleration caused by the rotation of the coordinates.

If we assume that the only real forces acting on the atmosphere are the pressure gradient force, gravitation, and friction, we can rewrite Newton's second law (2.3) with the aid of (2.7) as

$$\frac{D\mathbf{U}}{Dt} = -2\mathbf{\Omega} \times \mathbf{U} - \frac{\mathbf{1}}{\rho}\nabla p + \mathbf{g} + \mathbf{F}_r \tag{2.8}$$

where $\mathbf{F}_r$ designates the frictional force (see Section 1.4.3), and the centrifugal force has been combined with gravitation in the gravity term $\mathbf{g}$ (see Section 1.5.2). Equation (2.8) is the statement of Newton's second law for motion relative to a rotating coordinate frame. It states that the acceleration following the relative motion in the rotating frame equals the sum of the Coriolis force, the pressure gradient force, effective gravity, and friction. This form of the momentum equation is basic to most work in dynamic meteorology.

2.3 COMPONENT EQUATIONS IN SPHERICAL COORDINATES

For purposes of theoretical analysis and numerical prediction, it is necessary to expand the vectorial momentum equation (2.8) into its scalar components. Since the departure of the shape of the earth from sphericity is entirely negligible for meteorological purposes, it is convenient to expand (2.8) in spherical coordinates so that the (level) surface of the earth corresponds to a coordinate surface. The coordinate axes are then (λ, ϕ, z), where λ is longitude, ϕ is latitude, and z is the vertical distance above the surface of the earth. If the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are now taken to be directed eastward, northward, and upward, respectively, the relative velocity becomes

$$\mathbf{U} \equiv \mathbf{i}u + \mathbf{j}v + \mathbf{k}w$$

where the components u, v, and w are defined as

$$u \equiv r\cos\phi\,\frac{D\lambda}{Dt}, \qquad v \equiv r\frac{D\phi}{Dt}, \qquad w \equiv \frac{Dz}{Dt} \tag{2.9}$$

Here, r is the distance to the center of the earth, which is related to z by $r = a + z$, where a is the radius of the earth. Traditionally, the variable r in (2.9) is replaced by the constant a. This is a very good approximation, as $z \ll a$ for the regions of the atmosphere with which meteorologists are concerned. For notational simplicity,

it is conventional to define x and y as eastward and northward distance, such that $Dx = a\cos\phi\ D\lambda$ and $Dy = a\ D\phi$. Thus, the horizontal velocity components are $u \equiv Dx/Dt$ and $v \equiv Dy/Dt$ in the eastward and northward directions, respectively. The (x, y, z) coordinate system defined in this way is not, however, a Cartesian coordinate system because the directions of the **i**, **j**, **k** unit vectors are not constant, but are functions of position on the spherical earth. This position dependence of the unit vectors must be taken into account when the acceleration vector is expanded into its components on the sphere. Thus, we write

$$\frac{D\mathbf{U}}{Dt} = \mathbf{i}\frac{Du}{Dt} + \mathbf{j}\frac{Dv}{Dt} + \mathbf{k}\frac{Dw}{Dt} + u\frac{D\mathbf{i}}{Dt} + v\frac{D\mathbf{j}}{Dt} + w\frac{D\mathbf{k}}{Dt} \tag{2.10}$$

In order to obtain the component equations, it is necessary first to evaluate the rates of change of the unit vectors following the motion.

We first consider $D\mathbf{i}/Dt$. Expanding the total derivative as in (2.1) and noting that **i** is a function only of x (i.e., an eastward-directed vector does not change its orientation if the motion is in the north–south or vertical directions), we get

$$\frac{D\mathbf{i}}{Dt} = u\frac{\partial\mathbf{i}}{\partial x}$$

From Fig. 2.1 we see by similarity of triangles,

$$\lim_{\delta x\to 0}\frac{|\delta\mathbf{i}|}{\delta x} = \left|\frac{\partial\mathbf{i}}{\partial x}\right| = \frac{1}{a\cos\phi}$$

and that the vector $\partial\mathbf{i}/\partial x$ is directed toward the axis of rotation. Thus, as is illustrated in Fig. 2.2,

$$\frac{\partial\mathbf{i}}{\partial x} = \frac{1}{a\cos\phi}(\mathbf{j}\ \sin\phi - \mathbf{k}\cos\phi)$$

Therefore

$$\frac{D\mathbf{i}}{Dt} = \frac{u}{a\cos\phi}(\mathbf{j}\ \sin\phi - \mathbf{k}\cos\phi) \tag{2.11}$$

Considering now $D\mathbf{j}/Dt$, we note that **j** is a function only of x and y. Thus, with the aid of Fig. 2.3 we see that for eastward motion $|\delta\mathbf{j}| = \delta x/(a/\tan\phi)$. Because the vector $\partial\mathbf{j}/\partial x$ is directed in the negative x direction, we have then

$$\frac{\partial\mathbf{j}}{\partial x} = -\frac{\tan\phi}{a}\mathbf{i}$$

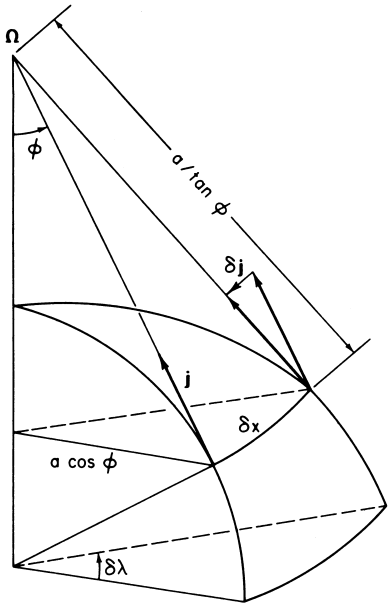


Fig. 2.3 Dependence of unit vector **j** on longitude.

From Fig. 2.4 it is clear that for northward motion $|\delta\mathbf{j}| = \delta\phi$, but $\delta y = a\delta\phi$, and $\delta\mathbf{j}$ is directed downward so that

$$\frac{\partial \mathbf{j}}{\partial y} = -\frac{\mathbf{k}}{a}$$

Hence,

$$\frac{D\mathbf{j}}{Dt} = -\frac{u\tan\phi}{a}\,\mathbf{i} - \frac{v}{a}\,\mathbf{k} \tag{2.12}$$

Finally, by similar arguments it can be shown that

$$\frac{D\mathbf{k}}{Dt} = \mathbf{i}\frac{u}{a} + \mathbf{j}\frac{v}{\mathbf{a}} \tag{2.13}$$

Substituting (2.11)–(2.13) into (2.10) and rearranging the terms, we obtain the spherical polar coordinate expansion of the acceleration following the relative motion

$$\begin{aligned}\frac{D\mathbf{U}}{Dt} = &\left(\frac{Du}{Dt} - \frac{uv\tan\phi}{a} + \frac{uw}{a}\right)\mathbf{i} + \left(\frac{Dv}{Dt} + \frac{u^2\tan\phi}{a} + \frac{vw}{a}\right)\mathbf{j} \\ &+ \left(\frac{Dw}{Dt} - \frac{u^2+v^2}{a}\right)\mathbf{k}\end{aligned} \tag{2.14}$$

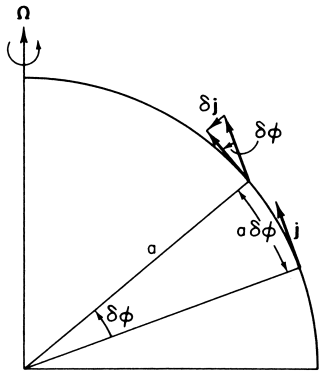


Fig. 2.4 Dependence of unit vector **j** on latitude.

We next turn to the component expansion of the force terms in (2.8). The Coriolis force is expanded by noting that $\mathbf{\Omega}$ has no component parallel to **i** and that its components parallel to **j** and **k** are $2\Omega\cos\phi$ and $2\Omega\sin\phi$, respectively. Thus, using the definition of the vector cross product,

$$-2\mathbf{\Omega}\times\mathbf{U} = -2\Omega\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 0 & \cos\phi & \sin\phi\\ u & v & w\end{vmatrix}$$
$$= -(2\Omega w\cos\phi - 2\Omega v\sin\phi)\,\mathbf{i} - 2\Omega u\sin\phi\,\mathbf{j} + 2\Omega u\cos\phi\,\mathbf{k} \qquad (2.15)$$

The pressure gradient may be expressed as

$$\nabla p = \mathbf{i}\frac{\partial p}{\partial x} + \mathbf{j}\frac{\partial p}{\partial y} + \mathbf{k}\frac{\partial p}{\partial z} \qquad (2.16)$$

and gravity is conveniently represented as

$$\mathbf{g} = -g\,\mathbf{k} \qquad (2.17)$$

where g is a positive scalar ($g \cong 9.8$ m s^{-2} at the earth's surface). Finally, recall from (1.5) that

$$\mathbf{F}_r = \mathbf{i}F_{rx} + \mathbf{j}F_{ry} + \mathbf{k}F_{rz} \qquad (2.18)$$

Substituting (2.14)–(2.18) into the equation of motion (2.8) and equating all terms in the **i**, **j**, and **k** directions, respectively, we obtain

$$\frac{Du}{Dt} - \frac{uv\tan\phi}{a} + \frac{uw}{a} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + 2\Omega v\sin\phi - 2\Omega w\cos\phi + F_{rx} \qquad (2.19)$$

$$\frac{Dv}{Dt} + \frac{u^2\tan\phi}{a} + \frac{vw}{a} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - 2\Omega u\sin\phi + F_{ry} \qquad (2.20)$$

$$\frac{Dw}{Dt} - \frac{u^2 + v^2}{a} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g + 2\Omega u \cos\phi + F_{rz} \tag{2.21}$$

which are the eastward, northward, and vertical component momentum equations, respectively. The terms proportional to $1/a$ on the left-hand sides in (2.19)–(2.21) are called *curvature* terms; they arise due to the curvature of the earth.[1] Because they are nonlinear (i.e., they are quadratic in the dependent variables), they are difficult to handle in theoretical analyses. Fortunately, as shown in the next section, the curvature terms are unimportant for midlatitude synoptic scale motions. However, even when the curvature terms are neglected, (2.19)–(2.21) are still nonlinear partial differential equations, as can be seen by expanding the total derivatives into their local and advective parts:

$$\frac{Du}{Dt} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}$$

with similar expressions for Dv/Dt and Dw/Dt. In general the advective acceleration terms are comparable in magnitude to the local acceleration. The presence of nonlinear advection processes is one reason that dynamic meteorology is an interesting and challenging subject.

2.4 SCALE ANALYSIS OF THE EQUATIONS OF MOTION

Section 1.3 discussed the basic notion of scaling the equations of motion in order to determine whether some terms in the equations are negligible for motions of meteorological concern. Elimination of terms on scaling considerations not only has the advantage of simplifying the mathematics, but as shown in later chapters, the elimination of small terms in some cases has the very important property of completely eliminating or *filtering* an unwanted type of motion. The complete equations of motion [(2.19)–(2.21)] describe all types and scales of atmospheric motions. Sound waves, for example, are a perfectly valid class of solutions to these equations. However, sound waves are of negligible importance in dynamical meteorology. Therefore, it will be a distinct advantage if, as turns out to be true, we can neglect the terms that lead to the production of sound waves and filter out this unwanted class of motions.

In order to simplify (2.19)–(2.21) for synoptic scale motions, we define the following characteristic scales of the field variables based on observed values for midlatitude synoptic systems.

[1] It can be shown that when r is replaced by a as done here (the traditional approximation) the Coriolis terms proportional to $\cos\phi$ in (2.19) and (2.21) must be neglected if the equations are to satisfy angular momentum conservation.

$U \sim 10\,\mathrm{m\,s^{-1}}$	horizontal velocity scale
$W \sim 1\,\mathrm{cm\,s^{-1}}$	vertical velocity scale
$L \sim 10^6\,\mathrm{m}$	length scale [$\sim 1/(2\pi)$ wavelength]
$H \sim 10^4\,\mathrm{m}$	depth scale
$\delta P/\rho \sim 10^3\,\mathrm{m^2\,s^{-2}}$	horizontal pressure fluctuation scale
$L/U \sim 10^5\,\mathrm{s}$	time scale

Horizontal pressure fluctuation δP is normalized by the density ρ in order to produce a scale estimate that is valid at all heights in the troposphere, despite the approximate exponential decrease with height of both δP and ρ. Note that $\delta P/\rho$ has units of a geopotential. Referring back to (1.25) we see that indeed the magnitude of the fluctuation of $\delta P/\rho$ on a surface of constant height must equal the magnitude of the fluctuation of the geopotential on an isobaric surface. The time scale here is an advective time scale, which is appropriate for pressure systems that move at approximately the speed of the horizontal wind, as is observed for synoptic scale motions. Thus, L/U is the time required to travel a distance L at a speed U, and the substantial differential operator scales as $D/Dt \sim U/L$ for such motions.

It should be pointed out here that the synoptic scale vertical velocity is not a directly measurable quantity. However, as shown in Chapter 3, the magnitude of w can be deduced from knowledge of the horizontal, velocity field.

We can now estimate the magnitude of each term in (2.19) and (2.20) for synoptic scale motions at a given latitude. It is convenient to consider a disturbance centered at latitude $\phi_0 = 45°$ and introduce the notation

$$f_0 = 2\Omega \sin\phi_0 = 2\Omega \cos\phi_0 \cong 10^{-4} s^{-1}$$

Table 2.1 shows the characteristic magnitude of each term in (2.19) and (2.20) based on the scaling considerations given above. The molecular friction term is so small that it may be neglected for all motions except the smallest scale turbulent motions near the ground, where vertical wind shears can become very large and the molecular friction term must be retained, as discussed in Chapter 5.

Table 2.1 *Scale Analysis of the Horizontal Momentum Equations*

	A	B	C	D	E	F	G
$x-$ Eq.	$\frac{Du}{Dt}$	$-2\Omega v \sin\phi$	$+2\Omega w \cos\phi$	$+\frac{uw}{a}$	$-\frac{uv\tan\phi}{a}$	$=-\frac{1}{\rho}\frac{\partial p}{\partial x}$	$+F_{rx}$
$y-$ Eq.	$\frac{Dv}{Dt}$	$+2\Omega u \sin\phi$		$+\frac{vw}{a}$	$+\frac{u^2\tan\phi}{a}$	$=-\frac{1}{\rho}\frac{\partial p}{\partial y}$	$+F_{ry}$
Scales	U^2/L	$f_0 U$	$f_0 W$	$\frac{UW}{a}$	$\frac{U^2}{a}$	$\frac{\delta P}{\rho L}$	$\frac{\nu U}{H^2}$
(m s^{-2})	10^{-4}	10^{-3}	10^{-6}	10^{-8}	10^{-5}	10^{-3}	10^{-12}

2.4.1 Geostrophic Approximation and Geostrophic Wind

It is apparent from Table 2.1 that for midlatitude synoptic scale disturbances the Coriolis force (term B) and the pressure gradient force (term F) are in approximate balance. Retaining only these two terms in (2.19) and (2.20) gives as a first approximation the *geostrophic* relationship

$$-fv \approx -\frac{1}{\rho}\frac{\partial p}{\partial x}; \qquad fu \approx -\frac{1}{\rho}\frac{\partial p}{\partial y} \tag{2.22}$$

where $f \equiv 2\Omega\sin\phi$ is called the *Coriolis parameter*. The geostrophic balance is a *diagnostic* expression that gives the approximate relationship between the pressure field and horizontal velocity in large-scale extratropical systems. The approximation (2.22) contains no reference to time and therefore cannot be used to predict the evolution of the velocity field. It is for this reason that the geostrophic relationship is called a diagnostic relationship.

By analogy to the geostrophic approximation (2.22) it is possible to define a horizontal velocity field, $\mathbf{V}_g \equiv \mathbf{i}u_g + \mathbf{j}v_g$, called the *geostrophic wind*, which satisfies (2.22) identically. In vectorial form,

$$\mathbf{V}_g \equiv \mathbf{k} \times \frac{1}{\rho f}\nabla p \tag{2.23}$$

Thus, knowledge of the pressure distribution at any time determines the geostrophic wind. It should be kept clearly in mind that (2.23) always defines the geostrophic wind, but only for large-scale motions away from the equator should the geostrophic wind be used as an approximation to the actual horizontal wind field. For the scales used in Table 2.1 the geostrophic wind approximates the true horizontal velocity to within 10–15% in midlatitudes.

2.4.2 Approximate Prognostic Equations; the Rossby Number

To obtain prediction equations it is necessary to retain the acceleration (term A) in (2.19) and (2.20). The resulting approximate horizontal momentum equations are

$$\frac{Du}{Dt} = fv - \frac{1}{\rho}\frac{\partial p}{\partial x} = f\left(v - v_g\right) \tag{2.24}$$

$$\frac{Dv}{Dt} = -fu - \frac{1}{\rho}\frac{\partial p}{\partial y} = -f\left(u - u_g\right) \tag{2.25}$$

where (2.23) is used to rewrite the pressure gradient force in terms of the geostrophic wind. Because the acceleration terms in (2.24) and (2.25) are proportional to the difference between the actual wind and the geostrophic wind, they are about an

order of magnitude smaller than the Coriolis force and the pressure gradient force in agreement with our scale analysis. The fact that the horizontal flow is in approximate geostrophic balance is helpful for diagnostic analysis. However, it makes actual applications of these equations in weather prognosis difficult because acceleration (which must be measured accurately) is given by the small difference between two large terms. Thus, a small error in measurement of either velocity or pressure gradient will lead to very large errors in estimating the acceleration. This problem is discussed in some detail in Chapter 13.

A convenient measure of the magnitude of the acceleration compared to the Coriolis force may be obtained by forming the ratio of the characteristic scales for the acceleration and the Coriolis force terms: $(U^2/L)/(f_0U)$. This ratio is a nondimensional number called the *Rossby number* after the Swedish meteorologist C. G. Rossby (1898–1957) and is designated by

$$R_0 \equiv U/(f_0L)$$

Thus, the smallness of the Rossby number is a measure of the validity of the geostrophic approximation.

2.4.3 The Hydrostatic Approximation

A similar scale analysis can be applied to the vertical component of the momentum equation (2.21). Because pressure decreases by about an order of magnitude from the ground to the tropopause, the vertical pressure gradient may be scaled by P_0/H, where P_0 is the surface pressure and H is the depth of the troposphere. The terms in (2.21) may then be estimated for synoptic scale motions and are shown in Table 2.2. As with the horizontal component equations, we consider motions centered at 45° latitude and neglect friction. The scaling indicates that to a high degree of accuracy the pressure field is in *hydrostatic equilibrium*; that is, the pressure at any point is simply equal to the weight of a unit cross-section column of air above that point.

The above analysis of the vertical momentum equation is, however, somewhat misleading. It is not sufficient to show merely that the vertical acceleration is small compared to g. Because only that part of the pressure field that varies horizontally is directly coupled to the horizontal velocity field, it is actually necessary to show that the horizontally varying pressure component is itself in hydrostatic equilibrium with the horizontally varying density field. To do this it is convenient to first define a

Table 2.2 *Scale Analysis of the Vertical Momentum Equation*

z - Eq.	Dw/Dt	$-2\Omega u\cos\phi$	$-(u^2+v^2)/a$	$=-\rho^{-1}\partial p/\partial z$	$-g$	$+F_{rz}$
Scales	UW/L	f_0U	U^2/a	$P_0/(\rho H)$	g	νWH^{-2}
m s^{-2}	10^{-7}	10^{-3}	10^{-5}	10	10	10^{-15}

standard pressure $p_0(z)$, which is the horizontally averaged pressure at each height, and a corresponding standard density $\rho_0(z)$, defined so that $p_0(z)$ and $\rho_0(z)$ are in *exact* hydrostatic balance:

$$\frac{1}{\rho_0}\frac{dp_0}{dz} \equiv -g \tag{2.26}$$

We may then write the total pressure and density fields as

$$\begin{aligned} p(x, y, z, t) &= p_0(z) + p'(x, y, z, t) \\ \rho(x, y, z, t) &= \rho_0(z) + \rho'(x, y, z, t) \end{aligned} \tag{2.27}$$

where p' and ρ' are deviations from the standard values of pressure and density. For an atmosphere at rest, p' and ρ' would thus be zero. Using the definitions of (2.26) and (2.27) and assuming that ρ'/ρ_0 is much less than unity in magnitude so that $(\rho_0 + \rho')^{-1} \cong \rho_0^{-1}\,(1 - \rho'/\rho_0)$, we find that

$$\begin{aligned} -\frac{1}{\rho}\frac{\partial p}{\partial z} - g &= -\frac{1}{(\rho_0 + \rho')}\frac{\partial}{\partial z}\left(p_0 + p'\right) - g \\ &\approx \frac{1}{\rho_0}\left[\frac{\rho'}{\rho_0}\frac{dp_0}{dz} - \frac{\partial p'}{\partial z}\right] = -\frac{1}{\rho_0}\left[\rho' g + \frac{\partial p'}{\partial z}\right] \end{aligned} \tag{2.28}$$

For synoptic scale motions, the terms in (2.28) have the magnitudes

$$\frac{1}{\rho_0}\frac{\partial p'}{\partial z} \sim \left[\frac{\delta P}{\rho_0 H}\right] \sim 10^{-1}\,\mathrm{m\,s^{-2}}, \qquad \frac{\rho' g}{\rho_0} \sim 10^{-1}\,\mathrm{m\,s^{-2}}$$

Comparing these with the magnitudes of other terms in the vertical momentum equation (Table 2.2), we see that to a very good approximation the perturbation pressure field is in hydrostatic equilibrium with the perturbation density field so that

$$\frac{\partial p'}{\partial z} + \rho' g = 0 \tag{2.29}$$

Therefore, for synoptic scale motions, vertical accelerations are negligible and the vertical velocity cannot be determined from the vertical momentum equation. However, we show in Chapter 3 that it is, nevertheless, possible to deduce the vertical motion field indirectly.

2.5 THE CONTINUITY EQUATION

We turn now to the second of the three fundamental conservation principles, conservation of mass. The mathematical relationship that expresses conservation of mass for a fluid is called the *continuity equation*. This section develops the continuity equation using two alternative methods. The first method is based on an Eulerian control volume, whereas the second is based on a Lagrangian control volume.

2.5.1 An Eulerian Derivation

We consider a volume element $\delta x\ \delta y\ \delta z$ that is fixed in a Cartesian coordinate frame as shown in Fig. 2.5. For such a fixed control volume the net rate of mass inflow through the sides must equal the rate of accumulation of mass within the volume. The rate of inflow of mass through the left-hand face per unit area is

$$\left[\rho u - \frac{\partial}{\partial x}(\rho u)\frac{\delta x}{2}\right]$$

whereas the rate of outflow per unit area through the right-hand face is

$$\left[\rho u + \frac{\partial}{\partial x}(\rho u)\frac{\delta x}{2}\right]$$

Because the area of each of these faces is $\delta y \delta z$, the net rate of flow into the volume due to the x velocity component is

$$\left[\rho u - \frac{\partial}{\partial x}(\rho u)\frac{\delta x}{2}\right]\delta y \delta z - \left[\rho u + \frac{\partial}{\partial x}(\rho u)\frac{\delta x}{2}\right]\delta y \delta z$$
$$= -\frac{\partial}{\partial x}(\rho u)\delta x \delta y \delta z$$

Similar expressions obviously hold for the y and z directions. Thus, the net rate of mass inflow is

$$-\left[\frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial y}(\rho v) + \frac{\partial}{\partial z}(\rho w)\right]\delta x \delta y \delta z$$

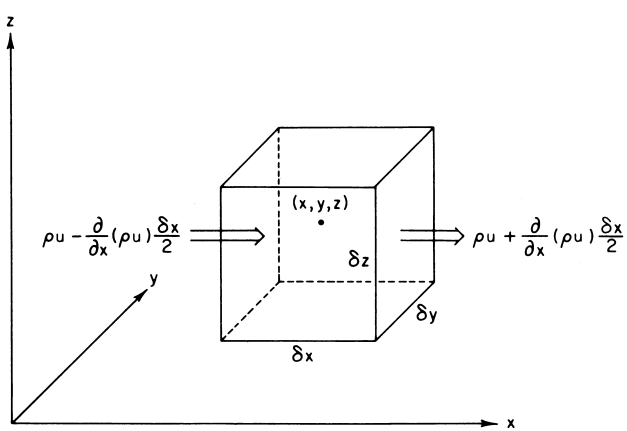


Fig. 2.5 Mass inflow into a fixed (Eulerian) control volume due to motion parallel to the x axis.

and the mass inflow per unit volume is just $-\nabla \bullet (\rho \mathbf{U})$, which must equal the rate of mass increase per unit volume. Now the increase of mass per unit volume is just the local density change $\partial\rho/\partial t$.Therefore,

$$\frac{\partial \rho}{\partial t} + \nabla \bullet (\rho \mathbf{U}) = 0 \tag{2.30}$$

Equation (2.30) is the mass divergence form of the continuity equation.

An alternative form of the continuity equation is obtained by applying the vector identity

$$\nabla \bullet (\rho \mathbf{U}) \equiv \rho \nabla \bullet \mathbf{U} + \mathbf{U} \bullet \nabla \rho$$

and the relationship

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{U} \bullet \nabla$$

to get

$$\frac{1}{\rho}\frac{D\rho}{Dt} + \nabla \bullet \mathbf{U} = 0 \tag{2.31}$$

Equation (2.31) is the velocity divergence form of the continuity equation. It states that the fractional rate of increase of the density *following the motion* of an air parcel is equal to minus the velocity divergence. This should be clearly distinguished from (2.30), which states that the *local* rate of change of density is equal to minus the mass divergence.

2.5.2 A Lagrangian Derivation

The physical meaning of divergence can be illustrated by the following alternative derivation of (2.31). Consider a control volume of fixed mass δM that moves with the fluid. Letting $\delta V = \delta x\ \delta y \delta z$ be the volume, we find that because $\delta M = \rho \delta V = \rho \delta x \delta y \delta z$ is conserved following the motion, we can write

$$\frac{1}{\delta M}\frac{D}{Dt}(\delta M) = \frac{1}{\rho \delta V}\frac{D}{Dt}(\rho \delta V) = \frac{1}{\rho}\frac{D\rho}{Dt} + \frac{1}{\delta V}\frac{D}{Dt}(\delta V) = 0 \tag{2.32}$$

but

$$\frac{1}{\delta V}\frac{D}{Dt}(\delta V) = \frac{1}{\delta x}\frac{D}{Dt}(\delta x) + \frac{1}{\delta y}\frac{D}{Dt}(\delta y) + \frac{1}{\delta z}\frac{D}{Dt}(\delta z)$$

Referring to Fig. 2.6, we see that the faces of the control volume in the y, z plane (designated A and B) are advected with the flow in the x direction at speeds $u_A = Dx/Dt$ and $u_B = D(x + \delta x)\,/\,Dt$, respectively. Thus, the difference in speeds of the two faces is $\delta u = u_B - u_A = D(x + \delta x)/Dt - Dx/Dt$ or $\delta u = D(\delta x)\,/Dt$. Similarly, $\delta v = D(\delta y)/Dt$ and $\delta w = D(\delta z)/Dt$. Therefore,

$$\lim_{\delta x, \delta y, \delta z \to 0}\left[\frac{1}{\delta V}\frac{D}{Dt}(\delta V)\right] = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = \nabla \bullet \mathbf{U}$$

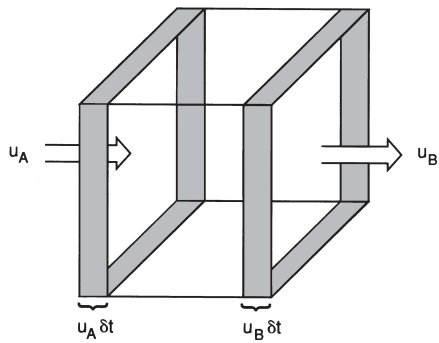


Fig. 2.6 Change in Lagrangian control volume (shown by shading) due to fluid motion parallel to the x axis.

so that in the limit $\delta V \to 0$, (2.32) reduces to the continuity equation (2.31); the divergence of the three-dimensional velocity field is equal to the fractional rate of change of volume of a fluid parcel in the limit $\delta V \to 0$. It is left as a problem for the student to show that the divergence of the *horizontal* velocity field is equal to the fractional rate of change of the horizontal area δA of a fluid parcel in the limit $\delta A \to 0$.

2.5.3 Scale Analysis of the Continuity Equation

Following the technique developed in Section 2.4.3, and again assuming that $|\rho'/\rho_0| \ll 1$, we can approximate the continuity equation (2.31) as

$$\underset{\text{A}}{\frac{1}{\rho_0}\left(\frac{\partial \rho'}{\partial t} + \mathbf{U} \bullet \nabla \rho'\right)} + \underset{\text{B}}{\frac{w}{\rho_0}\frac{d\rho_0}{dz}} + \underset{\text{C}}{\nabla \bullet \mathbf{U}} \approx 0 \tag{2.33}$$

where ρ' designates the local deviation of density from its horizontally averaged value, $\rho_0(z)$. For synoptic scale motions $\rho'/\rho_0 \sim 10^{-2}$ so that using the characteristic scales given in Section 2.4 we find that term A has magnitude

$$\frac{1}{\rho_0}\left(\frac{\partial \rho'}{\partial t} + \mathbf{U} \bullet \nabla \rho'\right) \sim \frac{\rho'}{\rho_0}\frac{U}{L} \approx 10^{-7}\text{s}^{-1}$$

For motions in which the depth scale H is comparable to the density scale height, $d \ln \rho_0/dz \sim H^{-1}$, so that term B scales as

$$\frac{w}{\rho_0}\frac{d\rho_0}{dz} \backsim \frac{W}{H} \approx 10^{-6}\text{s}^{-1}$$

Expanding term C in Cartesian coordinates, we have

$$\nabla \bullet \mathbf{U} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$$

For synoptic scale motions the terms $\partial u/\partial x$ and $\partial v/\partial y$ tend to be of equal magnitude but opposite sign. Thus, they tend to balance so that

$$\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \backsim 10^{-1}\frac{U}{L} \approx 10^{-6}\,\mathrm{s}^{-1}$$

and in addition

$$\frac{\partial w}{\partial z} \sim \frac{W}{H} \approx 10^{-6}\,\mathrm{s}^{-1}$$

Thus, terms B and C are each an order of magnitude greater than term A, and to a first approximation, terms B and C balance in the continuity equation. To a good approximation then

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} + w\frac{d}{dz}(\ln \rho_0) = 0$$

or, alternatively, in vector form

$$\nabla \bullet (\rho_0 \mathbf{U}) = 0 \tag{2.34}$$

Thus for synoptic scale motions the mass flux computed using the basic state density ρ_0 is nondivergent. This approximation is similar to the idealization of incompressibility, which is often used in fluid mechanics. However, an *incompressible* fluid has density constant following the motion:

$$\frac{D\rho}{Dt} = 0$$

Thus by (2.31) the velocity divergence vanishes ($\nabla \bullet \mathbf{U} = 0$) in an incompressible fluid, which is not the same as (2.34). Our approximation (2.34) shows that for purely horizontal flow the atmosphere behaves as though it were an incompressible fluid. However, when there is vertical motion the compressibility associated with the height dependence of ρ_0 must be taken into account.

2.6 THE THERMODYNAMIC ENERGY EQUATION

We now turn to the third fundamental conservation principle, the conservation of energy as applied to a moving fluid element. The first law of thermodynamics is usually derived by considering a system in thermodynamic equilibrium, that is, a system that is initially at rest and after exchanging heat with its surroundings and doing work on the surroundings is again at rest. For such a system the first law states that *the change in internal energy of the system is equal to the difference between the heat added to the system and the work done by the system.*

A Lagrangian control volume consisting of a specified mass of fluid may be regarded as a thermodynamic system. However, unless the fluid is at rest, it will not be in thermodynamic equilibrium. Nevertheless, the first law of thermodynamics still applies. To show that this is the case, we note that the total thermodynamic energy of the control volume is considered to consist of the sum of the internal energy (due to the kinetic energy of the individual molecules) and the kinetic energy due to the macroscopic motion of the fluid. The rate of change of this total thermodynamic energy is equal to the rate of diabatic heating plus the rate at which work is done on the fluid parcel by external forces.

If we let e designate the internal energy per unit mass, then the total thermodynamic energy contained in a Lagrangian fluid element of density ρ and volume δV is $\rho[e + (1/2)\mathbf{U} \bullet \mathbf{U}]\,\delta V$. The external forces that act on a fluid element may be divided into surface forces, such as pressure and viscosity, and body forces, such as gravity or the Coriolis force. The rate at which work is done on the fluid element by the x component of the pressure force is illustrated in Fig. 2.7. Recalling that pressure is a force per unit area and that the rate at which a force does work is given by the dot product of the force and velocity vectors, we see that the rate at which the surrounding fluid does work on the element due to the pressure force on the two boundary surfaces in the y, z plane is given by

$$(pu)_A \delta y \delta z - (pu)_B \delta y \delta z$$

(The negative sign is needed before the second term because the work done on the fluid element is positive if u is negative across face B.) Now by expanding in a Taylor series we can write

$$(pu)_B = (pu)_A + \left[\frac{\partial}{\partial x}(pu)\right]_A \delta x + \ldots$$

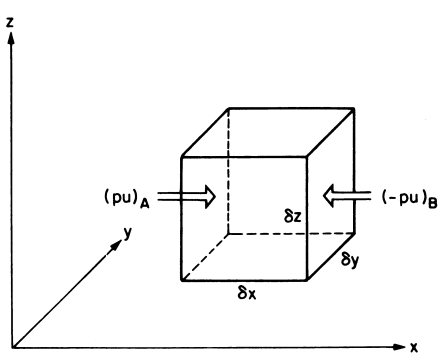


Fig. 2.7 Rate of working on a fluid element due to the x component of the pressure force.

Thus the net rate at which the pressure force does work due to the x component of motion is

$$\left[(pu)_A - (pu)_B\right]\delta y\delta z = -\left[\frac{\partial}{\partial x}(pu)\right]_A \delta V$$

where $\delta V = \delta x\delta y\delta z$.

Similarly, we can show that the net rates at which the pressure force does work due to the y and z components of motion are

$$-\left[\frac{\partial}{\partial y}(pv)\right]\delta V \text{ and } -\left[\frac{\partial}{\partial z}(pw)\right]\delta V$$

respectively. Hence, the total rate at which work is done by the pressure force is simply

$$-\boldsymbol{\nabla}\bullet(p\mathbf{U})\delta V$$

The only body forces of meteorological significance that act on an element of mass in the atmosphere are the Coriolis force and gravity. However, because the Coriolis force, $-2\Omega \times \mathbf{U}$, is perpendicular to the velocity vector, it can do no work. Thus the rate at which body forces do work on the mass element is just $\rho\,\mathbf{g}\bullet\mathbf{U}\,\delta V$.

Applying the principle of energy conservation to our Lagrangian control volume (neglecting effects of molecular viscosity), we thus obtain

$$\frac{D}{Dt}\left[\rho\left(e + \frac{1}{2}\mathbf{U}\bullet\mathbf{U}\right)\delta V\right] = -\boldsymbol{\nabla}\bullet(p\mathbf{U})\delta V + \rho\mathbf{g}\bullet\mathbf{U}\delta V + \rho J\delta V \tag{2.35}$$

Here J is the rate of heating per unit mass due to radiation, conduction, and latent heat release. With the aid of the chain rule of differentiation we can rewrite (2.35) as

$$\begin{aligned}\rho\delta V\frac{D}{Dt}\left(e + \frac{1}{2}\mathbf{U}\bullet\mathbf{U}\right) + \left(e + \frac{1}{2}\mathbf{U}\bullet\mathbf{U}\right)\frac{D(\rho\delta V)}{Dt} \\ = -\mathbf{U}\bullet\boldsymbol{\nabla}p\,\delta V - p\boldsymbol{\nabla}\bullet\mathbf{U}\,\delta V - \rho g w\,\delta V + \rho J\,\delta V\end{aligned} \tag{2.36}$$

where we have used $\mathbf{g} = -g\mathbf{k}$. Now from (2.32) the second term on the left in (2.36) vanishes so that

$$\rho\frac{De}{Dt} + \rho\frac{D}{Dt}\left(\frac{1}{2}\mathbf{U}\bullet\mathbf{U}\right) = -\mathbf{U}\bullet\boldsymbol{\nabla}p - p\boldsymbol{\nabla}\bullet\mathbf{U} - \rho g w + \rho J \tag{2.37}$$

This equation can be simplified by noting that if we take the dot product of $\mathbf{U}$ with the momentum equation (2.8) we obtain (neglecting friction)

$$\rho\frac{D}{Dt}\left(\frac{1}{2}\mathbf{U}\bullet\mathbf{U}\right) = -\mathbf{U}\bullet\boldsymbol{\nabla}p - \rho g w \tag{2.38}$$

Subtracting (2.38) from (2.37), we obtain

$$\rho \frac{De}{Dt} = -p\nabla \bullet \mathbf{U} + \rho J \tag{2.39}$$

The terms in (2.37) that were eliminated by subtracting (2.38) represent the balance of mechanical energy due to the motion of the fluid element; the remaining terms represent the thermal energy balance.

Using the definition of geopotential (1.15), we have

$$gw = g\frac{Dz}{Dt} = \frac{D\Phi}{Dt}$$

so that (2.38) can be rewritten as

$$\rho \frac{D}{Dt}\left(\frac{1}{2}\mathbf{U} \bullet \mathbf{U} + \Phi\right) = -\mathbf{U} \bullet \nabla p \tag{2.40}$$

which is referred to as the *mechanical energy equation*. The sum of the kinetic energy plus the gravitational potential energy is called the *mechanical energy*. Thus (2.40) states that following the motion, the rate of change of mechanical energy per unit volume equals the rate at which work is done by the pressure gradient force.

The thermal energy equation (2.39) can be written in more familiar form by noting from (2.31) that

$$\frac{1}{\rho}\nabla \bullet \mathbf{U} = -\frac{1}{\rho^2}\frac{D\rho}{Dt} = \frac{D\alpha}{Dt}$$

and that for dry air the internal energy per unit mass is given by $e = c_v T$, where $c_v (= 717 \text{ J kg}^{-1} \text{ K}^{-1})$ is the specific heat at constant volume. We then obtain

$$c_v \frac{DT}{Dt} + p\frac{D\alpha}{Dt} = J \tag{2.41}$$

which is the usual form of the thermodynamic energy equation. Thus the first law of thermodynamics indeed is applicable to a fluid in motion. The second term on the left, representing the rate of working by the fluid system (per unit mass), represents a conversion between thermal and mechanical energy. This conversion process enables the solar heat energy to drive the motions of the atmosphere.

2.7 THERMODYNAMICS OF THE DRY ATMOSPHERE

Taking the total derivative of the equation of state (1.14), we obtain

$$p\frac{D\alpha}{Dt} + \alpha\frac{Dp}{Dt} = R\frac{DT}{Dt}$$

Substituting for $pD\alpha/Dt$ in (2.41) and using $c_p = c_v + R$, where c_p (= 1004 J kg^{-1} K^{-1}) is the specific heat at constant pressure, we can rewrite the first law of thermodynamics as

$$c_p \frac{DT}{Dt} - \alpha \frac{Dp}{Dt} = J \tag{2.42}$$

Dividing through by T and again using the equation of state, we obtain the entropy form of the first law of thermodynamics:

$$c_p \frac{D \ln T}{Dt} - R \frac{D \ln p}{Dt} = \frac{J}{T} \equiv \frac{Ds}{Dt} \tag{2.43}$$

Equation (2.43) gives the rate of change of entropy per unit mass following the motion for a thermodynamically *reversible* process. A reversible process is one in which a system changes its thermodynamic state and then returns to the original state without changing its surroundings. For such a process the entropy s defined by (2.43) is a field variable that depends only on the state of the fluid. Thus Ds is a perfect differential, and Ds/Dt is to be regarded as a total derivative. However, "heat" is not a field variable, so that the heating rate J is not a total derivative.[2]

2.7.1 Potential Temperature

For an ideal gas undergoing an *adiabatic* process (i.e., a reversible process in which no heat is exchanged with the surroundings), the first law of thermodynamics can be written in differential form as

$$c_p D \ln T - R D \ln p = D\left(c_p \ln T - R \ln p\right) = 0$$

Integrating this expression from a state at pressure p and temperature T to a state in which the pressure is p_s and the temperature is θ, we obtain after taking the antilogarithm

$$\theta = T\,(p_s/p)^{R/c_p} \tag{2.44}$$

This relationship is referred to as Poisson's equation, and the temperature θ defined by (2.44) is called the *potential temperature*. θ is simply the temperature that a parcel of dry air at pressure p and temperature T would have if it were expanded or compressed adiabatically to a standard pressure p_s (usually taken to be 1000 hPa). Thus, every air parcel has a unique value of potential temperature, and this value is conserved for dry adiabatic motion. Because synoptic scale motions are approximately adiabatic outside regions of active precipitation, θ is a quasi-conserved quantity for such motions.

[2] For a discussion of entropy and its role in the second law of thermodynamics, see Curry and Webster (1999), for example.

Taking the logarithm of (2.44) and differentiating, we find that

$$c_p \frac{D \ln \theta}{Dt} = c_p \frac{D \ln T}{Dt} - R \frac{D \ln p}{Dt} \tag{2.45}$$

Comparing (2.43) and (2.45), we obtain

$$c_p \frac{D \ln \theta}{Dt} = \frac{J}{T} = \frac{Ds}{Dt} \tag{2.46}$$

Thus, for reversible processes, fractional potential temperature changes are indeed proportional to entropy changes. A parcel that conserves entropy following the motion must move along an isentropic (constant θ) surface.

2.7.2 The Adiabatic Lapse Rate

A relationship between the *lapse rate* of temperature (i.e., the rate of *decrease* of temperature with respect to height) and the rate of change of potential temperature with respect to height can be obtained by taking the logarithm of (2.44) and differentiating with respect to height. Using the hydrostatic equation and the ideal gas law to simplify the result gives

$$\frac{T}{\theta} \frac{\partial \theta}{\partial z} = \frac{\partial T}{\partial z} + \frac{g}{c_p} \tag{2.47}$$

For an atmosphere in which the potential temperature is constant with respect to height, the lapse rate is thus

$$-\frac{dT}{dz} = \frac{g}{c_p} \equiv \Gamma_d \tag{2.48}$$

Hence, the dry adiabatic lapse rate is approximately constant throughout the lower atmosphere.

2.7.3 Static Stability

If potential temperature is a function of height, the atmospheric lapse rate, $\Gamma \equiv -\partial \mathrm{T}/\partial z$, will differ from the adiabatic lapse rate and

$$\frac{T}{\theta} \frac{\partial \theta}{\partial z} = \Gamma_d - \Gamma \tag{2.49}$$

If $\Gamma < \Gamma_d$ so that θ increases with height, an air parcel that undergoes an adiabatic displacement from its equilibrium level will be positively buoyant when displaced

downward and negatively buoyant when displaced upward so that it will tend to return to its equilibrium level and the atmosphere is said to be statically stable or *stably stratified*.

Adiabatic oscillations of a fluid parcel about its equilibrium level in a stably stratified atmosphere are referred to as *buoyancy oscillations*. The characteristic frequency of such oscillations can be derived by considering a parcel that is displaced vertically a small distance δz without disturbing its environment. If the environment is in hydrostatic balance, $\rho_0 g = -dp_0/dz$, where p_0 and ρ_0 are the pressure and density of the environment. The vertical acceleration of the parcel is

$$\frac{Dw}{Dt} = \frac{D^2}{Dt^2}(\delta z) = -g - \frac{1}{\rho}\frac{\partial p}{\partial z} \tag{2.50}$$

where p and ρ are the pressure and density of the parcel. In the parcel method it is assumed that the pressure of the parcel adjusts instantaneously to the environmental pressure during the displacement: $p = p_0$. This condition must be true if the parcel is to leave the environment undisturbed. Thus with the aid of the hydrostatic relationship, pressure can be eliminated in (2.50) to give

$$\frac{D^2}{Dt^2}(\delta z) = g\left(\frac{\rho_0 - \rho}{\rho}\right) = g\frac{\theta}{\theta_0} \tag{2.51}$$

where (2.44) and the ideal gas law have been used to express the buoyancy force in terms of potential temperature. Here θ designates the deviation of the potential temperature of the parcel from its basic state (environmental) value $\theta_0(z)$. If the parcel is initially at level $z = 0$ where the potential temperature is $\theta_0(0)$, then for a small displacement δz we can represent the environmental potential temperature as

$$\theta_0(\delta z) \approx \theta_0(0) + \left(d\theta_0/dz\right)\delta z$$

If the parcel displacement is adiabatic, the potential temperature of the parcel is conserved. Thus, $\theta(\delta z) = \theta_0(0) - \theta_0(\delta z) = -(d\theta_0/dz)\delta z$, and (2.51) becomes

$$\frac{D^2}{Dt^2}(\delta z) = -N^2\delta z \tag{2.52}$$

where

$$N^2 = g\frac{d\ln\theta_0}{dz}$$

is a measure of the static stability of the environment. Equation (2.52) has a general solution of the form $\delta z = A\exp(iNt)$. Therefore, if $N^2 > 0$, the parcel will oscillate about its initial level with a period $\tau = 2\pi/N$. The corresponding frequency N is

the *buoyancy frequency*.[3] For average tropospheric conditions, $N \approx 1.2 \times 10^{-1}\ \mathrm{s}^{-1}$ so that the period of a buoyancy oscillation is about 8 min.

In the case of $N = 0$, examination of (2.52) indicates that no accelerating force will exist and the parcel will be in neutral equilibrium at its new level. However, if $N^2 < 0$ (potential temperature decreasing with height) the displacement will increase exponentially in time. We thus arrive at the familiar gravitational or static stability criteria for dry air:

$d\theta_0/dz > 0$	statically stable,
$d\theta_0/dz = 0$	statically neutral,
$d\theta_0/dz < 0$	statically unstable.

On the synoptic scale the atmosphere is always stably stratified because any unstable regions that develop are stabilized quickly by convective overturning. For a moist atmosphere, the situation is more complicated and discussion of that situation will be deferred until Chapter 11.

2.7.4 Scale Analysis of the Thermodynamic Energy Equation

If potential temperature is divided into a basic state $\theta_0(z)$ and a deviation $\theta\ (x, y, z, t)$ so that the total potential temperature at any point is given by $\theta_{\mathrm{tot}} = \theta_0\,(z) + \theta\ (x, y, z, t)$, the first law of thermodynamics (2.46) can be written approximately for synoptic scaling as

$$\frac{1}{\theta_0}\left(\frac{\partial\theta}{\partial t} + u\frac{\partial\theta}{\partial x} + v\frac{\partial\theta}{\partial y}\right) + w\frac{d\ln\theta_0}{dz} = \frac{J}{c_p T} \tag{2.53}$$

where we have used the facts that for $|\theta/\theta_0| \ll 1$, $\left|d\theta\big/dz\right| \ll d\theta_0\big/dz$, and

$$\ln\theta_{tot} = \ln\left[\theta_0\left(1 + \theta\big/\theta_0\right)\right] \approx \ln\theta_0 + \theta/\theta_0$$

Outside regions of active precipitation, diabatic heating is due primarily to net radiative heating. In the troposphere, radiative heating is quite weak so that typically $J\big/c_p \leq 1^\circ\mathrm{C\ d}^{-1}$ (except near cloud tops, where substantially larger cooling can occur due to thermal emission by the cloud particles). The typical amplitude of horizontal potential temperature fluctuations in a midlatitude synoptic system (above the boundary layer) is $\theta \sim 4^\circ$C. Thus,

$$\frac{T}{\theta_0}\left(\frac{\partial\theta}{\partial t} + u\frac{\partial\theta}{\partial x} + v\frac{\partial\theta}{\partial y}\right) \sim \frac{\theta U}{L} \sim 4^\circ\mathrm{C\ d}^{-1}$$

[3] N is often referred to as the Brunt–Väisälä frequency.

The cooling due to vertical advection of the basic state potential temperature (usually called the adiabatic cooling) has a typical magnitude of

$$w\left(\frac{T}{\theta_0}\frac{d\theta_0}{dz}\right) = w\left(\Gamma_d - \Gamma\right) \sim 4^\circ\text{C d}^{-1}$$

where $w \sim 1$ cm s^{-1} and $\Gamma_d - \Gamma$, the difference between dry adiabatic and actual lapse rates, is $\sim 4^\circ$C km^{-1}.

Thus, in the absence of strong diabatic heating, the rate of change of the perturbation potential temperature is equal to the adiabatic heating or cooling due to vertical motion in the statically stable basic state, and (2.53) can be approximated as

$$\left(\frac{\partial\theta}{\partial t} + u\frac{\partial\theta}{\partial x} + v\frac{\partial\theta}{\partial y}\right) + w\frac{d\theta_0}{dz} \approx 0 \tag{2.54}$$

Alternatively, if the temperature field is divided into a basic state $T_0\,(z)$ and a deviation $T\,(x, y, z, t)$, then since $\theta/\theta_0 \approx T/T_0$, (2.54) can be expressed to the same order of approximation in terms of temperature as

$$\left(\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y}\right) + w\left(\Gamma_d - \Gamma\right) \approx 0 \tag{2.55}$$

PROBLEMS

2.1. A ship is steaming northward at a rate of 10 km h^{-1}.The surface pressure increases toward the northwest at the rate of 5 Pa km^{-1}. What is the pressure tendency recorded at a nearby island station if the pressure aboard the ship decreases at a rate of 100 Pa/3 h?

2.2. The temperature at a point 50 km north of a station is 3°C cooler than at the station. If the wind is blowing from the northeast at 20 m s^{-1} and the air is being heated by radiation at the rate of 1°C h^{-1}, what is the local temperature change at the station?

2.3. Derive the relationship

$$\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}) = -\Omega^2\mathbf{R}$$

which was used in Eq. (2.7).

2.4. Derive the expression given in Eq. (2.13) for the rate of change of **k** following the motion.

2.5. Suppose a 1-kg parcel of dry air is rising at a constant vertical velocity. If the parcel is being heated by radiation at the rate of 10^{-1}W kg^{-1}, what must the speed of rise be to maintain the parcel at a constant temperature?

2.6. Derive an expression for the density ρ that results when an air parcel initially at pressure p_s and density ρ_s expands adiabatically to pressure p.

2.7. An air parcel that has a temperature of 20°C at the 1000-hPa level is lifted dry adiabatically. What is its density when it reaches the 500-hPa level?

2.8. Suppose an air parcel starts from rest at the 800-hPa level and rises vertically to 500 hPa while maintaining a constant 1°C temperature excess over the environment. Assuming that the mean temperature of the 800- to 500-hPa layer is 260 K, compute the energy released due to the work of the buoyancy force. Assuming that all the released energy is realized as kinetic energy of the parcel, what will the vertical velocity of the parcel be at 500 hPa?

2.9. Show that for an atmosphere with an adiabatic lapse rate (i.e., constant potential temperature) the geopotential height is given by

$$Z = H_\theta \left[1 - \left(p/p_0 \right)^{R/c_p} \right]$$

where p_0 is the pressure at $Z = 0$ and $\mathrm{H}_\theta \equiv \mathrm{c}_p\, \theta / g_0$ is the total geopotential height of the atmosphere.

2.10. In the *isentropic* coordinate system (see Section 4.6), potential temperature is used as the vertical coordinate. Because potential temperature in adiabatic flow is conserved following the motion, isentropic coordinates are useful for tracing the actual paths of travel of individual air parcels. Show that the transformation of the horizontal pressure gradient force from z to θ coordinates is given by

$$\frac{1}{\rho} \nabla_z p = \nabla_\theta M$$

where $M \equiv \mathrm{c}_p\, \mathrm{T} + \Phi$ is the *Montgomery streamfunction*.

2.11. French scientists have developed a high-altitude balloon that remains at constant potential temperature as it circles the earth. Suppose such a balloon is in the lower equatorial stratosphere where the temperature is isothermal at 200 K. If the balloon were displaced vertically from its equilibrium level by a small distance δz it would tend to oscillate about the equilibrium level. What is the period of this oscillation?

2.12. Derive the approximate thermodynamic energy equation (2.55) by using the scaling arguments of Sections 2.4 and 2.7.

MATLAB EXERCISES

M2.1. The MATLAB script **standard_T_p.m** defines and plots the temperature and the lapse rate associated with the U.S. Standard Atmosphere as

functions of height. Modify this script to compute the pressure and potential temperature and plot these in the same format used for temperature and lapse rate. Hint: to compute pressure, integrate the hydrostatic equation from the surface upward in increments of δz. Show that if we define a mean scale height for the layer between z and $z + \delta z$ by letting $H = R\left[T(z) + T(z+\delta z)\right]/(2g)$, then $p(z+\delta z) = p(z)\exp\left[-\delta z/H\right]$. (Note that as you move upward layer by layer you must use the local height-dependent value of H in this formula.)

M2.2. The MATLAB script **thermo_profile.m** is a simple script to read in data giving pressure and temperature for a tropical mean sounding. Run this script to plot temperature versus pressure for data in the file **tropical_temp.dat**. Use the hypsometric equation to compute the geopotential height corresponding to each pressure level of the data file. Compute the corresponding potential temperature and plot graphs of the temperature and potential temperature variations with pressure and with geopotential height.

Suggested References

Salby, *Fundamentals of Atmospheric Physics*, contains a thorough development of the basic conservation laws at the graduate level.

Pedlosky, *Geophysical Fluid Dynamics*, discusses the equations of motion for a rotating coordinate system and has a thorough discussion of scale analysis at a graduate level.

Curry and Webster, *Thermodynamics of Atmospheres and Oceans* contains an excellent treatment of atmospheric thermodynamics.

CHAPTER 3

Elementary Applications of the Basic Equations

In addition to the geostrophic wind, which was discussed in Chapter 2, there are other approximate expressions for the relationships among velocity, pressure, and temperature fields, which are useful in the analysis of weather systems. These are discussed most conveniently using a coordinate system in which pressure is the vertical coordinate. Thus, before introducing the elementary applications of the present chapter it is useful to present the dynamical equations in isobaric coordinates.

3.1 BASIC EQUATIONS IN ISOBARIC COORDINATES

3.1.1 The Horizontal Momentum Equation

The approximate horizontal momentum equations (2.24) and (2.25) may be written in vectorial form as

$$\frac{D\mathbf{V}}{Dt} + f\mathbf{k} \times \mathbf{V} = -\frac{1}{\rho}\boldsymbol{\nabla}_p \tag{3.1}$$

where $\mathbf{V} = \mathbf{i}u + \mathbf{j}v$ is the *horizontal* velocity vector. In order to express (3.1) in isobaric coordinate form, we transform the pressure gradient force using (1.20) and (1.21) to obtain

$$\frac{D\mathbf{V}}{Dt} + f\mathbf{k} \times \mathbf{V} = -\boldsymbol{\nabla}_p \Phi \tag{3.2}$$

where $\boldsymbol{\nabla}_p$is the horizontal gradient operator applied with pressure held constant.

Because p is the independent vertical coordinate, we must expand the total derivative as

$$\begin{aligned} \frac{D}{Dt} &\equiv \frac{\partial}{\partial t} + \frac{Dx}{Dt}\frac{\partial}{\partial x} + \frac{Dy}{Dt}\frac{\partial}{\partial y} + \frac{Dp}{Dt}\frac{\partial}{\partial p} \\ &= \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + \omega\frac{\partial}{\partial p} \end{aligned} \tag{3.3}$$

Here $\omega \equiv Dp/Dt$ (usually called the "omega" vertical motion) is the pressure change following the motion, which plays the same role in the isobaric coordinate system that $w \equiv Dz/Dt$ plays in height coordinates.

From (3.2) we see that the isobaric coordinate form of the geostrophic relationship is

$$f\mathbf{V}_g = \mathbf{k} \times \boldsymbol{\nabla}_p \Phi \tag{3.4}$$

One advantage of isobaric coordinates is easily seen by comparing (2.23) and (3.4). In the latter equation, density does not appear. Thus, a given geopotential gradient implies the same geostrophic wind at any height, whereas a given horizontal pressure gradient implies different values of the geostrophic wind depending on the density. Furthermore, if f is regarded as a constant, the horizontal divergence of the geostrophic wind at constant pressure is zero:

$$\boldsymbol{\nabla}_p \bullet \mathbf{V}_g = 0$$

3.1.2 The Continuity Equation

It is possible to transform the continuity equation (2.31) from height coordinates to pressure coordinates. However, it is simpler to directly derive the isobaric form by considering a Lagrangian control volume $\delta V = \delta x\ \delta y\ \delta z$ and applying the hydrostatic equation $\delta p = -\rho g \delta z$ (note that $\delta p < 0$) to express the volume element as $\delta V = -\delta x \delta y \delta p/(\rho g)$. The mass of this fluid element, which is conserved following the motion, is then $\delta M = \rho \delta V = -\delta x \delta y \delta p/g$. Thus,

$$\frac{1}{\delta M}\frac{D}{Dt}(\delta M) = \frac{g}{\delta x \delta y \delta p}\frac{D}{Dt}\left(\frac{\delta x \delta y \delta p}{g}\right) = 0$$

After differentiating, using the chain rule, and changing the order of the differential operators we obtain[1]

$$\frac{1}{\delta x}\delta\left(\frac{Dx}{Dt}\right)+\frac{1}{\delta y}\delta\left(\frac{Dy}{Dt}\right)+\frac{1}{\delta p}\delta\left(\frac{Dp}{Dt}\right)=0$$

or

$$\frac{\delta u}{\delta x}+\frac{\delta v}{\delta y}+\frac{\delta \omega}{\delta p}=0$$

Taking the limit $\delta x, \delta y, \delta p \to 0$ and observing that δx and δy are evaluated at constant pressure, we obtain the continuity equation in the isobaric system:

$$\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)_p+\frac{\partial \omega}{\partial p}=0 \tag{3.5}$$

This form of the continuity equation contains no reference to the density field and does not involve time derivatives. The simplicity of (3.5) is one of the chief advantages of the isobaric coordinate system.

3.1.3 The Thermodynamic Energy Equation

The first law of thermodynamics (2.42) can be expressed in the isobaric system by letting $Dp/Dt = \omega$ and expanding DT/Dt by using (3.3):

$$c_p\left(\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}+\omega\frac{\partial T}{\partial p}\right)-\alpha\omega=J$$

This may be rewritten as

$$\left(\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}\right)-S_p\omega=\frac{J}{c_p} \tag{3.6}$$

where, with the aid of the equation of state and Poisson's equation (2.44), we have

$$S_p \equiv \frac{RT}{c_p p}-\frac{\partial T}{\partial p}=-\frac{T}{\theta}\frac{\partial \theta}{\partial p} \tag{3.7}$$

which is the static stability parameter for the isobaric system. Using (2.49) and the hydrostatic equation, (3.7) may be rewritten as

$$S_p=\left(\Gamma_d-\Gamma\right)/\rho g$$

Thus, S_p is positive provided that the lapse rate is less than dry adiabatic. However, because density decreases approximately exponentially with height, S_p increases rapidly with height. This strong height dependence of the stability measure S_p is a minor disadvantage of isobaric coordinates.

[1] From now on g will be regarded as a constant.

3.2 BALANCED FLOW

Despite the apparent complexity of atmospheric motion systems as depicted on synoptic weather charts, the pressure (or geopotential height) and velocity distributions in meteorological disturbances are actually related by rather simple approximate force balances. In order to gain a qualitative understanding of the horizontal balance of forces in atmospheric motions, we idealize by considering flows that are steady state (i.e., time independent) and have no vertical component of velocity. Furthermore, it is useful to describe the flow field by expanding the isobaric form of the horizontal momentum equation (3.2) into its components in a so-called *natural* coordinate system.

3.2.1 Natural Coordinates

The natural coordinate system is defined by the orthogonal set of unit vectors **t**, **n**, and **k**. Unit vector **t** is oriented parallel to the horizontal velocity at each point; unit vector **n** is normal to the horizontal velocity and is oriented so that it is positive to the left of the flow direction; and unit vector **k** is directed vertically upward. In this system the horizontal velocity may be written $\mathbf{V} = V\mathbf{t}$ where V, the horizontal speed, is a nonnegative scalar defined by $V \equiv Ds/Dt$, where $s(x, y, t)$ is the distance along the curve followed by a parcel moving in the horizontal plane. The acceleration following the motion is thus

$$\frac{D\mathbf{V}}{Dt} = \frac{D(V\mathbf{t})}{Dt} = \mathbf{t}\frac{DV}{Dt} + V\frac{D\mathbf{t}}{Dt}$$

The rate of change of **t** following the motion may be derived from geometrical considerations with the aid of Fig. 3.1:

$$\delta\psi = \frac{\delta s}{|R|} = \frac{|\delta\mathbf{t}|}{|\mathbf{t}|} = |\delta\mathbf{t}|$$

Here R is the *radius of curvature* following the parcel motion, and we have used the fact that $|\mathbf{t}| = 1$. By convention, R is taken to be positive when the center of curvature is in the positive **n** direction. Thus, for $R > 0$, the air parcels turn toward the left following the motion, and for $R < 0$ the air parcels turn toward the right following the motion.

Noting that in the limit $\delta s \to 0$, $\delta\mathbf{t}$ is directed parallel to **n**, the above relationship yields $D\mathbf{t}/Ds = \mathbf{n}/R$. Thus,

$$\frac{D\mathbf{t}}{Dt} = \frac{D\mathbf{t}}{Ds}\frac{Ds}{Dt} = \frac{\mathbf{n}}{R}V$$

and

$$\frac{D\mathbf{V}}{Dt} = \mathbf{t}\frac{DV}{Dt} + \mathbf{n}\frac{V^2}{R} \tag{3.8}$$

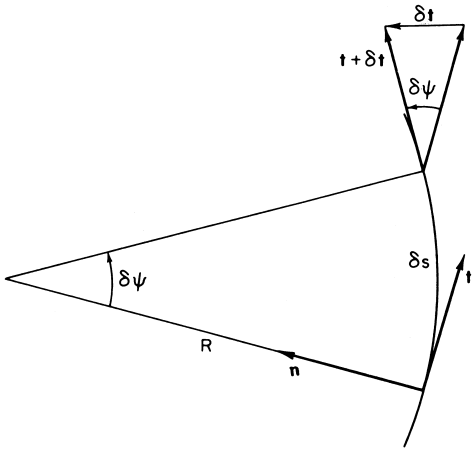


Fig. 3.1 Rate of change of the unit tangent vector **t** following the motion.

Therefore, the acceleration following the motion is the sum of the rate of change of speed of the air parcel and its centripetal acceleration due to the curvature of the trajectory. Because the Coriolis force always acts normal to the direction of motion, its natural coordinate form is simply

$$-f\,\mathbf{k} \times \mathbf{V} = -f\,V\mathbf{n}$$

whereas the pressure gradient force can be expressed as

$$-\boldsymbol{\nabla}_p \Phi = -\left(\mathbf{t}\frac{\partial \Phi}{\partial s} + \mathbf{n}\frac{\partial \Phi}{\partial n}\right)$$

The horizontal momentum equation may thus be expanded into the following component equations in the natural coordinate system:

$$\frac{DV}{Dt} = -\frac{\partial \Phi}{\partial s} \tag{3.9}$$

$$\frac{V^2}{R} + fV = -\frac{\partial \Phi}{\partial n} \tag{3.10}$$

Equations (3.9) and (3.10) express the force balances parallel to and normal to the direction of flow, respectively. For motion parallel to the geopotential height contours, $\partial \Phi / \partial s = 0$ and the speed is constant following the motion. If, in addition, the geopotential gradient normal to the direction of motion is constant along a trajectory, (3.10) implies that the radius of curvature of the trajectory is also constant. In that case the flow can be classified into several simple categories depending on the relative contributions of the three terms in (3.10) to the net force balance.

3.2.2 Geostrophic Flow

Flow in a straight line ($R \rightarrow \pm \infty$) parallel to height contours is referred to as *geostrophic motion*. In geostrophic motion the horizontal components of the Coriolis force and pressure gradient force are in exact balance so that $V = V_g$, where the geostrophic wind V_g is defined by[2]

$$fV_g = -\partial\Phi/\partial n \tag{3.11}$$

This balance is indicated schematically in Fig. 3.2. The actual wind can be in exact geostrophic motion only if the height contours are parallel to latitude circles. As discussed in Section 2.4.1, the geostrophic wind is generally a good approximation to the actual wind in extratropical synoptic-scale disturbances. However, in some of the special cases treated later this is not true.

3.2.3 Inertial Flow

If the geopotential field is uniform on an isobaric surface so that the horizontal pressure gradient vanishes, (3.10) reduces to a balance between Coriolis force and centrifugal force:

$$V^2/R + fV = 0 \tag{3.12}$$

Equation (3.12) may be solved for the radius of curvature

$$R = -V/f$$

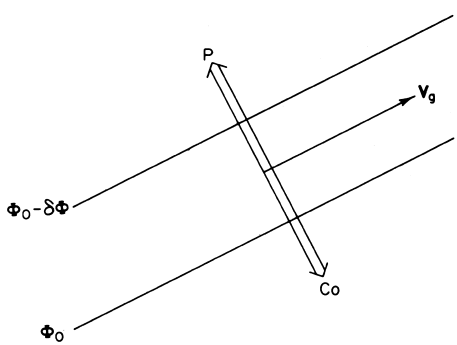


Fig. 3.2 Balance of forces for geostrophic equilibrium. The pressure gradient force is designated by *P* and the Coriolis force by *Co*.

[2] Note that although the actual speed V must always be positive in the natural coordinates, V_g, which is proportional to the height gradient normal to the direction of flow, may be negative, as in the "anomalous" low shown in Fig. 3.5c.

Since from (3.9), the speed must be constant in this case, the radius of curvature is also constant (neglecting the latitudinal dependence of f). Thus the air parcels follow circular paths in an anticyclonic sense.[3] The period of this oscillation is

$$P = \left| \frac{2\pi R}{V} \right| = \frac{2\pi}{|f|} = \frac{\frac{1}{2} day}{|\sin \phi|} \tag{3.13}$$

P is equivalent to the time that is required for a Foucault pendulum to turn through an angle of 180°. Hence, it is often referred to as one-half *pendulum day*.

Because both the Coriolis force and the centrifugal force due to the relative motion are caused by inertia of the fluid, this type of motion is traditionally referred to as an *inertial oscillation*, and the circle of radius $|R|$ is called the inertia circle. It is important to realize that the "inertial flow" governed by (3.12) is not the same as inertial motion in an absolute reference frame. The flow governed by (3.12) is just the constant angular momentum oscillation referred to in Section 1.5.4. In this flow the force of gravity, acting orthogonal to the plane of motion, keeps the oscillation on a horizontal surface. In true inertial motion, all forces vanish and the motion maintains a uniform absolute velocity.

In the atmosphere, motions are nearly always generated and maintained by pressure gradient forces; the conditions of uniform pressure required for pure inertial flow rarely exist. In the oceans, however, currents are often generated by transient winds blowing across the surface, rather than by internal pressure gradients. As a result, significant amounts of energy occur in currents that oscillate with near inertial periods. An example recorded by a current meter near the island of Barbados is shown in Fig. 3.3.

3.2.4 Cyclostrophic Flow

If the horizontal scale of a disturbance is small enough, the Coriolis force may be neglected in (3.10) compared to the pressure gradient force and the centrifugal force. The force balance normal to the direction of flow is then

$$\frac{V^2}{R} = -\frac{\partial \Phi}{\partial n}$$

If this equation is solved for V, we obtain the speed of the *cyclostrophic wind*

$$V = \left(-R \frac{\partial \Phi}{\partial n} \right)^{1/2} \tag{3.14}$$

[3] Anticyclonic flow is a clockwise rotation in the Northern Hemisphere and counterclockwise in the Southern Hemisphere. Cyclonic flow has the opposite sense of rotation in each hemisphere.

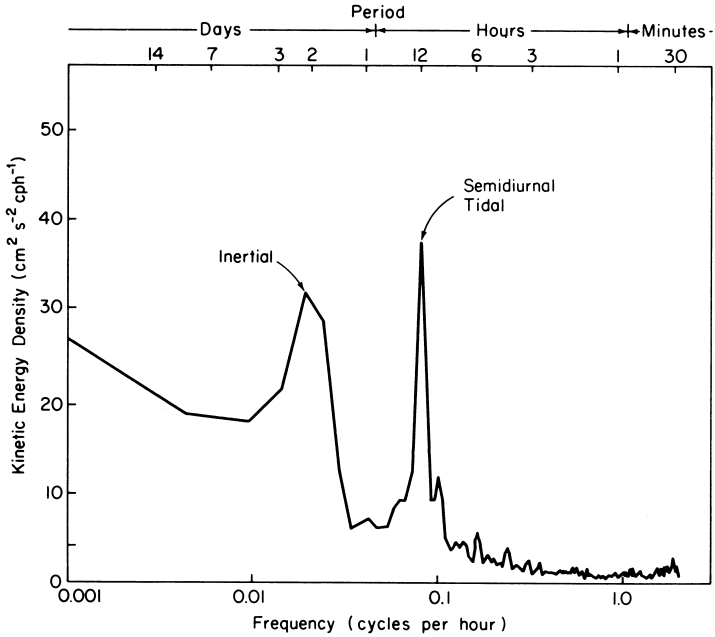


Fig. 3.3 Power spectrum of kinetic energy at 30-m depth in the ocean near Barbados (13°N). Ordinate shows kinetic energy density per unit frequency interval (cph^{-1} designates cycles per hour). This type of plot indicates the manner in which the total kinetic energy is partitioned among oscillations of different periods. Note the strong peak at 53 h, which is the period of an inertial oscillation at 13° latitude. [After Warsh et al., (1971.) Reproduced with permission of the American Meteorological Society.]

As indicated in Fig. 3.4, cyclostrophic flow may be either cyclonic or anticyclonic. In both cases the pressure gradient force is directed toward the center of curvature, and the centrifugal force away from the center of curvature.

The cyclostrophic balance approximation is valid provided that the ratio of the centrifugal force to the Coriolis force is large. This ratio $V/(fR)$ is equivalent to the Rossby number discussed in Section 2.4.2. As an example of cyclostrophic scale motion we consider a typical tornado. Suppose that the tangential velocity is $30\ \text{m s}^{-1}$ at a distance of 300 m from the center of the vortex. Assuming that $f = 10^{-4}\ \text{s}^{-1}$, the Rossby number is just $\text{Ro} = V/|fR| \approx 10^3$, which implies that the Coriolis force can be neglected in computing the balance of forces for a tornado. However, the majority of tornadoes in the Northern Hemisphere are observed to rotate in a cyclonic (counterclockwise) sense. This is apparently because they are embedded in environments that favor cyclonic rotation (see Section 9.6.1). Smaller scale vortices, however, such as dust devils and water spouts, do not have

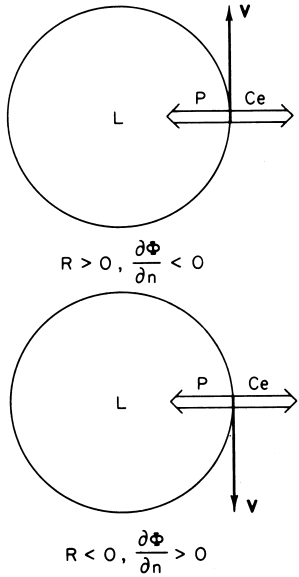


Fig. 3.4 Force balance in cyclostrophic flow: P designates the pressure gradient and Ce the centrifugal force.

a preferred direction of rotation. According to data collected by Sinclair (1965), they are observed to be anticyclonic as often as cyclonic.

3.2.5 The Gradient Wind Approximation

Horizontal frictionless flow that is parallel to the height contours so that the tangential acceleration vanishes ($DV/Dt = 0$) is called *gradient flow*. Gradient flow is a three-way balance among the Coriolis force, the centrifugal force, and the horizontal pressure gradient force. Like geostrophic flow, pure gradient flow can exist only under very special circumstances. It is always possible, however, to define a gradient wind, which at any point is just the wind component parallel to the height contours that satisfies (3.10). For this reason, (3.10) is commonly referred to as the gradient wind equation. Because (3.10) takes into account the centrifugal force due to the curvature of parcel trajectories, the gradient wind is often a better approximation to the actual wind than the geostrophic wind.

The gradient wind speed is obtained by solving (3.10) for V to yield

$$\begin{aligned} V &= -\frac{fR}{2} \pm \left(\frac{f^2R^2}{4} - R\frac{\partial \Phi}{\partial n} \right)^{1/2} \\ &= -\frac{fR}{2} \pm \left(\frac{f^2R^2}{4} + fRV_g \right)^{1/2} \end{aligned} \tag{3.15}$$

where in the lower expression (3.11) is used to express $\partial\Phi/\partial n$ in terms of the geostrophic wind.

Not all the mathematically possible roots of (3.15) correspond to physically possible solutions, as it is required that V be real and nonnegative. In Table 3.1 the various roots of (3.15) are classified according to the signs of R and $\partial\Phi/\partial n$ in order to isolate the physically meaningful solutions

The force balances for the four permitted solutions are illustrated in Fig. 3.5. Equation (3.15) shows that in cases of both regular and anomalous highs the pressure gradient is limited by the requirement that the quantity under the radical be nonnegative; that is,

$$\left| fV_g \right| = \left| \frac{\partial\Phi}{\partial n} \right| < \frac{|R| f^2}{4} \tag{3.16}$$

Thus, the pressure gradient in a high must approach zero as $|R| \to 0$. It is for this reason that the pressure field near the center of a high is always flat and the wind gentle compared to the region near the center of a low.

The absolute angular momentum about the axis of rotation for the circularly symmetric motions shown in Fig. 3.5 is given by $VR + fR^2/2$. From (3.15) it is verified readily that regular gradient wind balances have positive absolute angular momentum in the Northern Hemisphere, whereas anomalous cases have negative absolute angular momentum. Because the only source of negative absolute angular momentum is the Southern Hemisphere, the anomalous cases are unlikely to occur except perhaps close to the equator.

In all cases except the anomalous low (Fig. 3.5c) the horizontal components of the Coriolis and pressure gradient forces are oppositely directed. Such flow is

Table 3.1 *Classification of Roots of the Gradient Wind Equation in the Northern Hemisphere*

Sign $\partial\Phi/\partial n$	R > 0	R < 0
Positive ($V_g < 0$)	Positive root:[a] unphysical	Positive root: antibaric flow (anomalous low)
	Negative root: unphysical	Negative root: unphysical
Negative ($V_g > 0$)	Positive root: cyclonic flow (regular low)	Positive root: ($V > -fR/2$): anticyclonic flow (anomalous high)
	Negative root: unphysical	Negative root: ($V < -fR/2$): anticyclonic flow (regular high)

[a]The terms "positive root" and "negative root" in columns 2 and 3 refer to the sign taken in the final term in Eq. (3.15).

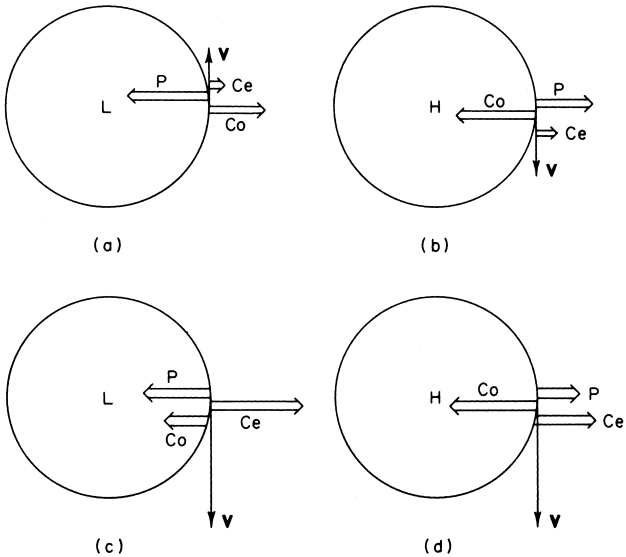


Fig. 3.5 Force balances in the Northern Hemisphere for the four types of gradient flow: (a) regular low (b) regular high (c) anomalous low (d) anomalous high.

called *baric*. The anomalous low is antibaric; the geostrophic wind V_g defined in (3.11) is negative for an anomalous low and is clearly not a useful approximation to the actual speed.[4] Furthermore, as shown in Fig. 3.5, gradient flow is cyclonic only when the centrifugal force and the horizontal component of the Coriolis force have the same sense ($Rf > 0$); it is anticyclonic when these forces have the opposite sense ($Rf < 0$). Since the direction of anticyclonic and cyclonic flow is reversed in the Southern Hemisphere, the requirement that $Rf > 0$ for cyclonic flow holds regardless of the hemisphere considered.

The definition of the geostrophic wind (3.11) can be used to rewrite the force balance normal to the direction of flow (3.10) in the form

$$V^2/R + fV - fV_g = 0$$

Dividing through by fV shows that the ratio of the geostrophic wind to the gradient wind is

$$\frac{V_g}{V} = 1 + \frac{V}{fR} \tag{3.17}$$

For normal cyclonic flow ($fR > 0$), V_g is larger than V, whereas for anticyclonic flow ($fR < 0$), V_g is smaller than V. Therefore, the geostrophic wind is an

[4] Remember that in the natural coordinate system the speed V is positive definite.

overestimate of the balanced wind in a region of cyclonic curvature and an underestimate in a region of anticyclonic curvature. For midlatitude synoptic systems, the difference between gradient and geostrophic wind speeds generally does not exceed 10–20%. [Note that the magnitude of $V/(fR)$ is just the Rossby number.] For tropical disturbances, the Rossby number is in the range of 1–10, and the gradient wind formula must be applied rather than the geostrophic wind. Equation (3.17) also shows that the antibaric anomalous low, which has $V_g < 0$, can exist only when $V/(fR) < -1$. Thus, antibaric flow is associated with small-scale intense vortices such as tornadoes.

3.3 TRAJECTORIES AND STREAMLINES

In the natural coordinate system used in the previous section to discuss balanced flow, $s(x, y, t)$ was defined as the distance along the curve in the horizontal plane traced out by the path of an air parcel. The path followed by a particular air parcel over a finite period of time is called the *trajectory* of the parcel. Thus, the radius of curvature R of the path s referred to in the gradient wind equation is the radius of curvature for a parcel trajectory. In practice, R is often estimated by using the radius of curvature of a geopotential height contour, as this can be estimated easily from a synoptic chart. However, the height contours are actually *streamlines* of the gradient wind (i.e., lines that are everywhere parallel to the instantaneous wind velocity).

It is important to distinguish clearly between streamlines, which give a "snapshot" of the velocity field at any instant, and trajectories, which trace the motion of individual fluid parcels over a finite time interval. In Cartesian coordinates, horizontal trajectories are determined by the integration of

$$\frac{Ds}{Dt} = V(x, y, t) \tag{3.18}$$

over a finite time span for each parcel to be followed, whereas streamlines are determined by the integration of

$$\frac{dy}{dx} = \frac{v(x, y, t_0)}{u(x, y, t_0)} \tag{3.19}$$

with respect to x at time t_0. (Note that since a streamline is parallel to the velocity field, its slope in the horizontal plane is just the ratio of the horizontal velocity components.) Only for *steady-state* motion fields (i.e., fields in which the local rate of change of velocity vanishes) do the streamlines and trajectories coincide. However, synoptic disturbances are not steady-state motions. They generally move at speeds of the same order as the winds that circulate about them. In order to gain an appreciation for the possible errors involved in using the curvature of

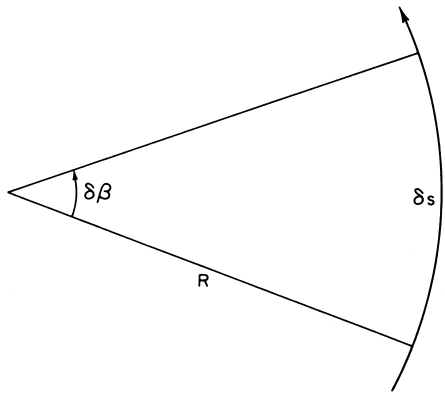


Fig. 3.6 Relationship between the change in angular direction of the wind $\delta\beta$ and the radius of curvature R.

the streamlines instead of the curvature of the trajectories in the gradient wind equation, it is necessary to investigate the relationship between the curvature of the trajectories and the curvature of the streamlines for a moving pressure system.

We let $\beta(x, y, t)$ designate the angular direction of the wind at each point on an isobaric surface, and R_t and R_s designate the radii of curvature of the trajectories and streamlines, respectively. Then, from Fig. 3.6, $\delta s = R\ \delta\beta$ so that in the limit $\delta s \to 0$

$$\frac{D\beta}{Ds} = \frac{1}{R_t} \qquad \text{and} \qquad \frac{\partial\beta}{\partial s} = \frac{1}{R_s} \tag{3.20}$$

where $D\beta/Ds$ means the rate of change of wind direction along a trajectory (positive for counterclockwise turning) and $\partial\beta/\partial s$ is the rate of change of wind direction along a streamline at any instant. Thus, the rate of change of wind direction following the motion is

$$\frac{D\beta}{Dt} = \frac{D\beta}{Ds}\frac{Ds}{Dt} = \frac{V}{R_t} \tag{3.21}$$

or, after expanding the total derivative,

$$\frac{D\beta}{Dt} = \frac{\partial\beta}{\partial t} + V\frac{\partial\beta}{\partial s} = \frac{\partial\beta}{\partial t} + \frac{V}{R_s} \tag{3.22}$$

Combining (3.21) and (3.22), we obtain a formula for the local turning of the wind:

$$\frac{\partial\beta}{\partial t} = V\left(\frac{1}{R_t} - \frac{1}{R_s}\right) \tag{3.23}$$

Equation (3.23) indicates that the trajectories and streamlines will coincide only when the local rate of change of the wind direction vanishes.

In general, midlatitude synoptic systems move eastward as a result of advection by upper level westerly winds. In such cases there is a local turning of the wind due to the motion of the system even if the shape of the height contour pattern remains constant as the system moves. The relationship between R_t and R_s in such a situation can be determined easily for an idealized circular pattern of height contours moving at a constant velocity **C**. In this case the local turning of the wind is entirely due to the motion of the streamline pattern so that

$$\frac{\partial \beta}{\partial t} = -\mathbf{C} \bullet \nabla \beta = -C \frac{\partial \beta}{\partial s} \cos \gamma = -\frac{C}{R_s} \cos \gamma$$

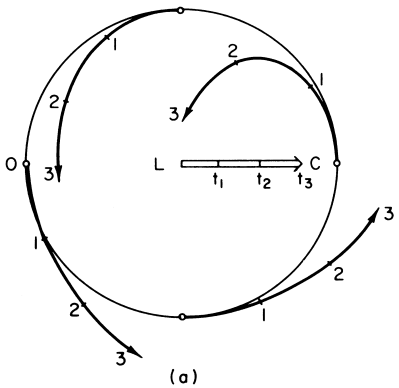
where γ is the angle between the streamlines (height contours) and the direction of motion of the system. Substituting the above into (3.23) and solving for R_t with the aid of (3.20), we obtain the desired relationship between the curvature of the streamlines and the curvature of the trajectories:

$$R_t = R_s \left(1 - \frac{C \cos \gamma}{V} \right)^{-1} \tag{3.24}$$

Equation (3.24) can be used to compute the curvature of the trajectory anywhere on a moving pattern of streamlines. In Fig. 3.7 the curvatures of the trajectories for parcels initially located due north, east, south, and west of the center of a cyclonic system are shown both for the case of a wind speed greater than the speed of movement of the height contours and for the case of a wind speed less than the speed of movement of the height contours. In these examples the plotted trajectories are based on a geostrophic balance so that the height contours are equivalent to streamlines. It is also assumed for simplicity that the wind speed does not depend on the distance from the center of the system. In the case shown in Fig. 3.7b there is a region south of the low center where the curvature of the trajectories is opposite that of the streamlines. Because synoptic-scale pressure systems usually move at speeds comparable to the wind speed, the gradient wind speed computed on the basis of the curvature of the height contours is often no better an approximation to the actual wind speed than the geostrophic wind. In fact, the actual gradient wind speed will vary along a height contour with the variation of the trajectory curvature.

3.4 THE THERMAL WIND

The geostrophic wind must have vertical shear in the presence of a horizontal temperature gradient, as can be shown easily from simple physical considerations based on hydrostatic equilibrium. Since the geostrophic wind (3.4) is proportional to the geopotential gradient on an isobaric surface, a geostrophic wind directed along the positive y axis that increases in magnitude with height requires that the



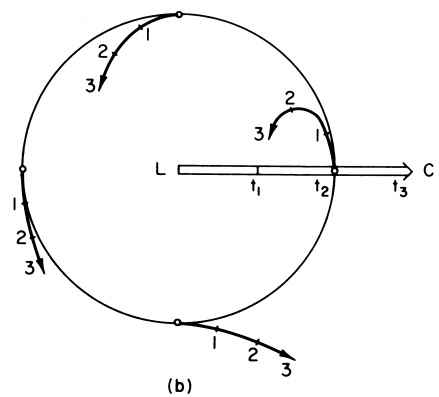


Fig. 3.7 Trajectories for moving circular cyclonic circulation systems in the Northern Hemisphere with (a) $V = 2C$ and (b) $2V = C$. Numbers indicate positions at successive times. The L designates a pressure minimum.

slope of the isobaric surfaces with respect to the x axis also must increase with height as shown in Fig. 3.8. According to the hypsometric Eq. (1.21), the thickness δz corresponding to a pressure interval δp is

$$\delta z \approx -g^{-1} RT \delta \ln p \tag{3.25}$$

Thus, the thickness of the layer between two isobaric surfaces is proportional to the temperature in the layer. In Fig. 3.8 the mean temperature T_1 of the column denoted by δz_1 must be less than the mean temperature T_2 for the column denoted by δz_2. Hence, an increase with height of a positive x directed pressure gradient must be associated with a positive x directed temperature gradient. The air in a vertical column at x_2, because it is warmer (less dense), must occupy a greater depth for a given pressure drop than the air at x_1.

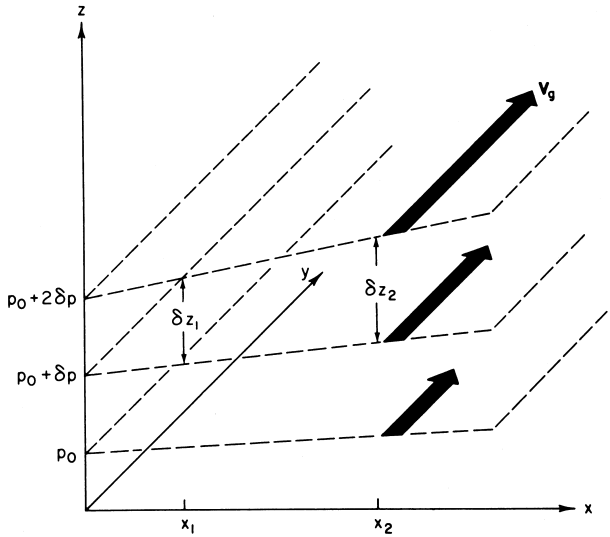


Fig. 3.8 Relationship between vertical shear of the geostrophic wind and horizontal thickness gradients. (Note that $\delta p < 0$.)

Equations for the rate of change with height of the geostrophic wind components are derived most easily using the isobaric coordinate system. In isobaric coordinates the geostrophic wind (3.4) has components given by

$$v_g = \frac{1}{f}\frac{\partial \Phi}{\partial x} \qquad \text{and} \qquad u_g = -\frac{1}{f}\frac{\partial \Phi}{\partial y} \tag{3.26}$$

where the derivatives are evaluated with pressure held constant. Also, with the aid of the ideal gas law we can write the hydrostatic equation as

$$\frac{\partial \Phi}{\partial p} = -\alpha = -\frac{RT}{p} \tag{3.27}$$

Differentiating (3.26) with respect to pressure and applying (3.27), we obtain

$$p\frac{\partial v_g}{\partial p} \equiv \frac{\partial v_g}{\partial \ln p} = -\frac{R}{f}\left(\frac{\partial T}{\partial x}\right)_p \tag{3.28}$$

$$p\frac{\partial u_g}{\partial p} \equiv \frac{\partial u_g}{\partial \ln p} = \frac{R}{f}\left(\frac{\partial T}{\partial y}\right)_p \tag{3.29}$$

or in vectorial form

$$\frac{\partial \mathbf{V}_g}{\partial \ln p} = -\frac{R}{f}\mathbf{k} \times \nabla_p T \tag{3.30}$$

Equation (3.30) is often referred to as the *thermal wind* equation. However, it is actually a relationship for the vertical wind *shear* (i.e., the rate of change of the geostrophic wind with respect to ln p). Strictly speaking, the term *thermal wind* refers to the vector difference between geostrophic winds at two levels. Designating the thermal wind vector by $\mathbf{V}_T$, we may integrate (3.30) from pressure level p_0 to level p_1 ($p_1 < p_0$) to get

$$\mathbf{V}_T \equiv \mathbf{V}_g(p_1) - \mathbf{V}_g(p_0) = -\frac{R}{f}\int_{p_0}^{p_1} \left(\mathbf{k} \times \boldsymbol{\nabla}_p T\right) d \ln p \tag{3.31}$$

Letting $\langle T \rangle$ denote the mean temperature in the layer between pressure p_0 and p_1, the x and y components of the thermal wind are thus given by

$$u_T = -\frac{R}{f}\left(\frac{\partial \langle T \rangle}{\partial y}\right)_p \ln\left(\frac{p_0}{p_1}\right); \quad v_T = \frac{R}{f}\left(\frac{\partial \langle T \rangle}{\partial x}\right)_p \ln\left(\frac{p_0}{p_1}\right) \tag{3.32}$$

Alternatively, we may express the thermal wind for a given layer in terms of the horizontal gradient of the geopotential difference between the top and the bottom of the layer:

$$u_T = -\frac{1}{f}\frac{\partial}{\partial y}(\Phi_1 - \Phi_0); \quad v_T = \frac{1}{f}\frac{\partial}{\partial x}(\Phi_1 - \Phi_0) \tag{3.33}$$

The equivalence of (3.32) and (3.33) can be verified readily by integrating the hydrostatic equation (3.27) vertically from p_0 to p_1 after replacing T by the mean $\langle T \rangle$. The result is the hypsometric equation (1.22):

$$\Phi_1 - \Phi_0 \equiv gZ_T = R\langle T \rangle \ln\left(\frac{p_0}{p_1}\right) \tag{3.34}$$

The quantity Z_T is the *thickness* of the layer between p_0 and p_1 measured in units of geopotential meters. From (3.34) we see that the thickness is proportional to the mean temperature in the layer. Hence, lines of equal Z_T (isolines of thickness) are equivalent to the isotherms of mean temperature in the layer.

The thermal wind equation is an extremely useful diagnostic tool, which is often used to check analyses of the observed wind and temperature fields for consistency. It can also be used to estimate the mean horizontal temperature advection in a layer as shown in Fig. 3.9. It is clear from the vector form of the thermal wind relation:

$$\mathbf{V}_T = \frac{1}{f}\mathbf{k} \times \boldsymbol{\nabla}(\Phi_1 - \Phi_0) = \frac{g}{f}\mathbf{k} \times \boldsymbol{\nabla} Z_T = \frac{R}{f}\mathbf{k} \times \boldsymbol{\nabla}\langle T \rangle \ln\left(\frac{p_0}{p_1}\right) \tag{3.35}$$

that the thermal wind blows parallel to the isotherms (lines of constant thickness) with the warm air to the right facing downstream in the Northern Hemisphere.

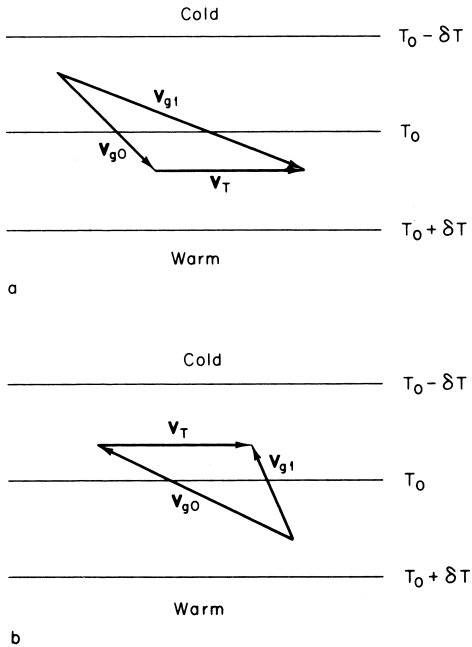


Fig. 3.9 Relationship between turning of geostrophic wind and temperature advection: (a) backing of the wind with height and (b) veering of the wind with height.

Thus, as is illustrated in Fig. 3.9a, a geostrophic wind that turns counterclockwise with height (backs) is associated with cold-air advection. Conversely, as shown in Fig. 3.9b, clockwise turning (veering) of the geostrophic wind with height implies warm advection by the geostrophic wind in the layer. It is therefore possible to obtain a reasonable estimate of the horizontal temperature advection and its vertical dependence at a given location solely from data on the vertical profile of the wind given by a single sounding. Alternatively, the geostrophic wind at any level can be estimated from the mean temperature field, provided that the geostrophic velocity is known at a single level. Thus, for example, if the geostrophic wind at 850 hPa is known and the mean horizontal temperature gradient in the layer 850–500 hPa is also known, the thermal wind equation can be applied to obtain the geostrophic wind at 500 hPa.

3.4.1 Barotropic and Baroclinic Atmospheres

A *barotropic* atmosphere is one in which the density depends only on the pressure, $\rho = \rho(p)$, so that isobaric surfaces are also surfaces of constant density. For an ideal gas, the isobaric surfaces will also be isothermal if the atmosphere is barotropic. Thus, $\nabla_p T = 0$ in a barotropic atmosphere, and the thermal wind equation (3.30) becomes $\partial \mathbf{V}_g / \partial \ln p = 0$, which states that the geostrophic wind

is independent of height in a barotropic atmosphere. Thus, barotropy provides a very strong constraint on the motions in a rotating fluid; the large-scale motion can depend only on horizontal position and time, not on height.

An atmosphere in which density depends on both the temperature and the pressure, $\rho = \rho(p, T)$, is referred to as a *baroclinic* atmosphere. In a baroclinic atmosphere the geostrophic wind generally has vertical shear, and this shear is related to the horizontal temperature gradient by the thermal wind equation (3.30). Obviously, the baroclinic atmosphere is of primary importance in dynamic meteorology. However, as shown in later chapters, much can be learned by study of the simpler barotropic atmosphere.

3.5 VERTICAL MOTION

As mentioned previously, for synoptic-scale motions the vertical velocity component is typically of the order of a few centimeters per second. Routine meteorological soundings, however, only give the wind speed to an accuracy of about a meter per second. Thus, in general the vertical velocity is not measured directly but must be inferred from the fields that are measured directly.

Two commonly used methods for inferring the vertical motion field are the kinematic method, based on the equation of continuity, and the adiabatic method, based on the thermodynamic energy equation. Both methods are usually applied using the isobaric coordinate system so that $\omega(p)$ is inferred rather than $w(z)$. These two measures of vertical motion can be related to each other with the aid of the hydrostatic approximation.

Expanding Dp/Dt in the (x, y, z) coordinate system yields

$$\omega \equiv \frac{Dp}{Dt} = \frac{\partial p}{\partial t} + \mathbf{V} \bullet \nabla p + w\left(\frac{\partial p}{\partial z}\right) \tag{3.36}$$

Now, for synoptic-scale motions, the horizontal velocity is geostrophic to a first approximation. Therefore, we can write $\mathbf{V} = \mathbf{V}_g + \mathbf{V}_a$, where $\mathbf{V}_a$ is the *ageostrophic* wind and $|\mathbf{V}_a| \ll |\mathbf{V}_g|$. However, $\mathbf{V}_g = (\rho f)^{-1}\mathbf{k} \times \nabla p$, so that $\mathbf{V}_g \bullet \nabla p = 0$. Using this result plus the hydrostatic approximation, (3.36) may be rewritten as

$$\omega = \frac{\partial p}{\partial t} + \mathbf{V}_a \bullet \nabla p - g\rho w \tag{3.37}$$

Comparing the magnitudes of the three terms on the right in (3.37), we find that for synoptic-scale motions

$$\partial p/\partial t \sim 10 \text{ hPa d}^{-1}$$
$$\mathbf{V}_a \bullet \nabla p \sim \left(1 \text{ m s}^{-1}\right)\left(1 \text{ Pa km}^{-1}\right) \sim 1 \text{ hPa d}^{-1}$$
$$g\rho w \sim 100 \text{ hPa d}^{-1}$$

Thus, it is quite a good approximation to let

$$\omega = -\rho g w \tag{3.38}$$

3.5.1 The Kinematic Method

One method of deducing the vertical velocity is based on integrating the continuity equation in the vertical. Integration of (3.5) with respect to pressure from a reference level p_s to any level p yields

$$\begin{aligned} \omega(p) &= \omega(p_s) - \int_{p_s}^{p} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)_p dp \\ &= \omega(p_s) + (p_s - p) \left(\frac{\partial \langle u \rangle}{\partial x} + \frac{\partial \langle v \rangle}{\partial y} \right)_p \end{aligned} \tag{3.39}$$

Here the angle brackets denote a pressure-weighted vertical average:

$$\langle\ \rangle \equiv (p - p_s)^{-1} \int_{p_s}^{p} ()dp$$

With the aid of (3.38), the averaged form of (3.39) can be rewritten as

$$w(z) = \frac{\rho(z_s)\, w(z_s)}{\rho(z)} - \frac{p_s - p}{\rho(z) g} \left(\frac{\partial \langle u \rangle}{\partial x} + \frac{\partial \langle v \rangle}{\partial y} \right) \tag{3.40}$$

where z and z_s are the heights of pressure levels p and p_s, respectively.

Application of (3.40) to infer the vertical velocity field requires knowledge of the horizontal divergence. In order to determine the horizontal divergence, the partial derivatives $\partial u/\partial x$ and $\partial v/\partial y$ are generally estimated from the fields of u and v by using *finite difference* approximations (see Section 13.3.1). For example, to determine the divergence of the horizontal velocity at the point x_0, y_0 in Fig. 3.10 we write

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \approx \frac{u(x_0 + d) - u(x_0 - d)}{2d} + \frac{v(y_0 + d) - v(y_0 - d)}{2d} \tag{3.41}$$

However, for synoptic-scale motions in midlatitudes, the horizontal velocity is nearly in geostrophic equilibrium. Except for the small effect due to the variation of the Coriolis parameter (see Problem 3.19), the geostrophic wind is nondivergent; that is, $\partial u/\partial x$ and $\partial v/\partial y$ are nearly equal in magnitude but opposite in sign. Thus, the horizontal divergence is due primarily to the small departures of the wind from geostrophic balance (i.e., the ageostrophic wind). A 10% error in evaluating one of the wind components in (3.41) can easily cause the estimated divergence to be in error by 100%. For this reason, the continuity equation method is not recommended for estimating the vertical motion field from observed horizontal winds.

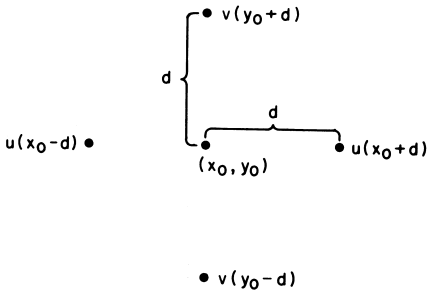


Fig. 3.10 Grid for estimation of the horizontal divergence.

3.5.2 The Adiabatic Method

A second method for inferring vertical velocities, which is not so sensitive to errors in the measured horizontal velocities, is based on the thermodynamic energy equation. If the diabatic heating J is small compared to the other terms in the heat balance, (3.6) yields

$$\omega = S_p^{-1}\left(\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y}\right) \tag{3.42}$$

Because temperature advection can usually be estimated quite accurately in midlatitudes by using geostrophic winds, the adiabatic method can be applied when only geopotential and temperature data are available. A disadvantage of the adiabatic method is that the local rate of change of temperature is required. Unless observations are taken at close intervals in time, it may be difficult to accurately estimate $\partial T/\partial t$ over a wide area. This method is also rather inaccurate in situations where strong diabatic heating is present, such as storms in which heavy rainfall occurs over a large area. Chapter 6 presents an alternative method for estimating ω, based on the so-called *omega equation*, that does not suffer from these difficulties.

3.6 SURFACE PRESSURE TENDENCY

The development of a negative surface *pressure tendency* is a classic warning of an approaching cyclonic weather disturbance. A simple expression that relates the surface pressure tendency to the wind field, and hence in theory might be used as the basis for short-range forecasts, can be obtained by taking the limit $p \to 0$ in (3.39) to get

$$\omega(p_s) = -\int_0^{p_s} (\nabla \bullet \mathbf{V})\, dp \tag{3.43}$$

followed by substituting from (3.37) to yield

$$\frac{\partial p_s}{\partial t} \approx -\int_0^{p_s} (\nabla \bullet \mathbf{V})\, dp \tag{3.44}$$

Here we have assumed that the surface is horizontal so that $w_s = 0$ and have neglected advection by the ageostrophic surface velocity in accord with the scaling arguments in Section 3.5.1.

According to (3.44), the surface pressure tendency at a given point is determined by the total convergence of mass into the vertical column of atmosphere above that point. This result is a direct consequence of the hydrostatic assumption, which implies that the pressure at a point is determined solely by the weight of the column of air above that point. Temperature changes in the air column will affect the heights of upper level pressure surfaces, but not the surface pressure.

Although, as stated earlier, the tendency equation might appear to have potential as a forecasting aid, its utility is severely limited due to the fact that, as discussed in Section 3.5.1, $\nabla \bullet \mathbf{V}$ is difficult to compute accurately from observations because it depends on the ageostrophic wind field. In addition, there is a strong tendency for vertical compensation. Thus, when there is convergence in the lower troposphere there is divergence aloft, and vice versa. The net integrated convergence or divergence is then a small residual in the vertical integral of a poorly determined quantity.

Nevertheless, (3.44) does have qualitative value as an aid in understanding the origin of surface pressure changes, and the relationship of such changes to the horizontal divergence. This can be illustrated by considering (as one possible example) the development of a thermal cyclone. We suppose that a heat source generates a local warm anomaly in the midtroposphere (Fig. 3.11a). Then according to the hypsometric equation (3.34), the heights of the upper level pressure surfaces are raised above the warm anomaly, resulting in a horizontal pressure gradient force

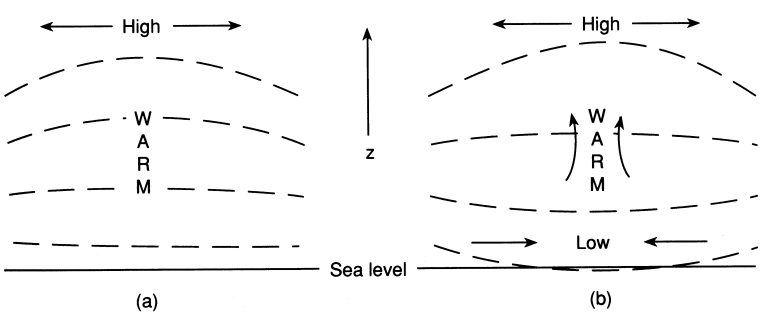


Fig. 3.11 Adjustment of surface pressure to a midtropospheric heat source. Dashed lines indicate isobars. (a) Initial height increase at upper level pressure surface. (b) Surface response to upper level divergence.

at the upper levels, which drives a divergent upper level wind. By (3.44) this upper level divergence will initially cause the surface pressure to decrease, thus generating a surface low below the warm anomaly (Fig. 3.11b). The horizontal pressure gradient associated with the surface low then drives a low-level convergence and vertical circulation, which tends to compensate the upper level divergence. The degree of compensation between upper divergence and lower convergence will determine whether the surface pressure continues to fall, remains steady, or rises.

The thermally driven circulation of the above example is by no means the only type of circulation possible (e.g., cold core cyclones are important synoptic-scale features). However, it does provide insight into how dynamical processes at upper levels are communicated to the surface and how the surface and upper troposphere are dynamically connected through the divergent circulation. This subject is considered in detail in Chapter 6.

Equation (3.44) is a lower boundary condition that determines the evolution of pressure at constant height. If the isobaric coordinate system of dynamical equations (3.2), (3.5), (3.6), and (3.27) is used as the set of governing equations, the lower boundary condition should be expressed in terms of the evolution of geopotential (or geopotential height) at constant pressure. Such an expression can be obtained simply by expanding $D\Phi/Dt$ in isobaric coordinates

$$\frac{\partial \Phi}{\partial t} = -\mathbf{V}_a \bullet \nabla \Phi - \omega \frac{\partial \Phi}{\partial p}$$

and substituting from (3.27) and (3.43) to get

$$\frac{\partial \Phi_s}{\partial t} \approx -\frac{RT_s}{p_s} \int_0^{p_s} (\nabla \bullet \mathbf{V})\, dp \tag{3.45}$$

where we have again neglected advection by the ageostrophic wind.

In practice the boundary condition (3.45) is difficult to use because it should be applied at pressure p_s, which is itself changing in time and space. In simple models it is usual to assume that p_s is constant (usually 1000 hPa) and to let $\omega = 0$ at p_s. For modern forecast models, an alternative coordinate system is generally employed in which the lower boundary is always a coordinate surface. This approach is described in Section 10.3.1.

PROBLEMS

3.1. An aircraft flying a heading of 60° (i.e., 60° to the east of north) at air speed 200 m s^{-1} moves relative to the ground due east (90°) at 225 m s^{-1}. If the plane is flying at constant pressure, what is its rate of change in altitude (in

meters per kilometer horizontal distance) assuming a steady pressure field, geostrophic winds, and $f = 10^{-4}\ \mathrm{s}^{-1}$?

3.2. The actual wind is directed 30° to the right of the geostrophic wind. If the geostrophic wind is 20 m s^{-1}, what is the rate of change of wind speed? Let $f = 10^{-4}\ \mathrm{s}^{-1}$.

3.3. A tornado rotates with constant angular velocity ω. Show that the surface pressure at the center of the tornado is given by

$$p = p_0 \exp\left(\frac{-\omega^2 r_0^2}{2RT}\right)$$

where p_0 is the surface pressure at a distance r_0 from the center and T is the temperature (assumed constant). If the temperature is 288 K and pressure and wind speed at 100 m from the center are 1000 hPa and 100 m s^{-1}, respectively, what is the central pressure?

3.4. Calculate the geostrophic wind speed (m s^{-1}) on an isobaric surface for a geopotential height gradient of 100 m per 1000 km and compare with all possible gradient wind speeds for the same geopotential height gradient and a radius of curvature of ± 500 km. Let $f = 10^{-4}\ \mathrm{s}^{-1}$.

3.5. Determine the maximum possible ratio of the normal anticyclonic gradient wind speed to the geostrophic wind speed for the same pressure gradient.

3.6. Show that the geostrophic balance in isothermal coordinates may be written

$$f\mathbf{V}_g = \mathbf{k} \times \boldsymbol{\nabla}_T\,(RT \ln p + \Phi)$$

3.7. Determine the radii of curvature for the trajectories of air parcels located 500 km to the east, north, south, and west of the center of a circular low-pressure system, respectively. The system is moving eastward at 15 m s^{-1}. Assume geostrophic flow with a uniform tangential wind speed of 15 m s^{-1}.

3.8. Determine the normal gradient wind speeds for the four air parcels of Problem 3.7 using the radii of curvature computed in Problem 3.7 Compare these speeds with the geostrophic speed. (Let $f = 10^{-4}\ \mathrm{s}^{-1}$.) Use the gradient wind speeds calculated here to recompute the radii of curvature for the four air parcels referred to in Problem 3.7. Use these new estimates of the radii of curvature to recompute the gradient wind speeds for the four air parcels. What fractional error is made in the radii of curvature by using the geostrophic wind approximation in this case? [Note that further iterations could be carried out but would converge rapidly.]

3.9. Show that as the pressure gradient approaches zero the gradient wind reduces to the geostrophic wind for a normal anticyclone and to inertial flow (Section 3.2.3) for an anomalous anticyclone.

3.10. The mean temperature in the layer between 750 and 500 hPa decreases eastward by 3° C per 100 km. If the 750-hPa geostrophic wind is from the southeast at 20 m s^{-1}, what is the geostrophic wind speed and direction at 500 hPa? Let $f = 10^{-4}$ s^{-1}.

3.11. What is the mean temperature advection in the 750- to 500-hPa layer in Problem 3.10?

3.12. Suppose that a vertical column of the atmosphere at 43°N is initially isothermal from 900 to 500 hPa. The geostrophic wind is 10 m s^{-1} from the south at 900 hPa, 10 m s^{-1} from the west at 700 hPa, and 20 m s^{-1} from the west at 500 hPa. Calculate the mean horizontal temperature gradients in the two layers 900-700 hPa and 700-500 hPa. Compute the rate of advective temperature change in each layer. How long would this advection pattern have to persist in order to establish a dry adiabatic lapse rate between 600 and 800 hPa? (Assume that the lapse rate is constant between 900 and 500 hPa and that the 800- to 600-hPa layer thickness is 2.25 km.)

3.13. An airplane pilot crossing the ocean at 45°N latitude has both a pressure altimeter and a radar altimeter, the latter measuring his absolute height above the sea. Flying at an air speed of 100 m s^{-1} he maintains altitude by referring to his pressure altimeter set for a sea level pressure of 1013 hPa. He holds an indicated 6000-m altitude. At the beginning of a 1-h period he notes that his radar altimeter reads 5700 m, and at the end of the hour he notes that it reads 5950 m. In what direction and approximately how far has he drifted from his heading?

3.14. Work out a gradient wind classification scheme equivalent to Table 3.1 for the Southern Hemisphere ($f < 0$) case.

3.15. In the geostrophic momentum approximation (Hoskins, 1975) the gradient wind formula for steady circular flow (3.17) is replaced by the approximation

$$V V_g R^{-1} + f V = f V_g$$

Compare the speeds V computed using this approximation with those obtained in Problem 3.8 using the gradient wind formula.

3.16. How large can the ratio $V_g/(fR)$ be before the geostrophic momentum approximation differs from the gradient wind approximation by 10% for cyclonic flow?

3.17. The planet Venus rotates about its axis so slowly that to a reasonable approximation the Coriolis parameter may be set equal to zero. For steady, frictionless motion parallel to latitude circles the momentum equation (2.20) then reduces to a type of cyclostrophic balance:

$$\frac{u^2 \tan\phi}{a} = -\frac{1}{\rho}\frac{\partial p}{\partial y}$$

By transforming this expression to isobaric coordinates, show that the thermal wind equation in this case can be expressed in the form

$$\omega_r^2\,(p_1) - \omega_r^2\,(p_0) = \frac{-R\ln(p_0/p_1)}{(a\sin\phi\cos\phi)}\frac{\partial\,\langle T\rangle}{\partial y}$$

where R is the gas constant, a is the radius of the planet, and $\omega_r \equiv \mathrm{u}/(a\cos\phi)$ is the relative angular velocity. How must $\langle T\rangle$(the vertically averaged temperature) vary with respect to latitude in order that ω_r be a function only of pressure? If the zonal velocity at about 60 km height above the equator ($p_1 = 2.9 \times 10^5$ Pa) is 100 m s^{-1} and the zonal velocity vanishes at the surface of the planet ($p_0 = 9.5 \times 10^6$ Pa), what is the vertically averaged temperature difference between the equator and the pole, assuming that ω_r depends only on pressure? The planetary radius is $a = 6100$ km, and the gas constant is $R = 187$ J kg^{-1} K^{-1}.

3.18. Suppose that during the passage of a cyclonic storm the radius of curvature of the isobars is observed to be 800 km at a station where the wind is veering (turning clockwise) at a rate of 10° per hour. What is the radius of curvature of the trajectory for an air parcel that is passing over the station? (The wind speed is 20 m s^{-1}.)

3.19. Show that the divergence of the geostrophic wind in isobaric coordinates on the spherical earth is given by

$$\nabla \bullet \mathbf{V}_g = -\frac{1}{fa}\frac{\partial\Phi}{\partial x}\left(\frac{\cos\phi}{\sin\phi}\right) = -v_g\left(\frac{\cot\phi}{a}\right)$$

(Use the spherical coordinate expression for the divergence operator given in Appendix C.)

3.20. The following wind data were received from 50 km to the east, north, west, and south of a station, respectively: 90°, 10 m s^{-1}; 120°, 4 m s^{-1}; 90°, 8 m s^{-1}; and 60°, 4 m s^{-1}. Calculate the approximate horizontal divergence at the station.

3.21. Suppose that the wind speeds given in Problem 3.20 are each in error by $\pm$ 10%. What would be the percentage error in the calculated horizontal divergence in the worst case?

3.22. The divergence of the horizontal wind at various pressure levels above a given station is shown in the following table. Compute the vertical velocity

Pressure (hPa)	$\nabla \bullet \mathbf{V} (\times 10^{-5}\ \mathrm{s}^{-1})$
1000	+0.9
850	+0.6
700	+0.3
500	0.0
300	−0.6
100	−1.0

at each level assuming an isothermal atmosphere with temperature 260 K and letting $w = 0$ at 1000 hPa.

3.23. Suppose that the lapse rate at the 850-hPa level is 4 K km^{-1}. If the temperature at a given location is decreasing at a rate of 2 K h^{-1}, the wind is westerly at 10 m s^{-1}, and the temperature decreases toward the west at a rate of 5 K/100 km, compute the vertical velocity at the 850-hPa level using the adiabatic method.

MATLAB EXERCISES

M3.1. For the situations considered in problems M2.1 and M2.2, make further modifications in the MATLAB scripts to compute the vertical profiles of density and of the static stability parameter S_p defined in (3.7). Plot these in the interval from $z = 0$ to $z = 15$ km. You will need to approximate the vertical derivative in S_p using a finite difference approximation (see Section 13.3.1).

M3.2. The objective of this exercise is to gain an appreciation for the difference between trajectories and streamlines in synoptic-scale flows. An idealized representation of a midlatitude synoptic disturbance in an atmosphere with no zonal mean flow is given by the simple sinusoidal pattern of geopotential,

$$\Phi(x, y, t) = \Phi_0 + \Phi' \sin[k(x - ct)] \cos\ ly$$

where $\Phi_0(p)$ is the standard atmosphere geopotential dependent only on pressure, Φ' is the magnitude of the geopotential wave disturbance, c is the phase speed of zonal propagation of the wave pattern, and k and l are the wave numbers in the x and y directions, respectively. If it is assumed that the flow is in geostrophic balance, then the geopotential is proportional to the stream function. Let the zonal and meridional wave numbers be equal ($k = l$) and define a perturbation wind amplitude $U' \equiv \Phi' k / f_0$, where f_0 is a constant Coriolis parameter. Trajectories are then given by the paths in (x, y) space obtained by solving the coupled set of ordinary differential equations:

$$\frac{Dx}{Dt} = u = -f_0^{-1}\frac{\partial \Phi}{\partial y} = +U' \sin[k(x - ct)] \sin ly$$
$$\frac{Dy}{Dt} = v = +f_0^{-1}\frac{\partial \Phi}{\partial x} = +U' \cos[k(x - ct)] \cos ly$$

[Note that U' (taken to be a positive constant) denotes the amplitude of both the x and y components of the disturbance wind.] The MATLAB script **trajectory_1.m** provides an accurate numerical solution to these equations for the special case in which the zonal mean wind vanishes. Three separate trajectories are plotted in the code. Run this script letting $U' = 10$ m s^{-1} for cases with $c = 5$, 10, and 15 m s^{-1}. Describe the behavior of the three trajectories for each of these cases. Why do the trajectories have their observed dependence on the phase speed c at which the geopotential height pattern propagates?

M3.3. The MATLAB script **trajectory_2.m** generalizes the case of problem M3.2 by adding a mean zonal flow. In this case the geopotential distribution is specified to be $\Phi(x, y, t) = \Phi_0 - f_0 \overline{U} l^{-1} \sin ly + \Phi' \sin[k(x - ct)] \cos ly$. (a) Solve for the latitudinal dependence of the mean zonal wind for this case. (b) Run the script with the initial x position specified as $x = -2250$ km and $U' = 15$ m s^{-1}. Do two runs, letting $\overline{U} = 10$ m s^{-1}, $c = 5$ m s^{-1} and $\overline{U} = 5$ m s^{-1}, $c = 10$ m s^{-1}, respectively. Determine the zonal distance that the ridge originally centered at $x = -2250$ km has propagated in each case. Use this information to briefly explain the characteristics (i.e., the shapes and lengths) of the 4-day trajectories for each of these cases. (c) What combination of initial position, $\overline{U}$ and c, will produce a straight-line trajectory?

M3.4. The MATLAB script **trajectory_3.m** can be used to examine the dispersion of a cluster of N parcels initially placed in a circle of small radius for a geopotential distribution representing the combination of a zonal mean jet plus a propagating wave with NE to SW tilt of trough and ridge lines.

The user must input the mean zonal wind amplitude, $\overline{U}$, the disturbance horizontal wind, U', the propagation speed of the waves, c, and the initial y position of the center of the cluster. (a) Run the model for three cases specifying $y = 0$, $U' = 15$ m s^{-1}, and $\overline{U} = 10, 12,$ and 15 m s^{-1}, respectively. Compute 20-day trajectories for all parcel clusters. Give an explanation for the differences in the dispersion of the parcel cluster for these three cases. (b) For the situation with $c = 10$ m s^{-1}, $U' = 15$ m s^{-1}, and $\overline{U} = 12$ m s^{-1}, run three additional cases with the initial y specified to be 250, 500, and 750 km. Describe how the results differ from the case with $y = 0$ km and $\overline{U} = 12$ m s^{-1} and give an explanation for the differences in these runs.

CHAPTER 4

Circulation and Vorticity

In classical mechanics the principle of conservation of angular momentum is often invoked in the analysis of motions that involve rotation. This principle provides a powerful constraint on the behavior of rotating objects. Analogous conservation laws also apply to the rotational field of a fluid. However, it should be obvious that in a continuous medium, such as the atmosphere, the definition of "rotation" is subtler than that for rotation of a solid object.

Circulation and vorticity are the two primary measures of rotation in a fluid. Circulation, which is a scalar integral quantity, is a *macroscopic* measure of rotation for a finite area of the fluid. Vorticity, however, is a vector field that gives a *microscopic* measure of the rotation at any point in the fluid.

4.1 THE CIRCULATION THEOREM

The *circulation*, C, about a closed contour in a fluid is defined as the line integral evaluated along the contour of the component of the velocity vector that is locally tangent to the contour:

$$C \equiv \oint \mathbf{U} \bullet d\mathbf{l} = \oint |\mathbf{U}| \cos \alpha \, dl$$

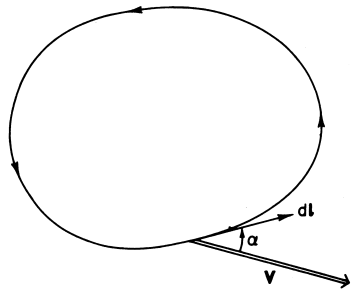


Fig. 4.1 Circulation about a closed contour.

where $\mathbf{l}(s)$ is a position vector extending from the origin to the point $s(x, y, z)$ on the contour C, and $d\mathbf{l}$ represents the limit of $\delta\mathbf{l} = \mathbf{l}(s+\delta s) - \mathbf{l}(s)$ as $\delta s \to 0$. Hence, as indicated in Fig. 4.1, $d\mathbf{l}$ is a displacement vector locally tangent to the contour. By convention the circulation is taken to be positive if $C > 0$ for counterclockwise integration around the contour.

That circulation is a measure of rotation is demonstrated readily by considering a circular ring of fluid of radius R in solid-body rotation at angular velocity $\mathbf{\Omega}$ about the z axis. In this case, $\mathbf{U} = \mathbf{\Omega} \times \mathbf{R}$, where $\mathbf{R}$ is the distance from the axis of rotation to the ring of fluid. Thus the circulation about the ring is given by

$$C \equiv \oint \mathbf{U} \bullet d\mathbf{l} = \int_0^{2\pi} \Omega R^2 d\lambda = 2\Omega\pi R^2$$

In this case the circulation is just 2π times the angular momentum of the fluid ring about the axis of rotation. Alternatively, note that $C/(\pi R^2) = 2\Omega$ so that the circulation divided by the area enclosed by the loop is just twice the angular speed of rotation of the ring. Unlike angular momentum or angular velocity, circulation can be computed without reference to an axis of rotation; it can thus be used to characterize fluid rotation in situations where "angular velocity" is not defined easily.

The circulation theorem is obtained by taking the line integral of Newton's second law for a closed chain of fluid particles. In the absolute coordinate system the result (neglecting viscous forces) is

$$\oint \frac{D_a \mathbf{U}_a}{Dt} \bullet d\mathbf{l} = -\oint \frac{\nabla_p \bullet d\mathbf{l}}{\rho} - \oint \nabla\Phi \bullet d\mathbf{l} \tag{4.1}$$

where the gravitational force $\mathbf{g}$ is represented as the gradient of the geopotential Φ, defined so that $-\nabla\Phi = \mathbf{g} = -g\mathbf{k}$. The integrand on the left-hand side can be

rewritten as[1]

$$\frac{D_a\mathbf{U}_a}{Dt} \bullet d\mathbf{l} = \frac{D}{Dt}(\mathbf{U}_a \bullet d\mathbf{l}) - \mathbf{U}_a \bullet \frac{D_a}{Dt}(d\mathbf{l})$$

or after observing that since $\mathbf{l}$ is a position vector, $D_a\mathbf{l}/Dt \equiv \mathbf{U}_a$,

$$\frac{D_a\mathbf{U}_a}{Dt} \bullet d\mathbf{l} = \frac{D}{Dt}(\mathbf{U}_a \bullet d\mathbf{l}) - \mathbf{U}_a \bullet d\mathbf{U}_a \tag{4.2}$$

Substituting (4.2) into (4.1) and using the fact that the line integral about a closed loop of a perfect differential is zero, so that

$$\oint \nabla\Phi \bullet d\mathbf{l} = \oint d\Phi = 0$$

and noting that

$$\oint \mathbf{U}_a \bullet d\mathbf{U}_a = \frac{1}{2}\oint d(\mathbf{U}_a \bullet \mathbf{U}_a) = 0$$

we obtain the circulation theorem:

$$\frac{DC_a}{Dt} = \frac{D}{Dt}\oint \mathbf{U}_a \bullet d\mathbf{l} = -\oint \rho^{-1}dp \tag{4.3}$$

The term on the right-hand side in (4.3) is called the solenoidal term. For a barotropic fluid, the density is a function only of pressure, and the solenoidal term is zero. Thus, in a barotropic fluid the absolute circulation is conserved following the motion. This result, called *Kelvin's circulation theorem*, is a fluid analog of angular momentum conservation in solid-body mechanics.

For meteorological analysis, it is more convenient to work with the relative circulation C rather than the absolute circulation, as a portion of the absolute circulation, C_e, is due to the rotation of the earth about its axis. To compute C_e, we apply Stokes' theorem to the vector $\mathbf{U}_e$, where $\mathbf{U}_e = \mathbf{\Omega} \times \mathbf{r}$ is the velocity of the earth at the position $\mathbf{r}$:

$$C_e = \oint \mathbf{U}_e \bullet d\mathbf{l} = \int_A\int (\nabla \times \mathbf{U}_e) \bullet \mathbf{n}dA$$

[1] Note that for a scalar $D_a/Dt = D/Dt$ (i.e., the rate of change following the motion does not depend on the reference system). For a vector, however, this is not the case, as was shown in Section 2.1.1.

where A is the area enclosed by the contour and the unit normal $\mathbf{n}$ is defined by the counterclockwise sense of the line integration using the "right-hand screw rule." Thus, for the contour of Fig. 4.1, $\mathbf{n}$ would be directed out of the page. If the line integral is computed in the horizontal plane, $\mathbf{n}$ is directed along the local vertical (see Fig. 4.2). Now, by a vector identity (see Appendix C)

$$\boldsymbol{\nabla} \times \mathbf{U}_e = \boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \mathbf{r}) = \boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \mathbf{R}) = \boldsymbol{\Omega}\boldsymbol{\nabla} \bullet \mathbf{R} = 2\boldsymbol{\Omega}$$

so that $(\boldsymbol{\nabla} \times \mathbf{U}_e) \bullet \mathbf{n} = 2\Omega \sin\phi \equiv f$ is just the Coriolis parameter. Hence, the circulation in the horizontal plane due to the rotation of the earth is

$$C_e = 2\Omega\langle\sin\phi\rangle\ A = 2\Omega A_e$$

where $\langle\sin\phi\rangle$ denotes an average over the area element A and A_e is the projection of A in the equatorial plane as illustrated in Fig. 4.2. Thus, the relative circulation may be expressed as

$$C = C_a - C_e = C_a - 2\Omega A_e \tag{4.4}$$

Differentiating (4.4) following the motion and substituting from (4.3) we obtain the Bjerknes circulation theorem:

$$\frac{DC}{Dt} = -\oint \frac{dp}{\rho} - 2\Omega\frac{DA_e}{Dt} \tag{4.5}$$

For a barotropic fluid, (4.5) can be integrated following the motion from an initial state (designated by subscript 1) to a final state (designated by subscript 2), yielding the circulation change

$$C_2 - C_1 = -2\Omega\ (A_2 \sin\phi_2 - A_1 \sin\phi_1) \tag{4.6}$$

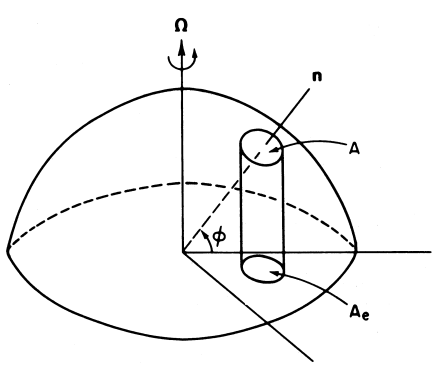


Fig. 4.2 Area A_e subtended on the equatorial plane by horizontal area A centered at latitude ϕ.

Equation (4.6) indicates that in a barotropic fluid the relative circulation for a closed chain of fluid particles will be changed if either the horizontal area enclosed by the loop changes or the latitude changes. Furthermore, a negative absolute circulation in the Northern Hemisphere can develop only if a closed chain of fluid particles is advected across the equator from the Southern Hemisphere. The anomalous gradient wind balances discussed in Section 3.2.5 are examples of systems with negative absolute circulations (see Problem 4.6).

Example. Suppose that the air within a circular region of radius 100 km centered at the equator is initially motionless with respect to the earth. If this circular air mass were moved to the North Pole along an isobaric surface preserving its area, the circulation about the circumference would be

$$C = -2\Omega\pi r^2[\sin(\pi/2) - \sin(0)]$$

Thus the mean tangential velocity at the radius $r = 100$ km would be

$$V = C/(2\pi r) = -\Omega r \approx -7 \text{ m s}^{-1}$$

The negative sign here indicates that the air has acquired anticyclonic relative circulation.

In a baroclinic fluid, circulation may be generated by the pressure-density solenoid term in (4.3). This process can be illustrated effectively by considering the development of a sea breeze circulation, as shown in Fig. 4.3. For the situation depicted, the mean temperature in the air over the ocean is colder than the mean temperature over the adjoining land. Thus, if the pressure is uniform at ground level, the isobaric surfaces above the ground will slope downward toward the

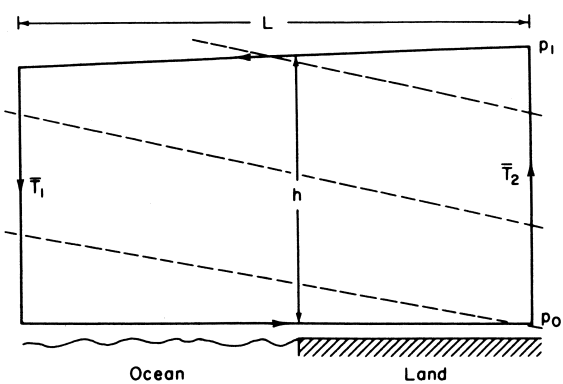


Fig. 4.3 Application of the circulation theorem to the sea breeze problem. The closed heavy solid line is the loop about which the circulation is to be evaluated. Dashed lines indicate surfaces of constant density.

ocean while the isopycnic surfaces (surfaces of constant density) will slope downward toward the land. To compute the acceleration as a result of the intersection of the pressure-density surfaces, we apply the circulation theorem by integrating around a circuit in a vertical plane perpendicular to the coastline. Substituting the ideal gas law into (4.3) we obtain

$$\frac{DC_a}{Dt} = -\oint RT d \ln p$$

For the circuit shown in Fig. 4.3 there is a contribution to the line integral only for the vertical segments of the loop, as the horizontal segments are taken at constant pressure. The resulting rate of increase in the circulation is

$$\frac{DC_a}{Dt} = R \ln\left(\frac{p_0}{p_1}\right)(\overline{T}_2 - \overline{T}_1) > 0$$

Letting $\langle v \rangle$ be the mean tangential velocity along the circuit, we find that

$$\frac{D\langle v \rangle}{Dt} = \frac{R \ln(p_0/p_1)}{2(h+L)}(\overline{T}_2 - \overline{T}_1) \tag{4.7}$$

If we let $p_0 = 1000$ hPa, $p_1 = 900$ hPa, $\overline{T}_2 - \overline{T}_1 = 10°$ C, $L = 20$ km, and $h = 1$ km, (4.7) yields an acceleration of about 7×10^{-3} m s^{-2}. In the absence of frictional retarding forces, this would produce a wind speed of 25 m s^{-1} in about 1 h. In reality, as the wind speed increases, the frictional force reduces the acceleration rate, and temperature advection reduces the land–sea temperature contrast so that a balance is obtained between the generation of kinetic energy by the pressure-density solenoids and frictional dissipation.

4.2 VORTICITY

Vorticity, the microscopic measure of rotation in a fluid, is a vector field defined as the curl of velocity. The absolute vorticity $\boldsymbol{\omega}_a$ is the curl of the absolute velocity, whereas the relative vorticity $\boldsymbol{\omega}$ is the curl of the relative velocity:

$$\boldsymbol{\omega}_a \equiv \boldsymbol{\nabla} \times \mathbf{U}_a, \qquad \boldsymbol{\omega} \equiv \boldsymbol{\nabla} \times \mathbf{U}$$

so that in Cartesian coordinates,

$$\boldsymbol{\omega} = \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$

In large-scale dynamic meteorology, we are in general concerned only with the vertical components of absolute and relative vorticity, which are designated by η and ζ, respectively.

$$\eta \equiv \mathbf{k} \bullet (\nabla \times \mathbf{U}_a), \qquad \zeta \equiv \mathbf{k} \bullet (\nabla \times \mathbf{U})$$

In the remainder of this book, η and ζ are referred to as absolute and relative vorticities, respectively, without adding the explicit modifier "vertical component of." Regions of positive ζ are associated with cyclonic storms in the Northern Hemisphere; regions of negative ζ are associated with cyclonic storms in the Southern Hemisphere. Thus, the distribution of relative vorticity is an excellent diagnostic for weather analysis. Absolute vorticity tends to be conserved following the motion at midtropospheric levels; this conservation property is the basis for the simplest dynamical forecast scheme discussed in Chapter 13.

The difference between absolute and relative vorticity is *planetary vorticity*, which is just the local vertical component of the vorticity of the earth due to its rotation; $\mathbf{k} \bullet \nabla \times \mathbf{U}_e = 2\Omega \sin\phi \equiv f$. Thus, $\eta = \zeta + f$ or, in Cartesian coordinates,

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \qquad \eta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + f$$

The relationship between relative vorticity and relative circulation C discussed in the previous section can be clearly seen by considering an alternative approach in which the vertical component of vorticity is defined as the circulation about a closed contour in the horizontal plane divided by the area enclosed, in the limit where the area approaches zero:

$$\zeta \equiv \lim_{A \to 0} \left(\oint \mathbf{V} \bullet d\mathbf{l} \right) A^{-1} \tag{4.8}$$

This latter definition makes explicit the relationship between circulation and vorticity discussed in the introduction to this chapter. The equivalence of these two definitions of ζ is shown easily by considering the circulation about a rectangular element of area $\delta x \delta y$ in the (x, y) plane as shown in Fig. 4.4. Evaluating $\mathbf{V} \bullet d\mathbf{l}$ for each side of the rectangle in Fig. 4.4 yields the circulation

$$\begin{aligned} \delta C &= u\delta x + \left(v + \frac{\partial v}{\partial x}\delta x \right) \delta y - \left(u + \frac{\partial u}{\partial y}\delta y \right) \delta x - v\delta y \\ &= \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \delta x \delta y \end{aligned}$$

Dividing through by the area $\delta A = \delta x \delta y$ gives

$$\frac{\delta C}{\delta A} = \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \equiv \zeta$$

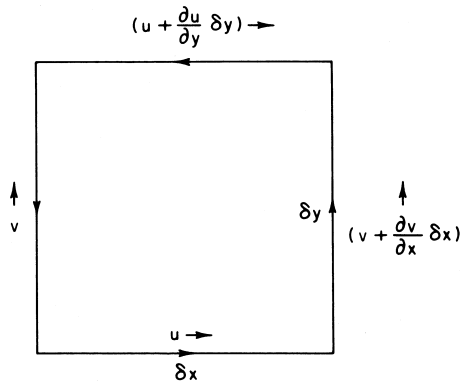


Fig. 4.4 Relationship between circulation and vorticity for an area element in the horizontal plane.

In more general terms the relationship between vorticity and circulation is given simply by Stokes' theorem applied to the velocity vector:

$$\oint \mathbf{U} \bullet d\mathbf{l} = \iint_A (\boldsymbol{\nabla} \times \mathbf{U}) \bullet \mathbf{n} dA$$

Here A is the area enclosed by the contour and $\mathbf{n}$ is a unit normal to the area element dA (positive in the right-hand sense). Thus, Stokes' theorem states that the circulation about any closed loop is equal to the integral of the normal component of vorticity over the area enclosed by the contour. Hence, for a finite area, circulation divided by area gives the *average* normal component of vorticity in the region. As a consequence, the vorticity of a fluid in solid-body rotation is just twice the angular velocity of rotation. Vorticity may thus be regarded as a measure of the local angular velocity of the fluid.

4.2.1 Vorticity in Natural Coordinates

Physical interpretation of vorticity is facilitated by considering the vertical component of vorticity in the natural coordinate system (see Section 3.2.1). If we compute the circulation about the infinitesimal contour shown in Fig. 4.5, we obtain[2]

$$\delta C = V\left[\delta s + d(\delta s)\right] - \left(V + \frac{\partial V}{\partial n}\delta n\right)\delta s$$

[2] Recall that n is a coordinate in the horizontal plane perpendicular to the local flow direction with positive values to the left of an observer facing downstream.

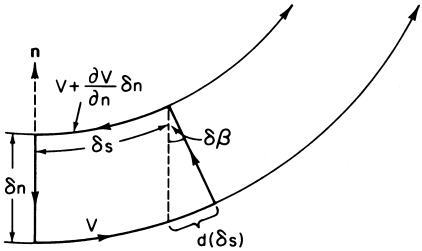


Fig. 4.5 Circulation for an infinitesimal loop in the natural coordinate system.

However, from Fig. 4.5, $d(\delta s) = \delta\beta\delta n$, where $\delta\beta$ is the angular change in the wind direction in the distance δs. Hence,

$$\delta C = \left(-\frac{\partial V}{\partial n} + V\frac{\delta\beta}{\delta s}\right)\delta n\ \delta s$$

or, in the limit $\delta n, \delta s \to 0$

$$\zeta = \lim_{\delta n, \delta s \to 0} \frac{\delta C}{(\delta n\ \delta s)} = -\frac{\partial V}{\partial n} + \frac{V}{R_s} \tag{4.9}$$

where R_s is the radius of curvature of the streamlines [Eq. (3.20)]. It is now apparent that the net vertical vorticity component is the result of the sum of two parts: (1) the rate of change of wind speed normal to the direction of flow $-\partial V/\partial n$, called the *shear* vorticity; and (2) the turning of the wind along a streamline V/R_s, called the *curvature* vorticity. Thus, even straight-line motion may have vorticity if the speed changes normal to the flow axis. For example, in the jet stream shown schematically in Fig. 4.6a, there will be cyclonic relative vorticity north of the velocity maximum and anticyclonic relative vorticity to the south (Northern Hemisphere conditions) as is recognized easily when the turning of a small paddle wheel placed in the flow is considered. The lower of the two paddle wheels in Fig. 4.6a will turn in a clockwise direction (anticyclonically) because the wind force on the blades north of its axis of rotation is stronger than the force on the blades to the south of the axis. The upper wheel will, of course, experience a counterclockwise (cyclonic) turning. Thus, the poleward and equatorward sides of a westerly jetstream are referred to as the cyclonic and anticyclonic shear sides, respectively.

Conversely, curved flow may have zero vorticity provided that the shear vorticity is equal and opposite to the curvature vorticity. This is the case in the example shown in Fig. 4.6b where a frictionless fluid with zero relative vorticity upstream flows around a bend in a canal. The fluid along the inner boundary on the curve flows faster in just the right proportion so that the paddle wheel does not turn.

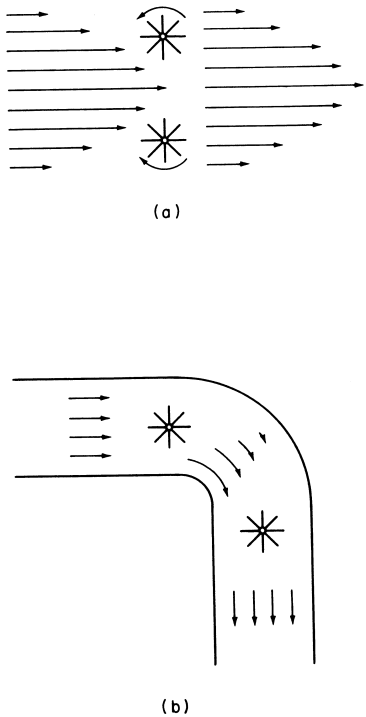


Fig. 4.6 Two types of two-dimensional flow: (a) linear shear flow with vorticity and (b) curved flow with zero vorticity.

4.3 POTENTIAL VORTICITY

With the aid of the ideal gas law (1.17), the definition of potential temperature (2.44) can be expressed as a relationship between pressure and density for a surface of constant θ:

$$\rho = p^{c_v/c_p}(R\theta)^{-1}(p_s)^{R/c_p}$$

Hence, on an isentropic surface, density is a function of pressure alone, and the solenoidal term in the circulation theorem (4.3) vanishes;

$$\oint \frac{dp}{\rho} \propto \oint dp^{(1-c_v/c_p)} = 0$$

Thus, for adiabatic flow the circulation computed for a closed chain of fluid parcels on a constant θ surface reduces to the same form as in a barotropic fluid; that is, it satisfies Kelvin's circulation theorem, which may be expressed as

$$\frac{D}{Dt}(C + 2\Omega\, \delta A \sin\phi) = 0 \tag{4.10}$$

where C is evaluated for a closed loop encompassing the area δA on an isentropic surface. If the isentropic surface is approximately horizontal, and it is recalled from (4.8) that $C \approx \zeta \delta A$, then for an infinitesimal parcel of air (4.10) implies that

$$\delta A(\zeta_\theta + f) = \text{Const} \tag{4.11}$$

where ζ_θ designates the vertical component of relative vorticity evaluated on an isentropic surface and $f = 2\Omega \sin\phi$ is the Coriolis parameter.

Suppose that the parcel of (4.11) is confined between potential temperature surfaces θ_0 and $\theta_0 + \delta\theta$, which are separated by a pressure interval $-\delta p$ as shown in Fig. 4.7. The mass of the parcel, $\delta M = (-\delta p/g)\delta A$, must be conserved following the motion. Therefore,

$$\delta A = -\frac{\delta M g}{\delta p} = \left(-\frac{\delta\theta}{\delta p}\right)\left(\frac{\delta M g}{\delta\theta}\right) = \text{Const} \times g\left(-\frac{\delta\theta}{\delta p}\right)$$

as both δM and $\delta\theta$ are constants. Substituting into (4.11) to eliminate δA and taking the limit $\delta p \to 0$, we obtain

$$P \equiv (\zeta_\theta + f)\left(-g\frac{\partial\theta}{\partial p}\right) = \text{Const} \tag{4.12}$$

The quantity P [units: K kg^{-1} m^2 s^{-1}] is the isentropic coordinate form of *Ertel's potential vorticity*.[3] It is defined with a minus sign so that its value is normally positive in the Northern Hemisphere.

According to (4.12), potential vorticity is conserved following the motion in adiabatic frictionless flow. The term *potential vorticity* is used, as shown later, in connection with several other mathematical expressions. In essence, however,

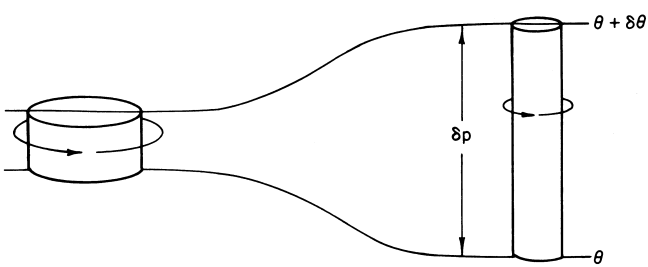


Fig. 4.7 A cylindrical column of air moving adiabatically, conserving potential vorticity.

[3] Named for the German meteorologist Hans Ertel. A more general form of Ertel's potential vorticity is discussed, for example, in Gill (1982). Potential vorticity is often expressed in the potential vorticity unit (PVU), where 1 PVU = 10^{-6} K kg^{-1} m^2 s^{-1}.

potential vorticity is always in some sense a measure of the ratio of the absolute vorticity to the effective depth of the vortex. In (4.12), for example, the effective depth is just the differential distance between potential temperature surfaces measured in pressure units $(-\partial\theta/\partial p)$.

In a homogeneous incompressible fluid, potential vorticity conservation takes a somewhat simpler form. In this case, because density is a constant, the horizontal area must be inversely proportional to the depth, h, of the fluid parcel:

$$\delta A = M(\rho h)^{-1} = \text{Const}/h$$

where h is the depth of the parcel. Substituting to eliminate δA in (4.11) yields

$$(\zeta + f)/h = \eta/h = \text{Const} \tag{4.13}$$

where ζ is here evaluated at constant height.

If the depth, h, is constant, (4.13) states that absolute vorticity is conserved following the motion. Conservation of absolute vorticity following the motion provides a strong constraint on the flow, as can be shown by a simple example. Suppose that at a certain point (x_0, y_0) the flow is in the zonal direction and the relative vorticity vanishes so that $\eta(x_0, y_0) = f_0$. Then, if absolute vorticity is conserved, the motion at any point along a parcel trajectory that passes through (x_0, y_0) must satisfy $\zeta + f = f_0$. Because f increases toward the north, trajectories that curve northward in the downstream direction must have $\zeta = f_0 - f < 0$, whereas trajectories that curve southward must have $\zeta = f_0 - f > 0$. However, as indicated in Fig. 4.8, if the flow is westerly, northward curvature downstream implies $\zeta > 0$, whereas southward curvature implies $\zeta < 0$. Thus, westerly zonal flow must remain purely zonal if absolute vorticity is to be conserved following the motion. The easterly flow case, also shown in Fig. 4.8, is just the opposite. Northward and southward curvatures are associated with negative and positive relative vorticities, respectively. Hence, an easterly current can curve either to the north or to the south and still conserve absolute vorticity.

When the depth of the fluid changes following the motion, it is potential vorticity that is conserved. However, again (4.13) indicates that westerly and easterly flows behave differently. The situation for westerly flow impinging on an infinitely long

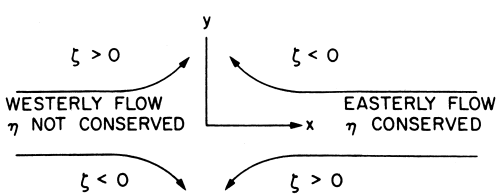


Fig. 4.8 Absolute vorticity conservation for curved flow trajectories.

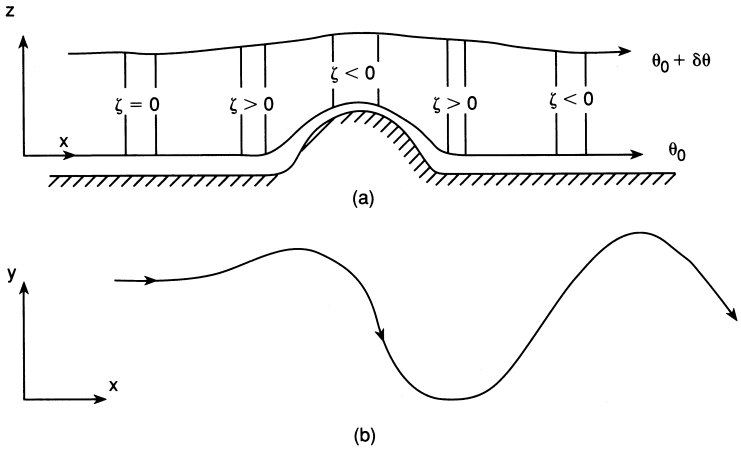


Fig. 4.9 Schematic view of westerly flow over a topographic barrier: (a) the depth of a fluid column as a function of x and (b) the trajectory of a parcel in the (x, y) plane.

topographic barrier is shown in Fig. 4.9. In Fig. 4.9a, a vertical cross section of the flow is shown. We suppose that upstream of the mountain barrier the flow is a uniform zonal flow so that $\zeta = 0$. If the flow is adiabatic, each column of air of depth h confined between the potential temperature surfaces θ_0 and $\theta_0 + \delta\theta$ remains between those surfaces as it crosses the barrier. For this reason, a potential temperature surface θ_0 near the ground must approximately follow the ground contours. A potential temperature surface $\theta_0 + \delta\theta$ several kilometers above the ground will also be deflected vertically. However, due to pressure forces produced by interaction of the flow with the topographic barrier, the vertical displacement at upper levels is spread horizontally; it extends upstream and downstream of the barrier and has smaller amplitude in the vertical than the displacement near the ground (see Figs. 4.9 and 4.10).

As a result of the vertical displacement of the upper level isentropes, there is a vertical stretching of air columns upstream of the topographic barrier. (For motions of large horizontal scale, the upstream stretching is quite small.) This stretching causes h to increase, and hence from (4.13) ζ must become positive in order to conserve potential vorticity. Thus, an air column turns cyclonically as it approaches the mountain barrier. This cyclonic curvature causes a poleward drift so that f also increases, which reduces the change in ζ required for potential vorticity conservation. As the column begins to cross the barrier, its vertical extent decreases; the relative vorticity must then become negative. Thus, the air column will acquire anticyclonic vorticity and move southward as shown in the x, y plane profile in Fig. 4.9b. When the air column has passed over the barrier and returned to its original depth, it will be south of its original latitude so that f will be smaller and the relative vorticity must be positive. Thus, the trajectory must have cyclonic

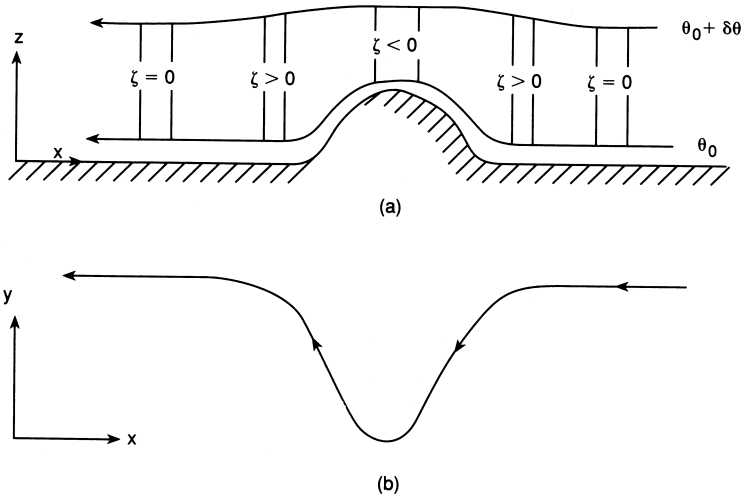


Fig. 4.10 As in Fig. 4.9, but for easterly flow.

curvature and the column will be deflected poleward. When the parcel returns to its original latitude, it will still have a poleward velocity component and will continue poleward gradually, acquiring anticyclonic curvature until its direction is again reversed. The parcel will then move downstream, conserving potential vorticity by following a wave-like trajectory in the horizontal plane. Therefore, steady westerly flow over a large-scale ridge will result in a cyclonic flow pattern immediately to the east of the barrier (the lee side trough) followed by an alternating series of ridges and troughs downstream.

The situation for easterly flow impinging on a mountain barrier is quite different. As indicated schematically in Fig. 4.10b, upstream stretching leads to a cyclonic turning of the flow, as in the westerly case. For easterly flow this cyclonic turning creates an equatorward component of motion. As the column moves westward and equatorward over the barrier, its depth contracts and its absolute vorticity must then decrease so that potential vorticity can be conserved. This reduction in absolute vorticity arises both from development of anticyclonic relative vorticity and from a decrease in f due to the equatorward motion. The anticyclonic relative vorticity gradually turns the column so that when it reaches the top of the barrier it is headed westward. As it continues westward down the barrier, conserving potential vorticity, the process is simply reversed with the result that some distance downstream from the mountain barrier the air column again is moving westward at its original latitude. Thus, the dependence of the Coriolis parameter on latitude creates a dramatic difference between westerly and easterly flow over large-scale topographic barriers. In the case of a westerly wind, the barrier generates a wave-like disturbance in the streamlines that extends far downstream. However, in the

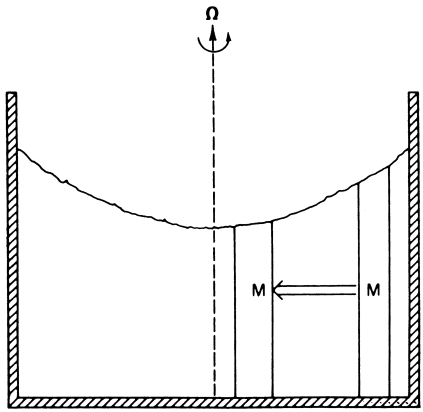


Fig. 4.11 Dependence of depth on radius in a rotating cylindrical vessel.

case of an easterly wind, the disturbance in the streamlines damps out away from the barrier.

The situation described above, which assumes that the mountain barrier is infinitely long in the meridional direction, is highly idealized. In reality, because static stability tends to suppress vertical motions, large-scale flows in statically stable environments are blocked by topographic barriers and are forced to flow around mountains rather than over them. However, whether fluid columns go over or around topographic barriers, the potential vorticity conservation constraint still must be satisfied.

The Rossby potential vorticity conservation law, (4.13), indicates that in a barotropic fluid, a change in the depth is dynamically analogous to a change in the Coriolis parameter. This can be demonstrated easily in a rotating cylindrical vessel filled with water. For solid-body rotation the equilibrium shape of the free surface, determined by a balance between the radial pressure gradient and centrifugal forces, is parabolic. Thus, as shown in Fig. 4.11, if a column of fluid moves radially outward, it must stretch vertically. According to (4.13), the relative vorticity must then increase to keep the ratio $(\zeta + f)/h$ constant. The same result would apply if a column of fluid on a rotating sphere were moved equatorward without a change in depth. In this case, ζ would have to increase to offset the decrease of f. Therefore, in a barotropic fluid, a decrease of depth with increasing latitude has the same effect on the relative vorticity as the increase of the Coriolis force with latitude.

4.4 THE VORTICITY EQUATION

The previous section discussed the time evolution of the vertical component of vorticity for the special case of adiabatic frictionless flow. This section uses the equations of motion to derive an equation for the time rate of change of vorticity without limiting the validity to adiabatic motion.

4.4.1 Cartesian Coordinate Form

For motions of synoptic scale, the vorticity equation can be derived using the approximate horizontal momentum equations (2.24) and (2.25). We differentiate the zonal component equation with respect to y and the meridional component equation with respect to x:

$$\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}-fv=-\frac{1}{\rho}\frac{\partial p}{\partial x}\right) \tag{4.14}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}+fu=-\frac{1}{\rho}\frac{\partial p}{\partial y}\right) \tag{4.15}$$

Subtracting (4.14) from (4.15) and recalling that $\zeta = \partial v/\partial x - \partial u/\partial y$, we obtain the vorticity equation

$$\begin{aligned}\frac{\partial \zeta}{\partial t}+u\frac{\partial \zeta}{\partial x}+v\frac{\partial \zeta}{\partial y}+w\frac{\partial \zeta}{\partial z}+(\zeta+f)\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)=&\\ +\left(\frac{\partial w}{\partial x}\frac{\partial v}{\partial z}-\frac{\partial w}{\partial y}\frac{\partial u}{\partial z}\right)+v\frac{df}{dy}=\frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial x}\frac{\partial p}{\partial y}-\frac{\partial \rho}{\partial y}\frac{\partial p}{\partial x}\right)&\end{aligned} \tag{4.16}$$

Using the fact that the Coriolis parameter depends only on y so that $Df/Dt = v(df/dy)$, (4.16) may be rewritten in the form

$$\begin{aligned}\frac{D}{Dt}(\zeta+f)=&-(\zeta+f)\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)\\ &-\left(\frac{\partial w}{\partial x}\frac{\partial v}{\partial z}-\frac{\partial w}{\partial y}\frac{\partial u}{\partial z}\right)+\frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial x}\frac{\partial p}{\partial y}-\frac{\partial \rho}{\partial y}\frac{\partial p}{\partial x}\right)\end{aligned} \tag{4.17}$$

Equation (4.17) states that the rate of change of the absolute vorticity following the motion is given by the sum of the three terms on the right, called the divergence term, the tilting or twisting term, and the solenoidal term, respectively.

The concentration or dilution of vorticity by the divergence field [the first term on the right in (4.17)] is the fluid analog of the change in angular velocity resulting from a change in the moment of inertia of a solid body when angular momentum is conserved. If the horizontal flow is divergent, the area enclosed by a chain of fluid parcels will increase with time and if circulation is to be conserved, the average absolute vorticity of the enclosed fluid must decrease (i.e., the vorticity will be diluted). If, however, the flow is convergent, the area enclosed by a chain of fluid parcels will decrease with time and the vorticity will be concentrated. This mechanism for changing vorticity following the motion is very important in synoptic-scale disturbances.

The second term on the right in (4.17) represents vertical vorticity generated by the tilting of horizontally oriented components of vorticity into the vertical by a

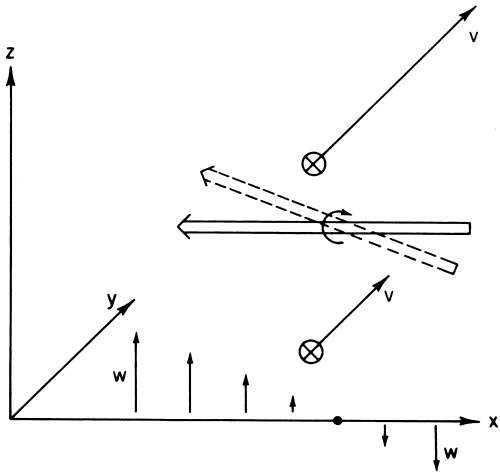


Fig. 4.12 Vorticity generation by the tilting of a horizontal vorticity vector (double arrow).

nonuniform vertical motion field. This mechanism is illustrated in Fig. 4.12, which shows a region where the y component of velocity is increasing with height so that there is a component of shear vorticity oriented in the negative x direction as indicated by the double arrow. If at the same time there is a vertical motion field in which w decreases with increasing x, advection by the vertical motion will tend to tilt the vorticity vector initially oriented parallel to x so that it has a component in the vertical. Thus, if $\partial v/\partial z > 0$ and $\partial w/\partial x < 0$, there will be a generation of positive vertical vorticity.

Finally, the third term on the right in (4.17) is just the microscopic equivalent of the solenoidal term in the circulation theorem (4.5). To show this equivalence, we may apply Stokes' theorem to the solenoidal term to get

$$-\oint \alpha dp \equiv -\oint \alpha \nabla p \bullet d\mathbf{l} = -\iint_A \nabla \times (\alpha \nabla p) \bullet \mathbf{k} dA$$

where A is the horizontal area bounded by the curve $\mathbf{l}$. Applying the vector identity $\nabla \times (\alpha \nabla p) \equiv \nabla \alpha \times \nabla p$, the equation becomes

$$-\oint \alpha dp = -\iint_A (\nabla \alpha \times \nabla p) \bullet \mathbf{k} dA$$

However, the solenoidal term in the vorticity equation can be written

$$-\left(\frac{\partial \alpha}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial \alpha}{\partial y}\frac{\partial p}{\partial x}\right) = -(\nabla \alpha \times \nabla p) \bullet \mathbf{k}$$

Thus, the solenoidal term in the vorticity equation is just the limit of the solenoidal term in the circulation theorem divided by the area when the area goes to zero.

4.4.2 The Vorticity Equation in Isobaric Coordinates

A somewhat simpler form of the vorticity equation arises when the motion is referred to the isobaric coordinate system. This equation can be derived in vector form by operating on the momentum equation (3.2) with the vector operator $\mathbf{k} \bullet \nabla \times$, where ∇ now indicates the horizontal gradient on a surface of constant pressure. However, to facilitate this process it is desirable to first use the vector identity

$$(\mathbf{V} \bullet \nabla)\mathbf{V} = \nabla \left(\frac{\mathbf{V} \bullet \mathbf{V}}{2} \right) + \zeta \mathbf{k} \times \mathbf{V} \tag{4.18}$$

where $\zeta = \mathbf{k} \bullet (\nabla \times \mathbf{V})$, to rewrite (3.2) as

$$\frac{\partial \mathbf{V}}{\partial t} = -\nabla \left(\frac{\mathbf{V} \bullet \mathbf{V}}{2} + \Phi \right) - (\zeta + f)\mathbf{k} \times \mathbf{V} - \omega \frac{\partial \mathbf{V}}{\partial p} \tag{4.19}$$

We now apply the operator $\mathbf{k} \bullet \nabla \times$ to (4.19). Using the facts that for any scalar A, $\nabla \times \nabla A = 0$ and for any vectors $\mathbf{a}$, $\mathbf{b}$,

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = (\nabla \bullet \mathbf{b})\, \mathbf{a} - (\mathbf{a} \bullet \nabla)\, \mathbf{b} - (\nabla \bullet \mathbf{a})\, \mathbf{b} + (\mathbf{b} \bullet \nabla)\, \mathbf{a} \tag{4.20}$$

we can eliminate the first term on the right and simplify the second term so that the resulting vorticity equation becomes

$$\frac{\partial \zeta}{\partial t} = -\mathbf{V} \bullet \nabla\, (\zeta + f) - \omega \frac{\partial \zeta}{\partial p} - (\zeta + f)\, \nabla \bullet \mathbf{V} + \mathbf{k} \bullet \left(\frac{\partial \mathbf{V}}{\partial p} \times \nabla \omega \right) \tag{4.21}$$

Comparing (4.17) and (4.21), we see that in the isobaric system there is no vorticity generation by pressure-density solenoids. This difference arises because in the isobaric system, horizontal partial derivatives are computed with p held constant so that the vertical component of vorticity is $\zeta = (\partial v/\partial x - \partial u/\partial y)_p$ whereas in height coordinates it is $\zeta = (\partial v/\partial x - \partial u/\partial y)_z$. In practice the difference is generally unimportant because as shown in the next section, the solenoidal term is usually sufficiently small so that it can be neglected for synoptic-scale motions.

4.4.3 Scale Analysis of the Vorticity Equation

In Section 2.4 the equations of motion were simplified for synoptic-scale motions by evaluating the order of magnitude of various terms. The same technique can also be applied to the vorticity equation. Characteristic scales for the field variables

based on typical observed magnitudes for synoptic-scale motions are chosen as follows:

$U \sim 10\,\mathrm{m\,s^{-1}}$	horizontal scale
$W \sim 1\,\mathrm{cm\,s^{-1}}$	vertical scale
$L \sim 10^6\,\mathrm{m}$	length scale
$H \sim 10^4\,\mathrm{m}$	depth scale
$\delta p \sim 10\,\mathrm{hPa}$	horizontal pressure scale
$\rho \sim 1\,\mathrm{kg\,m^{-3}}$	mean density
$\delta\rho/\rho \sim 10^{-2}$	fractional density fluctuation
$L/U \sim 10^5\,\mathrm{s}$	time scale
$f_0 \sim 10^{-4}\,\mathrm{s^{-1}}$	Coriolis parameter
$\beta \sim 10^{-11}\,\mathrm{m^{-1}\,s^{-1}}$	"beta" parameter

Again we have chosen an advective time scale because the vorticity pattern, like the pressure pattern, tends to move at a speed comparable to the horizontal wind speed. Using these scales to evaluate the magnitude of the terms in (4.16), we first note that

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \lesssim \frac{U}{L} \sim 10^{-5}\mathrm{s}^{-1}$$

where the inequality in this expression means less than or equal to in order of magnitude. Thus,

$$\zeta/f_0 \lesssim U/(f_0\,L) \equiv \mathrm{Ro} \sim 10^{-1}$$

For midlatitude synoptic-scale systems, the relative vorticity is often small (order Rossby number) compared to the planetary vorticity. For such systems, ζ may be neglected compared to f in the divergence term in the vorticity equation

$$(\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \approx f\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)$$

This approximation does not apply near the center of intense cyclonic storms. In such systems $|\zeta/f| \sim 1$, and the relative vorticity should be retained.

The magnitudes of the various terms in (4.16) can now be estimated as

$$\frac{\partial \zeta}{\partial t}, u\frac{\partial \zeta}{\partial x}, v\frac{\partial \zeta}{\partial y} \sim \frac{U^2}{L^2} \sim 10^{-10}\ \mathrm{s}^{-2}$$

$$w\frac{\partial \zeta}{\partial z} \sim \frac{WU}{HL} \sim 10^{-11}\ \mathrm{s}^{-2}$$

$$v\frac{df}{dy} \sim U\beta \sim 10^{-10}\ \mathrm{s}^{-2}$$
$$f\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \lesssim \frac{f_0 U}{L} \sim 10^{-9}\ \mathrm{s}^{-2}$$
$$\left(\frac{\partial w}{\partial x}\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}\frac{\partial u}{\partial z}\right) \lesssim \frac{WU}{HL} \sim 10^{-11}\ \mathrm{s}^{-2}$$
$$\frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y}\frac{\partial p}{\partial x}\right) \lesssim \frac{\delta\rho\delta p}{\rho^2 L^2} \sim 10^{-11}\ \mathrm{s}^{-2}$$

The inequality is used in the last three terms because in each case it is possible that the two parts of the expression might partially cancel so that the actual magnitude would be less than indicated. In fact, this must be the case for the divergence term (the fourth in the list) because if $\partial u/\partial x$ and $\partial v/\partial y$ were not nearly equal and opposite, the divergence term would be an order of magnitude greater than any other term and the equation could not be satisfied. Therefore, scale analysis of the vorticity equation indicates that synoptic-scale motions must be quasi-nondivergent. The divergence term will be small enough to be balanced by the vorticity advection terms only if

$$\left|\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)\right| \lesssim 10^{-6}\ \mathrm{s}^{-1}$$

so that the horizontal divergence must be small compared to the vorticity in synoptic-scale systems. From the aforementioned scalings and the definition of the Rossby number, we see that

$$\left|\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \Big/ f_0\right| \lesssim \mathrm{Ro}^2$$

and

$$\left|\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \Big/ \zeta\right| \lesssim \mathrm{Ro}$$

Thus the ratio of the horizontal divergence to the relative vorticity is the same magnitude as the ratio of relative vorticity to planetary vorticity.

Retaining only the terms of order $10^{-10}\ \mathrm{s}^{-2}$ in the vorticity equation yields the approximate form valid for synoptic-scale motions,

$$\frac{D_h(\zeta + f)}{Dt} = -f\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \tag{4.22a}$$

where

$$\frac{D_h}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}$$

As mentioned earlier, (4.22a) is not accurate in intense cyclonic storms. For these the relative vorticity should be retained in the divergence term:

$$\frac{D_h(\zeta + f)}{Dt} = -(\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \tag{4.22b}$$

Equation (4.22a) states that the change of absolute vorticity following the horizontal motion on the synoptic scale is given approximately by the concentration or dilution of planetary vorticity caused by the convergence or divergence of the horizontal flow, respectively. In (4.22b), however, it is the concentration or dilution of absolute vorticity that leads to changes in absolute vorticity following the motion.

The form of the vorticity equation given in (4.22b) also indicates why cyclonic disturbances can be much more intense than anticyclones. For a fixed amplitude of convergence, relative vorticity will increase, and the factor $(\zeta + f)$ becomes larger, which leads to even higher rates of increase in the relative vorticity. For a fixed rate of divergence, however, relative vorticity will decrease, but when $\zeta \to -f$, the divergence term on the right approaches zero and the relative vorticity cannot become more negative no matter how strong the divergence. (This difference in the potential intensity of cyclones and anticyclones was discussed in Section 3.2.5 in connection with the gradient wind approximation.)

The approximate forms given in (4.22a) and (4.22b) do not remain valid, however, in the vicinity of atmospheric fronts. The horizontal scale of variation in frontal zones is only ~100 km, and the vertical velocity scale is ~10 cm s^{-1}. For these scales, vertical advection, tilting, and solenoidal terms all may become as large as the divergence term.

4.5 VORTICITY IN BAROTROPIC FLUIDS

A model that has proved useful for elucidating some aspects of the horizontal structure of large-scale atmospheric motions is the barotropic model. In the most general version of this model, the atmosphere is represented as a homogeneous incompressible fluid of variable depth, $h(x, y, t) = z_2 - z_1$, where z_2 and z_1 are the heights of the upper and lower boundaries, respectively. In this model, a special form of potential vorticity is conserved following the motion. A simpler situation arises if the fluid depth is constant. In that case it is absolute vorticity that is conserved following the motion.

4.5.1 The Barotropic (Rossby) Potential Vorticity Equation

For a homogeneous incompressible fluid, the continuity equation (2.31) simplifies to $\nabla \bullet \mathbf{U} = 0$ or, in Cartesian coordinates,

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

so that the vorticity equation (4.22b) may be written

$$\frac{D_h(\zeta + f)}{Dt} = (\zeta + f)\left(\frac{\partial w}{\partial z}\right) \tag{4.23}$$

Section 3.4 showed that the thermal wind in a barotropic fluid vanishes so that the geostrophic wind is independent of height. Letting the vorticity in (4.23) be approximated by the geostrophic vorticity ζ_g and the wind by the geostrophic wind (u_g, v_g), we can integrate vertically from z_1 to z_2 to get

$$h\frac{D_h\left(\zeta_g + f\right)}{Dt} = \left(\zeta_g + f\right)\left[w(z_2) - w(z_1)\right] \tag{4.24}$$

However, because $w \equiv Dz/Dt$ and $h \equiv h(x, y, t)$,

$$w(z_2) - w(z_1) = \frac{Dz_2}{Dt} - \frac{Dz_1}{Dt} = \frac{D_h h}{Dt} \tag{4.25}$$

Substituting from (4.25) into (4.24) we get

$$\frac{1}{\left(\zeta_g + f\right)}\frac{D_h\left(\zeta_g + f\right)}{Dt} = \frac{1}{h}\frac{D_h h}{Dt}$$

or

$$\frac{D_h \ln\left(\zeta_g + f\right)}{Dt} = \frac{D_h \ln h}{Dt}$$

which implies that

$$\frac{D_h}{Dt}\left(\frac{\zeta_g + f}{h}\right) = 0 \tag{4.26}$$

This is just the potential vorticity conservation theorem for a barotropic fluid, which was first obtained by C. G. Rossby [see (4.13)]. The quantity conserved following the motion in (4.26) is the *Rossby potential vorticity*.

4.5.2 The Barotropic Vorticity Equation

If the flow is purely horizontal ($w = 0$), as is the case for barotropic flow in a fluid of constant depth, the divergence term vanishes in (4.23) and we obtain the *barotropic vorticity equation*

$$\frac{D_h\left(\zeta_g + f\right)}{Dt} = 0 \tag{4.27}$$

which states that absolute vorticity is conserved following the horizontal motion. More generally, absolute vorticity is conserved for any fluid layer in which the divergence of the horizontal wind vanishes, without the requirement that the flow be geostrophic. For horizontal motion that is nondivergent ($\partial u/\partial x + \partial v/\partial y = 0$), the flow field can be represented by a *streamfunction* $\psi\ (x,\ y)$ defined so that the velocity components are given as $u = -\partial\psi/\partial y,\ \ v = +\partial\psi/\partial x$. The vorticity is then given by

$$\zeta = \partial v/\partial x - \partial u/\partial y = \partial^2\psi/\partial x^2 + \partial^2\psi/\partial y^2 \equiv \nabla^2\psi$$

Thus, the velocity field and the vorticity can both be represented in terms of the variation of the single scalar field $\psi\ (x,\ y)$, and (4.27) can be written as a prognostic equation for vorticity in the form:

$$\frac{\partial}{\partial t}\nabla^2\psi = -\mathbf{V}_\psi \bullet \nabla\left(\nabla^2\psi + f\right) \tag{4.28}$$

where $\mathbf{V}_\psi \equiv \mathbf{k} \times \nabla\psi$ is a nondivergent horizontal wind. Equation (4.28) states that the local tendency of relative vorticity is given by the advection of absolute vorticity. This equation can be solved numerically to predict the evolution of the streamfunction, and hence of the vorticity and wind fields (see Section 13.4). Because the flow in the midtroposphere is often nearly nondivergent on the synoptic scale, (4.28) provides a surprisingly good model for short-term forecasts of the synoptic-scale 500-hPa flow field.

4.6 THE BAROCLINIC (ERTEL) POTENTIAL VORTICITY EQUATION

Section 4.3 used the circulation theorem and mass continuity to show that Ertel's potential vorticity, $P \equiv (\zeta_\theta + f)(-g\partial\theta/\partial p)$, is conserved following the motion in adiabatic flow. If diabatic heating or frictional torques are present, P is no longer conserved. An equation governing the rate of change of P following the motion in such circumstances can, however, be derived fairly simply starting from the equations of motion in their isentropic coordinate form.

4.6.1 Equations of Motion in Isentropic Coordinates

If the atmosphere is stably stratified so that potential temperature θ is a monotonically increasing function of height, θ may be used as an independent vertical coordinate. The vertical "velocity" in this coordinate system is just $\dot{\theta} \equiv D\theta/Dt$. Thus, adiabatic motions are two dimensional when viewed in an isentropic coordinate frame. An infinitesimal control volume in isentropic coordinates with cross-sectional area δA and vertical extent $\delta\theta$ has a mass

$$\delta M = \rho\delta A\delta z = \delta A\left(-\frac{\delta p}{g}\right) = \frac{\delta A}{g}\left(-\frac{\partial p}{\partial\theta}\right)\delta\theta = \sigma\delta A\delta\theta \tag{4.29}$$

Here the "density" in $(x,\ y,\ \theta)$ space (i.e., as shown in Fig. 4.7 the quantity that when multiplied by the "volume" element $\delta A\delta\theta$ yields the mass element δM) is defined as

$$\sigma \equiv -g^{-1}\partial p/\partial\theta \tag{4.30}$$

The horizontal momentum equation in isentropic coordinates may be obtained by transforming the isobaric form (4.19) to yield

$$\frac{\partial\mathbf{V}}{\partial t} + \nabla_\theta\left(\frac{\mathbf{V}\bullet\mathbf{V}}{2} + \Psi\right) + (\zeta_\theta + f)\,\mathbf{k}\times\mathbf{V} = -\dot{\theta}\frac{\partial\mathbf{V}}{\partial\theta} + \mathbf{F}_r \tag{4.31}$$

where ∇_θ is the gradient on an isentropic surface, $\zeta_\theta \equiv \mathbf{k}\bullet\nabla_\theta\times\mathbf{V}$ is the isentropic relative vorticity originally introduced in (4.11), and $\Psi \equiv c_pT + \Phi$ is the Montgomery streamfunction (see Problem 10 in Chapter 2). We have included a frictional term $\mathbf{F}_r$ on the right-hand side, along with the diabatic vertical advection term. The continuity equation can be derived with the aid of (4.29) in a manner analogous to that used for the isobaric system in Section 3.1.2. The result is

$$\frac{\partial\sigma}{\partial t} + \nabla_\theta\bullet(\sigma\mathbf{V}) = -\frac{\partial}{\partial\theta}\left(\sigma\dot{\theta}\right) \tag{4.32}$$

The Ψ and σ fields are linked through the pressure field by the hydrostatic equation, which in the isentropic system takes the form

$$\frac{\partial\Psi}{\partial\theta} = \Pi(p) \equiv c_p\left(\frac{p}{p_s}\right)^{R/c_p} = c_p\frac{T}{\theta} \tag{4.33}$$

where Π is called the *Exner function*. Equations (4.30)–(4.33) form a closed set for prediction of $\mathbf{V}$, σ, Ψ, and p, provided that $\dot{\theta}$ and $\mathbf{F}_r$ are known.

4.6.2 The Potential Vorticity Equation

If we take $\mathbf{k} \bullet \nabla_\theta \times$(4.31) and rearrange the resulting terms, we obtain the isentropic vorticity equation:

$$\frac{\widetilde{D}}{Dt}(\zeta_\theta + f) + (\zeta_\theta + f)\,\nabla_\theta \bullet \mathbf{V} = \mathbf{k} \bullet \nabla_\theta \times \left(\mathbf{F}_r - \dot{\theta}\frac{\partial \mathbf{V}}{\partial \theta}\right) \tag{4.34}$$

where

$$\frac{\widetilde{D}}{Dt} = \frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla_\theta$$

is the total derivative following the horizontal motion on an isentropic surface.

Noting that $\sigma^{-2}\partial\sigma/\partial t = -\partial\sigma^{-1}/\partial t$, we can rewrite (4.32) in the form

$$\frac{\widetilde{D}}{Dt}\left(\sigma^{-1}\right) - \left(\sigma^{-1}\right)\nabla_\theta \bullet \mathbf{V} = \sigma^{-2}\frac{\partial}{\partial \theta}\left(\sigma\dot{\theta}\right) \tag{4.35}$$

Multiplying each term in (4.34) by σ^{-1} and in (4.35) by $(\zeta_\theta + f)$ and adding, we obtain the desired conservation law:

$$\frac{\widetilde{D}P}{Dt} = \frac{\partial P}{\partial t} + \mathbf{V} \bullet \nabla_\theta P = \frac{P}{\sigma}\frac{\partial}{\partial \theta}\left(\sigma\dot{\theta}\right) + \sigma^{-1}\mathbf{k} \bullet \nabla_\theta \times \left(\mathbf{F}_r - \dot{\theta}\frac{\partial \mathbf{V}}{\partial \theta}\right) \tag{4.36}$$

where $P \equiv (\zeta_\theta + f)/\sigma$ is the Ertel potential vorticity defined in (4.12). If the diabatic and frictional terms on the right-hand side of (4.36) can be evaluated, it is possible to determine the evolution of P following the *horizontal* motion on an isentropic surface. When the diabatic and frictional terms are small, potential vorticity is approximately conserved following the motion on isentropic surfaces.

Weather disturbances that have sharp gradients in dynamical fields, such as jets and fronts, are associated with large anomalies in the Ertel potential vorticity. In the upper troposphere such anomalies tend to be advected rapidly under nearly adiabatic conditions. Thus, the potential vorticity anomaly patterns are conserved materially on isentropic surfaces. This material conservation property makes potential vorticity anomalies particularly useful in identifying and tracing the evolution of meteorological disturbances.

4.6.3 Integral Constraints on Isentropic Vorticity

The isentropic vorticity equation (4.34) can be written in the form

$$\frac{\partial \zeta_\theta}{\partial t} = -\nabla_\theta \bullet \left[(\zeta_\theta + f)\mathbf{V}\right] + \mathbf{k} \bullet \nabla_\theta \times \left(\mathbf{F}_r - \dot{\theta}\frac{\partial \mathbf{V}}{\partial \theta}\right) \tag{4.37}$$

Using the fact that any vector **A** satisfies the relationship

$$\mathbf{k} \bullet (\boldsymbol{\nabla}_\theta \times \mathbf{A}) = \boldsymbol{\nabla}_\theta \bullet (\mathbf{A} \times \mathbf{k})$$

we can rewrite (4.37) in the form

$$\frac{\partial \zeta_\theta}{\partial t} = -\boldsymbol{\nabla}_\theta \bullet \left[(\zeta_\theta + f)\mathbf{V} - \left(\mathbf{F}_r - \dot{\theta}\frac{\partial \mathbf{V}}{\partial \theta} \right) \times \mathbf{k} \right] \tag{4.38}$$

Equation (4.38) expresses the remarkable fact that isentropic vorticity can only be changed by the divergence or convergence of the horizontal flux vector in brackets on the right-hand side. The vorticity cannot be changed by vertical transfer across the isentropes. Furthermore, integration of (4.38) over the area of an isentropic surface and application of the divergence theorem (Appendix C.2) show that for an isentrope that does not intersect the surface of the earth the global average of ζ_θ is constant. Furthermore, integration of ζ_θ over the sphere shows that the global average ζ_θ is exactly zero. Vorticity on such an isentrope is neither created nor destroyed; it is merely concentrated or diluted by horizontal fluxes along the isentropes.

PROBLEMS

4.1. What is the circulation about a square of 1000 km on a side for an easterly (i.e., westward flowing) wind that decreases in magnitude toward the north at a rate of 10 m s^{-1} per 500 km? What is the mean relative vorticity in the square?

4.2. A cylindrical column of air at 30°N with radius 100 km expands to twice its original radius. If the air is initially at rest, what is the mean tangential velocity at the perimeter after expansion?

4.3. An air parcel at 30°N moves northward conserving absolute vorticity. If its initial relative vorticity is 5×10^{-5} s^{-1}, what is its relative vorticity upon reaching 90°N?

4.4. An air column at 60°N with $\zeta = 0$ initially stretches from the surface to a fixed tropopause at 10 km height. If the air column moves until it is over a mountain barrier 2.5 km high at 45°N, what is its absolute vorticity and relative vorticity as it passes the mountain top assuming that the flow satisfies the barotropic potential vorticity equation?

4.5. Find the average vorticity within a cylindrical annulus of inner radius 200 km and outer radius 400 km if the tangential velocity distribution is given by $V = A/r$, where $A = 10^6$ m^2 s^{-1} and r is in meters. What is the average vorticity within the inner circle of radius 200 km?

4.6. Show that the anomalous gradient wind cases discussed in Section 3.2.5 have negative absolute circulation in the Northern Hemisphere and hence have negative average absolute vorticity.

4.7. Compute the rate of change of circulation about a square in the (x, y) plane with corners at (0, 0), $(0, L)$, (L, L), and $(L, 0)$ if temperature increases eastward at a rate of 1°C per 200 km and pressure increases northward at a rate of 1 hPa per 200 km. Let $L = 1000$ km and the pressure at the point (0, 0) be 1000 hPa.

4.8. Verify the identity (4.18) by expanding the vectors in Cartesian components.

4.9. Derive a formula for the dependence of depth on radius for an incompressible fluid in solid-body rotation in a cylindrical tank with a flat bottom and free surface at the upper boundary. Let H be the depth at the center of the tank, Ω the angular velocity of rotation of the tank, and a the radius of the tank.

4.10. By how much does the relative vorticity change for a column of fluid in a rotating cylinder if the column is moved from the center of the tank to a distance 50 cm from the center? The tank is rotating at the rate of 20 revolutions per minute, the depth of the fluid at the center is 10 cm, and the fluid is initially in solid-body rotation.

4.11. A cyclonic vortex is in cyclostrophic balance with a tangential velocity profile given by the expression $V = V_{0(}r/r_0)^n$ where V_0 is the tangential velocity component at the distance r_0 from the vortex center. Compute the circulation about a streamline at radius r, the vorticity at radius r, and the pressure at radius r. (Let p_0 be the pressure at r_0 and assume that density is a constant.)

4.12. A westerly zonal flow at 45° is forced to rise adiabatically over a north–south-oriented mountain barrier. Before striking the mountain, the westerly wind increases linearly toward the south at a rate of 10 m s^{-1} per 1000 km. The crest of the mountain range is at 800 hPa and the tropopause, located at 300 hPa, remains undisturbed. What is the initial relative vorticity of the air? What is its relative vorticity when it reaches the crest if it is deflected 5° latitude toward the south during the forced ascent? If the current assumes a uniform speed of 20 m s^{-1} during its ascent to the crest, what is the radius of curvature of the streamlines at the crest?

4.13. A cylindrical vessel of radius a and constant depth H rotating at an angular velocity Ω about its vertical axis of symmetry is filled with a homogeneous, incompressible fluid that is initially at rest with respect to the vessel. A volume of fluid V is then withdrawn through a point sink at the center of the cylinder, thus creating a vortex. Neglecting friction, derive an expression for the resulting relative azimuthal velocity as a function of radius (i.e., the

velocity in a coordinate system rotating with the tank). Assume that the motion is independent of depth and that $V \ll \pi a^2 H$. Also compute the relative vorticity and the relative circulation.

4.14. (a) How far must a zonal ring of air initially at rest with respect to the earth's surface at 60° latitude and 100-km height be displaced latitudinally in order to acquire an easterly (east to west) component of 10 m s^{-1} with respect to the earth's surface? (b) To what height must it be displaced vertically in order to acquire the same velocity? Assume a frictionless atmosphere.

4.15. The horizontal motion within a cylindrical annulus with permeable walls of inner radius 10 cm, outer radius 20 cm, and 10-cm depth is independent of height and azimuth and is represented by the expressions $u = 7 - 0.2r$, $v = 40 + 2r$, where u and v are the radial and tangential velocity components in cm s^{-1}, positive outward and counterclockwise, respectively, and r is distance from the center of the annulus in centimeters. Assuming an incompressible fluid, find
(a) the circulation about the annular ring,
(b) the average vorticity within the annular ring,
(c) the average divergence within the annular ring, and
(d) the average vertical velocity at the top of the annulus if it is zero at the base.

4.16. Prove that, as stated below Eq. (4.38), the globally averaged isentropic vorticity on an isentropic surface that does not intersect the ground must be zero. Show that the same result holds for the isobaric vorticity on an isobaric surface.

MATLAB EXERCISES

M4.1. Section 4.5.2 showed that for nondivergent horizontal motion, the flow field can be represented by a *streamfunction* $\psi\ (x, y)$, and the vorticity is then given by $\zeta = \partial^2\psi/\partial x^2 + \partial^2\psi/\partial y^2 \equiv \nabla^2\psi$. Thus, if the vorticity is represented by a single sinusoidal wave distribution in both x and y, the streamfunction has the same spatial distribution as the vorticity and the opposite sign, as can be verified easily from the fact that the second derivative of a sine is proportional to minus the same sine function. An example is shown in the MATLAB script **vorticity_1.m**. However, when the vorticity pattern is localized in space, the space scales of the streamfunction and vorticity are much different. This latter situation is illustrated in the MATLAB script **vorticity_demo.m**, which shows the streamfunction corresponding to a point source of vorticity at (x, y) = (0, 0). For this problem you must modify the code in **vorticity_1.m** by specifying

$\zeta(x, y) = \exp\left[-b\left(x^2 + y^2\right)\right]$ where b is a constant. Run the model for several values of b from $b = 1 \times 10^{-4}$ km^{-2} to 4×10^{-7} km^{-2}. Show in a table or line plot the dependence of the ratio of the horizontal scales on which vorticity and the streamfunction decay to one-half of their maximum values as a function of the parameter b. Note that for geostrophic motions (with constant Coriolis parameter), the streamfunction defined here is proportional to geopotential height. What can you conclude from this exercise about the information content of a 500-hPa height map versus that of a map of the 500-hPa vorticity field?

M4.2. The MATLAB scripts **geowinds_1.m** (to be used when the mapping toolbox is available) and **geowinds_2.m** (to be used if no mapping toolbox is available) contain contour plots showing the observed 500-hPa height and horizontal wind fields for November 10, 1998 in the North American sector. Also shown is a color plot of the magnitude of the 500-hPa wind with height contours superposed. Using centered difference formulas (see Section 13.3.1), compute the geostrophic wind components, the magnitude of the geostrophic wind, the relative vorticity, the vorticity of the geostrophic wind, and the vorticity minus vorticity of the geostrophic wind. Following the models of Figs. 1 and 2, superpose these fields on the maps of the 500-hPa height field. Explain the distribution and the sign of regions with large differences between vorticity and geostrophic vorticity in terms of the force balances that were studied in Chapter 3.

Suggested References

Williams and Elder, *Fluid Physics for Oceanographers and Physicists*, provides an introduction to vorticity dynamics at an elementary level. This book also provides a good general introduction to fluid dynamics.

Acheson, *Elementary Fluid Dynamics*, provides a good introduction to vorticity at a graduate level.

Hoskins et al. (1985) provide an advanced discussion of Ertel potential vorticity and its uses in diagnosis and prediction of synoptic-scale disturbances.

CHAPTER 5

The Planetary Boundary Layer

The *planetary boundary layer* is that portion of the atmosphere in which the flow field is strongly influenced directly by interaction with the surface of the earth. Ultimately this interaction depends on molecular viscosity. It is, however, only within a few millimeters of the surface, where vertical shears are very intense, that molecular diffusion is comparable to other terms in the momentum equation. Outside this *viscous sublayer* molecular diffusion is not important in the boundary layer equations for the mean wind, although it is still important for small-scale turbulent eddies. However, viscosity still has an important indirect role; it causes the velocity to vanish at the surface. As a consequence of this *no-slip boundary condition,* even a fairly weak wind will cause a large-velocity shear near the surface, which continually leads to the development of turbulent eddies. These turbulent motions have spatial and temporal variations at scales much smaller than those resolved by the meteorological observing network. Such shear-induced eddies, together with convective eddies caused by surface heating, are very effective in transferring momentum to the surface and transferring heat (latent and sensible) away from the surface at rates many orders of magnitude faster than can be done by molecular processes. The depth of the planetary boundary layer produced by this turbulent transport may range from as little as 30 m in conditions of large

static stability to more than 3 km in highly convective conditions. For average midlatitude conditions the planetary boundary layer extends through the lowest kilometer of the atmosphere and thus contains about 10% of the mass of the atmosphere.

The dynamical structure of the flow in the planetary boundary layer is not produced directly by viscosity. Rather, it is largely determined by the fact that the atmospheric flow is turbulent. In the *free atmosphere* (i.e., the region above the planetary boundary layer), this turbulence can be ignored in an approximate treatment of synoptic-scale motions, except perhaps in the vicinity of jet streams, fronts, and convective clouds. However, in the boundary layer the dynamical equations of the previous chapters must be modified to properly represent the effects of turbulence.

5.1 ATMOSPHERIC TURBULENCE

Turbulent flow contains irregular quasi-random motions spanning a continuous spectrum of spatial and temporal scales. Such eddies cause nearby air parcels to drift apart and thus mix properties such as momentum and potential temperature across the boundary layer. Unlike the large-scale rotational flows discussed in earlier chapters, which have depth scales that are small compared to their horizontal scales, the turbulent eddies of concern in the planetary boundary layer tend to have similar scales in the horizontal and vertical. The maximum eddy length scale is thus limited by the boundary layer depth to be about 10^3 m. The minimum length scale (10^{-3} m) is that of the smallest eddies that can exist in the presence of diffusion by molecular friction.

Even when observations are taken with very short temporal and spatial separations, a turbulent flow will always have scales that are unresolvable because they have frequencies greater than the observation frequency and spatial scales smaller than the scale of separation of the observations. Outside the boundary layer, in the free atmosphere, the problem of unresolved scales of motion is usually not a serious one for the diagnosis or forecasting of synoptic and larger scale circulations (although it is crucial for the mesoscale circulations discussed in Chapter 9). The eddies that contain the bulk of the energy in the free atmosphere are resolved by the synoptic network. However, in the boundary layer, unresolved turbulent eddies are of critical importance. Through their transport of heat and moisture away from the surface they maintain the surface energy balance, and through their transport of momentum to the surface they maintain the momentum balance. The latter process dramatically alters the momentum balance of the large-scale flow in the boundary layer so that geostrophic balance is no longer an adequate approximation to the large-scale wind field. It is this aspect of boundary layer dynamics that is of primary importance for dynamical meteorology.

5.1.1 The Boussinesq Approximation

The basic state (standard atmosphere) density varies across the lowest kilometer of the atmosphere by only about 10%, and the fluctuating component of density deviates from the basic state by only a few percentage points. These circumstances might suggest that boundary layer dynamics could be modeled by setting density constant and using the theory of homogeneous incompressible fluids. This is, however, generally not the case. Density fluctuations cannot be totally neglected because they are essential for representing the buoyancy force (see the discussion in Section 2.7.3).

Nevertheless, it is still possible to make some important simplifications in dynamical equations for application in the boundary layer. The *Boussinesq approximation* is a form of the dynamical equations that is valid for this situation. In this approximation, density is replaced by a constant mean value, ρ_0, everywhere except in the buoyancy term in the vertical momentum equation. The horizontal momentum equations (2.24) and (2.25) can then be expressed in Cartesian coordinates as

$$\frac{Du}{Dt} = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} + fv + F_{rx} \tag{5.1}$$

and

$$\frac{Dv}{Dt} = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} - fu + F_{ry} \tag{5.2}$$

while, with the aid of (2.28) and (2.51) the vertical momentum equation becomes

$$\frac{Dw}{Dt} = -\frac{1}{\rho_0}\frac{\partial p}{\partial z} + g\frac{\theta}{\theta_0} + F_{rz} \tag{5.3}$$

Here, as in Section 2.7.3, θ designates the departure of potential temperature from its basic state value $\theta_0(z)$.[1] Thus, the total potential temperature field is given by $\theta_{tot} = \theta(x, y, z, t) + \theta_0(z)$, and the adiabatic thermodynamic energy equation has the form of (2.54):

$$\frac{D\theta}{Dt} = -w\frac{d\theta_0}{dz} \tag{5.4}$$

Finally, the continuity equation (2.34) under the Boussinesq approximation is

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{5.5}$$

[1] The reason that we have not used the notation θ' to designate the fluctuating part of the potential temperature field will become apparent in Section 5.1.2.

5.1.2 Reynolds Averaging

In a turbulent fluid, a field variable such as velocity measured at a point generally fluctuates rapidly in time as eddies of various scales pass the point. In order that measurements be truly representative of the large-scale flow, it is necessary to average the flow over an interval of time long enough to average out small-scale eddy fluctuations, but still short enough to preserve the trends in the large-scale flow field. To do this we assume that the field variables can be separated into slowly varying mean fields and rapidly varying turbulent components. Following the scheme introduced by Reynolds, we then assume that for any field variables, w and θ, for example, the corresponding means are indicated by overbars and the fluctuating components by primes. Thus, $w = \overline{w}+w'$ and $\theta = \overline{\theta}+\theta'$. By definition, the means of the fluctuating components vanish; the product of a deviation with a mean also vanishes when the time average is applied. Thus,

$$\overline{w'\overline{\theta}} = \overline{w'}\,\overline{\theta} = 0$$

where we have used the fact that a mean variable is constant over the period of averaging. The average of the product of deviation components (called the *covariance*) does not generally vanish. Thus, for example, if on average the turbulent vertical velocity is upward where the potential temperature deviation is positive and downward where it is negative, the product $\overline{w'\theta'}$ is positive and the variables are said to be correlated positively. These averaging rules imply that the average of the product of two variables will be the product of the average of the means plus the product of the average of the deviations:

$$\overline{w\theta} = \overline{(\overline{w} + w')(\overline{\theta} + \theta')} = \overline{w}\overline{\theta} + \overline{w'\theta'}$$

Before applying the Reynolds decomposition to (5.1)–(5.4), it is convenient to rewrite the total derivatives in each equation in flux form. For example, the term on the left in (5.1) may be manipulated with the aid of the continuity equation (5.5) and the chain rule of differentiation to yield

$$\begin{aligned}\frac{Du}{Dt} &= \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} + u\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)\\ &= \frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} + \frac{\partial uw}{\partial z}\end{aligned} \tag{5.6}$$

Separating each dependent variable into mean and fluctuating parts, substituting into (5.6), and averaging then yields

$$\overline{\frac{Du}{Dt}} = \frac{\partial \bar{u}}{\partial t} + \frac{\partial}{\partial x}\left(\bar{u}\bar{u} + \overline{u'u'}\right) + \frac{\partial}{\partial y}\left(\bar{u}\bar{v} + \overline{u'v'}\right) + \frac{\partial}{\partial z}\left(\bar{u}\bar{w} + \overline{u'w'}\right) \tag{5.7}$$

Noting that the mean velocity fields satisfy the continuity equation (5.5), we can rewrite (5.7), as

$$\overline{\frac{Du}{Dt}} = \frac{\bar{D}\bar{u}}{Dt} + \frac{\partial}{\partial x}\left(\overline{u'u'}\right) + \frac{\partial}{\partial y}\left(\overline{u'v'}\right) + \frac{\partial}{\partial z}\left(\overline{u'w'}\right) \tag{5.8}$$

where

$$\frac{\overline{D}}{Dt} = \frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x} + \overline{v}\frac{\partial}{\partial y} + \overline{w}\frac{\partial}{\partial z}$$

is the rate of change following the mean motion.

The mean equations thus have the form

$$\frac{\bar{D}\bar{u}}{Dt} = -\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial x} + f\bar{v} - \left[\frac{\partial \overline{u'u'}}{\partial x} + \frac{\partial \overline{u'v'}}{\partial y} + \frac{\partial \overline{u'w'}}{\partial z}\right] + \bar{F}_{rx} \tag{5.9}$$

$$\frac{\bar{D}\bar{v}}{Dt} = -\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial y} - f\bar{u} - \left[\frac{\partial \overline{u'v'}}{\partial x} + \frac{\partial \overline{v'v'}}{\partial y} + \frac{\partial \overline{v'w'}}{\partial z}\right] + \bar{F}_{ry} \tag{5.10}$$

$$\frac{\bar{D}\bar{w}}{Dt} = -\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial z} + g\frac{\bar{\theta}}{\theta_0} - \left[\frac{\partial \overline{u'w'}}{\partial x} + \frac{\partial \overline{v'w'}}{\partial y} + \frac{\partial \overline{w'w'}}{\partial z}\right] + \bar{F}_{rz} \tag{5.11}$$

$$\frac{\bar{D}\bar{\theta}}{Dt} = -\bar{w}\frac{d\theta_0}{dz} - \left[\frac{\partial \overline{u'\theta'}}{\partial x} + \frac{\partial \overline{v'\theta'}}{\partial y} + \frac{\partial \overline{w'\theta'}}{\partial z}\right] \tag{5.12}$$

$$\frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} + \frac{\partial \bar{w}}{\partial z} = 0 \tag{5.13}$$

The various covariance terms in square brackets in (5.9)–(5.12) represent turbulent fluxes. For example, $\overline{w'\theta'}$ is a vertical turbulent heat flux in kinematic form. Similarly $\overline{w'u'} = \overline{u'w'}$ is a vertical turbulent flux of zonal momentum. For many boundary layers the magnitudes of the turbulent flux divergence terms are of the same order as the other terms in (5.9)–(5.12). In such cases, it is not possible to neglect the turbulent flux terms even when only the mean flow is of direct interest. Outside the boundary layer the turbulent fluxes are often sufficiently weak so that the terms in square brackets in (5.9)–(5.12) can be neglected in the analysis of large-scale flows. This assumption was implicitly made in Chapters 3 and 4.

The complete equations for the mean flow (5.9)–(5.13), unlike the equations for the total flow (5.1)–(5.5), and the approximate equations of Chapters 3 and 4, are not a closed set, as in addition to the five unknown mean variables $\overline{u}$, $\overline{v}$, $\overline{w}$, $\overline{\theta}$, $\overline{p}$, there are unknown turbulent fluxes. To solve these equations, *closure* assumptions must be made to approximate the unknown fluxes in terms of the five known mean state variables. Away from regions with horizontal inhomogeneities

(e.g., shorelines, towns, forest edges), we can simplify by assuming that the turbulent fluxes are horizontally homogeneous so that it is possible to neglect the horizontal derivative terms in square brackets in comparison to the terms involving vertical differentiation.

5.2 TURBULENT KINETIC ENERGY

Vortex stretching and twisting associated with turbulent eddies always tend to cause turbulent energy to flow toward the smallest scales, where it is dissipated by viscous diffusion. Thus, there must be continuing production of turbulence if the turbulent kinetic energy is to remain statistically steady. The primary source of boundary layer turbulence depends critically on the structure of the wind and temperature profiles near the surface. If the lapse rate is unstable, boundary layer turbulence is convectively generated. If it is stable, then instability associated with wind shear must be responsible for generating turbulence in the boundary layer. The comparative roles of these processes can best be understood by examining the budget for turbulent kinetic energy.

To investigate the production of turbulence, we subtract the component mean momentum equations (5.9)–(5.11) from the corresponding unaveraged equations (5.1)–(5.3). We then multiply the results by u', v', w', respectively, add the resulting three equations, and average to obtain the turbulent kinetic energy equation. The complete statement of this equation is quite complicated, but its essence can be expressed symbolically as

$$\frac{\bar{D}\,(TKE)}{Dt} = MP + BPL + TR - \varepsilon \tag{5.14}$$

where $TKE \equiv (\overline{u'^2} + \overline{v'^2} + \overline{w'^2})/2$ is the turbulent kinetic energy per unit mass, *MP* is the mechanical production, *BPL* is the buoyant production or loss, *TR* designates redistribution by transport and pressure forces, and ε designates frictional dissipation. ε is always positive, reflecting the dissipation of the smallest scales of turbulence by molecular viscosity.

The buoyancy term in (5.14) represents a conversion of energy between mean flow potential energy and turbulent kinetic energy. It is positive for motions that lower the center of mass of the atmosphere and negative for motions that raise it. The buoyancy term has the form[2]

$$BPL \equiv \overline{w'\theta'}\left(g\,/\theta_0\right)$$

[2] In practice, buoyancy in the boundary layer is modified by the presence of water vapor, which has a density significantly lower than that of dry air. The potential temperature should be replaced by virtual potential temperature in (5.14) in order to include this effect. (See, for example, Curry and Webster, 1999, p.67.)

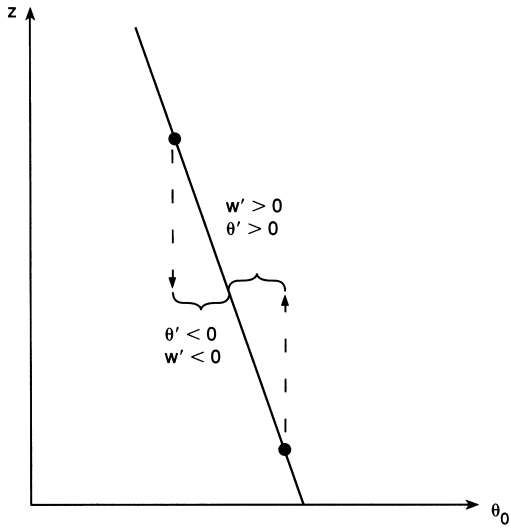


Fig. 5.1 Correlation between vertical velocity and potential temperature perturbations for upward or downward parcel displacements when the mean potential temperature $\theta_0(z)$ decreases with height.

Positive buoyancy production occurs when there is heating at the surface so that an unstable temperature lapse rate (see Section 2.7.2) develops near the ground and spontaneous convective overturning can occur. As shown in the schematic of Fig. 5.1, convective eddies have positively correlated vertical velocity and potential temperature fluctuations and hence provide a source of turbulent kinetic energy and positive heat flux. This is the dominant source in a convectively unstable boundary layer. For a statically stable atmosphere, *BPL* is negative, which tends to reduce or eliminate turbulence.

For both statically stable and unstable boundary layers, turbulence can be generated mechanically by dynamical instability due to wind shear. This process is represented by the mechanical production term in (5.14), which represents a conversion of energy between mean flow and turbulent fluctuations. This term is proportional to the shear in the mean flow and has the form

$$MP \equiv -\overline{u'w'}\frac{\partial \bar{u}}{\partial z} - \overline{v'w'}\frac{\partial \bar{v}}{\partial z} \tag{5.15}$$

MP is positive when the momentum flux is directed down the gradient of the mean momentum. Thus, if the mean vertical shear near the surface is westerly ($\partial \bar{u} / \partial z > 0$), then $\overline{u'w'} < 0$ for $MP > 0$.

In a statically stable layer, turbulence can exist only if mechanical production is large enough to overcome the damping effects of stability and viscous dissipation.

This condition is measured by a quantity called the *flux Richardson number*. It is defined as

$$Rf \equiv -BPL/MP$$

If the boundary layer is statically unstable, then Rf < 0 and turbulence is sustained by convection. For stable conditions, Rf will be greater than zero. Observations suggest that only when Rf is less than about 0.25 (i.e., mechanical production exceeds buoyancy damping by a factor of 4) is the mechanical production intense enough to sustain turbulence in a stable layer. Since *MP* depends on the shear, it always becomes large close enough to the surface. However, as the static stability increases, the depth of the layer in which there is net production of turbulence shrinks. Thus, when there is a strong temperature inversion, such as produced by nocturnal radiative cooling, the boundary layer depth may be only a few decameters, and vertical mixing is strongly suppressed.

5.3 PLANETARY BOUNDARY LAYER MOMENTUM EQUATIONS

For the special case of horizontally homogeneous turbulence above the viscous sublayer, molecular viscosity and horizontal turbulent momentum flux divergence terms can be neglected. The mean flow horizontal momentum equations (5.9) and (5.10) then become

$$\frac{\bar{D}\bar{u}}{Dt} = -\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial x} + f\bar{v} - \frac{\partial \overline{u'w'}}{\partial z} \tag{5.16}$$

$$\frac{\bar{D}\bar{v}}{Dt} = -\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial y} - f\bar{u} - \frac{\partial \overline{v'w'}}{\partial z} \tag{5.17}$$

In general, (5.16) and (5.17) can only be solved for $\overline{u}$ and $\overline{v}$ if the vertical distribution of the turbulent momentum flux is known. Because this depends on the structure of the turbulence, no general solution is possible. Rather, a number of approximate semiempirical methods are used.

For midlatitude synoptic-scale motions, Section 2.4 showed that to a first approximation the inertial acceleration terms [the terms on the left in (5.16) and (5.17)] could be neglected compared to the Coriolis force and pressure gradient force terms. Outside the boundary layer, the resulting approximation was then simply geostrophic balance. In the boundary layer the inertial terms are still small compared to the Coriolis force and pressure gradient force terms, but the turbulent flux terms must be included. Thus, to a first approximation, planetary boundary layer equations express a three-way balance among the Coriolis force, the pressure gradient force, and the turbulent momentum flux divergence:

$$f\left(\bar{v} - \bar{v}_g\right) - \frac{\partial \overline{u'w'}}{\partial z} = 0 \tag{5.18}$$

$$-f\left(\bar{u}-\bar{u}_g\right)-\frac{\partial \overline{v'w'}}{\partial z}=0 \tag{5.19}$$

where (2.23) is used to express the pressure gradient force in terms of geostrophic velocity.

5.3.1 Well-Mixed Boundary Layer

If a convective boundary layer is topped by a stable layer, turbulent mixing can lead to formation of a well-mixed layer. Such boundary layers occur commonly over land during the day when surface heating is strong and over oceans when the air near the sea surface is colder than the surface water temperature. The tropical oceans typically have boundary layers of this type.

In a well-mixed boundary layer, the wind speed and potential temperature are nearly independent of height, as shown schematically in Fig. 5.2, and to a first approximation it is possible to treat the layer as a slab in which the velocity and potential temperature profiles are constant with height and turbulent fluxes vary linearly with height. For simplicity, we assume that the turbulence vanishes at the top of the boundary layer. Observations indicate that the surface momentum flux can be represented by a *bulk aerodynamic formula*[3]

$$\left(\overline{u'w'}\right)_s=-C_d\left|\overline{\mathbf{V}}\right|\bar{u}, \qquad \text{and} \qquad \left(\overline{v'w'}\right)_s=-C_d\left|\overline{\mathbf{V}}\right|\bar{v}$$

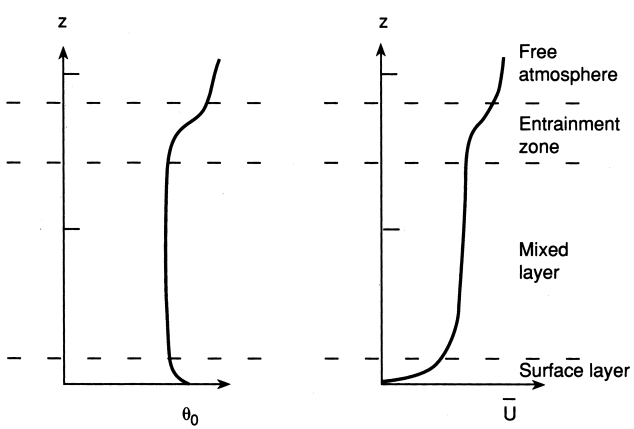


Fig. 5.2 Mean potential temperature, θ_0, and mean zonal wind, U, profiles in a well-mixed boundary layer. Adapted from Stull (1988).

[3] The turbulent momentum flux is often represented in terms of an "eddy stress" by defining, for example, $\tau_{ex}=\rho_o\overline{u'w'}$. We prefer to avoid this terminology to eliminate possible confusion with molecular friction.

where C_d is a nondimensional *drag coefficient*, $|\overline{\mathbf{V}}| = (\overline{u}^2 + \overline{v}^2)^{1/2}$, and the subscript s denotes surface values (referred to the standard anemometer height). Observations show that C_d is of order 1.5×10^{-3} over oceans, but may be several times as large over rough ground.

The approximate planetary boundary layer equations (5.18) and (5.19) can then be integrated from the surface to the top of the boundary layer at $z = h$ to give

$$f\left(\bar{v} - \bar{v}_g\right) = -\left(\overline{u'w'}\right)_s / h = C_d \left|\overline{\mathbf{V}}\right| \bar{u} / h \tag{5.20}$$

$$-f\left(\bar{u} - \bar{u}_g\right) = -\left(\overline{v'w'}\right)_s / h = C_d \left|\overline{\mathbf{V}}\right| \bar{v} / h \tag{5.21}$$

Without loss of generality we can choose axes such that $\overline{v}_g = 0$. Then (5.20) and (5.21) can be rewritten as

$$\bar{v} = \kappa_s \left|\overline{\mathbf{V}}\right| \bar{u}; \qquad \bar{u} = \bar{u}_g - \kappa_s \left|\overline{\mathbf{V}}\right| \bar{v}; \tag{5.22}$$

where $\kappa_s \equiv C_d / (fh)$. Thus, in the mixed layer the wind speed is less than the geostrophic speed and there is a component of motion directed toward lower pressure (i.e., to the left of the geostrophic wind in the Northern Hemisphere and to the right in the Southern Hemisphere) whose magnitude depends on κ_s. For example, if $\bar{u}_g = 10$ m s^{-1} and $\kappa_s = 0.05$ m^{-1} s, then $\bar{u} = 8.28$ m s^{-1}, $\bar{v} = 3.77$ m s^{-1}, and $\left|\overline{\mathbf{V}}\right| = 9.10$ m s^{-1} at all heights within this idealized slab mixed layer.

It is the work done by the flow toward lower pressure that balances the frictional dissipation at the surface. Because boundary layer turbulence tends to reduce wind speeds, the turbulent momentum flux terms are often referred to as boundary layer *friction*. It should be kept in mind, however, that the forces involved are due to turbulence, not molecular viscosity.

Qualitatively, the cross isobar flow in the boundary layer can be understood as a direct result of the three-way balance among the pressure gradient force, the Coriolis force, and turbulent drag:

$$f\mathbf{k} \times \overline{\mathbf{V}} = -\frac{1}{\rho_0}\nabla \bar{p} - \frac{C_d}{h}\left|\overline{\mathbf{V}}\right| \overline{\mathbf{V}} \tag{5.23}$$

This balance is illustrated in Fig. 5.3. Because the Coriolis force is always normal to the velocity and the turbulent drag is a retarding force, their sum can exactly balance the pressure gradient force only if the wind is directed toward lower pressure. Furthermore, it is easy to see that as the turbulent drag becomes increasingly dominant, the cross isobar angle must increase.

5.3.2 The Flux–Gradient Theory

In neutrally or stably stratified boundary layers, the wind speed and direction vary significantly with height. The simple slab model is no longer appropriate; some

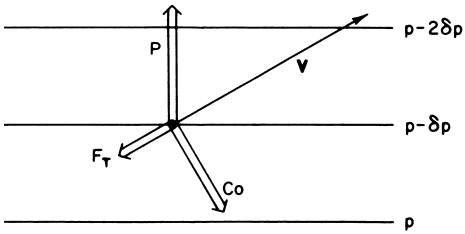


Fig. 5.3 Balance of forces in the well-mixed planetary boundary layer: P designates the pressure gradient force, Co the Coriolis force, and F_T the turbulent drag.

means is needed to determine the vertical dependence of the turbulent momentum flux divergence in terms of mean variables in order to obtain closed equations for the boundary layer variables. The traditional approach to this closure problem is to assume that turbulent eddies act in a manner analogous to molecular diffusion so that the flux of a given field is proportional to the local gradient of the mean. In this case the turbulent flux terms in (5.18) and (5.19) are written as

$$\overline{u'w'} = -K_m \left(\frac{\partial \bar{u}}{\partial z}\right); \ \overline{v'w'} = -K_m \left(\frac{\partial \bar{v}}{\partial z}\right)$$

and the potential temperature flux can be written as

$$\overline{\theta'w'} = -K_h \left(\frac{\partial \bar{\theta}}{\partial z}\right)$$

where $K_m(\mathrm{m^2s^{-1}})$ is the *eddy viscosity* coefficient and $K_h$is the eddy diffusivity of heat. This closure scheme is often referred to as *K theory*.

The K theory has many limitations. Unlike the molecular viscosity coefficient, eddy viscosities depend on the flow rather than the physical properties of the fluid and must be determined empirically for each situation. The simplest models have assumed that the eddy exchange coefficient is constant throughout the flow. This approximation may be adequate for estimating the small-scale diffusion of passive tracers in the free atmosphere. However, it is a very poor approximation in the boundary layer where the scales and intensities of typical turbulent eddies are strongly dependent on the distance to the surface as well as the static stability. Furthermore, in many cases the most energetic eddies have dimensions comparable to the boundary layer depth, and neither the momentum flux nor the heat flux is proportional to the local gradient of the mean. For example, in much of the mixed layer, heat fluxes are positive even though the mean stratification may be very close to neutral.

5.3.3 The Mixing Length Hypothesis

The simplest approach to determining a suitable model for the eddy diffusion coefficient in the boundary layer is based on the mixing length hypothesis introduced by the famous fluid dynamicist L. Prandtl. This hypothesis assumes that a parcel of fluid displaced vertically will carry the mean properties of its original level for a characteristic distance ξ' and then will mix with its surroundings just as an average molecule travels a mean free path before colliding and exchanging momentum with another molecule. By further analogy to the molecular mechanism, this displacement is postulated to create a turbulent fluctuation whose magnitude depends on ξ' and the gradient of the mean property. For example,

$$\theta' = -\xi' \frac{\partial \bar{\theta}}{\partial z};\ u' = -\xi' \frac{\partial \bar{u}}{\partial z};\ v' = -\xi' \frac{\partial \bar{v}}{\partial z}$$

where it must be understood that $\xi' > 0$ for upward parcel displacement and $\xi' < 0$ for downward parcel displacement. For a conservative property, such as potential temperature, this hypothesis is reasonable provided that the eddy scales are small compared to the mean flow scale or that the mean gradient is constant with height. However, the hypothesis is less justified in the case of velocity, as pressure gradient forces may cause substantial changes in the velocity during an eddy displacement.

Nevertheless, if we use the mixing length hypothesis, the vertical turbulent flux of zonal momentum can be written as

$$-\overline{u'w'} = \overline{w'\xi'} \frac{\partial \bar{u}}{\partial z} \tag{5.24}$$

with analogous expressions for the momentum flux in the meridional direction and the potential temperature flux. In order to estimate w' in terms of mean fields, we assume that the vertical stability of the atmosphere is nearly neutral so that buoyancy effects are small. The horizontal scale of the eddies should then be comparable to the vertical scale so that $|w'| \sim |\mathbf{V}'|$ and we can set

$$w' \approx \xi' \left| \frac{\partial \overline{\mathbf{V}}}{\partial z} \right|$$

where $\mathbf{V}'$ and $\overline{\mathbf{V}}$ designate the turbulent and mean parts of the horizontal velocity field, respectively. Here the absolute value of the mean velocity gradient is needed because if $\xi' > 0$, then $w > 0$ (i.e., upward parcel displacements are associated with upward eddy velocities). Thus the momentum flux can be written

$$-\overline{u'w'} = \overline{\xi'^2} \left| \frac{\partial \overline{\mathbf{V}}}{\partial z} \right| \frac{\partial \bar{u}}{\partial z} = K_m \frac{\partial \bar{u}}{\partial z} \tag{5.25}$$

where the eddy viscosity is now defined as $K_m = \overline{\xi'^2}\left|\partial\overline{\mathbf{V}}/\partial z\right| = \overline{\ell^2}\left|\partial\overline{\mathbf{V}}/\partial z\right|$ and the mixing length,

$$\ell \equiv \left(\overline{\xi'^2}\right)^{1/2}$$

is the root mean square parcel displacement, which is a measure of average eddy size. This result suggests that larger eddies and greater shears induce greater turbulent mixing.

5.3.4 The Ekman Layer

If the flux-gradient approximation is used to represent the turbulent momentum flux divergence terms in (5.18) and (5.19), and the value of K_m is taken to be constant, we obtain the equations of the classical Ekman layer:

$$K_m \frac{\partial^2 u}{\partial z^2} + f\left(v - v_g\right) = 0 \tag{5.26}$$

$$K_m \frac{\partial^2 v}{\partial z^2} - f\left(u - u_g\right) = 0 \tag{5.27}$$

where we have omitted the overbars because all fields are Reynolds averaged.

The Ekman layer equations (5.26) and (5.27) can be solved to determine the height dependence of the departure of the wind field in the boundary layer from geostrophic balance. In order to keep the analysis as simple as possible, we assume that these equations apply throughout the depth of the boundary layer. The boundary conditions on u and v then require that both horizontal velocity components vanish at the ground and approach their geostrophic values far from the ground:

$$\begin{aligned} &u = 0,\ v = 0 \text{ at } z = 0 \\ &u \to u_g,\ v \to v_g \text{ as } z \to \infty \end{aligned} \tag{5.28}$$

To solve (5.26) and (5.27), we combine them into a single equation by first multiplying (5.27) by $i = (-1)^{1/2}$ and then adding the result to (5.26) to obtain a second-order equation in the complex velocity, $(u + iv)$:

$$K_m \frac{\partial^2 (u + iv)}{\partial z^2} - if(u + iv) = -if\left(u_g + iv_g\right) \tag{5.29}$$

For simplicity, we assume that the geostrophic wind is independent of height and that the flow is oriented so that the geostrophic wind is in the zonal direction ($v_g = 0$). Then the general solution of (5.29) is

$$(u + iv) = A\exp\left[(if/K_m)^{1/2} z\right] + B\exp\left[-(if/K_m)^{1/2} z\right] + u_g \tag{5.30}$$

It can be shown that $\sqrt{i} = (1+i)/\sqrt{2}$. Using this relationship and applying the boundary conditions of (5.28), we find that for the Northern Hemisphere ($f > 0$), $A = 0$ and $B = -u_g$. Thus

$$u + iv = -u_g \exp\left[-\gamma\,(1+i)\,z\right] + u_g$$

where $\gamma = (f/2K_m)^{1/2}$.

Applying the Euler formula $\exp(-i\theta) = \cos\theta - i\sin\theta$ and separating the real from the imaginary part, we obtain for the Northern Hemisphere

$$u = u_g\left(1 - e^{-\gamma z}\cos\gamma z\right),\ v = u_g e^{-\gamma z}\sin\gamma z \tag{5.31}$$

This solution is the famous Ekman spiral named for the Swedish oceanographer V. W. Ekman, who first derived an analogous solution for the surface wind drift current in the ocean. The structure of the solution (5.31) is best illustrated by a hodograph as shown in Fig. 5.4, where the zonal and meridional components of the wind are plotted as functions of height. The heavy solid curve traced out on Fig. 5.4 connects all the points corresponding to values of u and v in (5.31) for values of γz increasing as one moves away from the origin along the spiral. Arrows show the velocities for various values of γz marked at the arrow points. When $z = \pi/\gamma$, the wind is parallel to and nearly equal to the geostrophic value. It is conventional to designate this level as the top of the Ekman layer and to define the layer depth as $De \equiv \pi/\gamma$.

Observations indicate that the wind approaches its geostrophic value at about 1 km above the surface. Letting $De = 1$ km and $f = 10^{-4}\ \mathrm{s}^{-1}$, the definition of γ implies that $K_m \approx 5\ \mathrm{m}^2\ \mathrm{s}^{-1}$. Referring back to (5.25) we see that for a characteristic boundary layer velocity shear of $|\delta\mathbf{V}/\delta z| \sim 5\times10^{-3}\ \mathrm{s}^{-1}$, this value of K_m implies a mixing length of about 30 m, which is small compared to the depth of the boundary layer, as it should be if the mixing length concept is to be useful.

Qualitatively the most striking feature of the Ekman layer solution is that, like the mixed layer solution of Section 5.3.1, it has a boundary layer wind component

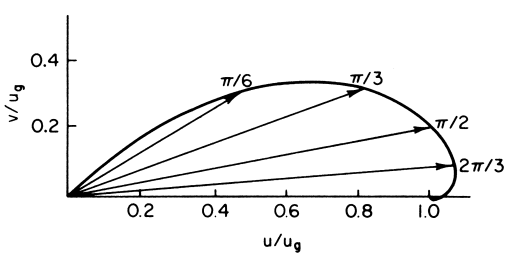


Fig. 5.4 Hodograph of wind components in the Ekman spiral solution. Arrows show velocity vectors for several levels in the Ekman layer, whereas the spiral curve traces out the velocity variation as a function of height. Points labeled on the spiral show the values of γz, which is a nondimensional measure of height.

directed toward lower pressure. As in the mixed layer case, this is a direct result of the three-way balance among the pressure gradient force, the Coriolis force, and the turbulent drag.

The ideal Ekman layer discussed here is rarely, if ever, observed in the atmospheric boundary layer. Partly this is because turbulent momentum fluxes are usually not simply proportional to the gradient of the mean momentum. However, even if the flux–gradient model were correct, it still would not be proper to assume a constant eddy viscosity coefficient, as in reality K_m must vary rapidly with height near the ground. Thus, the Ekman layer solution should not be carried all the way to the surface.

5.3.5 The Surface Layer

Some of the inadequacies of the Ekman layer model can be overcome if we distinguish a *surface layer* from the remainder of the planetary boundary layer. The surface layer, whose depth depends on stability, but is usually less than 10% of the total boundary layer depth, is maintained entirely by vertical momentum transfer by the turbulent eddies; it is not directly dependent on the Coriolis or pressure gradient forces. Analysis is facilitated by supposing that the wind close to the surface is directed parallel to the x axis. The kinematic turbulent momentum flux can then be expressed in terms of a *friction velocity*, u_*, which is defined as[4]

$$u_*^2 \equiv \left|\overline{(u'w')_s}\right|$$

Measurements indicate that the magnitude of the surface momentum flux is of the order 0.1 $\mathrm{m^2\ s^{-2}}$. Thus the friction velocity is typically of the order 0.3 $\mathrm{m\ s^{-1}}$.

According to the scale analysis in Section 2.4, the Coriolis and pressure gradient force terms in (5.16) have magnitudes of about 10^{-3} $\mathrm{m\ s^{-2}}$ in midlatitudes. The momentum flux divergence in the surface layer cannot exceed this magnitude or it would be unbalanced. Thus, it is necessary that

$$\frac{\delta\left(u_*^2\right)}{\delta z} \leq 10^{-3}\ \mathrm{m\ s^{-2}}$$

For $\delta z = 10$ m, this implies that $\delta(u_*^2) \leq 10^{-2}$ $\mathrm{m^2\ s^{-2}}$, so that the change in the vertical momentum flux in the lowest 10 m of the atmosphere is less than 10% of the surface flux. To a first approximation it is then permissible to assume that in the lowest several meters of the atmosphere the turbulent flux remains constant at its surface value, so that with the aid of (5.25)

$$K_m \frac{\partial \bar{u}}{\partial z} = u_*^2 \tag{5.32}$$

[4] Thus the surface eddy stress (see footnote 3) is equal to $\rho_0 u_*^2$.

where we have parameterized the surface momentum flux in terms of the eddy viscosity coefficient. In applying K_m in the Ekman layer solution, we assumed a constant value throughout the boundary layer. Near the surface, however, the vertical eddy scale is limited by the distance to the surface. Thus, a logical choice for the mixing length is $\ell = kz$ where k is a constant. In that case $K_m = (kz)^2 \left|\partial\bar{u} \big/ \partial z\right|$. Substituting this expression into (5.32) and taking the square root of the result gives

$$\partial\bar{u} \big/ \partial z = u_*/(kz) \tag{5.33}$$

Integrating with respect to z yields the *logarithmic wind profile*

$$\bar{u} = \left(u_* \big/ k\right) \ln\left(z \big/ z_0\right) \tag{5.34}$$

where z_0, the *roughness length*, is a constant of integration chosen so that $\bar{u} = 0$ at $z = z_0$. The constant k in (5.34) is a universal constant called *von Karman's constant*, which has an experimentally determined value of $k \approx 0.4$. The roughness length z_0 varies widely depending on the physical characteristics of the surface. For grassy fields, typical values are in the range of 1–4 cm. Although a number of assumptions are required in the derivation of (5.34), many experimental programs have shown that the logarithmic profile provides a generally satisfactory fit to observed wind profiles in the surface layer.

5.3.6 The Modified Ekman Layer

As pointed out earlier, the Ekman layer solution is not applicable in the surface layer. A more satisfactory representation for the planetary boundary layer can be obtained by combining the logarithmic surface layer profile with the Ekman spiral. In this approach the eddy viscosity coefficient is again treated as a constant, but (5.29) is applied only to the region above the surface layer and the velocity and shear at the bottom of the Ekman layer are matched to those at the top of the surface layer. The resulting modified Ekman spiral provides a somewhat better fit to observations than the classical Ekman spiral. However, observed winds in the planetary boundary layer generally deviate substantially from the spiral pattern. Both transience and baroclinic effects (i.e., vertical shear of the geostrophic wind in the boundary layer) may cause deviations from the Ekman solution. However, even in steady-state barotropic situations with near neutral static stability, the Ekman spiral is seldom observed.

It turns out that the Ekman layer wind profile is generally unstable for a neutrally buoyant atmosphere. The circulations that develop as a result of this instability have horizontal and vertical scales comparable to the depth of the boundary layer. Thus, it is not possible to parameterize them by a simple flux–gradient relationship. However, these circulations do in general transport considerable momentum vertically. The net result is usually to decrease the angle between the boundary layer wind

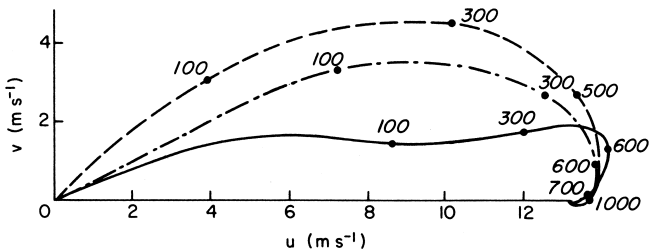


Fig. 5.5 Mean wind hodograph for Jacksonville, Florida ($\cong 30^\circ$ N), April 4, 1968 (solid line) compared with the Ekman spiral (dashed line) and the modified Ekman spiral (dash-dot line) computed with $De \cong 1200$ m. Heights are shown in meters. (Adapted from Brown, 1970. Reproduced with permission of the American Meteorological Society.)

and the geostrophic wind from that characteristic of the Ekman spiral. A typical observed wind hodograph is shown in Fig. 5.5. Although the detailed structure is rather different from the Ekman spiral, the vertically integrated horizontal mass transport in the boundary layer is still directed toward lower pressure. As shown in the next section, this fact is of primary importance for synoptic and larger scale motions.

5.4 SECONDARY CIRCULATIONS AND SPIN DOWN

Both the mixed-layer solution (5.22) and the Ekman spiral solution (5.31) indicate that in the planetary boundary layer the horizontal wind has a component directed toward lower pressure. As suggested by Fig. 5.6, this implies mass convergence in a cyclonic circulation and mass divergence in an anticyclonic circulation, which by mass continuity requires vertical motion out of and into the boundary layer, respectively. In order to estimate the magnitude of this induced vertical motion, we note that if $v_g = 0$ the cross isobaric mass transport per unit area at any level in the boundary layer is given by $\rho_0 v$. The net mass transport for a column of unit width extending vertically through the entire layer is simply the vertical integral of $\rho_0 v$. For the mixed layer, this integral is simply $\rho_0 v h$(kg m^{-1} s^{-1}), where h is the layer depth. For the Ekman spiral, it is given by

$$M = \int_0^{De} \rho_0 v dz = \int_0^{De} \rho_0 u_g \exp(-\pi z/De) \sin(\pi z/De) dz \tag{5.35}$$

where $De = \pi/\gamma$ is the Ekman layer depth defined in Section 5.3.4.

Integrating the mean continuity equation (5.13) through the depth of the boundary layer gives

$$w(De) = -\int_0^{De} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz \tag{5.36}$$

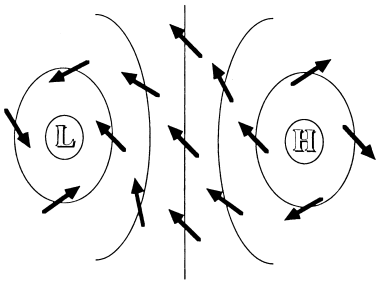


Fig. 5.6 Schematic surface wind pattern (arrows) associated with high- and low-pressure centers in the Northern Hemisphere. Isobars are shown by thin lines, and L and H designate high- and low-pressure centers, respectively. After Stull (1988).

where we have assumed that $w(0) = 0$. Assuming again that $v_g = 0$ so that u_g is independent of x, we find after substituting from (5.31) into (5.36) and comparing with (5.35) that the mass transport at the top of the Ekman layer is given by

$$\rho_0 w(De) = -\frac{\partial M}{\partial y} \tag{5.37}$$

Thus, the mass flux out of the boundary layer is equal to the convergence of the cross isobar mass transport in the layer. Noting that $-\partial u_g/\partial y = \zeta_g$ is just the geostrophic vorticity in this case, we find after integrating (5.35) and substituting into (5.37) that [5]

$$w(De) = \zeta_g \left(\frac{1}{2\gamma}\right) = \zeta_g \left|\frac{K_m}{2f}\right|^{1/2} \left(\frac{f}{|f|}\right) \tag{5.38}$$

where we have neglected the variation of density with height in the boundary layer and have assumed that $1 + e^{-\pi} \approx 1$. Hence, we obtain the important result that the vertical velocity at the top of the boundary layer is proportional to the geostrophic vorticity. In this way the effect of boundary layer fluxes is communicated directly to the free atmosphere through a forced *secondary circulation* that usually dominates over turbulent mixing. This process is often referred to as *boundary layer pumping*. It only occurs in rotating fluids and is one of the fundamental distinctions between rotating and nonrotating flow. For a typical synoptic-scale system with $\zeta_g \sim 10^{-5}\ \mathrm{s}^{-1}$, $f \sim 10^{-4}\ \mathrm{s}^{-1}$, and $De \sim 1$ km, the vertical velocity given by (5.38) is of the order of a few millimeters per second.

An analogous boundary layer pumping is responsible for the decay of the circulation created when a cup of tea is stirred. Away from the bottom and sides of the

[5] The ratio of the Coriolis parameter to its absolute value is included so that the formula will be valid in both hemispheres.

cup there is an approximate balance between the radial pressure gradient and the centrifugal force of the spinning fluid. However, near the bottom, viscosity slows the motion and the centrifugal force is not sufficient to balance the radial pressure gradient. (Note that the radial pressure gradient is independent of depth, as water is an incompressible fluid.) Therefore, radial inflow takes place near the bottom of the cup. Because of this inflow, the tea leaves always are observed to cluster near the center at the bottom of the cup if the tea has been stirred. By continuity of mass, the radial inflow in the bottom boundary layer requires upward motion and a slow compensating outward radial flow throughout the remaining depth of the tea. This slow outward radial flow approximately conserves angular momentum, and by replacing high angular momentum fluid by low angular momentum fluid serves to spin down the vorticity in the cup far more rapidly than could mere diffusion.

The characteristic time for the secondary circulation to spin down an atmospheric vortex is illustrated most easily in the case of a barotropic atmosphere. For synoptic-scale motions the barotropic vorticity equation (4.24) can be written approximately as

$$\frac{D\zeta_g}{Dt} = -f\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = f\frac{\partial w}{\partial z} \tag{5.39}$$

where we have neglected ζ_g compared to f in the divergence term and have also neglected the latitudinal variation of f. Recalling that the geostrophic vorticity is independent of height in a barotropic atmosphere, (5.39) can be integrated easily from the top of the Ekman layer ($z = De$) to the tropopause ($z = H$) to give

$$\frac{D\zeta_g}{Dt} = +f\left[\frac{w(H) - w(De)}{(H - De)}\right] \tag{5.40}$$

Substituting for $w(De)$ from (5.38), assuming that $w(H) = 0$ and that $H \gg De$, (5.40) may be written as

$$\frac{D\zeta_g}{Dt} = -\left|\frac{fK_m}{2H^2}\right|^{1/2}\zeta_g \tag{5.41}$$

This equation may be integrated in time to give

$$\zeta_g(t) = \zeta_g(0)\exp(-t/\tau_e) \tag{5.42}$$

where $\zeta_g(0)$ is the value of the geostrophic vorticity at time $t = 0$, and $\tau_e \equiv H|2/(fK_m)|^{1/2}$ is the time that it takes the vorticity to decrease to e^{-1} of its original value.

This *e-folding* time scale is referred to as the barotropic *spin-down time*. Taking typical values of the parameters as follows: $H \equiv 10\,\text{km}$, $f = 10^{-4}\ \text{s}^{-1}$, and $K_m = 10\ \text{m}^2\ \text{s}^{-1}$, we find that $\tau_e \approx 4$ days. Thus, for midlatitude synoptic-scale disturbances in a barotropic atmosphere, the characteristic spin-down time is a few days. This decay time scale should be compared to the time scale for ordinary

viscous diffusion. For viscous diffusion the time scale can be estimated from scale analysis of the diffusion equation

$$\frac{\partial u}{\partial t} = K_m \frac{\partial^2 u}{\partial z^2} \tag{5.43}$$

If τ_d is the diffusive time scale and H is a characteristic vertical scale for diffusion, then from the diffusion equation

$$U \big/ \tau_d \sim K_m U \Big/ H^2$$

so that $\tau_d \sim H^2/K_m$. For the above values of H and K_m, the diffusion time scale is thus about 100 days. Hence, in the absence of convective clouds the spin-down process is a far more effective mechanism for destroying vorticity in a rotating atmosphere than eddy diffusion. Cumulonimbus convection can produce rapid turbulent transports of heat and momentum through the entire depth of the troposphere. These must be considered together with boundary layer pumping for intense systems such as hurricanes.

Physically the spin-down process in the atmospheric case is similar to that described for the teacup, except that in synoptic-scale systems it is primarily the Coriolis force that balances the pressure gradient force away from the boundary, not the centrifugal force. Again the role of the secondary circulation driven by forces resulting from boundary layer drag is to provide a slow radial flow in the interior that is superposed on the azimuthal circulation of the vortex above the boundary layer. This secondary circulation is directed outward in a cyclone so that the horizontal area enclosed by any chain of fluid particles gradually increases. Since the circulation is conserved, the azimuthal velocity at any distance from the vortex center must decrease in time or, from another point of view, the Coriolis force for the outward-flowing fluid is directed clockwise, and this force thus exerts a torque opposite to the direction of the circulation of the vortex. Fig. 5.7 shows a qualitative sketch of the streamlines of this secondary flow.

It should now be obvious exactly what is meant by the term *secondary circulation*. It is simply a circulation superposed on the primary circulation (in this case the azimuthal circulation of the vortex) by the physical constraints of the system. In the case of the boundary layer, viscosity is responsible for the presence of the secondary circulation. However, other processes, such as temperature advection and diabatic heating, may also lead to secondary circulations, as shown later.

The above discussion has concerned only the neutrally stratified barotropic atmosphere. An analysis for the more realistic case of a stably stratified baroclinic atmosphere is more complicated. However, qualitatively the effects of stratification may be easily understood. The buoyancy force (see Section 2.7.3) will act to suppress vertical motion, as air lifted vertically in a stable environment will be

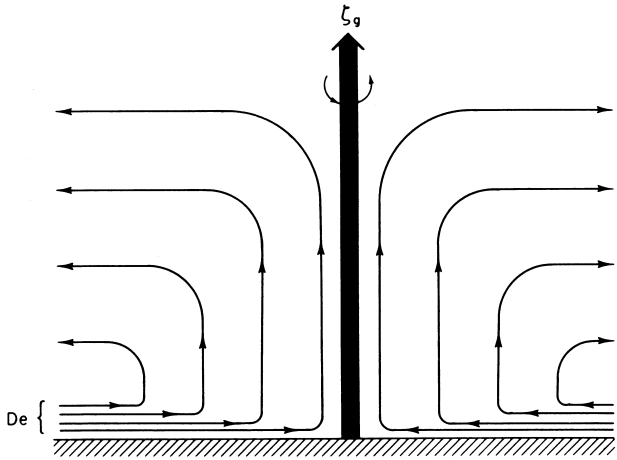


Fig. 5.7 Streamlines of the secondary circulation forced by frictional convergence in the planetary boundary layer for a cyclonic vortex in a barotropic atmosphere. The circulation extends throughout the full depth of the vortex.

denser than the environmental air. As a result the interior secondary circulation will decrease with altitude at a rate proportional to the static stability. This vertically varying secondary flow, shown in Fig. 5.8, will rather quickly spin down the vorticity at the top of the Ekman layer without appreciably affecting the higher levels. When the geostrophic vorticity at the top of the boundary layer is reduced to zero,

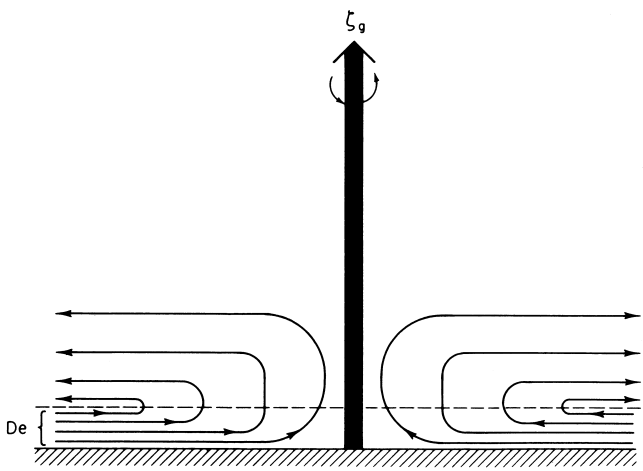


Fig. 5.8 Streamlines of the secondary circulation forced by frictional convergence in the planetary boundary layer for a cyclonic vortex in a stably stratified baroclinic atmosphere. The circulation decays with height in the interior.

the pumping action of the Ekman layer is eliminated. The result is a baroclinic vortex with a vertical shear of the azimuthal velocity that is just strong enough to bring ζ_g to zero at the top of the boundary layer. This vertical shear of the geostrophic wind requires a radial temperature gradient that is in fact produced during the spin-down phase by adiabatic cooling of the air forced out of the Ekman layer. Thus, the secondary circulation in the baroclinic atmosphere serves two purposes: (1) it changes the azimuthal velocity field of the vortex through the action of the Coriolis force and (2) it changes the temperature distribution so that a thermal wind balance is always maintained between the vertical shear of the azimuthal velocity and the radial temperature gradient.

PROBLEMS

5.1. Verify by direct substitution that the Ekman spiral expression (5.31) is indeed a solution of the boundary layer equations (5.26) and (5.27) for the case $v_g = 0$.

5.2. Derive the Ekman spiral solution for the more general case where the geostrophic wind has both x and y components (u_g and v_g, respectively), which are independent of height.

5.3. Letting the Coriolis parameter and density be constants, show that (5.38) is correct for the more general Ekman spiral solution obtained in Problem 5.2.

5.4. For laminar flow in a rotating cylindrical vessel filled with water (molecular kinematic viscosity $\nu = 0.01\ \mathrm{cm^2\ s^{-1}}$), compute the depth of the Ekman layer and the spin-down time if the depth of the fluid is 30 cm and the rotation rate of the tank is 10 revolutions per minute. How small would the radius of the tank have to be in order that the time scale for viscous diffusion from the side walls be comparable to the spin-down time?

5.5. Suppose that at 43° N the geostrophic wind is westerly at $15\ \mathrm{m\ s^{-1}}$. Compute the net cross isobaric transport in the planetary boundary layer using both the mixed layer solution (5.22) and the Ekman layer solution (5.31). You may let $|\mathbf{V}| = u_g$ in (5.22), $h = De = 1$ km, $\kappa_s = 0.05\ \mathrm{m^{-1}s}$, and $\rho = 1\ \mathrm{kg\ m^{-3}}$.

5.6. Derive an expression for the wind-driven surface Ekman layer in the ocean. Assume that the wind stress τ_w is constant and directed along the x axis. The continuity of turbulent momentum flux at the air–sea interface ($z = 0$) requires that the wind stress divided by air density must equal the oceanic turbulent momentum flux at $z = 0$. Thus, if the flux–gradient theory is used, the boundary condition at the surface becomes

$$\rho_0 K \frac{\partial u}{\partial z} = \tau_w, \quad \rho_0 K \frac{\partial v}{\partial z} = 0, \quad \text{at } z = 0$$

where K is the eddy viscosity in the ocean (assumed constant). As a lower boundary condition assume that $u, v \to 0$ as $z \to -\infty$. If $K = 10^{-3}\ \mathrm{m^2\ s^{-1}}$, what is the depth of the surface Ekman layer at 43°N latitude?

5.7. Show that the vertically integrated mass transport in the wind-driven oceanic surface Ekman layer is directed 90° to the right of the surface wind stress in the Northern Hemisphere. Explain this result physically.

5.8. A homogeneous barotropic ocean of depth H = 3 km has a zonally symmetric geostrophic jet whose profile is given by the expression

$$u_g = U \exp\left[-(y/L)^2\right]$$

where $U = 1\mathrm{m\ s^{-1}}$ and $L = 200$ km are constants. Compute the vertical velocity produced by convergence in the Ekman layer at the ocean bottom and show that the meridional profile of the secondary cross-stream motion forced in the interior is the same as the meridional profile of u_g. What are the maximum values of $\bar{v}$ in the interior and $\bar{w}$ if $K = 10^{-3}\mathrm{m^2\ s^{-1}}$ and $f = 10^{-4}\ \mathrm{s^{-1}}$ (Assume that w and the eddy stress vanish at the surface.)

5.9. Using the approximate zonally averaged momentum equation

$$\frac{\partial \bar{u}}{\partial t} \cong f\bar{v}$$

compute the spin-down time for the zonal jet in Problem 5.8.

5.10. Derive a formula for the vertical velocity at the top of the planetary boundary layer using the mixed layer expression (5.22). Assume that $|\overline{\mathrm{V}}| = 5\ \mathrm{m\ s^{-1}}$ is independent of x and y and that $\bar{u}_g = \bar{u}_g(y)$. If $h = 1$ km and $\kappa_s = 0.05$, what value must K_m have if this result is to agree with the vertical velocity derived from the Ekman layer solution at 43° latitude with $De = 1$ km?

5.11. Show that $K_m = kzu_*$in the surface layer.

MATLAB EXERCISES

M5.1. The MATLAB script **mixed_layer_wind1.m** uses a simple iterative technique to solve for u and v in (5.22) with u_g in the range 1–20 m s^{-1} for the case $v_g = 0$ and $\kappa_s = 0.05\ \mathrm{m^{-1}}$ s. If you run the script, you will observe that this iterative technique fails for u_g greater than 19 m s^{-1}. An alternative method, which works for a wide range of specified geostrophic winds, utilizes the natural coordinate system introduced in Section 3.2.1.

(a) Show that in the natural coordinate system the force balances parallel and perpendicular to the velocity vector in the mixed layer model (see Fig. 5.3) can be expressed respectively as

$$f\kappa_s V^2 = f u_g \cos\beta, \quad \text{and} \quad fV = f u_g \sin\beta$$

where it is here assumed that the pressure gradient force is directed northward so that $f u_g = \left|\rho_0^{-1}\nabla p\right|$, and β designates the angle between the pressure gradient force vector and the mixed layer velocity, $\boldsymbol{V}$. (Other notation as defined in Section 5.3.1.)
(b) Use MATLAB to solve for V and β given u_g in the range of 1 to 50 m s^{-1} and plot V and β versus u_g. *Hint*: Solve for V in the above two equations and for each value of u_g vary β until the two solutions for V agree.

M5.2. Suppose that the geopotential distribution at the top of the mixed layer can be expressed in the form $\Phi(x, y) = \Phi_0 - f_0 U_0 y + A \sin kx \sin ly$, where $\Phi_0 = 9800\,\mathrm{m^2\,s^{-2}}$, $f_0 = 10^{-4}\,\mathrm{s^{-1}}$, $U_0 = 5\,\mathrm{m\,s^{-1}}$, $A = 1500\,\mathrm{m^2\,s^{-2}}$, $k = \pi L^{-1}$, and, $l = \pi L^{-1}$, where $L = 6000$ km. (a) Use the technique given in the demonstration script **mixed_layer_wind1.m** to determine the wind distribution in the mixed layer for this situation for the case where $\kappa_s = 0.05$. (b) Using the formula obtained in Problem 5.10, compute the vertical velocity distribution at the top of the mixed layer for this distribution of geopotential when the depth of the mixed layer is 1 km. (The MATLAB script **mixed_layer_wind_2.m** has a template that you can use to contour the vertical velocity field and vorticity fields.)

M5.3. For the geopotential distribution of Problem M5.2, use the Ekman layer theory to derive the pattern of vertical velocity at the top of the boundary layer. Assume that $K_m = 10\,\mathrm{m^2\,s^{-1}}$ and again use MATLAB to contour the fields of vorticity and vertical velocity. Explain why the vertical velocity patterns derived from the mixed layer and Ekman theories differ for this situation.

Suggested References

Arya, *Introduction to Micrometeorology,* gives an excellent introduction to boundary layer dynamics and turbulence at the beginning undergraduate level.

Garratt, *The Atmospheric Boundary Layer* is an excellent graduate-level introduction to the physics of the atmospheric boundary layer.

Stull, *An Introduction to Boundary Layer Meteorology*, provides a comprehensive and very nicely illustrated treatment of all aspects of the subject at the beginning graduate level.

CHAPTER 6

Synoptic-Scale Motions I: Quasi-Geostrophic Analysis

A primary goal of dynamic meteorology is to interpret the observed structure of large-scale atmospheric motions in terms of the physical laws governing the motions. These laws, which express the conservation of momentum, mass, and energy, completely determine the relationships among the pressure, temperature, and velocity fields. As we saw in Chapter 2, these governing laws are quite complicated even when the hydrostatic approximation (which is valid for all large-scale meteorological systems) is applied. For extratropical synoptic-scale motions, however, the horizontal velocities are approximately geostrophic (see Section 2.4). Such motions, which are usually referred to as *quasi-geostrophic*, are simpler to analyze than are tropical disturbances or planetary scale disturbances. They are also the major systems of interest in traditional short-range weather forecasting, and are thus a reasonable starting point for dynamical analysis.

This chapter shows that for extratropical synoptic-scale systems, the twin requirements of hydrostatic and geostrophic balance constrain the baroclinic motions so that to a good approximation the structure and evolution of the three-dimensional velocity field are determined by the distribution of geopotential height on isobaric surfaces. The equations that express these relationships constitute the quasi-geostrophic system. Before developing this system of equations, it is useful to

briefly summarize the observed structure of midlatitude synoptic systems and the mean circulations in which they are embedded. We then develop the quasi-geostrophic momentum and thermodynamic energy equations and show how these can be manipulated to form the *quasi-geostrophic potential vorticity equation* and the *omega equation*. The former equation provides a method for predicting the evolution of the geopotential field, given its initial three-dimensional distribution; the latter provides a method for diagnosing the vertical motion from the known distribution of geopotential. In both cases, alternative versions of the equations are discussed to help elucidate the dynamical processes responsible for the development and evolution of synoptic-scale systems.

6.1 THE OBSERVED STRUCTURE OF EXTRATROPICAL CIRCULATIONS

Atmospheric circulation systems depicted on a synoptic chart rarely resemble the simple circular vortices discussed in Chapter 3. Rather, they are generally highly asymmetric in form, with the strongest winds and largest temperature gradients concentrated along narrow bands called *fronts*. Also, such systems generally are highly baroclinic; the amplitudes and phases of the geopotential and velocity perturbations both change substantially with height. Part of this complexity is due to the fact that these synoptic systems are not superposed on a uniform mean flow, but are embedded in a slowly varying planetary scale flow that is itself highly baroclinic. Furthermore, this planetary scale flow is influenced by *orography* (i.e., by large-scale terrain variations) and continent-ocean heating contrasts so that it is highly longitude dependent. Therefore, it is not accurate to view synoptic systems as disturbances superposed on a zonal flow that varies only with latitude and height. As shown in Chapter 8, however, such a point of view can be useful as a first approximation in theoretical analyses of synoptic-scale wave disturbances.

Zonally averaged cross sections do provide some useful information on the gross structure of the planetary scale circulation in which synoptic-scale eddies are embedded. Figure 6.1 shows meridional cross sections of the longitudinally averaged zonal wind and temperature for the solstice seasons of (a) December, January and February (DJF) and (b) June, July and August (JJA). These sections extend from approximately sea level (1000 hPa) to about 32 km altitude (10 hPa). Thus the troposphere and lower stratosphere are shown. This chapter is concerned with the structure of the wind and temperature fields in the troposphere. The stratosphere is discussed in Chapter 12.

The average pole to equator temperature gradient in the Northern Hemisphere troposphere is much larger in winter than in summer. In the Southern Hemisphere the difference between summer and winter temperature distributions is smaller, due mainly to the large thermal inertia of the oceans, together with the greater fraction of the surface that is covered by oceans in the Southern Hemisphere. Since the mean zonal wind and temperature fields satisfy the thermal wind relationship

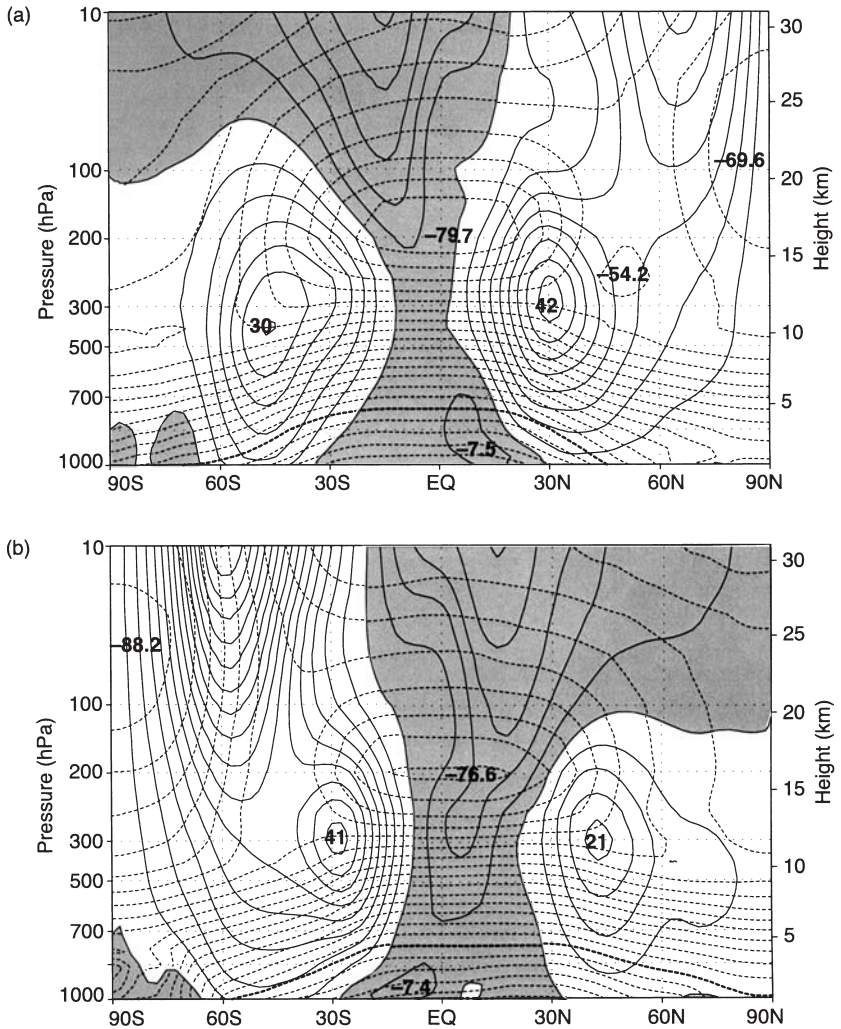


Fig. 6.1 Meridional cross sections of longitudinally and time-averaged zonal wind (solid contours, interval of m s^{-1}) and temperature (dashed contours, interval of 5 K) for December–February (a) and June–August (b). Easterly winds are shaded and 0° C isotherm is darkened. Wind maxima shown in m s^{-1}, temperature minima shown in °C. (Based on NCEP/NCAR reanalyses; after Wallace, 2003.)

(3.30) to a high degree of accuracy, the seasonal cycle in zonal wind speeds is similar to that of the meridional temperature gradient. In the Northern Hemisphere the maximum zonal wind speed in the winter is twice as large as in the summer, whereas in the Southern Hemisphere the difference between winter and summer

zonal wind maxima is much smaller. Furthermore, in both seasons the core of maximum zonal wind speed (called the mean *jetstream* axis) is located just below the *tropopause* (the boundary between the troposphere and stratosphere) at the latitude where the thermal wind integrated through the troposphere is a maximum. In both hemispheres, this is about 30° during winter, but moves poleward to 40°–45° during summer.

That the zonally averaged meridional cross sections of Fig. 6.1 are not representative of the mean wind structure at all longitudes can be seen in Fig. 6.2, which shows the distribution of the time-averaged zonal wind component for DJF on the 200-hPa surface in the Northern Hemisphere. It is clear from Fig. 6.2 that at some longitudes there are very large deviations of the time-mean zonal flow from its longitudinally averaged distribution. In particular there are strong zonal wind maxima (jets) near 30°N just east of the Asian and North American continents and north of the Arabian peninsula; distinct minima occur in the eastern Pacific and eastern Atlantic. Synoptic-scale disturbances tend to develop preferentially in the regions of maximum time-mean zonal winds associated with the western Pacific and western Atlantic jets and to propagate downstream along *storm tracks* that approximately follow the jet axes.

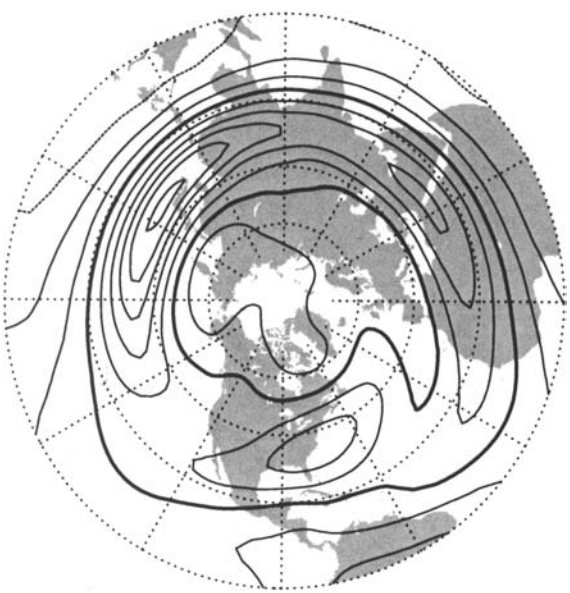


Fig. 6.2 Mean zonal wind at the 200-hPa level for December–February averaged for years 1958–1997. Contour interval 10 m s^{-1} (heavy contour, 20 m s^{-1}). (Based on NCEP/NCAR reanalyses; after Wallace, 2003.)

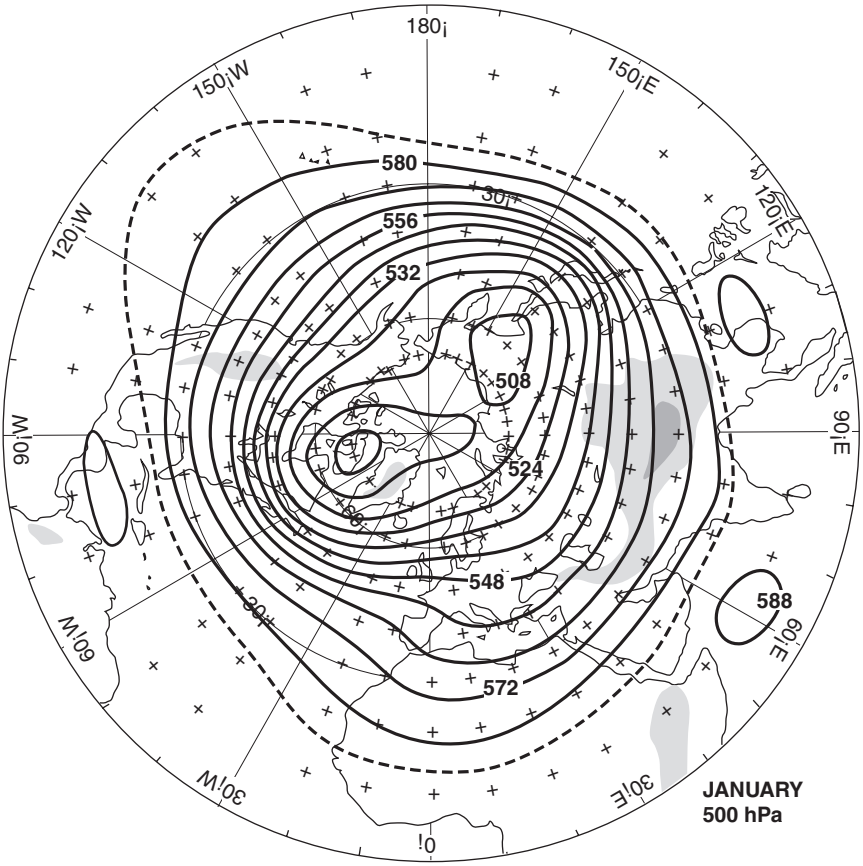


Fig. 6.3 Mean 500-hPa contours in January in the Northern Hemisphere. Heights shown in tens of meters. (After Palmén and Newton, 1969.)

The large departure of the northern winter climatological jetstream from zonal symmetry can also be readily inferred from examination of Fig. 6.3, which shows the mean 500-hPa geopotential contours for January in the Northern Hemisphere. Even after averaging the height field for a month, very striking departures from zonal symmetry remain. These are clearly linked to the distributions of continents and oceans. The most prominent asymmetries are the troughs to the east of the American and Asian continents. Referring back to Fig. 6.2, we see that the intense jet at 35˚N and 140˚E is a result of the semipermanent trough in that region. Thus, it is apparent that the mean flow in which synoptic systems are embedded should really be regarded as a longitude-dependent time-averaged flow.

In addition to its longitudinal dependence, the planetary scale flow also varies from day to day due to its interactions with transient synoptic-scale disturbances.

In fact, observations show that the transient planetary scale flow amplitude is comparable to that of the time-mean. As a result, monthly mean charts tend to smooth out the actual structure of the instantaneous jetstream since the position and intensity of the jet vary. Thus, at any time the planetary scale flow in the region of the jetstream has much greater baroclinicity than indicated on time-averaged charts. This point is illustrated schematically in Fig. 6.4, which shows a latitude–height cross section through an observed jetstream over North America. Fig. 6.4a shows the zonal wind and potential temperature, whereas Fig. 6.4b shows the potential temperature and Ertel potential vorticity. The 2 PVU contour of potential vorticity approximately marks the tropopause.

As illustrated in Fig. 6.4a, the axis of the jetstream tends to be located above a narrow sloping zone of strong potential temperature gradients called the *polar-frontal* zone. This is the zone that in general separates the cold air of polar origin from warm tropical air. The occurrence of an intense jet core above this zone of large magnitude potential temperature gradients is, of course, not mere coincidence, but rather a consequence of the thermal wind balance.

The potential temperature contours in Fig. 6.4 illustrate the strong static stability in the stratosphere. They also illustrate the fact that isentropes (constant θ surfaces) cross the tropopause in the vicinity of the jet so that air can move between the troposphere and the stratosphere without diabatic heating or cooling. The strong gradient of Ertel potential vorticity at the tropopause, however, provides a strong resistance to cross-tropopause flow along the isentropes. Note, however, that in the frontal region the potential vorticity surfaces are displaced substantially downward so that the frontal zone is characterized by a strong positive anomaly in potential vorticity associated with the strong relative vorticity on the poleward side of the jet and the strong static stability on the cold air side of the frontal zone.

It is a common observation in fluid dynamics that jets in which strong velocity shears occur may be unstable with respect to small perturbations. By this is meant that any small disturbance introduced into the jet will tend to amplify, drawing energy from the jet as it grows. Most synoptic-scale systems in midlatitudes appear to develop as the result of an instability of the jetstream flow. This instability, called *baroclinic instability*, depends on the meridional temperature gradient, particularly at the surface. Hence, through the thermal wind relationship, baroclinic instability depends on vertical shear and tends to occur in the region of the polar frontal zone.

Baroclinic instability is not, however, identical to frontal instability, as most baroclinic instability models describe only geostrophically scaled motions, whereas disturbances in the vicinity of strong frontal zones must be highly ageostrophic. As shown in Chapter 9, baroclinic disturbances may themselves act to intensify preexisting temperature gradients and hence generate frontal zones.

The stages in the development of a typical baroclinic cyclone that develops as a result of baroclinic instability are shown schematically in Fig. 6.5. In the stage of rapid development there is a cooperative interaction between the upper

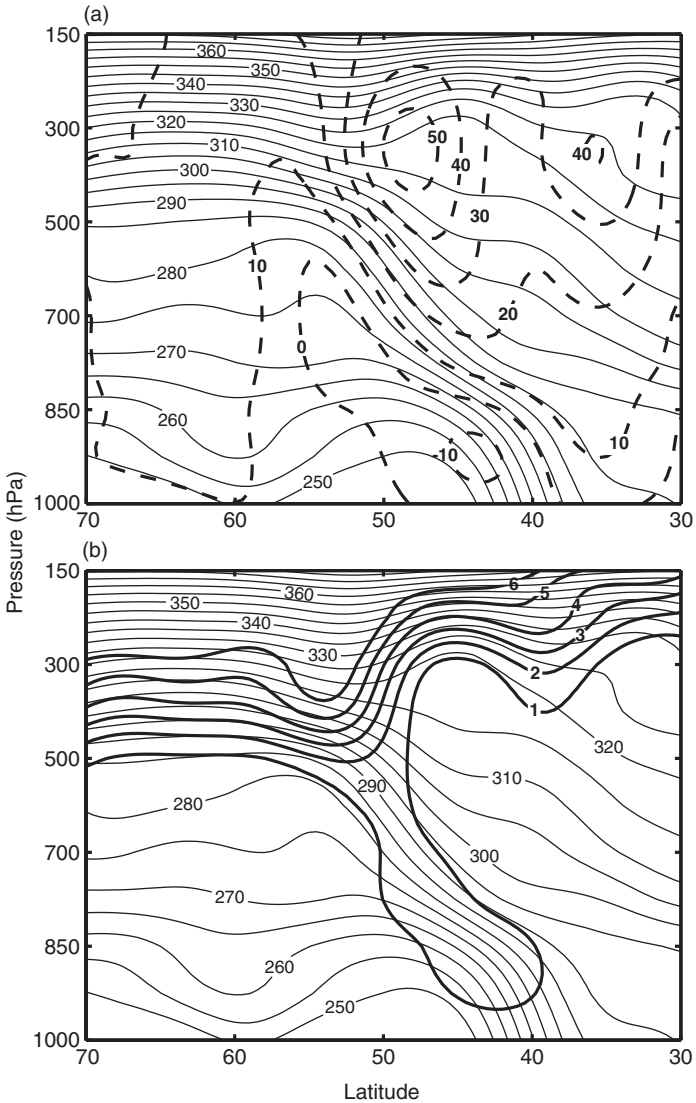


Fig. 6.4 Latitude–height cross sections through a cold front at 80W longitude on 00 UTC January 14, 1999. (a) Potential temperature contours (thin solid lines, K) and zonal wind isotachs (dashed lines, m s^{-1}). (b) Thin solid contours as in (a), heavy dashed contours, Ertel potential vorticity labeled in PVU (1PVU = 10^{-6} K kg^{-1} m^2 s^{-1}).

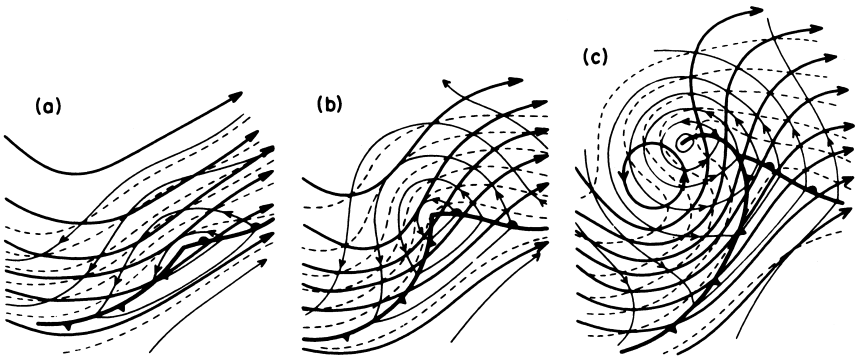


Fig. 6.5 Schematic 500-hPa contours (heavy solid lines), 1000-hPa contours (thin lines), and 1000–500 hPa thickness (dashed) for a developing baroclinic wave at three stages of development. (After Palmén and Newton, 1969.)

level and surface flows; strong cold advection is seen to occur west of the trough at the surface, with weaker warm advection to the east. This pattern of thermal advection is a direct consequence of the fact that the trough at 500 hPa lags (lies to the west of) the surface trough so that the mean geostrophic wind in the 1000- to 500-hPa layer is directed across the 1000- to 500-hPa thickness lines toward larger thickness west of the surface trough and toward smaller thickness east of the surface trough. This dependence of the phase of the disturbance on height is better illustrated by Fig. 6.6, which shows a schematic downstream (or west–east) cross section through an idealized developing baroclinic system. Throughout the troposphere the trough and ridge axes tilt westward (or upstream) with height,[1] whereas the axes of warmest and coldest air are observed to have the opposite tilt. As shown later, the westward tilt of the troughs and ridges is necessary in order that the mean flow transfers potential energy to the developing wave. In the mature stage (not shown in Fig. 6.5) the troughs at 500 and 1000 hPa are nearly in phase. As a consequence, the thermal advection and energy conversion are quite weak.

6.2 THE QUASI-GEOSTROPHIC APPROXIMATION

The main goal of this chapter is to show how the observed structure of midlatitude synoptic systems can be interpreted in terms of the constraints imposed on synoptic-scale motions by the dynamical equations. Specifically we show that for motions that are hydrostatic and nearly geostrophic, the three-dimensional flow field is determined approximately by the isobaric distribution of geopotential [$\Phi(x, y, p, t)$]. For this analysis, it is convenient to use the isobaric coordinate system both

[1] In reality, the phase tilts tend to be concentrated below the 700-hPa level.

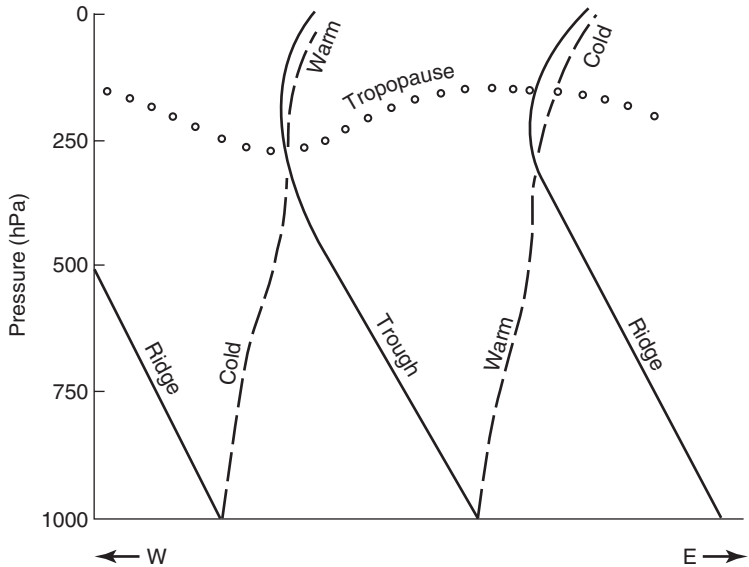


Fig. 6.6 West–east cross section through a developing baroclinic wave. Solid lines are trough and ridge axes; dashed lines are axes of temperature extrema; the chain of open circles denotes the tropopause.

because meteorological measurements are generally referred to constant pressure surfaces and because the dynamical equations are somewhat simpler in isobaric coordinates than in height coordinates. Thus, use of the isobaric coordinate system simplifies the development of approximate prognostic and diagnostic equations.

6.2.1 Scale Analysis in Isobaric Coordinates

The dynamical equations in isobaric coordinates were developed in Section 3.1 and, for reference, are repeated here. The horizontal momentum equation, the hydrostatic equation, the continuity equation, and the thermodynamic energy equation may be expressed as

$$\frac{D\mathbf{V}}{Dt} + f\mathbf{k} \times \mathbf{V} = -\nabla\Phi \tag{6.1}$$

$$\frac{\partial\Phi}{\partial p} = -\alpha = -RT/p \tag{6.2}$$

$$\nabla \bullet \mathbf{V} + \frac{\partial\omega}{\partial p} = 0 \tag{6.3}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla\right) T - S_p\omega = J/c_p \tag{6.4}$$

Here the total derivative in (6.1) is defined by

$$\frac{D}{Dt} \equiv \left(\frac{\partial}{\partial t}\right)_p + (\mathbf{V} \bullet \boldsymbol{\nabla})_p + \omega \frac{\partial}{\partial p} \tag{6.5}$$

where $\omega \equiv \mathrm{D}p/Dt$ is the pressure change following the motion, and in (6.4) $S_p \equiv -T\partial \ln\theta \big/ \partial p$ is the static stability parameter [$S_p \approx 5 \times 10^{-4}$ K Pa^{-1} in the midtroposphere].

These equations, although simplified by use of the hydrostatic approximation and by neglect of some small terms that appear in the complete spherical coordinate form, still contain several terms that are of secondary significance for midlatitude synoptic-scale systems. They can be simplified further by recalling from Section 2.4 that the horizontal flow is nearly geostrophic and that the magnitude of the ratio of vertical velocity to horizontal velocity is of order 10^{-3}.

We first separate the horizontal velocity into geostrophic and ageostrophic parts by letting

$$\mathbf{V} = \mathbf{V}_g + \mathbf{V}_a \tag{6.6}$$

where the geostrophic wind is defined as

$$\mathbf{V}_g \equiv f_0^{-1}\mathbf{k} \times \boldsymbol{\nabla}\Phi \tag{6.7}$$

and the ageostrophic wind, $\mathbf{V}_a$, is just the difference between the total horizontal wind and the geostrophic wind. We have here assumed that the meridional length scale, L, is small compared to the radius of the earth so that the geostrophic wind (6.7) may be defined using a constant reference latitude value of the Coriolis parameter.[2]

For the systems of interest $|\mathbf{V}_g| \gg |\mathbf{V}_a|$. More precisely,

$$|\mathbf{V}_a| \big/ |\mathbf{V}_g| \sim \mathrm{O}\,(\mathrm{Ro})$$

That is, the ratio of the magnitudes of the ageostrophic and geostrophic winds is the same order of magnitude as the Rossby number introduced in Section 2.4.2.

The momentum can then be approximated to O(Ro) by its geostrophic value, and the rate of change of momentum or temperature following the horizontal motion can be approximated to the same order by the rate of change following the geostrophic wind. Thus, in (6.5) $\mathbf{V}$ can be replaced by $\mathbf{V}_g$ and the vertical advection, which arises only from the ageostrophic flow, can be neglected. The rate of change

[2] This definition of the geostrophic wind will be referred to as *constant-f* (CF) geostrophy, whereas the definition given in (3.4) will be called *variable-f* (VF) geostrophy. The CF geostrophic wind is nondivergent, whereas the VF geostrophic wind has a divergent portion (see Problem 3.19). The interpretation of the ageostrophic wind depends strongly on which type of geostrophy is used, as explained in Blackburn (1985).

of momentum following the total motion is then approximately equal to the rate of change of the geostrophic momentum following the geostrophic wind:

$$\frac{D\mathbf{V}}{Dt} \approx \frac{D_g\mathbf{V}_g}{Dt}$$

where

$$\frac{D_g}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{V}_g \bullet \nabla = \frac{\partial}{\partial t} + u_g\frac{\partial}{\partial x} + v_g\frac{\partial}{\partial y} \tag{6.8}$$

Although a constant f_0 can be used in defining $\mathbf{V}_g$, it is still necessary to retain the dynamical effect of the variation of the Coriolis parameter with latitude in the Coriolis force term in the momentum equation. This variation can be approximated by expanding the latitudinal dependence of f in a Taylor series about a reference latitude ϕ_0 and retaining only the first two terms to yield

$$f = f_0 + \beta y \tag{6.9}$$

where $\beta \equiv (df/dy)_{\phi_0} = 2\Omega\cos\phi_0/a$ and $y = 0$ at ϕ_0. This approximation is usually referred to as the *midlatitude β-plane* approximation. For synoptic-scale motions, the ratio of the first two terms in the expansion of f has an order of magnitude

$$\frac{\beta L}{f_0} \sim \frac{\cos\phi_0}{\sin\phi_0}\frac{L}{a} \sim \mathrm{O}\,(\mathrm{Ro}) \ll 1$$

This justifies letting the Coriolis parameter have a constant value f_0 in the geostrophic approximation and approximating its variation in the Coriolis force term by (6.9).

From (6.1) the acceleration following the motion is equal to the difference between the Coriolis force and the pressure gradient force. This difference depends on the departure of the actual wind from the geostrophic wind. Thus, it is not permissible to simply replace the horizontal velocity by its geostrophic value in the Coriolis term. Rather, we use (6.6), (6.7), and (6.9) to write

$$\begin{aligned} f\mathbf{k}\times\mathbf{V} + \nabla\Phi &= (f_0 + \beta y)\,\mathbf{k}\times\left(\mathbf{V}_g + \mathbf{V}_a\right) - f_0\mathbf{k}\times\mathbf{V}_g \\ &\approx f_0\mathbf{k}\times\mathbf{V}_a + \beta y\mathbf{k}\times\mathbf{V}_g \end{aligned} \tag{6.10}$$

where we have used the geostrophic relation (6.7) to eliminate the pressure gradient force and neglected the ageostrophic wind compared to the geostrophic wind in the term proportional to βy. The approximate horizontal momentum equation thus has the form

$$\frac{D_g\mathbf{V}_g}{Dt} = -f_0\mathbf{k}\times\mathbf{V}_a - \beta y\mathbf{k}\times\mathbf{V}_g \tag{6.11}$$

Each term in (6.11) is thus O(Ro) compared to the pressure gradient force, whereas terms neglected are O(Ro^2) or smaller.

The geostrophic wind defined in (6.7) is nondivergent. Thus,

$$\nabla \bullet \mathbf{V} = \nabla \bullet \mathbf{V}_a = \frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y}$$

and the continuity equation (6.3) can be rewritten as

$$\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y} + \frac{\partial \omega}{\partial p} = 0 \tag{6.12}$$

which shows that if the geostrophic wind is defined by (6.7), ω is determined only by the ageostrophic part of the wind field.

In the thermodynamic energy equation (6.4), the horizontal advection can be approximated by its geostrophic value. However, the vertical advection is not neglected, but forms part of the adiabatic heating and cooling term. This term must be retained because the static stability is usually large enough on the synoptic scale so that the adiabatic heating or cooling due to vertical motion is of the same order as the horizontal temperature advection, despite the smallness of the vertical velocity. The adiabatic heating and cooling term can be somewhat simplified by dividing the total temperature field, T_{tot}, into a basic state (standard atmosphere) portion that depends only on pressure, $T_0(p)$, plus a deviation from the basic state, $T(x, y, p, t)$, as was done for potential temperature in Section 2.7.4. Thus,

$$T_{tot}\,(x, y, p, t) = T_0\,(p) + T\,(x, y, p, t)$$

Now because $|dT_0/dp| \gg |\partial T/\partial p|$, only the basic state portion of the temperature field need be included in the static stability term, and the quasi-geostrophic thermodynamic energy equation may be expressed in the form

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \bullet \nabla\right) T - \left(\frac{\sigma p}{R}\right) \omega = \frac{J}{c_p} \tag{6.13a}$$

where $\sigma \equiv -RT_0 p^{-1} d\ln\theta_0/dp$ and θ_0 is the potential temperature corresponding to the basic state temperature T_0 [$\sigma \approx 2.5 \times 10^{-6}\mathrm{m}^2\,\mathrm{Pa}^{-2}\mathrm{s}^{-2}$ in the midtroposphere]. From (6.2) the thermodynamic energy equation can be expressed in terms of the geopotential field. The result is

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \bullet \nabla\right)\left(-\frac{\partial \Phi}{\partial p}\right) - \sigma\omega = \frac{\kappa J}{p} \tag{6.13b}$$

where $\kappa \equiv R/c_p$.

Equations (6.7), (6.11), (6.12), and (6.13b) constitute the quasi-geostrophic equations. These form a complete set in the dependent variables Φ, $\mathbf{V}_g$, $\mathbf{V}_a$, and ω (provided that the diabatic heating rate is known). This form of the equations is not, however, suitable as a prediction system. As shown in the next section, it proves

useful to replace (6.11) by an equation for the evolution of the vorticity of the geostrophic wind, in which case only the divergent part of the ageostrophic wind plays a role in the dynamics. The quasi-geostrophic momentum equation (6.11) can still be used diagnostically, however, to determine the nondivergent part of $\mathbf{V}_a$, for a known distribution of $\mathbf{V}_g$ (see Section 6.4.3).

6.2.2 The Quasi-Geostrophic Vorticity Equation

Just as the horizontal momentum can be approximated to O(Ro) by its geostrophic value, the vertical component of vorticity can also be approximated geostrophically. In Cartesian coordinates the components of (6.7) are

$$f_0 v_g = \frac{\partial \Phi}{\partial x}, \quad f_0 u_g = -\frac{\partial \Phi}{\partial y} \tag{6.14}$$

Thus, the geostrophic vorticity, $\zeta_g = \mathbf{k} \bullet \nabla \times \mathbf{V}_g$, can be expressed as

$$\zeta_g = \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} = \frac{1}{f_0}\left(\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2}\right) = \frac{1}{f_0}\nabla^2 \Phi \tag{6.15}$$

Equation (6.15) can be used to determine $\zeta_g(x, y)$ from a known field $\Phi(x, y)$. Alternatively, (6.15) can be solved by inverting the Laplacian operator to determine Φ from a known distribution of ζ_g, provided that suitable conditions on Φ are specified on the boundaries of the region in question. This *invertibility* is one reason why vorticity is such a useful forecast diagnostic; if the evolution of the vorticity can be predicted, then inversion of (6.15) yields the evolution of the geopotential field, from which it is possible to determine the geostrophic wind and temperature distributions. Since the Laplacian of a function tends to be a maximum where the function itself is a minimum, positive vorticity implies low values of geopotential and vice versa, as illustrated for a simple sinusoidal disturbance in Fig. 6.7.

The quasi-geostrophic vorticity equation can be obtained from the x and y components of the quasi-geostrophic momentum equation (6.11), which can be expressed, respectively, as

$$\frac{D_g u_g}{Dt} - f_0 v_a - \beta y v_g = 0 \tag{6.16}$$

and

$$\frac{D_g v_g}{Dt} + f_0 u_a + \beta y u_g = 0 \tag{6.17}$$

Taking $\partial(6.17)/\partial x - \partial(6.16)/\partial y$, and using the fact that the divergence of the geostrophic wind defined in (6.7) vanishes, immediately yields the vorticity equation

$$\frac{D_g \zeta_g}{Dt} = -f_0\left(\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y}\right) - \beta v_g \tag{6.18}$$

which should be compared with (4.22b).

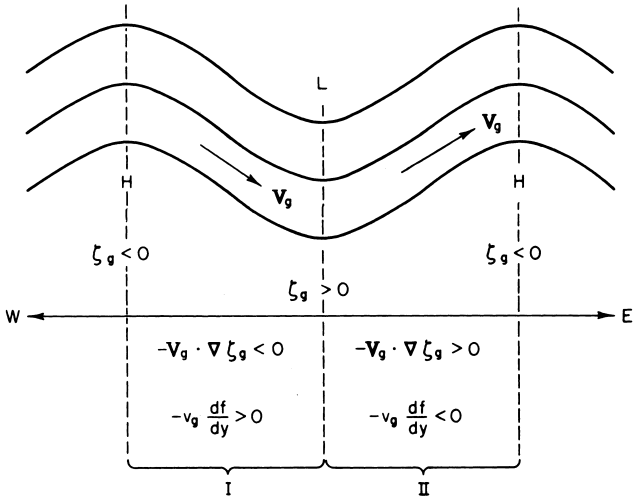


Fig. 6.7 Schematic 500-hPa geopotential field showing regions of positive and negative advections of relative and planetary vorticity.

Noting that because f depends only on y so that $D_g f/Dt = \mathbf{V}_g \bullet \nabla f = \beta v_g$ and that the divergence of the ageostrophic wind can be eliminated in favor of ω using (6.12), we can rewrite (6.18) as

$$\frac{\partial \zeta_g}{\partial t} = -\mathbf{V}_g \bullet \nabla \left(\zeta_g + f\right) + f_0 \frac{\partial \omega}{\partial p} \tag{6.19}$$

which states that the local rate of change of geostrophic vorticity is given by the sum of the advection of the absolute vorticity by the geostrophic wind plus the concentration or dilution of vorticity by stretching or shrinking of fluid columns (the divergence effect).

The vorticity tendency due to vorticity advection [the first term on the right in (6.19)] may be rewritten as

$$-\mathbf{V}_g \bullet \nabla \left(\zeta_g + f\right) = -\mathbf{V}_g \bullet \nabla \zeta_g - \beta v_g$$

The two terms on the right represent the geostrophic advections of relative vorticity and planetary vorticity, respectively. For disturbances in the westerlies, these two effects tend to have opposite signs, as illustrated schematically in Fig. 6.7 for an idealized 500-hPa flow.

In region I upstream of the 500-hPa trough, the geostrophic wind is directed from the relative vorticity minimum at the ridge toward the relative vorticity maximum at the trough so that $-V_g \bullet \nabla \zeta_g < 0$. However, at the same time, because $v_g < 0$ in region I, the geostrophic wind has its y component directed down the gradient of

planetary vorticity so that $-\beta v_g > 0$. Hence, in region I the advection of relative vorticity tends to decrease the local vorticity, whereas the advection of planetary vorticity tends to increase the local vorticity. Similar arguments (but with reversed signs) apply to region II. Therefore, advection of relative vorticity tends to move the vorticity pattern and hence the troughs and ridges eastward (downstream). However, advection of planetary vorticity tends to move the troughs and ridges westward against the advecting wind field. The latter motion is called *retrograde* motion or *retrogression*.

The net effect of advection on the evolution of the vorticity pattern depends on which type of vorticity advection dominates. In order to compare the magnitudes of the relative and planetary vorticity advections, we consider an idealized geopotential distribution on a midlatitude β-plane consisting of the sum of a zonally averaged part, which depends linearly on y, and a zonally varying part (representing a synoptic wave disturbance) that has a sinusoidal dependence in x and y:

$$\Phi(x, y) = \Phi_0 - f_0 U y + f_0\ A \sin kx \cos ly \tag{6.20}$$

Here, $y = a(\phi - \phi_0)$, where a is the radius of the earth and ϕ_0 is the latitude at which f_0 is evaluated. The parameters Φ_0, U, and A depend only on pressure, and the *wave numbers* k and l are defined as $k = 2\pi/L_x$ and $l = 2\pi/L_y$ with L_x and L_y the wavelengths in the x and y directions, respectively. The geostrophic wind components are then given by

$$u_g = -\frac{1}{f_0}\frac{\partial \Phi}{\partial y} = U + u'_g = U + lA \sin kx \sin ly$$
$$v_g = \frac{1}{f_0}\frac{\partial \Phi}{\partial x} = v'_g = +kA \cos kx \cos ly$$

where $\left(u'_g,\ v'_g\right)$ is the geostrophic wind due to the synoptic wave disturbance. The geostrophic vorticity is then simply

$$\zeta_g = f_0^{-1}\, \nabla^2 \Phi = -\left(k^2 + l^2\right) A \sin kx\ \cos ly$$

With the aid of these relations it can be shown that in this simple case the advection of relative vorticity by the wave component of the geostrophic wind vanishes:

$$u'_g \partial \zeta_g \Big/ \partial x + v'_g \partial \zeta_g \Big/ \partial y = 0$$

so that the advection of relative vorticity is simply

$$-u_g \frac{\partial \zeta_g}{\partial x} - v_g \frac{\partial \zeta_g}{\partial y} = -U \frac{\partial \zeta_g}{\partial x} = +kU \left(k^2 + l^2\right) A \cos kx \cos ly$$

and the advection of planetary vorticity can be expressed as

$$-\beta v_g = -\beta k A \cos kx \cos ly$$

The horizontal distributions of geopotential, relative vorticity, and advection of relative vorticity for this case are shown in Fig. 6.8. For a disturbance with a fixed amplitude of geopotential disturbance, the amplitude of the vorticity increases as the square of the wave number or, inversely, as the square of the horizontal scale. As a consequence, the advection of relative vorticity dominates over planetary vorticity advection for short waves ($L_x < 3000$ km), whereas for long waves ($L_x > 10,000$ km) the planetary vorticity advection tends to dominate. Therefore, as a general rule, short-wavelength synoptic-scale systems should move eastward in a westerly zonal flow, whereas long planetary waves should tend to retrogress

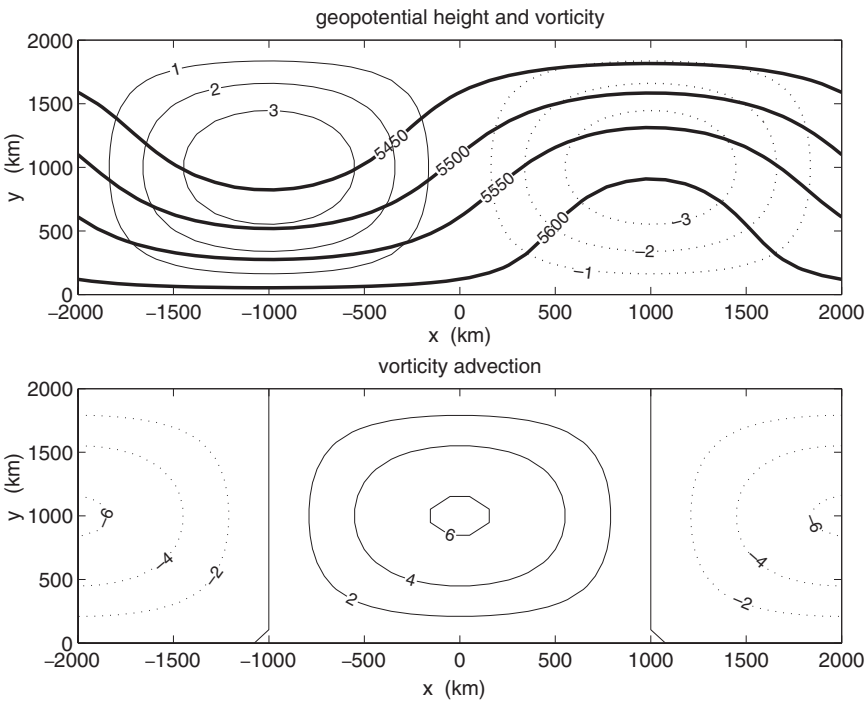


Fig. 6.8 (Top) Geopotential height in units of m, and relative vorticity in units of $10^{-5}\,\mathrm{s}^{-1}$ for the sinusoidal disturbance of equation (6.20). Here $\Phi_0 = 5.5 \times 10^4\,\mathrm{m}^2\,\mathrm{s}^{-2}$, $f_0 = 10^{-4}\,\mathrm{s}^{-1}$, $f_0 A = 800\,\mathrm{m}^2\,\mathrm{s}^{-2}$, $U = 10\,\mathrm{m\,s}^{-1}$, and $k = l = (\pi/2) \times 10^{-6}\,\mathrm{m}^{-1}$. (Bottom) Advection of relative vorticity in units of $10^{-10}\,\mathrm{s}^{-2}$ for the disturbance shown above.

(move westward against the mean flow).[3] Waves of intermediate wavelength may be quasi-stationary or move eastward much slower than the mean geostrophic wind speed. Since positive maxima in relative vorticity are associated with cyclonic disturbances, regions of positive vorticity advection, which can be estimated easily from upper level maps, are commonly used as aids in forecasting synoptic-scale weather disturbances.

Vorticity advection does not alone determine the evolution of meteorological systems. A change in the vertical shear of the horizontal wind associated with differential (i.e., height dependent) vorticity advection will drive an ageostrophic vertical circulation, which adiabatically adjusts the horizontal temperature gradient to maintain thermal wind balance. The convergence and divergence fields associated with this vertical circulation will not only modify the effects of vorticity advection at upper levels, but will force changes in the vorticity distribution in the lower troposphere where advection may be very weak.

In an analogous manner, thermal advection, which is often strong near the surface, does not merely force changes in the temperature in the lower troposphere. Rather, it will induce a vertical circulation, which through its associated divergence and convergence patterns will alter the vorticity fields both near the surface and aloft so that thermal wind balance is maintained.

The vertical circulation induced by quasi-geostrophic differential vorticity advection and thermal advection is generally an order of magnitude larger than that induced by boundary layer pumping (5.38). Thus, it is reasonable to neglect boundary layer effects to a first approximation in quasi-geostrophic theory.

6.3 QUASI-GEOSTROPHIC PREDICTION

Although, as explained above, the ageostrophic vertical motion plays an essential role in the maintenance of thermal wind balance as the flow evolves, the evolution of the geostrophic circulation can actually be determined without explicitly determining the distribution of ω. Defining the geopotential tendency $\chi \equiv \partial\Phi/\partial t$, and recalling that the order of partial differentiation may be reversed, the geostrophic vorticity equation (6.19) can be expressed as

$$\frac{1}{f_0}\nabla^2\chi = -\mathbf{V}_g \bullet \nabla\left(\frac{1}{f_0}\nabla^2\Phi + f\right) + f_0\frac{\partial\omega}{\partial p} \tag{6.21}$$

[3] Observed long waves tend to remain stationary rather than to retrogress. This is believed to be a result of processes such as modification of the Rossby phase speed by potential vorticity gradients associated with the mean flow, nonlinear interactions with transient short waves, forcing due to topographic influences, and diabatic heating contrasts associated with land–sea differences, as was mentioned previously in Section 6.1.

where we have used (6.15) to write the geostrophic vorticity and its tendency in terms of the Laplacian of geopotential. Thus, since by (6.7) the geostrophic wind can be expressed in terms of Φ, the right-hand side of (6.21) depends on the dependent variables Φ, and ω alone. An analogous equation dependent on these two variables can be obtained from the thermodynamic energy equation (6.13b) by multiplying through by f_0/σ and differentiating with respect to p. Using the definition of χ given above, the result can be expressed as

$$\frac{\partial}{\partial p}\left(\frac{f_0}{\sigma}\frac{\partial \chi}{\partial p}\right) = -\frac{\partial}{\partial p}\left[\frac{f_0}{\sigma}\mathbf{V}_g \bullet \nabla\left(\frac{\partial \Phi}{\partial p}\right)\right] - f_0\frac{\partial \omega}{\partial p} - f_0\frac{\partial}{\partial p}\left(\frac{\kappa J}{\sigma p}\right) \tag{6.22}$$

where σ was defined below (6.13a).

The ageostrophic vertical motion, ω, has equal and opposite effects on the left-hand sides in (6.21) and (6.22). Vertical stretching $\left(\partial\omega/\partial p > 0\right)$ forces a positive tendency in the geostrophic vorticity (6.21) and a negative tendency of equal magnitude in the term on the left side in (6.22). The left side of (6.22) can be interpreted as the local rate of change of a normalized *static stability anomaly* (i.e., a measure of the departure of static stability from S_p, its standard atmosphere value). This can be demonstrated by substituting from the hydrostatic equation (6.2) to give for the left side of (6.22)

$$\frac{\partial}{\partial p}\left(\frac{f_0}{\sigma}\frac{\partial \chi}{\partial p}\right) = -Rf_0\frac{\partial}{\partial p}\left(\frac{1}{\sigma p}\frac{\partial T}{\partial t}\right) = -f_0\frac{\partial}{\partial p}\left(\frac{1}{S_p}\frac{\partial T}{\partial t}\right) \approx -\frac{\partial}{\partial t}\left(\frac{f_0}{S_p}\frac{\partial T}{\partial p}\right)$$

where we have used the fact that $S_p = p\sigma/R$ varies only slowly with height in the troposphere. Because T is the departure of temperature from its standard atmosphere value, the expression

$$f_0 S_p^{-1}\partial T/\partial p$$

is proportional to the local static stability anomaly divided by the standard atmosphere static stability. Multiplication by f_0 gives this expression the same units as vorticity.

The left side of (6.22) is negative when the lapse rate tendency is positive (i.e., when the static stability tendency is negative). As was shown in Fig. 4.7, an air column that moves adiabatically from a region of high static stability to a region of low static stability is stretched vertically $\left(\partial\omega/\partial p > 0\right)$ so that the upper portion of the column cools adiabatically relative to the lower portion. Thus, the relative vorticity in (6.21) and the normalized static stability anomaly in (6.22) are changed by equal and opposite amounts. For this reason the normalized static stability anomaly is referred to as the *stretching vorticity.*

Purely geostrophic motion ($\omega = 0$) is a solution to (6.21) and (6.22) only in very special situations such as barotropic flow (no pressure dependence) or zonally

symmetric flow (no x dependence). More general purely geostrophic flows cannot satisfy both these equations simultaneously, as there are then two independent equations in a single unknown so that the system is over-determined. Thus, it should be clear that the role of the vertical motion distribution must be to maintain consistency between the geopotential tendencies required by vorticity advection in (6.21) and thermal advection in (6.22).

6.3.1 Geopotential Tendency

If for simplicity we set $J = 0$ and eliminate ω between (6.21) and (6.22), we obtain an equation that determines the local rate of change of geopotential in terms of the three-dimensional distribution of the Φ field:

$$\underbrace{\left[\nabla^2 + \frac{\partial}{\partial p}\left(\frac{f_0^2}{\sigma}\frac{\partial}{\partial p}\right)\right]}_{\text{A}}\chi = -\underbrace{f_0\mathbf{V}_g \bullet \nabla\left(\frac{1}{f_0}\nabla^2\Phi + f\right)}_{\text{B}} - \underbrace{\frac{\partial}{\partial p}\left[-\frac{f_0^2}{\sigma}\mathbf{V}_g \bullet \nabla\left(-\frac{\partial\Phi}{\partial p}\right)\right]}_{\text{C}} \tag{6.23}$$

This equation is often referred to as the *geopotential tendency equation*. It provides a relationship between the local geopotential tendency (term A) and the distributions of vorticity advection (term B) and thickness advection (term C). If the distribution of Φ is known at a given time, then terms B and C may be regarded as known forcing functions, and (6.23) is a linear partial differential equation in the unknown χ.

Although (6.23) appears to be quite complicated, a qualitative notion of its implications can be gained if we note that term A involves second derivatives in space of the χ field, and is thus generally proportional to minus χ. Term B is proportional to the advection of absolute vorticity; it is usually the dominant forcing term in the upper troposphere. As the discussion in Section 6.2.2 indicated, for short waves term B is negative in region I of Fig. 6.7 (upstream of the upper level trough). Thus, because the sign of the geopotential tendency is opposite to that of term B in this case, χ will be positive and a ridge will tend to develop. This ridging is, of course, necessary for the development of a negative geostrophic vorticity. Similar arguments, but with the signs reversed, apply to region II downstream from the upper level trough where falling geopotential heights are associated with a positive relative vorticity advection. It is also important to note that the vorticity advection term is zero along both the trough and ridge axes, as both $\nabla\,\zeta_g$ and v_g are zero at the axes. Thus, vorticity advection cannot change the strength of this

type of disturbance at the levels where the advection is occurring, but only acts to propagate the disturbance horizontally and (as shown in the next section) to spread it vertically.

The major mechanism for amplification or decay of midlatitude synoptic systems is contained in term C of (6.23). This term involves the rate of change with pressure of the horizontal thickness advection. (If we had retained the diabatic heating term, it would have contributed in a similar fashion.) The thickness advection tends to be largest in magnitude in the lower troposphere beneath the 500-hPa trough and ridge lines in a developing baroclinic wave. Now because $-\partial\Phi/\partial p$ is proportional to temperature, the thickness advection is proportional to the temperature advection. Thus, term C in (6.23) is proportional to minus the rate of change of temperature advection with respect to pressure (i.e., plus the rate of change with respect to height). This term is sometimes referred to as *differential temperature advection*.

Differential temperature advection enhances upper level height anomalies in developing disturbances. Below the 500-hPa ridge there is strong warm advection associated with the warm front, whereas below the 500-hPa trough there is strong cold advection associated with the cold front. The former increases thickness, thus builds the upper level ridge; the latter decreases thickness, thus deepens the upper level trough. Above the 500-hPa level the temperature gradient is usually weaker, and the isotherms often become nearly parallel to the height lines so that thermal advection tends to be small. Thus, in contrast to term B in (6.23), the forcing term C is concentrated in the lower troposphere, but again, the geopotential tendency response will not be limited to the levels of forcing, but is spread in the vertical so that, as noted above, in developing waves differential temperature advection will deepen upper level troughs and build upper level ridges.

In the region of warm advection $-\mathbf{V}_g \bullet \nabla(-\partial\Phi/\partial_p) > 0$, as $\mathbf{V}_g$ has a component down the temperature gradient. However, as explained above, the warm advection decreases with height (increases with pressure) so that $\partial\left[-\mathbf{V}_g \bullet \nabla\,(-\partial\Phi/\partial p)\right]/\partial p > 0$. Conversely, beneath the 500-hPa trough where there is cold advection decreasing with height, the opposite signs obtain. Thus, along the 500-hPa trough and ridge axes where the vorticity advection is zero, the tendency equation states that for a developing wave

$$\chi = \frac{\partial\Phi}{\partial t} \sim \left\{ \frac{\partial}{\partial p}\left[-\mathbf{V}_g \bullet \nabla\left(-\frac{\partial\Phi}{\partial p}\right)\right] \quad \begin{array}{l} > 0 \text{ at the ridge} \\ < 0 \text{ at the trough} \end{array} \right.$$

Therefore, as indicated in Fig. 6.9, the effect of cold advection below the 500-hPa trough is to *deepen* the trough in the upper troposphere, and the effect of warm advection below the 500-hPa ridge is to *build* the ridge in the upper troposphere. Hence, differential temperature advection, even if limited to the lower troposphere, intensifies the upper-level troughs and ridges in a developing system.

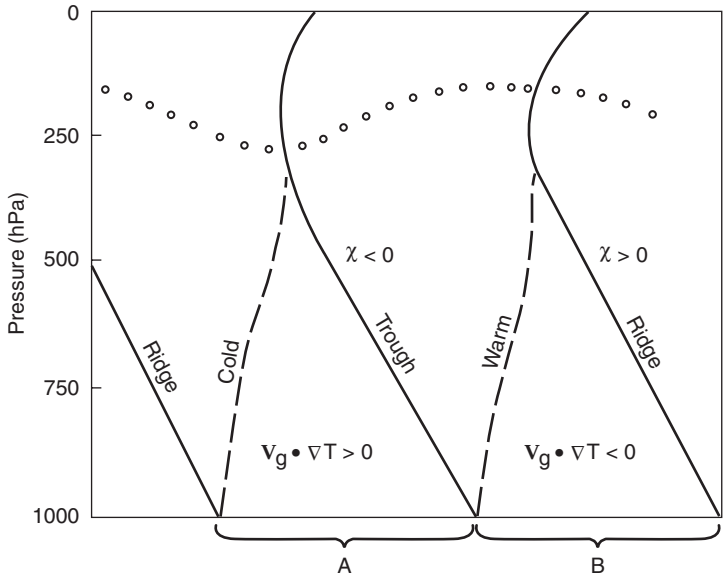


Fig. 6.9 East–west section through a developing synoptic disturbance showing the relationship of temperature advection to the upper level height tendencies. A and B designate, respectively, regions of cold advection and warm advection in the lower troposphere.

Qualitatively, the effects of differential temperature advection are easily understood since the advection of cold air into the air column below the 500-hPa trough reduces the thickness of that column, and hence lowers the height of the 500-hPa surface unless there is a compensating rise in the surface pressure. Obviously warm advection into the air column below the ridge has the opposite effect.

To summarize, in the absence of diabatic heating the horizontal temperature advection must be nonzero in order that a midlatitude synoptic system intensifies through baroclinic processes. As shown in Chapter 8, the temperature advection pattern described above indirectly implies conversion of potential energy to kinetic energy.

6.3.2 The Quasi-Geostrophic Potential Vorticity Equation

The geopotential tendency equation stated in the form (6.23) is useful for physical motivation of processes leading to geopotential changes (and hence upper level troughing and ridging), as the tendency χ is related to the easily comprehended processes of vorticity and temperature advection. However, this form of the equation actually conceals its true character as a conservation equation for a field commonly referred to as *quasi-geostrophic potential vorticity*. To put (6.23)

in conservation form we simplify the right-hand side by using the chain rule of differentiation to write term C as

$$-\mathbf{V}_g \bullet \nabla \frac{\partial}{\partial p}\left(\frac{f_0^2}{\sigma}\frac{\partial \Phi}{\partial p}\right) - \frac{f_0^2}{\sigma}\frac{\partial \mathbf{V}_g}{\partial p} \bullet \nabla \frac{\partial \Phi}{\partial p}$$

However, from the thermal wind relation, $f_0 \partial \mathbf{V}_g/\partial p = \mathbf{k} \times \nabla(\partial \Phi/\partial p)$, so that $\partial \mathbf{V}_g/\partial p$ is perpendicular to $\nabla(\partial \Phi/\partial p)$, and the second term in the above expression vanishes. The first term (which is proportional to the geostrophic advection of the normalized static stability anomaly) can be combined with term B in (6.23). Recalling the definition of χ and dividing through by f_0, (6.23) can then be expressed simply in the form of a conservation equation:

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \bullet \nabla\right) q = \frac{D_g q}{Dt} = 0 \tag{6.24}$$

where q is the quasi-geostrophic potential vorticity defined by

$$q \equiv \left[\frac{1}{f_0}\nabla^2 \Phi + f + \frac{\partial}{\partial p}\left(\frac{f_0}{\sigma}\frac{\partial \Phi}{\partial p}\right)\right] \tag{6.25}$$

The three parts of (6.25), reading from left to right, are the geostrophic *relative* vorticity, the *planetary* vorticity, and the *stretching* vorticity. As a parcel moves about in the atmosphere, the geostrophic relative vorticity, the planetary vorticity, and the stretching vorticity terms may each change. However, according to (6.24) their sum is conserved following the geostrophic motion. This conservation is consistent with (6.21) and (6.22), which show that for adiabatic flow, changes in the absolute vorticity are balanced by opposite changes in the stretching vorticity.

The scalar q is proportional to a linearized form of the Ertel potential vorticity, P, discussed in Section 4.3. However, unlike Ertel's potential vorticity, q has units of vorticity and is determined solely by the isobaric geopotential distribution. Note that whereas Ertel's potential vorticity is conserved following the adiabatic motion on an isentropic surface, q is conserved under adiabatic conditions following the geostrophic motion on an isobaric surface. For this reason some authors prefer the term *pseudo-potential vorticity*. However, we will generally refer to q as the quasi-geostrophic potential vorticity.

According to (6.25) quasi-geostrophic potential vorticity tends to be proportional to minus the geopotential. A local increase in q is associated with trough development, whereas a decrease in q is associated with ridge development. Because q is a conserved quantity following the geostrophic motion, we can diagnose the tendency purely from the geostrophic advection of q. Furthermore, (6.24) shows that the tendency will be zero (i.e., the flow will be steady) provided that the geostrophic wind is everywhere parallel to lines of constant q (i.e., that $\mathbf{V}_g \bullet \nabla q = 0$).

Given an initial distribution of Φ and suitable boundary conditions, (6.24) and (6.25) can be used to forecast the evolution of the geopotential field. For simplicity it is often assumed that the lower boundary corresponds to $p_0 = 1000$ hPa and that ω vanishes at this surface so that from (6.13b) the lower boundary condition for adiabatic motion is simply that temperature (or thickness) is conserved following the geostrophic motion on the 1000-hPa surface:

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \bullet \nabla\right)\left(-\frac{\partial \Phi}{\partial p}\right)\bigg|_{p=p_0} = 0 \tag{6.26}$$

6.3.3 Potential Vorticity Inversion

Equations (6.24)–(6.26), together with suitable lateral and upper boundary conditions on Φ, completely prescribe the evolution of the geostrophic flow. Thus, under adiabatic conditions the evolution of large-scale midlatitude meteorological systems is completely determined by the twin constraints of quasi-geostrophic potential vorticity conservation following the geostrophic flow in the interior and temperature conservation following the geostrophic flow along the lower boundary.

If boundary effects can be neglected, the distribution of potential vorticity alone determines the geopotential distribution, and hence the distributions of temperature and geostrophic wind. The process of recovering the geopotential field from the potential vorticity field is referred to as potential vorticity inversion. Because Φ is related to q by a second-order differential operator, the geopotential disturbance associated with a localized anomaly in q will extend horizontally and vertically beyond the region of anomalous q. This remote influence can be illustrated by a simple example.

For mathematical simplicity we assume that the static stability parameter, σ is constant in height and introduce a stretched vertical coordinate defined as $\hat{z} \equiv p\left(\sigma^{1/2}/f_0\right)$. Then (6.25) can be expressed as

$$q = f + q' = f + \frac{1}{f_0}\left(\frac{\partial^2 \Phi'}{\partial x^2} + \frac{\partial^2 \Phi'}{\partial y^2} + \frac{\partial^2 \Phi'}{\partial \hat{z}^2}\right) \tag{6.27}$$

where q' and Φ' designate the potential vorticity and geopotential anomalies, respectively.

We now suppose that the potential vorticity anomaly consists of a compact ball of constant positive potential vorticity, q_0, with spherical shape in the $\left(x, y, \hat{z}\right)$ coordinate system. Then, if the ball is centered at the origin, q'can be represented in terms of the radial coordinate $r \equiv \left(x^2 + y^2 + \hat{z}^2\right)^{1/2}$ as

$$q' = \frac{1}{f_0 r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \Phi'}{\partial r}\right) \tag{6.28}$$

where

$$q' = \left\{ \begin{array}{lll} q_0 & for & r \leq b \\ 0 & for & r > b \end{array} \right\}.$$

It is easily verified that the solution of (6.28) satisfying the requirement of continuity in Φ' and $\partial\Phi'/\partial r$ at $r = b$ is

$$\Phi' = \left\{ \begin{array}{ll} -f_0 q_0 \left(b^2/2 - r^2/6\right) & \text{for } r \leq b \\ -f_0 q_0 b^3/(3r) & \text{for } r > b \end{array} \right\} \tag{6.29}$$

Figure 6.10 shows the pattern of potential temperature and geostrophic wind induced by this compact positive potential vorticity anomaly. The distribution of the θ surfaces indicates that static stability is enhanced in the ball of positive potential vorticity and is reduced both above and below the ball. Because potential vorticity is conserved (and has zero anomaly outside the ball), positive vorticity anomalies are induced above and below the ball of potential vorticity to compensate the reduced stability, and hence the geostrophic flow induced by the potential vorticity anomaly extends well beyond the boundary of the anomaly.

6.3.4 Vertical Coupling Through Potential Vorticity

As illustrated above, a potential vorticity anomaly at one level yields a nonzero geopotential anomaly (and hence nonzero geostrophic winds) at other levels. This

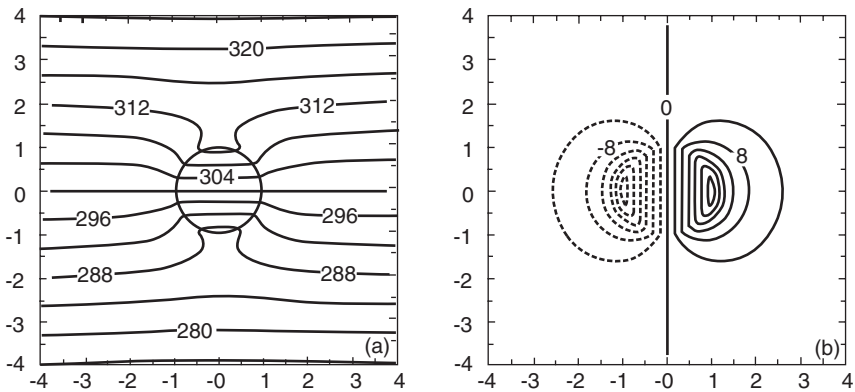


Fig. 6.10 Vertical cross sections showing potential temperature (left) and geostrophic wind (right) induced by a ball of constant potential vorticity of nondimensional radius unity. A constant standard atmosphere potential vorticity is added to the anomaly induced by the ball of potential vorticity. (Right) Solid contours indicate flow into the page and dashed contours indicate flow out of the page. The horizontal and scaled vertical distances are normalized by the radius of the potential vorticity ball. For a ball of 250-km radius the horizontal distance shown is 2000 km and the vertical distance is 20 km (assuming that the buoyancy frequency is 100 times the Coriolis frequency). (Adapted from Thorpe and Bishop, 1995)

has important consequences for the influence of upper level disturbances on the lower troposphere. The induction of geopotential changes in the lower troposphere by an upper level potential vorticity change can be demonstrated in a simple example.

Suppose that the flow in the upper troposphere is characterized by the geopotential distribution of (6.20). The potential vorticity can then be expressed as $q = f + Q(p)\sin kx \cos ly$, where $Q = -\left(k^2 + l\right)^2 A + \left(f_0/\sigma\right)\partial^2 A/\partial p^2$, and σ is assumed to be constant.[4] If the latitudinal variation of f is neglected, the quasi-geostrophic potential vorticity equation can then be expressed as

$$\frac{\partial q}{\partial t} = -\mathbf{V}_g \bullet \nabla q = -U\frac{\partial q}{\partial x} = -kUQ\cos kx \cos ly$$

so that again setting $\chi = \partial\Phi/\partial t$,

$$\frac{1}{f_0}\nabla^2\chi + \frac{f_0}{\sigma}\frac{\partial^2\chi}{\partial p^2} = -kUQ\cos kx \cos ly \tag{6.30}$$

which can be solved for the geopotential tendency. Letting

$$\chi(x, y, p, t) = \mathrm{X}(p, t)\cos kx \cos ly$$

and substituting into (6.30) yields an equation for the vertical dependence of the geopotential tendency:

$$\frac{d^2X}{dp^2} - \lambda^2 X = -\frac{\sigma}{f_0}kUQ \tag{6.31}$$

where $\lambda^2 \equiv \left(k^2 + l^2\right)\sigma f_0^{-2}$. Equation (6.31) shows that potential vorticity advection at a given altitude will generate a response in the geopotential tendency whose vertical scale (measured in pressure units) is λ^{-1}. Thus, for example, upper-level vorticity advection associated with disturbances of large horizontal scale (small k) will produce geopotential tendencies that extend down to the surface with little loss of amplitude, whereas for disturbances of small horizontal scale the response is confined close to the levels of forcing as shown in Fig. 6.11. In mathematical terms, the differential operator in (6.31) spreads the response in the vertical so that forcing at one altitude influences other altitudes.

[4] Actually, σ varies substantially with pressure even in the troposphere. However, the qualitative discussion in this section would not be changed if we were to include this additional complication.

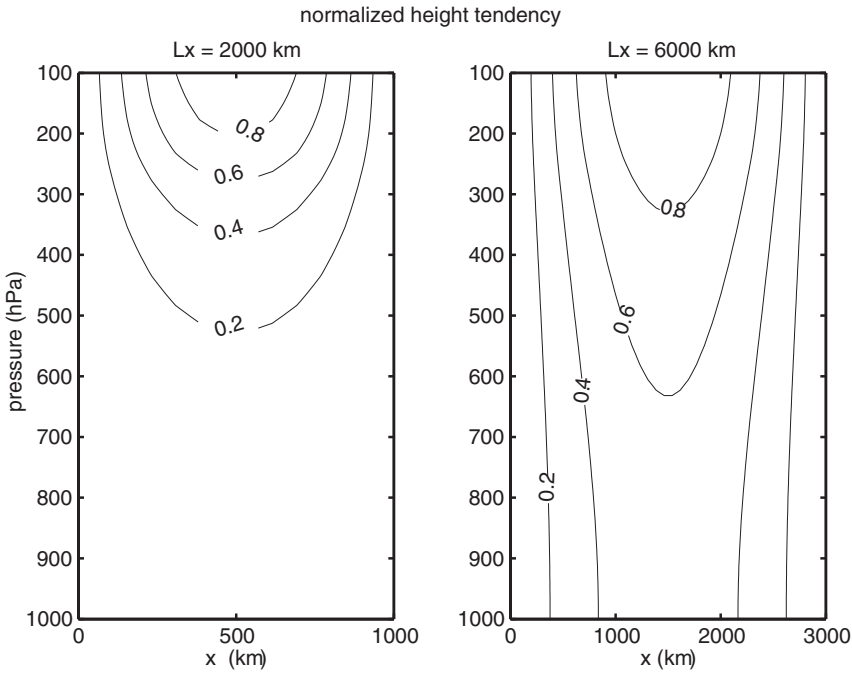


Fig. 6.11 Vertical spread of geopotential height tendencies forced by advection of a potential vorticity anomaly that is confined above the 250-hPa level. The displayed height tendencies are normalized by their value at the 100-hPa level. Solutions for zonal wavelength 2000 km (left) and 6000 km (right). Here, $l = 0, k = 2\pi/L_x$, $f_0 = 10^{-4}\text{s}^{-1}$, and $\sigma = 2 \times 10^{-6}\text{m}^2\,\text{Pa}^{-2}\,\text{s}^{-2}$.

6.4 DIAGNOSIS OF THE VERTICAL MOTION

As shown in the previous section, prediction of the evolution of the geopotential distribution for adiabatic flow does not depend explicitly on the ageostrophic motion. Nevertheless, the vertical velocity, which is strongly correlated with disturbed weather conditions, is an important diagnostic field.

6.4.1 The Traditional Omega Equation

Because ζ_g and $\mathbf{V}_g$are both defined in terms of $\Phi(x, y, p, t)$, the vorticity equation (6.19) can be used to diagnose the ω field provided that the fields of both Φ and $\partial\Phi/\partial t$ are known. The former is a primary product of operational weather analysis. However, because upper level analyses are generally available only twice per day, the latter can only be crudely approximated from observations by taking differences over 12 h. Despite this limitation, the vorticity equation method of estimating ω is usually more accurate than the continuity equation method discussed

in Section 3.5.1. However, neither of these methods of estimating ω uses the information available in the thermodynamic energy equation. An alternative method of estimating the vertical motion, called the *omega equation*, which utilizes both the vorticity equation and the thermodynamic equation, is developed in this section.

To obtain the omega equation we take the horizontal Laplacian of (6.13b) to yield

$$\nabla^2 \frac{\partial \chi}{\partial p} = -\nabla^2 \left[\mathbf{V}_g \bullet \nabla \left(\frac{\partial \Phi}{\partial p} \right) \right] - \sigma \nabla^2 \omega - \frac{\kappa}{p} \nabla^2 J \tag{6.32}$$

We next differentiate (6.21) with respect to p, yielding

$$\frac{\partial}{\partial p} \left(\nabla^2 \chi \right) = -f_0 \frac{\partial}{\partial p} \left[\mathbf{V}_g \bullet \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) \right] + f_0^2 \frac{\partial^2 \omega}{\partial p^2} \tag{6.33}$$

Because the order of the operators on the left-hand side in (6.32) and (6.33) may be reversed, the result of subtracting (6.32) from (6.33) is to eliminate χ. After some rearrangement of terms, we obtain the traditional *omega equation*

$$\begin{aligned} \underbrace{\left(\nabla^2 + \frac{f_0^2}{\sigma} \frac{\partial^2}{\partial p^2} \right)}_{\text{A}} \omega = {} & \frac{f_0}{\sigma} \frac{\partial}{\partial p} \underbrace{\left[\mathbf{V}_g \bullet \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) \right]}_{\text{B}} \\ & + \underbrace{\frac{1}{\sigma} \nabla^2 \left[\mathbf{V}_g \bullet \nabla \left(-\frac{\partial \Phi}{\partial p} \right) \right]}_{\text{C}} - \underbrace{\frac{\kappa}{\sigma p} \nabla^2 J}_{\text{D}} \end{aligned} \tag{6.34}$$

Although terms B and C in the omega equation (6.34) apparently have clear interpretations as separate physical processes, in practice there is often a significant amount of cancellation between them. They also are not invariant under a Galilean transformation of the zonal coordinate. That is, adding a constant mean zonal velocity will change the magnitude of each of these terms without changing the net forcing of the vertical motion. For these reasons, an alternative approximate form of the omega equation is often applied in synoptic analyses.

Employing the chain rule of differentiation yields for term B:

$$\frac{f_0}{\sigma} \left[\frac{\partial \mathbf{V}_g}{\partial p} \bullet \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) \right] + \frac{1}{\sigma} \mathbf{V}_g \bullet \nabla \left(\frac{\partial \nabla^2 \Phi}{\partial p} \right) \tag{6.35a}$$

and for term C:

$$-\frac{1}{\sigma} \left[\left(\nabla^2 u_g \right) \frac{\partial}{\partial x} \left(\frac{\partial \Phi}{\partial p} \right) + \left(\nabla^2 v_g \right) \frac{\partial}{\partial y} \left(\frac{\partial \Phi}{\partial p} \right) \right] - \frac{1}{\sigma} \mathbf{V}_g \bullet \nabla \left(\frac{\partial \nabla^2 \Phi}{\partial p} \right) \tag{6.35b}$$

The second terms in (6.35a) and (6.35b) are equal and opposite (hence the cancellation referred to above), whereas the first term in (6.35b) is generally much smaller than the first term in (6.35a). Thus, for adiabatic flow, the omega equation (6.34) can be expressed approximately as

$$\underbrace{\left(\nabla^2 + \frac{f_0^2}{\sigma}\frac{\partial^2}{\partial p^2}\right)}_{\mathrm{A}}\omega \approx \frac{f_0}{\sigma}\left[\frac{\partial \mathbf{V}_g}{\partial p}\bullet\nabla\left(\frac{1}{f_0}\nabla^2\Phi + f\right)\right] \tag{6.36}$$

Equation (6.36) involves only derivatives in space. It is, therefore, a diagnostic equation for the field of ω in terms of the instantaneous Φ field. The omega equation, unlike the continuity equation, provides a method of estimating ω that does not depend on observations of the ageostrophic wind. In fact, direct wind observations are not required at all; nor does the omega equation require information on the vorticity tendency, as required in the vorticity equation method, or information on the temperature tendency, as required in the adiabatic method discussed in Section 3.5.2. Only observations of Φ at a single time are needed to determine the ω field using (6.36). The terms in (6.36), however, do employ higher order derivatives than are involved in the other methods of estimating ω. Accurately estimating such terms from noisy observational data can be difficult.

The terms in (6.36) can be physically interpreted by first observing that the differential operator on the left is very similar to the operator in term A of the tendency equation (6.23) and hence tends to spread the response to a localized forcing. Because the forcing in (6.36) tends to be a maximum in the midtroposphere, and ω is required to vanish at the upper and lower boundaries, for qualitative discussion it is permissible to assume that ω has sinusoidal behavior not only in the horizontal, but also in the vertical:

$$\omega = W_0 \sin(\pi p/p_0)\ \sin kx\ \ \sin ly \tag{6.37}$$

we can then write

$$\left(\nabla^2 + \frac{f_0^2}{\sigma}\frac{\partial^2}{\partial p^2}\right)\omega \approx -\left[k^2 + l^2 + \frac{1}{\sigma}\left(\frac{f_0\pi}{p_0}\right)^2\right]\omega$$

which shows that the left side of (6.36) is proportional to $-\omega$. However, ω is proportional to $-w$ so $\omega < 0$ implies *upward* vertical motion, and the left-hand side of (6.36) is proportional to the vertical velocity. Thus, upward motion is forced where the right-hand side of (6.36) is positive and downward motion is forced where it is negative.

The right side of (6.36) represents the advection of absolute vorticity by the thermal wind. To understand the role of this advective term in forcing vertical

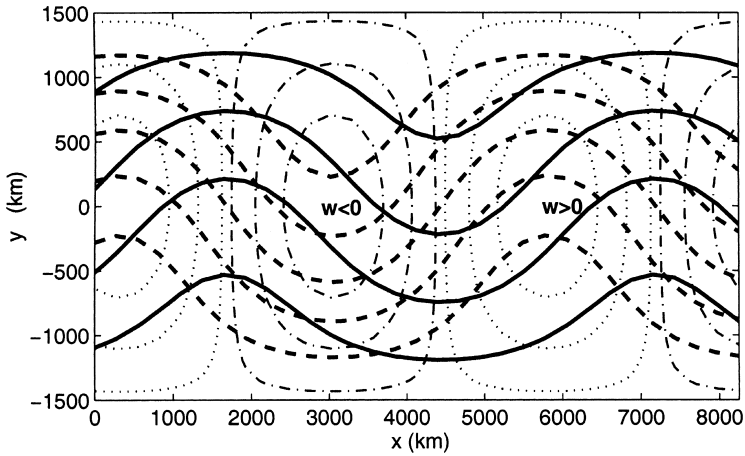


Fig. 6.12 Schematic 500-hPa height contours (solid lines), isotherms (dashed lines), and vertical motion field ($w > 0$ dash-dot lines, $w < 0$ dotted lines) for a developing synoptic-scale system. Upward motion occurs where vorticity decreases moving left to right along an isotherm, and downward motion occurs where vorticity decreases moving left to right along an isotherm.

motion, we again consider an idealized developing baroclinic system. Figure 6.12 indicates schematically the geopotential contours, isotherms, and vertical motion field at 500 hPa for such a system. Due to the westward tilt of the system with height, the 500-hPa geopotential field leads the isotherm pattern. Because the thermal wind is parallel to the isotherms, the right side of (6.36) can be estimated from the change of absolute vorticity along the isotherms. At 500 hPa the positive relative vorticity advection is a *maximum* above the surface low, whereas negative relative vorticity advection is strongest above the surface high. Thus, for a short-wave system where relative vorticity advection is larger than the planetary vorticity advection, the pattern of vertical motion at 500 hPa forced by the advection of vorticity by the thermal wind has

$$w \propto \left\{ \left[-\frac{\partial \mathbf{V}_g}{\partial z} \bullet \nabla \left(\zeta_g + f \right) \right] \begin{array}{l} < 0 \text{ east of the 500-hPa ridge} \\ > 0 \text{ east of the 500-hPa trough} \end{array} \right.$$

Thus, there is rising motion above the surface low and subsidence above the surface high. This pattern of vertical motion is in fact just what is required to produce the thickness tendencies in the 500- to 1000-hPa layer above the surface highs and lows. For example, above the surface low there is positive vorticity associated with negative geopotential deviations, as vorticity is proportional to the Laplacian of geopotential. Increasing vorticity thus implies a falling geopotential ($\chi < 0$). Hence the 500- to 1000-hPa thickness is decreasing in that region. Because horizontal temperature advection is small above the center of the surface low, the only

way to cool the atmosphere as required by the thickness tendency is by adiabatic cooling through the vertical motion field. Thus, the vertical motion maintains a hydrostatic temperature field (i.e., a field in which temperature and thickness are proportional) in the presence of differential vorticity advection. Without this compensating vertical motion, either the vorticity changes at 500 hPa could not remain geostrophic or the temperature changes in the 500- to 1000-hPa layer could not remain hydrostatic.

To summarize, we have shown as a result of scaling arguments that for synoptic-scale motions where vorticity is constrained to be geostrophic and temperature is constrained to be hydrostatic, the vertical motion field is determined uniquely by the geopotential field. Further, we have shown that this vertical motion field is just that required to ensure that changes in vorticity will be geostrophic and changes in temperature will be hydrostatic. These constraints, whose importance can hardly be overemphasized, are elaborated in the next subsection.

6.4.2 The Q Vector

In order to better appreciate the essential role of the divergent ageostrophic motion in quasi-geostrophic flow, it is useful to examine separately the rates of change, following the geostrophic wind, of the vertical shear of the geostrophic wind and of the horizontal temperature gradient.

On the midlatitude β-plane the quasi-geostrophic prediction equations may be expressed simply as

$$\frac{D_g u_g}{Dt} - f_0 v_a - \beta y v_g = 0 \tag{6.38}$$

$$\frac{D_g v_g}{Dt} + f_0 u_a + \beta y u_g = 0 \tag{6.39}$$

$$\frac{D_g T}{Dt} - \frac{\sigma p}{R}\omega = \frac{J}{c_p} \tag{6.40}$$

These are coupled by the thermal wind relationship

$$f_0 \frac{\partial u_g}{\partial p} = \frac{R}{p}\frac{\partial T}{\partial y}, \quad f_0 \frac{\partial v_g}{\partial p} = -\frac{R}{p}\frac{\partial T}{\partial x} \tag{6.41a,b}$$

or in vector form:

$$\left(f_0 \mathbf{k} \times \frac{\partial \mathbf{V}_g}{\partial p}\right) = \frac{R}{p}\boldsymbol{\nabla} T \tag{6.42}$$

Equations for the evolution of the thermal wind components are obtained by taking partial derivatives with respect to p in (6.38) and (6.39), multiplying through by f_0, and applying the chain rule of differentiation in the advective part of the total derivative to obtain

$$\frac{D_g}{Dt}\left(f_0\frac{\partial u_g}{\partial p}\right) = -f_0\left[\frac{\partial u_g}{\partial p}\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial p}\frac{\partial u_g}{\partial y}\right] + f_0^2\frac{\partial v_a}{\partial p} + f_0\beta y\frac{\partial v_g}{\partial p} \tag{6.43a}$$

$$\frac{D_g}{Dt}\left(f_0\frac{\partial v_g}{\partial p}\right) = -f_0\left[\frac{\partial u_g}{\partial p}\frac{\partial v_g}{\partial x} + \frac{\partial v_g}{\partial p}\frac{\partial v_g}{\partial y}\right] - f_0^2\frac{\partial u_a}{\partial p} - f_0\beta y\frac{\partial u_g}{\partial p} \tag{6.43b}$$

However, by the thermal wind relations (6.41a) and (6.41b), the first terms on the right-hand side in each of these may be expressed, respectively, as

$$-f_0\left[\frac{\partial u_g}{\partial p}\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial p}\frac{\partial u_g}{\partial y}\right] = -\frac{R}{p}\left[\frac{\partial T}{\partial y}\frac{\partial u_g}{\partial x} - \frac{\partial T}{\partial x}\frac{\partial u_g}{\partial y}\right]$$

$$-f_0\left[\frac{\partial u_g}{\partial p}\frac{\partial v_g}{\partial x} + \frac{\partial v_g}{\partial p}\frac{\partial v_g}{\partial y}\right] = -\frac{R}{p}\left[\frac{\partial T}{\partial y}\frac{\partial v_g}{\partial x} - \frac{\partial T}{\partial x}\frac{\partial v_g}{\partial y}\right]$$

Using the fact that the divergence of the geostrophic wind vanishes,

$$\partial u_g/\partial x + \partial v_g/\partial y = 0 \tag{6.44}$$

The above terms can be expressed, respectively, as

$$Q_2 \equiv -\frac{R}{p}\left[\frac{\partial u_g}{\partial y}\frac{\partial T}{\partial x} + \frac{\partial v_g}{\partial y}\frac{\partial T}{\partial y}\right] = -\frac{R}{p}\frac{\partial \mathbf{V}_g}{\partial y}\bullet\nabla T \tag{6.45a}$$

$$Q_1 \equiv -\frac{R}{p}\left[\frac{\partial u_g}{\partial x}\frac{\partial T}{\partial x} + \frac{\partial v_g}{\partial x}\frac{\partial T}{\partial y}\right] = -\frac{R}{p}\frac{\partial \mathbf{V}_g}{\partial x}\bullet\nabla T \tag{6.45b}$$

If we now take partial derivatives of (6.40) with respect to x and y, multiply the results by Rp^{-1}, and again apply the chain rule of differentiation to the advection terms, we obtain

$$\frac{D_g}{Dt}\left(\frac{R}{p}\frac{\partial T}{\partial x}\right) = -\frac{R}{p}\left(\frac{\partial u_g}{\partial x}\frac{\partial T}{\partial x} + \frac{\partial v_g}{\partial x}\frac{\partial T}{\partial y}\right) + \sigma\frac{\partial \omega}{\partial x} + \frac{\kappa}{p}\frac{\partial J}{\partial x} \tag{6.46a}$$

$$\frac{D_g}{Dt}\left(\frac{R}{p}\frac{\partial T}{\partial y}\right) = -\frac{R}{p}\left(\frac{\partial u_g}{\partial y}\frac{\partial T}{\partial x} + \frac{\partial v_g}{\partial y}\frac{\partial T}{\partial y}\right) + \sigma\frac{\partial \omega}{\partial y} + \frac{\kappa}{p}\frac{\partial J}{\partial y} \tag{6.46b}$$

Using the definitions of (6.45a,b), we can rewrite (6.43a,b) and (6.46a,b) as

$$\frac{D_g}{Dt}\left(f_0\frac{\partial u_g}{\partial p}\right) = -Q_2 + f_0^2\frac{\partial v_a}{\partial p} + f_0\beta y\frac{\partial v_g}{\partial p} \tag{6.47}$$

$$\frac{D_g}{Dt}\left(\frac{R}{p}\frac{\partial T}{\partial y}\right) = Q_2 + \sigma\frac{\partial \omega}{\partial y} + \frac{\kappa}{p}\frac{\partial J}{\partial y} \tag{6.48}$$

$$\frac{D_g}{Dt}\left(f_0\frac{\partial v_g}{\partial p}\right) = Q_1 - f_0^2\frac{\partial u_a}{\partial p} - f_0\beta y\frac{\partial u_g}{\partial p} \tag{6.49}$$

$$\frac{D_g}{Dt}\left(\frac{R}{p}\frac{\partial T}{\partial x}\right) = Q_1 + \sigma\frac{\partial \omega}{\partial x} + \frac{\kappa}{p}\frac{\partial J}{\partial x} \tag{6.50}$$

Suppose that $Q_2 > 0$ and that the thermal wind is westerly ($\partial u_g/\partial p < 0$ and $\partial T/\partial y < 0$). Then from (6.47), Q_2 forces an increase in the westerly shear following the geostrophic motion ($\partial u_g/\partial p$ becomes more negative). However, from (6.48), $Q_2 > 0$ forces a positive change in the meridional temperature gradient following the geostrophic motion ($\partial T/\partial y$ becomes less negative). Q_2 thus tends to destroy the thermal wind balance between the vertical shear of the zonal wind and the meridional temperature gradient. Similarly, Q_1 destroys the thermal wind balance between vertical shear of the meridional wind and the zonal temperature gradient. An ageostrophic circulation is thus required to keep the flow in approximate thermal wind balance.

Subtracting (6.47) from (6.48) and using (6.41a) to eliminate the total derivative gives

$$\sigma\frac{\partial \omega}{\partial y} - f_0^2\frac{\partial v_a}{\partial p} - f_0\beta y\frac{\partial v_g}{\partial p} = -2Q_2 - \frac{\kappa}{p}\frac{\partial J}{\partial y} \tag{6.51}$$

Similarly, adding (6.50) to (6.49) and using (6.41b) to eliminate the total derivative gives

$$\sigma\frac{\partial \omega}{\partial x} - f_0^2\frac{\partial u_a}{\partial p} - f_0\beta y\frac{\partial u_g}{\partial p} = -2Q_1 - \frac{\kappa}{p}\frac{\partial J}{\partial x} \tag{6.52}$$

If we now take $\partial(6.52)/\partial x + \partial(6.49)/\partial y$ and use (6.12) to eliminate the ageostrophic wind, we obtain the **Q** vector form of the omega equation:

$$\sigma\,\nabla^2\omega + f_0^2\frac{\partial^2\omega}{\partial p^2} = -2\nabla\bullet\mathbf{Q} + f_0\beta\frac{\partial v_g}{\partial p} - \frac{\kappa}{p}\nabla^2 J \tag{6.53}$$

where

$$\mathbf{Q} \equiv (Q_1, Q_2) = \left(-\frac{R}{p}\frac{\partial \mathbf{V}_g}{\partial x}\bullet\nabla T, -\frac{R}{p}\frac{\partial \mathbf{V}_g}{\partial y}\bullet\nabla T\right) \tag{6.54}$$

Equation (6.54) shows that vertical motion is forced by the sum of the divergence of **Q**, the Laplacian of the diabatic heating, and a term related to the β effect that is generally small for synoptic-scale motions. Unlike the traditional form of the omega equation, the **Q** vector form does not have forcing terms that partly cancel. The forcing of ω for adiabatic flow can be represented simply by the pattern of the **Q** vector. By the arguments of the last subsection, the left-hand side in (6.54) is proportional to the vertical velocity. Hence, a convergent **Q** forces ascent, and a divergent **Q** forces descent.

The direction and magnitude of the **Q** vector at a given point on a weather map can be estimated by referring the motion to a Cartesian coordinate system in which the x axis is parallel to the local isotherm with cold air on the left. Then (6.51) can be simplified to give

$$\mathbf{Q} = -\frac{R}{p}\left(\frac{\partial T}{\partial y}\right)\left(\frac{\partial v_g}{\partial x}\mathbf{i} - \frac{\partial u_g}{\partial x}\mathbf{j}\right)$$

where we have again used the fact that $\partial u_g/\partial x = -\partial v_g/\partial y$. From the rules for cross multiplication of unit vectors, the above expression for **Q** can be rewritten as

$$\mathbf{Q} = -\frac{R}{p}\left|\frac{\partial T}{\partial y}\right|\left(\mathbf{k} \times \frac{\partial \mathbf{V}_g}{\partial x}\right) \tag{6.55}$$

Thus, the **Q** vector can be obtained by evaluating the vectorial change of $\mathbf{V}_g$ along the isotherm (with cold air on the left), rotating this change vector by 90° clockwise, and multiplying the resulting vector by $|\partial T/\partial y|$.

The **Q** vector, and hence the forcing of vertical motion, can be estimated with the aid of (6.55) from observations of Φ and T on a single isobaric surface. Examples for two simple cases, both of which have temperature decreasing toward the north, are shown in Figs. 6.13 and 6.14. Figure 6.13 shows an idealized pattern of cyclones and anticyclones in a slightly perturbed westerly thermal wind. Near the center of the low, the geostrophic wind change moving eastward along the isotherm (i.e., with cold air to the left) is from northerly to southerly. Thus the geostrophic wind change vector points northward, and a 90° clockwise rotation produces a **Q** vector parallel to the thermal wind. In the highs, by the same reasoning, the **Q** vectors are antiparallel to the thermal wind. The pattern of $\nabla \bullet \mathbf{Q}$ thus yields descent in the region of cold air advection west of the trough, and ascent in the warm air advection region east of the trough.

In the situation shown in Fig. 6.14, the geostrophic flow is confluent so that the geostrophic wind increases eastward along the isotherms. In this case the vectorial

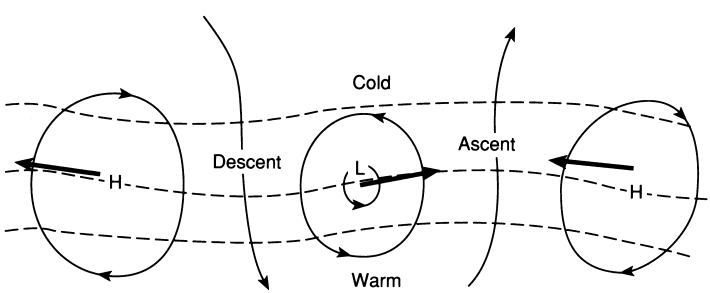


Fig. 6.13 **Q** vectors (bold arrows) for an idealized pattern of isobars (solid) and isotherms (dashed) for a family of cyclones and anticyclones. (After Sanders and Hoskins, 1990.)

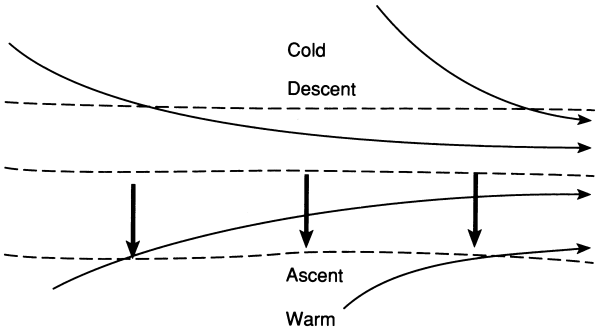


Fig. 6.14 Orientation of **Q** vectors (bold arrows) for confluent (jet entrance) flow. Dashed lines are isotherms. (After Sanders and Hoskins, 1990.)

change in $\mathbf{V}_g$ is parallel to the isotherms so that the **Q** vectors are normal to the isotherms and are directed up the temperature gradient. Again rising motion occurs where the **Q** vectors are convergent. Because such rising motion must imply vorticity stretching in the column below, cyclonic vorticity will tend to increase below a region of upper level convergent **Q** vectors.

6.4.3 The Ageostrophic Circulation

In the traditional form of quasi-geostrophic theory given in Section 6.3, the ageostrophic velocity component is not explicitly determined. Rather, its role in the secondary vertical circulation is implicitly included through diagnostic determination of the ω vertical motion field. Some dynamical aspects of the ageostrophic motion are not, however, obvious from the analysis of vertical motion alone. In particular, in some synoptic situations, advection by the ageostrophic wind may be important in the evolution of the temperature and vorticity fields.

Because the ageostrophic wind generally has both irrotational and nondivergent components, the total ageostrophic flow field cannot be obtained from knowledge of the divergence alone. Rather, it is necessary to use the quasi-geostrophic momentum equation (6.11). If for simplicity we neglect the β effect and solve (6.11) for the ageostrophic wind, we obtain

$$\mathbf{V}_a = \frac{1}{f_0}\mathbf{k} \times \frac{D_g\mathbf{V}_g}{Dt} = f_0^{-1}\left(\mathbf{k} \times \frac{\partial \mathbf{V}_g}{\partial t}\right) + f_0^{-1}\mathbf{k} \times \left[\left(\mathbf{V}_g \bullet \nabla\right)\mathbf{V}_g\right] \qquad (6.56)$$

which shows that in the Northern Hemisphere the ageostrophic wind vector is directed to the left of the geostrophic acceleration following the geostrophic motion.

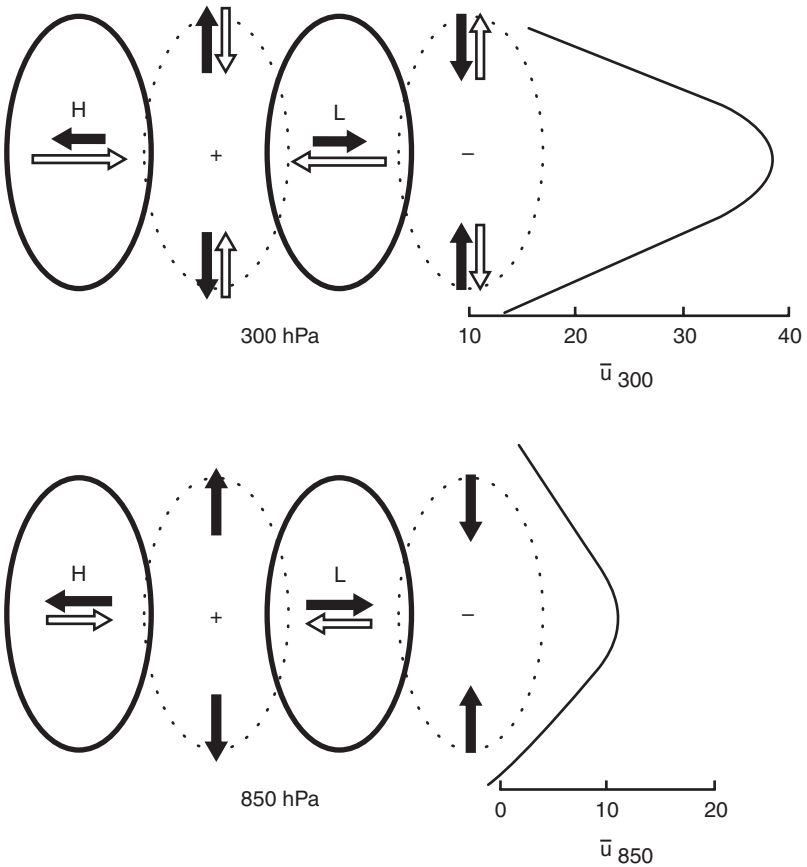


Fig. 6.15 Isallobaric and advective contributions to the ageostrophic wind for baroclinic waves in westerlies. Solid ellipses indicate perturbation geopotential patterns at 300 and 850 hPa. Dashed ellipses show a geopotential tendency pattern with positive and negative tendencies indicated by + and – signs, respectively. The mean zonal flow distribution in which the waves are embedded is indicated on the right for each level. Solid arrows show the isallobaric part of the ageostrophic wind, and open arrows show the advective part. (Adapted from Lim et al., 1991.)

The forcing of the ageostrophic wind can be divided conveniently into the two terms on the right in (6.56). The first term is referred to as the *isallobaric* wind. It can be shown easily to be proportional to the gradient of the geopotential tendency (see Problem 13) and is directed down the gradient of the tendency field. This contribution to $\mathbf{V}_a$ is shown schematically for a baroclinic wave disturbance by the black arrows in Fig. 6.15. The second term on the right in (6.54) may be called

the advective part of the ageostrophic wind. For baroclinic waves in the jetstream the advective term is dominated by zonal advection so that

$$\mathbf{k} \times \left[\left(\mathbf{V}_g \bullet \nabla\right) \mathbf{V}_g\right] \approx \overline{u} \frac{\partial}{\partial x} \left(\mathbf{k} \times \mathbf{V}_g\right) = -f_0^{-1} \overline{u} \frac{\partial}{\partial x} \left(\nabla \Phi\right)$$

where $\overline{u}$ is the mean zonal flow, and we have used the definition of the geostrophic wind (6.7). The advective contribution to the ageostrophic flow is shown by the open arrows in Fig. 6.15. Note that due to the strong jetstream at 300 hPa at the center of the waves the advective contribution dominates over the isallobaric contribution. On the flanks of the waves at 300 hPa the two contributions are of comparable amplitude so that the net ageostrophic wind is small. At the 850-hPa level, however, the two contributions nearly cancel at the center of the perturbations, whereas the advective contribution is nearly zero on the flanks. The net result is that the ageostrophic motion for baroclinic waves is primarily zonal in the upper troposphere and primarily meridional in the lower troposphere.

6.5 IDEALIZED MODEL OF A BAROCLINIC DISTURBANCE

Section 6.2 showed that for synoptic-scale systems the fields of vertical motion and geopotential tendency are determined to a first approximation by the three-dimensional distribution of geopotential. The results of our diagnostic analyses using the geopotential tendency and omega equations can now be combined to illustrate the essential structural characteristics of a developing baroclinic wave.

In Fig. 6.16 the relationship of the vertical motion field to the 500- and 1000-hPa geopotential fields is illustrated schematically for a developing baroclinic wave. Also indicated are the physical processes that give rise to the vertical circulation in various regions.

Additional structural features, including those that can be diagnosed with the tendency equation, are summarized in Table 6.1. In Table 6.1 the signs of various physical parameters are indicated for vertical columns located at the position of (A) the 500-hPa trough, (B) the surface low, and (C) the 500-hPa ridge. It can be seen from Table 6.1 that in all cases the vertical motion and divergence fields act to keep the temperature changes hydrostatic and vorticity changes geostrophic in order to preserve thermal wind balance.

Following the nomenclature of Chapter 5, we may regard the vertical and divergent ageostrophic motions as constituting a secondary circulation imposed by the simultaneous constraints of geostrophic and hydrostatic balance. The secondary circulation described in this chapter is, however, completely independent of the circulation driven by boundary layer pumping. In fact, it is observed that in midlatitude synoptic-scale systems, the vertical velocity forced by frictional convergence in the boundary layer is generally much smaller than the vertical velocity due to

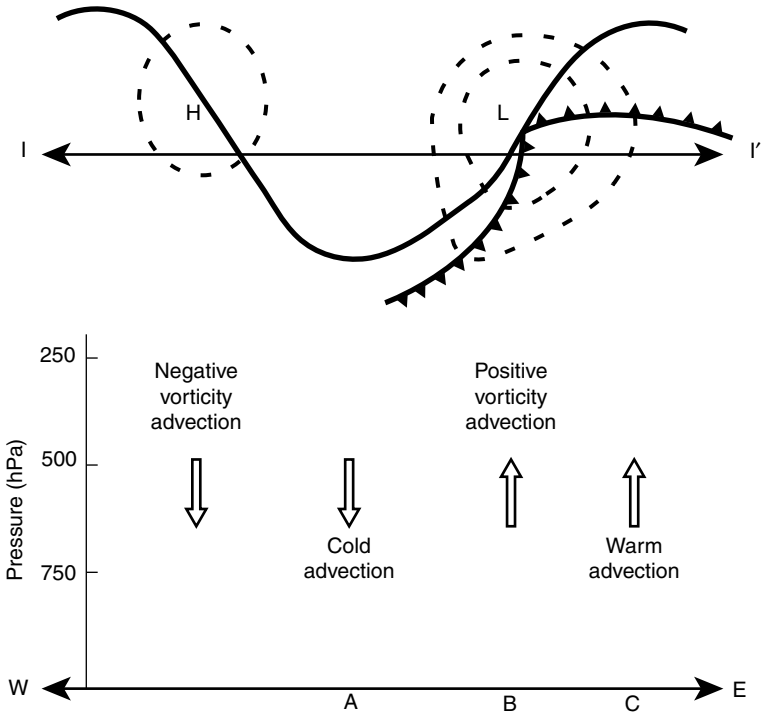


Fig. 6.16 Secondary circulation associated with a developing baroclinic wave: (top) schematic 500-hPa contour (solid line), 1000-hPa contours (dashed lines), and surface fronts; (bottom) vertical profile through line II' indicating the vertical motion field.

differential vorticity advection. For this reason we have neglected boundary layer friction in developing the equations of the quasi-geostrophic system.

It is also of interest to note that the secondary circulation in a developing baroclinic system always acts to oppose the horizontal advection fields. Thus, the divergent motions tend partly to cancel the vorticity advection, and the adiabatic temperature changes due to vertical motion tend to cancel partly the thermal advection. This tendency of the secondary flow to cancel partly the advective changes has important implications for the flow evolution. These are discussed in Chapter 8.

It should now be clear that a secondary divergent circulation is necessary to satisfy the twin constraints of geostrophic and hydrostatic balance for a baroclinic system. Without such a circulation, geostrophic advection tends to destroy the thermal wind balance. The secondary circulation is itself forced, however, by slight departures from geostrophy. Referring again to Fig. 6.16, we see that in the region of the upper-level trough (column A), cold advection causes the geopotential height to fall and thus intensifies the horizontal pressure gradient. The wind, therefore,

Table 6.1 *Characteristics of a Developing Baroclinic Disturbance*

Physical parameter	A 500-hPa trough	B Surface low	C 500-hPa ridge
$\partial(\delta\Phi)/\partial t$ (500–1000 hPa)	Negative (thickness advection partly canceled by adiabatic warming)	Negative (adiabatic cooling)	Positive (thickness advection partly canceled by adiabatic cooling)
w (500 hPa)	Negative	Positive	Positive
$\partial\Phi/\partial t$ (500 hpa)	Negative (differential thickness advection)	Negative (vortity advection)	Positive (differential thickness advection)
$\partial\zeta_g/\partial t$ (1000 hPa)	Negative (divergence)	Positive (convergence)	Positive (convergence)
$\partial\zeta_g/\partial t$ (500 hPa)	Positive (convergence)	Positive (advection partly canceled by divergence)	Negative (divergence)

becomes slightly subgeostrophic and experiences an acceleration across the isobars toward lower pressure. This cross isobaric ageostrophic wind component is responsible for the convergence spins up the vorticity in the upper troposphere so that it adjusts geostrophically to the new geopotential distribution. In terms of the momentum balance, the cross isobaric flow is accelerated by the pressure gradient force so that the wind speed adjusts back toward geostrophic balance. In the region of the upper level ridge, analogous arguments apply, but in this case the ageostrophic flow leads to a divergent secondary circulation. In both cases, as shown in Chapter 8, the ageostrophic flow toward lower pressure is associated with conversion of energy from potential energy to kinetic energy.

PROBLEMS

6.1. Show that the static stability parameter $\sigma = -\alpha\partial \ln\theta/\partial p$ may be written in terms of Φ as

$$\sigma = \frac{\partial^2\Phi}{\partial p^2} - \frac{1}{p}\frac{\partial\Phi}{\partial p}\left(\frac{R}{c_p} - 1\right) = \frac{1}{p^2}\left(\frac{\partial}{\partial \ln p} - \frac{R}{c_p}\right)\frac{\partial\Phi}{\partial \ln p}$$

6.2. Show that for an isothermal atmosphere σ, as defined in Problem 6.1, varies inversely as the square of the pressure.

6.3. Suppose that on the 500-hPa surface the relative vorticity at a certain location at 45° N latitude is increasing at a rate of $3 \times 10^{-6}\ \mathrm{s}^{-1}$ per 3 h. The wind

is from the southwest at 20 m s^{-1} and the relative vorticity decreases toward the northeast at a rate of 4×10^{-6} s^{-1} per 100 km. Use the quasi-geostrophic vorticity equation to estimate the horizontal divergence at this location on a β-plane.

6.4. Given the following expression for the geopotential field

$$\Phi = \Phi_0(p) + cf_0 \left\{ -y[\cos(\pi p/p_0) + 1] + k^{-1} \sin k(x - ct) \right\}$$

where Φ_0 is a function of p alone, c is a constant speed, k a zonal wave number, and $p_0 = 1000$ hPa (a) Obtain expressions for the corresponding geostrophic wind and relative vorticity fields. (b) Obtain an expression for the relative vorticity advection. (c) Use the quasi-geostrophic vorticity equation to obtain the horizontal divergence field consistent with this Φ field. (Assume that $df/dy = 0$.) (d) Assuming that $\omega(p_0) = 0$, obtain an expression for $\omega(x, y, p, t)$ by integrating the continuity equation with respect to pressure. (e) Sketch the geopotential fields at 750 and 250 hPa. Indicate regions of maximum divergence and convergence and positive and negative vorticity advection.

6.5. For the geopotential distribution of Problem 6.4, obtain an expression for ω by using the adiabatic form of the thermodynamic energy equation (6.13b). Assume that σ is a constant. For what value of k does this expression for ω agree with that obtained in Problem 6.4?

6.6. As an additional check on the results of Problems 6.4 and 6.5, use the approximate omega equation (6.36) to obtain an expression for ω. Note that the three expressions for ω agree only for one value of k. Thus, for given values of σ and f_0, the geopotential field $\Phi(x, y, p, t)$ of Problem 6.4 is consistent with quasi-geostrophic dynamics only for one value of the zonal wave number.

6.7. Suppose that the geopotential distribution at a certain time has the form

$$\Phi(x, y, p) = \Phi_0(p) - f_0 U_0 y \cos(\pi p/p_0) + f_0 c k^{-1} \sin kx$$

where U_0 is a constant zonal speed and all other constants are as in Problem 6.4. Assuming that f and σ are constants, show by evaluating the terms on the right-hand side of the tendency equation (6.23) that $\chi = 0$ provided that $k^2 = \sigma^{-1}(f_0 \pi/p_0)^2$. Make qualitative sketches of the geopotential fields at 750 and 250 hPa for this case. Indicate regions of maximum positive and negative vorticity advection at each level. (Note: the wavelength corresponding to this value of k is called the *radius of deformation*.)

6.8. For the geopotential field of Problem 6.7, use the omega equation (6.36) to find an expression for ω for the conditions in which $\chi = 0$. *Hint*: Let $\omega = W_0 \cos kx \sin(\pi p/p_0)$ where W_0 is a constant to be determined. Sketch a cross section in the x, p plane indicating trough and ridge lines, vorticity maxima and minima, vertical motion and divergence patterns, and locations of maximum cold and warm temperature advection.

6.9. Given the following expression for the geopotential field,

$$\Phi(x, y, p) = \Phi_0(p) + f_0\left[-Uy + k^{-1}V\cos(\pi p/p_0)\sin k(x - ct)\right]$$

where U, V, and c are constant speeds, use the quasi-geostrophic vorticity equation (6.19) to obtain an estimate of ω. Assume that $df/dy = \beta$ is a constant (*not zero*) and that ω vanishes for $p = p_0$.

6.10. For the conditions given in Problem 6.9, use the adiabatic thermodynamic energy equation to obtain an alternative estimate for ω. Determine the value of c for which this estimate of ω agrees with that found in Problem 6.9.

6.11. For the conditions given in Problem 6.9, use the omega equation (6.36) to obtain an expression for ω. Verify that this result agrees with the results of Problems 6.9 and 6.10. Sketch the phase relationship between Φ and ω at 250 and 750 hPa. What is the amplitude of ω if $\beta = 2 \times 10^{-11}\ \mathrm{m^{-1}\,s^{-1}}$, $U = 25\ \mathrm{m\,s^{-1}}$, $V = 8\ \mathrm{m\,s^{-1}}$, $k = 2\pi/(10^4\ \mathrm{km})$, $f_0 = 10^{-4}\ \mathrm{s^{-1}}$, $\sigma = 2 \times 10^{-6}\ \mathrm{Pa^{-2}\,m^2\,s^{-2}}$, and $p_0 = 1000$ hPa?

6.12. Compute the **Q** vector distributions corresponding to the geopotential fields given in Problems 6.4 and 6.7.

6.13. Show that the isallobaric wind may be expressed in the form

$$\mathbf{V}_{isall} = -f_0^{-2}\boldsymbol{\nabla}\chi$$

where $\chi = \partial\Phi/\partial t$.

MATLAB EXERCISES

M6.1. Suppose that the 500-hPa geopotential field has the following structure (expressed in $\mathrm{m^2\,s^{-1}}$):

$$\Phi = 5.5 \times 10^4 - U_0 f_0 y + V f_0 k^{-1}\sin[k(x - ct)]\cos ly$$

where $U_0 = 30$ m/s, $f_0 = 1.0 \times 10^{-4}\ \mathrm{s^{-1}}$, $V = 25\ \mathrm{m\,s^{-1}}$, $c = 25\ \mathrm{m\,s^{-1}}$, $k = (2\pi/L_x)$, and $l = 2\pi/L_y$ where $L_x = 6 \times 10^6$ m and $L_y =$

4×10^6 m. The MATLAB script **Chapter_6_plot1.m** produces a contour plot of the corresponding geopotential height field and provides a template that can be modified to produce vector and contour plots of various fields computed for this geopotential distribution in the domain $-3000\,\text{km} \leq x \leq +3000\,\text{km}$, $-1000 \leq y \leq 1000\,\text{km}$. The template shows a contour plot of Φ and a quiver plot of the zonal mean wind U_0. Derive the symbolic expression for each of the fields in parts (a)–(d) and plot the horizontal distributions in MATLAB. (a) Solve for the x and y components of the geostrophic wind; plot as a quiver plot for time $t = 0$. (b) Solve for the geostrophic relative vorticity and plot as a contour plot for time $t = 0$. (c) Solve for the relative vorticity advection by the mean zonal wind and plot as a contour plot for time $t = 0$. (d) Assuming that $\beta = 1.67 \times 10^{-11}\ \text{m}^{-1}\ \text{s}^{-1}$, solve for the planetary vorticity advection and plot as a contour plot for time $t = 0$.

M6.2. For the situation of Problem M6.1: (a) Use the quasi-geostrophic vorticity equation to solve for the divergence field at 500 hPa and plot the vorticity tendency, absolute vorticity advection, and divergence fields as contour plots for time $t = 0$ using MATLAB. (b) What is the phase difference between the divergence field and the vorticity field? Explain in physical terms (referring to the vorticity balance) why the divergence field has the phase relationship to vorticity that you observe.

M6.3. Equations (6.16) and (6.17) can be solved for the ageostrophic wind (u_a, v_a) if the geostrophic wind components are known. For the situation of Problem M6.1, solve for the ageostrophic wind. For simplicity you may approximate the advective terms by letting $\mathbf{V}_g \bullet \nabla u_g \approx U \partial u_g / \partial x$ and $\mathbf{V}_g \bullet \nabla v_g \approx U \partial v_g / \partial x$. Note that the ageostrophic wind has both a divergent and a nondivergent part. (The nondivergent part of the ageostrophic wind in this example is proportional to the β-effect.) Compute the divergence of the ageostrophic wind and plot as a contour map in MATLAB. Compare with the results of Problem M6.2. Plot the divergent part of the ageostrophic wind as a quiver plot in MATLAB for time $t = 0$. Compute the vorticity of the ageostrophic wind and plot as a contour map in MATLAB. Overlay this plot with a quiver plot showing the nondivergent part of the ageostrophic wind. Explain why vorticity has an ageostrophic component in this case.

M6.4. The MATLAB script **pv_anomaly.m** contains code to compute and plot the vertical distribution of the geopotential anomaly induced by quasi-geostrophic potential vorticity advection confined to above the 250-hPa level. The example given in Fig. 6.11 is for an infinite meridional scale. Modify the code to compute the dependence of the ratio of the response at the 600-hPa level to the response at the 100-hPa level for zonal wavelength 6000 km as the meridional scale is varied from 12,000 to 1000 km. Plot your results as a line graph in MATLAB.

M6.5. Consider a modification of the *stationary* geopotential field given in Problem 6.7:

$$\Phi(x, y, p) = \Phi_0(p) - f_0 U_0 y \cos\left(\pi p/p_0\right) + f_0 V k^{-1} \sin(kx)\cos(ly)$$

where Φ_0 is the standard atmosphere geopotential, $f_0 = 1.0 \times 10^{-4}\,\mathrm{s}^{-1}$, $U_0 = 35\,\mathrm{m\,s}^{-1}$, $V = 20\,\mathrm{m\,s}^{-1}$, $k = (2\pi/L)$ with $L = 6.0 \times 10^6\,\mathrm{m}$, $l = 1.5k$, and $p_0 = 1000\,\mathrm{hPa}$. Note that in this case the mean zonal wind is negative in the lower troposphere and positive in the upper troposphere. Again the domain of interest is $-3000\,\mathrm{km} \le x \le +3000\,\mathrm{km}$, $-1000 \le y \le 1000\,\mathrm{km}$. (a) Neglecting the β-effect, show by evaluating the terms on the right-hand side of the tendency equation (6.23) that the tendency vanishes provided that $K^2 \equiv k^2 + l^2 = \sigma^{-1}(f_0\pi/p_0)^2$ (where the static stability σ is treated here as a constant whose magnitude is given by this formula). (b) Use MATLAB to plot maps of the geopotential height ($Z = \Phi/g$) distributions at the 750- and 250-hPa levels assuming that the mean *geopotential height* is 2500 m at 750 hPa and 10,000 m at 250 hPa. (c) For the conditions given above, find an expression for the horizontal divergence and overlay maps of the divergence fields at 750 and 250 hPa on the geopotential height maps of part (b). Explain the phase relationships between the divergence and height fields at each level.

M6.6. For the geopotential distribution of the previous problem, use the traditional omega equation (6.34) to find an expression for ω for the conditions in which the geopotential tendency vanishes $\left[\text{i.e., for } K^2 = \sigma^{-1}(f_0\pi/p_0)^2\right]$ as shown in the problem earlier. The best approach is to use a trial solution of the form $\omega = W_0 \cos(kx)\cos(ly)\sin\left(\pi p/p_0\right)$ where W_0 is a constant to be determined. If you substitute this trial form into the left side of (6.34) you can solve for W_0. (a) Use MATLAB to plot vertical cross sections of the divergence, the vertical motion, and the vorticity fields for this situation at $y = 0$. The script **Chapter_6_plot2.m** provides a template for contour plotting vertical cross sections for this situation. (b) Derive an expression for the thickness advection for this case. Use MATLAB to plot a vertical cross section of the thickness advection. Describe the phase relationship between thickness advection and vertical velocity and relate these to the balance in the thermodynamic energy equation (6.21).

Suggested References

Wallace and Hobbs, *Atmospheric Science: An Introductory Survey,* has an excellent introductory level description of the observed structure and evolution of midlatitude synoptic-scale disturbances.

Bluestein, *Synoptic-Dynamic Meteorology in Midlatitudes, Vol. II,* has a comprehensive treatment of midlatitude synoptic disturbances at the graduate level.

Blackburn (1985) *Interpretation of Ageostrophic Winds and Implications for Jetstream Maintenance,* discusses the differences between variable f (VF) and constant f (CF) ageostrophic motion.

Durran and Snellman (1987) illustrate the application of both traditional and **Q** vector forms of the omega equation in diagnosing the vertical motion of an observed system.

Sanders and Hoskins (1990) show how the distribution of **Q** vectors can easily be visualized on realistic weather maps.

Pedlosky, *Geophysical Fluid Dynamics*, *2nd Edition,* presents a detailed formal derivation of the quasi-geostrophic system with applications to both the atmosphere and the oceans.

CHAPTER 7

Atmospheric Oscillations: Linear Perturbation Theory

Chapter 13 discusses numerical techniques for solving the equations governing large-scale atmospheric motions. If the objective is to produce an accurate forecast of the circulation at some future time, a detailed numerical model based on the primitive equations and including processes such as latent heating, radiative transfer, and boundary layer drag should produce the best results. However, the inherent complexity of such a model generally precludes any simple interpretation of the physical processes that produce the predicted circulation. If we wish to gain physical insight into the fundamental nature of atmospheric motions, it is helpful to employ simplified models in which certain processes are omitted and compare the results with those of more complete models. This is, of course, just what was done in deriving the quasi-geostrophic model. However, the quasi-geostrophic potential vorticity equation is still a complicated nonlinear equation that must be solved numerically. It is difficult to gain an appreciation for the processes that produce the wave-like character observed in many meteorological disturbances through the study of numerical integrations alone.

This chapter discusses the *perturbation method*, a simple technique that is useful for qualitative analysis of atmospheric waves. We then use this method to examine

several types of waves in the atmosphere. In Chapter 8 the perturbation theory is used to study the development of synoptic-wave disturbances.

7.1 THE PERTURBATION METHOD

In the perturbation method, all field variables are divided into two parts, a *basic state* portion, which is usually assumed to be independent of time and longitude, and a perturbation portion, which is the local deviation of the field from the basic state. Thus, for example, if $\overline{u}$ designates a time and longitude-averaged zonal velocity and u' is the deviation from that average, then the complete zonal velocity field is $u(x,t) = \overline{u} + u'(x,t)$. In that case, for example, the inertial acceleration $u\partial u/\partial x$ can be written

$$u\frac{\partial u}{\partial x} = \left(\overline{u} + u'\right)\frac{\partial}{\partial x}\left(\overline{u} + u'\right) = \overline{u}\frac{\partial u'}{\partial x} + u'\frac{\partial u'}{\partial x}$$

The basic assumptions of perturbation theory are that the basic state variables must themselves satisfy the governing equations when the perturbations are set to zero, and the perturbation fields must be small enough so that all terms in the governing equations that involve products of the perturbations can be neglected. The latter requirement would be met in the above example if $|u'/\overline{u}| \ll 1$ so that

$$\left|\overline{u}\partial u'/\partial x\right| \gg \left|u'\partial u'/\partial x\right|$$

If terms that are products of the perturbation variables are neglected, the nonlinear governing equations are reduced to linear differential equations in the perturbation variables in which the basic state variables are specified coefficients. These equations can then be solved by standard methods to determine the character and structure of the perturbations in terms of the known basic state. For equations with constant coefficients the solutions are sinusoidal or exponential in character. Solution of perturbation equations then determines such characteristics as the propagation speed, vertical structure, and conditions for growth or decay of the waves. The perturbation technique is especially useful in studying the stability of a given basic state flow with respect to small superposed perturbations. This application is the subject of Chapter 8.

7.2 PROPERTIES OF WAVES

Wave motions are oscillations in field variables (such as velocity and pressure) that propagate in space and time. In this chapter we are concerned with linear sinusoidal wave motions. Many of the mechanical properties of such waves are also features of a familiar system, the linear harmonic oscillator. An important

property of the harmonic oscillator is that the period, or time required to execute a single oscillation, is independent of the amplitude of the oscillation. For most natural vibratory systems, this condition holds only for oscillations of sufficiently small amplitude. The classical example of such a system is the simple pendulum (Fig. 7.1) consisting of a mass M suspended by a massless string of length l, free to perform small oscillations about the equilibrium position $\theta = 0$. The component of the gravity force parallel to the direction of motion is $-Mg\sin\theta$. Thus, the equation of motion for the mass M is

$$Ml\frac{d^2\theta}{dt^2} = -Mg\sin\theta$$

Now for small displacements, $\sin\theta \approx \theta$ so that the governing equation becomes

$$\frac{d^2\theta}{dt^2} + \nu^2\theta = 0 \tag{7.1}$$

where $\nu^2 \equiv g/l$. The harmonic oscillator equation (7.1) has the general solution

$$\theta = \theta_1\cos\nu t + \theta_2\sin\nu t = \theta_0\cos(\nu t - \alpha)$$

where $\theta_1, \theta_2, \theta_0$, and α are constants determined by the initial conditions (see Problem 7.1) and ν is the frequency of oscillation. The complete solution can thus be expressed in terms of an amplitude θ_0 and a phase $\phi(t) = \nu t - \alpha$. The phase varies linearly in time by a factor of 2π radians per wave period.

Propagating waves can also be characterized by their amplitudes and phases. In a propagating wave, however, phase depends not only on time, but on one or more space variables as well. Thus, for a one-dimensional wave propagating in the x direction, $\phi(x,t) = kx - \nu t - \alpha$. Here the *wave number*, k, is defined as 2π divided by the wavelength. For propagating waves the phase is constant for an

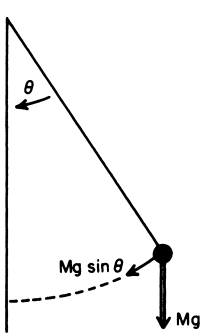


Fig. 7.1 A simple pendulum.

observer moving at the *phase speed* $c \equiv \nu/k$. This may be verified by observing that if phase is to remain constant following the motion,

$$\frac{D\phi}{Dt} = \frac{D}{Dt}(kx - \nu t - \alpha) = k\frac{Dx}{Dt} - \nu = 0$$

Thus, $Dx/Dt = c = \nu/k$ for phase to be constant. For $\nu > 0$ and $k > 0$ we have $c > 0$. In that case if $\alpha = 0$, $\phi = k(x - ct)$, so that x must increase with increasing t for ϕ to remain constant. Phase then propagates in the positive direction as illustrated for a sinusoidal wave in Fig. 7.2.

7.2.1 Fourier Series

The representation of a perturbation as a simple sinusoidal wave might seem an oversimplification since disturbances in the atmosphere are never purely sinusoidal. It can be shown, however, that any reasonably well-behaved function of longitude can be represented in terms of a zonal mean plus a *Fourier* series of sinusoidal components:

$$f(x) = \sum_{s=1}^{\infty} (A_s \sin k_s x + B_s \cos k_s x) \tag{7.2}$$

where $k_s = 2\pi s/L$ is the *zonal* wave number (units m^{-1}), L is the distance around a latitude circle, and s, the *planetary* wave number, is an integer designating the number of waves around a latitude circle. The coefficients A_s are calculated by multiplying both sides of (7.2) by $\sin(2\pi nx/L)$, where n is an integer, and integrating around a latitude circle. Applying the orthogonality relationships

$$\int_0^L \sin\frac{2\pi sx}{L} \sin\frac{2\pi nx}{L} dx = \left\{ \begin{array}{ll} 0, & s \neq n \\ L/2, & s = n \end{array} \right\}$$

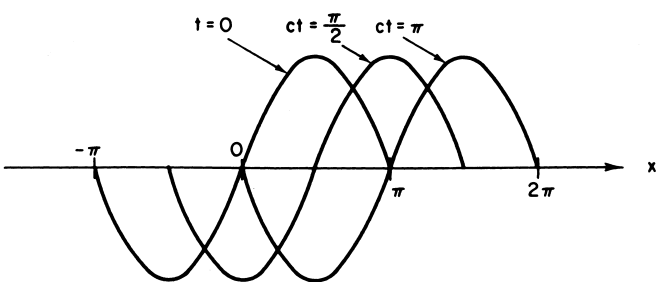


Fig. 7.2 A sinusoidal wave traveling in the positive x direction at speed c. (Wave number is assumed to be unity.)

we obtain

$$A_s = \frac{2}{L}\int_0^L f(x)\sin\frac{2\pi sx}{L}dx$$

In a similar fashion, multiplying both sides in (7.2) by $\cos(2\pi nx/L)$ and integrating gives

$$B_s = \frac{2}{L}\int_0^L f(x)\cos\frac{2\pi sx}{L}dx$$

A_s and B_s are called the *Fourier coefficients*, and

$$f_s(x) = A_s \sin k_s x + B_s \cos k_s x \tag{7.3}$$

is called the sth *Fourier component* or sth harmonic of the function $f(x)$. If the Fourier coefficients are computed for, say, the longitudinal dependence of the (observed) geopotential perturbation, the largest amplitude Fourier components will be those for which s is close to the observed number of troughs or ridges around a latitude circle. When only qualitative information is desired, it is usually sufficient to limit the analysis to a single typical Fourier component and assume that the behavior of the actual field will be similar to that of the component. The expression for a Fourier component may be written more compactly by using complex exponential notation. According to the Euler formula

$$\exp(i\phi) = \cos\phi + i\sin\phi$$

where $i \equiv (-1)^{1/2}$ is the imaginary unit. Thus, we can write

$$\begin{aligned} f_s(x) &= \mathrm{Re}[C_s \exp(ik_s x)] \\ &= \mathrm{Re}[C_s \cos k_s x + iC_s \sin k_s x] \end{aligned} \tag{7.4}$$

where Re[ ] denotes "real part of" and $C_s$is a complex coefficient. Comparing (7.3) and (7.4) we see that the two representations of $f_s(x)$ are identical, provided that

$$B_s = \mathrm{Re}[C_s] \quad \text{and} \quad A_s = -\mathrm{Im}[C_s]$$

where Im[ ] stands for "imaginary part of." This exponential notation will generally be used for applications of the perturbation theory below and also in Chapter 8.

7.2.2 Dispersion and Group Velocity

A fundamental property of linear oscillators is that the frequency of oscillation ν depends only on the physical characteristics of the oscillator, not on the motion itself. For propagating waves, however, ν generally depends on the wave number of the perturbation as well as the physical properties of the medium. Thus, because

$c = \nu/k$, the phase speed also depends on the wave number except in the special case where $\nu \propto k$. For waves in which the phase speed varies with k, the various sinusoidal components of a disturbance originating at a given location are at a later time found in different places, that is, they are dispersed. Such waves are referred to as *dispersive*, and the formula that relates ν and k is called a *dispersion relationship*. Some types of waves, such as acoustic waves, have phase speeds that are independent of the wave number. In such *nondispersive* waves a spatially localized disturbance consisting of a number of Fourier wave components (a *wave group*) will preserve its shape as it propagates in space at the phase speed of the wave.

For dispersive waves, however, the shape of a wave group will not remain constant as the group propagates. Because the individual Fourier components of a wave group may either reinforce or cancel each other, depending on the relative phases of the components, the energy of the group will be concentrated in limited regions as illustrated in Fig. 7.3. Furthermore, the group generally broadens in the course of time, that is, the energy is *dispersed*.

When waves are dispersive, the speed of the wave group is generally different from the average phase speed of the individual Fourier components. Hence, as shown in Fig. 7.4, individual wave components may move either more rapidly or more slowly than the wave group as the group propagates along. Surface waves in deep water (such as a ship wake) are characterized by dispersion in which individual wave crests move twice as fast as the wave group. In synoptic-scale atmospheric disturbances, however, the group velocity exceeds the phase velocity. The resulting downstream development of new disturbances will be discussed later.

An expression for the *group velocity*, which is the velocity at which the observable disturbance (and hence the energy) propagates, can be derived as follows: We consider the superposition of two horizontally propagating waves of equal amplitude but slightly different wavelengths with wave numbers and frequencies differing by $2\delta k$ and $2\delta\nu$, respectively. The total disturbance is thus

$$\Psi(x,t) = \exp\{i[(k+\delta k)x - (\nu+\delta\nu)t]\} + \exp\{i[(k-\delta k)x - (\nu-\delta\nu)t]\}$$

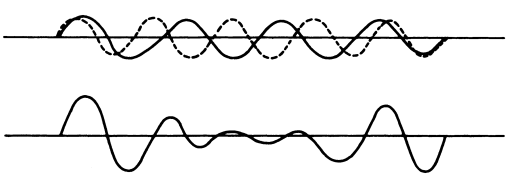


Fig. 7.3 Wave groups formed from two sinusoidal components of slightly different wavelengths. For nondispersive waves, the pattern in the lower part of the diagram propagates without change of shape. For dispersive waves, the shape of the pattern changes in time.

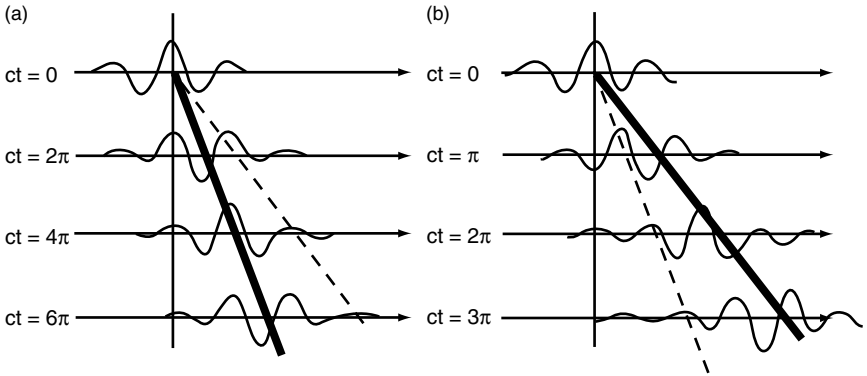


Fig. 7.4 Schematic showing propagation of wave groups: (a) group velocity less than phase speed and (b) group velocity greater than phase speed. Heavy lines show group velocity, and light lines show phase speed.

where for brevity the Re[ ] notation in (7.4) is omitted, and it is understood that only the real part of the right-hand side has physical meaning. Rearranging terms and applying the Euler formula gives

$$\begin{aligned}\Psi &= \left[e^{i(\delta kx-\delta\nu t)} + e^{-i(\delta kx-\delta\nu t)}\right] e^{i(kx-\nu t)} \\ &= 2\cos(\delta kx - \delta\nu t)\, e^{i(kx-\nu t)}\end{aligned} \tag{7.5}$$

The disturbance (7.5) is the product of a high-frequency *carrier wave* of wavelength $2\pi/k$ whose phase speed, ν/k, is the average for the two Fourier components, and a low-frequency *envelope* of wavelength $2\pi/\delta k$ that travels at the speed $\delta\nu/\delta k$. Thus, in the limit as $\delta k \to 0$, the horizontal velocity of the envelope, or *group velocity*, is just

$$c_{gx} = \partial\nu/\partial k$$

Thus, the wave energy propagates at the group velocity. This result applies generally to arbitrary wave envelopes provided that the wavelength of the wave group, $2\pi/\delta k$, is large compared to the wavelength of the dominant component, $2\pi/k$.

7.3 SIMPLE WAVE TYPES

Waves in fluids result from the action of restoring forces on fluid parcels that have been displaced from their equilibrium positions. The restoring forces may be due to compressibility, gravity, rotation, or electromagnetic effects. This section considers the two simplest examples of linear waves in fluids: acoustic waves and shallow water gravity waves.

7.3.1 Acoustic or Sound Waves

Sound waves, or acoustic waves, are *longitudinal waves*. That is, they are waves in which the particle oscillations are parallel to the direction of propagation. Sound is propagated by the alternating adiabatic compression and expansion of the medium. As an example, Fig. 7.5 shows a schematic section along a tube that has a diaphragm at its left end. If the diaphragm is set into vibration, the air adjacent to it will be alternately compressed and expanded as the diaphragm moves inward and outward. The resulting oscillating pressure gradient force will be balanced by an oscillating acceleration of the air in the adjoining region, which will cause an oscillating pressure oscillation further into the tube, and so on. The result of this continual adiabatic increase and decrease of pressure through alternating compression and rarefaction is, as shown in Fig. 7.5, a sinusoidal pattern of pressure and velocity perturbations that propagates to the right down the tube. Individual air parcels do not, however, have a net rightward motion; they only oscillate back and forth while the pressure pattern moves rightward at the speed of sound.

To introduce the perturbation method we consider the problem illustrated by Fig. 7.5, that is, one-dimensional sound waves propagating in a straight pipe parallel to the x axis. To exclude the possibility of *transverse* oscillations (i.e., oscillations in which the particle motion is at right angles to the direction of phase propagation), we assume at the outset that $v = w = 0$. In addition, we eliminate all dependence on y and z by assuming that $u = u(x,t)$. With these restrictions the momentum equation, continuity equation, and thermodynamic energy equation for adiabatic motion are, respectively,

$$\frac{Du}{Dt} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0 \tag{7.6}$$

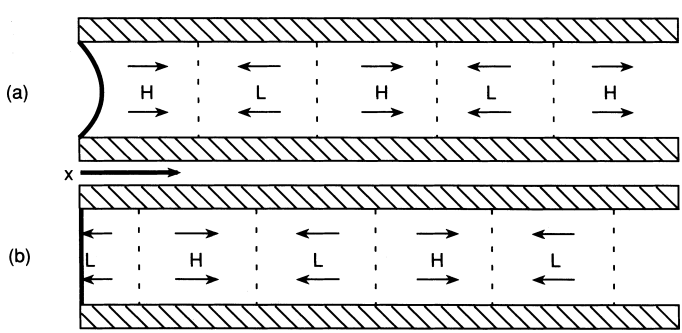


Fig. 7.5 Schematic diagram illustrating the propagation of a sound wave in a tube with a flexible diaphragm at the left end. Labels H and L designate centers of high and low perturbation pressure. Arrows show velocity perturbations. (b) The situation 1/4 period later than in (a) for propagation in the positive x direction.

$$\frac{D\rho}{Dt} + \rho\frac{\partial u}{\partial x} = 0 \tag{7.7}$$

$$\frac{D\ln\theta}{Dt} = 0 \tag{7.8}$$

where for this case $D/Dt = \partial/\partial t + u\partial/\partial x$. Recalling from (2.44) and the ideal gas law that potential temperature may be expressed as

$$\theta = \left(p/\rho R\right)\left(p_s/p\right)^{R/c_p}$$

where $p_s = 1000$ hPa, we may eliminate θ in (7.8) to give

$$\frac{1}{\gamma}\frac{D\ln p}{Dt} - \frac{D\ln\rho}{Dt} = 0 \tag{7.9}$$

where $\gamma = c_p/c_v$. Eliminating ρ between (7.7) and (7.9) gives

$$\frac{1}{\gamma}\frac{D\ln p}{Dt} + \frac{\partial u}{\partial x} = 0 \tag{7.10}$$

The dependent variables are now divided into constant basic state portions (denoted by overbars) and perturbation portions (denoted by primes):

$$\begin{aligned} u(x,t) &= \overline{u} + u'(x,t) \\ p(x,t) &= \overline{p} + p'(x,t) \\ \rho(x,t) &= \overline{\rho} + \rho'(x,t) \end{aligned} \tag{7.11}$$

Substituting (7.11) into (7.6) and (7.10) we obtain

$$\frac{\partial}{\partial t}\left(\overline{u}+u'\right) + \left(\overline{u}+u'\right)\frac{\partial}{\partial x}\left(\overline{u}+u'\right) + \frac{1}{(\overline{\rho}+\rho')}\frac{\partial}{\partial x}\left(\overline{p}+p'\right) = 0$$

$$\frac{\partial}{\partial t}\left(\overline{p}+p'\right) + \left(\overline{u}+u'\right)\frac{\partial}{\partial x}\left(\overline{p}+p'\right) + \gamma\left(\overline{p}+p'\right)\frac{\partial}{\partial x}\left(\overline{u}+u'\right) = 0$$

We next observe that provided $\left|\rho'/\overline{\rho}\right| \ll 1$ we can use the binomial expansion to approximate the density term as

$$\frac{1}{(\overline{\rho}+\rho')} = \frac{1}{\overline{\rho}}\left(1+\frac{\rho'}{\overline{\rho}}\right)^{-1} \approx \frac{1}{\overline{\rho}}\left(1-\frac{\rho'}{\overline{\rho}}\right)$$

Neglecting products of the perturbation quantities and noting that the basic state fields are constants, we obtain the linear perturbation equations[1]

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)u'+\frac{1}{\overline{\rho}}\frac{\partial p'}{\partial x}=0 \tag{7.12}$$

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)p'+\gamma\overline{p}\frac{\partial u'}{\partial x}=0 \tag{7.13}$$

Eliminating u' by operating on (7.13) with $\left(\partial/\partial t+\overline{u}\partial/\partial x\right)$ and substituting from (7.12), we get[2]

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)^2 p'-\frac{\gamma\overline{p}}{\overline{\rho}}\frac{\partial^2 p'}{\partial x^2}=0 \tag{7.14}$$

which is a form of the standard *wave equation* familiar from electromagnetic theory. A simple solution representing a plane sinusoidal wave propagating in x is

$$p'=A\exp\left[ik\left(x-ct\right)\right] \tag{7.15}$$

where for brevity we omit the Re{ } notation, but it is to be understood that only the real part of (7.15) has physical significance. Substituting the assumed solution (7.15) into (7.14), we find that the phase speed c must satisfy

$$(-ikc+ik\overline{u})^2-\left(\gamma\overline{p}/\overline{\rho}\right)(ik)^2=0$$

where we have canceled out the factor $A\exp\left[ik\left(x-ct\right)\right]$, which is common to both terms. Solving for c gives

$$c=\overline{u}\pm(\gamma\overline{p}/\overline{\rho})^{1/2}=\overline{u}\pm\left(\gamma R\overline{T}\right)^{1/2} \tag{7.16}$$

Therefore (7.15) is a solution of (7.14), provided that the phase speed satisfies (7.16). According to (7.16) the speed of wave propagation relative to the zonal current is $c-\overline{u}=\pm c_s$, where $c_s\equiv\left(\gamma R\overline{T}\right)^{1/2}$ is called the *adiabatic speed of sound*.

[1] It is not necessary that the perturbation velocity be small compared to the mean velocity for linearization to be valid. It is only required that quadratic terms in the perturbation variables be small compared to the dominant linear terms in (7.12) and (7.13).

[2] Note that the squared differential operator in the first term expands in the usual way as

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)^2=\frac{\partial^2}{\partial t^2}+2\overline{u}\frac{\partial^2}{\partial t\partial x}+\overline{u}^2\frac{\partial^2}{\partial x^2}$$

The mean zonal velocity here plays only a role of *Doppler shifting* the sound wave so that the frequency relative to the ground corresponding to a given wave number k is

$$\nu = kc = k\,(\overline{u} \pm c_s)$$

Thus, in the presence of a wind, the frequency as heard by a fixed observer depends on the location of the observer relative to the source. If $\overline{u} > 0$ the frequency of a stationary source will appear to be higher for an observer to the east (downstream) of the source ($c = \overline{u} + c_s$) than for an observer to the west (upstream) of the source $c = \overline{u} - c_s$.

7.3.2 Shallow Water Gravity Waves

As a second example of pure wave motion we consider the horizontally propagating oscillations known as shallow water waves. Shallow water gravity waves can exist only if the fluid has a free surface or an internal density discontinuity. As shown in the previous subsection, in acoustic waves the restoring force is parallel to the direction of propagation of the wave. In shallow water gravity waves, however, the restoring force is in the vertical so that it is transverse to the direction of propagation.

The mechanism for propagation of gravity waves can be understood by considering water in a channel extending in the x direction with an oscillating paddle at the origin. The back-and-forth oscillations of the paddle generate alternating upward and downward perturbations in the free surface height, which produce alternating positive and negative accelerations. These, in turn, lead to alternating patterns of fluid convergence and divergence. The net result is a sinusoidal disturbance of the free surface height, which moves toward the right, and has perturbation velocity and free surface height exactly in phase as shown in Fig. 7.6. A similar sort of disturbance could be set up moving toward the left, but in that case the velocity and free surface height perturbations would be exactly 180° out of phase.

As a specific example we consider a fluid system consisting of two homogeneous incompressible layers of differing density as shown in Fig. 7.7. Waves may propagate along the interface between the two layers. The assumption of incompressibility is sufficient to exclude sound waves, and we can thus isolate the gravity waves. If the density of the lower layer ρ_1 is greater than the density of the upper layer ρ_2, the system is stably stratified. Because both ρ_1 and ρ_2 are constants, the horizontal pressure gradient in each layer is independent of height if the pressure is hydrostatic. This may be verified by differentiating the hydrostatic approximation with respect to x:

$$\frac{\partial}{\partial z}\left(\frac{\partial p}{\partial x}\right) = -\frac{\partial \rho}{\partial x} g = 0$$

For simplicity, we assume that there is no horizontal pressure gradient in the upper layer. The pressure gradient in the lower layer can be obtained by vertical

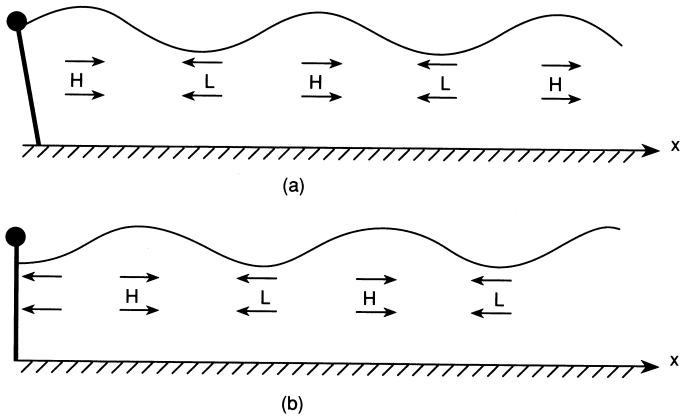


Fig. 7.6 Propagation of a surface gravity wave in a water channel generated by a paddle at the left end. Symbols as in Fig. 7.5.

integration of the hydrostatic equation. For the points A and B shown in Fig. 7.7 we find, respectively,

$$p + \delta p_1 = p + \rho_1 g \delta z = p + \rho_1 g \left(\partial h / \partial x\right) \delta x$$
$$p + \delta p_2 = p + \rho_2 g \delta z = p + \rho_2 g \left(\partial h / \partial x\right) \delta x$$

where $\partial h / \partial x$ is the slope of the interface. Taking the limit $\delta x \to 0$, we obtain the pressure gradient in the lower layer

$$\lim_{\delta x \to 0} \left[\frac{(p + \delta p_1) - (p + \delta p_2)}{\delta x} \right] = g \delta \rho \frac{\partial h}{\partial x}$$

where $\delta \rho = \rho_1 - \rho_2$.

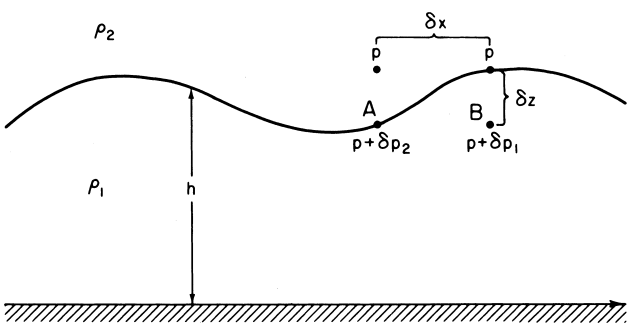


Fig. 7.7 A two-layer fluid system.

We assume that the motion is two dimensional in the x, z plane. The x momentum equation for the lower layer is then

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} = -\frac{g\delta\rho}{\rho_1}\frac{\partial h}{\partial x} \tag{7.17}$$

whereas the continuity equation is

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \tag{7.18}$$

Now since the pressure gradient in (7.17) is independent of z, u will also be independent of z provided that $u \neq u(z)$ initially. Thus, (7.18) can be integrated vertically from the lower boundary $z = 0$ to the interface $z = h$ to yield

$$w\,(h) - w\,(0) = -h\big(\partial u/\partial x\big)$$

However, $w(h)$ is just the rate at which the interface height is changing,

$$w\,(h) = \frac{Dh}{Dt} = \frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x}$$

and $w(0) = 0$ for a flat lower boundary. Hence, the vertically integrated continuity equation can be written

$$\frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + h\frac{\partial u}{\partial x} = \frac{\partial h}{\partial t} + \frac{\partial}{\partial x}\,(hu) = 0 \tag{7.19}$$

Equations (7.17) and (7.19) are a closed set in the variables u and h. We now apply the perturbation technique by letting

$$u = \overline{u} + u', \quad h = H + h'$$

where $\overline{u}$ as before is a constant basic state zonal velocity and H is the mean depth of the lower layer. The perturbation forms of (7.17) and (7.19) are then

$$\frac{\partial u'}{\partial t} + \overline{u}\frac{\partial u'}{\partial x} + \frac{g\delta\rho}{\rho_1}\frac{\partial h'}{\partial x} = 0 \tag{7.20}$$

$$\frac{\partial h'}{\partial t} + \overline{u}\frac{\partial h'}{\partial x} + H\frac{\partial u'}{\partial x} = 0 \tag{7.21}$$

where we assume that $H \gg |h'|$ so that products of the perturbation variables can be neglected.

Eliminating u' between (7.20) and (7.21) yields

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)^2 h' - \frac{gH\delta\rho}{\rho_1}\frac{\partial^2 h'}{\partial x^2} = 0 \tag{7.22}$$

which is a wave equation similar in form to (7.14). It is easily verified by direct substitution that (7.22) has a solution of the form

$$h' = A\exp[ik(x-ct)]$$

where the phase speed c satisfies the relationship

$$c = \overline{u} \pm \left(gH\delta\rho/\rho_1\right)^{1/2} \tag{7.23}$$

If the upper and lower layers are air and water, respectively, then $\delta\rho \approx \rho_1$ and the phase speed formula simplifies to

$$c = \overline{u} \pm \sqrt{gH}$$

The quantity $\sqrt{gH}$ is called the *shallow water* wave speed. It is a valid approximation only for waves whose wavelengths are much greater than the depth of the fluid. This restriction is necessary in order that the vertical velocities be small enough so that the hydrostatic approximation is valid. For an ocean depth of 4 km, the shallow water gravity wave speed is ≈ 200 m s^{-1}. Thus, long waves on the ocean surface travel very rapidly. It should be emphasized again that this theory applies only to waves of wavelength much greater than H. Such long waves are not ordinarily excited by the wind stresses, but may be produced by very large-scale disturbances such as earthquakes.[3]

Shallow water gravity waves may also occur at interfaces within the ocean where there is a very sharp density gradient (diffusion will always prevent formation of a true density discontinuity). In particular, the surface water is separated from the deep water by a narrow region of sharp density contrast called the *thermocline*. If the horizontal pressure gradient vanishes in the layer above the thermocline, then (7.22) governs the displacement, h', of the thermocline from its mean height H. If the density changes by an amount $\delta\rho/\rho_1 \approx 0.01$, across the thermocline, then from (7.23) the wave speed for waves traveling along the thermocline will be only one-tenth of the surface wave speed for a fluid of the same depth.[4]

[3] Long waves excited by underwater earthquakes or volcanic eruptions are called *tsunamis*.

[4] Gravity waves propagating along an internal density discontinuity are sometimes referred to as *internal* waves. We will, however, reserve that terminology for the vertically propagating waves considered in Section 7.4.

7.4 INTERNAL GRAVITY (BUOYANCY) WAVES

We now consider the nature of gravity wave propagation in the atmosphere. Atmospheric gravity waves can only exist when the atmosphere is stably stratified so that a fluid parcel displaced vertically will undergo buoyancy oscillations (see Section 2.7.3). Because the buoyancy force is the restoring force responsible for gravity waves, the term *buoyancy wave* is actually more appropriate as a name for these waves. However, in this text we will generally use the traditional name *gravity wave*.

In a fluid, such as the ocean, which is bounded both above and below, gravity waves propagate primarily in the horizontal plane since vertically traveling waves are reflected from the boundaries to form standing waves. However, in a fluid that has no upper boundary, such as the atmosphere, gravity waves may propagate vertically as well as horizontally. In vertically propagating waves the phase is a function of height. Such waves are referred to as *internal* waves. Although internal gravity waves are not generally of great importance for synoptic-scale weather forecasting (and indeed are nonexistent in the filtered quasi-geostrophic models), they can be important in mesoscale motions. For example, they are responsible for the occurrence of mountain *lee waves*. They also are believed to be an important mechanism for transporting energy and momentum into the middle atmosphere, and are often associated with the formation of clear air turbulence (CAT).

7.4.1 Pure Internal Gravity Waves

For simplicity we neglect the Coriolis force and limit our discussion to two-dimensional internal gravity waves propagating in the x, z plane. An expression for the frequency of such waves can be obtained by modifying the parcel theory developed in Section 2.7.3.

Internal gravity waves are transverse waves in which the parcel oscillations are parallel to the phase lines as indicated in Fig. 7.8. A parcel displaced a distance δs along a line tilted at an angle α to the vertical as shown in Fig. 7.8 will undergo a vertical displacement $\delta z = \delta s \cos\alpha$. For such a parcel the *vertical* buoyancy force per unit mass is just $-N^2\delta z$, as was shown in (2.52). Thus, the component of the buoyancy force parallel to the tilted path along which the parcel oscillates is just

$$-N^2\delta z\cos\alpha = -N^2\left(\delta s\cos\alpha\right)\cos\alpha = -\left(N\cos\alpha\right)^2\delta s$$

The momentum equation for the parcel oscillation is then

$$\frac{d^2\left(\delta s\right)}{dt^2} = -\left(N\cos\alpha\right)^2\delta s \tag{7.24}$$

which has the general solution $\delta s = \exp\left[\pm i\left(N\cos\alpha\right)t\right]$. Thus, the parcels execute a simple harmonic oscillation at the frequency $\nu = N\cos\alpha$. This frequency

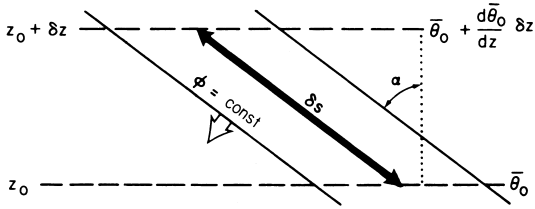


Fig. 7.8 Parcel oscillation path (heavy arrow) for pure gravity waves with phase lines tilted at an angle α to the vertical.

depends only on the static stability (measured by the buoyancy frequency N) and the angle of the phase lines to the vertical.

The above heuristic derivation can be verified by considering the linearized equations for two-dimensional internal gravity waves. For simplicity, we employ the *Boussinesq approximation*, in which density is treated as a constant except where it is coupled with gravity in the buoyancy term of the vertical momentum equation. Thus, in this approximation the atmosphere is considered to be incompressible, and local density variations are assumed to be small perturbations of the constant basic state density field. Because the vertical variation of the basic state density is neglected except where coupled with gravity, the Boussinesq approximation is only valid for motions in which the vertical scale is less than the atmospheric scale height $H(\approx 8km)$.

Neglecting effects of rotation, the basic equations for two-dimensional motion of an incompressible atmosphere may be written as

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0 \tag{7.25}$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + w\frac{\partial w}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial z} + g = 0 \tag{7.26}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \tag{7.27}$$

$$\frac{\partial \theta}{\partial t} + u\frac{\partial \theta}{\partial x} + w\frac{\partial \theta}{\partial z} = 0 \tag{7.28}$$

where the potential temperature θ is related to pressure and density by

$$\theta = \frac{p}{\rho R}\left(\frac{p_s}{p}\right)^{\kappa}$$

which after taking logarithms on both sides yields

$$\ln\theta = \gamma^{-1}\ln p - \ln\rho + \text{constant} \tag{7.29}$$

We now linearize (7.25)–(7.29) by letting

$$\begin{aligned} \rho &= \rho_0 + \rho' & u &= \overline{u} + u' \\ p &= \overline{p}(z) + p' & w &= w' \\ \theta &= \overline{\theta}(z) + \theta' \end{aligned} \tag{7.30}$$

where the basic state zonal flow $\overline{u}$ and the density ρ_0 are both assumed to be constant. The basic state pressure field must satisfy the hydrostatic equation

$$d\overline{p}/dz = -\rho_0 g \tag{7.31}$$

while the basic state potential temperature must satisfy (7.29) so that

$$\ln \overline{\theta} = \gamma^{-1} \ln \overline{p} - \ln \rho_0 + \text{constant} \tag{7.32}$$

The linearized equations are obtained by substituting from (7.30) into (7.25)–(7.29) and neglecting all terms that are products of the perturbation variables. Thus, for example, the last two terms in (7.26) are approximated as

$$\begin{aligned} \frac{1}{\rho}\frac{\partial p}{\partial z} + g &= \frac{1}{\rho_0 + \rho'}\left(\frac{d\overline{p}}{dz} + \frac{\partial p'}{\partial z}\right) + g \\ &\approx \frac{1}{\rho_0}\frac{d\overline{p}}{dz}\left(1 - \frac{\rho'}{\rho_0}\right) + \frac{1}{\rho_0}\frac{\partial p'}{\partial z} + g = \frac{1}{\rho_0}\frac{\partial p'}{\partial z} + \frac{\rho'}{\rho_0} g \end{aligned} \tag{7.33}$$

where (7.31) has been used to eliminate $\overline{p}$. The perturbation form of (7.29) is obtained by noting that

$$\ln\left[\overline{\theta}\left(1 + \frac{\theta'}{\overline{\theta}}\right)\right] = \gamma^{-1} \ln\left[\overline{p}\left(1 + \frac{p'}{\overline{p}}\right)\right] - \ln\left[\rho_0\left(1 + \frac{\rho'}{\rho_0}\right)\right] + \text{const.} \tag{7.34}$$

Now, recalling that $\ln(ab) = \ln(a) + \ln(b)$ and that $\ln(1 + \varepsilon) \approx \varepsilon$ for any $\varepsilon \ll 1$, we find with the aid of (7.32) that (7.34) may be approximated by

$$\frac{\theta'}{\overline{\theta}} \approx \frac{1}{\gamma}\frac{p'}{\overline{p}} - \frac{\rho'}{\rho_0}$$

Solving for ρ' yields

$$\rho' \approx -\rho_0 \frac{\theta'}{\overline{\theta}} + \frac{p'}{c_s^2} \tag{7.35}$$

where $c_s^2 \equiv \overline{p}\gamma/\rho_0$ is the square of the speed of sound. For buoyancy wave motions $\left|\rho_0\theta'/\overline{\theta}\right| \gg \left|p'/c_s^2\right|$; that is, density fluctuations due to pressure changes

are small compared with those due to temperature changes. Therefore, to a first approximation,

$$\theta'/\overline{\theta} = -\rho'/\rho_0 \tag{7.36}$$

Using (7.33) and (7.36), the linearized version of the set (7.25)–(7.28), we can write as

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)u' + \frac{1}{\rho_0}\frac{\partial p'}{\partial x} = 0 \tag{7.37}$$

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)w' + \frac{1}{\rho_0}\frac{\partial p'}{\partial z} - \frac{\theta'}{\overline{\theta}}g = 0 \tag{7.38}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0 \tag{7.39}$$

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\theta' + w'\frac{d\overline{\theta}}{dz} = 0 \tag{7.40}$$

Subtracting $\partial(7.37)/\partial z$ from $\partial(7.38)/\partial x$, we can eliminate p' to obtain

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\left(\frac{\partial w'}{\partial x} - \frac{\partial u'}{\partial z}\right) - \frac{g}{\overline{\theta}}\frac{\partial \theta'}{\partial x} = 0 \tag{7.41}$$

which is just the y component of the vorticity equation.

With the aid of (7.39) and (7.40), u' and θ' can be eliminated from (7.41) to yield a single equation for w':

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)^2\left(\frac{\partial^2 w'}{\partial x^2} + \frac{\partial^2 w'}{\partial z^2}\right) + N^2\frac{\partial^2 w'}{\partial x^2} = 0 \tag{7.42}$$

where $N^2 \equiv g\, d\ln\overline{\theta}/dz$ is the square of the buoyancy frequency, which is assumed to be constant.[5]

Equation (7.42) has harmonic wave solutions of the form

$$w' = \mathrm{Re}\left[\hat{w}\exp(i\phi)\right] = w_r\cos\phi - w_i\sin\phi \tag{7.43}$$

where $\hat{w} = w_r + iw_i$ is a complex amplitude with real part w_r and imaginary part w_i, and $\phi = kx + mz - \nu t$ is the phase, which is assumed to depend linearly on z as well as on x and t. Here the horizontal wave number k is real because the solution is always sinusoidal in x. The vertical wave number $m = m_r + im_i$ may,

[5] Strictly speaking, N^2 cannot be exactly constant if ρ_0 is constant. However, for shallow disturbances the variation of N^2 with height is unimportant.

however, be complex, in which case m_r describes sinusoidal variation in z and m_i describes exponential decay or growth in z depending on whether m_i is positive or negative. When m is real, the total wave number may be regarded as a vector $\boldsymbol{\kappa} \equiv (k, m)$, directed perpendicular to lines of constant phase, and in the direction of phase increase, whose components, $k = 2\pi/L_x$ and $m = 2\pi/L_z$, are inversely proportional to the horizontal and vertical wavelengths, respectively. Substitution of the assumed solution into (7.42) yields the dispersion relationship

$$(\nu - \overline{u}k)^2 \left(k^2 + m^2\right) - N^2 k^2 = 0$$

so that

$$\hat{\nu} \equiv \nu - \overline{u}k = \pm Nk/\left(k^2 + m^2\right)^{1/2} = \pm Nk/|\boldsymbol{\kappa}| \tag{7.44}$$

where $\hat{\nu}$, the *intrinsic frequency*, is the frequency relative to the mean wind. Here, the plus sign is to be taken for eastward phase propagation and the minus sign for westward phase propagation, relative to the mean wind.

If we let $k > 0$ and $m < 0$, then lines of constant phase tilt eastward with increasing height as shown in Fig. 7.9 (i.e., for $\phi = kx + mz$ to remain constant as x increases, z must also increase when $k > 0$ and $m < 0$). The choice of the positive root in (7.44) then corresponds to eastward and downward phase propagation relative to the mean flow with horizontal and vertical phase speeds (relative to the

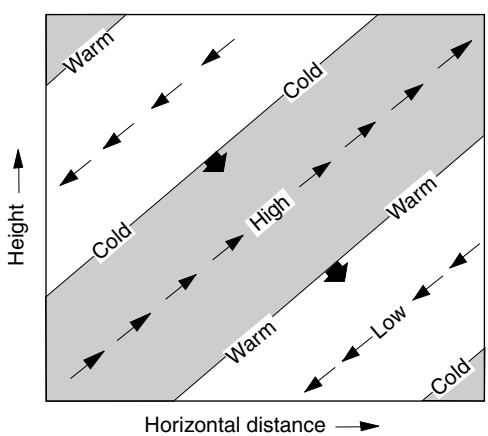


Fig. 7.9 Idealized cross section showing phases of pressure, temperature, and velocity perturbations for an internal gravity wave. Thin arrows indicate the perturbation velocity field, blunt solid arrows the phase velocity. Shading shows regions of upward motion.

mean flow) given by $c_x = \hat{\nu}/k$ and , $c_z = \hat{\nu}/m$ respectively.[6] The components of the group velocity, $c_{gx}$and c_{gz}, however, are given by

$$c_{gx} = \frac{\partial \nu}{\partial k} = \overline{u} \pm \frac{Nm^2}{\left(k^2 + m^2\right)^{3/2}} \tag{7.45a}$$

$$c_{gz} = \frac{\partial \nu}{\partial m} = \pm \frac{(-Nkm)}{\left(k^2 + m^2\right)^{3/2}} \tag{7.45b}$$

where the upper or lower signs are chosen in the same way as in (7.44). Thus, the vertical component of group velocity has a sign opposite to that of the vertical phase speed relative to the mean flow (downward phase propagation implies upward energy propagation). Furthermore, it is easily shown from (7.45) that the group velocity vector is parallel to lines of constant phase. Internal gravity waves thus have the remarkable property that group velocity is perpendicular to the direction of phase propagation. Because energy propagates at the group velocity this implies that energy propagates parallel to the wave crests and troughs, rather than perpendicular to them as in acoustic waves or shallow water gravity waves. In the atmosphere, internal gravity waves generated in the troposphere by cumulus convection, by flow over topography, and by other processes may propagate upward many scale heights into the middle atmosphere, even though individual fluid parcel oscillations may be confined to vertical distances much less than a kilometer.

Referring again to Fig. 7.9 it is evident that the angle of the phase lines to the local vertical is given by

$$\cos\alpha = L_z/(L_x^2 + L_z^2)^{1/2} = \pm k/(k^2 + m^2)^{1/2} = \pm k/\,|\boldsymbol{\kappa}|$$

Thus, $\hat{\nu} = \pm N \cos\alpha$ (i.e., gravity wave frequencies must be less than the buoyancy frequency) in agreement with the heuristic parcel oscillation model (7.24). The tilt of phase lines for internal gravity waves depends only on the ratio of the intrinsic wave frequency to the buoyancy frequency, and is independent of wavelength.

7.4.2 Topographic Waves

When air with mean wind speed $\overline{u}$ is forced to flow over a sinusoidal pattern of ridges under statically stable conditions, individual air parcels are alternately displaced upward and downward from their equilibrium levels and will thus undergo buoyancy oscillations as they move across the ridges as shown in Fig. 7.10. In

[6] Note that phase speed is not a vector. The phase speed in the direction perpendicular to constant phase lines (i.e., the blunt arrows in Fig. 7.9) is given by $\nu/(k^2 + m^2)^{1/2}$, which is not equal to $(c_x^2 + c_z^2)^{1/2}$.

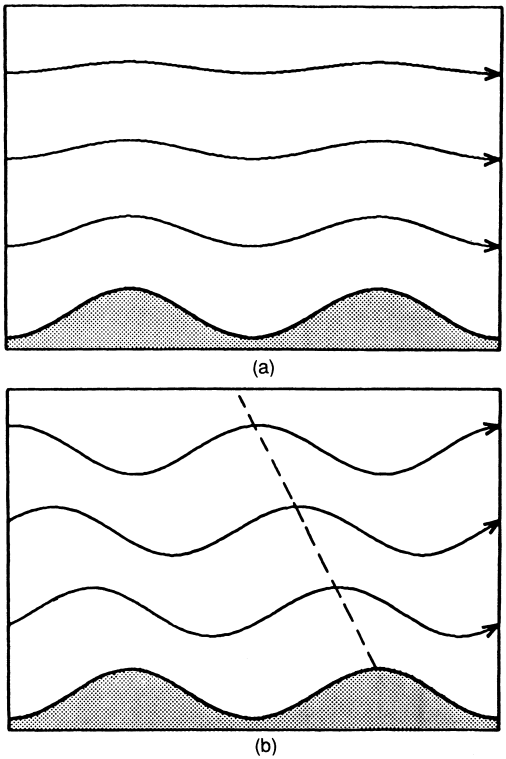


Fig. 7.10 Streamlines in steady flow over an infinite series of sinusoidal ridges for the narrow ridge case (a) and broad ridge case (b). The dashed line in (b) shows the phase of maximum upward displacement. (After Durran, 1990.)

this case there are solutions in the form of waves that are stationary relative to the ground [i.e, $\nu = 0$ in (7.43)]. For such stationary waves, w' depends only on (x, z) and (7.42) simplifies to

$$\left(\frac{\partial^2 w'}{\partial x^2} + \frac{\partial^2 w'}{\partial z^2}\right) + \frac{N^2}{\overline{u}^2} w' = 0 \tag{7.46}$$

Substituting from (7.43) into (7.46) then yields the dispersion relationship

$$m^2 = N^2 \Big/ \overline{u}^2 - k^2 \tag{7.47}$$

For given values of N, k, and $\overline{u}$, (7.47) determines the vertical structure. Then if $|\overline{u}| < N/k$, (7.47) shows that $m^2 > 0$ (i.e., m must be real) and solutions of (7.46) have the form of vertically propagating waves:

$$w' = \hat{w} \exp\left[i\,(kx + mz)\right]$$

Here we see from (7.44) that if we set $k > 0$ then for $\overline{u} > 0$ we have $\hat{v} < 0$ so that $m > 0$, whereas for $\overline{u} < 0$ we have $m < 0$. In the former situation the lower signs apply on the right in (7.45a,b), whereas in the latter the upper signs apply. In both cases the vertical phase propagation is downward relative to the mean flow, and vertical energy propagation is upward.

When $m^2 < 0, m = im_i$ is imaginary and the solution to (7.46) will have the form of vertically trapped waves:

$$w' = \hat{w} \exp(ikx) \exp(-m_i z)$$

Thus, vertical propagation is possible only when $|\overline{u}k|$, the magnitude of the frequency relative to the mean flow, is less than the buoyancy frequency. Stable stratification, wide ridges, and comparatively weak zonal flow provide favorable conditions for the formation of vertically propagating topographic waves (m real). Because the energy source for these waves is at the ground, they must transport energy upward. Hence, the phase speed relative to the mean zonal flow must have a downward component. Thus if $\overline{u} > 0$, lines of constant phase must tilt westward with height. When m is imaginary, however, the solution (7.43) has exponential behavior in the vertical with an exponential decay height of μ^{-1}, where $\mu = |m|$. Boundedness as $z \to \infty$ requires that we choose the solution with exponential decay away from the lower boundary.

In order to contrast the character of the solutions for real and imaginary m, we consider a specific example in which there is westerly mean flow over topography with a height profile given by

$$h(x) = h_M \cos kx$$

where h_M is the amplitude of the topography. Then because the flow at the lower boundary must be parallel to the boundary, the vertical velocity perturbation at the boundary is given by the rate at which the boundary height changes following the motion:

$$w'(x, 0) = \left(Dh/Dt\right)_{z=0} \approx \overline{u}\partial h/\partial x = -\overline{u}kh_M \sin kx$$

and the solution of (7.46) that satisfies this condition can be written

$$w(x, z) = \left\{ \begin{array}{ll} -\overline{u}h_M k e^{-\mu z} \sin kx, & \overline{u}k > N \\ -\overline{u}h_M k \sin(kx + mz), & \overline{u}k < N \end{array} \right\} \tag{7.48}$$

For fixed mean wind and buoyancy frequency, the character of the solution depends only on the horizontal scale of the topography. The two cases of (7.48) may be regarded as narrow ridge and wide ridge cases, respectively, since for specified

values of $\overline{u}$ and N the character of the solution is determined by the zonal wave number k. The streamline patterns corresponding to these cases for westerly flow are illustrated in Fig. 7.10. In the narrow ridge case (Fig. 7.10a), the maximum upward displacement occurs at the ridge tops, and the amplitude of the disturbance decays with height. In the wide ridge case (Fig. 7.10b), the line of maximum upward displacement tilts back toward the west ($m > 0$), and amplitude is independent of height consistent with an internal gravity wave propagating westward relative to the mean flow.

Alternatively, for fixed zonal wave number and buoyancy frequency the solution depends only on the speed of the mean zonal wind. As indicated in (7.48), only for mean zonal wind magnitudes less than the critical value N/k will vertical wave propagation occur.

Equation (7.46) was obtained for conditions of constant basic state flow. In reality, both the zonal wind $\overline{u}$ and the stability parameter N generally vary with height, and ridges are usually isolated rather than periodic. A wide variety of responses are possible depending on the shape of the terrain and wind and stability profiles. Under certain conditions, large-amplitude waves can be formed, which may generate severe downslope surface winds and zones of strong clear air turbulence. Such circulations are discussed further in Section 9.4.

7.5 GRAVITY WAVES MODIFIED BY ROTATION

Gravity waves with horizontal scales greater than a few hundred kilometers and periods greater than a few hours are hydrostatic, but they are influenced by the Coriolis effect and are characterized by parcel oscillations that are elliptical rather than straight lines as in the pure gravity wave case. This *elliptical polarization* can be understood qualitatively by observing that the Coriolis effect resists horizontal parcel displacements in a rotating fluid, but in a manner somewhat different from that in which the buoyancy force resists vertical parcel displacements in a statically stable atmosphere. In the latter case the resistive force is opposite to the direction of parcel displacement, whereas in the former it is at right angles to the horizontal parcel velocity.

7.5.1 Pure Inertial Oscillations

Section 3.2.3 showed that a parcel put into horizontal motion in a resting atmosphere with constant Coriolis parameter executes a circular trajectory in an anticyclonic sense. A generalization of this type of inertial motion to the case with a geostrophic mean zonal flow can be derived using a parcel argument similar to that used for the buoyancy oscillation in Section 2.7.3.

If the basic state flow is assumed to be a zonally directed geostrophic wind u_g, and it is assumed that the parcel displacement does not perturb the pressure field, the approximate equations of motion become

$$\frac{Du}{Dt} = fv = f\frac{Dy}{Dt} \tag{7.49}$$

$$\frac{Dv}{Dt} = f\left(u_g - u\right) \tag{7.50}$$

We consider a parcel that is moving with the geostrophic basic state motion at a position $y = y_0$. If the parcel is displaced across stream by a distance δy, we can obtain its new zonal velocity from the integrated form of (7.49):

$$u\left(y_0 + \delta y\right) = u_g\left(y_0\right) + f\delta y \tag{7.51}$$

The geostrophic wind at $y_0 + \delta y$ can be approximated as

$$u_g\left(y_0 + \delta y\right) = u_g\left(y_0\right) + \frac{\partial u_g}{\partial y}\delta y \tag{7.52}$$

Using (7.51) and (7.52) to evaluate (7.50) at $y_0 + \delta y$ yields

$$\frac{Dv}{Dt} = \frac{D^2\delta y}{Dt^2} = -f\left(f - \frac{\partial u_g}{\partial y}\right)\delta y = -f\frac{\partial M}{\partial y}\delta y \tag{7.53}$$

where we have defined the *absolute momentum*, $M \equiv fy - u_g$.

This equation is mathematically of the same form as (2.52), the equation for the motion of a vertically displaced particle in a stratified atmosphere. Depending on the sign of the coefficient on the right-hand side in (7.53), the parcel will either be forced to return to its original position or will accelerate further from that position. This coefficient thus determines the condition for *inertial instability*:

$$f\frac{\partial M}{\partial y} = f\left(f - \frac{\partial u_g}{\partial y}\right)\begin{cases} > 0 & \text{stable} \\ = 0 & \text{neutral} \\ < 0 & \text{unstable} \end{cases} \tag{7.54}$$

Viewed in an inertial reference frame, instability results from an imbalance between the pressure gradient and inertial forces for a parcel displaced radially in an axisymmetric vortex. In the Northern Hemisphere, where f is positive, the flow is inertially stable provided that the absolute vorticity of the basic flow, $\partial M/\partial y$, is positive. In the Southern Hemisphere, however, inertial stability requires that the absolute vorticity be negative. Observations show that for extratropical synoptic-scale systems the flow is always inertially stable, although near neutrality often

occurs on the anticyclonic shear side of upper level jet streaks. The occurrence of inertial instability over a large area would immediately trigger inertially unstable motions, which would mix the fluid laterally just as convection mixes it vertically, and reduce the shear until the absolute vorticity times f was again positive. (This explains why anticyclonic shears cannot become arbitrarily large.) Inertial instability is considered further in a more general context in Section 9.3.

7.5.2 Inertia–Gravity Waves

When the flow is both inertially and gravitationally stable, parcel displacements are resisted by both rotation and buoyancy. The resulting oscillations are called *inertia–gravity waves*. The dispersion relation for such waves can be analyzed using a variant of the parcel method applied in Section 7.4. We consider parcel oscillations along a slantwise path in the (y, z) plane as shown in Fig. 7.11. For a vertical displacement δz the buoyancy force component parallel to the slope of the parcel oscillation is $-N^2\delta z \cos\alpha$, and for a meridional displacement δy the Coriolis (inertial) force component parallel to the slope of the parcel path is $-f^2 \delta y \sin\alpha$, where we have assumed that the geostrophic basic flow is constant in latitude. Thus, the harmonic oscillator equation for the parcel (7.24) is modified to the form

$$\frac{D^2\delta s}{Dt^2} = -(f\sin\alpha)^2\,\delta s - (N\cos\alpha)^2\,\delta s \tag{7.55}$$

where δs is again the perturbation parcel displacement.

The frequency now satisfies the dispersion relationship

$$\nu^2 = N^2\cos^2\alpha + f^2\sin^2\alpha \tag{7.56}$$

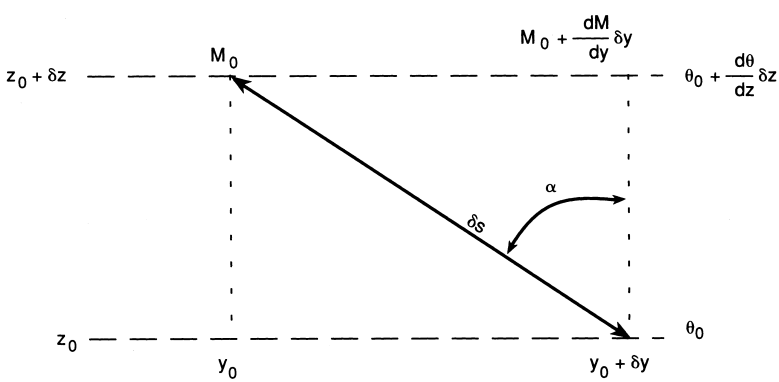


Fig. 7.11 Parcel oscillation path in meridional plane for an inertia–gravity wave. See text for definition of symbols.

Since in general $N^2 > f^2$ (7.56) indicates that inertia–gravity wave frequencies must lie in the range $f \leq |\nu| \leq N$. The frequency approaches N as the trajectory slope approaches the vertical, and approaches f as the trajectory slope approaches the horizontal. For typical midlatitude tropospheric conditions, *inertia*–gravity wave periods are in the approximate range of 12 min to 15 h. Rotational effects become important, however, only when the second term on the right in (7.56) is similar in magnitude to the first term. This requires that $\tan^2 \alpha \sim N^2/f^2 = 10^4$, in which case it is clear from (7.56) that $\nu \ll N$. Thus, only low-frequency gravity waves are modified significantly by the rotation of the earth, and these have very small parcel trajectory slopes.

The heuristic parcel derivation can again be verified by using the linearized dynamical equations. In this case, however, it is necessary to include rotation. The small parcel trajectory slopes of the relatively long period waves that are altered significantly by rotation imply that the horizontal scales are much greater than the vertical scales for these waves. Therefore, we may assume that the motions are in hydrostatic balance. If in addition we assume a motionless basic state, the linearized equations (7.37)–(7.40) are replaced by the set

$$\frac{\partial u'}{\partial t} - fv' + \frac{1}{\rho_0}\frac{\partial p'}{\partial x} = 0 \tag{7.57}$$

$$\frac{\partial v'}{\partial t} + fu' + \frac{1}{\rho_0}\frac{\partial p'}{\partial y} = 0 \tag{7.58}$$

$$\frac{1}{\rho_0}\frac{\partial p'}{\partial z} - \frac{\theta'}{\overline{\theta}}g = 0 \tag{7.59}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} = 0 \tag{7.60}$$

$$\frac{\partial \theta'}{\partial t} + w'\frac{d\overline{\theta}}{dz} = 0 \tag{7.61}$$

The hydrostatic relationship in (7.59) may be used to eliminate θ' in (7.61) to yield

$$\frac{\partial}{\partial t}\left(\frac{1}{\rho_0}\frac{\partial p'}{\partial z}\right) + N^2 w' = 0 \tag{7.62}$$

Letting

$$\left(u', v', w', p'/\rho_0\right) = \mathrm{Re}\left[\left(\hat{u}, \hat{v}, \hat{w}, \hat{p}\right)\exp i\,(kx + ly + mz - \nu t)\right]$$

and substituting into (7.57), (7.58), and (7.62), we obtain

$$\hat{u} = \left(\nu^2 - f^2\right)^{-1}(\nu k + ilf)\,\hat{p} \tag{7.63}$$

$$\hat{v} = \left(\nu^2 - f^2\right)^{-1} (\nu l - ikf)\, \hat{p} \tag{7.64}$$

$$\hat{w} = -\left(\nu m \Big/ N^2\right) \hat{p} \tag{7.65}$$

which with the aid of (7.60) yields the dispersion relation for hydrostatic waves.

$$\nu^2 = f^2 + N^2 \left(k^2 + l^2\right) m^{-2} \tag{7.66}$$

Because hydrostatic waves must have $(k^2 + l^2)/m^2 \ll 1$, (7.66) indicates that for vertical propagation to be possible (m real) the frequency must satisfy the inequality $|f| < |\nu| \ll N$. Equation (7.66) is just the limit of (7.56) when we let

$$\sin^2 \alpha \to 1, \cos^2\alpha = \left(k^2 + l^2\right) \Big/ m^2$$

which is consistent with the hydrostatic approximation.

If axes are chosen to make $l = 0$, it may be shown (see Problem 7.14) that the ratio of the vertical to horizontal components of group velocity is given by

$$\left|c_{gz}/c_{gx}\right| = |k/m| = \left(\nu^2 - f^2\right)^{1/2} / N \tag{7.67}$$

Thus, for fixed υ, inertia–gravity waves propagate more closely to the horizontal than pure internal gravity waves. However, as in the latter case the group velocity vector is again parallel to lines of constant phase.

Eliminating $\hat{p}$ between (7.63) and (7.64) for the case $l = 0$ yields the relationship $\hat{v} = -if\hat{u}/\nu$, from which it is easily verified that if $\hat{u}$ is real, the perturbation horizontal motions satisfy the relations

$$u' = \hat{u} \cos(kx + mz - \nu t), \quad v' = \hat{u}\left(f/\nu\right) \sin(kx + mz - \nu t) \tag{7.68}$$

so that the horizontal velocity vector rotates anticyclonically (that is, clockwise in the Northern Hemisphere) with time. As a result, parcels follow elliptical trajectories in a plane orthogonal to the wavenumber vector. Equations (7.68) also show that the horizontal velocity vector turns anticyclonically with height for waves with upward energy propagation (e. g., waves with $m < 0$ and $\nu < 0$). These characteristics are illustrated by the vertical cross section shown in Fig. 7.12. The anticyclonic turning of the horizontal wind with height and time is a primary method for identifying inertia–gravity oscillations in meteorological data.

7.6 ADJUSTMENT TO GEOSTROPHIC BALANCE

Chapter 6 showed that synoptic-scale motions in midlatitudes are in approximate geostrophic balance. Departures from this balance can lead to the excitation of inertia–gravity waves, which act to adjust the mass and momentum

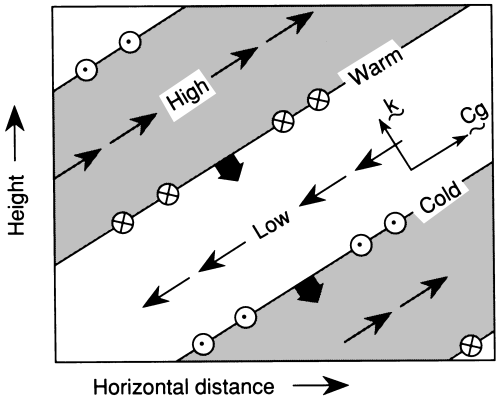


Fig. 7.12 Vertical section in a plane containing the wave vector $\boldsymbol{k}$ showing the phase relationships among velocity, geopotential, and temperature fluctuations in an upward propagating inertia–gravity wave with $m < 0$, $\nu > 0$, and $f > 0$ (Northern Hemisphere). Thin sloping lines denote the surfaces of constant phase (perpendicular to the wave vector), and thick arrows show the direction of phase propagation. Thin arrows show the perturbation zonal and vertical velocity fields. Meridional wind perturbations are shown by arrows pointed into the page (northward) and out of the page (southward). Note that the perturbation wind vector turns clockwise (anticyclonically) with height. (After Andrews et al., 1987.)

distributions so that the flow tends to return toward geostrophic balance. This section investigates the process by which geostrophic balance is achieved, that is, the *adjustment* process. For simplicity we utilize the prototype shallow water system; similar considerations apply to a continuously stratified atmosphere. For linearized disturbances about a basic state of no motion with a constant Coriolis parameter, f_0, the horizontal momentum and continuity equations are

$$\frac{\partial u'}{\partial t} - f_0 v' = -g\frac{\partial h'}{\partial x} \tag{7.69}$$

$$\frac{\partial v'}{\partial t} + f_0 u' = -g\frac{\partial h'}{\partial y} \tag{7.70}$$

$$\frac{\partial h'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0 \tag{7.71}$$

where h' is again the deviation from the mean depth H. Taking $\partial(7.69)/\partial x + \partial(7.70)/\partial y$ yields

$$\frac{\partial^2 h'}{\partial t^2} - c^2\left(\frac{\partial^2 h'}{\partial x^2} + \frac{\partial^2 h'}{\partial y^2}\right) + f_0 H\zeta' = 0 \tag{7.72}$$

here $c^2 \equiv g\mathrm{H}$ and $\zeta' = \partial v'/\partial x - \partial u'/\partial y$.

For $f_0 = 0$ (nonrotating system) the vorticity and height perturbations are uncoupled, and (7.72) yields a two-dimensional shallow water wave equation for h[compare with (7.22)]:

$$\frac{\partial^2 h'}{\partial t^2} - c^2 \left(\frac{\partial^2 h'}{\partial x^2} + \frac{\partial^2 h'}{\partial y^2} \right) = 0 \tag{7.73}$$

which has solutions of the form

$$h' = A \exp[i(kx + ly - \nu t)] \tag{7.74}$$

with $\nu^2 = c^2 \left(k^2 + l^2\right) = gH \left(k^2 + l^2\right)$. However, for $f_0 \neq 0$ the h' and ζ' fields are coupled through (7.72). For motions with time scales longer than $1/f_0$ (which is certainly true for synoptic-scale motions), the ratio of the first two terms in (7.72) is given by

$$\frac{\left|\partial^2 h'/\partial t^2\right|}{\left|c^2 \left(\partial^2 h'/\partial x^2 + \partial^2 h'/\partial y^2\right)\right|} \lesssim \frac{f_0^2 L^2}{gH}$$

which is small for $L \sim 1000$ km, provided that $H \gg 1$ km. Under such circumstances the time derivative term in (7.72) is small compared to the other two terms, and (7.72) states simply that the vorticity is in geostrophic balance.

If the flow is initially unbalanced, the complete equation (7.72) can be used to describe the approach toward geostrophic balance provided that we can obtain a second relationship between h' and ζ' taking

$$\partial\,(7.70)/\partial x - \partial\,(7.69)/\partial y$$

yields

$$\frac{\partial \zeta'}{\partial t} + f_0 \left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} \right) = 0 \tag{7.75}$$

which can be combined with (7.71) to give the linearized potential vorticity conservation law:

$$\frac{\partial \zeta'}{\partial t} - \frac{f_0}{H} \frac{\partial h'}{\partial t} = 0 \tag{7.76}$$

Thus, letting Q' designate the perturbation potential vorticity, we obtain from (7.76) the conservation relationship

$$Q'(x, y, t) = \zeta'/f_0 - h'/H = \text{Const.} \tag{7.77}$$

Hence, if we know the distribution of Q' at the initial time, we know Q' for all time:

$$Q'(x, y, t) = Q'(x, y, 0)$$

and the final adjusted state can be determined without solving the time-dependent problem.

This problem was first solved by Rossby in the 1930s and is often referred to as the Rossby adjustment problem. As a simplified, albeit somewhat unrealistic, example of the adjustment process, we consider an idealized shallow water system on a rotating plane with initial conditions

$$u', v' = 0; \qquad h' = -h_0 \mathrm{sgn}\,(x) \tag{7.78}$$

where $\mathrm{sgn}(x) = 1$ for $x > 0$ and $\mathrm{sgn}(x) = -1$ for $x < 0$. This corresponds to an initial step function in h' at $x = 0$, with the fluid motionless. Thus, from (7.77)

$$\left(\zeta'/f_0\right) - \left(h'/H\right) = \left(h_0/H\right) \mathrm{sgn}\,(x) \tag{7.79}$$

Using (7.79) to eliminate ζ' in (7.72) yields

$$\frac{\partial^2 h'}{\partial t^2} - c^2 \left(\frac{\partial^2 h'}{\partial x^2} + \frac{\partial^2 h'}{\partial y^2} \right) + f_0^2 h' = -f_0^2 h_0 \mathrm{sgn}\,(x) \tag{7.80}$$

which in the homogeneous case ($h_0 = 0$) yields the dispersion relation

$$\nu^2 = f_0^2 + c^2 \left(k^2 + l^2\right) = f_0^2 + \mathrm{g}H \left(k^2 + l^2\right) \tag{7.81}$$

This should be compared to (7.66).

Because initially h' is independent of y, it will remain so for all time. Thus, in the final steady state (7.80) becomes

$$-c^2 \frac{d^2 h'}{dx^2} + f_0^2 h' = -f_0^2 h_0 \,\mathrm{sgn}\,(x) \tag{7.82}$$

which has the solution

$$\frac{h'}{h_0} = \left\{ \begin{array}{ll} -1 + \exp\left(-x/\lambda_R\right) & \text{for } x > 0 \\ +1 - \exp\left(+x/\lambda_R\right) & \text{for } x < 0 \end{array} \right\} \tag{7.83}$$

where $\lambda_R \equiv f_0^{-1}\sqrt{\mathrm{g}H}$ is the Rossby *radius of deformation*. Hence, the radius of deformation may be interpreted as the horizontal length scale over which the height field adjusts during the approach to geostrophic equilibrium. For $|x| \gg \lambda_R$ the original h' remains unchanged. Substituting from (7.83) into (7.69)–(7.71) shows that the steady velocity field is geostrophic and nondivergent:

$$u' = 0, \quad \text{and} \quad v' = \frac{\mathrm{g}}{f_0} \frac{\partial h'}{\partial x} = -\frac{g h_0}{f_0 \lambda_R} \exp\left(-|x|/\lambda_R\right) \tag{7.84}$$

The steady-state solution (7.84) is shown in Fig. 7.13.

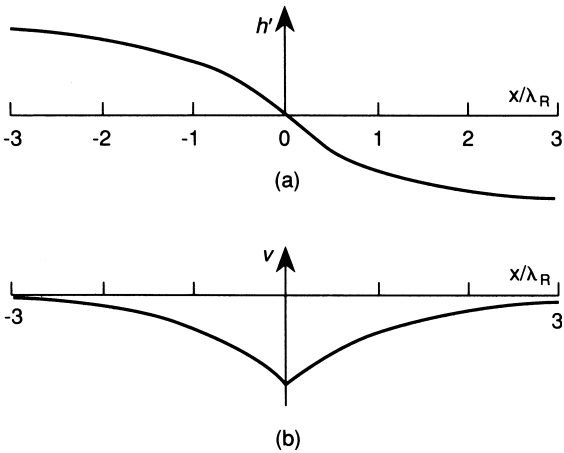


Fig. 7.13 The geostrophic equilibrium solution corresponding to adjustment from the initial state defined in (7.78). (a) Final surface elevation profiles; (b) the geostrophic velocity profile in the final state. (After Gill, 1982.)

Note that the result (7.84) could not be derived merely by setting $\partial/\partial t = 0$ in (7.69)–(7.71). That would yield geostrophic balance, and *any* distribution of h' would satisfy the equations:

$$f_0 u' = -g\frac{\partial h'}{\partial y}, \quad f_0 v' = g\frac{\partial h'}{\partial x}, \quad \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} = 0$$

Only by combining (7.69)–(7.71) to obtain the potential vorticity equation, and requiring the flow to satisfy potential vorticity conservation at all intermediate times, can the degeneracy of the geostrophic final state be eliminated. In other words, although any height field can satisfy the steady-state versions of (7.69)–(7.71), there is only one field that is consistent with a given initial state; this field can be found readily because it can be computed from the distribution of potential vorticity, which is conserved.

Although the final state can be computed without solving the time-dependent equation, if the evolution of the adjustment process is required, it is necessary to solve (7.80) subject to the initial conditions (7.78), which is beyond the scope of this discussion. We can, however, compute the amount of energy that is dispersed by gravity waves during the adjustment process. This only requires computing the energy change between initial and final states.

The potential energy per unit horizontal area is given by

$$\int_0^{h'} \rho g z dz = \rho g h'^2/2$$

Thus, the potential energy released per unit length in y during adjustment is

$$\int_{-\infty}^{+\infty} \frac{\rho g h_0^2}{2} dx - \int_{-\infty}^{+\infty} \frac{\rho g h'^2}{2} dx =$$

$$2\int_{0}^{+\infty} \frac{\rho g h_0^2}{2} \left[1 - \left(1 - e^{-x/\lambda_R}\right)^2\right] dx = \frac{3}{2}\rho g h_0^2 \lambda_R \tag{7.85}$$

In the nonrotating case ($\lambda_R \to \infty$) all potential energy available initially is released (converted to kinetic energy) so that there is an infinite energy release. (Energy is radiated away in the form of gravity waves, leaving a flat free surface extending to $|x| \to \infty$ as $t \to \infty$.)

In the rotating case only the finite amount given in (7.85) is converted to kinetic energy, and only a portion of this kinetic energy is radiated away. The rest remains in the steady geostrophic circulation. The kinetic energy in the steady-state per unit length is

$$2\int_{0}^{+\infty} \rho H \frac{v'^2}{2} dx = \rho H \left(\frac{g h_0}{f \lambda_R}\right)^2 \int_{0}^{+\infty} e^{-2x/\lambda_R} dx = \frac{1}{2}\rho g h_0^2 \lambda_R \tag{7.86}$$

Thus, in the rotating case a finite amount of potential energy is released, but only one-third of the potential energy released goes into the steady geostrophic mode. The remaining two-thirds is radiated away in the form of inertia–gravity waves.

This simple analysis illustrates the following points: (a) It is difficult to extract the potential energy of a rotating fluid. Although there is an infinite reservoir of potential energy in this example (because h' is finite as $|x| \to \infty$), only a finite amount is converted before geostrophic balance is achieved. (b) Conservation of potential vorticity allows one to determine the steady-state geostrophically adjusted velocity and height fields without carrying out a time integration. (c) The length scale for the steady solution is the Rossby radius λ_R.

The dynamics of the adjustment process plays an essential role in initialization and data assimilation in numerical prediction (see Section 13.7). For example, under some conditions the adjustment process may effectively damp out new height data inserted at a gridpoint, as the new data will generally be unbalanced and hence will tend to adjust toward geostrophic balance with the existing wind field.

7.7 ROSSBY WAVES

The wave type that is of most importance for large-scale meteorological processes is the *Rossby wave*, or *planetary wave*. In an inviscid barotropic fluid of constant depth (where the divergence of the horizontal velocity must vanish), the Rossby wave is an absolute vorticity-conserving motion that owes its existence to

the variation of the Coriolis parameter with latitude, the so-called β-effect. More generally, in a baroclinic atmosphere, the Rossby wave is a potential vorticity-conserving motion that owes its existence to the isentropic gradient of potential vorticity.

Rossby wave propagation can be understood in a qualitative fashion by considering a closed chain of fluid parcels initially aligned along a circle of latitude. Recall that the absolute vorticity η is given by $\eta = \zeta + f$, where ζ is the relative vorticity and f is the Coriolis parameter. Assume that $\zeta = 0$ at time t_0. Now suppose that at t_1, δy is the meridional displacement of a fluid parcel from the original latitude. Then at t_1 we have

$$(\zeta + f)_{t_1} = f_{t_0}$$

or

$$\zeta_{t_1} = f_{t_0} - f_{t_1} = -\beta \delta y \tag{7.87}$$

where $\beta \equiv df/dy$ is the planetary vorticity gradient at the original latitude.

From (7.87) it is evident that if the chain of parcels is subject to a sinusoidal meridional displacement under absolute vorticity conservation, the resulting perturbation vorticity will be positive for a southward displacement and negative for a northward displacement.

This perturbation vorticity field will induce a meridional velocity field, which advects the chain of fluid parcels southward west of the vorticity maximum and northward west of the vorticity minimum, as indicated in Fig. 7.14. Thus, the fluid parcels oscillate back and forth about their equilibrium latitude, and the pattern of vorticity maxima and minima propagates to the west. This westward propagating vorticity field constitutes a Rossby wave. Just as a positive vertical gradient of potential temperature resists vertical fluid displacements and provides the restoring force for gravity waves, the meridional gradient of absolute vorticity

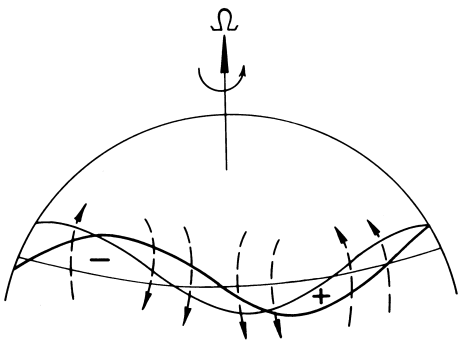


Fig. 7.14 Perturbation vorticity field and induced velocity field (dashed arrows) for a meridionally displaced chain of fluid parcels. Heavy wavy line shows original perturbation position; light line shows westward displacement of the pattern due to advection by the induced velocity.

resists meridional displacements and provides the restoring mechanism for Rossby waves.

The speed of westward propagation, c, can be computed for this simple example by letting $\delta y = a \sin[k(x - ct)]$, where a is the maximum northward displacement. Then $v = D(\delta y)/Dt = -kca \cos[k(x - ct)]$, and

$$\zeta = \partial v/\partial x = k^2 ca \sin[k(x - ct)]$$

Substitution for δy and ζ in (7.87) then yields

$$k^2 ca \sin[k(x - ct)] = -\beta a \sin[k(x - ct)]$$

or

$$c = -\beta/k^2 \tag{7.88}$$

Thus, the phase speed is westward relative to the mean flow and is inversely proportional to the square of the zonal wave number.

7.7.1 Free Barotropic Rossby Waves

The dispersion relationship for barotropic Rossby waves may be derived formally by finding wave-type solutions of the linearized barotropic vorticity equation. The barotropic vorticity equation (4.27) states that the vertical component of absolute vorticity is conserved following the horizontal motion. For a midlatitude β plane this equation has the form

$$\left(\frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}\right)\zeta + \beta v = 0 \tag{7.89}$$

We now assume that the motion consists of a constant basic state zonal velocity plus a small horizontal perturbation:

$$u = \overline{u} + u', \quad v = v', \quad \zeta = \partial v'/\partial x - \partial u'/\partial y = \zeta'$$

We define a perturbation streamfunction ψ' according to

$$u' = -\partial\psi'/\partial y, \quad v' = \partial\psi'/\partial x$$

from which $\zeta' = \nabla^2\psi'$. The perturbation form of (7.89) is then

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\nabla^2\psi' + \beta\frac{\partial\psi'}{\partial x} = 0 \tag{7.90}$$

where as usual we have neglected terms involving the products of perturbation quantities. We seek a solution of the form

$$\psi' = \mathrm{Re}\left[\Psi \exp(i\phi)\right]$$

where $\phi = kx + ly - \nu t$. Here k and l are wave numbers in the zonal and meridional directions, respectively. Substituting for ψ' in (7.90) gives

$$(-\nu + k\overline{u})\left(-k^2 - l^2\right) + k\beta = 0$$

which may immediately be solved for ν:

$$\nu = \overline{u}k - \beta k/K^2 \tag{7.91}$$

where $K^2 \equiv k^2 + l^2$ is the total horizontal wave number squared.

Recalling that $c = \nu/k$, we find that the zonal phase speed relative to the mean wind is

$$c - \overline{u} = -\beta/K^2 \tag{7.92}$$

which reduces to (7.88) when the mean wind vanishes and $l \to 0$. Thus, the Rossby wave zonal phase propagation is always *westward* relative to the mean zonal flow. Furthermore, the Rossby wave phase speed depends inversely on the square of the horizontal wavenumber. Therefore, Rossby waves are dispersive waves whose phase speeds increase rapidly with increasing wavelength.

This result is consistent with the discussion in Section 6.2.2 in which we showed that the advection of planetary vorticity, which tends to make disturbances retrogress, increasingly dominates over relative vorticity advection as the wavelength of a disturbance increases. Equation (7.92) provides a quantitative measure of this effect in cases where the disturbance is small enough in amplitude so that perturbation theory is applicable. For a typical midlatitude synoptic-scale disturbance, with similar meridional and zonal scales ($l \approx k$) and zonal wavelength of order 6000 km, the Rossby wave speed relative to the zonal flow calculated from (7.92) is approximately -8 m s^{-1}. Because the mean zonal wind is generally westerly and greater than 8 m s^{-1}, synoptic-scale Rossby waves usually move eastward, but at a phase speed relative to the ground that is somewhat less than the mean zonal wind speed.

For longer wavelengths the westward Rossby wave phase speed may be large enough to balance the eastward advection by the mean zonal wind so that the resulting disturbance is stationary relative to the surface of the earth. From (7.92) it is clear that the free Rossby wave solution becomes stationary when

$$K^2 = \beta/\overline{u} \equiv K_s^2 \tag{7.93}$$

The significance of this condition is discussed in the next subsection.

Unlike the phase speed, which is always westward relative to the mean flow, the zonal group velocity for a Rossby wave may be either eastward or westward relative to the mean flow, depending on the ratio of the zonal and meridional wave numbers (see Problem 7.20). Stationary Rossby modes (i.e., modes with $c = 0$) have zonal group velocities that are eastward relative to the ground. Synoptic-scale Rossby waves also tend to have zonal group velocities that are eastward relative to the ground. For synoptic waves, advection by the mean zonal wind is generally larger than the Rossby phase speed so that the phase speed is also eastward relative to the ground, but is slower than the zonal group velocity. As indicated in Fig. 7.4b, this implies that new disturbances tend to develop downstream of existing disturbances, which is an important consideration for forecasting.

It is possible to carry out a less restrictive analysis of free planetary waves using the perturbation form of the full primitive equations. In that case the structure of the free modes depends critically on the boundary conditions at the surface and the upper boundary. The results of such an analysis are mathematically complicated, but qualitatively yield waves with horizontal dispersion properties similar to those in the shallow water model. It turns out that the free oscillations allowed in a hydrostatic gravitationally stable atmosphere consist of eastward- and westward-moving gravity waves that are slightly modified by the rotation of the earth, and westward-moving Rossby waves that are slightly modified by gravitational stability. These free oscillations are the normal modes of oscillation of the atmosphere. As such, they are continually excited by the various forces acting on the atmosphere. Planetary scale-free oscillations, although they can be detected by careful observational studies, generally have rather weak amplitudes. Presumably this is because the forcing is quite weak at the large phase speeds characteristic of most such waves. An exception is the 16-day period zonal wave number 1 normal mode, which can be quite strong in the winter stratosphere.

7.7.2 Forced Topographic Rossby Waves

Although free propagating Rossby modes are only rather weakly excited in the atmosphere, forced stationary Rossby modes are of primary importance for understanding the planetary scale circulation pattern. Such modes may be forced by longitudinally dependent diabatic heating patterns or by flow over topography. Of particular importance for the Northern Hemisphere extratropical circulation are stationary Rossby modes forced by flow over the Rockies and the Himalayas. It is just the topographic Rossby wave that was described qualitatively in the discussion of streamline deflections in potential vorticity-conserving flows crossing mountain ranges in Section 4.3.

As the simplest possible dynamical model of topographic Rossby waves, we use the barotropic potential vorticity equation for a homogeneous fluid of variable depth (4.26). We assume that the upper boundary is at a fixed height H, and the

lower boundary is at the variable height $h_T(x, y)$ where $|h_T| \ll H$. We also use quasi-geostrophic scaling so that $|\zeta_g| \ll f_0$. We can then approximate (4.26) by

$$H\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla\right)(\zeta_g + f) = -f_0 \frac{Dh_T}{Dt} \tag{7.94}$$

Linearizing and applying the midlatitude β-plane approximation yields

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\zeta_g' + \beta v_g' = -\frac{f_0}{H}\overline{u}\frac{\partial h_T}{\partial x} \tag{7.95}$$

We now examine solutions of (7.95) for the special case of a sinusoidal lower boundary. We specify the topography to have the form

$$h_T(x, y) = \mathrm{Re}\left[h_0 \exp(ikx)\right]\cos ly \tag{7.96}$$

and represent the geostrophic wind and vorticity by the perturbation streamfunction

$$\psi(x, y) = \mathrm{Re}\left[\psi_0 \exp(ikx)\right]\cos ly \tag{7.97}$$

Then (7.95) has a steady-state solution with complex amplitude given by

$$\psi_0 = f_0 h_0 / \left[H\left(K^2 - K_s^2\right)\right] \tag{7.98}$$

The streamfunction is either exactly in phase (ridges over the mountains) or exactly out of phase (troughs over the mountains), with the topography depending on the sign of $K^2 - K_s^2$. For long waves ($K < K_s$) the topographic vorticity source in (7.95) is balanced primarily by the meridional advection of planetary vorticity (the β effect). For short waves ($K > K_s$) the source is balanced primarily by the zonal advection of relative vorticity.

The topographic wave solution (7.98) has the unrealistic characteristic that when the wave number exactly equals the critical wave number K_s the amplitude goes to infinity. From (7.93) it is clear that this singularity occurs at the zonal wind speed for which the free Rossby mode becomes stationary. Thus, it may be thought of as a resonant response of the barotropic system.

Charney and Eliassen (1949) used the topographic Rossby wave model to explain the winter mean longitudinal distribution of 500-hPa heights in Northern Hemisphere midlatitudes. They removed the resonant singularity by including boundary layer drag in the form of Ekman pumping, which for the barotropic vorticity equation is simply a linear damping of the relative vorticity [see (5.41)]. The vorticity equation thus takes the form

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\zeta_g' + \beta v_g' + r\zeta_g' = -\frac{f_0}{H}\overline{u}\frac{\partial h_T}{\partial x} \tag{7.99}$$

where $r \equiv \tau_e^{-1}$ is the inverse of the spin-down time defined in Section 5.4.

For steady flow, (7.99) has a solution with complex amplitude

$$\psi_0 = f_0 h_0 / \left[H \left(K^2 - K_s^2 - i\varepsilon \right) \right] \tag{7.100}$$

where $\varepsilon \equiv r K^2 (k\bar{u})^{-1}$. Thus, boundary layer drag shifts the phase of the response and removes the singularity at resonance. However, the amplitude is still a maximum for $K = K_s$. and the trough in the streamfunction occurs 1/4 cycle east of the mountain crest, in approximate agreement with observations.

By use of a Fourier expansion (7.99) can be solved for realistic distributions of topography. The results for an x-dependence of h_T given by a smoothed version of the earth's topography at 45°N, a meridional wave number corresponding to a latitudinal half-wavelength of 35°, $\tau_e = 5$ days, $\bar{u} = 17\ \mathrm{m\,s^{-1}}$, $f_0 = 10^{-4}\ \mathrm{s^{-1}}$, and $H = 8$ km are shown in Fig. 7.15. Despite its simplicity, the Charney–Eliassen model does a remarkable job of reproducing the observed 500-hPa stationary wave pattern in Northern Hemisphere midlatitudes.

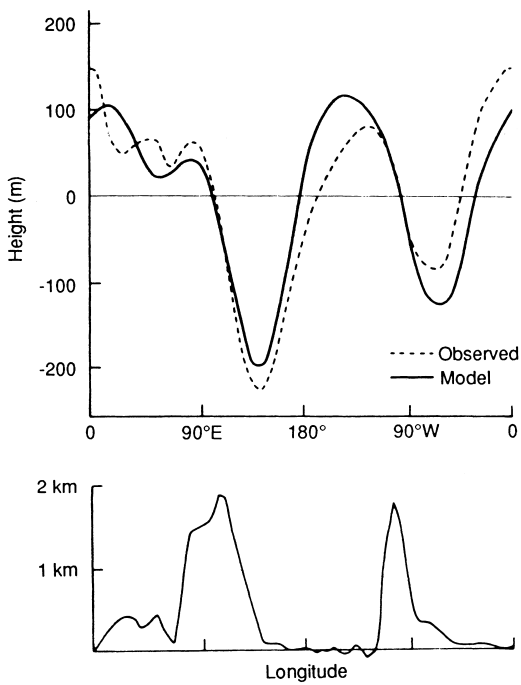


Fig. 7.15 (Top) Longitudinal variation of the disturbance geopotential height ($\equiv f_0\Psi/g$) in the Charney–Eliassen model for the parameters given in the text (solid line) compared with the observed 500-hPa height perturbations at 45°N in January (dashed line). (Bottom) Smoothed profile of topography at 45°N used in the computation. (After Held, 1983.)

PROBLEMS

7.1. Show that the Fourier component $F(x) = \mathrm{Re}\left[C \exp(imx)\right]$ can be written as

$$F(x) = |C| \cos m(x + x_0)$$

where $x_0 = m^{-1} \sin^{-1}\left(C_i / |C|\right)$ and C_i stands for the imaginary part of C.

7.2. In the study of atmospheric wave motions, it is often necessary to consider the possibility of amplifying or decaying waves. In such a case we might assume that a solution has the form

$$\psi = A \cos(kx - \nu t - kx_0) \exp(\alpha t)$$

where A is the initial amplitude, α the amplification factor, and x_0 the initial phase. Show that this expression can be written more concisely as

$$\psi = \mathrm{Re}\left[B e^{ik(x-ct)}\right]$$

where both B and c are complex constants. Determine the real and imaginary parts of B and c in terms of A, α, k, ν, and x_0.

7.3. Several of the wave types discussed in this chapter are governed by equations that are generalizations of the wave equation

$$\frac{\partial^2 \psi}{\partial t^2} = c^2 \frac{\partial^2 \psi}{\partial x^2}$$

This equation can be shown to have solutions corresponding to waves of arbitrary profile moving at the speed c in both positive and negative x directions. We consider an arbitrary initial profile of the field ψ ; $\psi = f(x)$ at $t = 0$. If the profile is translated in the positive x direction at speed c without change of shape, then $\psi = f(x')$, where x' is a coordinate moving at speed c so that $x = x' + ct$. Thus, in terms of the fixed coordinate x we can write $\psi = f(x - ct)$, corresponding to a profile that moves in the positive x direction at speed c without change of shape. Verify that $\psi = f(x - ct)$ is a solution for any arbitrary continuous profile $f(x - ct)$. *Hint*: Let $x - ct = x'$and differentiate f using the chain rule.

7.4. Assuming that the pressure perturbation for a one-dimensional acoustic wave is given by (7.15), find the corresponding solutions of the zonal wind and density perturbations. Express the amplitude and phase for u' and ρ' in terms of the amplitude and phase of p'.

7.5. Show that for isothermal motion ($DT/Dt = 0$) the acoustic wave speed is given by $(gH)^{1/2}$ where $H = RT/g$ is the scale height.

7.6. In Section 7.3.1 the linearized equations for acoustic waves were developed for the special situation of one-dimensional propagation in a horizontal tube. Although this situation does not appear to be directly applicable to the atmosphere, there is a special atmospheric mode, the *Lamb wave*, which is a horizontally propagating acoustic mode with no vertical velocity perturbation ($w' = 0$). Such oscillations have been observed following violent explosions such as volcanic eruptions and atmospheric nuclear tests. Using (7.12), (7.13) plus the linearized forms of the hydrostatic equation, and the continuity equation (7.7) derive the height dependence of the perturbation fields for the Lamb mode in an isothermal basic state atmosphere, assuming that the pressure perturbation at the lower boundary ($z = 0$) has the form (7.15). Determine the vertically integrated kinetic energy density per unit horizontal area for this mode.

7.7. If the surface height perturbation in a shallow water gravity wave is given by

$$h' = \mathrm{Re}\left[Ae^{ik(x-ct)}\right]$$

find the corresponding velocity perturbation $u'(x, t)$. Sketch the phase relationship between h' and u' for an eastward propagating wave.

7.8. Assuming that the vertical velocity perturbation for a two-dimensional internal gravity wave is given by (7.43), obtain the corresponding solution for the u', p', and θ' fields. Use these results to verify the approximation

$$\left|\rho_0\theta'/\overline{\theta}\right| \gg \left|p'/c_s^2\right|$$

which was used in (7.36).

7.9. For the situation in Problem 7.8, express the vertical flux of horizontal momentum, $\rho_0\overline{u'w'}$, in terms of the amplitude A of the vertical velocity perturbation. Hence, show that the momentum flux is positive for waves in which phase speed propagates eastward and downward.

7.10. Show that if (7.38) is replaced by the hydrostatic equation (i.e., the terms in w' are neglected) the resulting frequency equation for internal gravity waves is just the asymptotic limit of (7.44) for waves in which $|k| \ll |m|$.

7.11. (a) Show that the intrinsic group velocity vector in two-dimensional internal gravity waves is parallel to lines of constant phase. (b) Show that in the long-wave limit ($|k| \ll |m|$) the magnitude of the zonal component of the group velocity equals the magnitude of the zonal phase speed so that energy propagates one wavelength per wave period.

7.12. Determine the perturbation horizontal and vertical velocity fields for stationary gravity waves forced by flow over sinusoidally varying topography given the following conditions: the height of the ground is $h = h_0 \cos kx$ where $h_0 = 50$ m is a constant; $N = 2\times10^{-2}\mathrm{s}^{-1}$; $\overline{u} = 5\,\mathrm{m\,s}^{-1}$; and $k = 3 \times 10^{-3}\mathrm{m}^{-1}$. *Hint*: For small amplitude topography ($h_0 k \ll 1$) we can approximate the lower boundary condition by

$$w' = Dh/Dt = \overline{u}\partial h\big/\partial x \quad \text{at z} = 0.$$

7.13. For the topographic gravity wave problem discussed in Section 7.4.2, with vertical velocity given by (7.48) and $\overline{u}k < N$, find the zonal wind perturbation. Compute the vertical momentum flux $\overline{u'w'}$ and show that this flux is a maximum when $k^2\overline{u}^2 = N^2\big/2$. Determine the slope of the phase lines in the x, z plane for this case.

7.14. Verify the group velocity relationship for inertia–gravity waves given in (7.67).

7.15. Show that when $\overline{u} = 0$ the wave number vector $\boldsymbol{\kappa}$ for an internal gravity wave is perpendicular to the group velocity vector.

7.16. Using the linearized form of the vorticity equation (6.18) and the β-plane approximation, derive the Rossby wave speed for a homogeneous incompressible ocean of depth h. Assume a motionless basic state and small perturbations that depend only on x and t,

$$u = u'(x,t), \quad v = v'(x,t), \quad h = H + h'(x,t)$$

where H is the mean depth of the ocean. With the aid of the continuity equation for a homogeneous layer (7.21) and the geostrophic wind relationship $v' = gf_0^{-1}\partial h'/\partial x$, show that the perturbation potential vorticity equation can be written in the form

$$\frac{\partial}{\partial t}\left(\frac{\partial^2}{\partial x^2} - \frac{f_0^2}{gH}\right)h' + \beta\frac{\partial h'}{\partial x} = 0$$

and that $h' = h_0 e^{ik(x-ct)}$ is a solution provided that

$$c = -\beta\left(k^2 + f_0^2/gH\right)^{-1}$$

If the ocean is 4 km deep, what is the Rossby wave speed at latitude 45° for a wave of 10,000 km zonal wavelength?

7.17. Section 4.3 showed that for a homogeneous incompressible fluid a decrease in depth with latitude has the same dynamic effect as a latitudinal dependence of the Coriolis parameter. Thus, Rossby-type waves can be produced in a rotating cylindrical vessel if the depth of the fluid is dependent on the radial coordinate. To determine the Rossby wave speed formula for this *equivalent β-effect*, we assume that the flow is confined between rigid lids in an annular region whose distance from the axis or rotation is large enough so that the curvature terms in the equations can be neglected. We then can refer the motion to Cartesian coordinates with x directed azimuthally and y directed toward the axis of rotation. If the system is rotating at angular velocity Ω and the depth is linearly dependent on y,

$$H(y) = H_0 - \gamma y$$

show that the perturbation continuity equation can be written as

$$H_0 \left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} \right) - \gamma v' = 0$$

and that the perturbation quasi-geostrophic vorticity equation is thus

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

where ψ' is the perturbation geostrophic streamfunction and $\beta = 2\Omega\gamma / H_0$. What is the Rossby wave speed in this situation for waves of wavelength 100 cm in both the x and y directions if $\Omega = 1\ \mathrm{s}^{-1}$, $H_0 = 20$ cm, and $\gamma = 0.05$? *Hint*: Assume that the velocity field is geostrophic except in the divergence term.

7.18. Show by scaling arguments that if the horizontal wavelength is much greater than the depth of the fluid, two-dimensional surface gravity waves will be hydrostatic so that the "shallow water" approximation applies.

7.19. The linearized form of the quasi-geostrophic vorticity equation (6.18) can be written as

$$\left(\frac{\partial}{\partial t} + \overline{u} \frac{\partial}{\partial x} \right) \nabla^2 \psi' + \beta \frac{\partial \psi'}{\partial x} = -f_0 \nabla \bullet \mathbf{V}$$

Suppose that the horizontal divergence field is given by

$$\nabla \bullet \mathbf{V} = A \cos [k(x - ct)]$$

where A is a constant. Find a solution for the corresponding relative vorticity field. What is the phase relationship between vorticity and divergence? For what value of c does the vorticity become infinite?

7.20. Derive an expression for the group velocity of a barotropic Rossby wave with dispersion relation (7.91). Show that for stationary waves the group velocity always has an eastward zonal component relative to the earth. Hence, Rossby wave energy propagation must be downstream of topographic sources.

MATLAB EXERCISES

M7.1. (a) The MATLAB script **phase_demo.m** shows that the Fourier series $F(x) = A\sin(kx) + B\cos(kx)$ is equivalent to the form $F(x) = \text{Re}\,[C\exp(ikx)]$ where A and B are real coefficients and C is a complex coefficient. Modify the MATLAB script to confirm that the expression $F(x) = |C|\cos k(x + x_0) = |C|\cos(kx + \alpha)$ represents the same Fourier series where, $kx_0 \equiv \alpha = \sin^{-1}(C_i/|C|)$ and C_i stands for the imaginary part of $C$and α is the "phase" defined in the MATLAB script. Plot your results as the third subplot in the script. (b) By running the script for several input phase angles (such as 0, 30, 60, and 90°), determine the relationship between α and the location of the maximum of $F(x)$.

M7.2. In this problem you will examine the formation of a "wave envelope" for a combination of dispersive waves. The example is that of a deep water wave in which the group velocity is 1/2 of the phase velocity. The MATLAB script **grp_vel_3.m** has code to show the wave height field at four different times for a group composed of various numbers of waves with differing wave numbers and frequencies. Study the code and determine the period and wavelength of the carrier wave. Then run the script several times varying the number of wave modes from 4 to 32. Determine the half-width of the envelope at time $t = 0$ (top line on graph) as a function of the number of modes in the group. The half-width is here defined as two times the distance from the point of maximum amplitude ($x = 0$) to the point along the envelope where the amplitude is 1/2 the maximum. You can estimate this from the graph using the **ginput** command to determine the distance. Use MATLAB to plot a curve of the half-width versus number of wave modes.

M7.3. Consider stationary gravity waves forced by flow over a sinusoidal lower boundary for a case in which the static stability decreases rapidly with height at about the 6-km level. Thus, the buoyancy frequency is altitude dependent and the simple analytic solution (7.48) no longer applies. The MATLAB script named **linear_grav_wave_1.m** gives a highly accurate numerical solution for this situation. (a) Describe the qualitative change in the wave behavior as the zonal wavelength is changed over the range of 10 to 100 km. Be sure to comment on momentum flux as well as on the vertical velocity. (b) Determine as accurately as you can what the minimum

zonal wavelength is for vertical propagation to occur in the upper part of the domain (above the 6-km level). (c) Determine how the amplitudes of the momentum flux and momentum flux convergence at $z = 6$ km change as the zonal wavelength is increased from 20 to 100 km. Do enough different wavelengths so that you can plot graphs of the dependence of the magnitudes of the momentum flux and zonal force on the wavelength of the sinusoidal topography. Use MATLAB to plot these two graphs.

M7.4. The script **geost_adjust_1.m** together with the function **yprim_adj_1.m.** illustrates one-dimensional geostrophic adjustment of the velocity field in a barotropic model for a sinusoidally varying initial height field. The equations are a simplification of (7.69), (7.70), and (7.71) in the text for the case of no y dependence. Initially $u' = v' = 0,\ \phi' \equiv gh' = 9.8\cos(kx)$. The final balanced wind in this case will have only a meridional component. Run the program **geost_adjust_1.m** for the cases of latitude 30 and 60° choosing a time value of at least 10 days. For each of these cases choose values of wavelength of 2000, 4000, 6000, and 8000 km (a total of eight runs). Construct a table showing the initial and final values of the fields u', v', ϕ' and the ratio of the final energy to the initial energy. You can read the values from the MATLAB graphs by using the **ginput** command or, for greater accuracy, add lines in the MATLAB code to print out the values needed. Modify the MATLAB script to determine the partition of final state energy per unit mass between the kinetic energy $(v'^2/2)$ and the potential energy $\phi'^2/(2gH)$. Compute the ratio of final kinetic energy to potential energy for each of your eight cases and show this in a table.

M7.5. The script **geost_adjust_2.m** together with the function **yprim_adj_2.m**. extends Problem M7.4 by using Fourier expansion to examine the geostrophic adjustment for an isolated initial height disturbance of the form $h_0(x) = -h_m \Big/ \left[1 + (x/L)^2\right]$. The version given here uses 64 Fourier modes and employs a fast Fourier transform algorithm (FFT). (There are 128 modes in the FFT, but only one-half of them provide real information.) In this case you may run the model for only 5 days of integration time (it requires a lot of computation compared to the previous case). Choose an initial zonal scale of the disturbance of 500 km and run the model for latitudes of 15, 30, 45, 60, 75, and 90° (six runs). Study the animations for each case. Note that the zonal flow is entirely in a gravity wave mode that propagates away from the initial disturbance. The meridional flow has a propagating gravity wave component, but also a geostrophic part (cyclonic flow). Use the **ginput** command to estimate the zonal scale of the final geostrophic flow (v component) by measuring the distance from the negative velocity maximum just west of the center to the positive maximum just east of the center. Plot a curve showing the zonal scale as a function of

latitude. Compare this scale with the Rossby radius of deformation defined in the text. (Note that $gH = 400\ m^2\ s^{-2}$ in this example.)

M7.6. This problem examines the variation of phase velocity for Rossby waves as the zonal wavelength is varied. Run the MATLAB program **rossby_1.m** with zonal wavelengths specified as 5000, 10,000 and 20,000 km. For each of these cases try different values of the mean zonal wind until you find the mean wind for which the Rossby wave is approximately stationary.

M7.7. The MATLAB script **rossby_2.m** shows an animation of the Rossby waves generated by a vorticity disturbance initially localized in the center of the domain, with mean wind zero. Run the script and note that the waves excited have westward phase speeds, but that disturbances develop on the eastward side of the original disturbance. By following the development of these disturbances, make a crude estimate of the characteristic wavelength and the group velocity for the disturbances appearing to the east of the original disturbance at time $t = 7.5$ days. (Wavelength can be estimated by using **ginput** to measure the distance between adjacent troughs.) Compare your estimate with the group velocity formula derived from (7.91). Can you think of a reason why your estimate for group velocity may differ from that given in the formula?

M7.8. The MATLAB script **rossby_3.m** gives the surface height and meridional velocity disturbances for topographic Rossby waves generated by flow over an isolated ridge. The program uses a Fourier series approach to the solution. Ekman damping with a 2-day damping time is included to minimize the effect of waves propagating into the mountain from upstream (but this cannot be entirely avoided). Run this program for input zonal mean winds from 10 to 100 m/s at 10-m/s intervals. For each run use **ginput** to estimate the scale of the leeside trough by measuring the zonal distance between the minimum and maximum in the meridional velocity. (This should be equal approximately to one-half the wavelength of the dominant disturbance.) Compare your results in each case with the zonal wavelength for resonance given by solving (7.93) to determine the resonant $L_x = 2\pi/k$, where $K^2 = k^2 + l^2$ and $l = \pi/8 \times 10^6$ in units of 1/m. (Do not expect exact agreement because the actual disturbance corresponds to the sum over many separate zonal wavelengths.)

Suggested References

Hildebrand, *Advanced Calculus for Applications*, is one of many standard textbooks that discuss the mathematical techniques used in this chapter, including the representation of functions in Fourier series and the general properties of the wave equation.

Gill, *Atmosphere-Ocean Dynamics*, has a very complete treatment of gravity, inertia-gravity, and Rossby waves, with particular emphasis on observed oscillations in the oceans.

Smith (1979) discusses many aspects of waves generated by flow over mountains.

Chapman and Lindzen, *Atmospheric Tides*: *Thermal and Gravitational*, is the classic reference on both observational and theoretical aspects of tides, a class of atmospheric motions for which the linear perturbations method has proved to be particularly successful.

Scorer, *Natural Aerodynamics*, contains an excellent qualitative discussion on many aspects of waves generated by barriers such as lee waves.

Nappo, *An Introduction to Atmospheric Gravity Waves* is an excellent introduction to the theory and observation of gravity waves in the atmosphere.

CHAPTER 8

Synoptic-Scale Motions II: Baroclinic Instability

Chapter 6 showed that the quasi-geostrophic system can qualitatively account for the observed relationships among the vorticity, temperature, and vertical velocity fields in midlatitude synoptic-scale systems. The diagnostic approach used in that chapter provided useful insights into the structure of synoptic-scale systems; it also demonstrated the key role of potential vorticity in dynamical analysis. It did not, however, provide quantitative information on the origins, growth rates, and propagation speeds of such disturbances. This chapter shows how linear perturbation analysis can be used to obtain such information.

The development of synoptic-scale weather disturbances is often referred to as *cyclogenesis*, a term that emphasizes the role of relative vorticity in developing synoptic-scale systems. This chapter analyzes the processes that lead to cyclogenesis. Specifically, we discuss the role of dynamical instability of the mean flow in accounting for the growth of synoptic-scale disturbances. We show that the quasi-geostrophic equations can, indeed, provide a reasonable theoretical basis for understanding the development of synoptic-scale storms, although as discussed in Section 9.2, ageostrophic effects must be included to model the development of fronts and subsynoptic-scale storms.

8.1 HYDRODYNAMIC INSTABILITY

A zonal mean flow field is said to be hydrodynamically unstable if a small disturbance introduced into the flow grows spontaneously, drawing energy from the mean flow. It is useful to divide fluid instabilities into two types: parcel instability and wave instability. The simplest example of a parcel instability is the convective overturning that occurs when a fluid parcel is displaced vertically in a statically unstable fluid (see Section 2.7.3). Another example is inertial instability, which occurs when a parcel is displaced radially in an axisymmetric vortex with negative absolute vorticity in the Northern Hemisphere or positive absolute vorticity in the Southern Hemisphere. This instability was discussed in Section 7.5.1. A more general type of parcel instability, called *symmetric instability*, may also be significant in weather disturbances; this is discussed in Section 9.3.

Most of the instabilities of importance in meteorology, however, are associated with wave propagation; they cannot be related easily to the behavior of individual fluid parcels. The wave instabilities important for synoptic-scale meteorology generally occur in the form of zonally asymmetric perturbations to a zonally symmetric basic flow field. In general the basic flow is a jetstream that has both horizontal and vertical mean-flow shears. *Barotropic instability* is a wave instability associated with the horizontal shear in a jet-like current. Barotropic instabilities grow by extracting kinetic energy from the mean-flow field. *Baroclinic instability*, however, is associated with vertical shear of the mean flow. Baroclinic instabilities grow by converting potential energy associated with the mean horizontal temperature gradient that must exist to provide thermal wind balance for the vertical shear in the basic state flow. In neither of these instability types does the parcel method provide a satisfactory stability criterion. A more rigorous approach is required in which a linearized version of the governing equations is analyzed to determine the structure and amplification rate for the various wave modes supported by the system.

As indicated in Problem 2 of Chapter 7, the traditional approach to instability analysis is to assume that a small perturbation consisting of a single Fourier wave mode of the form $\exp[ik(x - ct)]$ is introduced into the flow and to determine the conditions for which the phase velocity c has an imaginary part. This technique, which is called the *normal modes* method, is applied in the next section to analyze the stability of a baroclinic current.

An alternative method of instability analysis is the *initial value* approach. This method is motivated by the recognition that in general the perturbations from which storms develop cannot be described as single normal mode disturbances, but may have a complex structure. The initial growth of such disturbances may strongly depend on the potential vorticity distribution in the initial disturbance. On the time scale of a day or two, such growth can be quite different from that of a normal mode of similar scale, although in the absence of nonlinear interactions the fastest-growing normal mode disturbance must eventually dominate.

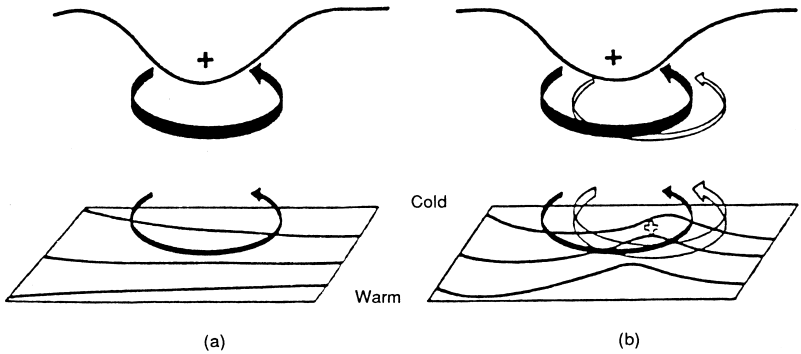


Fig. 8.1 A schematic picture of cyclogenesis associated with the arrival of an upper-level positive vorticity perturbation over a lower-level baroclinic region. (a) Lower-level cyclonic vorticity induced by the upper-level vorticity anomaly. The circulation induced by the vorticity anomaly is shown by the solid arrows, and potential temperature contours are shown at the lower boundary. The advection of potential temperature by the induced lower-level circulation leads to a warm anomaly slightly east of the upper-level vorticity anomaly. This in turn will induce a cyclonic circulation as shown by the open arrows in (b). The induced upper-level circulation will reinforce the original upper-level anomaly and can lead to amplification of the disturbance. (After Hoskins et al., 1985.)

A strong dependence of cyclogenesis on initial conditions occurs when a large-amplitude upper-level potential vorticity anomaly is advected into a region where there is a preexisting meridional temperature gradient at the surface. In that case, as shown schematically in Fig. 8.1, the circulation induced by the upper-level anomaly (which extends downward, as discussed in Section 6.3) leads to temperature advection at the surface; this induces a potential vorticity anomaly near the surface, which in turn reinforces the upper-level anomaly. Under some conditions the surface and upper-level potential vorticity anomalies can become locked in phase so that the induced circulations produce a very rapid amplification of the anomaly pattern. Detailed discussion of the initial value approach to cyclogenesis is beyond the scope of this text. Here we concentrate primarily on the simplest normal mode instability models.

8.2 NORMAL MODE BAROCLINIC INSTABILITY: A TWO-LAYER MODEL

Even for a highly idealized mean-flow profile, the mathematical treatment of baroclinic instability in a continuously stratified atmosphere is rather complicated. Before considering such a model we first focus on the simplest model that can incorporate baroclinic processes. The atmosphere is represented by two discrete layers bounded by surfaces numbered 0, 2, and 4 (generally taken to be the 0-, 500-, and 1000-hPa surfaces, respectively) as shown in Fig. 8.2. The quasi-geostrophic

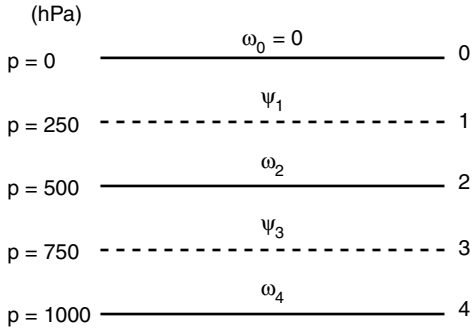


Fig. 8.2 Arrangement of variables in the vertical for the two-level baroclinic model.

vorticity equation for the midlatitude β plane is applied at the 250- and 750-hPa levels, designated by 1 and 3 in Fig. 8.2, whereas the thermodynamic energy equation is applied at the 500-hPa level, designated by 2 in Fig. 8.2.

Before writing out the specific equations of the two-layer model, it is convenient to define a *geostrophic streamfunction*, $\psi \equiv \Phi/f_0$. Then the geostrophic wind (6.7) and the geostrophic vorticity (6.15) can be expressed respectively as

$$\mathbf{V}_\psi = \mathbf{k} \times \boldsymbol{\nabla}\psi, \quad \zeta_\mathbf{g} = \nabla^2\psi \tag{8.1}$$

The quasi-geostrophic vorticity equation (6.19) and the hydrostatic thermodynamic energy equation (6.13) can then be written in terms of ψ and ω as

$$\frac{\partial}{\partial t}\nabla^2\psi + \mathbf{V}_\psi \bullet \boldsymbol{\nabla}\left(\nabla^2\psi\right) + \beta\frac{\partial\psi}{\partial x} = f_0\frac{\partial\omega}{\partial p} \tag{8.2}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial\psi}{\partial p}\right) = -\mathbf{V}_\psi \bullet \boldsymbol{\nabla}\left(\frac{\partial\psi}{\partial p}\right) - \frac{\sigma}{f_0}\omega \tag{8.3}$$

We now apply the vorticity equation (8.2) at the two levels designated as 1 and 3, which are at the middle of the two layers. To do this we must estimate the divergence term $\partial\omega/\partial p$ at these levels using finite difference approximations to the vertical derivatives:

$$\left(\frac{\partial\omega}{\partial p}\right)_1 \approx \frac{\omega_2 - \omega_0}{\delta p}, \left(\frac{\partial\omega}{\partial p}\right)_3 \approx \frac{\omega_4 - \omega_2}{\delta p} \tag{8.4}$$

where $\delta p = 500$ hPa is the pressure interval between levels 0–2 and 2–4, and subscript notation is used to designate the vertical level for each dependent variable.

The resulting vorticity equations are

$$\frac{\partial}{\partial t}\nabla^2\psi_1 + \mathbf{V}_1 \bullet \boldsymbol{\nabla}\left(\nabla^2\psi_1\right) + \beta\frac{\partial\psi_1}{\partial x} = \frac{f_0}{\delta p}\omega_2 \tag{8.5}$$

$$\frac{\partial}{\partial t}\nabla^2\psi_3 + \mathbf{V}_3 \bullet \boldsymbol{\nabla}\left(\nabla^2\psi_3\right) + \beta\frac{\partial\psi_3}{\partial x} = -\frac{f_0}{\delta p}\omega_2 \tag{8.6}$$

where we have used the fact that $\omega_0 = 0$, and assumed that $\omega_4 = 0$, which is approximately true for a level-lower boundary surface.

We next write the thermodynamic energy equation (8.3) at level 2. Here we must evaluate $\partial\psi/\partial p$ using the difference formula

$$(\partial\psi/\partial p)_2 \approx (\psi_3 - \psi_1)/\delta p$$

The result is

$$\frac{\partial}{\partial t}(\psi_1 - \psi_3) = -\mathbf{V}_2 \bullet \boldsymbol{\nabla}(\psi_1 - \psi_3) + \frac{\sigma\delta p}{f_0}\omega_2 \tag{8.7}$$

The first term on the right-hand side in (8.7) is the advection of the 250- to 750-hPa thickness by the wind at 500 hPa. However, ψ_2, the 500-hPa streamfunction, is not a predicted field in this model. Therefore, ψ_2 must be obtained by linearly interpolating between the 250- and 750-hPa levels:

$$\psi_2 = (\psi_1 + \psi_3)/2$$

If this interpolation formula is used, (8.5)–(8.7) become a closed set of prediction equations in the variables ψ_1, ψ_3, and ω_2.

8.2.1 Linear Perturbation Analysis

To keep the analysis as simple as possible, we assume that the streamfunctions ψ_1 and ψ_3 consist of basic state parts that depend linearly on y alone, plus perturbations that depend only on x and t. Thus, we let

$$\begin{aligned}\psi_1 &= -U_1 y + \psi_1'(x,t)\\ \psi_3 &= -U_3 y + \psi_3'(x,t)\\ \omega_2 &= \omega_2'(x,t)\end{aligned} \tag{8.8}$$

The zonal velocities at levels 1 and 3 are then constants with the values U_1 and U_3, respectively. Hence, the perturbation field has meridional and vertical velocity components only.

Substituting from (8.8) into (8.5)–(8.7) and linearizing yields the perturbation equations

$$\left(\frac{\partial}{\partial t}+U_1\frac{\partial}{\partial x}\right)\frac{\partial^2\psi_1'}{\partial x^2}+\beta\frac{\partial\psi_1'}{\partial x}=\frac{f_0}{\delta p}\omega_2' \tag{8.9}$$

$$\left(\frac{\partial}{\partial t}+U_3\frac{\partial}{\partial x}\right)\frac{\partial^2\psi_3'}{\partial x^2}+\beta\frac{\partial\psi_3'}{\partial x}=-\frac{f_0}{\delta p}\omega_2' \tag{8.10}$$

$$\left(\frac{\partial}{\partial t}+U_m\frac{\partial}{\partial x}\right)\left(\psi_1'-\psi_3'\right)-U_T\frac{\partial}{\partial x}\left(\psi_1'+\psi_3'\right)=\frac{\sigma\delta p}{f_0}\omega_2' \tag{8.11}$$

where we have linearly interpolated to express $\mathbf{V}_2$ in terms of ψ_1 and ψ_3, and have defined

$$U_m\equiv\left(U_1+U_3\right)/2,\quad U_T\equiv\left(U_1-U_3\right)/2$$

Thus, U_m and U_T are, respectively, the vertically averaged mean zonal wind and the mean thermal wind.

The dynamical properties of this system are more clearly expressed if (8.9)–(8.11) are combined to eliminate ω_2'. We first note that (8.9) and (8.10) can be rewritten as

$$\left[\frac{\partial}{\partial t}+\left(U_m+U_T\right)\frac{\partial}{\partial x}\right]\frac{\partial^2\psi_1'}{\partial x^2}+\beta\frac{\partial\psi_1'}{\partial x}=\frac{f_0}{\delta p}\omega_2' \tag{8.12}$$

$$\left[\frac{\partial}{\partial t}+\left(U_m-U_T\right)\frac{\partial}{\partial x}\right]\frac{\partial^2\psi_3'}{\partial x^2}+\beta\frac{\partial\psi_3'}{\partial x}=-\frac{f_0}{\delta p}\omega_2' \tag{8.13}$$

We now define the barotropic and baroclinic perturbations as

$$\psi_m\equiv\left(\psi_1'+\psi_3'\right)/2;\qquad \psi_T\equiv\left(\psi_1'-\psi_3'\right)/2 \tag{8.14}$$

Adding (8.12) and (8.13) and using the definitions in (8.14) yield

$$\left[\frac{\partial}{\partial t}+U_m\frac{\partial}{\partial x}\right]\frac{\partial^2\psi_m}{\partial x^2}+\beta\frac{\partial\psi_m}{\partial x}+U_T\frac{\partial}{\partial x}\left(\frac{\partial^2\psi_T}{\partial x^2}\right)=0 \tag{8.15}$$

while subtracting (8.13) from (8.12) and combining with (8.11) to eliminate ω_2' yield

$$\left[\frac{\partial}{\partial t}+U_m\frac{\partial}{\partial x}\right]\left(\frac{\partial^2\psi_T}{\partial x^2}-2\lambda^2\psi_T\right)+\beta\frac{\partial\psi_T}{\partial x}+U_T\frac{\partial}{\partial x}\left(\frac{\partial^2\psi_m}{\partial x^2}+2\lambda^2\psi_m\right)=0 \tag{8.16}$$

where $\lambda^2\equiv f_0^2/[\sigma(\delta p)^2]$. Equations (8.15) and (8.16) govern the evolution of the barotropic (vertically averaged) and baroclinic (thermal) perturbation vorticities, respectively.

As in Chapter 7 we assume that wave-like solutions exist of the form

$$\psi_m = A\,e^{ik(x-ct)},\ \ \psi_{\mathrm{T}} = B\,e^{ik(x-ct)} \tag{8.17}$$

Substituting these assumed solutions into (8.15) and (8.16) and dividing through by the common exponential factor, we obtain a pair of simultaneous linear algebraic equations for the coefficients A, B:

$$ik\left[\left(c-U_m\right)k^2+\beta\right]A - ik^3U_T\,B = 0 \tag{8.18}$$

$$ik\left[\left(c-U_m\right)\left(k^2+2\lambda^2\right)+\beta\right]B - ikU_T\left(k^2-2\lambda^2\right)A = 0 \tag{8.19}$$

Because this set is homogeneous, nontrivial solutions will exist only if the determinant of the coefficients of A and B is zero. Thus the phase speed c must satisfy the condition

$$\begin{vmatrix} \left(c-U_m\right)k^2+\beta & -k^2U_T \\ -U_T\left(k^2-2\lambda^2\right) & \left(c-U_m\right)\left(k^2+2\lambda^2\right)+\beta \end{vmatrix} = 0$$

which gives a quadratic dispersion equation in c:

$$\begin{aligned}\left(c-U_m\right)^2k^2\left(k^2+2\lambda^2\right)+2\left(c-U_m\right)\beta\left(k^2+\lambda^2\right)+\\ \left[\beta^2+U_T^2k^2\left(2\lambda^2-k^2\right)\right]=0\end{aligned} \tag{8.20}$$

which is analogous to the linear wave dispersion equations developed in Chapter 7. The dispersion relationship in (8.20) yields for the phase speed

$$c = U_m - \frac{\beta\left(k^2+\lambda^2\right)}{k^2\left(k^2+2\lambda^2\right)} \pm \delta^{1/2} \tag{8.21}$$

where

$$\delta \equiv \frac{\beta^2\lambda^4}{k^4\left(k^2+2\lambda^2\right)^2} - \frac{U_T^2\left(2\lambda^2-k^2\right)}{\left(k^2+2\lambda^2\right)}$$

We have now shown that (8.17) is a solution for the system (8.15) and (8.16) only if the phase speed satisfies (8.21). Although (8.21) appears to be rather complicated, it is immediately apparent that if $\delta < 0$ the phase speed will have an imaginary part and the perturbations will amplify exponentially. Before discussing the general physical conditions required for exponential growth it is useful to consider two special cases.

As the first special case we let $U_T = 0$ so that the basic state thermal wind vanishes and the mean flow is barotropic. The phase speeds in this case are

$$c_1 = U_m - \beta k^{-2} \tag{8.22}$$

and

$$c_2 = U_m - \beta \left(k^2 + 2\lambda^2\right)^{-1} \tag{8.23}$$

These are real quantities that correspond to the free (normal mode) oscillations for the two-level model with a barotropic basic state current. The phase speed c_1 is simply the dispersion relationship for a barotropic Rossby wave with no y dependence (see Section 7.7). Substituting from (8.22) for c in (8.18) and (8.19) we see that in this case $B = 0$ so that the perturbation is barotropic in structure. The expression of (8.23), however, may be interpreted as the phase speed for an internal baroclinic Rossby wave. Note that c_2 is a dispersion relationship analogous to the Rossby wave speed for a homogeneous ocean with a free surface, which was given in Problem 7.16. However, in the two-level model, the factor $2\lambda^2$ appears in the denominator in place of the f_0^2/gH for the oceanic case. In each of these cases there is vertical motion associated with the Rossby wave so that static stability modifies the wave speed. It is left as a problem for the reader to show that if c_2 is substituted into (8.18) and (8.19), the resulting fields of ψ_1 and ψ_3 are 180° out of phase so that the perturbation is baroclinic, although the basic state is barotropic. Furthermore, the ω_2' field is 1/4 cycle out of phase with the 250-hPa geopotential field, with the maximum upward motion occurring west of the 250-hPa trough.

This vertical motion pattern may be understood if we note that $c_2 - U_m < 0$, so that the disturbance pattern moves westward *relative to the mean wind*. Now, viewed in a coordinate system moving with the mean wind the vorticity changes are due only to the planetary vorticity advection and the convergence terms, while the thickness changes must be caused solely by the adiabatic heating or cooling due to vertical motion. Hence, there must be rising motion west of the 250-hPa trough in order to produce the thickness changes required by the westward motion of the system.

Comparing (8.22) and (8.23) we see that the phase speed of the baroclinic mode is generally much less than that of the barotropic mode since for average midlatitude tropospheric conditions $\lambda^2 \approx 2 \times 10^{-12}\text{m}^{-2}$, which is approximately equal to k^2 for zonal wavelength of 4300 km.[1]

As the second special case, we assume that $\beta = 0$. This case corresponds, for example, to a laboratory situation in which the fluid is bounded above and below by

[1] The presence of the free internal Rossby wave should actually be regarded as a weakness of the two-level model. Lindzen et al. (1968) have shown that this mode does not correspond to any free oscillation of the real atmosphere. Rather, it is a spurious mode resulting from the use of the upper boundary condition $\omega = 0$ at $p = 0$, which formally turns out to be equivalent to putting a lid at the top of the atmosphere.

rotating horizontal planes so that the gravity and rotation vectors are everywhere parallel. In such a situation

$$c = U_m \pm U_T \left(\frac{k^2 - 2\lambda^2}{k^2 + 2\lambda^2} \right)^{1/2} \tag{8.24}$$

For waves with zonal wave numbers satisfying $k^2 < 2\lambda^2$, (8.24) has an imaginary part. Thus, all waves longer than the critical wavelength $L_c = \sqrt{2}\pi/\lambda$ will amplify. From the definition of λ we can write

$$L_c = \delta p \pi \, (2\sigma)^{1/2} / f_0$$

For typical tropospheric conditions $(2\sigma)^{1/2} \approx 2 \times 10^{-3}\,\mathrm{N^{-1}m^3\,s^{-1}}$. Thus, with $\delta p = 500$ hPa and $f_0 = 10^{-4}\,\mathrm{s^{-1}}$ we find that $L_c \approx 3000$ km. It is also clear from this formula that the critical wavelength for baroclinic instability increases with static stability. The role of static stability in stabilizing the shorter waves can be understood qualitatively as follows: For a sinusoidal perturbation, the relative vorticity, and hence the differential vorticity advection, increases with the square of the wave number. However, as shown in Chapter 6, a secondary vertical circulation is required to maintain hydrostatic temperature changes and geostrophic vorticity changes in the presence of differential vorticity advection. Thus, for a geopotential perturbation of fixed amplitude the relative strength of the accompanying vertical circulation must increase as the wavelength of the disturbance decreases. Because static stability tends to resist vertical displacements, the shortest wavelengths will thus be stabilized.

It is also of interest that with $\beta = 0$ the critical wavelength for instability does not depend on the magnitude of the basic state thermal wind U_T. The growth rate, however, does depend on U_T. According to (8.17) the time dependence of the disturbance solution has the form exp(–*ikct*). Thus, the exponential growth rate is $\alpha = kc_i$, where c_i designates the imaginary part of the phase speed. In the present case

$$\alpha = kU_T \left(\frac{2\lambda^2 - k^2}{2\lambda^2 + k^2} \right)^{1/2} \tag{8.25}$$

so that the growth rate increases linearly with the mean thermal wind.

Returning to the general case where all terms are retained in (8.21), the stability criterion is understood most easily by computing the *neutral curve*, which connects all values of U_T and k for which $\delta = 0$ so that the flow is *marginally stable*. From (8.21), the condition $\delta = 0$ implies that

$$\frac{\beta^2 \lambda^4}{k^4 \left(2\lambda^2 + k^2\right)} = U_T^2 \left(2\lambda^2 - k^2\right) \tag{8.26}$$

This complicated relationship between U_T and k can best be displayed by solving (8.26) for $k^4/2\lambda^4$, yielding

$$k^4/\left(2\lambda^4\right) = 1 \pm \left[1 - \beta^2/\left(4\lambda^4 U_T^2\right)\right]^{1/2}$$

In Fig. 8.3 the nondimensional quantity $k^2/2\lambda^2$, which is proportional to the square of the zonal wave number, is plotted against the nondimensional parameter $2\lambda^2 U_T/\beta$, which is proportional to the thermal wind. As indicated in Fig. 8.3, the neutral curve separates the unstable region of the U_T, k plane from the stable region. It is clear that the inclusion of the β effect serves to stabilize the flow, for now unstable roots exist only for $|U_T| > \beta/(2\lambda^2)$. In addition, the minimum value of U_T required for unstable growth depends strongly on k. Thus, the β effect strongly stabilizes the long wave end of the wave spectrum ($k \to 0$). Again the flow is always stable for waves shorter than the critical wavelength $L_c = \sqrt{2}\pi/\lambda$.

This long wave stabilization associated with the β effect is caused by the rapid westward propagation of long waves (i.e., Rossby wave propagation), which occurs only when the β effect is included in the model. It can be shown that baroclinically unstable waves always propagate at a speed that lies between maximum and minimum mean zonal wind speeds. Thus, for the two-level model in the usual midlatitude case where $U_1 > U_3 > 0$, the real part of the phase speed satisfies the inequality $U_3 < c_r < U_1$ for unstable waves. In a continuous atmosphere, this would imply that there must be a level where $U = c_r$. Such a level is called a *critical*

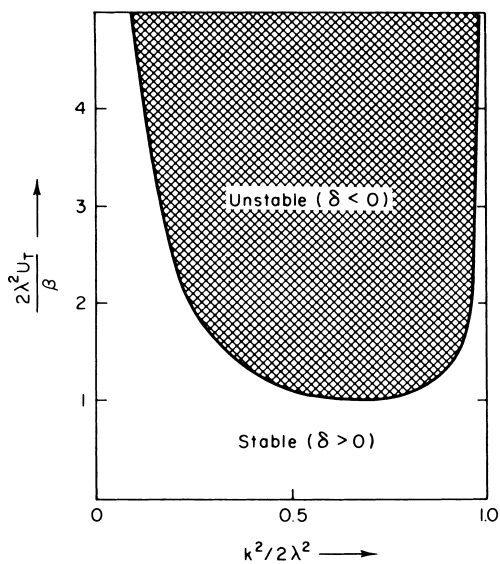


Fig. 8.3 Neutral stability curve for the two-level baroclinic model.

level by theoreticians and a *steering level* by synopticians. For long waves and weak basic state wind shear, the solution given by (8.21) will have $c_r < U_3$, there is no steering level, and unstable growth cannot occur.

Differentiating (8.26) with respect to k and setting $dU_T/dk = 0$ we find that the minimum value of U_T for which unstable waves may exist occurs when $k^2 = \sqrt{2}\lambda^2$. This wave number corresponds to the wave of maximum instability. Wave numbers of observed disturbances should be close to the wave number of maximum instability, for if U_T were gradually raised from zero the flow would first become unstable for perturbations of wave number $k = 2^{1/4}\lambda$. Those perturbations would then amplify and in the process remove energy from the mean thermal wind, thereby decreasing U_T and stabilizing the flow. Under normal conditions of static stability the wavelength of maximum instability is about 4000 km, which is close to the average wavelength for midlatitude synoptic systems. Furthermore, the thermal wind required for marginal stability at this wavelength is only about $U_T \approx 4\ \mathrm{m\,s^{-1}}$, which implies a shear of $8\ \mathrm{m\,s^{-1}}$ between 250 and 750 hPa. Shears greater than this are certainly common in middle latitudes for the zonally averaged flow. Therefore, the observed behavior of midlatitude synoptic systems is consistent with the hypothesis that such systems can originate from infinitesimal perturbations of a baroclinically unstable basic current. Of course in the real atmosphere many other factors may influence the development of synoptic systems, e.g., instabilities due to lateral shear in the jetstream, nonlinear interactions of finite amplitude perturbations, and the release of latent heat in precipitating systems. However, observational studies, laboratory simulations, and numerical models all suggest that baroclinic instability is a primary mechanism for synoptic-scale wave development in middle latitudes.

8.2.2 Vertical Motion in Baroclinic Waves

Since the two-level model is a special case of the quasi-geostrophic system, the physical mechanisms responsible for forcing vertical motion should be those discussed in Section 6.4. Thus, the forcing of vertical motion can be expressed in terms of the sum of the forcing by thermal advection (evaluated at level 2) plus the differential vorticity advection (evaluated as the difference between the vorticity advection at level 1 and that at level 3). Alternatively, the forcing of vertical motion can be expressed in terms of the divergence of the **Q** vector.

The **Q** vector form of the omega equation for the two-level model can be derived simply from (6.35). We first estimate the second term on the left-hand side by finite differencing in p. Using (8.4) and again letting $\omega_0 = \omega_4 = 0$, we obtain

$$\frac{\partial^2\omega}{\partial p^2} \approx \frac{(\partial\omega/\partial p)_3 - (\partial\omega/\partial p)_1}{\delta p} \approx -\frac{2\omega_2}{(\delta p)^2}$$

and observe that temperature in the two-level model is represented as

$$\frac{RT}{p} = -\frac{\partial \Phi}{\partial p} \approx \frac{f_0}{\delta p}(\psi_1 - \psi_3)$$

Thus, (6.35) becomes

$$\sigma\left(\nabla^2 - 2\lambda^2\right)\omega_2 = -2\boldsymbol{\nabla} \bullet \mathbf{Q} \tag{8.27}$$

where

$$\mathbf{Q} = \frac{f_0}{\delta p}\left[-\frac{\partial \mathbf{V_2}}{\partial x} \bullet \boldsymbol{\nabla}(\psi_1 - \psi_3)\,, \quad -\frac{\partial \mathbf{V_2}}{\partial y} \bullet \boldsymbol{\nabla}(\psi_1 - \psi_3)\right]$$

In order to examine the forcing of vertical motion in baroclinically unstable waves, we linearize (8.27) by specifying the same basic state and perturbation variables as in (8.8). For this situation, in which the mean zonal wind and the perturbation streamfunctions are independent of y, the $\mathbf{Q}$ vector has only an x component:

$$Q_1 = \frac{f_0}{\delta p}\left[\frac{\partial^2 \psi_2'}{\partial x^2}(U_1 - U_3)\right] = \frac{2 f_0}{\delta p} U_T \zeta_2'$$

The pattern of the $\mathbf{Q}$ vector in this case is similar to that of Fig. 6.13, with eastward pointing $\mathbf{Q}$ centered at the trough and westward pointing $\mathbf{Q}$ centered at the ridge. This is consistent with the fact that $\mathbf{Q}$ represents the change of temperature gradient forced by geostrophic motion alone. In this simple model the temperature gradient is entirely due to the vertical shear of the mean zonal wind $[U_T \propto -\partial T/\partial y]$ and the shear of the perturbation meridional velocity tends to advect warm air poleward east of the 500-hPa trough and cold air equatorward west of the 500-hPa trough so that there is a tendency to produce a component of temperature gradient directed eastward at the trough.

The forcing of vertical motion by the $\mathbf{Q}$ vector in the linearized model is from (8.27)

$$\left(\frac{\partial^2}{\partial x^2} - 2\lambda^2\right)\omega_2' = -\frac{4 f_0}{\sigma \delta p} U_T \frac{\partial \zeta_2'}{\partial x} \tag{8.28}$$

Observing that

$$\left(\frac{\partial^2}{\partial x^2} - 2\lambda^2\right)\omega_2' \propto -\omega_2'$$

we may interpret (8.28) physically by noting that

$$w_2' \propto -\omega_2' \propto -U_T \frac{\partial \zeta_2'}{\partial x} \propto -v_2' \frac{\partial \overline{T}}{\partial y}$$

Thus, sinking motion is forced by negative advection of disturbance vorticity by the basic state thermal wind (or alternatively, cold advection of the basic state thermal field by the perturbation meridional wind), while rising motion is forced by advection of the opposite sign.

We now have the information required to diagram the structure of a baroclinically unstable disturbance in the two-level model. The lower part of Fig. 8.4 shows schematically the phase relationship between the geopotential field and the divergent secondary motion field for the usual midlatitude situation where $U_T > 0$. Linear interpolation has been used between levels so that the trough and ridge axes are straight lines tilted back toward the west with height. In this example the ψ_1 field lags the ψ_3 field by about 65° in phase so that the trough at 250 hPa lies 65° in phase west of the 750-hPa trough. At 500 hPa the perturbation thickness field lags the geopotential field by one-quarter wavelength as shown in the top part of Fig. 8.4 and the thickness and vertical motion fields are in phase. Note that the temperature advection by the perturbation meridional wind is in phase with the 500-hPa thickness field so that the advection of the basic state temperature by the

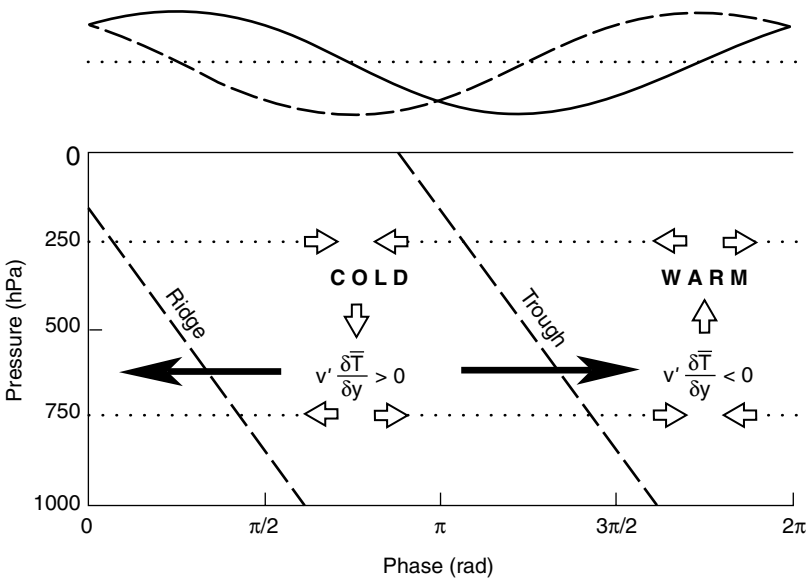


Fig. 8.4 Structure of an unstable baroclinic wave in the two-level model. (Top) Relative phases of the 500-hPa perturbation geopotential (solid line) and temperature (dashed line). (Bottom) Vertical cross section showing phases of geopotential, meridional temperature advection, ageostrophic circulation (open arrows), **Q** vectors (solid arrows), and temperature fields for an unstable baroclinic wave in the two-level model.

perturbation wind acts to intensify the perturbation thickness field. This tendency is also illustrated by the zonally oriented **Q** vectors shown at the 500 hPa-level in Fig. 8.4.

The pattern of vertical motion forced by the divergence of the **Q** vector, as shown in Fig. 8.4, is associated with a divergence–convergence pattern that contributes a positive vorticity tendency near the 250-hPa trough and a negative vorticity tendency near the 750-hPa ridge, with opposite tendencies at the 250-hPa ridge and 750-hPa trough. Since in all cases these vorticity tendencies tend to increase the extreme values of vorticity at the troughs and ridges, this secondary circulation system will act to increase the strength of the disturbance.

The total vorticity change at each level is, of course, determined by the sum of vorticity advection and vortex stretching due to the divergent circulation. The relative contributions of these processes are indicated schematically in Figs. 8.5 and 8.6, respectively. As can be seen in Fig. 8.5, vorticity advection leads the vorticity field by one-quarter wavelength. Since in this case the basic state wind increases with height, the vorticity advection at 250 hPa is larger than that at 750 hPa. If no other processes influenced the vorticity field, the effect of this differential vorticity advection would be to move the upper-level trough and ridge pattern eastward more rapidly than the lower-level pattern. Thus, the westward tilt of the trough–ridge pattern would quickly be destroyed. The maintenance of this tilt in the presence of differential vorticity advection is due to the concentration of vorticity by vortex stretching associated with the divergent secondary circulation. Referring to Fig. 8.6, we see that concentration of vorticity by the divergence effect lags the vorticity field by about 65° at 250 hPa and leads the vorticity field by about 65° at 750 hPa. As a result, the net vorticity tendencies ahead of the

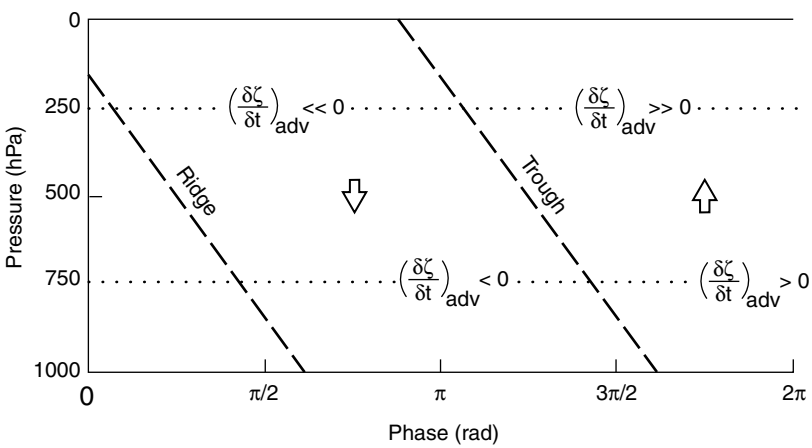


Fig. 8.5 Vertical cross section showing phase of vorticity change due to vorticity advection for an unstable baroclinic wave in the two-level model.

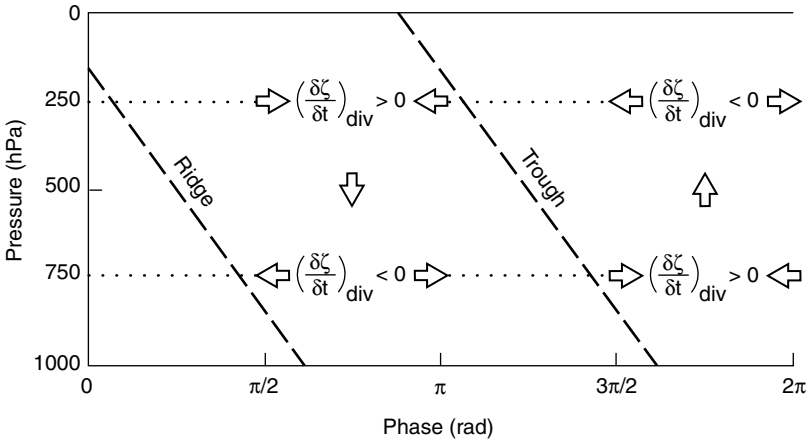


Fig. 8.6 Vertical cross section showing phase of vorticity change due to divergence–convergence for an unstable baroclinic wave in the two-level model.

vorticity maxima and minima are less than the advective tendencies at the upper level and greater than the advective tendencies at the lower level:

$$\left|\frac{\partial \zeta_1'}{\partial t}\right| < \left|\frac{\partial \zeta_1'}{\partial t}\right|_{\text{adv}}, \quad \left|\frac{\partial \zeta_3'}{\partial t}\right| > \left|\frac{\partial \zeta_3'}{\partial t}\right|_{\text{adv}}$$

Furthermore, vorticity concentration by the divergence effect will tend to amplify the vorticity perturbations in the troughs and ridges at both the 250- and 750-hPa levels as required for a growing disturbance.

8.3 THE ENERGETICS OF BAROCLINIC WAVES

The previous section showed that under suitable conditions a vertically sheared geostrophically balanced basic state flow is unstable to small wave-like perturbations with horizontal wavelengths in the range of observed synoptic scale systems. Such baroclinically unstable perturbations will amplify exponentially by drawing energy from the mean flow. This section considers the energetics of linearized baroclinic disturbances and shows that the potential energy of the mean flow is the energy source for baroclinically unstable perturbations.

8.3.1 Available Potential Energy

Before discussing the energetics of baroclinic waves, it is necessary to consider the energy of the atmosphere from a more general point of view. For all practical

purposes, the total energy of the atmosphere is the sum of internal energy, gravitational potential energy, and kinetic energy. However, it is not necessary to consider separately the variations of internal and gravitational potential energy because in a hydrostatic atmosphere these two forms of energy are proportional and may be combined into a single term called the *total potential energy*. The proportionality of internal and gravitational potential energy can be demonstrated by considering these forms of energy for a column of air of unit horizontal cross section extending from the surface to the top of the atmosphere.

If we let dE_I be the internal energy in a vertical section of the column of height dz, then from the definition of internal energy [see (2.4)]

$$dE_I = \rho c_v T dz$$

so that the internal energy for the entire column is

$$E_I = c_v \int_0^\infty \rho T dz \tag{8.29}$$

However, the gravitational potential energy for a slab of thickness dz at a height z is just

$$dE_P = \rho g z dz$$

so that the gravitational potential energy in the entire column is

$$E_P = \int_0^\infty \rho g z dz = -\int_{p_0}^0 z dp \tag{8.30}$$

where we have substituted from the hydrostatic equation to obtain the last integral in (8.30). Integrating (8.30) by parts and using the ideal gas law we obtain

$$E_P = \int_0^\infty p dz = R \int_0^\infty \rho T dz \tag{8.31}$$

Comparing (8.29) and (8.31) we see that $c_v E_P = R E_I$. Thus, the total potential energy may be expressed as

$$E_P + E_I = \left(c_p/c_v\right) E_I = \left(c_p/R\right) E_P \tag{8.32}$$

Therefore, in a hydrostatic atmosphere the total potential energy can be obtained by computing either E_I or E_P alone.

The total potential energy is not a very useful measure of energy in the atmosphere because only a very small fraction of the total potential energy is available for conversion to kinetic energy in storms. To demonstrate qualitatively why most of the total potential energy is unavailable, we consider a simple model atmosphere that initially consists of two equal masses of dry air separated by a vertical partition as shown in Fig. 8.7. The two air masses are at uniform potential temperatures θ_1 and θ_2, respectively, with $\theta_1 < \theta_2$. The ground level pressure on each side of the partition is taken to be 1000 hPa. We wish to compute the maximum kinetic energy that can be realized by an adiabatic rearrangement of mass within the same volume when the partition is removed.

Now for an adiabatic process, total energy is conserved:

$$E_K + E_P + E_I = \text{constant}$$

where E_K denotes the kinetic energy. If the air masses are initially at rest $E_K = 0$. Thus, if we let primed quantities denote the final state

$$E'_K + E'_P + E'_I = E_P + E_I$$

so that with the aid of (8.32) the kinetic energy realized by removal of the partition may be expressed as

$$E'_K = \left(c_p/c_v\right)\left(E_I - E'_I\right)$$

Because θ is conserved for an adiabatic process, the two air masses cannot mix. It is clear that E'_I will be a minimum (designated by E''_I) when the masses are rearranged so that the air at potential temperature θ_1 lies entirely beneath the air at potential temperature θ_2 with the 500-hPa surface as the horizontal boundary between the two masses. In that case the total potential energy $(c_p/c_v)E''_I$ is not

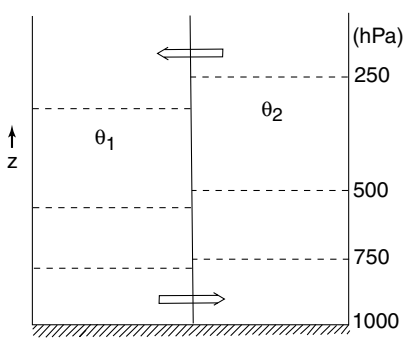


Fig. 8.7 Two air masses of differing potential temperature separated by a vertical partition. Dashed lines indicate isobaric surfaces. Arrows show direction of motion when the partition is removed.

available for conversion to kinetic energy because no adiabatic process can further reduce E_I''.

The *available potential energy* (APE) can now be defined as the difference between the total potential energy of a closed system and the minimum total potential energy that could result from an adiabatic redistribution of mass. Thus, for the idealized model given earlier, the APE, which is designated by the symbol P, is

$$P = \left(c_p/c_v\right)\left(E_I - E_I''\right) \tag{8.33}$$

which is equivalent to the maximum kinetic energy that can be realized by an adiabatic process.

Lorenz (1960) showed that available potential energy is given approximately by the volume integral over the entire atmosphere of the variance of potential temperature on isobaric surfaces. Thus, letting $\overline{\theta}$ designate the average potential temperature for a given pressure surface and θ' the local deviation from the average, the average available potential energy per unit volume satisfies the proportionality

$$\overline{P} \propto V^{-1} \int \left(\overline{\theta'^2}/\bar{\theta}^2\right) dV$$

where V designates the total volume. For the quasi-geostrophic model, this proportionality is an exact measure of the available potential energy, as shown in the following subsection.

Observations indicate that for the atmosphere as a whole

$$\overline{P}/\left[\left(c_p/c_v\right)\overline{E}_I\right] \sim 5 \times 10^{-3}, \quad \overline{K}/\overline{P} \sim 10^{-1}$$

Thus only about 0.5% of the total potential energy of the atmosphere is available, and of the available portion only about 10% is actually converted to kinetic energy. From this point of view the atmosphere is a rather inefficient heat engine.

8.3.2 Energy Equations for the Two-Layer Model

In the two-layer model of Section 8.2, the perturbation temperature field is proportional to $\psi_1' - \psi_3'$, the 250- to 750-hPa thickness. Thus, in view of the discussion in the previous section, we anticipate that the available potential energy in this case will be proportional to $(\psi_1' - \psi_3')^2$. To show that this in fact must be the case, we derive the energy equations for the system in the following manner: We first multiply (8.9) by $-\psi_1$, (8.10) by $-\psi_3$, and (8.11) by $(\psi_1' - \psi_3')$. We then integrate the resulting equations over one wavelength of the perturbation in the zonal direction.

The resulting zonally averaged[2] terms are denoted by overbars as done previously in Chapter 7:

$$\overline{(\ )} = L^{-1}\int_0^L (\ )\,dx$$

where L is the wavelength of the perturbation. Thus, for the first term in (8.9) we have, after multiplying by $-\psi_1$, averaging and differentiating by parts:

$$-\overline{\psi_1'\frac{\partial}{\partial t}\left(\frac{\partial^2\psi_1'}{\partial x^2}\right)} = -\frac{\partial}{\partial x}\overline{\left[\psi_1'\frac{\partial}{\partial x}\left(\frac{\partial\psi_1'}{\partial t}\right)\right]} + \overline{\frac{\partial\psi_1'}{\partial x}\frac{\partial}{\partial t}\left(\frac{\partial\psi_1'}{\partial x}\right)}$$

The first term on the right-hand side vanishes because it is the integral of a perfect differential in x over a complete cycle. The second term on the right can be rewritten in the form

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\frac{\partial\psi_1'}{\partial x}\right)^2}$$

which is just the rate of change of the perturbation kinetic energy per unit mass averaged over a wavelength. Similarly, $-\psi_1'$ times the advection term on the left in (8.9) can be written after integration in x as

$$\begin{aligned}-U_1\overline{\psi_1'\frac{\partial^2}{\partial x^2}\left(\frac{\partial\psi_1'}{\partial x}\right)} &= -U_1\frac{\partial}{\partial x}\overline{\left[\psi_1'\frac{\partial}{\partial x}\left(\frac{\partial\psi_1'}{\partial x}\right)\right]} + U_1\overline{\frac{\partial\psi_1'}{\partial x}\frac{\partial^2\psi_1'}{\partial x^2}}\\ &= \frac{U_1}{2}\frac{\partial}{\partial x}\overline{\left(\frac{\partial\psi_1'}{\partial x}\right)^2} = 0\end{aligned}$$

Thus, the advection of kinetic energy vanishes when integrated over a wavelength. Evaluating the various terms in (8.10) and (8.11) in the same manner after multiplying through by $-\psi_3'$ and $(\psi_1' - \psi_3')$, respectively, we obtain the following set of perturbation energy equations:

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\frac{\partial\psi_1'}{\partial x}\right)^2} = -\frac{f_0}{\delta p}\overline{\omega_2'\psi_1'} \tag{8.34}$$

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\frac{\partial\psi_3'}{\partial x}\right)^2} = +\frac{f_0}{\delta p}\overline{\omega_2'\psi_3'} \tag{8.35}$$

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\psi_1' - \psi_3'\right)^2} = U_T\overline{\left(\psi_1' - \psi_3'\right)\frac{\partial}{\partial x}\left(\psi_1' + \psi_3'\right)} + \frac{\sigma\,\delta p}{f_0}\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} \tag{8.36}$$

where as before $U_T \equiv (U_1 - U_3)/2$.

[2] A zonal average generally designates the average around an entire circle of latitude. However, for a disturbance consisting of a single sinusoidal wave of wave number $k = m/(a\cos\phi)$, where m is an integer, the average over a wavelength is identical to a zonal average.

Defining the perturbation kinetic energy to be the sum of the kinetic energies of the 250- and 750-hPa levels:

$$K' \equiv (1/2)\left[\overline{\left(\partial\psi_1'/\partial x\right)^2} + \overline{\left(\partial\psi_3'/\partial x\right)^2}\right]$$

we find by adding (8.34) and (8.35) that

$$dK'/dt = -\left(f_0/\delta p\right)\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} = -\left(2f_0/\delta p\right)\overline{\omega_2'\psi_T} \tag{8.37}$$

Thus, the rate of change of perturbation kinetic energy is proportional to the correlation between perturbation thickness and vertical motion.

If we now define the perturbation available potential energy as

$$P' \equiv \lambda^2\overline{\left(\psi_1' - \psi_3'\right)^2}/2$$

we obtain from (8.36)

$$\begin{aligned} dP'/dt &= \lambda^2 U_T\overline{\left(\psi_1' - \psi_3'\right)\partial\left(\psi_1' + \psi_3'\right)/\partial x} + \left(f_0/\delta p\right)\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} \\ &= 4\lambda^2 U_T\overline{\psi_T\,\partial\psi_m/\partial x} + \left(2f_0/\delta p\right)\overline{\omega_2'\psi_T} \end{aligned} \tag{8.38}$$

The last term in (8.38) is just equal and opposite to the kinetic energy source term in (8.37). This term clearly must represent a conversion between potential and kinetic energy. If on average the vertical motion is positive ($\omega_2' < 0$) where the thickness is greater than average ($\psi_1' - \psi_3' > 0$) and vertical motion is negative where thickness is less than average, then

$$\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} = 2\overline{\omega_2'\psi_T} < 0$$

and the perturbation potential energy is being converted to kinetic energy. Physically, this correlation represents an overturning in which cold air aloft is replaced by warm air from below, a situation that clearly tends to lower the center of mass and hence the potential energy of the perturbation. However, the available potential energy and kinetic energy of a disturbance can still grow simultaneously, provided that the potential energy generation due to the first term in (8.38) exceeds the rate of potential energy conversion to kinetic energy.

The potential energy generation term in (8.38) depends on the correlation between the perturbation thickness ψ_T and the meridional velocity at 500 hPa, $\partial\psi_m/\partial$x. In order to understand the role of this term, it is helpful to consider a particular sinusoidal wave disturbance. Suppose that the barotropic and baroclinic parts of the disturbance can be written, respectively, as

$$\psi_m = A_m\cos k\left(x - ct\right) \quad \text{and} \quad \psi_T = A_T\cos k\left(x + x_0 - ct\right) \tag{8.39}$$

where x_0 designates the phase difference. Because ψ_m is proportional to the 500-hPa geopotential and ψ_T is proportional to the 500-hPa temperature (or 250-750-hPa thickness), the phase angle kx_0 gives the phase difference between geopotential and temperature fields at 500 hPa. Furthermore, A_m and A_T are measures of the amplitudes of the 500-hPa disturbance geopotential and thickness fields, respectively. Using the expressions in (8.39) we obtain

$$\begin{aligned}\overline{\psi_T \frac{\partial \psi_m}{\partial x}} &= -\frac{k}{L}\int_0^L A_T A_m \cos k\,(x + x_0 - ct)\sin k\,(x - ct)\,dx \\ &= \frac{k A_T A_m \sin kx_0}{L}\int_0^L [\sin k\,(x - ct)]^2\,dx \\ &= (A_T A_m k \sin kx_0)\,/2\end{aligned} \tag{8.40}$$

Substituting from (8.40) into (8.38) we see that for the usual midlatitude case of a westerly thermal wind ($U_T > 0$) the correlation in (8.40) must be positive if the perturbation potential energy is to increase. Thus, kx_0 must satisfy the inequality $0 < kx_0 < \pi$. Furthermore, the correlation will be a positive maximum for $kx_0 = \pi/2$, that is, when the temperature wave lags the geopotential wave by 90° in phase at 500 hPa.

This case is shown schematically in Fig. 8.4. Clearly, when the temperature wave lags the geopotential by one-quarter cycle, the northward advection of warm air by the geostrophic wind east of the 500-hPa trough and the southward advection of cold air west of the 500-hPa trough are both maximized. As a result, cold advection is strong below the 250-hPa trough, and warm advection is strong below the 250-hPa ridge. In that case, as discussed previously in Section 6.3.1, the upper-level disturbance will intensify. It should also be noted here that if the temperature wave lags the geopotential wave, the trough and ridge axes will tilt westward with height, which, as mentioned in Section 6.1, is observed to be the case for amplifying midlatitude synoptic systems.

Referring again to Fig. 8.4 and recalling the vertical motion pattern implied by the omega equation (8.28), we see that the signs of the two terms on the right in (8.38) cannot be the same. In the westward tilting perturbation of Fig. 8.4, the vertical motion must be downward in the cold air behind the trough at 500 hPa. Hence, the correlation between temperature and vertical velocity must be positive in this situation; that is,

$$\overline{\omega_2' \psi_T} < 0$$

Thus, for quasi-geostrophic perturbations, a westward tilt of the perturbation with height implies both that the horizontal temperature advection will increase the available potential energy of the perturbation and that the vertical circulation will convert perturbation available potential energy to perturbation kinetic energy. Conversely, an eastward tilt of the system with height would change the signs of both terms on the right in (8.38).

Although the signs of the potential energy generation term and the potential energy conversion term in (8.38) are always opposite for a developing baroclinic wave, it is only the potential energy generation rate that determines the growth of the total energy $P' + K'$ of the disturbance. This may be proved by adding (8.37) and (8.38) to obtain

$$d\left(P' + K'\right)/dt = 4\lambda^2 U_T \overline{\psi_T \partial \psi_m / \partial x}$$

Provided the correlation between the meridional velocity and temperature is positive and $U_T > 0$, the total energy of the perturbation will increase. Note that the vertical circulation merely converts disturbance energy between the available potential and kinetic forms without affecting the total energy of the perturbation.

The rate of increase of the total energy of the perturbation depends on the magnitude of the basic state thermal wind U_T. This is, of course, proportional to the zonally averaged meridional temperature gradient. Because the generation of perturbation energy requires systematic poleward transport of warm air and equatorward transport of cold air, it is clear that baroclinically unstable disturbances tend to reduce the meridional temperature gradient and hence the available potential energy of the mean flow. This latter process cannot be described mathematically in terms of the linearized equations. However, from Fig. 8.8 we can see qualitatively that parcels that move poleward and upward with slopes less than the slope of the zonal mean potential temperature surface will become warmer than their surroundings, and vice versa for parcels moving downward and equatorward. For

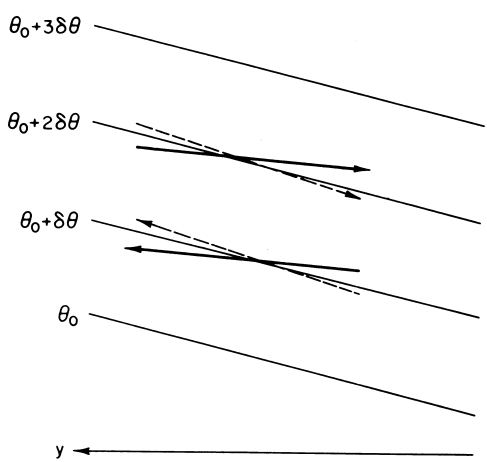


Fig. 8.8 Slopes of parcel trajectories relative to the zonal mean potential temperature surfaces for a baroclinically unstable disturbance (solid arrows) and for a baroclinically stable disturbance (dashed arrows).

such parcels the correlations between disturbance meridional velocity and temperature and between disturbance vertical velocity and temperature will both be positive as required for baroclinically unstable disturbances. For parcels that have trajectory slopes greater than the mean potential temperature slope, however, both of these correlations will be negative. Such parcels must then convert disturbance kinetic energy to disturbance available potential energy, which is in turn converted to zonal mean available potential energy. Therefore, in order that perturbations are able to extract potential energy from the mean flow, the perturbation parcel trajectories in the meridional plane must have slopes less than the slopes of the potential temperature surfaces, and a permanent rearrangement of air must take place for there to be a net heat transfer.

Since we have previously seen that poleward-moving air must rise and equatorward-moving air must sink, it is clear that the rate of energy generation can be greater for an atmosphere in which the meridional slope of the potential temperature surfaces is large. We can also see more clearly why baroclinic instability has a short-wave cutoff. As mentioned previously, the intensity of the vertical circulation must increase as the wavelength decreases. Thus, the slopes of the parcel trajectories must increase with decreasing wavelength, and for some critical wavelength the trajectory slopes will become greater than the slopes of the potential temperature surfaces. Unlike convective instability, where the most rapid amplification occurs for the smallest possible scales, baroclinic instability is most effective at an intermediate range of scales.

The energy flow for quasi-geostrophic perturbations is summarized in Fig. 8.9 by means of a block diagram. In this type of energy diagram each block represents a reservoir of a particular type of energy, and arrows designate the direction of energy flow. The complete energy cycle cannot be derived in terms of linear perturbation theory but will be discussed qualitatively in Chapter 10.

8.4 BAROCLINIC INSTABILITY OF A CONTINUOUSLY STRATIFIED ATMOSPHERE

In the two previous sections some basic aspects of baroclinic instability were elucidated in the context of a simple two-layer model. The dependence of growth rate on vertical shear and the existence of a short-wave cutoff were clearly demonstrated. The two-layer model, however, does have one severe constraint: it assumes that the altitude dependence of large-scale systems can be adequately represented with only two degrees of freedom in the vertical (i.e., the streamfunctions at the 250- and 750-hPa levels). Although most synoptic-scale systems in midlatitudes are observed to have vertical scales comparable to the depth of the troposphere, observed vertical structures do differ. Disturbances that are concentrated near the ground or near the tropopause can hardly be represented accurately in the two-layer model.

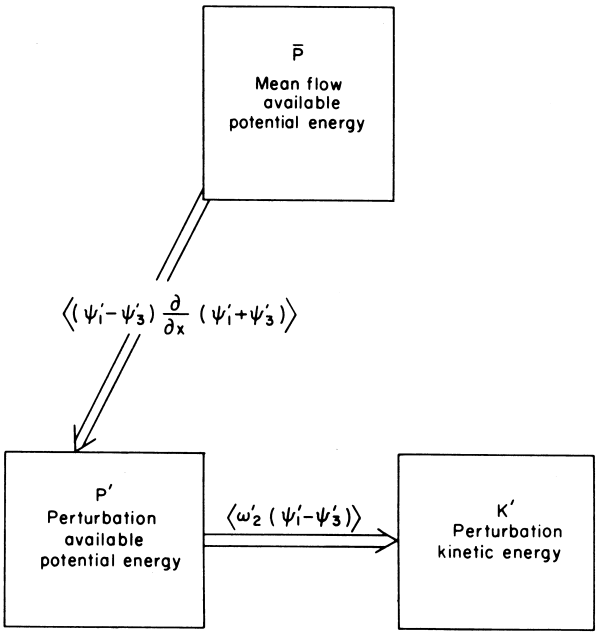


Fig. 8.9 Energy flow in an amplifying baroclinic wave.

An analysis of the structure of baroclinic modes for realistic mean zonal wind profiles is quite complex, and indeed can only be done by numerical methods. However, without obtaining specific normal mode solutions, it is possible to obtain necessary conditions for baroclinic or barotropic instability from an integral theorem first developed by Rayleigh. This theorem, which is discussed in Section 8.4.2, also shows how baroclinic instability is intimately related to the mean meridional gradient of potential vorticity and the mean meridional temperature gradient at the surface.

If a number of simplifying assumptions are made, it is possible to pose the stability problem for a continuously stratified atmosphere in a fashion that leads to a second-order differential equation for the vertical structure that can be solved by standard methods. This problem was originally studied by the British meteorologist Eady (1949) and, although mathematically similar to the two-layer model, provides additional insights. It is developed in Section 8.4.3.

8.4.1 Log-Pressure Coordinates

Derivation of the Rayleigh theorem and the Eady stability model is facilitated if we transform from the standard isobaric coordinates to a vertical coordinate based on

the logarithm of pressure. In the *log-pressure coordinates*, the vertical coordinate is defined as

$$z^* \equiv -H \ln (p/p_s) \tag{8.41}$$

where p_s is a standard reference pressure (usually taken to be 1000 hPa) and H is a standard scale height, $H \equiv RT_s/g$, with T_s a global average temperature. For the special case of an isothermal atmosphere at temperature T_s, z^* is exactly equal to geometric height, and the density profile is given by the reference density

$$\rho_0 \left(z^*\right) = \rho_s \exp \left(-z^*/H\right)$$

where ρ_s is the density at $z^* = 0$.

For an atmosphere with a realistic temperature profile, z^* is only approximately equivalent to the height, but in the troposphere the difference is usually quite small.

The vertical velocity in this coordinate system is

$$w^* \equiv Dz^*/dt$$

The horizontal momentum equation in the log-pressure system is the same as that in the isobaric system:

$$D\mathbf{V}/Dt + f\mathbf{k} \times \mathbf{V} = -\boldsymbol{\nabla}\Phi \tag{8.42}$$

However, the operator D/Dt is now defined as

$$D/Dt = \partial/\partial t + \mathbf{V} \bullet \boldsymbol{\nabla} + w^* \partial/\partial z^*$$

The hydrostatic equation $\partial\Phi/\partial p = -\alpha$ can be transformed to the log-pressure system by multiplying through by p and using the ideal gas law to get

$$\partial\Phi/\partial \ln p = -RT$$

which upon dividing through by $-H$ and using (8.41) gives

$$\partial\Phi/\partial z^* = RT/\mathrm{H} \tag{8.43}$$

The log-pressure form of the continuity equation can be obtained by transforming from the isobaric coordinate form (3.5). We first note that

$$w^* \equiv -(H/p)\, Dp/Dt = -H\omega/p$$

so that

$$\frac{\partial\omega}{\partial p} = -\frac{\partial}{\partial p}\left(\frac{pw^*}{H}\right) = \frac{\partial w^*}{\partial z^*} - \frac{w^*}{H} = \frac{1}{\rho_0}\frac{\partial\left(\rho_0 w^*\right)}{\partial z^*}$$

Thus, in log-pressure coordinates the continuity equation becomes simply

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{1}{\rho_0}\frac{\partial\left(\rho_0 w^*\right)}{\partial z^*} = 0 \tag{8.44}$$

It is left as a problem for the reader to show that the first law of thermodynamics (3.6) can be expressed in log-pressure form as

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \boldsymbol{\nabla}\right)\frac{\partial \Phi}{\partial z^*} + w^* N^2 = \frac{\kappa J}{H} \tag{8.45}$$

where

$$N^2 \equiv (R/H)\left(\partial T/\partial z^* + \kappa T/H\right)$$

is the buoyancy frequency squared (see Section 2.7.3) and $\kappa \equiv R/c_p$. Unlike the static stability parameter, S_p, in the isobaric form of the thermodynamic equation (3.6), the parameter N^2 varies only weakly with height in the troposphere; it can be assumed to be constant without serious error. This is a major advantage of the log-pressure formulation.

The quasi-geostrophic potential vorticity equation (6.24) has the same form as in the isobaric system, but with q defined as

$$q \equiv \nabla^2\psi + f + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\psi}{\partial z^*}\right) \tag{8.46}$$

where $\varepsilon \equiv f_0^2/N^2$.

8.4.2 Baroclinic Instability: The Rayleigh Theorem

We now examine the stability problem for a continuously stratified atmosphere on the midlatitude β plane. The linearized form of the quasi-geostrophic potential vorticity equation (6.24) can be expressed in log-pressure coordinates as

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)q' + \frac{\partial\overline{q}}{\partial y}\frac{\partial\psi'}{\partial x} = 0 \tag{8.47}$$

where

$$q' \equiv \nabla^2\psi' + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\psi'}{\partial z^*}\right) \tag{8.48}$$

and

$$\frac{\partial\overline{q}}{\partial y} = \beta - \frac{\partial^2\overline{u}}{\partial y^2} - \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\overline{u}}{\partial z^*}\right) \tag{8.49}$$

As in the two-layer model, boundary conditions are required at lower and upper boundary pressure surfaces. Assuming that the flow is adiabatic and that the vertical motion w^* vanishes at these boundaries, the linearized form of the thermodynamic energy equation (8.45) valid at horizontal boundary surfaces is simply

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)\frac{\partial\psi'}{\partial z^*}-\frac{\partial\psi'}{\partial x}\frac{\partial\overline{u}}{\partial z^*}=0 \tag{8.50}$$

The sidewall boundary conditions are

$$\partial\psi'/\partial x=0, \quad \text{hence, } \psi'=0 \quad at \quad y=\pm L \tag{8.51}$$

We now assume that the perturbation consists of a single zonal Fourier component propagating in the x direction:

$$\psi'(x,y,z,t)=\mathrm{Re}\left\{\Psi\big(y,z\big)\exp\big[ik\big(x-ct\big)\big]\right\} \tag{8.52}$$

where $\Psi(y,z)=\Psi_r+i\Psi_i$ is a complex amplitude, k is the zonal wave number, and $c=c_r+ic_i$ is a complex phase speed. Note that (8.52) can alternatively be expressed as

$$\psi'\big(x,y,z,t\big)=e^{kc_it}\big[\Psi_r\cos k\big(x-c_rt\big)-\Psi_i\sin k\big(x-c_rt\big)\big]$$

Thus, the relative magnitudes of Ψ_r and Ψ_i determine the phase of the wave for any y, z^*.

Substituting from (8.52) into (8.47) and (8.50) yields

$$(\overline{u}-c)\left[\frac{\partial^2\Psi}{\partial y^2}-k^2\Psi+\frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\Psi}{\partial z^*}\right)\right]+\frac{\partial\overline{q}}{\partial y}\Psi=0 \tag{8.53}$$

and

$$(\overline{u}-c)\frac{\partial\Psi}{\partial z^*}-\frac{\partial\overline{u}}{\partial z^*}\Psi=0 \quad \text{at } z^*=0 \tag{8.54}$$

If the upper boundary is taken to be a rigid lid at a finite height, as is sometimes done in theoretical studies, the condition (8.54) is appropriate at that boundary as well. Alternatively the upper boundary condition can be specified by requiring Ψ to remain finite as $z^*\to\infty$.

Equation (8.53), together with its boundary conditions, constitutes a linear boundary value problem for $\Psi(y,z^*)$. It is generally not simple to obtain solutions to (8.53) for realistic mean zonal wind profiles. Nevertheless, we can obtain some useful information on stability properties simply by analyzing the energetics of the system.

Dividing (8.53) by $(\overline{u} - c)$ and separating the resulting equation into real and imaginary parts we obtain

$$\frac{\partial^2 \Psi_r}{\partial y^2} + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0 \frac{\partial \Psi_r}{\partial z^*}\right) - \left[k^2 - \delta_r\left(\partial\overline{q}/\partial y\right)\right]\Psi_r - \delta_i \partial\overline{q}/\partial y \Psi_i = 0 \quad (8.55)$$

$$\frac{\partial^2 \Psi_i}{\partial y^2} + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0 \frac{\partial \Psi_i}{\partial z^*}\right) - \left[k^2 - \delta_r\left(\partial\overline{q}/\partial y\right)\right]\Psi_i + \delta_i \partial\overline{q}/\partial y \Psi_r = 0 \quad (8.56)$$

where

$$\delta_r = \frac{\overline{u} - c_r}{(\overline{u} - c_r)^2 + c_i^2}, \quad \text{and} \quad \delta_i = \frac{c_i}{(\overline{u} - c_r)^2 + c_i^2}$$

Similarly, dividing (8.54) through by $(\bar{u} - c)$ and separating into real and imaginary parts gives for the boundary condition at $z^* = 0$:

$$\frac{\partial \Psi_r}{\partial z^*} + \frac{\partial \overline{u}}{\partial z^*}\left(\delta_i \Psi_i - \delta_r \Psi_r\right) = 0; \quad \frac{\partial \Psi_i}{\partial z^*} - \frac{\partial \overline{u}}{\partial z^*}\left(\delta_r \Psi_i + \delta_i \Psi_r\right) = 0 \quad (8.57)$$

Multiplying (8.55) by Ψ_i, (8.56) by Ψ_r, and subtracting the latter from the former yields

$$\rho_0\left[\Psi_i \frac{\partial^2 \Psi_r}{\partial y^2} - \Psi_r \frac{\partial^2 \Psi_i}{\partial y^2}\right] + \left[\Psi_i \frac{\partial}{\partial z^*}\left(\varepsilon\rho_0 \frac{\partial \Psi_r}{\partial z^*}\right) - \Psi_r \frac{\partial}{\partial z^*}\left(\varepsilon\rho_0 \frac{\partial \Psi_i}{\partial z^*}\right)\right]$$
$$- \rho_0 \delta_i \left(\partial\overline{q}/\partial y\right)\left(\Psi_i^2 + \Psi_r^2\right) = 0 \quad (8.58)$$

Using the chain rule of differentiation, (8.58) can be expressed in the form

$$\rho_0 \frac{\partial}{\partial y}\left[\Psi_i \frac{\partial \Psi_r}{\partial y} - \Psi_r \frac{\partial \Psi_i}{\partial y}\right] + \frac{\partial}{\partial z^*}\left[\varepsilon\rho_0\left(\Psi_i \frac{\partial \Psi_r}{\partial z^*} - \Psi_r \frac{\partial \Psi_i}{\partial z^*}\right)\right]$$
$$- \rho_0 \delta_i \left(\partial\overline{q}/\partial y\right)\left(\Psi_i^2 + \Psi_r^2\right) = 0 \quad (8.59)$$

The first term in brackets in (8.59) is a perfect differential in y; the second term is a perfect differential in z^*. Thus, if (8.59) is integrated over the y, z^* domain the result can be expressed as

$$\int_0^\infty \left[\Psi_i \frac{\partial \Psi_r}{\partial y} - \Psi_r \frac{\partial \Psi_i}{\partial y}\right]_{-L}^{+L} \rho_0 dz^* + \int_{-L}^{+L}\left[\varepsilon\rho_0\left(\Psi_i \frac{\partial \Psi_r}{\partial z^*} - \Psi_r \frac{\partial \Psi_i}{\partial z^*}\right)\right]_0^\infty dy$$
$$= \int_{-L}^{+L}\int_0^\infty \rho_0 \delta_i \frac{\partial \overline{q}}{\partial y}\left(\Psi_i^2 + \Psi_r^2\right) dy dz^* \quad (8.60)$$

However, from (8.51) $\Psi_i = \Psi_r = 0$ at $y = \pm L$ so that the first integral in (8.60) vanishes. Furthermore, if Ψ remains finite as $z^* \to \infty$ the contribution to the second integral of (8.60) at the upper boundary vanishes. If we then use (8.57) to eliminate the vertical derivatives in this term at the lower boundary, (8.60) can be expressed as

$$c_i \left[\int_{-L}^{+L} \int_{0}^{\infty} \frac{\partial \overline{q}}{\partial y} \frac{\rho_0 |\Psi|^2}{|\overline{u} - c|^2} dy dz^* - \int_{-L}^{+L} \varepsilon \frac{\partial \overline{u}}{\partial z^*} \frac{\rho_0 |\Psi|^2}{|\overline{u} - c|^2} \bigg|_{z^*=0} dy \right] = 0 \qquad (8.61)$$

where $|\Psi|^2 = \Psi_r^2 + \Psi_i^2$ is the disturbance amplitude squared.

Equation (8.61) has important implications for the stability of quasi-geostrophic perturbations. For unstable modes, c_i must be nonzero, and thus the quantity in square brackets in (8.61) must vanish. Because $|\Psi|^2/|\overline{u} - c|^2$ is nonnegative, instability is possible only when $\partial \overline{u}/\partial z^*$ at the lower boundary and $(\partial \overline{q}/\partial y)$ in the whole domain satisfy certain constraints:

(a) If $\partial \overline{u}/\partial z^*$ at $z^* = 0$ (which by thermal wind balance implies that the meridional temperature gradient vanishes at the boundary), the second integral in (8.61) vanishes. Thus, the first integral must also vanish for instability to occur. This can occur only if $\partial \overline{q}/\partial y$ changes sign within the domain (i.e., $\partial \overline{q}/\partial y = 0$ somewhere). This is referred to as the *Rayleigh necessary condition* and is another demonstration of the fundamental role played by potential vorticity. Because $\partial \overline{q}/\partial y$ is normally positive, it is clear that in the absence of temperature gradients at the lower boundary, a region of negative meridional potential vorticity gradients must exist in the interior for instability to be possible.

(b) If $\partial \overline{q}/\partial y \geq 0$ everywhere, then it is necessary that $\partial \overline{u}/\partial z^* > 0$ somewhere at the lower boundary for $c_i > 0$.

(c) If $\partial \overline{u}/\partial z^* < 0$ everywhere at $z^* = 0$, then it is necessary that $\partial \overline{q}/\partial y < 0$ somewhere for instability to occur. Thus, there is an asymmetry between westerly and easterly shear at the lower boundary, with the former more favorable for baroclinic instability.

The basic state potential vorticity gradient in (8.49) can be written in the form

$$\frac{\partial \overline{q}}{\partial y} = \beta - \frac{\partial^2 \overline{u}}{\partial y^2} + \frac{\varepsilon}{H} \frac{\partial \overline{u}}{\partial z^*} - \varepsilon \frac{\partial^2 \overline{u}}{\partial z^{*2}} - \frac{\partial \varepsilon}{\partial z^*} \frac{\partial \overline{u}}{\partial z^*}$$

Because β is positive everywhere, if ε is constant a negative basic state potential vorticity gradient can occur only for strong positive mean flow curvature (i.e., $\partial^2 \overline{u}/\partial y^2$ or $\partial^2 \overline{u}/\partial z^{*2} >> 0$) or strong negative vertical shear ($\partial \overline{u}/\partial z^* << 0$).

Strong positive meridional curvature can occur at the core of an easterly jet or on the flanks of a westerly jet. Instability associated with such horizontal curvature is referred to as *barotropic instability*. The normal baroclinic instability in midlatitudes is associated with mean flows in which $\partial\overline{q}/\partial y > 0$ and $\partial\overline{u}/\partial z^* > 0$ at the ground. Hence, a mean meridional temperature gradient at the ground is essential for the existence of such instability. Baroclinic instability can also be excited at the tropopause due to the rapid decrease of ε with height if there is a sufficiently strong easterly mean wind shear to cause a local reversal in the mean potential vorticity gradient.

8.4.3 The Eady Stability Problem

This section analyzes the structures (eigenfunctions) and growth rates (eigenvalues) for unstable modes in the simplest possible model that satisfies the necessary conditions for instability in a continuous atmosphere given in the previous subsection. For simplicity we make the following assumptions:

(i) Basic state density constant (Boussinesq approximation).
(ii) f-plane geometry ($\beta = 0$).
(iii) $\partial\overline{u}/\partial z^* = \Lambda =$ constant.
(iv) Rigid lids at $z^* = 0$ and H.

These conditions are only a crude model of the atmosphere, but provide a first approximation for study of the dependence of vertical structure on horizontal scale and stability. Despite the zero mean potential vorticity in the domain, the Eady model satisfies the necessary conditions for instability discussed in the previous subsection because vertical shear of the basic state mean flow at the upper boundary provides an additional term in (8.61) that is equal and opposite to the lower boundary integral.

Using the aforementioned approximations, the quasi-geostrophic potential vorticity equation and thermodynamic energy equation are

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\left(\nabla^2\psi' + \varepsilon\frac{\partial^2\psi'}{\partial z^{*2}}\right) = 0 \tag{8.62}$$

and

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\frac{\partial\psi'}{\partial z^*} - \frac{\partial\psi'}{\partial x}\frac{\partial\overline{u}}{\partial z^*} + w^*\frac{N^2}{f_0} = 0 \tag{8.63}$$

respectively, where again $\varepsilon \equiv f_0^2/N^2$.

Letting

$$\begin{aligned} \psi'\left(x, y, z^*, t\right) &= \Psi\left(z^*\right)\cos ly \exp\left[ik\left(x - ct\right)\right] \\ \overline{u}\left(z^*\right) &= \Lambda z^* \end{aligned} \tag{8.64}$$

where as in the previous subsection $\Psi(z^*)$ is a complex amplitude and c a complex phase speed, and substituting from (8.64) into (8.62) we find that the vertical structure is given by the solution of the standard second-order differential equation

$$d^2\Psi \Big/ dz^{*2} - \alpha^2\Psi = 0 \tag{8.65}$$

where $\alpha^2 = \left(k^2 + l^2\right)/\varepsilon$. A similar substitution into (8.63) yields the boundary conditions

$$\left(\Lambda z^* - c\right) d\Psi/dz^* - \Psi\Lambda = 0, \quad \text{at } z^* = 0, H \tag{8.66}$$

valid for rigid horizontal boundaries ($w^* = 0$) at the surface ($z^* = 0$) and the tropopause ($z^* = H$).

The general solution of (8.65) can be written in the form

$$\Psi\left(z^*\right) = A \sinh \alpha z^* + B \cosh \alpha z^* \tag{8.67}$$

Substituting from (8.67) into the boundary conditions (8.66) for $z^* = 0$ and H yields a set of two linear homogeneous equations in the amplitude coefficients A and B:

$$-c\alpha A - B\Lambda = 0$$
$$\alpha\left(\Lambda H - c\right)\left(A \cosh \alpha H + B \sinh \alpha H\right) - \Lambda\left(A \sinh \alpha H + B \cosh \alpha H\right) = 0$$

As in the two-layer model a nontrivial solution exists only if the determinant of the coefficients of A and B vanishes. Again, this leads to a quadratic equation in the phase speed c. The solution (see Problem 8.12) has the form

$$c = \frac{\Lambda H}{2} \pm \frac{\Lambda H}{2}\left[1 - \frac{4 \cosh \alpha H}{\alpha H \sinh \alpha H} + \frac{4}{\alpha^2 H^2}\right]^{1/2} \tag{8.68}$$

Thus

$$c_i \neq 0 \quad \text{if} \quad 1 - \frac{4 \cosh \alpha H}{\alpha H \sinh \alpha H} + \frac{4}{\alpha^2 H^2} < 0$$

and the flow is then baroclinically unstable. When the quantity in square brackets in (8.68) is equal to zero, the flow is said to be *neutrally stable*. This condition occurs for $\alpha = \alpha_c$ where

$$\alpha_c^2 H^2/4 - \alpha_c H \left(\tanh \alpha_c H\right)^{-1} + 1 = 0 \tag{8.69}$$

Using the identity

$$\tanh \alpha_c H = 2 \tanh\left(\frac{\alpha_c H}{2}\right) \Big/ \left[1 + \tanh^2\left(\frac{\alpha_c H}{2}\right)\right]$$

we can factor (8.69) to yield

$$\left[\frac{\alpha_c H}{2} - \tanh\left(\frac{\alpha_c H}{2}\right)\right]\left[\frac{\alpha_c H}{2} - \coth\left(\frac{\alpha_c H}{2}\right)\right] = 0 \tag{8.70}$$

Thus, the critical value of α is given by $\alpha_c H/2 = / \coth(\alpha_c H/2)$, which implies $\alpha_c H \cong 2.4$. Hence instability requires $\alpha < \alpha_c$ or

$$\left(k^2 + l^2\right) < \left(\alpha_c^2 f_0^2 / N^2\right) \approx 5.76/L_R^2$$

where $L_R \equiv NH/f_0 \approx 1000$ km is the Rossby *radius of deformation* for a continuously stratified fluid [compare λ^{-1} defined just below (8.16)]. For waves with

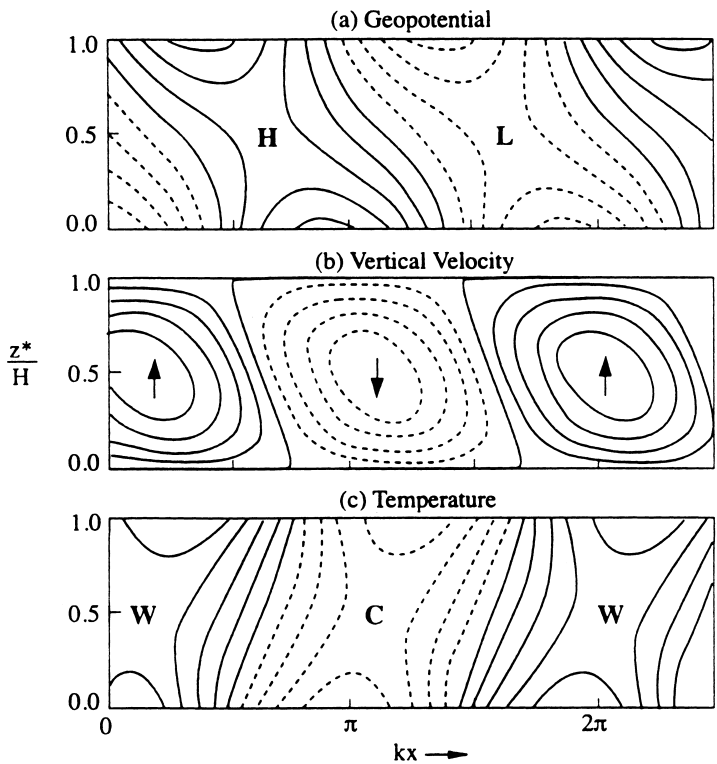


Fig. 8.10 Properties of the most unstable Eady wave. (a) Contours of perturbation geopotential height; H and L designate ridge and trough axes, respectively. (b) Contours of vertical velocity; up and down arrows designate axes of maximum upward and downward motion, respectively. (c) Contours of perturbation temperature; W and C designate axes of warmest and coldest temperatures, respectively. In all panels 1 and 1/4 wavelengths are shown for clarity.

equal zonal and meridional wave numbers ($k = l$), the wavelength of maximum growth rate turns out to be

$$L_m = 2\sqrt{2}\pi L_R/(H\alpha_m) \cong 5500 \text{ km}$$

where α_m is the value of α for which kc_i is a maximum.

Substituting this value of α into the solution for the vertical structure of the streamfunction (8.67) and using the lower boundary condition to express the coefficient B in terms of A, we can determine the vertical structure of the most unstable mode. As shown in Fig. 8.10, trough and ridge axes slope westward with height, in agreement with the requirements for extraction of available potential energy from the mean flow. The axes of the warmest and coldest air, however, tilt eastward with height, a result that could not be determined from the two-layer model where temperature was given at a single level. Furthermore, Figs. 8.10a and 8.10b show that east of the upper-level trough axis, where the perturbation meridional velocity is positive, the vertical velocity is also positive. Thus, parcel motion is poleward and upward in the region where $\theta' > 0$. Conversely, west of the upper-level trough axis parcel motion is equatorward and downward where $\theta' < 0$. Both cases are thus consistent with the energy converting parcel trajectory slopes shown in Fig. 8.8.

8.5 GROWTH AND PROPAGATION OF NEUTRAL MODES

As suggested earlier, baroclinic wave disturbances with certain favorable initial configurations may amplify rapidly even in the absence of baroclinic instability. The theory for the optimal initial perturbations that give rise to rapid transient growth is beyond the scope of this book. The basic idea of transient growth of neutral modes can be illustrated quite simply, however, in the two-layer model by considering disturbances that consist of waves with zonal wave numbers slightly higher than the short-wave cutoff for baroclinic instability. Similarly, neutral nodes of the two-layer model can also be used to demonstrate the important phenomenon of downstream development of disturbances due to the group velocity effect (see Fig. 7.4).

8.5.1 Transient Growth of Neutral Waves

If we neglect the β effect, let $U_m = 0$, and assume that $k^2 > 2\lambda^2$, the two-layer model of Section 8.2 has two oppositely propagating neutral solutions given by (8.24) with zonal phase speeds

$$c_1 = +U_T\mu; \quad c_2 = -U_T\mu. \tag{8.71}$$

where

$$\mu = \left[\frac{k^2 - 2\lambda^2}{k^2 + 2\lambda^2}\right]^{1/2}$$

Then from (8.17) a disturbance consisting of these two modes can be expressed as

$$\psi_m = A_1 \exp\left[ik\left(x - c_1 t\right)\right] + A_2 \exp\left[ik\left(x - c_2 t\right)\right] \tag{8.72a}$$

$$\psi_T = B_1 \exp\left[ik\left(x - c_1 t\right)\right] + B_2 \exp\left[ik\left(x - c_2 t\right)\right] \tag{8.72b}$$

but from (8.18)

$$c_1 A_1 - U_T B_1 = 0; \quad c_2 A_2 - U_T B_2 = 0$$

So that with the aid of (8.71)

$$B_1 = \mu A_1; \quad B_2 = -\mu A_2 \tag{8.73}$$

For an initial disturbance confined entirely to the upper level, it is easily verified that $\psi_m = \psi_T$ ($\psi_1 = 2\psi_m$ and $\psi_3 = 0$). Thus, initially $A_1 + A_2 = B_1 + B_2$, and substituting from (8.73) gives

$$A_2 = -A_1\left[(1-\mu)/(1+\mu)\right]$$

Hence, if A_1 is real, the streamfunctions of (8.72a,b) can be expressed as

$$\begin{aligned}\psi_m(x,t) &= A_1\left[\cos\left[k\left(x - \mu U_T t\right)\right] - \frac{(1-\mu)}{(1+\mu)}\cos\left[k\left(x + \mu U_T t\right)\right]\right] \\ &= \frac{2\mu A_1}{(1+\mu)}\left[\cos kx \cos\left(k\mu U_T t\right) + \frac{1}{\mu}\sin kx \sin\left(k\mu U_T t\right)\right]\end{aligned} \tag{8.74a}$$

$$\begin{aligned}\psi_T(x,t) &= \mu A_1\left[\cos\left[k\left(x - \mu U_T t\right)\right] + \frac{(1-\mu)}{(1+\mu)}\cos\left[k\left(x + \mu U_T t\right)\right]\right] \\ &= \frac{2\mu A_1}{\left(1+\mu\right)}\left[\cos kx \cos\left(k\mu U_T t\right) + \mu \sin kx \sin\left(k\mu U_T t\right)\right]\end{aligned} \tag{8.74b}$$

The first forms of the right side in (8.74a,b) show that for small μ the barotropic mode initially consists of two waves of nearly equal amplitude that are 180° out of phase so that they nearly cancel, whereas the baroclinic mode initially consists of two very weak waves that are in phase.

As time advances the two oppositely propagating barotropic modes begin to reinforce each other, leading to a maximum amplitude trough 90° to the east of the

initial trough at time $t_m = \pi/(2k\mu U_T)$. From (8.74) the baroclinic and barotropic modes at time t_m are then

$$\begin{aligned}(\psi_m)_{\max} &= 2A_1\,(1+\mu)^{-1}\sin kx \\ (\psi_T)_{\max} &= 2A_1\mu^2\,(1+\mu)^{-1}\sin kx\end{aligned} \tag{8.75a,b}$$

from which it is shown easily that

$$(\psi_1)_{\max} = 2A_1\left(1+\mu^2\right)(1+\mu)^{-1}\sin kx; \quad (\psi_3)_{\max} = 2A_1\,(1-\mu)\sin kx$$

so that for small μ the resulting disturbance is nearly barotropic. Thus, the initial disturbance not only amplifies, but spreads in the vertical. The time for growth to maximum amplitude is inversely proportional to the basic state thermal wind; the maximum amplitude, however, depends only on the initial amplitude and the parameter μ.

Figure 8.11 shows the initial and maximum amplitudes of the barotropic and baroclinic streamfunctions for $f_0 = 10^{-4}\,\mathrm{s}^{-1}$, $\sigma = 2\times10^{-6}\,\mathrm{m}^2\,\mathrm{Pa}^{-2}\,\mathrm{s}^{-2}$, $U_T = 35\,\mathrm{m\,s}^{-1}$, and a zonal wavelength of 3000 km. In this case the amplitude of the barotropic disturbance increases by a factor of 8 in about 48 h. Although the most unstable normal mode for these conditions amplifies by a similar amount in less than 1 day, if the initial upper level neutral disturbance has a velocity amplitude of a few meters per second, its growth on the time scale of a few days may dominate over a normal mode instability that grows exponentially from a much smaller initial perturbation.

As with normal mode baroclinic instability, the energy source for the transient amplification of the neutral modes is conversion of the available potential energy of the mean flow into disturbance potential energy by meridional temperature advection, followed immediately by conversion to disturbance kinetic energy by the secondary vertical circulation. This secondary circulation has a vertical velocity field that lags the upper level streamfunction field by 90° phase. Maximum upward motion and lower level convergence thus occur to the west of the initial upper level ridge, and maximum downward motion and lower level divergence occur west of the initial upper level trough. Vorticity tendencies associated with this convergence and divergence pattern partly balance the eastward vorticity advection at the upper level and also act to produce a lower level trough to the east of the initial upper level trough and lower level ridge to the east of the initial upper level ridge. The result is that a pattern of nearly barotropic troughs and ridges develops 90° to the east of the initial trough and ridge pattern. Unlike normal mode instability the growth in this case does not continue indefinitely. Rather, for a given value of μ the maximum growth occurs at the time for which $t_m = \pi/(2kU_T\mu)$. As the zonal wave number approaches the short-wave instability cutoff, the total amplification increases, but the amplification time also increases.

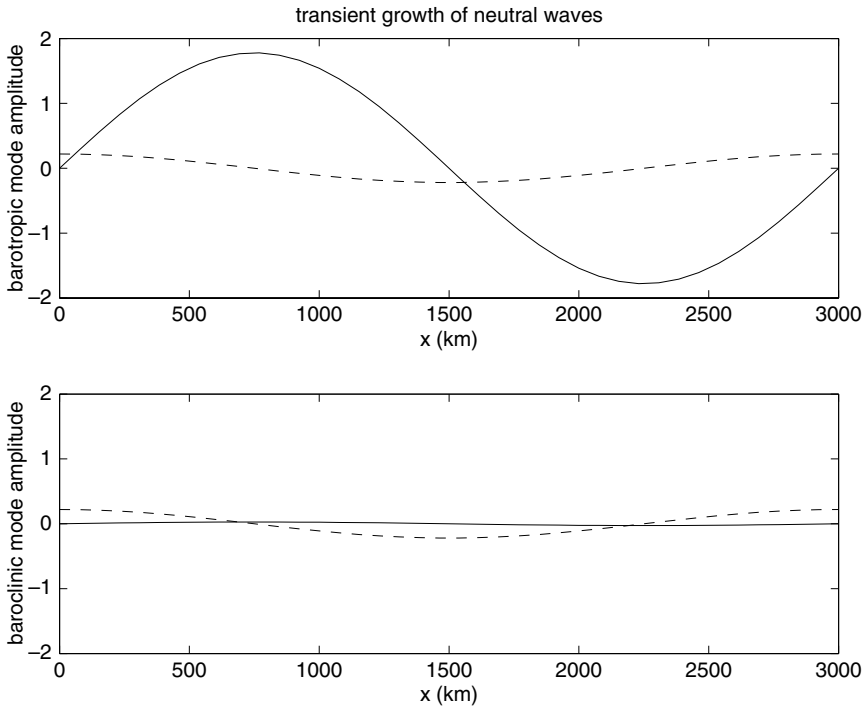


Fig. 8.11 Zonal distribution of the amplitude of a transient neutral disturbance initially confined entirely to the upper layer in the two-layer model. Dashed lines show the initial distribution of the baroclinic and barotropic modes; solid lines show the distribution at time of maximum amplification.

8.5.2 Downstream Development

It was noted in Section 7.2.2 that under some circumstances the zonal group velocity may exceed the zonal phase speed so that the energy of a wave group propagates faster than the individual wave disturbances. This dispersion effect was demonstrated for barotropic Rossby waves in Problem 7.19. The β effect is not necessary, however, for the development of new disturbances to occur downstream of the original wave disturbances on the synoptic scale. The two-layer channel model with constant Coriolis parameter can be used to demonstrate the process of dispersion and downstream development of neutral waves in a simple context.

Expressed in terms of wave frequency, the dispersion relation of (8.24) becomes

$$\nu = kU_m \pm k\mu U_T \tag{8.76}$$

where μ is defined below (8.71). The corresponding group velocity is then

$$c_{gx} = \frac{\partial \nu}{\partial k} = U_m \pm U_T \mu \left(1 + \frac{4k^2\lambda^2}{k^4 - 4\lambda^4}\right) \tag{8.77}$$

Comparison of (8.77) with (8.71) shows that relative to the mean zonal flow the group velocity exceeds the phase velocity by the factor in parentheses on the right side of (8.77) for both eastward and westward directed neutral modes. For example, if $k^2 = 2\lambda^2\left(1 + \sqrt{2}\right)$, the group velocity equals twice the phase speed, which is the situation shown schematically in Fig. 7.4.

PROBLEMS

8.1. Show using (8.25) that the maximum growth rate for baroclinic instability when $\beta = 0$ occurs for

$$k^2 = 2\lambda^2\left(\sqrt{2} - 1\right)$$

How long does it take the most rapidly growing wave to amplify by a factor of e^1 if $\lambda^2 = 2 \times 10^{-12}\ \mathrm{m}^{-1}$ and $U_T = 20\ \mathrm{m\ s}^{-1}$?

8.2. Solve for ψ_3' and ω_2' in terms of ψ_1' for a baroclinic Rossby wave whose phase speed satisfies (8.23). Explain the phase relationship among ψ_1', ψ_3', ω_2' and in terms of the quasi-geostrophic theory. (Note that $U_T = 0$ in this case.)

8.3. For the case $U_1 = -U_3$ and $k^2 = \lambda^2$ solve for ψ_3' and ω_2' in terms of ψ_1' for marginally stable waves [i.e., $\delta = 0$ in (8.21)].

8.4. For the case $\beta = 0$, $k^2 = \lambda^2$, and $U_m = U_T$ solve for ψ_3' and ω_2' in terms of ψ_1'. Explain the phase relationships among ω_2', ψ_1', and ψ_3' in terms of the energetics of quasi-geostrophic waves for the amplifying wave.

8.5. Suppose that a baroclinic fluid is confined between two rigid horizontal lids in a rotating tank in which $\beta = 0$ but friction is present in the form of linear drag proportional to the velocity (i.e., $\mathbf{Fr} = -\mu\mathbf{V}$). Show that the two-level model perturbation vorticity equations in Cartesian coordinates can be written as

$$\left(\frac{\partial}{\partial t} + U_1\frac{\partial}{\partial x} + \mu\right)\frac{\partial^2\psi_1'}{\partial x^2} - \frac{f}{\delta p}\omega_2' = 0$$

$$\left(\frac{\partial}{\partial t} + U_3\frac{\partial}{\partial x} + \mu\right)\frac{\partial^2\psi_3'}{\partial x^2} + \frac{f}{\delta p}\omega_2' = 0$$

where perturbations are assumed in the form given in (8.8). Assuming solutions of the form (8.17), show that the phase speed satisfies a relationship similar to (8.21) with β replaced everywhere by $i\mu k$, and that as a result the condition for baroclinic instability becomes

$$U_T > \mu \left(2\lambda^2 - k^2\right)^{-1/2}$$

8.6. For the case $\beta = 0$ determine the phase difference between the 250- and 750-hPa geopotential fields for the most unstable baroclinic wave (see Problem 8.1). Show that the 500-hPa geopotential and thickness fields (ψ_m, ψ_T) are 90° out of phase.

8.7. For the conditions of Problem 8.6, given that the amplitude of ψ_m is $A = 10^7\,\mathrm{m^2\,s^{-1}}$, solve the system (8.18) and (8.19) to obtain B. Let $\lambda^2 = 2.0 \times 10^{-12}\,\mathrm{m^{-2}}$ and $U_T = 15\,\mathrm{m\,s^{-1}}$. Use your results to obtain expressions for ψ_1' and ψ_3'.

8.8. For the situation of Problem 8.7, compute ω_2' using (8.28).

8.9. Compute the total potential energy per unit cross-sectional area for an atmosphere with an adiabatic lapse rate given that the temperature and pressure at the ground are $p = 10^5$ Pa and $T = 300$ K, respectively.

8.10. Consider two air masses at the uniform potential temperatures $\theta_1 = 320$ K and $\theta_2 = 340$ K that are separated by a vertical partition as shown in Fig. 8.7. Each air mass occupies a horizontal area of $10^4\,\mathrm{m^2}$ and extends from the surface ($p_0 = 10^5$ Pa) to the top of the atmosphere. What is the available potential energy for this system? What fraction of the total potential energy is available in this case?

8.11. For the unstable baroclinic wave that satisfies the conditions given in Problems 8.7 and 8.8, compute the energy conversion terms in (8.37) and (8.38) and hence obtain the instantaneous rates of change of the perturbation kinetic and available potential energies.

8.12. Starting with (8.62) and (8.63), derive the phase speed c for the Eady wave given in (8.68).

8.13. Unstable baroclinic waves play an important role in the global heat budget by transferring heat poleward. Show that for the Eady wave solution the poleward heat flux averaged over a wavelength,

$$\overline{v'T'} = \frac{1}{L}\int_0^L v'T'dx$$

is independent of height and is positive for a growing wave. How does the magnitude of the heat flux at a given instant change if the mean wind shear is doubled?

8.14. Assuming that the coefficient A in (8.67) is real, obtain an expression for the geostrophic streamfunction $\psi'(x, y, z^*, t)$ for the most unstable mode in the Eady stability problem for the case $k = l$. Use this result to derive an expression for the corresponding vertical velocity w^* in terms of A.

8.15. For the neutral baroclinic wave disturbance in the two-layer model given by (8.75a,b), derive the corresponding ω_2' field. Describe how the convergence and divergence fields associated with this secondary circulation influence the evolution of the disturbance.

8.16. For the situation of Problem 8.15, derive expressions for the conversion of zonal available potential energy to eddy available potential energy and the conversion of eddy available potential energy to eddy kinetic energy.

MATLAB EXERCISES

M8.1. By varying the input zonal wavelength in the MATLAB script **twolayer_model_1A.m**, find the shortest zonal wavelength for which the two-layer model is baroclinically unstable and the wavelength of maximum instability (i.e., most rapid growth rate) in the case where the basic state thermal wind is 15 ms^{-1}. The case given corresponds to the situation given in Section 8.2.1 in the textbook. The code of **twolayer_model_1B.m** differs only in that a finite meridional width is assumed with meridional wave number $m = \pi/(3000\text{ km})$. Compare the zonal wavelengths for short-wave cutoff and for maximum growth rate for these two cases.

M8.2. Use the MATLAB script **twolayer_model_2A.m** in order to examine the transient growth associated with a disturbance that is initially entirely confined to the 250-hPa level. Let the basic state thermal wind be 25 m s^{-1}. By examining zonal wavelengths shorter than the instability cutoff determined in Problem M8.1, find the zonal wavelength that gives the largest amplification of the disturbance. Repeat this exercise using the script **twolayer_model_2B.m**, which includes a meridional dependence with wave number $m = \pi/(3000\text{ km})$. [Note that the program will terminate if you choose a wavelength that is long enough to be baroclinically unstable.] Discuss the vertical structure of the transiently growing stable modes. What sort of situation in the real atmosphere does this solution crudely model?

M8.3. The MATLAB script **twolayer_model_3b.m** shows the propagation of a wave packet composed of the sum of nine eastward propagating neutral waves of the two-layer model, ranging in wave number from 0.6 to $1.6k$, where $k = 2\pi/L$, and $L = 1850\,\text{km}$. Run the script and estimate from the animation the characteristic phase velocity and the characteristic group velocity for the cases of $U_T = 15$ and $30\,\text{ms}^{-1}$.

M8.4. Suppose that a meridional dependence is included in the eddy fields in the two-layer model so that (8.17) becomes

$$\psi_m = A\cos(my)\,e^{ik(x-ct)};\;\; \psi_T = B\cos(my)\,e^{ik(x-ct)}$$

where $m = 2\pi/L_y$ and $L_y = 3000\,\text{km}$. Determine the zonal wavelength of the most unstable wave. Solve for the fields ψ_m, ψ_T, and ω_2' and plot (x, y) cross sections of these fields using the MATLAB script **contour_sample.m** as a template. *Hint:* The solutions in this case are similar to those of Problems 8.6–8.8 but with k^2 replaced by $k^2 + m^2$ everywhere.

M8.5. The MATLAB script **eady_model_1.m** shows vertical and meridional cross sections of the solution for the Eady model (Section 8.4.3) for the most unstable wave mode at time $t = 0$. Modify this script to plot the solution for the Eady wave corresponding to the neutral stability condition given in (8.69). Explain the vertical structure in this case in terms of quasi-geostrophic theory.

Suggested References

Charney (1947) is the classic paper on baroclinic instability. The mathematical level is advanced, but Charney's paper contains an excellent qualitative summary of the main results which is very readable.

Hoskins, McIntyre, and Robertson (1985) discuss cyclogenesis and baroclinic instability from a potential vorticity perspective.

Pedlosky, *Geophysical Fluid Dynamics* (2nd ed.), contains a thorough treatment of normal mode baroclinic instability at a mathematically advanced level.

Pierrehumbert and Swanson (1995) review numerous aspects of baroclinic instability, including spatiotemporal development.

CHAPTER 9

Mesoscale Circulations

The previous chapters focused primarily on the dynamics of synoptic and planetary scale circulations. Such large-scale motions are strongly influenced by the rotation of the earth so that outside the equatorial zone the Coriolis force dominates over inertia (i.e., the Rossby number is small). To a first approximation, as shown in Chapter 6, large-scale motions can be modeled by quasi-geostrophic theory.

The study of quasi-geostrophic motions has been a central theme of dynamic meteorology for many years. Not all important circulations fit into the quasi-geostrophic classification, however. Some have Rossby numbers of order unity, and some are hardly at all influenced by the rotation of the earth. Such circulations include a wide variety of phenomena. They all, however, are characterized by horizontal scales that are smaller than the synoptic scale (i.e., the *macroscale* of motion), but larger than the scale of an individual fair weather cumulus cloud (i.e., the *microscale*). Hence, they can be classified conveniently as *mesoscale* circulations. Most severe weather is associated with mesoscale motion systems. Thus, understanding of the mesoscale is of both scientific and practical importance.

9.1 ENERGY SOURCES FOR MESOSCALE CIRCULATIONS

Mesoscale dynamics is generally defined to include the study of motion systems that have horizontal scales in the range of about 10 to 1000 km. It includes circulations ranging from thunderstorms and internal gravity waves at the small end of the scale to fronts and hurricanes at the large end. Given the diverse nature of mesoscale systems, it is not surprising that there is no single conceptual framework, equivalent to the quasi-geostrophic theory, that can provide a unified model for the dynamics of the mesoscale. Indeed, the dominant dynamical processes vary enormously depending on the type of mesoscale circulation system involved.

Possible sources of mesoscale disturbances include instabilities that occur intrinsically on the mesoscale, forcing by mesoscale thermal or topographic sources, nonlinear transfer of energy from either macroscale or microscale motions, and interaction of cloud physical and dynamical processes.

Although instabilities associated with the mean velocity or thermal structure of the atmosphere are a rich source of atmospheric disturbances, most instabilities have their maximum growth rates either on the large scale (baroclinic and most barotropic instability) or on the small scale (convection and Kelvin–Helmholtz instability). Only symmetric instability (to be discussed in Section 9.3) appears to be an intrinsically mesoscale instability.

Mountain waves created by flow over individual peaks are generally regarded as small-scale phenomena. However, flow over large mountain ranges can produce orographic disturbances in the 10- to 100-km mesoscale range, whose characteristics depend on the mean wind and static stability profiles and the scale of the orography. Flow over mountain ranges, such as the Front Range of the Colorado Rockies, can under some conditions of mean flow and static stability lead to strong downslope wind storms.

Energy transfer from small scales to the mesoscale is a primary energy source for mesoscale convective systems. These may start as individual convective cells, which grow and combine to form thunderstorms, convective complexes such as squall lines and mesocyclones, and even hurricanes. Conversely, energy transfer from the large scale associated with temperature and vorticity advection in synoptic-scale circulations is responsible for the development of frontal circulations.

9.2 FRONTS AND FRONTOGENESIS

In the discussion of baroclinic instability in Chapter 8 the mean thermal wind U_T was taken to be a constant independent of the y coordinate. That assumption was necessary to obtain a mathematically simple model that retained the basic instability mechanism. It was pointed out in Section 6.1, however, that baroclinicity is

not distributed uniformly in the atmosphere. Rather, horizontal temperature gradients tend to be concentrated in baroclinic zones associated with tropospheric jet streams. Not surprisingly, the development of baroclinic waves is also concentrated in such regions.

We showed in Section 8.3 that the energetics of baroclinic waves require that they remove available potential energy from the mean flow. Thus, on average baroclinic wave development tends to weaken the meridional temperature gradient (i.e., reduce the mean thermal wind). The mean pole to equator temperature gradient is of course restored continually by differential solar heating, which maintains the time-averaged temperature gradient pattern. In addition there are transient dynamical processes that produce zones with greatly enhanced temperature gradients within individual baroclinic eddies. Such zones, which are particularly intense at the surface, are referred to as *fronts*. Processes that generate fronts are called *frontogenetic*. Frontogenesis usually occurs in association with developing baroclinic waves, which in turn are concentrated in the storm tracks associated with the time-mean jetstreams. Thus, even though on average baroclinic disturbances transport heat down the mean temperature gradient and tend to weaken the temperature difference between the polar and tropical regions, locally the flow associated with baroclinic disturbances may actually enhance the temperature gradient.

9.2.1 The Kinematics of Frontogenesis

A complete discussion of the dynamics of frontogenesis is beyond the scope of this text. A qualitative description of frontogenesis can be obtained, however, by considering the evolution of the temperature gradient when temperature is treated as a passive tracer in a specified horizontal flow field. Such an approach is referred to as *kinematic*; it considers the effects of advection on a field variable without reference to the underlying physical forces, or to any influence of the advected tracer on the flow field.

The influence of a purely geostrophic flow on the temperature gradient was given in terms of the **Q** vector in (6.46). If for simplicity we focus on the meridional temperature gradient, then from (6.46b) neglecting ageostrophic and diabatic effects gives

$$\frac{D_g}{Dt}\left(\frac{\partial T}{\partial y}\right) = -\left[\frac{\partial u_g}{\partial y}\frac{\partial T}{\partial x} - \frac{\partial u_g}{\partial x}\frac{\partial T}{\partial y}\right] \tag{9.1}$$

where we have used the fact that the geostrophic wind is nondivergent so that $\partial v_g/\partial y = -\partial u_g/\partial x$. The two terms within the brackets on the right in (9.1) can be interpreted as the forcing of the meridional temperature gradient by horizontal shear deformation and stretching deformation, respectively.

Horizontal shear has two effects on a fluid parcel; it tends to rotate the parcel (due to shear vorticity) and to deform the parcel through stretching parallel to the shear vector (i.e., along the x axis in Fig. 9.1a) and shrinking along the horizontal

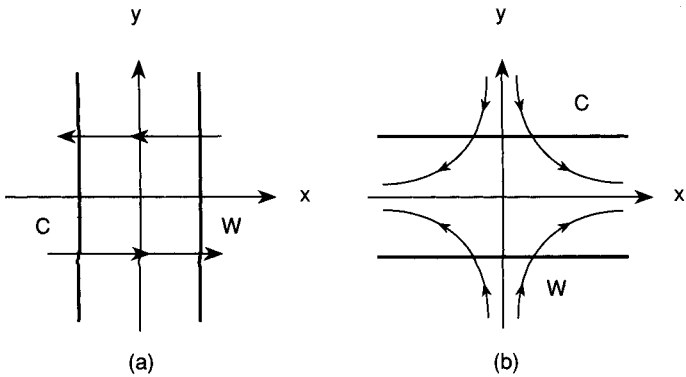


Fig. 9.1 Frontogenetic flow configurations: (a) horizontal shearing deformation and (b) horizontal stretching deformation.

direction perpendicular to the shear vector. Thus, the x-directed temperature gradient in Fig. 9.1a is both rotated into the positive y direction and intensified by the shear. Horizontal shear is an important frontogenetic mechanism in both cold and warm fronts. For example, in the schematic surface pressure chart of Fig. 9.2 the geostrophic wind has a northerly component west of point B and a southerly component east of point B. The resulting cyclonic shear will tend to rotate the isotherms and to concentrate them along the line of maximum shear passing through B. (Note the strong cold advection northwest of B and the weak thermal advection southeast of B.)

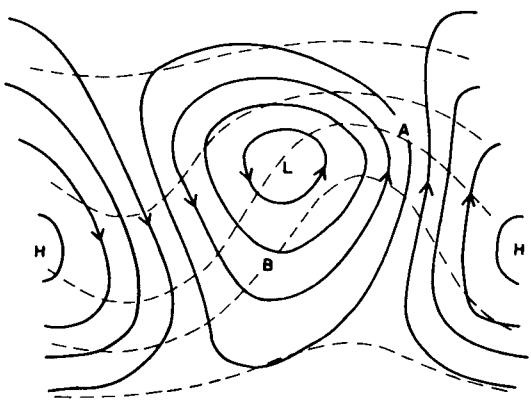


Fig. 9.2 Schematic surface isobars (solid lines) and isotherms (dashed lines) for a baroclinic wave disturbance. Arrows show direction of geostrophic wind. Horizontal stretching deformation intensifies the temperature gradient at A, and horizontal shear deformation intensifies the gradient at B. (After Hoskins and Bretherton, 1972. Reproduced with permission of the American Meteorological Society.)

Horizontal *stretching deformation* tends to advect the temperature field so that the isotherms become concentrated along the axis of *dilation* (the x axis in Fig. 9.1b), provided that the initial temperature field has a finite gradient along the axis of *contraction* (the y axis in Fig. 9.1b). That this effect is represented by the second term on the right in (9.1) can be verified by noting from Fig. 9.1b that $\partial T/\partial y < 0$ and $\partial u_g/\partial x > 0$.

The velocity field shown in Fig. 9.1b is a pure stretching deformation field. A pure deformation field is both irrotational and nondivergent. Thus, a parcel advected by a pure deformation field will merely have its shape changed in time, without any rotation or change in horizontal area. The deformation field of Fig. 9.1b has a streamfunction given by $\psi = -Kxy$, where K is a constant. Such a field is characterized by the rate at which advection changes the shape of a horizontal area element. This can be illustrated by considering the rectangular element with sides δx and δy. The shape can be represented by the ratio $\delta x/\delta y$, and the fractional rate of change of shape can thus be expressed as

$$\frac{1}{(\delta x/\delta y)}\frac{D(\delta x/\delta y)}{Dt} = \frac{1}{\delta x}\frac{D\delta x}{Dt} - \frac{1}{\delta y}\frac{D\delta y}{Dt} = \frac{\delta u}{\delta x} - \frac{\delta v}{\delta y} \approx \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}$$

It is easily verified that the fractional rate of change of $\delta x/\delta y$ for the velocity field in Fig. 9.1b equals $+2K$. Thus, a square parcel with sides parallel to the x and y axes would be deformed into a rectangle as the sides parallel to the x axis stretch and those parallel to the y axis contract in time at a constant rate.

Horizontal deformation at low levels is an important mechanism for the development of both cold and warm fronts. In the example of Fig. 9.2 the flow near point A has $\partial u/\partial x > 0$ and $\partial v/\partial y < 0$ so that there is a stretching deformation field present with its axis of contraction nearly orthogonal to the isotherms. This deformation field leads to strong warm advection south of point A and weak warm advection north of point A.

Although, as shown in Fig. 9.2, the low-level flow in the vicinity of a developing warm front may resemble a pure deformation field, the total flow in the upper troposphere in baroclinic disturbances seldom resembles that of a pure deformation field due to the presence of strong mean westerlies. Rather, a combination of mean flow plus horizontal stretching deformation produces a *confluent* flow as shown in Fig. 9.3. Such confluence acts to concentrate the cross-stream temperature gradient as parcels move downstream. Confluent regions are always present in the tropospheric jetstream due to the influence of quasi-stationary planetary-scale waves on the position and intensity of the jet. In fact, even a monthly mean 500-hPa chart (see Fig. 6.3) reveals two regions of large-scale confluence immediately to the east of the continents of Asia and North America. Observationally, these two regions are known to be regions of intense baroclinic wave development and frontogenesis.

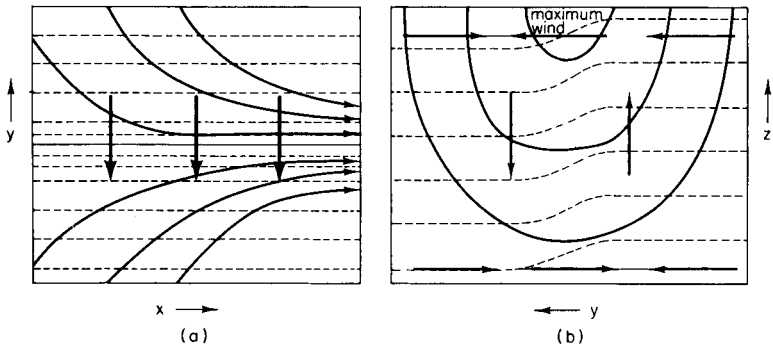


Fig. 9.3 (a) Horizontal streamlines, isotherms, and **Q** vectors in a frontogenetic confluence. (b) Vertical section across the confluence showing isotachs (solid), isotherms (dashed), and vertical and transverse motions (arrows). (After Sawyer, 1956.)

The mechanisms of horizontal shear and horizontal stretching deformation discussed above operate to concentrate the pole–equator temperature gradient on the synoptic scale (~1000 km). The time scale on which these processes operate can be estimated with the aid of (9.1). Suppose that the geostrophic wind consists of a pure deformation field so that $u_g = Kx$, and $v_g = -Ky$, and let T be a function of y only. Then (9.1) simplifies to

$$\frac{D_g}{Dt}\left(\frac{\partial T}{\partial y}\right) = K\frac{\partial T}{\partial y}$$

so that following the geostrophic motion

$$\frac{\partial T}{\partial y} = e^{Kt}\left(\frac{\partial T}{\partial y}\right)_{t=0}$$

Typically, $K \sim 10^{-5}\ \mathrm{s}^{-1}$ so that the temperature gradient amplifies by a factor of 10 in about 3 days. This is much slower than observed rates of atmospheric frontogenesis.

Thus, geostrophic deformation fields alone cannot cause the rapid frontogenesis often observed in extratropical systems, in which the temperature gradient can become concentrated in a zone of ~50 km width on a time scale of less than a day. This rapid reduction in scale is caused primarily by the frontogenetic character of the secondary circulation driven by the quasi-geostrophic synoptic-scale flow (moist processes may also be important in rapid frontogenesis).

The nature of the secondary flow may be deduced from the pattern of **Q** vectors illustrated in Fig. 9.3a. As discussed in Section 6.4.2, the divergence of **Q** forces a secondary ageostrophic circulation. For the situation of Fig. 9.3 this circulation is in the cross-frontal plane as illustrated in Fig. 9.3b. Advection of the temperature

field by this ageostrophic circulation tends to increase the horizontal temperature gradient at the surface on the warm side of the jet axis. Temperature advection by the upper-level secondary circulation, however, tends to concentrate the temperature gradient on the cold side of the jet axis. As a result, the frontal zone tends to slope toward the cold air side with height. The differential vertical motion associated with the ageostrophic circulation tends to weaken the front in the midtroposphere due to adiabatic temperature changes (adiabatic warming on the cold side of the front and adiabatic cooling on the warm side). For this reason fronts are most intense in the lower troposphere and near the tropopause.

The secondary circulation associated with frontogenesis is required to maintain the thermal wind balance between the along-front flow and the cross-front temperature gradient in the presence of advective processes that tend to destroy this balance. The concentration of the isotherms by the horizontal advection increases the cross-stream pressure gradient at upper levels, and thus requires an increase in the vertical shear of the along-jet flow in order to maintain thermal wind balance. The required upper-level acceleration of the jet is produced by the Coriolis force caused by the cross-jet ageostrophic wind, which develops in response to the increased cross-stream pressure gradient in the jetstream core. As the jet accelerates, cyclonic vorticity must be generated on the cold side of the jet axis and anticyclonic vorticity on the warm side. These vorticity changes require that the horizontal flow at the jetstream level be convergent on the cold side of the jet axis and divergent on the warm side of the jet axis. The resulting vertical circulation and low-level secondary ageostrophic motion required by mass continuity are indicated in Fig. 9.3b.

9.2.2 Semigeostrophic Theory

To analyze the dynamics of the frontogenetic motion fields discussed in the previous subsection, it is convenient to use the Boussinesq approximation introduced in Section 7.4, in which density is replaced by a constant reference value ρ_0 except where it appears in the buoyancy force. This approximation simplifies the equations of motion without affecting the main features of the results. It is also useful to replace the total pressure and density fields with deviations from their standard atmosphere values. Thus, we let $\Phi(x, y, z, t) = (p - p_0)/\rho_0$ designate the pressure deviation normalized by density and $\Theta = \theta - \theta_0$ designate the potential temperature deviation where $p_0(z)$ and $\theta_0(z)$ are the height-dependent standard atmosphere values of pressure and potential temperature, respectively.

With the above definitions, the horizontal momentum equations, thermodynamic energy equation, hydrostatic approximation, and continuity equation become

$$\frac{Du}{Dt} - fv + \frac{\partial \Phi}{\partial x} = 0 \tag{9.2}$$

$$\frac{Dv}{Dt} + fu + \frac{\partial \Phi}{\partial y} = 0 \tag{9.3}$$

$$\frac{D\Theta}{Dt} + w\frac{d\theta_0}{dz} = 0 \tag{9.4}$$

$$b \equiv \frac{g\Theta}{\theta_{00}} = \frac{\partial \Phi}{\partial z} \tag{9.5}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{9.6}$$

where b is the *buoyancy*, θ_{00} is a constant reference value of potential temperature, and

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}$$

From the discussion of the previous subsection it should be clear that the horizontal scale of variations parallel to a front is much larger than the cross-frontal scale. This scale separation suggests that to a first approximation we can model fronts as two-dimensional structures. For convenience we choose a coordinate system in which the front is stationary and take the cross-frontal direction to be parallel to the y axis. Then $L_x \gg L_y$, where L_x and L_y designate the along-front and cross-front length scales. Similarly, $U \gg V$ where U and V, respectively, designate the along-front and cross-front velocity scales. Fig. 9.4 shows these scales relative to the front.

Letting $U \sim 10\,\mathrm{m\,s^{-1}}$, $V \sim 1\,\mathrm{m\,s^{-1}}$, $L_x \sim 1000\,\mathrm{km}$, and $L_y \sim 100\,\mathrm{km}$, we find that it is possible to utilize the differing scales of the along-front and cross-front motion to simplify the dynamics. Assuming that $D/Dt \sim V/L_y$ (the cross-front advection time scale) and defining a Rossby number, $\mathrm{Ro} \equiv V/fL_y \ll 1$, the magnitude of the ratios of the inertial and Coriolis terms in the x and y components of the momentum equation can be expressed as

$$\frac{|Du/Dt|}{|fv|} \sim \frac{UV/L_y}{fV} \sim \mathrm{Ro}\left(\frac{U}{V}\right) \sim 1$$

$$\frac{|Dv/Dt|}{|fu|} \sim \frac{V^2/L_y}{fU} \sim \mathrm{Ro}\left(\frac{V}{U}\right) \sim 10^{-2}$$

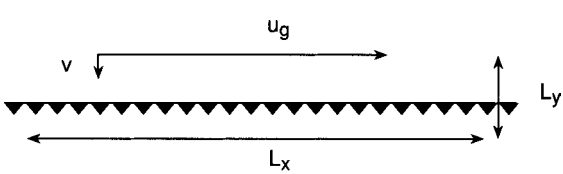


Fig. 9.4 Velocity and length scales relative to a front parallel to the x axis.

The along-front velocity is in geostrophic balance with the cross-front pressure gradient with error of order 1%, but geostrophy does not hold even approximately for the cross-front velocity. Therefore, if the geostrophic wind components are defined by

$$fu_g = -\partial\Phi/\partial y, \qquad fv_g = \partial\Phi/\partial x$$

and we separate the horizontal velocity field into geostrophic and ageostrophic parts, to a good approximation $u = u_g$, but $v = v_g + v_a$, where v_g and v_a are the same order of magnitude.

The x component of the horizontal momentum equation (9.2), the thermodynamic energy equation (9.4), and the continuity equation (9.6) for frontal scaling can thus be expressed as

$$\frac{Du_g}{Dt} - fv_a = 0 \tag{9.7}$$

$$\frac{Db}{Dt} + wN^2 = 0 \tag{9.8}$$

$$\frac{\partial v_a}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{9.9}$$

Here, (9.8) is obtained by using (9.5) to replace Θ by b in (9.4), and N is the buoyancy frequency defined in terms of potential temperature as

$$N^2 \equiv \frac{g}{\theta_{00}}\frac{\partial\theta_0}{\partial z}$$

Because the along-front velocity is in geostrophic balance, u_g and b are related by the thermal wind relationship:

$$f\frac{\partial u_g}{\partial z} = -\frac{\partial b}{\partial y} \tag{9.10}$$

Note that (9.7) and (9.8) differ from their quasi-geostrophic analogues; although zonal momentum is still approximated geostrophically, and advection parallel to the front is geostrophic, the advection of momentum and temperature across the front is due not only to the geostrophic wind, but to the ageostrophic (v_a, w) circulation:

$$\frac{D}{Dt} = \frac{D_g}{Dt} + \left(v_a\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\right)$$

where D_g/Dt was defined in (6.8). Replacement of momentum by its geostrophic value in (9.7) is referred to as the *geostrophic momentum* approximation, and the resulting set of prediction equations are called the *semigeostrophic* equations.[1]

[1] Some authors reserve this name for a version of the equations written in a transformed set of coordinates called geostrophic coordinates (e.g., Hoskins, 1975).

9.2.3 Cross-Frontal Circulation

Equations (9.7)–(9.10) form a closed set that can be used to determine the cross-frontal ageostrophic circulation in terms of the distribution of zonal wind or temperature. Suppose that the large-scale geostrophic flow is acting to intensify the north–south temperature gradient through deformation as shown in Fig. 9.3. As the temperature gradient increases, the vertical shear of the zonal wind must also increase to maintain geostrophic balance. This requires an increase of u_g in the upper troposphere, which must be produced by the Coriolis force associated with a cross-frontal ageostrophic circulation [see (9.7)]. The structure of this secondary circulation can be computed by deriving an equation analogous to the omega equation discussed in Section 6.4.2.

We first differentiate (9.8) with respect to y and use the chain rule to express the result as

$$\frac{D}{Dt}\left(\frac{\partial b}{\partial y}\right) = Q_2 - \frac{\partial v_a}{\partial y}\frac{\partial b}{\partial y} - \frac{\partial w}{\partial y}\left(N^2 + \frac{\partial b}{\partial z}\right) \tag{9.11}$$

where

$$Q_2 = -\frac{\partial u_g}{\partial y}\frac{\partial b}{\partial x} - \frac{\partial v_g}{\partial y}\frac{\partial b}{\partial y} \tag{9.12}$$

is just the y component of the **Q** vector discussed in Section 6.4.2, but expressed in the Boussinesq approximation.

Next we differentiate (9.7) with respect to z, again use the chain rule to rearrange terms, and the thermal wind equation (9.10) to replace $\partial u_g/\partial z$ by $\partial b/\partial y$ on the right-hand side. The result can then be written as

$$\frac{D}{Dt}\left(f\frac{\partial u_g}{\partial z}\right) = Q_2 + \frac{\partial v_a}{\partial z} f\left(f - \frac{\partial u_g}{\partial y}\right) + \frac{\partial w}{\partial z}\frac{\partial b}{\partial y} \tag{9.13}$$

Again, as shown in Section 6.4.2, the geostrophic forcing (given by Q_2) tends to destroy thermal wind balance by changing the temperature gradient and vertical shear parts of the thermal wind equation in equal but opposite senses. This tendency of geostrophic advection to destroy geostrophic balance is counteracted by the cross-frontal secondary circulation.

In this case the secondary circulation is a two-dimensional transverse circulation in the y, z plane. It can thus be represented in terms of a meridional streamfunction ψ_M defined so that

$$v_a = -\partial\psi_M/\partial z, \qquad w = \partial\psi_M/\partial y \tag{9.14}$$

which identically satisfies the continuity equation (9.9). Adding (9.11) and (9.13), and using the thermal wind balance (9.10) to eliminate the time derivative and (9.14) to eliminate v_a and w, we obtain the *Sawyer–Eliassen equation*

$$N_s^2\frac{\partial^2\psi_M}{\partial y^2} + F^2\frac{\partial^2\psi_M}{\partial z^2} + 2S^2\frac{\partial^2\psi_M}{\partial y\partial z} = 2Q_2 \tag{9.15}$$

where

$$N_s^2 \equiv N^2 + \frac{\partial b}{\partial z}, \quad F^2 \equiv f\left(f - \frac{\partial u_g}{\partial y}\right) = f\frac{\partial M}{\partial y}, \quad S^2 \equiv -\frac{\partial b}{\partial y} \tag{9.16}$$

where M is the absolute momentum defined below (7.53).

Equation (9.15) can be compared with the quasi-geostrophic version obtained by neglecting advection by the ageostrophic circulation in (9.7) and (9.8). This has the form

$$N^2\frac{\partial^2\psi_M}{\partial y^2} + f^2\frac{\partial^2\psi_M}{\partial z^2} = 2Q_2 \tag{9.17}$$

Thus, in the quasi-geostrophic case the coefficients in the differential operator on the left depend only on the standard atmosphere static stability, N, and the planetary vorticity, f, whereas in the semigeostrophic case they depend on the deviation of potential temperature from its standard profile through the N_s and S terms and the absolute vorticity through the F term.

An equation of the form (9.17), in which the coefficients of the derivatives on the left are positive, is referred to as an *elliptic boundary value problem*. It has a solution ψ_M that is uniquely determined by Q_2 plus the boundary conditions. For a situation such as that of Fig. 9.1b, with both $\partial v_g/\partial y$ and $\partial b/\partial y$ negative, the forcing term Q_2 is negative in the frontal region. The streamfunction in that case describes a circulation symmetric about the y axis, with rising on the warm side and sinking on the cold side. The semigeostrophic case (9.15) is also an elliptic boundary value problem provided that $N_s^2F^2 - S^4 > 0$. It can be shown (see Problem 9.1) that this condition requires that the Ertel potential vorticity be positive in the Northern Hemisphere and negative in the Southern Hemisphere, which is nearly always the case in the extratropics for an unsaturated atmosphere.

The spatial variation of the coefficients in (9.16), and the presence of the cross-derivative term, produce a distortion of the secondary circulation, as shown in Fig. 9.5. The frontal zone slopes toward the cold air side with height; there is an intensification of the cross-frontal flow near the surface in the region of large absolute vorticity on the warm air side of the front, and a tilting of the circulation with height.

The influence of the ageostrophic circulation on the time scale for frontogenesis can be illustrated by comparing the processes included in quasi-geostrophic and semigeostrophic frontogenesis. For semigeostrophic advection there is a positive feedback that reduces the time scale of frontogenesis greatly compared to that in the quasi-geostrophic case. As the temperature contrast increases, Q_2 increases, and the secondary circulation must also increase so that the amplification *rate* of $|\partial T/\partial y|$ increases with $|\partial T/\partial y|$ rather than remaining constant as in the quasi-geostrophic case. Because of this feedback, in the absence of frictional effects, the semigeostrophic model can produce an infinite temperature gradient at the surface in less than half a day.

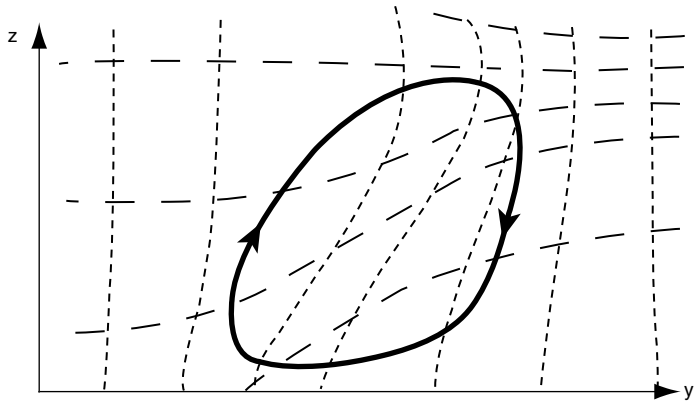


Fig. 9.5 Relationship of the ageostrophic circulation (heavy curve with arrows) in two-dimensional frontogenesis to the potential temperature field (long dashes) and absolute momentum field (short dashes). Cold air is on the right and warm air on the left. Note the tilt of the circulation toward the cold air side and the enhanced gradients of the absolute momentum and potential temperature fields in the frontal zone.

9.3 SYMMETRIC BAROCLINIC INSTABILITY

Observations indicate that mesoscale bands of cloud and precipitation commonly occur in association with synoptic-scale systems. These are often aligned with fronts, but are separate entities. In common with fronts, such features are generally associated with strong baroclinicity and have length scales parallel to the mean wind shear that are much larger than the scales across the wind shear. A plausible source for such features is a two-dimensional form of baroclinic instability known as *symmetric instability* (also referred to as slantwise convection).

For typical atmospheric conditions, buoyancy tends to stabilize air parcels against vertical displacements, and rotation tends to stabilize parcels with respect to horizontal displacements. Instability with respect to vertical displacements is referred to as hydrostatic (or simply, static) instability (see Section 2.7.3). For an unsaturated atmosphere, static stability requires that the local buoyancy frequency squared be positive ($N_s^2 > 0$). Instability with respect to horizontal displacements, however, is referred to as inertial instability (see Section 7.5.1). The condition for inertial stability (7.54) requires that F^2 as defined in (9.16) be positive.

If parcels are displaced along slantwise paths rather than vertical or horizontal paths, it is possible under certain conditions for the displacements to be unstable even when the conditions for ordinary static and inertial stability are separately satisfied. Such instability can occur only in the presence of vertical shear of the mean horizontal wind and may be regarded as a special form of baroclinic instability in which the perturbations are independent of the coordinate parallel to the

mean flow. Alternatively, as shown below, symmetric instability may be regarded as isentropic inertial instability.

For convenience in deriving the conditions for symmetric instability, we use the Boussinesq equations of the previous section and assume that the flow is independent of the x coordinate. The mean wind is directed along the x axis and is in thermal wind balance with the meridional temperature gradient:

$$f\partial u_g/\partial z = -\partial b/\partial y = -(g/\theta_{00})\,\partial\Theta/\partial y \tag{9.18}$$

where as before θ_{00} is a constant reference value of potential temperature.

Following Sections 2.7.3 and 7.5.1, we measure the stability with respect to vertical displacements by the distribution of total potential temperature $\theta = \theta_0 + \Theta$ and that with respect to horizontal displacements by the distribution of the absolute zonal momentum, $M \equiv fy - u_g$ (where $\partial M/\partial y = f - \partial u_g/\partial y$ is the zonal mean absolute vorticity).

For a barotropic flow, potential temperature surfaces are oriented horizontally and absolute momentum surfaces are oriented vertically in the meridional plane. When the mean flow is westerly and increases with height, however, the potential temperature and absolute momentum surfaces both slope upward toward the pole. The comparative strengths of the vertical and horizontal restoring forces in the midlatitude troposphere are given by the ratio $N_s^2/(f\partial M/\partial y)$. For typical tropospheric conditions this ratio is $\sim 10^4$. Thus, parcel motion in the plane orthogonal to the mean flow will remain much closer to θ surfaces than to M surfaces. It is thus natural to utilize isentropic coordinates to analyze parcel displacements. The arguments of Section 7.5.1 still apply, provided that derivatives with respect to y are taken at constant θ. The stability of such motions thus depends on the relative slope of the θ and M surfaces. Ordinarily the M surfaces will slope more than the θ surfaces (see Fig. 9.5), and parcel displacements are stable. However, when the θ surfaces slope more than the M surfaces so that

$$f(\partial M/\partial y)_\theta < 0 \tag{9.19}$$

the flow is unstable with respect to displacements along the θ surfaces. This situation may occur in regions of a very strong horizontal temperature gradient and weak vertical stability. The condition in (9.19) is similar to the criterion of (7.54) for inertial instability, except that the derivative of M is here taken along a sloping θ surface.

If (9.19) is multiplied by $-g(\partial\theta/\partial p)$, the criterion for symmetric instability can be expressed in terms of the distribution of Ertel potential vorticity (4.12) in the simple form

$$f\bar{P} < 0 \tag{9.20}$$

where $\bar{P}$ is the potential vorticity of the basic state geostrophic flow. Thus, if the initial state potential vorticity in the Northern Hemisphere is everywhere positive, then symmetric instability cannot develop through adiabatic motions, as potential vorticity is conserved following the motion and will always remain positive.

To demonstrate that (9.19) is the condition for symmetric instability, we consider the change in mean kinetic energy required for exchange of the tubes of fluid labeled 1 and 2 in Fig. 9.6. (These tubes are located at y_1 and $y_2 = y_1 + \delta y$, respectively, and are assumed to extend infinitely along the x axis so that the problem is two dimensional.) Because the tubes lie on the same potential temperature surface, they have the same available potential energy. Thus a spontaneous exchange of parcels is possible if δ(KE), the kinetic energy of the zonal flow after exchange minus that in the initial state, is negative. Otherwise some external source of energy is required to furnish the kinetic energy of the meridional and vertical motions required for the exchange.

Initially the motion of the tubes is parallel to the x axis and in geostrophic balance so that the absolute momentum for the two tubes is

$$\begin{aligned} M_1 &= fy_1 - u_1 = fy_1 - u_g(y_1, z), \\ M_2 &= fy_2 - u_2 = fy_1 + f\delta y - u_g(y_1 + \delta y, z + \delta z) \end{aligned} \tag{9.21}$$

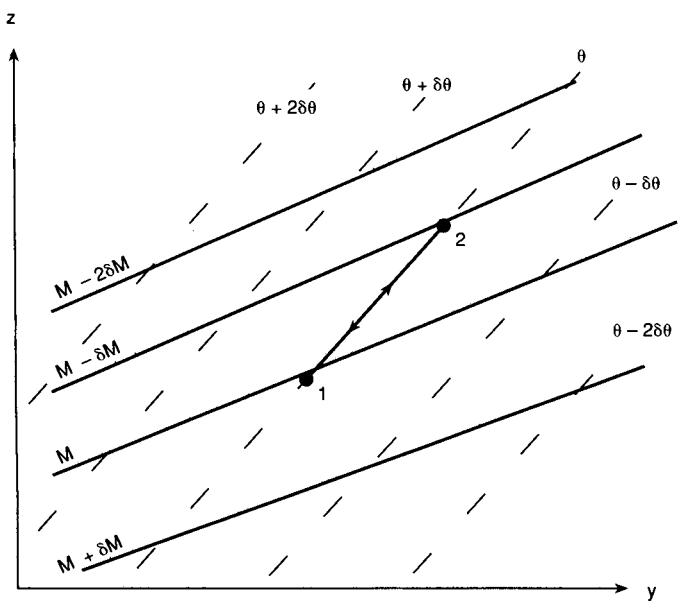


Fig. 9.6 Cross section showing isolines of absolute momentum and potential temperature for a symmetrically unstable basic state. Motion along the isentropic path between points labeled 1 and 2 is unstable, as M decreases with latitude along the path. See text for details.

Absolute momentum is conserved by the tubes so that after exchange the perturbation zonal velocities are given by

$$M_1' = fy_1 + f\delta y - u_1' = M_1, \quad M_2' = fy_1 - u_2' = M_2 \tag{9.22}$$

Eliminating M_1 and M_2 between (9.21) and (9.22) and solving for the disturbance zonal wind, we obtain

$$u_1' = f\delta y + u_1, \quad u_2' = -f\delta y + u_2$$

The difference in zonal kinetic energy between the final state and the initial state is given by

$$\begin{aligned} \delta\,(\mathrm{KE}) &= \frac{1}{2}\left(u_1'^2 + u_2'^2\right) - \frac{1}{2}\left(u_1^2 + u_2^2\right) \\ &= f\delta y\,(u_1 - u_2 + f\delta y) = f\delta y\,(M_2 - M_1)) \end{aligned} \tag{9.23}$$

Thus, δ(KE) is negative, and unforced meridional motion may occur, provided that $f(M_2 - M_1) < 0$. This is equivalent to the condition (9.19), as the tubes lie along the same θ surface.

To estimate the likelihood that conditions for symmetric instability may be satisfied, it is useful to express the stability criterion in terms of a mean-flow *Richardson number*. To do this we first note that the slope of an M surface can be estimated from noting that on an M surface

$$\delta M = \frac{\partial M}{\partial y}\delta y + \frac{\partial M}{\partial z}\delta z = 0$$

so that the ratio of δz to δy at constant M is

$$\left(\frac{\delta z}{\delta y}\right)_M = \left(-\frac{\partial M}{\partial y}\right)\Big/\left(\frac{\partial M}{\partial z}\right) = \left(f - \frac{\partial u_g}{\partial y}\right)\Big/\left(\frac{\partial u_g}{\partial z}\right) \tag{9.24}$$

Similarly, the slope of a potential temperature surface is

$$\left(\frac{\delta z}{\delta y}\right)_\theta = \left(-\frac{\partial \theta}{\partial y}\right)\Big/\left(\frac{\partial \theta}{\partial z}\right) = \left(f\frac{\partial u_g}{\partial z}\right)\Big/\left(\frac{g}{\theta_{00}}\frac{\partial \theta}{\partial z}\right) \tag{9.25}$$

where we have used the thermal wind relationship to express the meridional temperature gradient in terms of the vertical shear of the zonal wind. The ratio of (9.24) to (9.25) is simply

$$\left(\frac{\delta z}{\delta y}\right)_M\Big/\left(\frac{\delta z}{\delta y}\right)_\theta = f\left(f - \frac{\partial u_g}{\partial y}\right)\left(\frac{g}{\theta_{00}}\frac{\partial \theta}{\partial z}\right)\Big/\left[f^2\left(\frac{\partial u_g}{\partial z}\right)^2\right] = \frac{F^2 N_s^2}{S^4}$$

where the notation in the last term is defined in (9.16).

Recalling that symmetric instability requires that the slopes of the θ surfaces exceed those of the M surfaces, the necessary condition for instability of geostrophic flow parallel to the x axis becomes

$$\left(\frac{\delta z}{\delta y}\right)_M \Big/ \left(\frac{\delta z}{\delta y}\right)_\theta = f\left(f - \frac{\partial u_g}{\partial y}\right)\mathrm{Ri}\Big/ f^2 = \frac{F^2 N_s^2}{S^4} < 1 \tag{9.26}$$

where the mean-flow *Richardson number* Ri is defined as

$$\mathrm{Ri} \equiv \left(\frac{g}{\theta_{00}}\frac{\partial\theta}{\partial z}\right)\Bigg/\left(\frac{\partial u_g}{\partial z}\right)^2$$

Thus, if the relative vorticity of the mean flow vanishes ($\partial u_g/\partial y = 0$), Ri < 1 is required for instability.

The condition (9.26) can be related to (9.20) by observing that (9.26) requires that $F^2 N_s^2 - S^4 < 0$ for symmetric instability. As is to be shown in Problem 9.1,

$$F^2 N_s^2 - S^4 = \left(\rho f g \big/ \theta_{00}\right)\overline{P} \tag{9.27}$$

Because the large-scale potential vorticity, $\bar{P}$, is normally positive in the Northern Hemisphere and negative in the Southern Hemisphere, (9.27) is ordinarily positive in both hemispheres; thus, the condition for symmetric instability is rarely satisfied. If the atmosphere is saturated, however, the relevant static stability condition involves the lapse rate of the equivalent potential temperature, and neutral conditions with respect to symmetric instability may easily occur (see Section 9.5).

Finally, it is worth noting that the condition for stability with respect to symmetric displacements, $F^2 N_s^2 - S^4 > 0$, is identical to the condition that the Sawyer–Eliassen equation (9.15) be an elliptic boundary value problem. Thus, when the flow is stable with respect to symmetric baroclinic perturbations, a nonzero *forced* transverse circulation governed by (9.15) will exist when there is nonzero forcing, Q_2 [see (9.12)]. *Free* transverse oscillations, however, may occur in the absence of forcing. These require including the horizontal acceleration term in the y component of the momentum equation. The resulting equation for the transverse circulation has the form (see Appendix F)

$$\frac{\partial^2}{\partial t^2}\left(\frac{\partial^2\psi_M}{\partial z^2}\right) + N_s^2\frac{\partial^2\psi_M}{\partial y^2} + F^2\frac{\partial^2\psi_M}{\partial z^2} + 2S^2\frac{\partial^2\psi_M}{\partial y\partial z} = 0 \tag{9.28}$$

which should be compared with (9.15). When $F^2 N_s^2 - S^4 > 0$ solutions of (9.28) correspond to stable oscillations, while for $F^2 N_s^2 - S^4 < 0$ solutions are exponentially growing, corresponding to symmetric baroclinic instability.

9.4 MOUNTAIN WAVES

Section 7.5.2 showed that stably stratified air forced to flow over sinusoidally varying surface topography creates oscillations, which can be either vertically propagating or vertically decaying, depending on whether the intrinsic wave frequency relative to the mean flow is less than or greater than the buoyancy frequency. Most topographic features on the surface of the earth do not, however, consist of regularly repeating lines of ridges. In general the distance between large topographic barriers is large compared to the horizontal scales of the barriers. Also, the static stability and the basic state flow are not usually constants, as was assumed in Section 7.5.2, but may vary strongly with height. Furthermore, nonlinear modifications of mountain waves are sometimes associated with strong surface winds along the lee slopes of ridges. Thus, mountain waves are significant features of mesoscale meteorology.

9.4.1 Flow over Isolated Ridges

Just as flow over a periodic series of sinusoidal ridges can be represented by a single sinusoidal function, that is, by a single Fourier harmonic, flow over an isolated ridge can be approximated by the sum of a number of Fourier components (see Section 7.2.1). Thus, any zonally varying topography can be represented by a Fourier series of the form

$$h_M(x) = \sum_{s=1}^{\infty} \mathrm{Re}\left[h_s \exp(ik_s x)\right] \tag{9.29}$$

where h_s is the amplitude of the sth Fourier component of the topography. We can then express the solution to the wave equation (7.46) as the sum of Fourier components:

$$w(x, z) = \sum_{s=1}^{\infty} \mathrm{Re}\left\{W_s \exp\left[i\left(k_s x + m_s z\right)\right]\right\} \tag{9.30}$$

where $W_s = ik_s\bar{u}h_s$, and $m_s^2 = N^2/\bar{u}^2 - k_s^2$.

Individual Fourier modes will yield vertically propagating or vertically decaying contributions to the total solution (9.30) depending on whether m_s is real or imaginary. This in turn depends on whether k_s^2 is less than or greater than $N^2/\bar{u}^2$. Thus, each Fourier mode behaves just as the solution (7.48) for periodic sinusoidal topography. For a narrow ridge, Fourier components with wave numbers greater than $N/\bar{u}$ dominate in (9.29), and the resulting disturbance decays with height. For a broad ridge, components with wave numbers less than $N/\bar{u}$ dominate and the disturbance propagates vertically. In the wide mountain limit where $m_s^2 \approx N^2/\bar{u}^2$, the flow is periodic in the vertical with a vertical wavelength of $2\pi m_s^{-1}$, and phase lines tilt upstream with height as shown in Fig. 9.7.

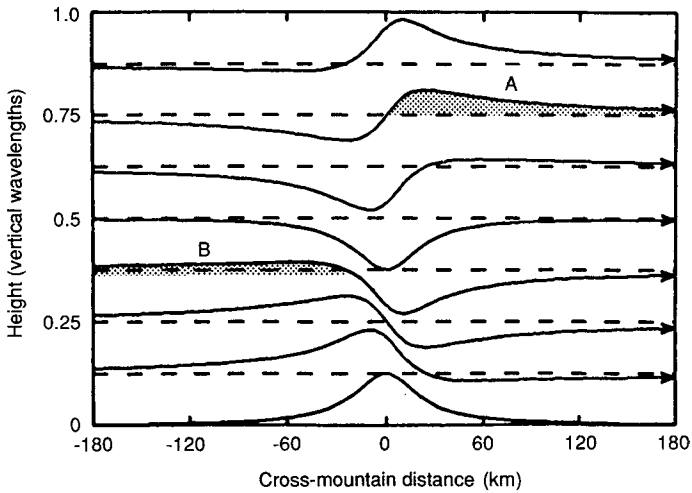


Fig. 9.7 Streamlines of flow over a broad isolated ridge showing upstream phase tilt with height. The pattern is periodic in height and one vertical wavelength is shown. Orographic clouds may form in the shaded areas where streamlines are displaced upward from equilibrium either upstream or downstream of the ridge if sufficient moisture is present. (After Durran, 1990.)

Vertically propagating gravity waves generated by flow over broad topography can produce clouds both upstream and downstream of the topography depending on variations of the moisture distribution with altitude. In the example shown in Fig. 9.7 the positions labeled A and B indicate regions where streamlines are displaced upward downstream and upstream of the ridge, respectively. If sufficient moisture is present, orographic clouds may then form in region A or B as suggested by the shading in Fig. 9.7.

9.4.2 Lee Waves

If $\bar{u}$ and N are allowed to vary in height, then (7.46) must be replaced by

$$\left(\frac{\partial^2 w'}{\partial x^2} + \frac{\partial^2 w'}{\partial z^2}\right) + l^2 w' = 0 \tag{9.31}$$

where the *Scorer parameter*, l, is defined as

$$l^2 = N^2 \Big/ \overline{u}^2 - \overline{u}^{-1} d^2\overline{u} \Big/ dz^2$$

and the condition for vertical propagation becomes $k_s^2 < l^2$.

If the mean cross mountain wind speed increases strongly with height, or there is a low-level stable layer so that N decreases strongly with height, there may

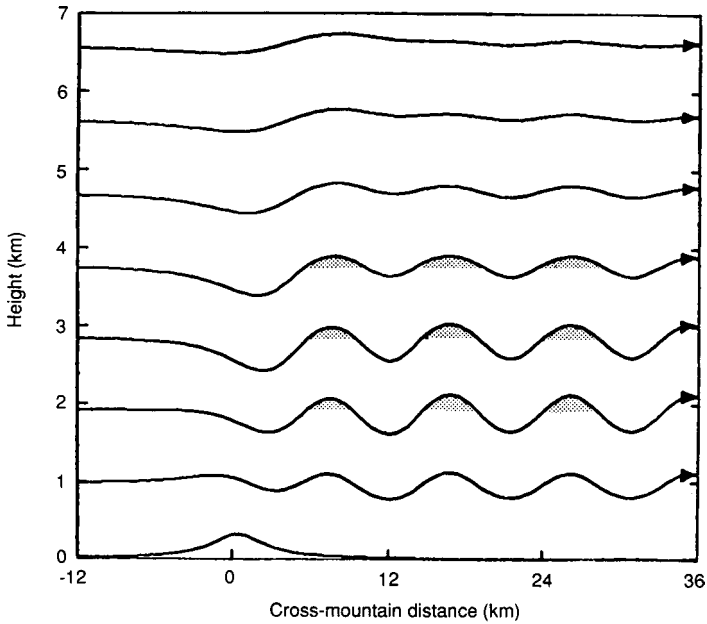


Fig. 9.8 Streamlines for trapped lee waves generated by flow over topography with vertical variation of the Scorer parameter. Shading shows locations where lee wave clouds may form. (After Durran, 1990.)

be a layer near the surface in which vertically propagating waves are permitted, which is topped by a layer in which the disturbance decays in the vertical. In that case vertically propagating waves in the lower layer are reflected when they reach the upper layer. Under some circumstances the waves may be repeatedly reflected from the upper layer and the surface downstream of the mountain, leading to a series of "trapped" lee waves as shown in Fig. 9.8.

Vertical variations in the Scorer parameter can also modify the amplitude of waves that are sufficiently long to propagate vertically through the entire troposphere. Amplitude enhancement leading to wave breaking and turbulent mixing can occur if there is a *critical level* where the mean flow goes to zero ($l \to \infty$).

9.4.3 Downslope Windstorms

Strong downslope surface winds are occasionally observed along the lee slopes of mountain ranges. Although partial reflection of vertically propagating linear gravity waves may produce enhanced surface winds under some conditions, it appears that nonlinear processes are essential to account for observed windstorms associated with stable flow over topography.

To demonstrate the role of nonlinearity, we assume that the troposphere has a stable lower layer of undisturbed depth h topped by a weakly stable upper layer and assume that the lower layer behaves as a barotropic fluid with a free surface $h(x,t)$. We assume that disturbances have zonal wavelengths much greater than the layer depth. The motion of the lower layer may then be described by the shallow water equations of Section 7.3.2, but with the lower boundary condition replaced by

$$w(x, h_M) = Dh_M \big/ Dt = u\partial h_M \big/ \partial x$$

where h_M again denotes the height of the topography.

We first examine the linear behavior of this model by considering steady flow over small-amplitude topography. The linearized shallow water equations (7.20) and (7.21) then become

$$\overline{u}\frac{\partial u'}{\partial x} + \frac{g\delta\rho}{\rho_1}\frac{\partial h'}{\partial x} = 0 \tag{9.32}$$

$$\overline{u}\frac{\partial\left(h' - h_M\right)}{\partial x} + H\frac{\partial u'}{\partial x} = 0 \tag{9.33}$$

Here $\delta\rho/\rho_1$ is the fractional change in density across the interface between the layers, $h' = h - H$, where H is the mean height of the interface and $h' - h_M$ is the deviation from H of the thickness of the lower layer.

The solutions for (9.32) and (9.33) can be expressed as

$$h' = -\frac{h_M\left(\overline{u}^2 \big/ c^2\right)}{\left(1 - \overline{u}^2 \big/ c^2\right)}, \quad u' = \frac{h_M}{H}\left(\frac{\overline{u}}{1 - \overline{u}^2 \big/ c^2}\right) \tag{9.34}$$

where $c^2 \equiv (gH\delta\rho/\rho_1)$ is the shallow water wave speed. The characteristics of the disturbance fields h' and u' depend on the magnitude of the mean-flow *Froude number*, defined by the relation $\mathrm{Fr}^2 = \overline{u}^2 \big/ c^2$. When $\mathrm{Fr} < 1$, the flow is referred to as *subcritical*. In subcritical flow, the shallow water gravity wave speed is greater than the mean-flow speed, and the disturbance height and wind fields are out of phase. The interface height disturbance is negative, and the velocity disturbance is positive over the topographic barrier as shown in Fig. 9.9a. When $\mathrm{Fr} > 1$, the flow is referred to as *supercritical*. In supercritical flow the mean flow exceeds the shallow water gravity wave speed. Gravity waves cannot play a role in establishing the steady-state adjustment between height and velocity disturbances because such waves are swept downstream from the ridge by the mean flow. In this case the fluid thickens and slows as it ascends over the barrier (Fig. 9.9b). It is also clear from (9.34) that for $\mathrm{Fr} \sim 1$ the perturbations are no longer small and the linear solution breaks down.

The nonlinear equations corresponding to (9.32) and (9.33) in the case $\delta\rho = \rho_1$ can be expressed as

$$u\frac{\partial u}{\partial x} + g\frac{\partial h}{\partial x} = 0 \tag{9.35}$$

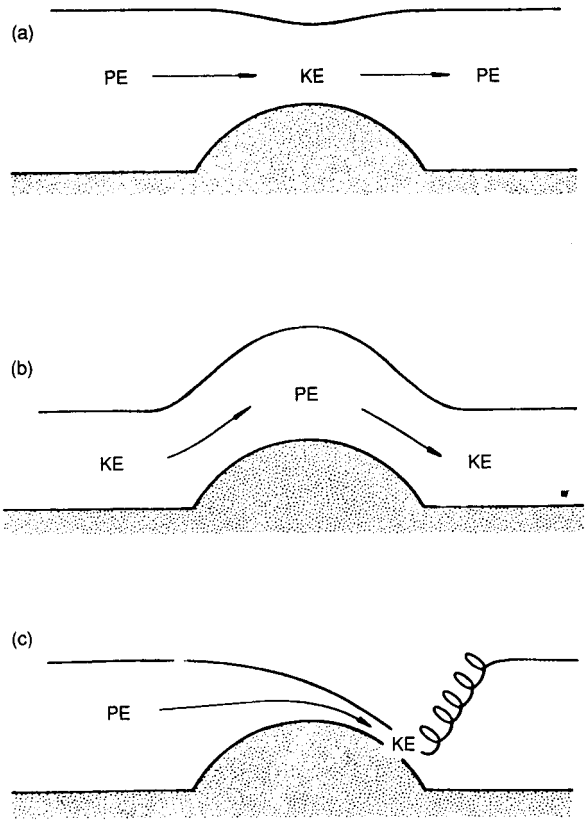


Fig. 9.9 Flow over an obstacle for a barotropic fluid with free surface. (a) Subcritical flow (Fr < 1 everywhere). (b) Supercritical flow (Fr > 1 everywhere). (c) Supercritical flow on lee slope with adjustment to subcritical flow at hydraulic jump near base of obstacle. (After Durran, 1990.)

$$\frac{\partial}{\partial x}\left[u\left(h - h_M\right)\right] = 0 \tag{9.36}$$

Equation (9.35) may be integrated immediately to show that the sum of kinetic and potential energy, $u^2/2 + gh$, is constant following the motion. Thus, energy conservation requires that if u increases h must decrease, and vice versa. In addition, (9.36) shows that the mass flux, $u(h - h_M)$, must also be conserved. The direction of the exchange between kinetic and potential energy in flow over a ridge is determined by the necessity that both (9.35) and (9.36) be satisfied.

Multiplying (9.35) by u and eliminating $\partial h/\partial x$ with the aid of (9.36) gives

$$\left(1 - \mathrm{Fr}^2\right)\frac{\partial u}{\partial x} = \frac{ug}{c^2}\frac{\partial h_M}{\partial x} \tag{9.37}$$

where the shallow water wave speed and the Froude number are now defined using the local thickness and velocity of the fluid:

$$c^2 \equiv g\,(h - h_M)\,; \quad \mathrm{Fr}^2 \equiv u^2/c^2$$

From (9.37) it is clear that the flow will accelerate on the upslope side of the ridge ($\partial u/\partial x > 0$ where $\partial h_M/\partial x > 0$) if the Froude number is less than unity, but will decelerate if the Froude number is greater than unity.

As a subcritical flow ascends the upslope side of a topographic barrier, Fr will tend to increase both from the increase in u and the decrease in c. If $\mathrm{Fr} = 1$ at the crest, then from (9.37) the flow will become supercritical and continue to accelerate as it descends the lee side until it adjusts back to the ambient subcritical conditions in a turbulent hydraulic jump as illustrated in Fig. 9.9c. In this case very high velocities can occur along the lee slope, as potential energy is converted into kinetic energy during the entire period that a fluid column traverses the barrier. Although conditions in the continuously stratified atmosphere are clearly more complex than in the shallow water hydraulic model, numerical simulations have demonstrated that the hydraulic model provides a reasonable conceptual model for the primary processes occurring in downslope windstorms.

9.5 CUMULUS CONVECTION

Mesoscale storms associated with cumulus convection represent a large fraction of all meteorologically important mesoscale circulations. Before considering such systems it is necessary to examine a few of the essential thermodynamic and dynamical aspects of individual cumulus clouds. The subject of cumulus convection is extremely complex to treat theoretically. Much of this difficulty stems from the fact that cumulus clouds have a complex internal structure. They are generally composed of a number of short-lived individual rising towers, which are produced by elements of ascending buoyant air called *thermals*. Rising thermals *entrain* environmental air and thus modify the cloud air through mixing. Thermals are nonhydrostatic, nonsteady, and highly turbulent. The buoyancy of an individual thermal (i.e., the difference between its density and the density of the environment) depends on a number of factors, including the environmental lapse rate, the rate of dilution by entrainment, and drag by the weight of liquid water in cloud droplets. A detailed discussion of the dynamics of thermal convection is beyond the scope of this text. This section utilizes a simple one-dimensional cloud model and focuses primarily on the thermodynamic aspects of moist convection. Convective storm dynamics is considered in Section 9.6.

9.5.1 Equivalent Potential Temperature

We have previously applied the parcel method to discuss the vertical stability of a dry atmosphere. We found that the stability of a dry parcel with respect to a vertical displacement depends on the lapse rate of potential temperature in the environment such that a parcel displacement is stable, provided that $\partial\theta/\partial z > 0$ (i.e., the actual lapse rate is less than the adiabatic lapse rate). The same condition also applies to parcels in a moist atmosphere when the relative humidity is less than 100%. If, however, a parcel of moist air is forced to rise, it will eventually become saturated at a level called the lifting condensation level (LCL). A further forced rise will then cause condensation and latent heat release, and the parcel will then cool at the saturated adiabatic lapse rate. If the environmental lapse rate is greater than the saturated adiabatic lapse rate, and the parcel is forced to continue to rise, it will reach a level at which it becomes buoyant relative to its surroundings. It can then freely accelerate upward. The level at which this occurs is called the level of free convection (LFC).

Discussion of parcel dynamics in a moist atmosphere is facilitated by defining a thermodynamic field called the *equivalent potential temperature*. Equivalent potential temperature, designated by θ_e, is the potential temperature that a parcel of air would have if all its moisture were condensed and the resultant latent heat used to warm the parcel. The temperature of an air parcel can be brought to its equivalent potential value by raising the parcel from its original level until all the water vapor in the parcel has condensed and fallen out and then compressing the parcel adiabatically to a pressure of 1000 hPa. Because the condensed water is assumed to fall out, the temperature increase during the compression will be at the dry adiabatic rate and the parcel will arrive back at its original level with a temperature that is higher than its original temperature. Thus, the process is irreversible. Ascent of this type, in which all condensation products are assumed to fall out, is called *pseudoadiabatic* ascent. (It is not a truly adiabatic process because the liquid water that falls out carries a small amount of heat with it.)

A complete derivation of the mathematical expression relating θ_e to the other variables of state is rather involved and will be relegated to Appendix D. For most purposes, however, it is sufficient to use an approximate expression for θ_e that can be immediately derived from the entropy form of the first law of thermodynamics (2.46). If we let q_s denote the mass of water vapor per unit mass of dry air in a saturated parcel (q_s is called the saturation mixing ratio), then the rate of diabatic heating per unit mass is

$$J = -L_c \frac{Dq_s}{Dt}$$

where L_c is the latent heat of condensation. Thus, from the first law of thermodynamics

$$c_p \frac{D \ln \theta}{Dt} = -\frac{L_c}{T} \frac{Dq_s}{Dt} \tag{9.38}$$

For a saturated parcel undergoing pseudoadiabatic ascent the rate of change in q_s following the motion is much larger than the rate of change in T or L_c. Therefore,

$$d \ln \theta \approx -d \left(L_c q_s \big/ c_p T \right) \tag{9.39}$$

Integrating (9.39) from the initial state (θ, q_s, T) to a state where $q_s \approx 0$ we obtain

$$\ln \left(\theta \big/ \theta_e \right) \approx -L_c q_s \big/ c_p T$$

where θ_e, the potential temperature in the final state, is approximately the equivalent potential temperature defined earlier. Thus, θ_e, for a saturated parcel is given by

$$\theta_e \approx \theta \exp \left(L_c q_s \big/ c_p T \right) \tag{9.40}$$

The expression in (9.40) may also be used to compute θ_e for an unsaturated parcel provided that the temperature used in the formula is the temperature that the parcel would have if expanded adiabatically to saturation (i.e., T_{LCL}) and the saturation mixing ratio is replaced by the *actual* mixing ratio of the initial state. Thus, equivalent potential temperature is conserved for a parcel during both dry adiabatic and pseudoadiabatic displacements.

An alternative to θ_e, which is sometimes used in studies of convection, is the *moist static energy*, defined as $h \equiv s + L_c q$, where $s \equiv c_p T + gz$ is the *dry static energy*. It can be shown (Problem 9.7) that

$$c_p T d \ln \theta_e \approx dh \tag{9.41}$$

Hence, moist static energy is approximately conserved when θ_e is conserved.

9.5.2 The Pseudoadiabatic Lapse Rate

The first law of thermodynamics (9.38) can be used to derive a formula for the rate of change of temperature with respect to height for a saturated parcel undergoing pseudoadiabatic ascent. Using the definition of θ (2.44), we can rewrite (9.38) for vertical ascent as

$$\frac{d \ln T}{dz} - \frac{R}{c_p} \frac{d \ln p}{dz} = -\frac{L_c}{c_p T} \frac{dq_s}{dz}$$

which upon noting that $q_s \equiv q_s\,(T, p)$ and applying the hydrostatic equation and equation of state can be expressed as

$$\frac{dT}{dz} + \frac{g}{c_p} = -\frac{L_c}{c_p} \left[\left(\frac{\partial q_s}{\partial T} \right)_p \frac{dT}{dz} - \left(\frac{\partial q_s}{\partial p} \right)_T \rho g \right]$$

Thus, as is shown in detail in Appendix D, following the ascending saturated parcel

$$\Gamma_s \equiv -\frac{dT}{dz} = \Gamma_d \frac{\left[1 + L_c q_s \big/ (RT)\right]}{\left[1 + \varepsilon L_c^2 q_s \big/ \left(c_p R T^2\right)\right]} \tag{9.42}$$

where $\varepsilon = 0.622$ is the ratio of the molecular weight of water to that of dry air, $\Gamma_d \equiv g/c_p$ is the dry adiabatic lapse rate, and Γ_s is the *pseudoadiabatic lapse rate,* which is always less than Γ_d. Observed values of Γ_s range from $\sim$4 K km^{-1} in warm humid air masses in the lower troposphere to $\sim$ 6–7 K km^{-1} in the midtroposphere.

9.5.3 Conditional Instability

Section 2.7.3 showed that for dry adiabatic motions the atmosphere is statically stable provided that the lapse rate is less than the dry adiabatic lapse rate (i.e., the potential temperature increases with height). If the lapse rate Γ lies between dry adiabatic and pseudoadiabatic values ($\Gamma_s < \Gamma < \Gamma_d$), the atmosphere is stably stratified with respect to dry adiabatic displacements but unstable with respect to pseudoadiabatic displacements. Such a situation is referred to as *conditional instability* (i.e., the instability is conditional to saturation of the air parcel).

The conditional stability criterion can also be expressed in terms of the gradient of a field variable θ_e^*, defined as the equivalent potential temperature of a hypothetically saturated atmosphere that has the thermal structure of the actual atmosphere.[2] Thus,

$$d \ln \theta_e^* = d \ln \theta + d\left(L_c q_s \big/ c_p T\right) \tag{9.43}$$

where T is the actual temperature, not the temperature after adiabatic expansion to saturation as in (9.40). To derive an expression for conditional instability we consider the motion of a saturated parcel in an environment in which the potential temperature is θ_0 at the level z_0. At the level $z_0 - \delta z$ the undisturbed environmental air thus has potential temperature

$$\theta_0 - \left(\partial\theta \big/ \partial z\right) \delta z$$

Suppose a saturated parcel that has the environmental potential temperature at $z_0 - \delta z$ is raised to the level z_0. When it arrives at z_0 the parcel will have the potential temperature

$$\theta_1 = \left(\theta_0 - \frac{\partial\theta}{\partial z}\delta z\right) + \delta\theta$$

[2] Note that θ_e^* is *not* the same as θ_e except in a saturated atmosphere.

where $\delta\theta$ is the change in parcel potential temperature due to condensation during ascent through vertical distance δz. Assuming a pseudoadiabatic ascent, we see from (9.39) that

$$\frac{\delta\theta}{\theta} \approx -\delta\left(\frac{L_c q_s}{c_p T}\right) \approx -\frac{\partial}{\partial z}\left(\frac{L_c q_s}{c_p T}\right)\delta z$$

so that the buoyancy of the parcel when it arrives at z_0 is proportional to

$$\frac{(\theta_1 - \theta_0)}{\theta_0} \approx -\left[\frac{1}{\theta}\frac{\partial\theta}{\partial z} + \frac{\partial}{\partial z}\left(\frac{L_c q_s}{c_p T}\right)\right]\delta z \approx -\frac{\partial \ln \theta_e^*}{\partial z}\delta z$$

where the last expression arises from the application of (9.43).

The saturated parcel will be warmer than its environment at z_0 provided that $\theta_1 > \theta_0$. Thus the conditional stability criterion for a saturated parcel is

$$\frac{\partial\theta_e^*}{\partial z}\begin{cases} < 0 & \text{conditionally unstable} \\ = 0 & \text{saturated neutral} \\ > 0 & \text{conditionally stable} \end{cases} \tag{9.44}$$

In Fig. 9.10 the vertical profiles of θ,θ_e, and θ_e^* for a typical sounding in the vicinity of an extratropical thunderstorm are shown. It is obvious from Fig. 9.10 that the sounding is conditionally unstable in the lower troposphere. However, this observed profile does not imply that convective overturning will occur spontaneously. The release of conditional instability requires not only that $\partial\theta_e^*/\partial z < 0$, but also parcel saturation at the environmental temperature of the level where the convection begins (i.e., the parcel must reach the LFC). The mean relative humidity in the troposphere is well below 100%, even in the boundary layer. Thus, low-level convergence with resultant forced layer ascent or vigorous vertical turbulent mixing in the boundary layer is required to produce saturation.

The amount of ascent necessary to raise a parcel to its LFC can be estimated simply from Fig. 9.10. A parcel rising pseudoadiabatically from a level $z_0 - \delta z$ will conserve the value of θ_e characteristic of the environment at $z_0 - \delta z$. However, the buoyancy of a parcel depends only on the difference in density between the parcel and its environment. Thus, in order to compute the buoyancy of the parcel at z_0, it is not correct simply to compare θ_e of the environment at z_0 to $\theta_e(z_0 - \delta z)$ because if the environment is unsaturated, the difference in θ_e between the parcel and the environment may be due primarily to the difference in mixing ratios, not to any temperature (density) difference. To estimate the buoyancy of the parcel $\theta_e(z_0 - \delta z)$ should instead be compared to $\theta_e^*(z_0)$, which is the equivalent potential temperature that the environment at z_0 would have if it were isothermally brought

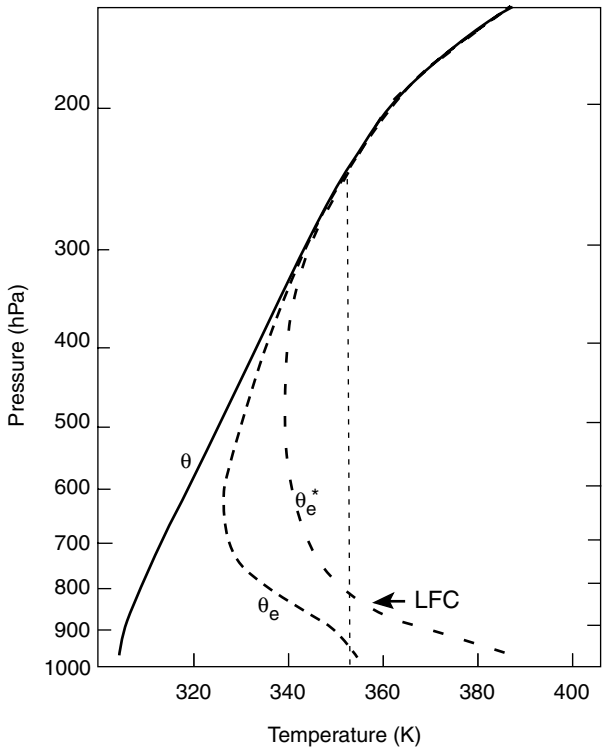


Fig. 9.10 Schematic sounding for a conditionally unstable environment characteristic of Midwestern North America thunderstorm conditions showing the vertical profiles of potential temperature, θ, equivalent potential temperature, θ_e, and the equivalent temperature, θ_e^*, of a hypothetically saturated atmosphere with the same temperature profile. Dotted line shows θ_e for a nonentraining parcel raised from the surface. Arrow denotes the LFC for the parcel.

to saturation. The parcel will thus become buoyant when raised to the level z_0 if $\theta_e(z_0 - \delta z) > \theta_e^*(z_0)$, for then the parcel temperature will exceed the temperature of the environment at z_0. From Fig. 9.10 we see that θ_e for a parcel raised from about 960 hPa will intersect the θ_e^* curve near 850 hPa, whereas a parcel raised from any level much above 850 hPa will not intersect θ_e^* no matter how far it is forced to ascend. It is for this reason that low-level convergence is usually required to initiate convective overturning over the oceans. Only air near the surface has a sufficiently high value of θ_e to become buoyant when it is forcibly raised. Convection over continental regions, however, can be initiated without significant boundary layer convergence, as strong surface heating can produce positive parcel buoyancy all the way to the surface. Sustained deep convection, however, requires mean low-level moisture convergence.

9.5.4 Convective Available Potential Energy (CAPE)

Development of convective storms depends on the presence of environmental conditions favorable for the occurrence of deep convection. Several indices have been developed to measure the susceptibility of a given temperature and moisture profile to the occurrence of deep convection. A particularly useful measure is the *convective available potential energy*. CAPE provides a measure of the maximum possible kinetic energy that a statically unstable parcel can acquire (neglecting effects of water vapor and condensed water on the buoyancy), assuming that the parcel ascends without mixing with the environment and adjusts instantaneously to the local environmental pressure.

The momentum equation for such a parcel is (2.51), which can be rewritten following the vertical motion of the parcel as

$$\frac{Dw}{Dt} = \frac{Dz}{Dt}\frac{Dw}{Dz} = w\frac{Dw}{Dz} = b' \tag{9.45}$$

where $b'(z)$ is again the buoyancy given by

$$b' = g\frac{\left(\rho_{env} - \rho_{parcel}\right)}{\rho_{parcel}} = g\frac{\left(T_{parcel} - T_{env}\right)}{T_{env}} \tag{9.46}$$

and T_{env} designates the temperature of the environment. If (9.45) is integrated vertically from the level of free convection, z_{LFC}, to the level of neutral buoyancy, z_{LNB}, following the motion of the parcel the result is

$$\frac{w_{\max}^2}{2} = \int_{z_{LFC}}^{z_{LNB}} g\left(\frac{T_{parcel} - T_{env}}{T_{env}}\right) dz \equiv B \tag{9.47}$$

Here, B is the maximum kinetic energy per unit mass that a buoyant parcel could obtain by ascending from a state of rest at the level of free convection to the level of neutral buoyancy near the tropopause (see Fig. 9.10). This is an overestimate of the actual kinetic energy realized by a nonentraining parcel, as the negative buoyancy contribution of liquid water reduces the effective buoyancy, especially in the tropics.

In a typical tropical oceanic sounding, parcel temperature excesses of 1–2 K may occur over a depth of 10–12 km. A typical value of CAPE is then $B \approx 500\ \mathrm{m^2\ s^{-2}}$. In severe storm conditions in the Midwest of North America, however, parcel temperature excesses can be 7–10 K (see Fig. 9.10) and $B \approx 2000$–$3000\ \mathrm{m^2\ s^{-2}}$. Observed updrafts in the latter case (up to 50 m s^{-1}) are much stronger than in the former case (5–10 m s^{-1}). The small value of CAPE in the mean tropical environment is the major reason that updraft velocities in tropical

cumulonimbus are observed to be much smaller than those in midlatitude thunderstorms.

9.5.5 Entrainment

In the previous subsection it was assumed that convective cells rise without mixing with environmental air so that they maintain constant θ_e during their rise. In reality, however, rising saturated air parcels tend to be diluted by entraining, or mixing in, some of the relatively dry environmental air. If the air in the environment is unsaturated, some of the liquid water in the rising parcel must be evaporated to maintain saturation in the convective cell as air from the environment is entrained. The evaporative cooling caused by entrainment will reduce the buoyancy of the convective parcel (i.e., lower its θ_e). Thus, the equivalent potential temperature in an entraining convection cell will decrease with height rather than remaining constant. Similar considerations hold for any other conservable variable[3] in which environmental values differ from cloud values; entrainment will modify the in-cloud vertical profiles.

Denoting the amount of an arbitrary conservable variable per unit mass of air by A, the vertical dependence of A in an entraining convective cell can be estimated by assuming that to a first approximation the cell can be modeled as a steady-state jet as shown in Fig. 9.11. Thus, in a time increment δt a mass m of saturated cloud air with an amount of the arbitrary variable given by $m A_{cld}$ mixes with a mass

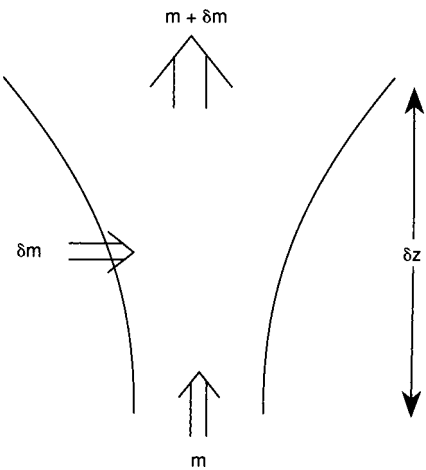


Fig. 9.11 An entraining jet model of cumulus convection. See text for explanation.

[3] A conservable variable is one that is conserved following the motion in the absence of sources and sinks (e.g., a chemical trace constituent).

δm of entrained environmental air, which has an amount of the arbitrary variable given by $\delta m A_{env}$. The change in the value of A within the cloud, δA_{cld}, is then given by the mass balance relationship

$$(m + \delta m)(A_{cld} + \delta A_{cld}) = m A_{cld} + \delta m A_{env} + \left(DA_{cld} / Dt\right)_S m \delta t \tag{9.48}$$

where $\left(DA_{cld} / Dt\right)_S$ designates the rate of change of A_{cld} due to sources and sinks unrelated to entrainment. Dividing through by δt in (9.48), neglecting the second-order term, and rearranging yields

$$\frac{\delta A_{cld}}{\delta t} = \left(\frac{DA_{cld}}{Dt}\right)_S - \frac{1}{m}\frac{\delta m}{\delta t}(A_{cld} - A_{env}) \tag{9.49}$$

Noting that in the time increment δt an ascending parcel rises a distance $\delta z = w\delta t$, where w is the speed of ascent of the parcel, we can eliminate δt in (9.49) to obtain an equation for the vertical dependence of A_{cld} in a continuously entraining convective cell:

$$w\frac{dA_{cld}}{dz} = \left(\frac{DA_{cld}}{Dt}\right)_S - w\lambda(A_{cld} - A_{env}) \tag{9.50}$$

where we have defined the *entrainment rate*; $\lambda \equiv d \ln m / dz$.

Letting $A_{cld} = \ln \theta_e$ and noting that θ_e is conserved in the absence of entrainment, (9.50) yields

$$\begin{aligned}\left(\frac{d \ln \theta_e}{dz}\right)_{cld} &= -\lambda\left[(\ln \theta_e)_{cld} - (\ln \theta_e)_{env}\right] \\ &\approx -\lambda\left[\frac{L_c}{c_p T}(q_s - q_{env}) + \ln\left(\frac{T_{cld}}{T_{env}}\right)\right]\end{aligned} \tag{9.51}$$

where (9.40) is used to obtain the latter form of the right-hand side. Thus an entraining convective cell is less buoyant than a nonentraining cell. Letting $A_{cld} = w$ in (9.50), applying (9.45), and neglecting the pressure contribution to buoyancy, we find that the height dependence of kinetic energy per unit mass is given by

$$\frac{d}{dz}\left(\frac{w^2}{2}\right) = g\left(\frac{T_{cld} - T_{env}}{T_{env}}\right) - \lambda w^2 \tag{9.52}$$

An entraining cell will undergo less acceleration than a nonentraining cell not only because the buoyancy is reduced, but also because of the drag exerted by mass entrainment.

Equations (9.51) and (9.52), together with suitable relations for the cloud moisture variables, can be used to determine the vertical profile of cloud variables. Such

one-dimensional cloud models have been very popular in the past. Unfortunately, observed cloud properties, such as maximum cloud depth and cloud water concentration, cannot simultaneously be satisfactorily predicted by this type of model. In reality, pressure perturbations generated by the convective cells are important in the momentum budget, and entrainment does not occur instantaneously, but in a sporadic manner that allows some cells to rise through most of the troposphere with very little dilution. Although more sophisticated one-dimensional models can partly overcome these deficiencies, some of the most important aspects of thunderstorm dynamics (e.g., the influence of vertical shear of the environmental wind) can only be included in multidimensional models.

9.6 CONVECTIVE STORMS

Convective storms can take a large variety of forms. These range in scale from isolated thunderstorms involving a single convective cloud (or *cell*) to mesoscale convective complexes consisting of ensembles of multicelled thunderstorms. Here we distinguish three primary types: the single cell, the multicell, and the supercell storm. As shown in the previous section, convective available potential energy measures whether thermodynamical conditions are favorable for the development of cumulus convection. CAPE, therefore, provides a guide to the strength of convection. It does not, however, provide any notion of the most likely type of mesoscale organization. It turns out, as suggested above, that storm type also depends on the vertical shear of the lower tropospheric environment.

When vertical shear is weak ($< 10\ \mathrm{m\,s^{-1}}$ below 4 km), single cell storms occur. These tend to be short lived (~ 30 min) and move with the mean wind in the lowest 8 km. When there is moderate vertical shear (~ 10–$20\ \mathrm{m\,s^{-1}}$ below 4 km), multicell storms arise in which individual cells have lifetimes of ~ 30 min, but the storm lifetime may be many hours. In multicell storms the downdrafts induced by the evaporation of precipitation form a dome of cold outflowing air at the surface. New cells tend to develop along the *gust front* where the cold outflow lifts conditionally unstable surface air. When vertical shear is large ($> 20\mathrm{m\,s^{-1}}$ below 4 km), the strong tilting of convective cells tends to delay storm development even in a thermodynamically favorable environment so that an hour or more may be required for the initial cell to completely develop. This development may be followed by a split into two storms, which move to the left and right of the mean wind. Usually the left-moving storm dies rapidly while the right-moving storm slowly evolves into a rotating circulation with a single updraft core and trailing downdrafts, as discussed in the next section. Such supercell storms often produce heavy rain, hail, and damaging tornadoes. Ensembles of multicell or supercell storms are often organized along lines referred to as *squall lines*, which may move in a different direction than the individual thunderstorms.

9.6.1 Development of Rotation in Supercell Thunderstorms

The supercell thunderstorm is of particular dynamical interest because of its tendency to develop a rotating mesocyclone from an initially nonrotating environment. The dominance of cyclonic rotation in such systems might suggest that the Coriolis force plays a role in supercell dynamics. However, it can be shown readily that the rotation of the earth is not relevant to development of rotation in supercell storms.

Although a quantitative treatment of supercell dynamics requires that the density stratification of the atmosphere be taken into account, for purposes of understanding the processes that lead to development of rotation in such systems and to the dominance of the right-moving cell, it is sufficient to use the Boussinesq approximation. The Euler momentum equation and continuity equation may then be expressed as

$$\frac{D\mathbf{U}}{Dt} = \frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \bullet \boldsymbol{\nabla})\,\mathbf{U} = -\frac{1}{\rho_0}\boldsymbol{\nabla} p + b\mathbf{k}$$

$$\boldsymbol{\nabla} \bullet \mathbf{U} = 0$$

Here, $\mathbf{U} \equiv \mathbf{V} + \mathbf{k}w$ is the three-dimensional velocity, $\boldsymbol{\nabla}$ is the three-dimensional del operator, ρ_0 is the constant basic state density, p is the deviation of pressure from its horizontal mean, and $b \equiv -g\rho'/\rho_0$ is the total buoyancy.

It is convenient to rewrite the momentum equation using the vector identity

$$(\mathbf{U} \bullet \boldsymbol{\nabla})\,\mathbf{U} = \boldsymbol{\nabla}\left(\frac{\mathbf{U} \bullet \mathbf{U}}{2}\right) - \mathbf{U} \times (\boldsymbol{\nabla} \times \mathbf{U})$$

to obtain

$$\frac{\partial \mathbf{U}}{\partial t} = -\boldsymbol{\nabla}\left(\frac{p}{\rho_0} + \frac{\mathbf{U} \bullet \mathbf{U}}{2}\right) + \mathbf{U} \times \boldsymbol{\omega} + b\mathbf{k} \tag{9.53}$$

Taking $\boldsymbol{\nabla} \times$ (9.53) and recalling that the curl of the gradient vanishes, we obtain the three-dimensional vorticity equation

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{U} \times \boldsymbol{\omega}) + \boldsymbol{\nabla} \times (b\mathbf{k}) \tag{9.54}$$

Letting $\zeta = \mathbf{k} \bullet \boldsymbol{\omega}$ be the vertical component of vorticity and taking $\mathbf{k} \bullet$ (9.54), we obtain an equation for the tendency of ζ in a nonrotating reference frame:

$$\frac{\partial \zeta}{\partial t} = \mathbf{k} \bullet \nabla \times (\mathbf{U} \times \boldsymbol{\omega}) \tag{9.55}$$

Note that buoyancy only affects the horizontal vorticity components.

We now consider a flow consisting of a single convective updraft embedded in a basic state westerly flow that depends on z alone. Linearizing about this basic state by letting

$$\boldsymbol{\omega} = \mathbf{j}d\overline{u}\,/dz + \boldsymbol{\omega}'(x, y, z, t)\,, \quad \mathbf{U} = \mathbf{i}\overline{u} + \mathbf{U}'(x, y, z, t)$$

and noting that the linearized form of the right-hand side in (9.55) becomes

$$\mathbf{k} \bullet \boldsymbol{\nabla} \times (\mathbf{U} \times \boldsymbol{\omega}) = -\mathbf{k} \bullet \boldsymbol{\nabla} \times \left(\mathbf{i}w'd\overline{u}\,/dz + \mathbf{j}\overline{u}\zeta'\right)$$

we find that the linearized vorticity tendency is

$$\frac{\partial \zeta'}{\partial t} = -\overline{u}\frac{\partial \zeta'}{\partial x} + \frac{\partial w'}{\partial y}\frac{d\overline{u}}{dz} \tag{9.56}$$

The first term on the right in (9.56) is just the advection by the basic state flow. The second term represents tilting of horizontal shear vorticity into the vertical by differential vertical motion.

Because $d\overline{u}/dz$ is positive, the vorticity tendency due to this tilting will be positive to the south of the updraft core and negative to the north of the updraft core. As a result, a counterrotating vortex pair is established with cyclonic rotation to the south and anticyclonic rotation to the north of the initial updraft, as shown in Fig. 9.12a. Eventually, the development of negative buoyancy due to precipitation loading generates an upper level downdraft at the position of the initial updraft and the storm splits as shown in Fig. 9.12b. New updraft cores are established centered on the counterrotating vortex pair.

In order to understand the generation of updrafts in the vortices on the flanks of the storm, we examined the perturbation pressure field. A diagnostic equation for the disturbance pressure is obtained by taking $\boldsymbol{\nabla} \bullet$ (9.53) to yield

$$\nabla^2\left(\frac{p}{\rho_0}\right) = -\nabla^2\left(\frac{\mathbf{U} \bullet \mathbf{U}}{2}\right) + \boldsymbol{\nabla} \bullet (\mathbf{U} \times \boldsymbol{\omega}) + \frac{\partial b}{\partial z} \tag{9.57}$$

The first two terms on the right in (9.57) represent dynamical forcing, whereas the last term represents buoyancy forcing. Observations and numerical models suggest that the buoyancy forcing in (9.57) produces pressure perturbations that tend partly to compensate the buoyancy force in the vertical momentum equation. Dynamically forced pressure perturbations, however, may generate substantial vertical accelerations.

In order to compute the dynamical contribution to the disturbance pressure gradient force in either the right or left side vortex, we use cylindrical coordinates (r, λ, z) centered on the axis of rotation of either vortex and assume that to a first

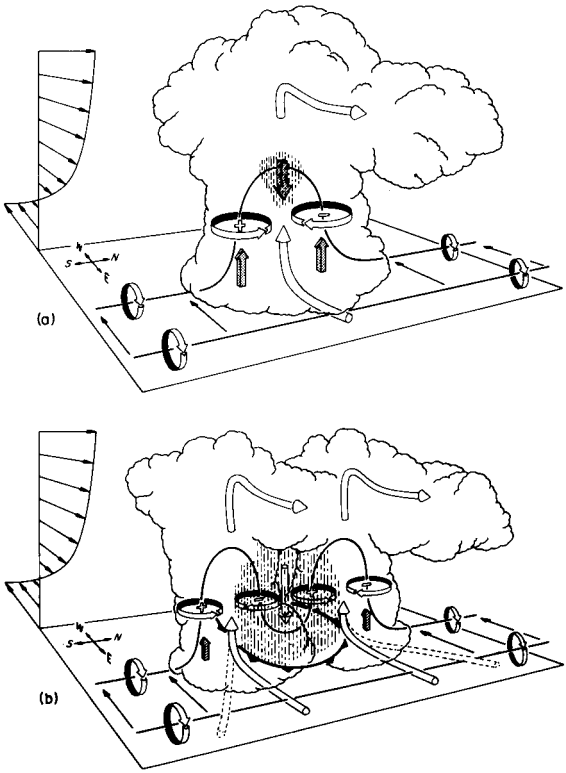


Fig. 9.12 Development of rotation and splitting in a supercell storm with westerly mean wind shear (shown by storm relative wind arrows in the upper left corner of each panel). Cylindrical arrows show the direction of cloud relative air flow. Heavy solid lines show vortex lines with a sense of rotation shown by circular arrows. Plus and minus signs indicate cyclonic and anticyclonic rotation caused by vortex tube tilting. Shaded arrows represent updraft and downdraft growth. Vertical dashed lines denote regions of precipitation. (a) In the initial stage, the environmental shear vorticity is tilted and stretched into the vertical as it is swept into the updraft. (b) In the splitting stage, downdraft forms between the new updraft cells. Barbed line at surface indicates downdraft outflow at surface. (After Klemp, 1987.)

approximation the azimuthal velocity v_λ (positive for cyclonic flow) is independent of λ. In this system the storm relative horizontal motion and vertical component of vorticity are given approximately by

$$\mathbf{U}' \approx \mathbf{j}_\lambda v_\lambda, \quad \mathbf{k} \bullet \boldsymbol{\omega}' = \zeta' \approx r^{-1} \partial (r v_\lambda) \big/ \partial r$$

where $\mathbf{j}_\lambda$ is the unit vector in the azimuthal direction (positive counterclockwise) and r is the distance from the axis of the vortex. Letting $\mathbf{i}_\lambda$ be the unit vector in the

radial direction, we then have

$$\mathbf{U} \times \boldsymbol{\omega} \approx \mathbf{i}_\lambda \frac{v_\lambda}{r} \frac{\partial}{\partial r} (r v_\lambda)$$

Assuming that the vertical scale is much larger than the radial scale, the Laplacian in cylindrical coordinates can be approximated by

$$\nabla^2 \approx \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right)$$

Thus, from (9.57) the dynamical component of the pressure perturbation in the vortices (designated p_{dyn}) can be expressed as

$$\frac{1}{r} \frac{\partial}{\partial r} \left(\frac{r}{\rho_0} \frac{\partial p_{dyn}}{\partial r} \right) \approx -\frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial \left(v_\lambda^2 / 2 \right)}{\partial r} \right] + \frac{1}{r} \frac{\partial}{\partial r} \left[v_\lambda \frac{\partial (r v_\lambda)}{\partial r} \right] = \frac{1}{r} \frac{\partial v_\lambda^2}{\partial r} \tag{9.58}$$

Integrating (9.58) with respect to r, we obtain the equation of cyclostrophic balance (see Section 3.2.4):

$$\rho_0^{-1} \partial p_{dyn} \big/ \partial r \; \partial r \approx v_\lambda^2 / r \tag{9.59}$$

Hence, there is a pressure minimum at the vortex center irrespective of whether the rotation is cyclonic or anticyclonic. The strong midtropospheric rotation induced by vortex tube twisting and stretching creates a "centrifugal pump" effect, which causes a negative dynamical pressure perturbation centered in the vortices in the midtroposphere. This in turn produces an upward directed dynamical contribution to the vertical component of the pressure gradient force, and thus provides an upward acceleration, which produces updrafts in the cores of the counterrotating vortices as depicted in Fig. 9.12. These updrafts are separated by a downdraft that leads to a splitting of the storm and the development of two new centers of convection that move to the right and left of the original storm (Fig. 9.12b).

As discussed in this section, the tilting and stretching of horizontal vorticity associated with the vertical shear of the basic state wind can account for the development of mesoscale rotating supercells. This process does not, however, appear to be able to produce the large vorticities observed in the tornadoes that often accompany supercell thunderstorms. Numerical simulations suggest that these tend to involve tilting and stretching of especially strong horizontal vorticity produced by horizontal gradients in buoyancy that occur near the surface along the *gust front* where negatively buoyant outdrafts produced by convective downdrafts meet moist warm boundary layer air.

9.6.2 The Right-Moving Storm

When the environmental wind shear is unidirectional, as in the case discussed above (see also Fig. 9.13a), the anticyclonic (left moving) and cyclonic (right

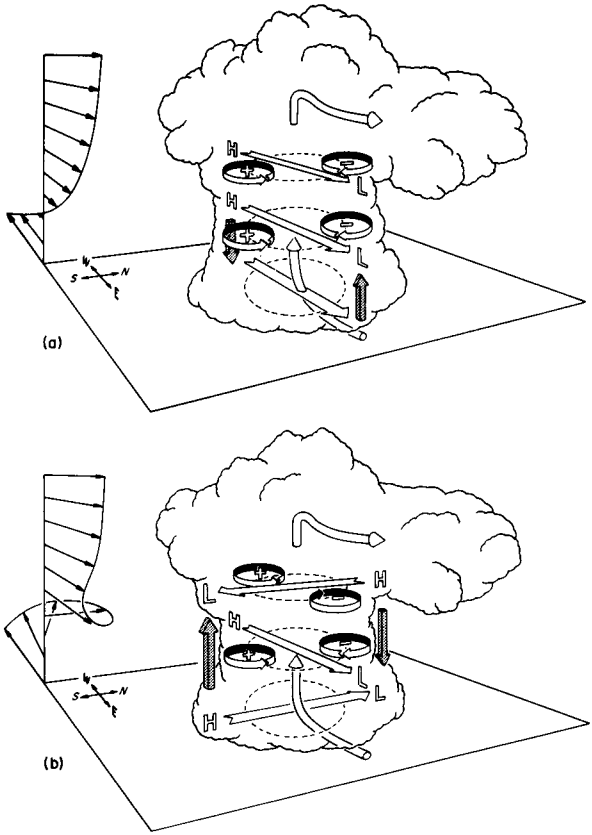


Fig. 9.13 Pressure and vertical vorticity perturbations produced by interaction of the updraft with environmental wind shear in a supercell storm. (a) Wind shear does not change direction with height. (b) Wind shear turns clockwise with height. Broad open arrows designate the shear vectors. H and L designate high and low dynamical pressure perturbations, respectively. Shaded arrows show resulting disturbance vertical pressure gradients. (After Klemp, 1987.)

moving) updraft cores are equally favored. In most severe storms in the central United States, however, the mean flow turns anticyclonically with height; this directional shear in the environment favors the right-moving storm center, while inhibiting the left-moving center. Thus, right-moving storms are observed far more than left-moving storms.

The dominance of the right-moving storm can be understood qualitatively by again considering the dynamical pressure perturbations. We define the basic state wind shear vector $\overline{\mathbf{S}} \equiv \partial \overline{\mathbf{V}} / \partial z$, which is assumed to turn clockwise with height.

Noting that the basic state horizontal vorticity in this case is

$$\overline{\boldsymbol{\omega}} = \mathbf{k} \times \overline{\mathbf{S}} = -\mathbf{i}\partial\overline{v}\,/\partial z + \mathbf{j}\partial\overline{u}\,/\partial z$$

we see that there is a contribution to the dynamic pressure in (9.57) of the form

$$\boldsymbol{\nabla} \bullet \left(\mathbf{U}' \times \overline{\boldsymbol{\omega}}\right) \approx -\boldsymbol{\nabla} \bullet \left(w'\overline{\mathbf{S}}\right)$$

From (9.57) the sign of the pressure perturbation due to this effect may be determined by noting that

$$\nabla^2 p_{dyn} \sim -p_{dyn} \sim -\frac{\partial}{\partial x}\left(w'S_x\right) - \frac{\partial}{\partial y}\left(w'S_y\right) \tag{9.60}$$

which shows that there is a positive dynamical pressure perturbation upshear of the cell and a negative perturbation downshear (analogous to the positive pressure perturbation upwind and negative perturbation downwind of an obstacle). The resulting pattern of dynamical pressure perturbations is shown in Fig. 9.13. In the case of unidirectional shear (Fig. 9.13a), the induced pressure pattern favors updraft growth on the leading edge of the storm. However, when the shear vector rotates clockwise with height as in Fig. 9.13b, (9.60) shows that a dynamical pressure disturbance pattern is induced in which there is an upward directed vertical pressure gradient force on the flank of the cyclonically rotating cell and a downward directed pressure gradient force on the flank of the anticyclonic cell. Thus, in the presence of clockwise rotation of the environmental shear stronger updrafts are favored in the right-moving cyclonic vortex to the south of the initial updraft.

9.7 HURRICANES

Hurricanes, which are also referred to as tropical cyclones and typhoons in some parts of the world, are intense vortical storms that develop over the tropical oceans in regions of very warm surface water. Although the winds and convective clouds observed in hurricanes are not truly axisymmetric about the vortex center, the fundamental aspects of hurricane dynamics can be modeled by idealizing the hurricane as an axisymmetric vortex. Typically hurricanes have radial scales of several hundred kilometers, similar to those of some midlatitude synoptic systems. However, the horizontal scale of the region of intense convection and strong winds in a hurricane is typically only about 100 km in radius. Thus, it is reasonable to classify hurricanes as mesoscale systems.

Unlike the rotating convective storms treated in the previous section, the hurricane vortex cannot be understood without including the rotation of the earth in the vorticity balance. The rapid rotation observed in hurricanes is produced by concentration of the vertical component of absolute vorticity by vortex stretching, not

by tilting horizontal vorticity into the vertical. Maximum tangential wind speeds in these storms range typically from 50 to 100 m s^{-1}. For such high velocities and relatively small scales, the centrifugal force term cannot be neglected compared to the Coriolis force. Thus, to a first approximation, the azimuthal velocity in a steady-state hurricane is in gradient wind balance with the radial pressure gradient force. Hydrostatic balance also holds on the hurricane scale, which implies that the vertical shear of the azimuthal velocity is a function of the radial temperature gradient.

9.7.1 Dynamics of Mature Hurricanes

An expression for the thermal wind in an axisymmetric hurricane can be easily derived starting from the gradient wind balance in cylindrical coordinates (see Appendix C). This can be written as

$$\frac{v_\lambda^2}{r} + f v_\lambda = \frac{\partial \Phi}{\partial r} \tag{9.61}$$

where r is the radial distance from the axis of the storm (positive outward) and v_λ is the tangential velocity (positive for anticlockwise flow). Alternatively, it is sometimes useful to use the absolute angular momentum $M_\lambda \equiv v_\lambda r + f r^2 / 2$ in place of v_λ because above the boundary layer M_λ is nearly conserved following the motion. (It thus plays the same role in cylindrically symmetric flow as the absolute momentum defined in Section 9.3 plays in linearly symmetric flow.) In terms of M_λ the gradient wind balance can be expressed as (see Problem 9.9)

$$\frac{M_\lambda^2}{r^3} - \frac{f^2 r}{4} = \frac{\partial \Phi}{\partial r} \tag{9.62}$$

We can eliminate Φ in (9.62) with the aid of the hydrostatic equation in log-pressure coordinates introduced in Section 8.4.1:

$$\partial \Phi / \partial z^* = RT / H$$

to obtain a relationship between the radial temperature gradient and the vertical shear of the absolute angular momentum:

$$\frac{1}{r^3} \frac{\partial M_\lambda^2}{\partial z^*} = \frac{R}{H} \frac{\partial T}{\partial r} \tag{9.63}$$

The cyclonic flow in a hurricane is observed to be a maximum near the top of the boundary layer. Above the boundary layer $\partial M_\lambda / \partial z^* < 0$, which by (9.63) implies that $\partial T / \partial r < 0$. Thus, a temperature maximum must occur at the center of the storm. This is consistent with the observation that hurricanes are *warm core* systems.

If we assume that for a typical hurricane the vertical scale is the scale height H, the azimuthal velocity scale is $U \sim 50\ \mathrm{m\ s^{-1}}$, the horizontal scale is $L \sim 100$ km, and $f \sim 5 \times 10^{-5}\ \mathrm{s^{-1}}$ (corresponding to a latitude of about 20°), we find from (9.63) that the radial temperature fluctuation must have a magnitude

$$\delta T \sim \left(UL/R\right)\left(f + 2U/L\right) \sim 10°\mathrm{C}$$

This strong radial variation in temperature is one of the most important thermodynamical characteristics of hurricanes.

The kinetic energy of hurricanes is maintained in the presence of boundary layer dissipation by the conversion of latent heat energy acquired from the underlying ocean. This potential energy conversion is carried out by a transverse secondary circulation associated with the hurricane, as shown schematically in Fig. 9.14. This circulation consists of boundary layer inflow into a region of enhanced convection surrounding the storm center that is referred to as the *eyewall*, ascent within convective cloud towers that tend to be concentrated in the narrow outward-sloping eyewall, radial outflow in a thin layer near the tropopause, and gentle subsidence at a large radius. Observations show that evaporation of water from the sea surface into the inward flowing air in the boundary layer causes a large increase in θ_e as the air approaches the eyewall region. Within the eyewall the θ_e and M_λ surfaces coincide so that parcel ascent in the eyewall (along the path labeled 1 in Fig. 9.14) is neutral with respect to conditional symmetric instability, and thus does not require external forcing. The eyewall surrounds a central eye of radius 5–50 km that is often calm and nearly cloud free.

The energetics of the steady-state hurricane can be viewed as an example of a Carnot cycle heat engine in which heat is absorbed (in the form of water vapor) from the ocean at temperature T_s and expelled by radiative cooling to space at

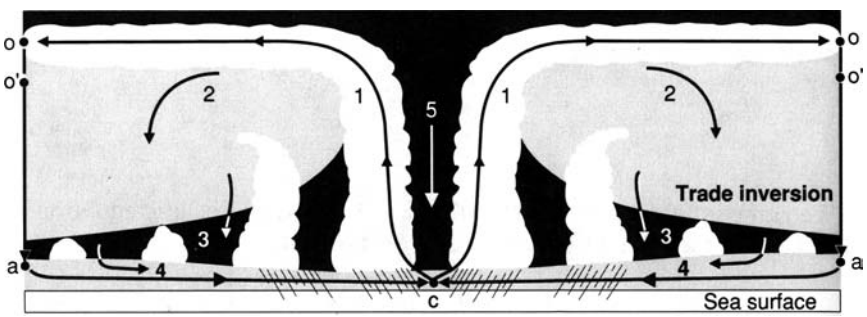


Fig. 9.14 Schematic cross section of secondary meridional circulation in a mature hurricane. Air spirals in toward the eye (region 5) in the boundary layer (region 4), ascends along constant M surfaces in the eyewall cloud (region 1), and slowly subsides and dries in regions 2 and 3. (After Emanuel, 1988.)

temperature T_0 at the top of the storm. Because $T_s \approx 300$ K and $T_0 \approx 200$ K, the efficiency of the heat engine is very high.

9.7.2 Hurricane Development

The origin of tropical cyclones is still a matter of uncertainty. It is not clear under what conditions a weak tropical disturbance can be transformed into a hurricane. Although there are many tropical disturbances each year, only rarely does one develop into a hurricane. Thus, the development of a hurricane must require rather special conditions. Many theoretical investigations of this problem have assumed the initial existence of a small-amplitude cylindrically symmetric disturbance and examined the conditions under which unstable amplification of the disturbance can occur. As shown in Chapter 8, this sort of linear stability theory is quite successful in accounting for the development of extratropical baroclinic disturbances.

In the tropics, however, the only well-documented linear instability is conditional instability. This instability has its maximum growth rate on the scale of an individual cumulus cloud. Therefore, it cannot explain the synoptic-scale organization of the motion. Observations indicate, moreover, that the mean tropical atmosphere is not saturated, even in the planetary boundary layer. Thus, a parcel must undergo a considerable amount of forced ascent before it reaches its LFC and becomes positively buoyant. Such forced parcel ascent can only be caused by small-scale motions, such as turbulent plumes in the boundary layer. The efficacy of boundary layer turbulence in producing parcel ascent to the LFC clearly depends on the temperature and humidity of the boundary layer environment. In the tropics it is difficult to initialize deep convection unless the boundary layer is brought toward saturation and destabilized, which may occur if there is large-scale (or mesoscale) ascent in the boundary layer. Thus, convection tends to be concentrated in regions of large-scale low-level convergence. This concentration does not arise because the large-scale convergence directly "forces" the convection, but rather because it preconditions the environment to be favorable for parcel ascent to the LFC.

Cumulus convection and the large-scale environmental motion are thus often viewed as cooperatively interacting. In this viewpoint, diabatic heating due to latent heat released by cumulus clouds produces a large-scale (or mesoscale) cyclonic disturbance; this disturbance, in turn, through boundary layer pumping, drives the low-level moisture convergence necessary to maintain an environment favorable for the development of cumulus convection. There have been attempts to formalize these ideas into a linear stability theory [often referred to as *conditional instability of the second kind* (CISK)], which attributes hurricane growth to the organized interaction between the cumulus scale and the large-scale moisture convergence. This interaction process is indicated schematically in Fig. 9.15a. The CISK model for hurricane development has not been very successful, as there is little evidence

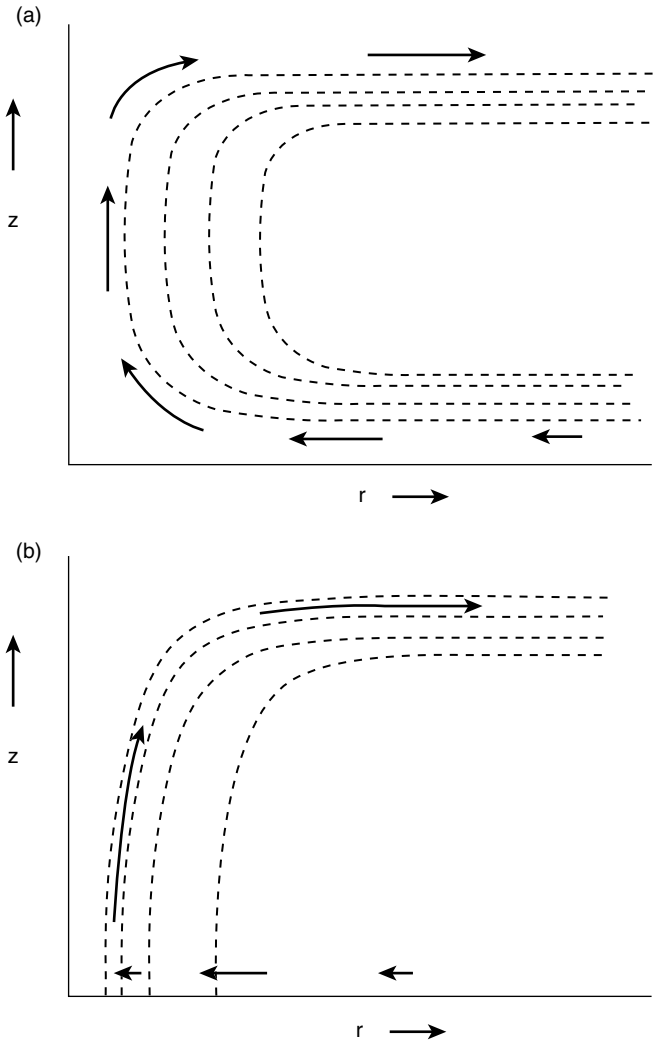


Fig. 9.15 Meridional cross sections showing the relationship between surfaces of constant saturation θ_e (dashed contours; values decreasing as r increases) and meridional circulation (arrows) in CISK (a) and WISHE (b) theories for hurricane development. (a) Frictionally induced boundary layer convergence moistens the environment and destabilizes it through layer ascent. This enables small-scale plumes to reach their levels of free convection easily and to produce cumulonimbus clouds. Diabatic heating due to the resulting precipitation drives the large-scale circulation and thus maintains the large-scale convergence. (b) The saturation θ_e is tied to the θ_e of the boundary layer. The warm core occurs because enhanced surface fluxes of latent heat increase θ_e there. (After Emanuel, 2000.)

that such interaction leads to a growth rate maximum on the observed scale of hurricanes.

In recent years a dramatically different view of the stability of the tropical atmosphere has come into vogue. This view, which is referred to as *wind-induced surface heat exchange* (WISHE), is based on air–sea interactions. According to the WISHE view, illustrated schematically in Fig. 9.15b, the potential energy for hurricanes arises from the thermodynamic disequilibrium between the atmosphere and the underlying ocean. The efficacy of air–sea interaction in providing potential energy to balance frictional dissipation depends on the rate of transfer of latent heat from the ocean to the atmosphere. This is a function of surface wind speed; strong surface winds, which produce a rough sea surface, can increase the evaporation rate greatly. Thus, hurricane development depends on the presence of a finite amplitude initiating disturbance, such as an equatorial wave, to provide the winds required to produce strong evaporation. Given a suitable initial disturbance, a feedback may occur in which an increase in inward spiraling surface winds increases the rate of moisture transfer from the ocean, which by bringing the boundary layer toward saturation increases the intensity of the convection, which further increases the secondary circulation.

The air–sea interaction theory is consistent with observations that hurricanes can develop only in the presence of very warm ocean surface temperatures. For surface temperatures less than 26°C, the converging flow in the boundary layer apparently cannot acquire a high enough equivalent potential temperature to sustain the intense transverse circulation needed to maintain the hurricane. Thus, it appears that hurricanes do not arise from linear instability associated with latent heating of the tropical atmosphere, but develop from preexisting large-scale disturbances under rather special conditions that permit a rapid flux of moisture from the ocean surface. The role of convection is then not to provide an internal heat source, but rather to rapidly establish a moist adiabatic lapse rate, tied to the boundary layer θ_e. This produces a warm core structure in the region of enhanced surface saturation θ_e (Fig. 9.15b).

PROBLEMS

9.1. Show by transforming from θ coordinates to height coordinates that the Ertel potential vorticity P is proportional to $F^2 N_s^2 - S^4$. [See (9.27).]

9.2. Show that (9.17), the quasi-geostrophic version of the equation for the streamfunction of the cross-frontal circulation, is equivalent to the omega equation (6.53) in the Boussinesq approximation for adiabatic flow independent of the x coordinate.

9.3. Starting with the linearized Boussinesq equations for a basic state zonal flow that is a function of height, derive (9.31) and verify the form given for the Scorer parameter.

9.4. Show that for stationary flow over an isolated ridge in the broad ridge limit ($k_s \ll m_s$), the group velocity vector is directed upward so that energy cannot propagate up or downstream of the ridge.

9.5. An air parcel at 920 hPa with temperature 20°C is saturated (mixing ratio 16 g kg^{-1}). Compute θ_e for the parcel.

9.6. Suppose that the mass of air in an entraining cumulus updraft increases exponentially with height so that $m = m_0 e^{z/H}$ where $H = 8$ km and m_0 is the mass at a reference level. If the updraft speed is 3 m s^{-1} at 2 km height, what is its value at a height of 8 km, assuming that the updraft has zero net buoyancy?

9.7. Verify the approximate relationship between moist static energy and θ_e given by (9.41).

9.8. The azimuthal velocity component in some hurricanes is observed to have a radial dependence given by $v_\lambda = V_0(r_0/r)^2$ for distances from the center given by $r \geq r_0$. Letting $V_0 = 50\,\mathrm{m\,s^{-1}}$ and $r_0 = 50\,\mathrm{km}$, find the total geopotential difference between the far field ($r \to \infty$) and $r = r_0$, assuming gradient wind balance and $f_0 = 5 \times 10^{-5}$ s^{-1}. At what distance from the center does the Coriolis force equal the centrifugal force?

9.9. Starting with (9.61) derive the angular momentum form of the gradient wind balance for an axisymmetric vortex given by (9.62).

MATLAB EXERCISES

M9.1. The MATLAB script **surface_front_1.m** demonstrates the concentration of an initial temperature gradient by an imposed deformation field. The temperature field is initially taken to have a linear decrease in y: $T(y) = 300 - 10^{-2}y$, where y is in kilometers. The horizontal wind is given by the streamfunction $\psi(x, y) = -15k^{-1}\sin(kx)\sin(my)$, where $k = 2\pi/3 \times 10^6$ and $m = \pi/3 \times 10^6$. Evaluate and contour plot the deformation field for this streamfunction, where deformation is defined as $\partial u/\partial x - \partial v/\partial y$. Run the model and observe where the temperature gradient increases. Modify the code to plot the maximum temperature gradient as a function of time. For how long does the gradient increase approximately exponentially? What is the e-folding time for increase of the temperature gradient at the origin?

M9.2. The MATLAB script **profile_1.m** and the function **Tmoist.m** use pressure, temperature, and humidity data in the file **sounding.dat** to calculate and graph vertical profiles of temperature, dewpoint, and the temperature corresponding to pseudoadiabatic ascent from the lifting condensation level

for a parcel lifted from the lowest level of the sounding. Modify this script to plot profiles of the potential temperature, the equivalent potential temperature, and the saturated potential temperature. *Hint:* Temperature at saturation can be calculated by inverting the expression for saturation vapor pressure, which is obtainable by integrating (D.4):

$$e_s = e_{s,tr} \exp\left[\frac{L_v}{R_v}\left(\frac{1}{T_{tr}} - \frac{1}{T}\right)\right]$$

where $e_{s,tr} = 6.11\text{hPa}$, $T_{tr} = 273.16\text{K}$, and $L_v = 2.5 \times 10^6\ \text{J kg}^{-1}$ are the saturation vapor pressure at the triple point temperature, the triple point temperature, and the latent heat of vaporization, respectively.

M9.3. For the thermodynamic sounding of Problem M9.2 compute the CAPE and the profile of vertical velocity for a parcel lifted from the lowest layer in the absence of entrainment. What is the maximum vertical velocity, and what is the distance that the parcel overshoots its level of neutral buoyancy?

M9.4. The MATLAB script **lee_wave_1.m** uses the Fourier series approach to solve for flow over an isolated ridge for the case of constant N and constant mean zonal flow. For the case of $\bar{u} = 50\ \text{m s}^{-1}$ determine by running the script for various ridge widths what width is required for significant vertical propagation (i.e., at what ridge width do phase lines of vertical velocity begin to tilt with height)?

M9.5. The script **lee_wave_2.m** plots an approximate analytic solution for flow over a mountain in the wide ridge limit. The height profile is given by

$$h(x) = h_0 L^2 \Big/ \left(L^2 + x^2\right)$$

where L is the width scale of the ridge and h_0 is the ridge height. For a fixed ridge height of 2 km, vary the input zonal mean wind and comment on the dependence of the contoured solutions. (Note that the vertical scale is given in vertical wavelengths as in Fig. 9.7 in the text.) Modify the script to vary the mountain height and determine what fraction of the vertical wavelength the mountain height must have in order that wave breaking occur. (The onset of wave breaking occurs when the streamlines become vertical.)

Suggested References

Durran (1990) reviews the dynamics of mountain waves and downslope windstorms.

Eliassen (1990) discusses the secondary circulation within frontal zones.

Emanuel, *Atmospheric Convection* is a graduate-level text covering the subject of convection, including symmetric instability in great detail.

Emanuel (2000) provides a qualitative explanation of the WISHE theory for hurricane development.
Hoskins (1982) discusses the semigeostrophic theory of frontogenesis and summarizes results from numerical models.
Houze, *Cloud Dynamics* discusses convective storms at the graduate level.
Klemp (1987) describes the dynamics of tornadic thunderstorms.

CHAPTER 10

The General Circulation

In its broadest sense the *general circulation* of the atmosphere is usually considered to include the totality of motions that characterizes the global scale atmospheric flow. Specifically, the study of the general circulation is concerned with the dynamics of climate—that is, with the temporally averaged structures of the fields of wind, temperature, humidity, precipitation, and other meteorological variables. The general circulation may thus be considered to consist of the flow averaged in time over a period sufficiently long to remove the random variations associated with individual weather systems, but short enough to retain monthly and seasonal variations.

In the past, both observational and theoretical studies of the general circulation concentrated on the dynamics of the zonally averaged flow. The time-averaged circulation is, however, highly dependent on longitude due to longitudinally asymmetric forcing by orography and land–sea heating contrasts. The longitudinally dependent components of the general circulation may be separated into *quasi-stationary circulations*, which vary little in time, *monsoonal circulations*, which are seasonally reversing, and various subseasonal and interannual components, which together account for *low-frequency variability*. A complete understanding of the physical basis for the general circulation requires an explanation not only

for the zonally averaged circulation, but the longitudinally and time-varying components as well.

Nevertheless, to introduce the study of the general circulation, it proves useful to isolate those processes that maintain the zonal-mean flow (i.e., the flow averaged around latitude circles). This approach follows naturally from the linear wave studies of previous chapters in which flow fields were split into zonal-mean and longitudinally dependent eddy components. In this chapter, however, we concentrate not on the development and motion of the eddies, but on the influence of the eddies on the structure of the zonal-mean circulation. Focusing on the zonal mean allows us to isolate those features of the circulation that are not dependent on continentality, and should thus be common to all thermally driven rotating fluid systems. In particular, we discuss the angular momentum and energy budgets of the zonally averaged flow. We also show that the mean meridional circulation (i.e., the circulation consisting of the zonal-mean vertical and meridional velocity components) satisfies a diagnostic equation analogous to the omega equation of Section 6.4.1, but with the forcing determined by the distributions of diabatic heating and eddy heat and momentum fluxes.

Following our discussion of the zonal-mean circulation we consider the longitudinally varying time averaged circulation. In this chapter the primary emphasis is on extratropical aspects of the circulation; these can be discussed within the framework of quasi-geostrophic theory. The general circulation of the tropics is considered in Chapter 11.

10.1 THE NATURE OF THE PROBLEM

Theoretical speculation on the nature of the general circulation has quite a long history. Perhaps the most important early work on the subject was that of the 18th-century Englishman, George Hadley. Hadley, in seeking a cause for the trade wind circulation, realized that this circulation must be a form of thermal convection driven by the difference in solar heating between equatorial and polar regions. He visualized the general circulation as consisting of a zonally symmetric overturning in which the heated equatorial air rises and flows poleward where it cools, sinks, and flows equatorward again. At the same time, the Coriolis force deflects the poleward moving air at the upper levels to the east, and the equatorward moving air near the surface to the west. The latter is, of course, consistent with the observed surface winds in the trade wind zone, which are northeasterly in the Northern Hemisphere and southeasterly in the Southern Hemisphere. This type of circulation is now called a *Hadley circulation* or Hadley cell.

Although a circulation consisting of Hadley cells extending from equator to pole in each hemisphere is mathematically possible in the sense that such a circulation would not violate the laws of physics, the observed Hadley circulation is confined to

the tropics. Evidence from a number of studies indicates that for conditions existing in the earth's atmosphere a symmetric hemispheric-wide Hadley circulation would be baroclinically unstable. If such a circulation were to become established by some mechanism, it would quickly break down outside the tropics as baroclinic eddies developed and modified the zonal-mean circulation through their heat and momentum fluxes.

The observed general circulation thus cannot be understood purely in terms of zonally symmetric processes. Rather, it can be thought of qualitatively as developing through three-dimensional interactions among radiative and dynamical processes. In the mean the net solar energy absorbed by the atmosphere and the earth must equal the infrared energy radiated back to space by the planet. The annually averaged solar heating is, however, strongly dependent on latitude, with a maximum at the equator and minima at the poles. The outgoing infrared radiation, however, is only weakly latitude dependent. Thus, there is a net radiation surplus in the equatorial region and a deficit in the polar region. This differential heating warms the equatorial atmosphere relative to higher latitudes and creates a pole-to-equator temperature gradient. Hence it produces a growing store of zonal-mean available potential energy. At some point the westerly thermal wind (which must develop if the motion is to be balanced geostrophically in the presence of the pole-to-equator temperature gradient) becomes baroclinically unstable. As shown in chapter 8, the resulting baroclinic waves transport heat poleward. These waves will intensify until their heat transport (together with the heat transported by planetary waves and ocean currents) is sufficient to balance the radiation deficit in the polar regions so that the pole-to-equator temperature gradient ceases to grow. At the same time, these perturbations convert potential energy into kinetic energy, thereby maintaining the kinetic energy of the atmosphere against the effects of frictional dissipation.

From a thermodynamic point of view, the atmosphere may be regarded as a "heat engine," which absorbs net heat at relatively warm temperatures in the tropics (primarily in the form of latent heat due to evaporation from the sea surface) and gives up heat at relatively cool temperatures in the extratropics. In this manner net radiation generates available potential energy, which is in turn partially converted to kinetic energy, which does work to maintain the circulation against frictional dissipation. Only a small fraction of the solar energy input actually gets converted to kinetic energy. Thus, from an engineer's viewpoint the atmosphere is a rather inefficient heat engine. However, if due account is taken of the many constraints operating on atmospheric motions, it appears that the atmosphere may in fact generate kinetic energy about as efficiently as dynamically possible.

The above qualitative discussion suggests that the gross features of the general circulation outside the tropics can be understood on the basis of quasi-geostrophic theory since, as we have previously seen, baroclinic instability is contained within

the quasi-geostrophic framework. In view of this fact, and to keep the equations as simple as possible, our discussion of the zonally averaged and longitudinally varying components of the circulation outside the tropics will concentrate on those aspects that can be qualitatively represented by the quasi-geostrophic equations on the midlatitude β plane.

It should be recognized that a quasi-geostrophic model cannot provide a complete theory of the general circulation because in the quasi-geostrophic theory a number of assumptions are made concerning scales of motion that restrict in advance the possible types of solutions. Quantitative modeling of the general circulation requires complicated numerical models based on the primitive equations in spherical coordinates. The ultimate objective of these modeling efforts is to simulate the general circulation so faithfully that the climatological consequences of any change in the external parameters (such as the atmospheric concentration of carbon dioxide) can be predicted accurately. Present models can provide fairly accurate simulations of the current climate and plausible predictions of the response of the climate system to changes in external conditions. However, uncertainties in the representations of a number of physical processes, particularly clouds and precipitation, limit the confidence that can be placed in quantitative climate change predictions based on such models.

10.2 THE ZONALLY AVERAGED CIRCULATION

The observed global distribution of the longitudinally averaged zonal wind for the two solstice seasons was shown in Fig. 6.1. Although there are important interhemispheric differences, the flow in both hemispheres is characterized by westerly jets with maximum zonal velocities at about 30°–35° latitude near the tropopause. These westerlies vary in longitude, especially in the Northern Hemisphere (see Fig. 6.2), but there is a significant zonally symmetric component, which we will refer to as the *mean zonal wind*.

Although departures of the time-averaged flow from zonal symmetry are important aspects of the general circulation, especially in the Northern Hemisphere, it is useful to first obtain an understanding of the dynamics of the zonally symmetric component before investigating the three-dimensional time mean circulation. This section examines the dynamics of zonally symmetric motions using quasi-geostrophic theory and the log-pressure coordinate system introduced in Chapter 8 and shows that the meridional circulation associated with an axially symmetric vortex is dynamically analogous to the secondary divergent circulation in a baroclinic wave.

In the log-pressure coordinate system, the x and y components of the momentum equation, the hydrostatic approximation, the continuity equation, and the thermodynamic energy equation can be written as

$$Du/Dt - fv + \partial\Phi/\partial x = X \tag{10.1}$$

$$Dv/Dt + fu + \partial\Phi/\partial y = Y \tag{10.2}$$

$$\partial\Phi/\partial z = H^{-1}RT \tag{10.3}$$

$$\partial u/\partial x + \partial v/\partial y + \rho_0^{-1}\partial\left(\rho_0 w\right)/\partial z = 0 \tag{10.4}$$

$$DT\,Dt + \left(\kappa T\big/H\right)w = J\big/c_p \tag{10.5}$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}$$

and X and Y designate the zonal and meridional components of drag due to small-scale eddies.

For convenience in this and the following chapters we have dropped the asterisk notation used in Chapter 8 to distinguish the log-pressure coordinate from geometric height. Thus, z here designates the log-pressure variable defined in Section (8.4.1).

Analysis of the zonally averaged circulation involves study of the interaction of longitudinally varying disturbances (referred to as eddies, and denoted by primed variables) with the longitudinally averaged flow (referred to as the mean flow and denoted by overbars). Thus, any variable A is expanded in the form $A = \overline{A} + A'$. This sort of average is an *Eulerian mean*, as it is evaluated at fixed latitude, height, and time. Eulerian mean equations are obtained by taking zonal averages of (10.1)–(10.5). Such averaging is facilitated by using (10.4) to expand the material derivative for any variable A in flux form as follows:

$$\begin{aligned}\rho_0\frac{DA}{Dt} &= \rho_0\left(\frac{\partial}{\partial t} + \mathbf{V}\bullet\boldsymbol{\nabla} + w\frac{\partial}{\partial z}\right)A + A\left[\boldsymbol{\nabla}\bullet\left(\rho_0\mathbf{V}\right) + \frac{\partial}{\partial z}\left(\rho_0 w\right)\right] \\ &= \frac{\partial}{\partial t}\left(\rho_0 A\right) + \frac{\partial}{\partial x}\left(\rho_0 A u\right) + \frac{\partial}{\partial y}\left(\rho_0 A v\right) + \frac{\partial}{\partial z}\left(\rho_0 A w\right)\end{aligned} \tag{10.6}$$

where we recall that $\rho_0 = \rho_0(z)$.

Applying the zonal averaging operator then gives

$$\rho_0\overline{\frac{DA}{Dt}} = \frac{\partial}{\partial t}\left(\rho_0\overline{A}\right) + \frac{\partial}{\partial y}\left[\rho_0\left(\overline{A}\,\overline{v} + \overline{A'v'}\right)\right] + \frac{\partial}{\partial z}\left[\rho_0\left(\overline{A}\,\overline{w} + \overline{A'w'}\right)\right] \tag{10.7}$$

Here we have used the fact that $\partial\,\overline{(\,)}/\partial x = 0$, as quantities with the overbar are independent of x. We have also used the fact that for any variables a and b

$$\overline{ab} = \overline{(\overline{a} + a')\left(\overline{b} + b'\right)} = \overline{\overline{a}\overline{b}} + \overline{\overline{a}b'} + \overline{a'\overline{b}} + \overline{a'b'} = \overline{a}\overline{b} + \overline{a'b'}$$

which follows from that facts that $\overline{a}$ and $\overline{b}$ are independent of x and $\overline{a'} = \overline{b'} = 0$ so that, for example, $\overline{\overline{a}b'} = \overline{a}\overline{b'} = 0$.

The terms involving $(\overline{v}, \overline{w})$ on the right in (10.7) can be rewritten in advective form with the aid of the zonal mean of (10.4):

$$\partial\overline{v}/\partial y + \rho_0^{-1}\partial(\rho_0\overline{w})/\partial z = 0 \tag{10.8}$$

Applying the chain rule of differentiation to the mean terms on the right in (10.7) and substituting from (10.8) we can rewrite (10.7) as

$$\rho_0\frac{\overline{DA}}{Dt} = \frac{\overline{D}}{Dt}\left(\rho_0\overline{A}\right) + \frac{\partial}{\partial y}\left[\rho_0\left(\overline{A'v'}\right)\right] + \frac{\partial}{\partial z}\left[\rho_0\left(\overline{A'w'}\right)\right] \tag{10.9}$$

where

$$\frac{\overline{D}}{Dt} \equiv \frac{\partial}{\partial t} + \overline{v}\frac{\partial}{\partial y} + \overline{w}\frac{\partial}{\partial z} \tag{10.10}$$

is the rate of change following the mean meridional motion $(\overline{v}, \overline{w})$.

10.2.1 The Conventional Eulerian Mean

Applying the averaging scheme of (10.9) to (10.1) and (10.5) we obtain the zonal-mean zonal momentum and thermodynamic energy equations for quasi-geostrophic motions on the midlatitude β plane:

$$\partial\overline{u}/\partial t - f_0\overline{v} = -\partial\left(\overline{u'v'}\right)/\partial y + \overline{X} \tag{10.11}$$

$$\partial\overline{T}/\partial t + N^2HR^{-1}\overline{w} = -\partial\left(\overline{v'T'}\right)/\partial y + \overline{J}/c_p \tag{10.12}$$

where N is the buoyancy frequency defined by

$$N^2 \equiv \frac{R}{H}\left(\frac{\kappa T_0}{H} + \frac{dT_0}{dz}\right)$$

In (10.11) and (10.12), consistent with quasi-geostrophic scaling, we have neglected advection by the ageostrophic mean meridional circulation and vertical eddy flux divergences. It is easily confirmed that for quasi-geostrophic scales these terms are small compared to the retained terms (see Problem 10.4). We have included the zonally averaged turbulent drag in (10.11) because stresses due to unresolved eddies (such as gravity waves) may be important not only in the boundary layer, but also in the upper troposphere and in the middle atmosphere.

Similar scaling shows that the zonal average of the meridional momentum equation (10.2) can be approximated accurately by geostrophic balance,

$$f_0\overline{u} = -\partial\overline{\Phi}/\partial y$$

This can be combined with the hydrostatic approximation (10.3) to give the thermal wind relation

$$f_0\partial\overline{u}/\partial z + RH^{-1}\partial\overline{T}/\partial y = 0 \tag{10.13}$$

This relationship between zonal-mean wind and potential temperature distributions imposes a strong constraint on the ageostrophic mean meridional circulation $(\overline{v}, \overline{w})$. In the absence of a mean meridional circulation the eddy momentum flux divergence in (10.11) and eddy heat flux divergence in (10.12) would tend separately to change the mean zonal wind and temperature fields, and hence would destroy thermal wind balance. The pressure gradient force that results from any small departure of the mean zonal wind from geostrophic balance will, however, drive a mean meridional circulation, which adjusts the mean zonal wind and temperature fields so that (10.13) remains valid. In many situations this compensation allows the mean zonal wind to remain unchanged even in the presence of large eddy heat and momentum fluxes. The mean meridional circulation thus plays exactly the same role in the zonal-mean circulation that the secondary divergent circulation plays in the synoptic-scale quasi-geostrophic system. In fact, for steady-state mean flow conditions the $(\overline{v}, \overline{w})$ circulation must just balance the eddy forcing plus diabatic heating so that the balances in (10.11) and (10.12) are as follows:

Coriolis force $(f_0\overline{v}) \approx$ divergence of eddy momentum fluxes

Adiabatic cooling $\approx$ diabatic heating plus convergence of eddy heat fluxes

Analysis of observations shows that outside the tropics these balances appear to be approximately true above the boundary layer. Thus, changes in the zonal- mean flow arise from small imbalances between the forcing terms and the mean meridional circulation.

The Eulerian mean meridional circulation can be determined in terms of the forcing from an equation similar to the omega equation of Section 6.4. Before deriving this equation it is useful to observe that the mean meridional mass circulation is nondivergent in the meridional plane. Thus, it can be represented in terms of a meridional mass transport stream function, which identically satisfies the continuity equation (10.8), by letting

$$\rho_0\overline{v} = -\frac{\partial\overline{\chi}}{\partial z}; \quad \rho_0\overline{w} = \frac{\partial\overline{\chi}}{\partial y} \tag{10.14}$$

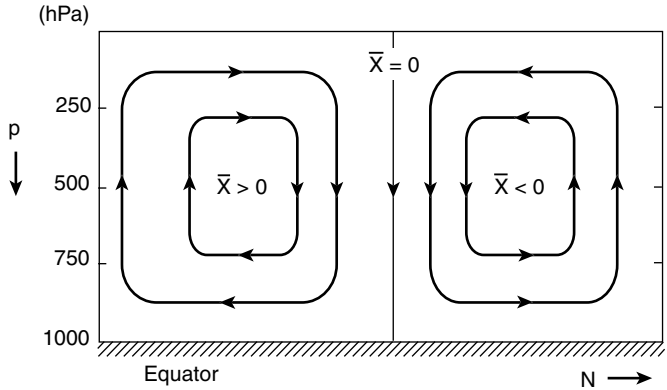


Fig. 10.1 Relationship of the Eulerian mean meridional stream function to vertical and meridional motion.

The relationship of the sign of the streamfunction $\overline{\chi}$ to the sense of the mean meridional circulation is shown schematically in Fig. 10.1.

The diagnostic equation for $\overline{\chi}$ is derived by first taking

$$f_0 \frac{\partial}{\partial z}\,(10.11) + \frac{R}{H}\frac{\partial}{\partial y}\,(10.12)$$

and then using (10.13) to eliminate the time derivatives and (10.14) to express the mean meridional circulation in terms of $\overline{\chi}$. The resulting elliptic equation has the form

$$\begin{aligned}\frac{\partial^2 \overline{\chi}}{\partial y^2} + \frac{f_0^2}{N^2}\rho_0 \frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial \overline{\chi}}{\partial z}\right) = \frac{\rho_0}{N^2}\Bigg[\frac{\partial}{\partial y}\left(\frac{\kappa \overline{J}}{H} - \frac{R}{H}\frac{\partial}{\partial y}\left(\overline{v'T'}\right)\right) \\ - f_0\left(\frac{\partial^2}{\partial z \partial y}\left(\overline{u'v'}\right) - \frac{\partial \overline{X}}{\partial z}\right)\Bigg]\end{aligned} \tag{10.15}$$

Equation (10.15) can be used to qualitatively diagnose the mean meridional circulation. Because $\overline{\chi}$ must vanish on the boundaries it can be represented by a double Fourier series in y and z. Hence, the elliptic operator on the left-hand side of (10.15) is approximately proportional to $-\overline{\chi}$ and (10.15) states qualitatively that

$$\begin{aligned}\overline{\chi} \propto -\frac{\partial}{\partial y}\,(\text{diabatic heating}) + \frac{\partial^2}{\partial y^2}\,(\text{large-scale eddy heat flux}) \\ + \frac{\partial^2}{\partial y \partial z}\,(\text{large-scale eddy momentum flux}) + \frac{\partial}{\partial z}\,(\text{zonal drag force})\end{aligned}$$

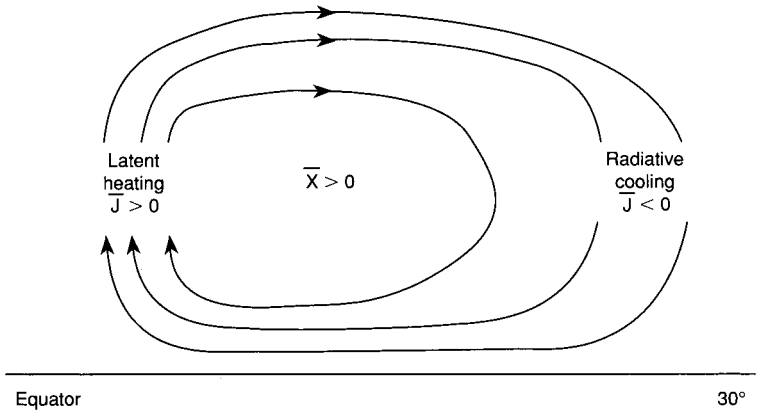


Fig. 10.2 Schematic Eulerian mean meridional circulation showing the streamfunction for a thermally direct Hadley cell.

Now, diabatic heating in the Northern Hemisphere decreases for increasing y. Thus, the first term on the right is positive and tends to force a mean meridional cell with $\overline{\chi} > 0$. This is referred to as a *thermally direct* cell, as warm air is rising and cool air sinking. It is this process that primarily accounts for the Hadley circulation of the tropics as illustrated in Fig. 10.2. For an idealized Hadley cell in the absence of eddy sources, the differential diabatic heating would be balanced only by adiabatic cooling near the equator and adiabatic warming at higher latitudes.

In the extratropical Northern Hemisphere, poleward eddy heat fluxes due both to transient synoptic-scale eddies and to stationary planetary waves tend to transfer heat poleward, producing a maximum poleward heat flux $\overline{v'T'}$ in the lower troposphere at about 50° latitude as shown in Fig. 10.3. Because $\overline{\chi}$ is proportional to the second derivative of $\overline{v'T'}$, which should be negative where $\overline{v'T'} > 0$, the heat flux forcing term should tend to produce a mean meridional cell with $\overline{\chi} < 0$ centered in the lower troposphere at midlatitudes. Thus, the eddy heat flux tends to drive an indirect meridional circulation.

The existence of this indirect meridional circulation can be understood in terms of the need to maintain geostrophic and hydrostatic balance. North of the latitude where $\overline{v'T'}$ is a maximum there is a convergence of eddy heat flux, while equatorward of that latitude there is a divergence. Thus, eddy heat transport tends to reduce the pole-to-equator mean temperature gradient. If the mean zonal flow is to remain geostrophic, the thermal wind must then also decrease. In the absence of eddy momentum transport, this decrease in the thermal wind can only be produced by the Coriolis torque due to a mean meridional circulation with the sense of that in Fig. 10.4. At the same time, it is not surprising to find that the vertical mean motions required by continuity oppose the temperature tendency associated

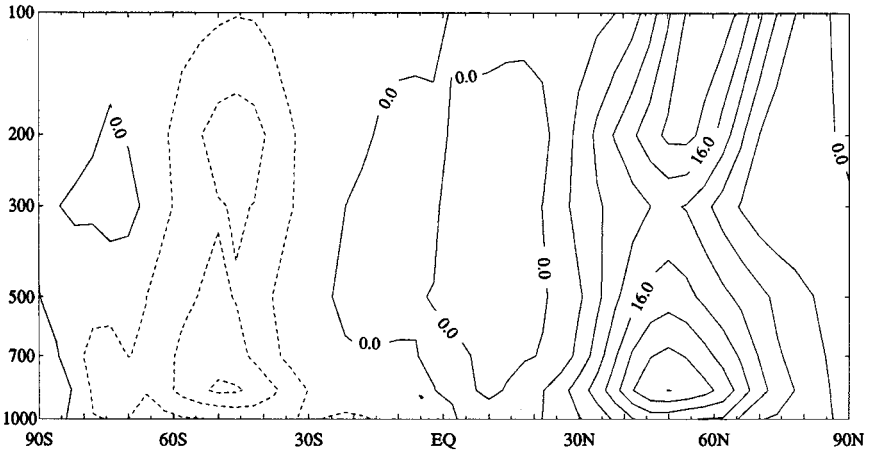


Fig. 10.3 Observed northward eddy heat flux distribution (°C m s^{-1}) for Northern Hemisphere winter. (Adapted from Schubert et al., 1990.)

with the eddy heat flux by producing adiabatic warming in the region of eddy heat flux divergence, and adiabatic cooling in the region of eddy heat flux convergence.

The next to last forcing term in (10.15) is proportional to the vertical gradient of the horizontal eddy momentum flux convergence. However, it can be shown (Problem 10.5) that

$$-\frac{\partial^2}{\partial z \partial y}\left(\overline{u'v'}\right) = +\frac{\partial}{\partial z}\left(\overline{v'\zeta'}\right)$$

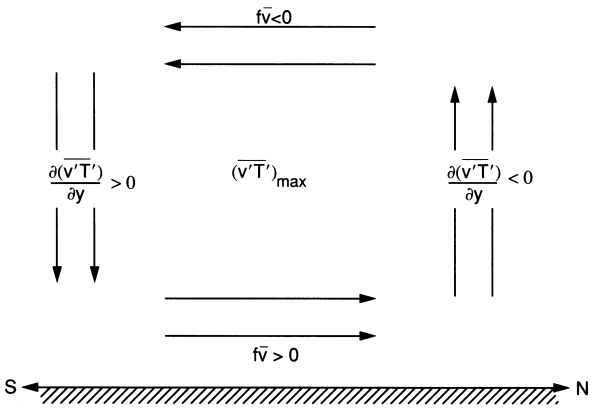


Fig. 10.4 Schematic Eulerian mean meridional circulation forced by poleward heat fluxes.

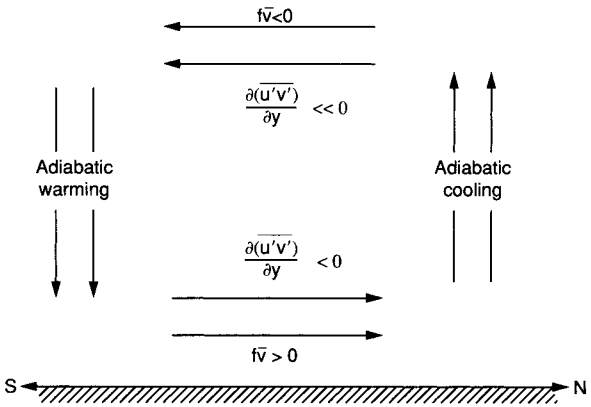


Fig. 10.5 Schematic Eulerian mean meridional circulation forced by a vertical gradient in eddy momentum flux convergence.

Thus, this term is proportional to the vertical derivative of the meridional vorticity flux. To interpret this eddy forcing physically, we suppose as shown in Fig. 10.5 that the momentum flux convergence (or the vorticity flux) is positive and increasing with height. This will be true in the Northern Hemisphere troposphere poleward of the core of the jet stream, as $\overline{u'v'}$ tends to be poleward and to reach its maximum near the tropopause at about 30° (at the core of the mean jetstream) as shown in Fig. 10.6. For this configuration of momentum flux $\partial^2\overline{u'v'}/\partial y\partial z < 0$ in the midlatitude troposphere, which again drives a mean meridional cell with $\overline{\chi} < 0$. From (10.11) it is clear that the Coriolis force of this induced indirect meridional

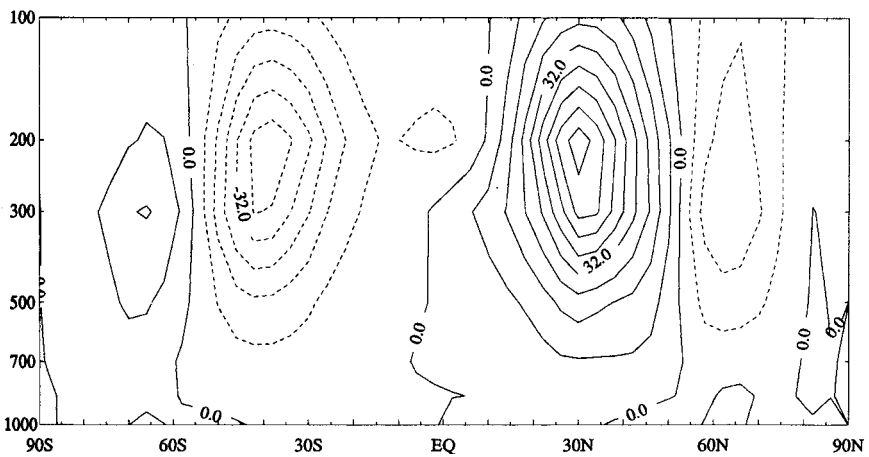


Fig. 10.6 Observed northward eddy momentum flux distribution ($m^2\ s^{-2}$) for Northern Hemisphere winter. (Adapted from Schubert et al., 1990.)

circulation is required to balance the acceleration due to the momentum flux convergence, which would otherwise increase the vertical shear of the mean zonal wind and destroy the thermal wind balance.

Thus, the combined eddy heat flux and the eddy momentum flux distributions tend to drive mean meridional cells in each hemisphere with rising motion poleward of 45° and sinking motion equatorward of 45°. This eddy forcing more than compensates the direct diabatic drive at midlatitudes and is responsible for the observed thermally indirect *Ferrel cell*.

The resultant observed climatology of the Eulerian mean meridional circulation is shown in Fig. 10.7. It consists primarily of tropical Hadley cells driven by diabatic heating and eddy-driven midlatitude Ferrel cells. There are also minor thermally direct cells at polar latitudes. The meridional circulation in the winter is much stronger than that in the summer, especially in the Northern Hemisphere. This reflects the seasonal variation both in the diabatic and in the eddy flux forcing terms in (10.15).

The zonal momentum balance in the upper troposphere in tropical and midlatitude cells is maintained by the balance between the Coriolis force caused by the mean meridional drift and the eddy momentum flux convergence. The heat balance is maintained by rising motion (adiabatic cooling) balancing the diabatic heating in the tropics and the eddy heat flux convergence at high latitudes, and subsidence (adiabatic warming) balancing the eddy heat flux divergence in the subtropics.

Because of the appearance of eddy flux terms in both mean momentum and thermodynamic energy equations, and the near cancellation of eddy and mean flow processes, it is rather inefficient to attempt to diagnose the net eddy forcing of the mean flow from the conventional Eulerian mean. It can be shown that similar

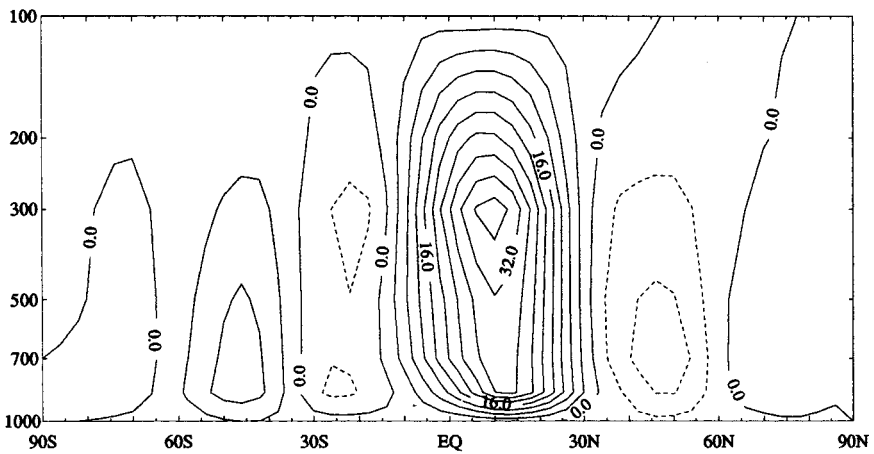


Fig. 10.7 Streamfunction (units: 10^2 kg m^{-1}s^{-1}) for the observed Eulerian mean meridional circulation for Northern Hemisphere winter, based on the data of Schubert et al. (1990).

eddy and mean flow compensation occurs in the Eulerian mean continuity equation for a long-lived tracer so that tracer transport calculations are also inefficient in this formulation.

10.2.2 The Transformed Eulerian Mean (TEM)

An alternative approach to analysis of the zonal-mean circulation, which provides a clearer diagnosis of eddy forcing and also provides a more direct view of transport processes in the meridional plane, is the *transformed Eulerian mean* formulation introduced by Andrews and McIntyre (1976). This transformation takes account of the fact that in (10.12) there tends to be a strong cancellation between the eddy heat flux convergence and adiabatic cooling, while the diabatic heating term is a small residual. Since in the mean an air parcel will rise to a higher equilibrium altitude only if its potential temperature is increased by diabatic heating, it is the r*esidual meridional circulation* associated with diabatic processes that is directly related to the mean meridional mass flow.

The TEM equations can be obtained from (10.11) and (10.12) by defining the residual circulation $\left(\overline{v}^*, \overline{w}^*\right)$ as follows:

$$\overline{v}^* = \overline{v} - \rho_0^{-1} R H^{-1} \partial \left(\rho_0 \overline{v'T'} \Big/ N^2\right) \Big/ \partial z \tag{10.16a}$$

$$\overline{w}^* = \overline{w} + R H^{-1} \partial \left(\overline{v'T'} \Big/ N^2\right) \Big/ \partial y \tag{10.16b}$$

The residual vertical velocity defined in this manner clearly represents that part of the mean vertical velocity whose contribution to adiabatic temperature change is not canceled by the eddy heat flux divergence.

Substituting from (10.16) into (10.11) and (10.12) to eliminate $(\overline{v}, \overline{w})$ yields the TEM equations

$$\partial \overline{u} \big/ \partial t - f_0 \overline{v}^* = +\rho_0^{-1} \nabla \bullet \mathbf{F} + \overline{X} \equiv \overline{G} \tag{10.17}$$

$$\partial \overline{T} \big/ \partial t + N^2 H R^{-1} \overline{w}^* = \overline{J} \big/ c_p \tag{10.18}$$

$$\partial \overline{v}^* \big/ \partial y + \rho_0^{-1} \partial \left(\rho_0 \overline{w}^*\right) \big/ \partial z = 0 \tag{10.19}$$

where $\mathbf{F} \equiv \mathbf{j} F_y + \mathbf{k} F_z$, the *Eliassen–Palm flux* (EP flux), is a vector in the meridional (y, z) plane, which for large-scale quasi-geostrophic eddies has the components

$$F_y = -\rho_0 \overline{u'v'}, \qquad F_z = \rho_0 f_0 R \overline{v'T'} \Big/ \left(N^2 H\right) \tag{10.20}$$

and $\overline{G}$ designates the total zonal force due to both large- and small-scale eddies.

The TEM formulation clearly shows that the eddy heat and momentum fluxes do not act separately to drive changes in the zonal-mean circulation, but only in

the combination given by the divergence of the EP flux. The fundamental role of the eddies is thus to exert a zonal force. This eddy forcing of the zonal-mean flow can be displayed conveniently by mapping the field of **F** and contouring the isolines of its divergence. When properly scaled by the basic state density, these contours give the zonal force per unit mass exerted by quasi-geostrophic eddies. The mean global EP flux divergence pattern for a Northern Hemisphere winter is shown in Fig. 10.8. Note that in most of the extratropical troposphere the EP flux is convergent so that the eddies exert a westward zonal force on the atmosphere. On the seasonal time scale the zonal force due to the EP flux divergence in (10.17) is nearly balanced by the eastward Coriolis force of the residual mean meridional circulation. Conditions for this balance are discussed in the next subsection.

The structure of the residual mean meridional circulation can be determined by defining a residual stream function

$$\overline{\chi}^* \equiv \overline{\chi} + \rho_0 \frac{R}{H} \frac{\overline{v'T'}}{N^2}$$

It can then be shown by direct substitution into (10.14) and (10.15) that

$$\rho_0 \overline{v}^* = -\frac{\partial \overline{\chi}^*}{\partial z}; \quad \rho_0 \overline{w}^* = \frac{\partial \overline{\chi}^*}{\partial y}$$

and

$$\frac{\partial^2 \overline{\chi}^*}{\partial y^2} + \rho_0 \frac{f_0^2}{N^2} \frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial \overline{\chi}^*}{\partial z}\right) = \frac{\rho_0}{N^2}\left[\frac{\partial}{\partial y}\left(\frac{\kappa \overline{J}}{H}\right) + f_0 \frac{\partial \overline{G}}{\partial z}\right] \tag{10.21}$$

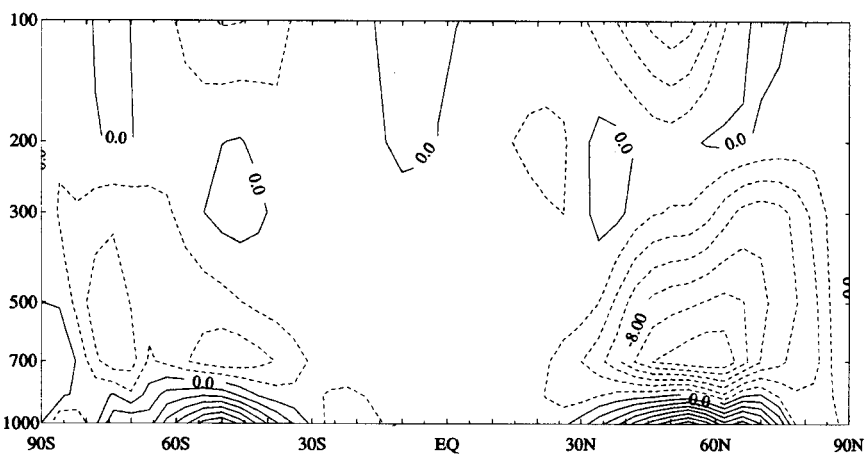


Fig. 10.8 Eliassen–Palm flux divergence divided by the standard density ρ_0 for Northern Hemisphere winter, based on the data of Schubert et al. (1990). (Units: m s^{-1} day^{-1}.)

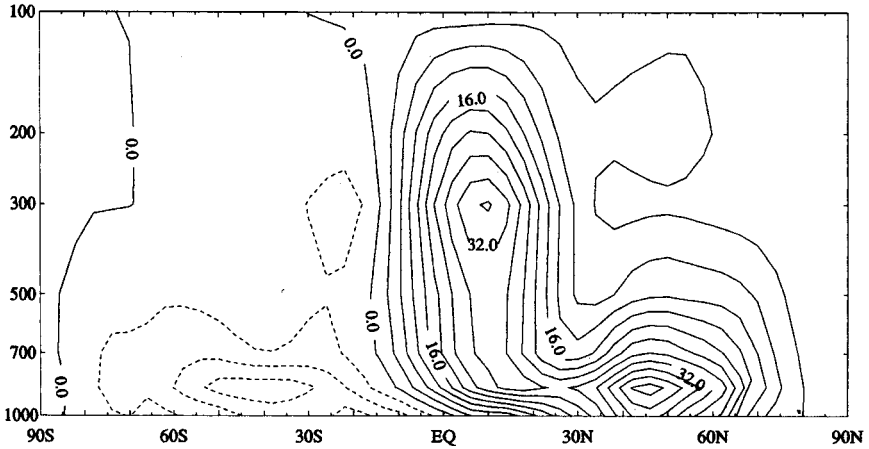


Fig. 10.9 Residual mean meridional stream function (units: 10^2 kg m^{-1} s^{-1}) for Northern Hemisphere winter, based on the data of Schubert et al. (1990).

The magnitudes of the diabatic and EP flux contributions to the source term on the right in (10.21) are generally larger in the winter hemisphere than in the summer hemisphere. In the Northern Hemisphere troposphere the source terms are generally negative, whereas in the Southern Hemisphere they are generally positive. This implies that $\overline{\chi}^*$ itself is positive in the Northern Hemisphere and negative in the Southern Hemisphere so that the residual meridional circulation consists of a single thermally direct overturning in each hemisphere, with the strongest cell in the winter hemisphere as shown in Fig. 10.9.

Unlike the conventional Eulerian mean, the residual mean vertical motion for time-averaged conditions is proportional to the rate of diabatic heating. It approximately represents the *diabatic circulation* in the meridional plane, that is, the circulation in which parcels that rise are diabatically heated and those that sink are diabatically cooled in order that their potential temperatures adjust to the local environment. The time-averaged residual mean meridional circulation thus approximates the mean motion of air parcels, and hence, unlike the conventional Eulerian mean, provides an approximation to the mean advective transport of trace substances.

10.2.3 The Zonal-Mean Potential Vorticity Equation

Further insight into the nature of the extratropical zonal-mean circulation can be obtained by zonally averaging the quasi-geostrophic potential vorticity equation (6.24) to obtain

$$\partial \overline{q} / \partial t = -\partial \left(\overline{q'v'} \right) / \partial y \tag{10.22}$$

where from (6.25) the zonal-mean potential vorticity is

$$\overline{q} = f_0 + \beta y + \frac{1}{f_0}\frac{\partial^2 \overline{\Phi}}{\partial y^2} + \frac{f_0}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\frac{\partial \overline{\Phi}}{\partial z}\right) \tag{10.23}$$

and the eddy potential vorticity is

$$q' = \frac{1}{f_0}\left(\frac{\partial^2 \Phi'}{\partial x^2} + \frac{\partial^2 \Phi'}{\partial y^2}\right) + \frac{f_0}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\frac{\partial \Phi'}{\partial z}\right) \tag{10.24}$$

The quantity $\overline{q'v'}$on the right-hand side in (10.22) is the divergence of the meridional flux of potential vorticity. According to (10.22), for adiabatic quasi-geostrophic flow the mean distribution of potential vorticity can be changed only if there is a nonzero flux of eddy potential vorticity. The zonal-mean potential vorticity, together with suitable boundary conditions on $\overline{\Phi}$, completely determines the distribution of zonal-mean geopotential, and hence the zonal-mean geostrophic wind and temperature distributions. Thus, eddy-driven mean-flow accelerations require nonzero potential vorticity fluxes.

It can be shown that the potential vorticity flux is related to the eddy momentum and heat fluxes. We first note that for quasi-geostrophic motions the eddy horizontal velocities in the flux terms are geostrophic:

$$f_0 v' = \partial \Phi' \big/ \partial x \ , \quad \text{and} \quad f_0 u' = -\partial \Phi' \big/ \partial y$$

Thus,

$$\overline{v'\frac{\partial^2 \Phi'}{\partial x^2}} = \frac{1}{f_0}\overline{\frac{\partial \Phi'}{\partial x}\frac{\partial^2 \Phi'}{\partial x^2}} = \frac{1}{2 f_0}\overline{\frac{\partial}{\partial x}\left(\frac{\partial \Phi'}{\partial x}\right)^2} = 0$$

where we have used the fact that a perfect differential in x vanishes when averaged zonally. Thus,

$$\overline{q'v'} = \overline{\frac{v'}{f_0}\frac{\partial^2 \Phi'}{\partial y^2}} + \frac{f_0}{\rho_0}\overline{v'\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\frac{\partial \Phi'}{\partial z}\right)}$$

We can use the chain rule of differentiation to rewrite the terms on the right in this expression as

$$\overline{\frac{v'}{f_0}\frac{\partial^2 \Phi'}{\partial y^2}} = \frac{1}{f_0^2}\left(\overline{\frac{\partial \Phi'}{\partial x}\frac{\partial^2 \Phi'}{\partial y^2}}\right) =$$

$$\frac{1}{f_0^2}\left[\frac{\partial}{\partial y}\left(\overline{\frac{\partial \Phi'}{\partial x}\frac{\partial \Phi'}{\partial y}}\right) - \frac{1}{2}\frac{\partial}{\partial x}\overline{\left(\frac{\partial \Phi'}{\partial y}\right)^2}\right] = -\frac{\partial}{\partial y}\left(\overline{u'v'}\right)$$

and

$$\frac{f_0}{\rho_0}\overline{v'\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\frac{\partial\Phi'}{\partial z}\right)} = \frac{1}{\rho_0}\left[\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\overline{\frac{\partial\Phi'}{\partial x}\frac{\partial\Phi'}{\partial z}}\right) - \frac{\rho_0}{2N^2}\frac{\partial}{\partial x}\overline{\left(\frac{\partial\Phi'}{\partial z}\right)^2}\right]$$
$$= \frac{f_0}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\overline{v'\frac{\partial\Phi'}{\partial z}}\right)$$

Thus,

$$\overline{q'v'} = -\frac{\partial\overline{u'v'}}{\partial y} + \frac{f_0}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\overline{v'\frac{\partial\Phi'}{\partial z}}\right) \tag{10.25}$$

so that it is not the momentum flux $\overline{u'v'}$ or heat flux $\overline{v'\partial\Phi'/\partial z}$ that drives net changes in the mean-flow distribution, but rather the combination given by the potential vorticity flux. Under some circumstances the eddy momentum flux and eddy heat flux may individually be large, but the combination in (10.25) actually vanishes. This cancellation effect makes the traditional Eulerian mean formulation a poor framework for analysis of mean-flow forcing by eddies.

Comparing (10.25) and (10.20) we see that the potential vorticity flux is proportional to the divergence of the EP flux vector:

$$\overline{q'v'} = \rho_0^{-1}\nabla\bullet\mathbf{F} \tag{10.26}$$

Thus, the contribution of large-scale motions to the zonal force in (10.17) equals the meridional flux of quasi-geostrophic potential vorticity. If the motion is adiabatic and the potential vorticity flux is nonzero, (10.22) shows that the mean-flow distribution must change in time. Thus, there cannot be complete compensation between Coriolis torque and zonal force terms in (10.17).

10.3 THE ANGULAR MOMENTUM BUDGET

The previous section used the quasi-geostrophic version of the zonal-mean equations to show that large-scale eddies play an essential part in the maintenance of the zonal-mean circulation in the extratropics. In particular, we contrasted the mean-flow forcing as represented by the conventional Eulerian mean and TEM formulations. This section expands our consideration of the momentum budget by considering the overall balance of *angular* momentum for the atmosphere and the earth combined. Thus, rather than simply considering the balance of momentum for a given latitude and height in the atmosphere, we must consider the transfer of angular momentum between the earth and the atmosphere, and the flow of angular momentum in the atmosphere.

It would be possible to utilize the complete spherical coordinate version of the TEM equations for this analysis, but we are concerned primarily with the angular momentum balance for a zonal ring of air extending from the surface to the top of the atmosphere. In that case it proves simpler to use the conventional Eulerian mean formulation. It also proves convenient to use a special vertical coordinate, called the *sigma coordinate*, in which the surface of the earth is a coordinate surface.

Because the average rotation rate of the earth is itself observed to be very close to constant, the atmosphere must also on the average conserve its angular momentum. The atmosphere gains angular momentum from the earth in the tropics where the surface winds are easterly (i.e., where the angular momentum of the atmosphere is less than that of the earth) and gives up angular momentum to the earth in middle latitudes where the surface winds are westerly. Thus, there must be a net poleward transport of angular momentum within the atmosphere; otherwise the torque due to surface friction would decelerate both easterlies and westerlies. Furthermore, the angular momentum given by the earth to the atmosphere in the belt of easterlies must just balance the angular momentum given up to the earth in the westerlies if the global angular momentum of the atmosphere is to remain constant.

In the equatorial regions the poleward angular momentum transport is divided between the advection of absolute angular momentum by the poleward flow in the axially symmetric Hadley circulation and transport by eddies. In midlatitudes, however, it is primarily the eddy motions that transport momentum poleward, and the angular momentum budget of the atmosphere must be qualitatively as shown in Fig. 10.10.

As Fig. 10.10 suggests, there is a maximum poleward flux of angular momentum at about 30° latitude and a maximum horizontal flux convergence at about 45°. This maximum in the flux convergence is a reflection of the strong energy conversion in upper-level westerlies and is the mechanism whereby the atmosphere can maintain

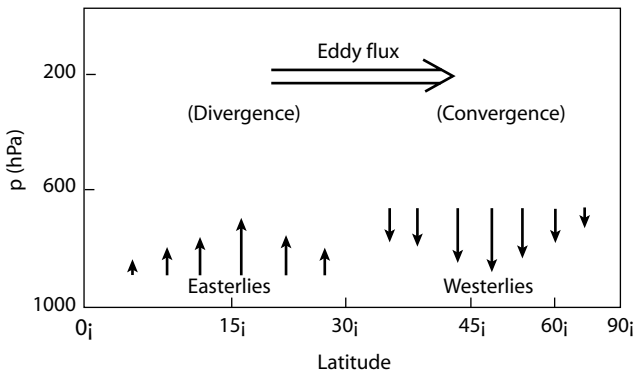


Fig. 10.10 Schematic mean angular momentum budget for the atmosphere–earth system.

a positive zonal wind in the middle latitudes despite the momentum lost to the surface.

It is convenient to analyze the momentum budget in terms of absolute angular momentum. The absolute angular momentum per unit mass of atmosphere is

$$M = (\Omega a \cos\phi + u)\, a \cos\phi$$

where as before a is the radius of the earth. The crucial role of zonal eddy drag in maintaining the observed latitudinal profile of the mean zonal wind can be seen by evaluating the mean zonal velocity that would arise if a zonal ring of air at rest on the equator were to be displaced poleward conserving M. In that case $u(\phi) = \Omega a \sin^2\phi \big/ \cos\phi$, so that in an angular momentum-conserving Hadley circulation, $u \approx 130\, m s^{-1}$ at 30° latitude, which is far greater than is observed.

Clearly, the absolute angular momentum must decrease as air parcels are advected poleward in the Hadley circulation. The absolute angular momentum of an individual air parcel can be changed only by torques caused by the zonal pressure gradient and eddy stresses. In isobaric coordinates, Newton's second law in its angular momentum form is thus

$$\frac{DM}{Dt} = -a\cos\phi \left[\frac{\partial \Phi}{\partial x} + g \frac{\partial \tau_E^x}{\partial p} \right] \tag{10.27}$$

where τ_E^x is the zonal component of the vertical eddy stress, and it is assumed that horizontal eddy stresses are negligible compared to the vertical eddy stress.

10.3.1 Sigma Coordinates

In neither the isobaric nor the log-pressure coordinate system does the lower boundary exactly coincide with a coordinate surface. In analytical studies it is usual to assume that the lower boundary can be approximated as a constant pressure surface and to apply the approximate condition

$$\omega(p_s) \approx -\rho_0 g w(z_0)$$

as the lower boundary condition. Here we have assumed that the height of the ground z_0 is coincident with the pressure surface p_s (where p_s is usually set equal to 1000 hPa). These assumptions are of course not strictly valid even when the ground is level. Pressure does change at the ground, but, more importantly, the height of the ground generally varies so that even if the pressure tendency were zero everywhere, the lower boundary condition should not be applied at a constant p_s. Rather, we should set $p_s = p_s(x, y)$. It is very inconvenient for mathematical analysis, however, to have a boundary condition that must be applied at a surface that is a function of the horizontal variables.

This problem can be overcome by defining a vertical coordinate that is proportional to pressure normalized by the surface pressure. The most common form of this coordinate is the *sigma coordinate,* defined as $\sigma \equiv p/p_s$, where $p_s(x, y, t)$ is the pressure at the surface. Thus, σ is a nondimensional-independent vertical coordinate that decreases upward from a value of $\sigma = 1$ at the ground to $\sigma = 0$ at the top of the atmosphere. In sigma coordinates the lower boundary condition will always apply exactly at $\sigma = 1$. Furthermore, the vertical σ velocity defined by

$$\dot{\sigma} \equiv D\sigma \big/ Dt$$

will always be zero at the ground even in the presence of sloping terrain. Thus, the lower boundary condition in the σ system is merely

$$\dot{\sigma} = 0 \text{ at } \sigma = 1$$

To transform the dynamical equations from the isobaric system to the σ system we first transform the pressure gradient force in a manner analogous to that shown in Section 1.6.3. Applying (1.28) with p replaced by Φ, s replaced by σ, and z replaced by p, we find that

$$\left(\partial\Phi \big/ \partial x\right)_\sigma = \left(\partial\Phi \big/ \partial x\right)_p + \sigma\left(\partial \ln p_s \big/ \partial x\right)\left(\partial\Phi \big/ \partial\sigma\right) \tag{10.28}$$

Because any other variable will transform in an analogous way, we can write the general transformation as

$$\nabla_p(\ ) = \nabla_\sigma(\ ) - \sigma\nabla \ln p_s \partial(\ ) \big/ \partial\sigma \tag{10.29}$$

Applying the transformation (10.29) to the momentum equation (3.2), we get

$$\frac{D\mathbf{V}}{Dt} + f\mathbf{k} \times \mathbf{V} = -\nabla\Phi + \sigma\nabla \ln p_s \frac{\partial\Phi}{\partial\sigma} \tag{10.30}$$

where ∇ is now applied holding σ constant, and the total differential is

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla + \dot{\sigma}\frac{\partial}{\partial\sigma} \tag{10.31}$$

The equation of continuity can be transformed to the σ system by first using (10.29) to express the divergence of the horizontal wind as

$$\nabla_p \bullet \mathbf{V} = \nabla_\sigma \bullet \mathbf{V} - \sigma\left(\nabla \ln p_s\right) \bullet \partial\mathbf{V} \big/ \partial\sigma \tag{10.32}$$

To transform the term $\partial\omega/\partial p$ we first note that since p_s does not depend on σ

$$\frac{\partial}{\partial p} = \frac{\partial}{\partial\left(\sigma p_s\right)} = \frac{1}{p_s}\frac{\partial}{\partial\sigma}$$

Thus, the continuity equation (3.5) can be written as

$$p_s\left(\nabla_p \bullet \mathbf{V}\right) + \partial\omega\big/\partial\sigma = 0 \tag{10.33}$$

Now the sigma vertical velocity can be written as

$$\dot{\sigma} = \left(\frac{\partial\sigma}{\partial t} + \mathbf{V}\bullet\nabla\sigma\right)_p + \omega\frac{\partial\sigma}{\partial p} = -\frac{\sigma}{p_s}\left(\frac{\partial p_s}{\partial t} + \mathbf{V}\bullet\nabla p_s\right) + \frac{\omega}{p_s}$$

Differentiating the above with respect to σ, eliminating $\partial\omega\big/\partial\sigma$ with (10.33), and rearranging yields the transformed continuity equation

$$\frac{\partial p_s}{\partial t} + \nabla\bullet\left(p_s\mathbf{V}\right) + p_s\frac{\partial\dot{\sigma}}{\partial\sigma} = 0 \tag{10.34}$$

With the aid of the equation of state and Poisson's equation (2.44) the hydrostatic approximation can be written in the sigma system as

$$\frac{\partial\Phi}{\partial\sigma} = -\frac{RT}{\sigma} = -\frac{R\theta}{\sigma}\left(p\big/p_0\right)^{\kappa} \tag{10.35}$$

where $p_0 = 1000\,\text{hPa}$.

Expanding the total derivative in (2.46), we may write the thermodynamic energy equation for sigma coordinates as

$$\frac{\partial\theta}{\partial t} + \mathbf{V}\bullet\nabla\theta + \dot{\sigma}\frac{\partial\theta}{\partial\sigma} = \frac{J}{c_p}\frac{\theta}{T} \tag{10.36}$$

10.3.2 The Zonal-Mean Angular Momentum

We now transform the angular momentum equation (10.27) into sigma coordinates with the aid of (10.28) and (10.35) to yield

$$\left(\frac{\partial}{\partial t} + \mathbf{V}\bullet\nabla + \dot{\sigma}\frac{\partial}{\partial\sigma}\right)M = -a\cos\phi\left(\frac{\partial\Phi}{\partial x} + \frac{RT}{p_s}\frac{\partial p_s}{\partial x} + \frac{g}{p_s}\frac{\partial\tau_E^x}{\partial\sigma}\right) \tag{10.37}$$

Multiplying the continuity equation (10.34) by M and adding the result to (10.37) multiplied by p_s we obtain the flux form of the angular momentum equation:[1]

$$\begin{aligned}\frac{\partial\left(p_s M\right)}{\partial t} = &-\nabla\bullet\left(p_s M\mathbf{V}\right) - \frac{\partial\left(p_s M\dot{\sigma}\right)}{\partial\sigma}\\ &- a\cos\phi\left[p_s\frac{\partial\Phi}{\partial x} + RT\frac{\partial p_s}{\partial x}\right] - ga\cos\phi\frac{\partial\tau_E^x}{\partial\sigma}\end{aligned} \tag{10.38}$$

[1] It may be shown (Problem 10.2) that in sigma coordinates the mass element $\rho_0 dxdydz$ takes the form $-g^{-1}p_s\mathrm{d}xdyd\sigma$. Thus, p_s in sigma space plays a role similar to density in physical space.

To obtain the zonal-mean angular momentum budget, we must average (10.38) in longitude. Using the spherical coordinate expansion for the horizontal divergence as given in Appendix C, we have

$$\nabla \bullet (p_s M\mathbf{V}) = \frac{1}{a\cos\phi}\left[\frac{\partial\,(p_s M u)}{\partial\lambda} + \frac{\partial\,(p_s M v\cos\phi)}{\partial\phi}\right] \tag{10.39}$$

We also observe that the bracketed term on the right in (10.38) can be rewritten as

$$\left[p_s\frac{\partial}{\partial x}(\Phi - RT) + \frac{\partial}{\partial x}(p_s RT)\right] \tag{10.40}$$

However, with the aid of the hydrostatic equation (10.35) we can write

$$(\Phi - RT) = \Phi + \sigma\partial\Phi\big/\partial\sigma = \partial\,(\sigma\Phi)\big/\partial\sigma$$

Thus, recalling that p_s does not depend on σ, we obtain

$$\left[p_s\frac{\partial\Phi}{\partial x} + RT\frac{\partial p_s}{\partial x}\right] = \left[\frac{\partial}{\partial\sigma}\left(p_s\sigma\frac{\partial\Phi}{\partial x}\right) + \frac{\partial}{\partial x}(p_s RT)\right] \tag{10.41}$$

Substituting from (10.39) and (10.41) into (10.38) and taking the zonal average gives

$$\begin{aligned}\frac{\partial\left(\overline{p_s M}\right)}{\partial t} = &-\frac{1}{\cos\phi}\frac{\partial}{\partial y}\left(\overline{p_s M v}\cos\phi\right)\\ &-\frac{\partial}{\partial\sigma}\left[\overline{p_s M\dot{\sigma}} + ga\cos\phi\left(\overline{\tau}_E^x\right) + (a\cos\phi)\,\overline{\sigma p_s\frac{\partial\Phi}{\partial x}}\right]\end{aligned} \tag{10.42}$$

The terms on the right in (10.42) represent the convergence of the horizontal flux of angular momentum and the convergence of the vertical flux of angular momentum, respectively.

Integrating (10.42) vertically from the surface of the earth ($\sigma = 1$) to the top of the atmosphere ($\sigma = 0$), and recalling that $\dot{\sigma} = 0$ for $\sigma = 0, 1$ we have

$$\begin{aligned}\int_0^1 g^{-1}\frac{\partial}{\partial t}\overline{p_s M}d\sigma = &-(g\cos\phi)^{-1}\int_0^1\frac{\partial}{\partial y}\left(\overline{p_s M v}\cos\phi\right)d\sigma\\ &-a\cos\phi\left[\left(\overline{\tau_E^x}\right)_{\sigma=1} + \overline{p_s\partial h\big/\partial x}\right]\end{aligned} \tag{10.43}$$

where $h(x, y) = g^{-1}\Phi(x, y, 1)$ is the height of the lower boundary ($\sigma = 1$), and we have assumed that the eddy stress vanishes at $\sigma = 0$.

Equation (10.43) expresses the angular momentum budget for a zonal ring of air of unit meridional width, extending from the ground to the top of the atmosphere. In the long-term mean the three terms on the right, representing the convergence of the meridional flux of angular momentum, the torque due to small scale turbulent fluxes at the surface, and the surface pressure torque, must balance. In the sigma coordinate system the surface pressure torque takes the particularly simple form $-\overline{p_s \partial h / \partial x}$. Thus, the pressure torque acts to transfer angular momentum from the atmosphere to the ground, provided that the surface pressure and the slope of the ground ($\partial h/\partial x$) are correlated positively. Observations indicate that this is generally the case in middle latitudes because there is a slight tendency for the surface pressure to be higher on the western sides of mountains than on the eastern sides (see Fig. 4.9). In midlatitudes of the Northern Hemisphere the surface pressure torque provides nearly one-half of the total atmosphere–surface momentum exchange; but in the tropics and the Southern Hemisphere the exchange is dominated by turbulent eddy stresses.

The role of eddy motions in providing the meridional angular momentum transport necessary to balance the surface angular momentum sinks can be elucidated best if we divide the flow into zonal-mean and eddy components by letting

$$M = \overline{M} + M' = \left(\Omega a \cos\phi + \overline{u} + u'\right) a \cos\phi$$
$$p_s\, v = \overline{(p_s\, v)} + (p_s\, v)'$$

where primes indicate deviations from the zonal mean. Thus, the meridional flux becomes

$$\overline{(p_s\, Mv)} = \left[\Omega a \cos\phi\, \overline{p_s\, v} + \overline{u}\, \overline{p_s\, v} + \overline{u'\, (p_s\, v)'}\right] a \cos\phi \tag{10.44}$$

The three meridional flux terms on the right in this expression are called the meridional Ω–momentum flux, the meridional drift, and the meridional eddy momentum flux, respectively.

The drift term is important in the tropics, but in midlatitudes it is small compared to the eddy flux and can be neglected in an approximate treatment. Furthermore, we can show that the meridional Ω–momentum flux does not contribute to the vertically integrated flux. Averaging the continuity equation (10.34) zonally and integrating vertically, we obtain

$$\overline{\frac{\partial p_s}{\partial t}} = -\,(\cos\phi)^{-1} \frac{\partial}{\partial y} \int_0^1 \overline{p_s v} \cos\phi\, d\sigma \tag{10.45}$$

Thus, for time-averaged flow [where the left-hand side of (10.45) vanishes] there is no net mass flow across latitude circles. The vertically integrated meridional

angular momentum flux is therefore given approximately by

$$\int_0^1 \overline{p_s M v} d\sigma \approx \int_0^1 \overline{u' (p_s v)'} a \cos\phi d\sigma \approx \int_0^1 a \cos\phi \overline{p}_s \overline{u'v'} d\sigma \qquad (10.46)$$

where we have assumed that the fractional change in p_s is small compared to the change in v' so that $(p_s v)' \approx \overline{p}_s v'$. The angular momentum flux is thus proportional to the negative of the meridional component of the EP flux given in (10.20).

In the Northern Hemisphere, as shown in Fig. 10.6, the eddy momentum flux in midlatitudes is positive and decreases in magnitude with latitude poleward of 30°. For quasi-geostrophic flow positive eddy momentum flux requires that the eddies be asymmetric in the horizontal plane with the trough and ridge axes tilting as indicated in Fig. 10.11. When the troughs and ridges on the average have southwest to northeast phase tilt, the zonal flow will be larger than average ($u' > 0$) where the meridional flow is poleward ($v' > 0$) and less than average ($u' < 0$) where the flow is equatorward. Thus, $\overline{u'v'} > 0$ and the eddies will systematically transport positive zonal momentum poleward.

As shown in (10.42) the total vertical momentum flux consists of the flux due to large-scale motions $\overline{p_s M \dot{\sigma}} \approx \Omega a \cos\phi \overline{p_s \dot{\sigma}}$, the flux due to the pressure torque, $(a \cos\phi) \overline{\sigma p_s \partial \Phi / \partial x}$, and the flux due to small-scale turbulent stresses, $ga \cos\phi \overline{\tau_E^x}$. As mentioned previously, the last two are responsible for the transfer of momentum from the earth to the atmosphere in the tropics and from the atmosphere to the earth in midlatitudes. Outside the planetary boundary layer, however, the vertical momentum transport in the troposphere is due primarily to the vertical Ω–momentum flux, $\Omega a \cos\phi \overline{p_s \dot{\sigma}}$.

An estimate of the annual average of the zonal-mean total surface torque is shown in Fig. 10.12. The total surface torque must be balanced by poleward fluxes of angular momentum in the atmosphere. Except for the belt within 10° of the equator, almost all of the poleward flux is accounted for by the eddy flux term given on the right-hand side of (10.46). Thus, the momentum budget and the energy cycle both depend critically on the transports by the eddies.

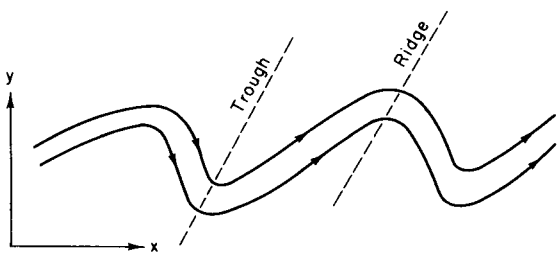


Fig. 10.11 Schematic streamlines for a positive eddy momentum flux.

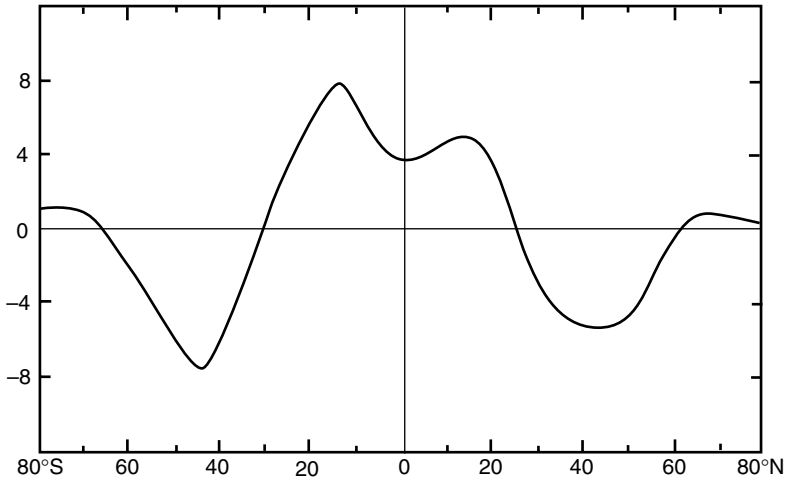


Fig. 10.12 Latitudinal profile of the annual mean eastward torque (surface friction plus mountain torque) exerted on the atmosphere in units of $10^{18}\,\mathrm{m}^2\,\mathrm{kg\,s}^{-2}$. (Adapted from Oort and Peixoto, 1983.)

10.4 THE LORENZ ENERGY CYCLE

The previous section discussed the interaction between zonally averaged flow and longitudinally varying eddy motions in terms of the angular momentum balance. It is also useful to examine the exchange of energy between the eddies and the mean flow. As in Section 10.2, we limit the analysis to quasi-geostrophic flow on a midlatitude β plane. Eulerian mean equations in log-pressure coordinates can then be written as

$$\partial\overline{u}\,/\partial t\,-\,f_0\overline{v} = -\partial\left(\overline{u'v'}\right)/\partial y + \overline{X} \tag{10.47}$$

$$f_0\overline{u} = -\partial\overline{\Phi}\,/\partial y \tag{10.48}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial\overline{\Phi}}{\partial z}\right) + \overline{w}N^2 = \frac{\kappa}{H}\overline{J} - \frac{\partial}{\partial y}\left(\overline{v'\frac{\partial\Phi'}{\partial z}}\right) \tag{10.49}$$

$$\partial\overline{v}\,/\partial y + \rho_0^{-1}\partial\left(\rho_0\overline{w}\right)/\partial z = 0 \tag{10.50}$$

Here we have used the hydrostatic approximation in (10.3) to express the temperature in (10.49) in terms of the differential geopotential thickness, and we have again neglected vertical eddy fluxes and advection by the mean meridional circulation in (10.47) and (10.49). We have, however, included a turbulent drag force, $\overline{X}$, in (10.47), as dissipation by unresolved turbulent eddies is an essential element in the energy balance.

In order to analyze the exchange of energy between mean flow and eddies, we require a similar set of dynamical equations for the eddy motion. For simplicity we assume that the eddies satisfy the following linearized set of equations[2] :

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)u'-\left(f_0-\frac{\partial\overline{u}}{\partial y}\right)v'=-\frac{\partial\Phi'}{\partial x}+X' \tag{10.51}$$

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)v'+f_0u'=-\frac{\partial\Phi'}{\partial y}+Y' \tag{10.52}$$

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)\frac{\partial\Phi'}{\partial z}+v'\frac{\partial}{\partial y}\left(\frac{\partial\overline{\Phi}}{\partial z}\right)+N^2w'=\frac{\kappa J'}{H} \tag{10.53}$$

$$\frac{\partial u'}{\partial x}+\frac{\partial v'}{\partial y}+\frac{1}{\rho_0}\frac{\partial\left(\rho_0w'\right)}{\partial z}=0 \tag{10.54}$$

where X' and Y' are the zonally varying components of drag due to unresolved turbulent motions.

We now define a global average

$$\langle\ \rangle\equiv A^{-1}\int_0^{\infty}\int_0^{D}\int_0^{L}(\ )\,dxdydz$$

where L is the distance around a latitude circle, D is the meridional extent of the midlatitude β plane, and A designates the total horizontal area of the β plane. Then for any quantity Ψ

$$\begin{aligned}
&\left\langle\partial\Psi\,/\partial x\right\rangle=0\\
&\left\langle\partial\Psi\,/\partial y\right\rangle=0, \quad \text{if } \Psi \text{ vanishes at } y=\pm D\\
&\left\langle\partial\Psi\,/\partial z\right\rangle=0, \quad \text{if } \Psi \text{ vanishes at } z=0 \text{ and } z\to\infty
\end{aligned}$$

An equation for the evolution of the mean flow kinetic energy can then be obtained by multiplying (10.47) by $\rho_0\overline{u}$ and (10.48) by $\rho_0\overline{v}$ and adding the results to get

$$\begin{aligned}
\rho_0\frac{\partial}{\partial t}\left(\frac{\overline{u}^2}{2}\right)&=-\rho_0\overline{v}\frac{\partial\overline{\Phi}}{\partial y}-\rho_0\overline{u}\frac{\partial}{\partial y}\left(\overline{u'v'}\right)+\rho_0\overline{u}\overline{X}\\
&=-\frac{\partial}{\partial y}\left(\rho_0\overline{v}\overline{\Phi}\right)+\rho_0\overline{\Phi}\frac{\partial\overline{v}}{\partial y}-\frac{\partial}{\partial y}\left(\rho_0\overline{u}\overline{u'v'}\right)+\rho_0\overline{u'v'}\frac{\partial\overline{u}}{\partial y}+\rho_0\overline{u}\overline{X}
\end{aligned}$$

[2] A similar analysis can be carried out for the nonlinear case.

Integrating over the entire volume we get

$$\frac{d}{dt}\left\langle \frac{\rho_0 \overline{u}^2}{2} \right\rangle = +\left\langle \rho_0 \overline{\Phi} \frac{\partial \overline{v}}{\partial y} \right\rangle + \left\langle \rho_0 \overline{u'v'} \frac{\partial \overline{u}}{\partial y} \right\rangle + \left\langle \rho_0 \overline{u} \overline{X} \right\rangle \tag{10.55}$$

where we have assumed that $\overline{v} = 0$ and $\overline{u'v'} = 0$ for $y = \pm D$. Terms on the right-hand side in (10.55) can be interpreted as the work done by the zonal-mean pressure force, the conversion of eddy kinetic energy to zonal-mean kinetic energy, and dissipation by the zonal-mean eddy stress. Alternatively, the first term on the right can be rewritten with the aid of the continuity equation to yield

$$\left\langle \rho_0 \overline{\Phi} \frac{\partial \overline{v}}{\partial y} \right\rangle = -\left\langle \overline{\Phi} \frac{\partial \rho_0 \overline{w}}{\partial z} \right\rangle = \left\langle \rho_0 \overline{w} \frac{\partial \overline{\Phi}}{\partial z} \right\rangle = \frac{R}{H} \left\langle \rho_0 \overline{w} \overline{T} \right\rangle$$

where we have assumed that $\rho_0 \overline{w} = 0$ at $z = 0$, and $z \to \infty$. Thus, averaged over the whole domain the pressure work term is proportional to the correlation between the zonal-mean vertical mass flux $\rho_0 \overline{w}$ and the zonal-mean temperature (or thickness). This term will be positive if on the average warm air is rising and cold air sinking, that is, if there is a conversion from potential to kinetic energy.

Section 8.3.1 showed that in the quasi-geostrophic system the available potential energy is proportional to the square of the deviation of temperature from a standard atmosphere profile divided by the static stability. In terms of differential thickness the zonal-mean available potential energy is defined as

$$\overline{P} \equiv \frac{1}{2} \left\langle \frac{\rho_0}{N^2} \left(\frac{\partial \overline{\Phi}}{\partial z} \right)^2 \right\rangle$$

Multiplying (10.49) through by $\rho_0 \left(\partial \overline{\Phi} / \partial z \right) / N^2$ and averaging over space gives

$$\begin{aligned} \frac{d}{dt} \left\langle \frac{\rho_0}{2N^2} \left(\frac{\partial \overline{\Phi}}{\partial z} \right)^2 \right\rangle = &-\left\langle \rho_0 \overline{w} \frac{\partial \overline{\Phi}}{\partial z} \right\rangle + \left\langle \frac{\rho_0 \kappa \bar{J}}{N^2 H} \left(\frac{\partial \overline{\Phi}}{\partial z} \right) \right\rangle \\ &- \left\langle \frac{\rho_0}{N^2} \frac{\partial \overline{\Phi}}{\partial z} \frac{\partial}{\partial y} \left(\overline{v' \frac{\partial \Phi'}{\partial z}} \right) \right\rangle \end{aligned} \tag{10.56}$$

The first term on the right is just equal and opposite to the first term on the right in (10.55), which confirms that this term represents a conversion between zonal-mean kinetic and potential energies. The second term involves the correlation between temperature and diabatic heating; it expresses the generation of zonal-mean potential energy by diabatic processes. The final term, which involves the

meridional eddy heat flux, expresses the conversion between zonal-mean and eddy potential energy.

That the second term on the right in (10.55) and the final term in (10.56) represent conversion between zonal-mean and eddy energies can be confirmed by performing analogous operations on the eddy equations (10.51)–(10.53) to obtain equations for the eddy kinetic and available potential energies:

$$\begin{aligned}\frac{d}{dt}\left\langle \rho_0 \frac{\overline{u'^2}+\overline{v'^2}}{2}\right\rangle = &+\left\langle \rho_0 \overline{\Phi'\left(\frac{\partial u'}{\partial x}+\frac{\partial v'}{\partial y}\right)}\right\rangle \\ &-\left\langle \rho_0 \overline{u'v'}\frac{\partial \overline{u}}{\partial y}\right\rangle + \left\langle \rho_0 \left(\overline{u'X'}+\overline{v'Y'}\right)\right\rangle\end{aligned} \tag{10.57}$$

$$\begin{aligned}\frac{d}{dt}\left\langle \frac{\rho_0}{2N^2}\overline{\left(\frac{\partial \Phi'}{\partial z}\right)^2}\right\rangle = &-\left\langle \rho_0 \overline{w'\frac{\partial \Phi'}{\partial z}}\right\rangle + \left\langle \frac{\rho_0 \kappa \overline{J'\partial \Phi'/\partial z}}{N^2 H}\right\rangle \\ &-\left\langle \frac{\rho_0}{N^2}\left(\frac{\partial^2 \overline{\Phi}}{\partial z \partial y}\right)\left(\overline{v'\frac{\partial \Phi'}{\partial z}}\right)\right\rangle\end{aligned} \tag{10.58}$$

If we set $w' = 0$ at $z = 0$, the first term on the right in (10.57) can be rewritten using the continuity equation (10.54) as

$$\left\langle \rho_0 \overline{\Phi'\left(\frac{\partial u'}{\partial x}+\frac{\partial v'}{\partial y}\right)}\right\rangle = -\left\langle \overline{\Phi'\frac{\partial(\rho_0 w')}{\partial z}}\right\rangle = \left\langle \rho_0 \overline{w'\left(\frac{\partial \Phi'}{\partial z}\right)}\right\rangle$$

which is equal to minus the first term on the right in (10.58). Thus, this term expresses the conversion between eddy kinetic and eddy potential energy for the Eulerian mean formulation. Similarly, the last term in (10.58) is equal to minus the last term in (10.56), and thus represents conversion between eddy and zonal-mean available potential energy.

The Lorenz energy cycle can be expressed compactly by defining zonal-mean and eddy kinetic and available potential energies:

$$\overline{K} \equiv \left\langle \rho_0 \frac{\overline{u}^2}{2}\right\rangle, \quad K' \equiv \left\langle \rho_0 \frac{\overline{u'^2}+\overline{v'^2}}{2}\right\rangle,$$

$$\overline{P} \equiv \frac{1}{2}\left\langle \frac{\rho_0}{N^2}\left(\frac{\partial \overline{\Phi}}{\partial z}\right)^2\right\rangle, \quad P' \equiv \frac{1}{2}\left\langle \frac{\rho_0}{N^2}\overline{\left(\frac{\partial \Phi'}{\partial z}\right)^2}\right\rangle$$

the energy transformations

$$\left[\overline{P}\bullet\overline{K}\right] \equiv \left\langle \rho_0 \overline{w}\frac{\partial \overline{\Phi}}{\partial z}\right\rangle, \quad \left[P'\bullet K'\right] \equiv \left\langle \rho_0 \overline{w'\frac{\partial \Phi'}{\partial z}}\right\rangle,$$

$$\left[K' \bullet \overline{K}\right] \equiv \left\langle \rho_0 \overline{u'v'} \frac{\partial \overline{u}}{\partial y} \right\rangle, \quad \left[P' \bullet \overline{P}\right] \equiv \left\langle \frac{\rho_0}{N^2} \overline{v' \frac{\partial \Phi'}{\partial z}} \frac{\partial^2 \overline{\Phi}}{\partial y \partial z} \right\rangle$$

and the sources and sinks

$$\overline{R} \equiv \left\langle \frac{\rho_0}{N^2} \frac{\kappa \overline{J}}{H} \frac{\partial \overline{\Phi}}{\partial z} \right\rangle, \quad R' \equiv \left\langle \frac{\rho_0}{N^2} \overline{\frac{\kappa J'}{H} \frac{\partial \Phi'}{\partial z}} \right\rangle,$$

$$\overline{\varepsilon} \equiv \left\langle \rho_0 \overline{u} \overline{X} \right\rangle, \quad \varepsilon' \equiv \left\langle \rho_0 \left(\overline{u'X'} + \overline{v'Y'} \right) \right\rangle$$

Equations (10.55)–(10.58) can then be expressed in the simple form

$$d\overline{K}/dt = \left[\overline{P} \bullet \overline{K}\right] + \left[K' \bullet \overline{K}\right] + \overline{\varepsilon} \tag{10.59}$$

$$d\overline{P}/dt = -\left[\overline{P} \bullet \overline{K}\right] + \left[P' \bullet \overline{P}\right] + \overline{R} \tag{10.60}$$

$$dK'/dt = \left[P' \bullet K'\right] - \left[K' \bullet \overline{K}\right] + \varepsilon' \tag{10.61}$$

$$dP'/dt = -\left[P' \bullet K'\right] - \left[P' \bullet \overline{P}\right] + R' \tag{10.62}$$

Here $[A \bullet B]$ designates conversion from energy form A to form B.

Adding (10.59)–(10.62), we obtain an equation for the rate of change of total energy (kinetic plus available potential):

$$d\left(\overline{K} + K' + \overline{P} + P'\right)/dt = \overline{R} + R' + \overline{\varepsilon} + \varepsilon' \tag{10.63}$$

For adiabatic inviscid flow the right side vanishes and the total energy $\overline{K} + K' + \overline{P} + P'$ is conserved. In this system the zonal-mean kinetic energy does not include a contribution from the mean meridional flow because the zonally averaged meridional momentum equation was replaced by the geostrophic approximation. (Likewise, use of the hydrostatic approximation means that neither the mean nor the eddy vertical motion is included in the total kinetic energy.) Thus, the quantities that are included in the total energy depend on the particular model used. For any model the definitions of energy must be consistent with the approximations employed.

In the long-term mean the left side of (10.63) must vanish. Thus, the production of available potential energy by zonal-mean and eddy diabatic processes must balance the mean plus eddy kinetic energy dissipation:

$$\overline{R} + R' = -\overline{\varepsilon} - \varepsilon' \tag{10.64}$$

Because solar radiative heating is a maximum in the tropics, where the temperatures are high, it is clear that $\overline{R}$, the generation of zonal-mean potential energy by the

zonal-mean heating, will be positive. For a *dry* atmosphere in which eddy diabatic processes are limited to radiation and diffusion R', the diabatic production of eddy available potential energy should be negative because the thermal radiation emitted to space from the atmosphere increases with increasing temperature and thus tends to reduce horizontal temperature contrasts in the atmosphere. For the earth's atmosphere, however, the presence of clouds and precipitation greatly alters the distribution of R'. Present estimates (see Fig. 10.13) suggest that in the Northern Hemisphere R' is positive and nearly half as large as $\overline{R}$. Thus, diabatic heating generates both zonal-mean and eddy available potential energy.

The equations (10.59)–(10.62) together provide a complete description of the quasi-geostrophic energy cycle from the conventional Eulerian mean point of view. The content of these equations is summarized by means of the four-box diagram of Fig. 10.13. In this diagram the squares represent reservoirs of energy and the arrows indicate sources, sinks, and conversions. The observed direction of the conversion terms in the troposphere for the Northern Hemisphere annual mean is indicated by arrows. It should be emphasized that the direction of the various conversions cannot be theoretically deduced by reference to the energy equations alone. It also should be emphasized that the conversion terms given here are a result of the particular type of zonal average model used. The analogous energy equations for the TEM equations have rather different conversions. Thus, the energy transformations given in the present analysis should not be regarded as fundamental properties of the atmosphere, but rather as properties of the Eulerian mean system.

Nevertheless, because the conventional Eulerian mean model is generally used as a basis for the study of baroclinic waves, the four-box energy diagram presented

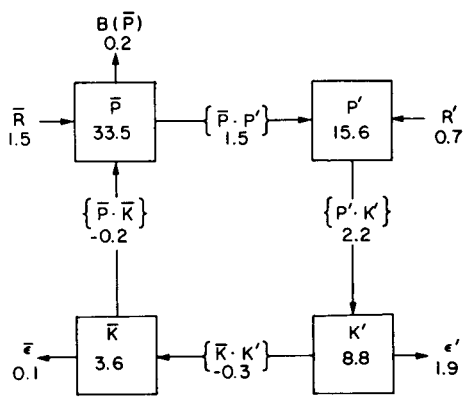


Fig. 10.13 The observed mean energy cycle for the Northern Hemisphere. Numbers in squares are energy amounts in units of 10^5 J m^{-2}. Numbers next to arrows are energy transformation rates in units of W m^{-2}. $B(\overline{p})$ represents a net energy flux into the Southern Hemisphere. Other symbols are defined in the text. (Adapted from Oort and Peixoto, 1974.)

here does provide a useful framework for considering the role of weather disturbances in maintenance of the general circulation. The observed energy cycle as summarized in Fig. 10.13 suggests the following qualitative picture:

1. The zonal-mean diabatic heating generates mean zonal available potential energy through a net heating of the tropics and cooling of the polar regions.
2. Baroclinic eddies transport warm air poleward, cold air equatorward, and transform the mean available potential energy to eddy available potential energy.
3. At the same time eddy available potential energy is transformed into eddy kinetic energy by the vertical motions in the eddies.
4. The zonal kinetic energy is maintained primarily by the conversions from eddy kinetic energy due to the correlation $\overline{u'v'}$. This is discussed further in the next section.
5. The energy is dissipated by surface and internal friction in the eddies and mean flow.

In summary, the observed atmospheric energy cycle as given by the Eulerian mean formulation is consistent with the notion that baroclinically unstable eddies are the primary disturbances responsible for the energy exchange in midlatitudes. It is through the eddy motions that the kinetic energy lost through turbulent stresses is replaced, and it is the eddies that are primarily responsible for the poleward heat transport to balance the radiation deficit in the polar regions. In addition to transient baroclinic eddies, forced stationary orographic waves and free Rossby waves may also contribute substantially to the poleward heat flux. The direct conversion of mean available potential energy to mean kinetic energy by symmetric overturning is, however, small and negative in middle latitudes, but positive in the tropics where it plays an important role in the maintenance of the mean Hadley circulation.

10.5 LONGITUDINALLY DEPENDENT TIME-AVERAGED FLOW

So far in this chapter we have concentrated on the zonally averaged component of the general circulation. For a planet with a longitudinally uniform surface the flow averaged over a season should be completely characterized by the zonally averaged component of the circulation since for such a hypothetical planet the statistics of zonally asymmetric transient eddies (i.e., weather disturbances) should be independent of longitude. On the earth, however, large-scale topography and continent–ocean heating contrasts provide strong forcing for longitudinally asymmetric planetary scale time-mean motions. Such motions, usually referred to as *stationary waves*, are especially strong in the Northern Hemisphere during the winter season.

Observations indicate that tropospheric stationary waves generally tend to have an equivalent barotropic structure; that is, wave amplitude generally increases

with height, but phase lines tend to be vertical. Although nonlinear processes may be significant in the formation and maintenance of stationary waves, the climatological stationary wave pattern can to a first approximation be described in terms of forced barotropic Rossby waves. When superposed on zonal-mean circulation, such waves produce local regions of enhanced and diminished time mean westerly winds, which strongly influence the development and propagation of transient weather disturbances. They thus represent essential features of the climatological flow.

10.5.1 Stationary Rossby Waves

The most significant of the time mean zonally asymmetric circulation features is the pattern of stationary planetary waves excited in the Northern Hemisphere by the flow over the Himalayas and the Rockies. It was shown in Section 7.7.2 that the quasi-stationary wave pattern along 45° latitude could be accounted for to a first approximation as the forced wave response when mean westerlies impinge on the large-scale topography. More detailed analysis suggests that zonally asymmetric heat sources also contribute to the forcing of the climatological stationary wave pattern. Some controversy remains, however, concerning the relative importance of heating and orography in forcing the observed stationary wave pattern. The two processes are difficult to separate because the pattern of diabatic heating is influenced by orography.

The discussion of topographic Rossby waves in Section 7.7.2 used a β-plane channel model in which it was assumed that wave propagation was parallel to latitude circles. In reality, however, large-scale topographic features and heat sources are confined in latitude as well as longitude, and the stationary waves excited by such forcing may propagate energy meridionally as well as zonally. For a quantitatively accurate analysis of the barotropic Rossby wave response to a local source it is necessary to utilize the barotropic vorticity equation in spherical coordinates and to include the latitudinal dependence of the mean zonal wind. The mathematical analysis for such a situation is beyond the scope of this book. It is possible, however, to obtain a qualitative notion of the nature of the wave propagation for this case by generalizing the β-plane analysis of Section 7.7. Thus, rather than assuming that propagation is limited to a channel of specified width, we assume that the β plane extends to plus and minus infinity in the meridional direction and that Rossby waves can propagate both northward and southward without reflection from artificial walls.

The free barotropic Rossby wave solution then has the form of (7.90) and satisfies the dispersion relation of (7.91) where l is the meridional wave number, which is now allowed to vary. From (7.93) it is clear that for a specified zonal wave number, k, the free solution is stationary for l given by

$$l^2 = \beta/\overline{u} - k^2 \tag{10.65}$$

Thus, for example, westerly flow over an isolated mountain that primarily excites a response at a given k will produce stationary waves with both positive and negative l satisfying (10.65). As remarked in Section 7.7.1, although Rossby wave phase propagation relative to the mean wind is always westward, this is not true of the group velocity. From (7.91) we readily find that the x and y components of group velocity are

$$c_{gx} = \frac{\partial \nu}{\partial k} = \overline{u} + \beta \frac{\left(k^2 - l^2\right)}{\left(k^2 + l^2\right)^2} \tag{10.66}$$

$$c_{gy} = \frac{\partial \nu}{\partial l} = \frac{2\beta k l}{\left(k^2 + l^2\right)^2} \tag{10.67}$$

For stationary waves, these may be expressed alternatively with the aid of (10.65) as

$$c_{gx} = \frac{2\overline{u}k^2}{\left(k^2 + l^2\right)}, \quad c_{gy} = \frac{2\overline{u}kl}{\left(k^2 + l^2\right)} \tag{10.68}$$

The group velocity vector for stationary Rossby waves is perpendicular to the wave crests. It always has an eastward zonal component and a northward or southward meridional component depending on whether l is positive or negative. The magnitude is given by

$$\left|\mathbf{c_g}\right| = 2\overline{u}\cos\alpha \tag{10.69}$$

(see Problem 10.9). Here, as shown in Fig. 10.14 for the case of positive l, α is the angle between lines of constant phase and the y axis.

Because energy propagates at the group velocity, (10.68) indicates that the stationary wave response to a localized topographic feature should consist of two wave trains, one with $l > 0$ extending eastward and northward and the other with $l < 0$ extending eastward and southward. An example computed using spherical geometry is given in Fig. 10.15. Although the positions of individual troughs and ridges remain fixed for stationary waves, the wave trains in this example do not decay in time, as the effects of dissipation are counteracted by energy propagation from the source at the Rossby wave group velocity.

For the climatological stationary wave distribution in the atmosphere the excitation comes from a number of sources, both topographic and thermal, distributed around the globe. Thus, it is not easy to trace out distinct paths of wave propagation. Nevertheless, detailed calculations using spherical geometry suggest that two-dimensional barotropic Rossby wave propagation provides a reasonable first approximation for the observed departure of the extratropical time-mean flow from zonal symmetry.

Rossby waves excited by isolated orographic features also play a significant role in the momentum budget. Letting the amplitude coefficient Ψ be real in (7.90), the

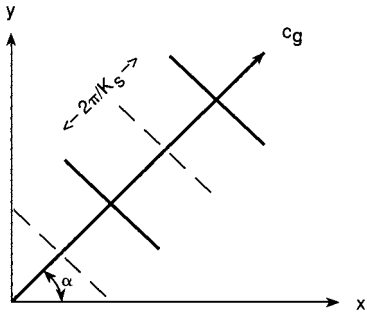


Fig. 10.14 Stationary plane Rossby wave in a westerly flow. Ridges (solid lines) and troughs (dashed lines) are oriented at an angle α to the y axis, and the group velocity relative to the ground, $\mathbf{c}_g$, is oriented at an angle α to the x axis. The wavelength is $2\,\pi/K_s$. (Adapted from Hoskins, 1983.)

meridional momentum flux can be expressed as

$$\overline{u'v'} = -\overline{\left(\partial\psi'/\partial x\right)\left(\partial\psi'/\partial y\right)} = -\Psi^2 kl/2$$

From (10.68) it is then verified readily that if $\overline{u} > 0$,

$$c_{gy} > 0 \text{ implies } \overline{u'v'} < 0$$
$$c_{gy} < 0 \text{ implies } \overline{u'v'} > 0$$

Thus, westerly momentum converges into the wave source region (where the energy flux is divergent). This eddy momentum flux convergence is necessary to balance the momentum lost to the surface through the pressure torque mechanism discussed in Section 10.3.

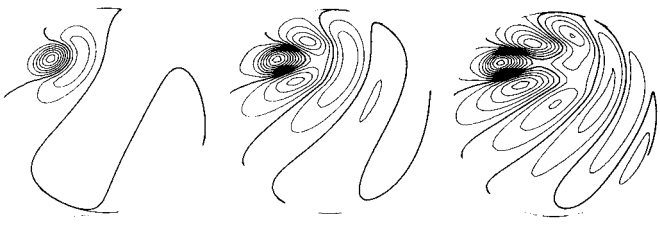


Fig. 10.15 The vorticity pattern generated on a sphere when a constant angular velocity westerly flow impinges on a circular forcing centered at 30°N and 45°W of the central point. Left to right, the response at 2, 4, and 6 days after switch on of the forcing. Five contour intervals correspond to the maximum vorticity response that would occur in 1 day if there were no wave propagation. Heavy lines correspond to zero contours. The pattern is drawn on a projection in which the sphere is viewed from infinity. (After Hoskins, 1983.)

10.5.2 Jetstream and Storm Tracks

When the longitudinally asymmetric geopotential anomalies associated with stationary waves are added to the zonal-mean geopotential distribution, the resulting time mean field includes local regions of enhanced meridional geopotential gradient that are manifested in the wind field of the Northern Hemisphere by the Asian and North American jetstreams. The existence of these two jets can be inferred from the January mean 500-hPa geopotential height field shown in Fig. 6.3. Note the strong meridional gradients in height associated with the troughs centered just east of the Asian and North American continents (the same features can be seen in annual mean charts, although with somewhat reduced intensity). The zonal flow associated with these semipermanent troughs is illustrated in Fig. 6.2. In addition to the two intense jet cores in the western Pacific and western Atlantic, there is a third weaker jet centered over North Africa and the Middle East. Figure 6.2 shows dramatically the large deviations from zonal symmetry in the jetstream structure. In midlatitudes the zonal wind speed varies by nearly a factor of three between the core of the Asian jet and the low wind speed area in western North America. Although, as was argued earlier, the climatological stationary wave distribution on which the Asian and North American jets are superposed is apparently forced primarily by orography, the structure of the jets also appears to be influenced by continent–ocean heating contrasts. Thus, the strong vertical shear in Asian and North American jets reflects a thermal wind balance consistent with the very strong meridional temperature gradients that occur in winter near the eastern edges of the Asian and North American continents due to the contrast between warm water to the southeast and cold land to the northwest. A satisfactory description of the jetstreams must account, however, not only for their thermal structure, but for the westerly acceleration that air parcels must experience as they enter the jet, and the deceleration as they leave the jet core.

To understand the momentum budget in the jetstreams and its relationship to the observed distribution of weather, we consider the zonal component of the momentum equation, which (if we neglect the β effect) may be written in the form of (6.38):

$$\frac{D_g u_g}{Dt} = f_0 \left(v - v_g\right) \equiv f_0 v_a \tag{10.70}$$

where v_a is the meridional component of the ageostrophic wind. This equation indicates that the westerly acceleration $(Du_g/Dt) > 0)$ that air parcels experience as they enter the jet can only be provided by a poleward ageostrophic wind component ($v_a > 0$), and conversely, the easterly acceleration that air parcels experience as they leave the jet requires an equatorward ageostrophic motion. This meridional flow, together with its accompanying vertical circulation, is illustrated in Fig. 10.16. Note that this secondary circulation is thermally direct upstream of the jet core. A magnitude of $v_a \sim 2-3$ ms^{-1} is required to account for the observed zonal wind

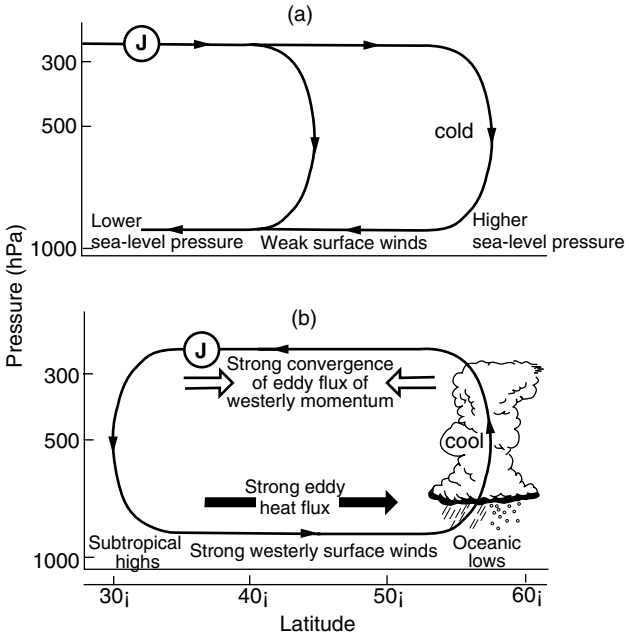


Fig. 10.16 Meridional cross sections showing the relationship between the time mean secondary meridional circulation (continuous thin lines with arrows) and the jet stream (denoted by J) at locations (a) upstream and (b) downstream from the jet stream cores. (After Blackmon et al., 1977. Reproduced with permission of the American Meteorological Society.)

acceleration. This is an order of magnitude stronger than the zonal-mean indirect cell (Ferrel cell) that prevails in midlatitudes. Downstream of the jet core, however, the secondary circulation is thermally indirect, but much stronger than the zonally averaged Ferrel cell. It is interesting to note that the vertical motion pattern on the poleward (cyclonic shear) side of the jet is similar to that associated with deep transient baroclinic eddies in the sense that subsidence occurs to the west of the stationary trough associated with the jet, and ascent occurs east of the trough (see, e.g., Fig. 6.12).

Because the growth rate of baroclinically unstable synoptic scale disturbances is proportional to the strength of the basic state thermal wind, it is not surprising that the Pacific and Atlantic jet streams are important source regions for storm development. Typically, transient baroclinic waves develop in the jet entrance region, grow as they are advected downstream, and decay in the jet exit region. The role of these transient eddies in maintenance of the jetstream structure is rather complex. Transient eddy heat fluxes, which are strong and poleward in the storm tracks, appear to act to weaken the climatological jets. The transient eddy vorticity

flux in the upper troposphere, however, appears to act to maintain the jet structure. In both cases the secondary ageostrophic circulation associated with the jet tends to partly balance the influence of the transient eddy fluxes in order to maintain the mean thermal and momentum balances.

10.6 LOW-FREQUENCY VARIABILITY

An understanding of the general circulation requires consideration not only of the zonal-mean and stationary wave components and their variations with the annual cycle, but also of irregular variability on time scales longer than that of individual transient eddies. The term *low-frequency variability* is generally used to describe such components of the general circulation. The observed spectrum of low-frequency variability ranges from weather anomalies lasting only 7–10 days to interannual variability on the scale of several years (see Section 11.1.6).

One possible cause of atmospheric low-frequency variability is forcing due to anomalies in sea surface temperature (SST), which themselves arise from coherent air–sea interaction. Because of the large thermal inertia of the oceanic surface mixed layer, such anomalies tend to have time scales much longer than those associated with subseasonal variations in the atmosphere; they are thought to be of greatest significance on the seasonal and interannual time scales.

It is believed, however, that substantial variability on subseasonal timescales can arise in midlatitudes in the absence of anomalous SST forcing as a result of internal nonlinear atmospheric dynamics, although SST anomalies may tend to favor the occurrence of certain types of variations. One example of internally generated low-frequency variability is the forcing of large-scale anomalies by potential vorticity fluxes of high-frequency transient waves. This process appears to be important in the maintenance of high-amplitude quasi-stationary wave disturbances called *blocking* patterns. Some types of blocking may also be related to special nonlinear wave patterns called *solitary waves*, in which damping by Rossby wave dispersion is balanced by intensification due to nonlinear advection. Although most internal mechanisms involve nonlinearity, there is some evidence that the longitudinally dependent time-mean flow may be unstable to linear barotropic normal modes that are stationary in space, but oscillate in time at low frequencies. Such modes, which are global in scale, may be responsible for some observed teleconnection patterns.

10.6.1 Climate Regimes

It has long been observed that extratropical circulation appears to alternate between a so-called high-index state, corresponding to a circulation with strong zonal flow and weak waves, and a "low-index" state with weak zonal flow and high-amplitude waves. This behavior suggests that more than one climate regime exists consistent with a given external forcing and that the observed climate may switch back and

forth between regimes in a chaotic fashion. Whether the high-index and low-index states actually correspond to distinct quasi-stable atmospheric climate regimes is a matter of controversy. The general notion of vacillation between two quasi-stable flow regimes can, however, be demonstrated quite convincingly in laboratory experiments. Such experiments are described briefly in Section 10.7.

The concept of climate regimes can also be demonstrated in a highly simplified model of the atmosphere developed by Charney and DeVore (1979). They examined the equilibrium mean states that can result when a damped topographic Rossby wave interacts with the zonal-mean flow. Their model is an extension of the topographic Rossby wave analysis given in Section 7.7.2. In this model the wave disturbance is governed by (7.99), which is the linearized form of the barotropic vorticity equation (7.94) with weak damping added. The zonal-mean flow is governed by the barotropic momentum equation

$$\frac{\partial \overline{u}}{\partial t} = -D\left(\overline{u}\right) - \kappa\left(\overline{u} - U_e\right) \tag{10.71}$$

where the first term on the right designates forcing by interaction between the waves and the mean flow, and the second term represents a linear relaxation toward an externally determined basic state flow, U_e.

The zonal mean equation (10.71) can be obtained from (7.94) by dividing the flow into zonal mean and eddy parts and taking the zonal average to get

$$\frac{\partial}{\partial t}\left(-\frac{\partial \overline{u}}{\partial y}\right) = -\frac{\partial}{\partial y}\left(\overline{v_g' \zeta_g'}\right) - \frac{f_0}{H}\frac{\partial}{\partial y}\left(\overline{v_g' h_T}\right)$$

which after integrating in y and adding the external forcing term yields (10.71) with

$$D\left(\overline{u}\right) = -\overline{v_g' \zeta_g'} - \left(f_0 / H\right)\overline{v_g' h_T} \tag{10.72}$$

As shown in Problem 10.5, the eddy vorticity flux [the first term on the right in (10.72)] is proportional to the divergence of the eddy momentum flux. The second term, which is sometimes referred to as the *form drag*, is the equivalent in the barotropic model of the surface pressure torque term in the angular momentum balance equation (10.43).

If h_T and the eddy geostrophic streamfunction are assumed to consist of single harmonic wave components in x and y, as given by (7.96) and (7.97), respectively, the vorticity flux vanishes, and with the aid of (7.100) the form drag can be expressed as

$$D\left(\overline{u}\right) = -\left(\frac{f_0}{H}\right)\overline{v_g' h_T} = \left(\frac{r K^2 f_0^2}{2\overline{u} H^2}\right)\frac{h_0^2 \cos^2 l y}{\left[\left(K^2 - K_s^2\right)^2 + \varepsilon^2\right]} \tag{10.73}$$

where r is the spin-down rate due to boundary layer dissipation, ε is as defined below (7.100), and K_s is the resonant stationary Rossby wave number defined in (7.93).

It is clear from (10.73) that the form drag will have a strong maximum when $\overline{u} = \beta / K^2$, as shown schematically in Fig. 10.17. The last term on the right in (10.71), however, will decrease linearly as $\overline{u}$ increases. Thus, for suitable parameters there will be three values of $\overline{u}$ (labeled A, B, and C in Fig. 10.17) for which the wave drag just balances the external forcing so that steady-state solutions may exist. By perturbing the solution about the points A, B, and C, it is shown easily (Problem 10.12) that point B corresponds to an unstable equilibrium, whereas the equilibria at points A and C are stable. Solution A corresponds to a low-index equilibrium, with high-amplitude waves analogous to a blocking regime. Solution C corresponds to a high-index equilibrium with strong zonal flow and weak waves. Thus, for this very simple model there are two possible "climates" associated with the same forcing.

The Charney–DeVore model is a highly oversimplified model of the atmosphere. Models that contain more degrees of freedom do not generally have multiple steady solutions. Rather, there is a tendency for the (unsteady) solutions to cluster about certain climate regimes and to shift between regimes in an unpredictable fashion. Such behavior is characteristic of a wide range of nonlinear dynamical systems and is popularly known as *chaos* (see Section 13.8).

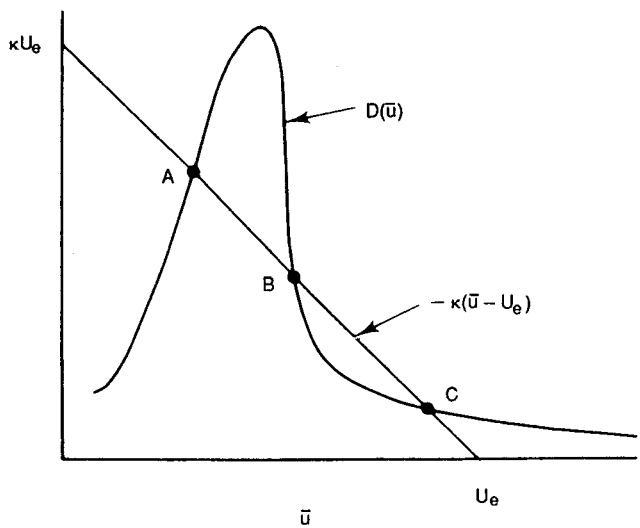


Fig. 10.17 Schematic graphical solution for steady-state solutions of the Charney–DeVore model. (Adapted from Held, 1983.)

10.6.2 Annular Modes

If a second meridional mode is added to the Charney–DeVore model (see MATLAB exercise M10.2), oscillatory solutions can be obtained that bear a qualitative resemblance to the observed leading modes of variability in the extratropical atmospheric circulation in both hemispheres. These observed modes, referred to as *annular modes*, are characterized by geopotential anomalies of opposite signs in the polar and midlatitude regions and corresponding mean zonal wind anomalies with opposite signs equatorward and poleward of about 45° latitude (Fig. 10.18). The

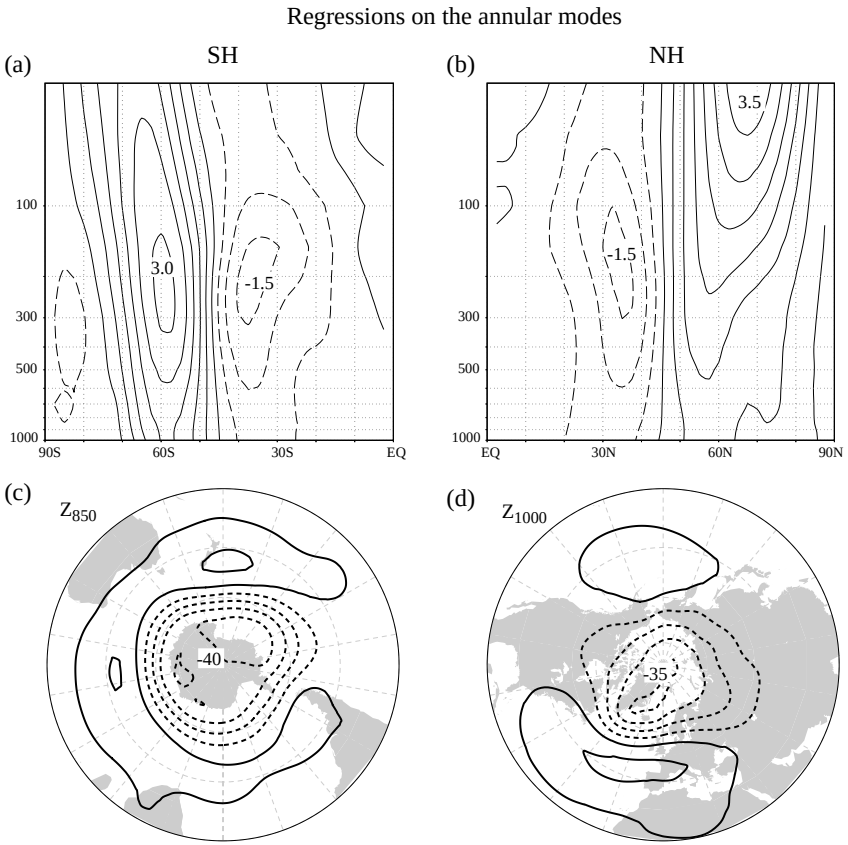


Fig. 10.18 (Top) Latitude–height cross sections showing typical amplitudes of annular mode anomalies for mean-zonal wind (ordinate labels are pressure in hPa, contour interval 0.5 m s^{-1}). (Bottom) Lower tropospheric geopotential height (contour interval 10 m). (Left) Southern Hemisphere; (right) Northern Hemisphere. Phase shown corresponds to the high index state (strong polar vortices). The low index state would have opposite signed anomalies. (After Thompson and Wallace, 2000.)

annular modes exist year round in the troposphere, but are strongest in the winter when they extend well into the stratosphere, especially in the Northern Hemisphere. The zonally symmetric mean-flow anomalies associated with the annular modes are apparently maintained by anomalous eddy momentum fluxes, which are themselves influenced by the zonally symmetric flow anomalies. Because these modes have their greatest influence at high latitudes, the northern and southern annular modes are sometimes referred to as the Arctic and Antarctic oscillations, respectively; it should be stressed, however, that they are not periodic oscillations, but rather represent two extremes of a broad distribution of climate states, with a wide range of associated time scales.

There is some evidence for downward propagation of the wintertime annular modes, suggesting that circulation changes in the stratosphere may precede annular mode changes in the troposphere. Dynamical linkages between the stratosphere and the troposphere are discussed further in Chapter 12.

10.6.3 Sea Surface Temperature Anomalies

Sea surface temperature anomalies influence the atmosphere through altering the flux of latent and sensible heat from the ocean, and thus providing anomalous heating patterns. The efficacy of such anomalies in exciting global scale responses depends on their ability to generate Rossby waves. A thermal anomaly can generate a Rossby wave response only by perturbing the vorticity field. This requires that the thermal anomaly produces an anomalous vertical motion field, which in turn produces anomalous vortex tube stretching.

For low-frequency disturbances the thermodynamic energy equation (10.5) may be approximated as

$$\mathbf{V} \bullet \nabla T + wN^2 HR^{-1} \approx J/c_p \tag{10.74}$$

Thus, diabatic heating can be balanced by horizontal temperature advection or by adiabatic cooling due to vertical motion. The ability of diabatic heating produced by a sea surface temperature anomaly to generate Rossby waves depends on which of these processes dominates. In the extratropics SST anomalies primarily generate low-level heating, which is balanced mainly by horizontal temperature advection. In the tropics, positive SST anomalies are associated with enhanced convection, and the resulting diabatic heating is balanced by adiabatic cooling. Tropical anomalies have their greatest effect in the Western Pacific where the average sea surface temperature is very high so that even a small positive anomaly can generate large increases in evaporation due to the exponential increase of saturation vapor pressure with temperature. By continuity of mass the upward motion in cumulonimbus convection requires convergence at low levels and divergence in the upper troposphere. The low-level convergence acts to sustain the convection by moistening and destabilizing the environment, whereas the upper-level divergence generates

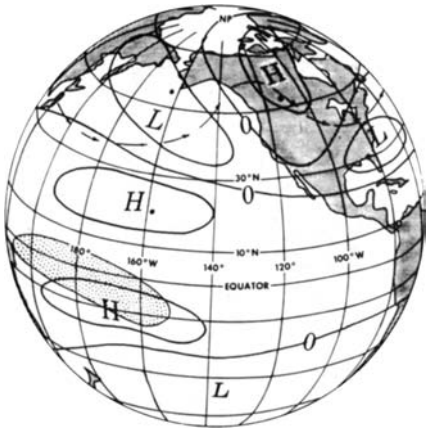


Fig. 10.19 The Pacific North American (PNA) pattern of middle and upper tropospheric height anomalies for Northern Hemisphere winter during an ENSO event in the tropical Pacific. The region of enhanced tropical precipitation is shown by shading. Arrows depict a 500-hPa streamline for the anomaly conditions. H and L designate anomaly highs and lows, respectively. The anomaly pattern propagates along a great circle path with an eastward component of group velocity, as predicted by stationary Rossby wave theory. (After Horel and Wallace, 1981. Reproduced with permission of the American Meteorological Society.)

a vorticity anomaly. If the mean flow is westerly in the region of upper level divergence, the forced vorticity anomaly will form a stationary Rossby wave train. The observed upper tropospheric height anomalies during the Northern Hemisphere winter associated with such an anomaly [referred to as the Pacific North America Pattern (PNA)] are shown schematically in Fig. 10.19. The pattern strongly suggests a train of stationary Rossby waves that emanates from the equatorial source region and follows a great circle path, as predicted by barotropic Rossby wave theory (Section 10.5.1). In this manner, tropical SST anomalies may generate low-frequency variability in the extratropics. It is possible that the effects of SST anomalies and of internal variability are not completely independent. In particular, it is more likely that the atmosphere will tend to preferentially reside in those regimes whose anomalous flow is correlated with the pattern in Fig 10.19 than would be the case in the absence of SST anomalies.

10.7 LABORATORY SIMULATION OF THE GENERAL CIRCULATION

We have shown that the gross features of the general circulation outside the tropics can be understood within the framework of the quasi-geostrophic system; the observed zonal-mean circulation and the atmospheric energy cycle are both rather

well described by quasi-geostrophic theory. The fact that the quasi-geostrophic model, in which the spherical earth is replaced by the β plane, can successfully model many of the essential features of the general circulation suggests that the fundamental properties of the general circulation are not dependent on parameters unique to a spherical planet, but may be common to all rotating differentially heated fluids. That the conjecture is in fact correct can be demonstrated in the laboratory with a rather simple apparatus.

In one group of experiments the apparatus consists of a cylindrical vessel that is rotated about its vertical axis. The vessel is heated at its rim and cooled at its center. These experiments are often referred to as *dishpan* experiments, as some of the first experiments of this type utilized an ordinary dishpan. The fluid in the dishpan experiments crudely represents one hemisphere in the atmosphere, with the rim of the dishpan corresponding to the equator and the center to the pole. Because the geometry is cylindrical rather than spherical, the β effect is not modeled for baroclinic motions in this type of system.[3] Thus the dishpan experiments omit dynamical effects of the atmospheric meridional vorticity gradient as well as the geometrical curvature terms, which were neglected in the quasi-geostrophic model.

For certain combinations of rotation and heating rates, the flow in the dishpan appears to be axially symmetric with a steady azimuthal flow that is in thermal wind equilibrium with the radial temperature gradient, and a superposed direct meridional circulation with rising motion near the rim and sinking near the center. This symmetric flow is usually called the *Hadley regime*, as the flow is essentially that of a Hadley cell.

For other combinations of rotation and heating rates, however, the observed flow is not symmetric. It consists, rather, of irregular wave-like fluctuations and meandering zonal jets. In such experiments, velocity tracers on the surface of the fluid reveal patterns very similar to those on midlatitude upper air weather charts, whereas tracers near the lower boundary reveal structures similar to atmospheric fronts. This type of flow is usually called the *Rossby regime*, although it should be noted that the observed waves are not Rossby waves because there is no β effect in the tank.

Experiments have also been carried out using an apparatus in which the fluid is contained in the annular region between two coaxial cylinders of different radii. In these *annulus* experiments, the walls of the inner and outer cylinders are held at constant temperatures so that a precisely controlled temperature difference can be maintained across the annular region. Very regular wave patterns can be obtained for certain combinations of rotation and heating. In some cases the wave patterns are steady, whereas in other cases they undergo regular periodic fluctuations called *vacillation cycles* (see Fig. 10.20). As mentioned in the last section, vacillation

[3] For barotropic motions the radial height gradient of the rotating fluid in the cylinder creates an "equivalent" β effect (see Section 4.3).

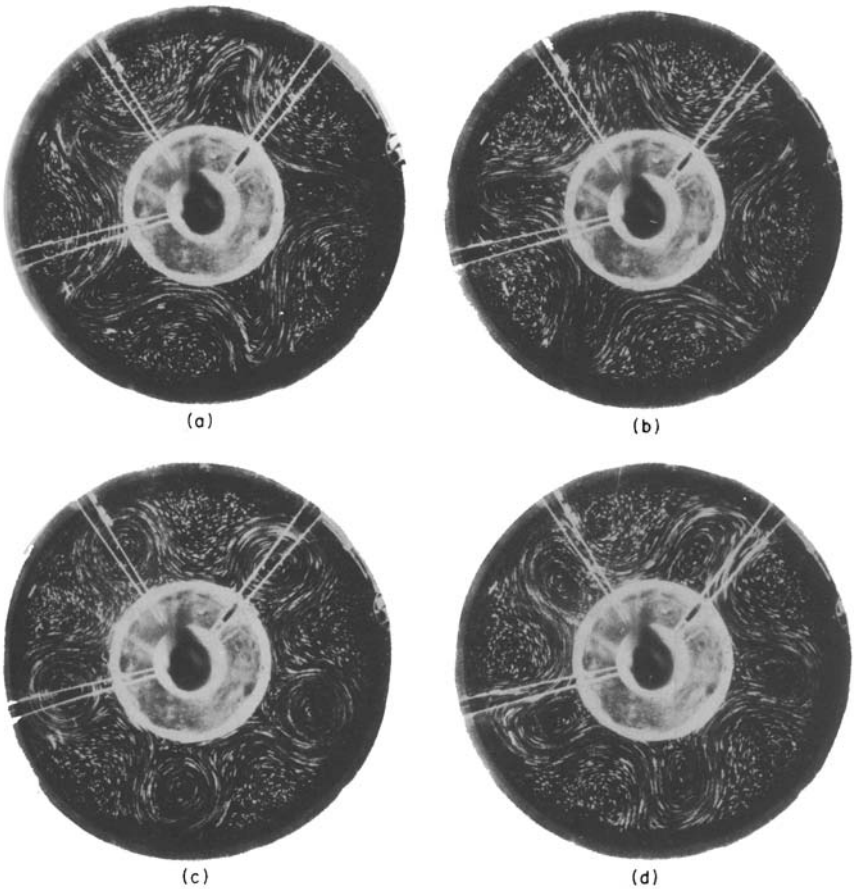


Fig. 10.20 Time exposures showing the motion of surface tracer particles in a rotating annulus. The four photographs illustrate various stages of a five-wave-tilted trough vacillation cycle. The period of the vacillation cycle is 16.25 revolutions, and the photographs are at intervals of four revolutions. (Photographs by Dave Fultz.)

is an example of internally generated low-frequency variability, which may be analogous to some atmospheric climate variability.

In order to determine whether the annulus experiments are really valid analogues of the atmospheric circulation or merely bear an accidental resemblance to atmospheric flow, it is necessary to analyze the experiments quantitatively. Mathematical analysis of the experiments can proceed from essentially the same equations that are applied to the atmosphere, except that cylindrical geometry replaces spherical geometry and temperature replaces potential temperature in the heat equation. (Because water is nearly incompressible, adiabatic temperature changes

are negligible following the motion.) In addition, the equation of state must be replaced by an appropriate measure of the relationship between temperature and density:

$$\rho = \rho_0 \left[1 - \varepsilon \left(T - T_0\right)\right] \tag{10.75}$$

where ε is the thermal expansion coefficient ($\varepsilon \approx 2 \times 10^{-4}\,\mathrm{K}^{-1}$ for water) and ρ_0 is the density at the mean temperature T_0.

Chapter 1 indicated that the character of the motion in a fluid is crucially dependent on the characteristic scales of parameters such as velocity, pressure perturbation, length, and time. In the laboratory the scales of these parameters are generally many orders of magnitude different from their scales in nature. It is still possible, however, to produce quasi-geostrophic motions in the laboratory provided that the motions are slow enough so that the flow is approximately hydrostatic and geostrophic. As shown in Section 2.4.2, the geostrophy of a flow does not depend on the absolute value of the scaling parameters, but rather on a nondimensional ratio of these parameters called the Rossby number. For annulus experiments, the maximum horizontal scale is set by the dimensions of the tank so that it is convenient to define the Rossby number as

$$\mathrm{Ro} \equiv U / \left[\Omega \left(b - a\right)\right]$$

where Ω is the angular velocity of the tank, U is a typical relative velocity of the fluid, and $b - a$ is the difference between the radius b of the outer wall and the radius a of the inner wall of the annular region.

Using the hydrostatic approximation

$$\partial p \big/ \partial z = -\rho g$$

together with the geostrophic relationship

$$2\Omega \mathbf{V}_g = \rho_0^{-1} \mathbf{k} \times \boldsymbol{\nabla} p$$

and the equation of state (10.75) we can obtain a *thermal wind* relationship in the form

$$\frac{\partial \mathbf{V}_g}{\partial z} = \frac{\varepsilon g}{2\Omega} \mathbf{k} \times \boldsymbol{\nabla} T \tag{10.76}$$

Letting U denote the scale of the geostrophic velocity, H the mean depth of fluid, and δT the radial temperature difference across the width of the annulus, we obtain from (10.76)

$$U \sim \varepsilon g H \delta T \,/\, \left[2\Omega \left(b - a\right)\right] \tag{10.77}$$

Substituting the value of U in (10.77) into the formula for the Rossby number yields the *thermal Rossby number*

$$\mathrm{Ro_T} \equiv \varepsilon g H \delta T \Big/ \left[2\Omega^2 \left(b - a\right)^2\right]$$

This nondimensional number is the best measure of the range of validity of quasi-geostrophic dynamics in annulus experiments. Provided that $\mathrm{Ro_T} \ll 1$ the quasi-geostrophic theory should be valid for motions in the annulus. For example, in a typical experiment

$$H \approx b - a \approx 10\ \mathrm{cm}, \quad \delta T \approx 10\ \mathrm{K}, \quad \text{and} \quad \Omega \approx 1\ \mathrm{s}^{-1}$$

Thus, $\mathrm{Ro_T} \approx 10^{-1}$.

Annulus experiments carried out over a wide range of rotation rates and temperature contrasts produce results that can be classified consistently on a log–log plot of the thermal Rossby number versus a nondimensional measure of the rotation rate :

$$\left(G^*\right)^{-1} \equiv (b - a)\,\Omega^2/g \tag{10.78}$$

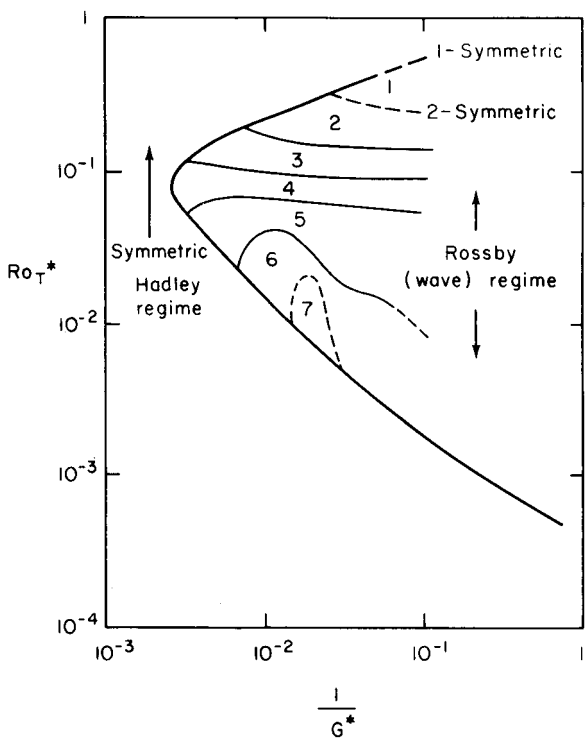


Fig. 10.21 Regime diagram for Fultz's annulus experiments. (After Phillips, 1963.)

Some results are summarized in Fig. 10.21. The heavy solid line separates the axially symmetric Hadley regime from the wavy Rossby regime. These results can best be understood qualitatively by considering an experiment in which the thermal Rossby number is increased slowly from zero by gradually imposing a temperature difference $T_b - T_a$. For small $T_b - T_a$ the motion is of the Hadley type, with weak horizontal and vertical temperature gradients in the fluid. As the horizontal temperature contrast is increased, however, the mean thermal wind must also increase, until at some critical value of Ro_T the flow becomes baroclinically unstable. According to the theory presented in Chapter 8, the wavelength of maximum instability is proportional to the ratio of the static stability to the square of the rotation rate. Thus, as can be seen in Fig. 10.21 the wave number observed when the flow becomes unstable decreases (i.e., wavelength increases) as the rotation rate is reduced. Furthermore, because baroclinic waves transport heat vertically as well as laterally, they will tend to increase the static stability of the fluid. Therefore, as the thermal Rossby number is increased within the Rossby regime, the increased heat transport by the waves will raise the static stability, and hence increase the wavelength of the wave of maximum instability. The flow then undergoes transitions in which the observed wave number decreases until finally the static stability becomes so large that the flow is stable to even the largest wave that can fit the tank. The flow then returns to a symmetric Hadley circulation stabilized by a high static stability, which is maintained by a vigorous direct meridional circulation. This regime is usually called the *upper* symmetric regime to distinguish it from the *lower* symmetric regime, which occurs for very weak heating.

Laboratory studies, despite their many idealizations, can model the most important of those features of the general circulation that are not dependent on the topography of the earth or continent–ocean heating contrasts. Specifically, we find, perhaps surprisingly, that the β effect (i.e., the planetary vorticity gradient) is not essential for the development of circulations that look very much like tropospheric synoptic systems. Thus, observed midlatitude waves should be regarded essentially as baroclinic waves modified by the β effect, not as Rossby waves in a baroclinic current.

In addition to demonstrating the primacy of baroclinic instability, the experiments also confirm that internal diabatic heating due to condensation or radiative processes is not an essential mechanism for simulation of large-scale midlatitude circulations. Laboratory experiments thus enable us to separate the essential mechanisms from second-order effects in a manner not easily accomplished by observation of the atmosphere itself. Also, because laboratory simulation experiments have typical rotation rates of 10 rpm, it is possible to model many years of "atmospheric" flow in a short time so that it is feasible to accumulate accurate statistics even for very low-frequency variability. In addition, because temperatures and velocities can be measured at uniform intervals, the experiments can provide excellent sets of data for testing numerical weather prediction models.

10.8 NUMERICAL SIMULATION OF THE GENERAL CIRCULATION

The previous section discussed the role of laboratory experiments in contributing to a qualitative understanding of the general circulation of the atmosphere. Although laboratory experiments can elucidate most of the gross features of the general circulation, there are many details that cannot possibly be duplicated in the laboratory. For example, the possible long-term climatic effects of aerosols and trace gases that are added to the atmosphere as a result of human activities could not possibly be predicted on the basis of laboratory experiments. As another example, the influence of an ice-free Arctic Ocean on the global climate would also be extremely difficult to simulate in the laboratory. Because all the conditions of the atmosphere cannot be duplicated in the laboratory, the only practical manner in which to quantitatively simulate the present climate, or to predict possible climate modifications resulting from intentional or unintentional human intervention, is by numerical simulation with the aid of supercomputers.

Atmospheric general circulation models (AGCMs) are similar to large-scale numerical weather prediction models (see Chapter 13) in that they attempt to explicitly simulate synoptic-scale weather disturbances. However, whereas weather prediction is an *initial value* problem, which requires that the evolution of the flow be computed from a specified initial state, general circulation modeling is a *boundary value* problem in which the average circulation is computed for specified external forcing conditions. In many AGCMs the sea surface temperature is treated as a specified forcing. In reality, of course, there are strong interactions between the atmosphere and the ocean: the winds drive currents, which influence the sea surface temperature distribution, which in turn influences the atmosphere.

Although the dynamical equations used in general circulation models are the same as those used in short-range numerical prediction models, general circulation models are generally more complex, as in simulations of time scales beyond a few days physical processes that are unimportant for the short-term evolution may become crucial. Thus, parameterizations are needed for physical processes such as surface heat and moisture fluxes, moist convection, turbulent mixing, and radiation. However, because a forecast of the flow evolution from a specific initial state is not needed, the problems associated with specifying initial data in forecast models do not arise in AGCMs. In recent years, as numerical prediction has been extended into the middle range of 1–2 weeks, more physical processes have been included in the forecast models, and the differences between global weather forecast models and general circulation models have become smaller.

10.8.1 The Development of AGCMs

The success of the quasi-geostrophic model in short-range prediction suggests that for simulating the gross features of the general circulation such a model

might be adequate provided that diabatic heating and frictional dissipation were included in a suitable manner. Phillips (1956) made the first attempt to model the atmospheric general circulation numerically. His experiment employed a two-level quasi-geostrophic forecast model modified so that it included boundary layer friction and latitudinally dependent radiative heating. The heating rates chosen by Phillips were based on estimates of the net diabatic heating rates necessary to balance the poleward heat transport at 45°N computed from observational data. Despite the severe limitations of this model, the simulated circulation in some respects resembled the observed extratropical circulation.

Phillips' experiment, although it was an extremely important advance in dynamic meteorology, suffered from a number of shortcomings as a general circulation model. Perhaps the gravest shortcoming in his model was the specification of diabatic heating as a fixed function of latitude only. In reality the atmosphere must to some degree determine the distribution of its own heat sources. This is true not only for condensation heating, which obviously depends on the distribution of vertical motion and water vapor, but also holds for radiative heating as well. Both net solar heating and net infrared heating are sensitive to the distribution of clouds, and infrared heating depends on the atmospheric temperature as well.

Another important limitation of Phillips' two-level model was that static stability could not be predicted but had to be specified as an external parameter. This limitation is a serious one because the static stability of the atmosphere is obviously controlled by the motions.

It is possible to design a quasi-geostrophic model in which the diabatic heating and static stability are motion dependent. Indeed such models have been used to some extent, especially in theoretical studies of the annulus experiments. To model the global circulation completely, however, requires a dynamical framework that is valid in the equatorial zone. Thus, it is desirable to base an AGCM on the global primitive equations. Because of its enormous complexity and many important applications, general circulation modeling has become a highly specialized activity that cannot possibly be covered adequately in a short space. Here we can only give a summary of the primary physical processes represented and present an example of an application in climate modeling. A brief discussion of the technical aspects of the formulation of numerical prediction models is given in Chapter 13.

10.8.2 Dynamical Formulation

Most general circulation models are based on the primitive equations in the σ-coordinate form introduced in Section 10.3.1. As was pointed out in that section, σ coordinates make it possible to retain the dynamical advantages of pressure coordinates, but simplify the specification of boundary conditions at the surface.

The minimum prediction equations for a σ coordinate GCM are the horizontal momentum equation (10.30), the mass continuity equation (10.34), the thermodynamic energy equation (10.36), and a moisture continuity equation, which can be expressed as

$$\frac{D}{Dt}(q_v) = P_v \tag{10.79}$$

where q_v is the water vapor mixing ratio and P_v is the sum of all sources and sinks. In addition we require the hydrostatic equation (10.35) to provide a diagnostic relationship between the geopotential and temperature fields. Finally, a relationship is needed to determine the evolution of the surface pressure $p_s(x, y, t)$. This is given by integrating (10.34) vertically and using the boundary conditions $\dot{\sigma} = 0$ at $\sigma = 0$ and 1 to obtain

$$\frac{\partial p_s}{\partial t} = -\int_0^1 \nabla \bullet (p_s \mathbf{V}) d\sigma \tag{10.80}$$

Vertical variations are generally represented by dividing the atmosphere into a number of levels and utilizing a finite difference grid. AGCMs typically have prediction levels at 1–3-km intervals extending from the surface to about a 30-km altitude. Some models, however, have many more levels extending nearly to the mesopause. Horizontal resolution of global models varies widely, from an effective grid size of several hundred kilometers, to less than 100 km.

10.8.3 Physical Processes and Parameterizations

The various types of surface and atmospheric processes represented in a typical AGCM and the interactions among these processes are shown schematically in Fig. 10.22. The most important classes of physical processes are (i) radiative, (ii) cloud and precipitation, and (iii) turbulent mixing and exchange.

As pointed out in Section 10.1, the fundamental process that drives the circulation of the atmosphere is the differential radiative heating that results from the fact that solar absorption at the surface is greater than long-wave emission to space at low latitudes, whereas long-wave emission dominates over solar absorption at high latitudes, at least in the winter hemisphere. The general circulation of the atmosphere and the oceans provides the meridional and vertical heat transfer required for balance.

Most of the solar radiation absorbed at the surface is used to evaporate water, and hence to moisten the atmosphere. Solar heating is realized in the atmosphere primarily in the form of latent heat release associated with convective clouds. The global distribution of evaporation clearly depends on the sea surface temperature, which is itself dependent on the general circulation of the oceans as well as on interactions with the atmosphere. That is why for detailed understanding of the

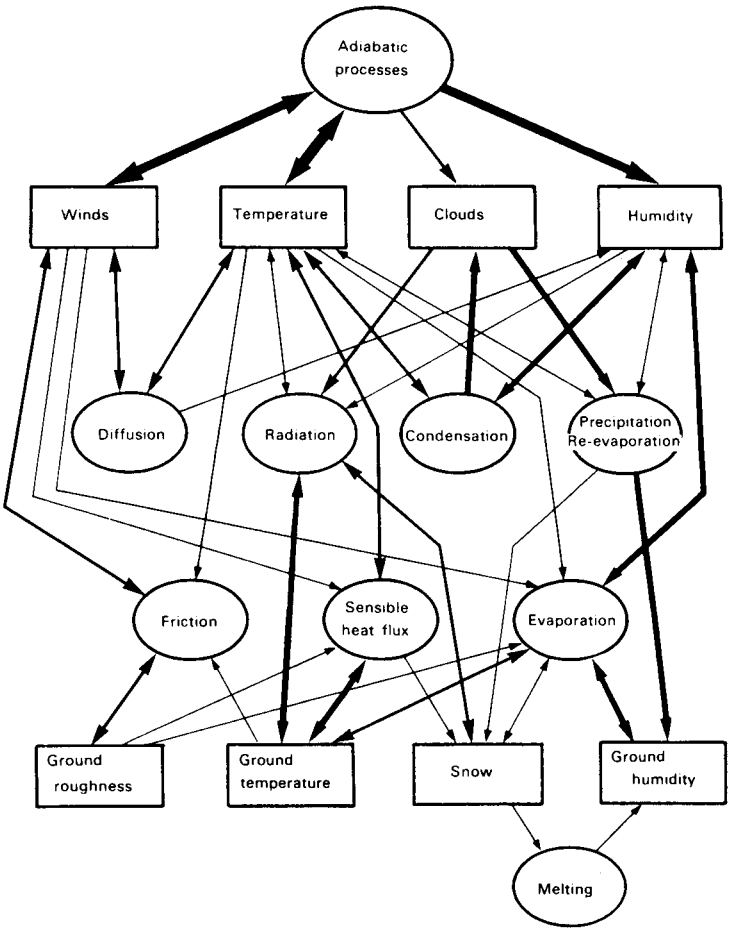
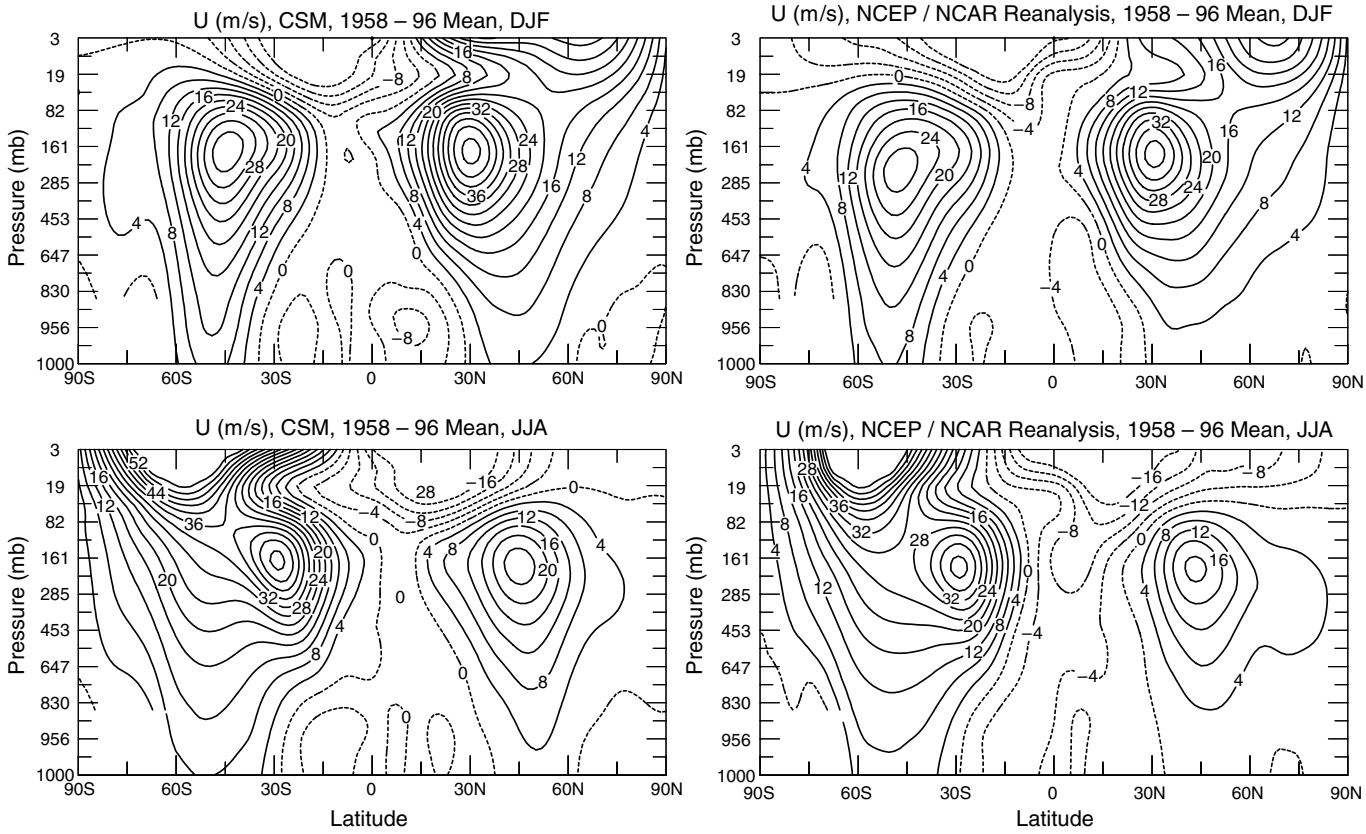


Fig. 10.22 Schematic illustration of processes commonly included in AGCMs and their interactions. The thickness of each arrow gives a rough indication of the importance of the interaction that the particular arrow represents. (After Simmons and Bengtsson, 1984.)

influence of large-scale air–sea interactions on climate it is important to run coupled atmosphere–ocean GCMs in which the sea surface temperature is predicted. At present, however, most AGCMs use externally specified monthly or seasonal mean sea surface temperatures. Over land, however, the surface temperature adjusts very quickly to changes in the fluxes of solar and infrared radiation and is determined from a surface energy-balance equation.

Atmospheric radiative heating by solar radiation and heating and cooling by long-wave thermal radiation are computed using radiative transfer models of varying sophistication. Zonally averaged distributions of radiatively important


U (m/s), CSM, 1958 – 96 Mean, DJF
U (m/s), NCEP / NCAR Reanalysis, 1958 – 96 Mean, DJF
U (m/s), CSM, 1958 – 96 Mean, JJA
U (m/s), NCEP / NCAR Reanalysis, 1958 – 96 Mean, JJA
Pressure (mb)
3
19
82
161
285
453
647
830
956
1000
90S
60S
30S
0
30N
60N
90N
Latitude

constituents such as carbon dioxide, ozone, and even cloudiness are often employed, but the more complete models do utilize model predicted zonally and time varying cloudiness in their radiation codes.

Boundary layer fluxes of momentum, heat, and moisture are parameterized in most AGCMs by the use of bulk aerodynamic formulas (see Section 5.3.1). Typically the fluxes are specified to be proportional to the magnitude of the horizontal velocity at the lowest atmospheric level times the difference between the field variable at the boundary and its value at the lowest atmospheric level. In some models the boundary layer is explicitly resolved by locating several prediction levels within the lowest 2 km and utilizing the model predicted boundary layer static stability in the parameterization of turbulent fluxes.

The hydrological cycle is usually represented by a combination of parameterization and explicit prediction. The water vapor mixing ratio is generally one of the explicitly predicted fields. The distributions of layer clouds and large-scale precipitation are then determined from the predicted distribution of humidity by requiring that when the predicted humidity exceeds 100% enough vapor is condensed to reduce the mixing ratio to saturation or less. Parameterizations in terms of the mean state thermal and humidity structure must be used to represent the distributions of convective clouds and precipitation.

10.8.4 The NCAR Climate System Model

The Climate System Model (CSM) developed at the National Center for Atmospheric Research (NCAR) is a coupled global climate model consisting of an AGCM, an ocean GCM, a land surface biophysics model, and a sea-ice model. The atmospheric component is a spectral model (see Section 13.5) with about 2.8° horizontal resolution and 18 layers in the vertical with the top layer centered at 4.8 hPa.

The ocean component of the CSM has 2.4° resolution in longitude and variable resolution in latitude, with smallest grid spacing of 1.2° at the equator and in the Arctic, and largest spacing of 2.3° at midlatitudes. In the vertical, the ocean model has 45 levels, with more than half these concentrated in the upper kilometer of the ocean. The land surface model runs on the same grid as the AGCM, whereas the sea-ice model runs on the ocean model grid.

Because characteristic timescales for the various components of the CSM are very different, the individual components are integrated to approximate equilibrium states with fixed external forcing, and then the components are coupled together and stepped forward until the coupled climate system equilibrates. Solstice season zonal mean winds simulated by the CSM for late 20th century conditions are shown in Fig. 10.23. Overall there is excellent agreement between the model and observations.

Fig. 10.23 Meridional cross sections of mean zonal wind ($m\,s^{-1}$) for the two solstice seasons. (Left) Results from CSM simulation; (right) observed climatology. (After Dai et al., 2001.)

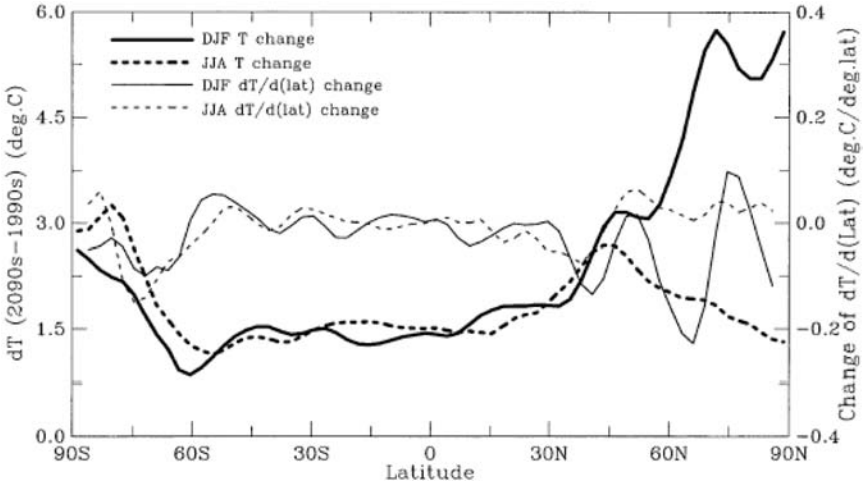


Fig. 10.24 Latitudinal distribution of predicted changes of the zonally averaged surface air temperature from the 1990s to the 2090s for Northern Hemisphere winter (DJF: heavy solid curve) and Northern Hemisphere summer (JJA: heavy dashed curve) for a scenario with approximate doubling of the atmospheric carbon dioxide concentration. Also shown as light curves (right side ordinate) are changes (2090s minus 1990s) of the pole to equator temperature gradient in °C per degree of latitude. (After Dai et al., 2001.)

The model has been used to assess likely responses of the climate system to the radiative forcing by increased greenhouse gas concentrations. The latitudinal distribution of zonally averaged surface temperature change predicted for the end of the 21st century under a scenario of approximate doubling of atmospheric carbon dioxide concentrations ("business as usual") is shown in Fig. 10.24. In common with other climate models, the CSM result suggests that the largest temperature changes will occur at high latitudes in the Northern Hemisphere winter, accompanied by a substantial reduction of sea ice. A global average warming of 1.9°C by the end of the century is predicted for the business as usual scenario. Because of the complexity of the physical processes involved, many uncertainties remain in such model predictions of climate change. Nevertheless, most climate scientists agree that increased carbon dioxide concentrations will lead to a warmer surface climate.

PROBLEMS

10.1. Starting with the isobaric version of the thermodynamic energy equation (2.42), derive the log-pressure version (10.5).

10.2. Show that in the σ-coordinate system a mass element $\rho_0 dx dy dz$ takes the form $-g^{-1} p_s dx dy d\sigma$.

10.3. Compute the mean zonal wind $\overline{u}$ at the 200-hPa level at 30°N under the assumptions that $\overline{u} = 0$ at the equator and that the absolute angular momentum is independent of latitude. What is the implication of this result for the role of eddy motions?

10.4. Show by scale analysis that advection by the mean meridional circulation can be neglected in the zonally averaged equations (10.11) and (10.12) for quasi-geostrophic motions.

10.5. Show that for quasi-geostrophic eddies the next to last term in square brackets on the right-hand side in (10.15) is proportional to the vertical derivative of the eddy meridional relative vorticity flux.

10.6. Starting from equations (10.16)–(10.19), derive the governing equation for the residual streamfunction (10.21).

10.7. Using the observed data given in Fig. 10.13, compute the time required for each possible energy transformation or loss to restore or deplete the observed energy stores. (A watt equals 1 J s^{-1}.)

10.8. Compute the surface torque per unit horizontal area exerted on the atmosphere by topography for the following distribution of surface pressure and surface height:

$$p_s = p_0 + \hat{p} \sin kx, \qquad h = \hat{h} \sin (kx - \gamma)$$

where $p_0 = 1000$ hPa, $\hat{p} = 10$ hPa, $\hat{h} = 2.5 \times 10^3$ m, $\gamma = \pi/6$ rad, and $k = 1/(a \cos \phi)$, where $\phi = \pi/4$ radians is the latitude, and a is the radius of the earth. Express the answer in kg s^{-2}.

10.9. Starting from (10.66) and (10.67) show that the group velocity relative to the ground for stationary Rossby waves is perpendicular to the wave crests and has a magnitude given by (10.69).

10.10. Derive the expression of (10.76) for the thermal wind in the dishpan experiments.

10.11. Consider a thermally stratified liquid contained in a rotating annulus of inner radius 0.8 m, outer radius 1.0 m, and depth 0.1 m. The temperature at the bottom boundary is held constant at T_0. The fluid is assumed to satisfy the equation of state (10.75) with $\rho_0 = 10^3$ kg m^{-3} and $\varepsilon = 2 \times 10^{-4}$ K^{-1}. If the temperature increases linearly with height along the outer radial boundary at a rate of 1°C cm^{-1} and is constant with height along the inner radial boundary, determine the geostrophic velocity at the upper boundary for a rotation rate of $\Omega = 1$ rad s^{-1}. (Assume that the temperature depends linearly on radius at each level.)

10.12. Show by considering $\partial\overline{u}/\partial t$ for small perturbations about the equilibrium points in Fig. 10.17 that point B is an unstable equilibrium point while points A and C are stable.

MATLAB EXERCISES

M10.1. The MATLAB script **topo_wave_1.m** uses finite differencing to solve the linearized vorticity equation (7.99) on a midlatitude β plane for the case of a constant mean zonal wind and forcing by a circular mountain with mountain height given by $h_T(x, y) = h_0 L^2 \left(L^2 + x^2 + y^2\right)^{-1}$, where h_0 and L are constants characterizing the mountain height and horizontal scale, respectively. Run the script for mean zonal winds of 5, 10, 15, and 20 m s^{-1}. In each case estimate the horizontal wavelength and group velocity of the northward and southward Rossby wave trains that form in the lee of the mountain. Compare your results with the corresponding expressions given in Section 10.5.1.

M10.2. The two-meridional-mode version of the Charney and Devore (1979) model described in Section 10.6.1 is given in the MATLAB script **C_D_model.m**. Solutions for this model may be steady state, periodic in time, or irregularly varying, depending on the forcing of the first mode zonal-mean streamfunction (zf(1) in the script) and the initial amplitude of this mode (zinit(1) in the script). Run this script for forcing values of zf(1) = 0.1, 0.2, 0.3, 0.4, 0.5. In each case do two runs with initial conditions zinit(1) = zf(1) and zinit(1) = 0.1, respectively. For each of the 10 cases, note whether the solution is steady, periodic, or irregular. Compute the time average of the zonal mean zonal wind for nondimensional time interval $2000 < t < 3000$ for each of these cases. Are your results consistent with Fig. 10.17 in the text? Note for each case whether the streamfunction tends to be in phase or out of phase with the topography and whether the results are qualitatively consistent with the solution for topographic Rossby waves given in Section 7.7.2.

M10.3. The MATLAB script **baroclinic_1.m** provides a simple illustration of the effect of baroclinic eddies on the mean flow. The script extends the two-level baroclinic instability model discussed in Section 8..2 by calculating the evolution of the mean zonal flow components U_m and U_T caused by meridional vorticity and heat fluxes associated with unstable baroclinic waves. The calculation is done for a midlatitude β plane with weak Ekman layer damping. The eddies are governed by the linearized model of Section 8.2 (with fixed zonal wavelength of 6000 km), but the zonal-mean flow is affected by eddy-mean flow interactions and evolves

as a result of eddy heat and momentum fluxes. In this model the mean thermal wind is relaxed toward a baroclinically unstable radiative equilibrium state (designated by U0rad in the code), which has a sinusoidal distribution in the meridional direction, and the mean wind components U_m and U_T are assumed to vanish at the northern and southern boundaries. Run the model for four cases with U0rad = 10, 20, 30, and 40 $\mathrm{m\,s^{-1}}$, respectively. In each case run the model long enough so that the eddy kinetic energy equilibrates. Show in a table the maximum values for U_m and U_T for each of these cases. Explain these results in terms of the baroclinic instability theory of Section 8.2.

Suggested References

Lorenz, *The Nature and Theory of the General Circulation of the Atmosphere*, although somewhat out of date, contains an excellent survey of the subject, including observational and theoretical aspects.

James, *Introduction to Circulating Atmospheres*, provides an excellent introduction to observational and theoretical aspects of the global general circulation.

Hide (1966) reviews a number of types of laboratory experiments with rotating fluids, including rotating annulus experiments.

Washington and Parkinson, *An Introduction to Three-Dimensional Climate Modeling*, provide an excellent text, which covers the physical basis and computational aspects of general circulation modeling.

Randall (ed.), *General Circulation Model Development*, contains articles describing all aspects of general circulation modeling.

CHAPTER 11

Tropical Dynamics

Throughout the previous chapters of this book, we have emphasized circulation systems of the extratropical regions (i.e., the regions poleward of about 30° latitude). This emphasis should not be regarded as an indication of a lack of interesting motion systems in the tropics, but is a result, rather, of the relative complexity of the dynamics of tropical circulations. There is no simple theoretical framework, analogous to quasi-geostrophic theory, that can be used to provide an overall understanding of large-scale tropical motions.

Outside the tropics, the primary energy source for synoptic-scale disturbances is the zonal available potential energy associated with the latitudinal temperature gradient. Observations indicate that latent heat release and radiative heating are usually secondary contributors to the energetics of extratropical synoptic-scale systems. In the tropics, however, the storage of available potential energy is small due to the very small temperature gradients in the tropical atmosphere. Latent heat release appears to be the primary energy source, at least for those disturbances that originate within the equatorial zone. Most latent heat release in the tropics occurs in association with convective cloud systems, although much of the actual precipitation falls from mesoscale regions of stratiform clouds within such systems; the cloud systems are themselves generally embedded in large-scale

circulations. The diabatic heating associated with tropical precipitation not only drives a local response in the atmospheric circulation, but through excitation of equatorial waves may also induce a remote response. Thus, there is a strong interaction among cumulus convection, mesoscale, and large-scale circulations, which is of primary importance for understanding tropical motion systems. Furthermore, the distribution of diabatic heating in the tropics is strongly influenced by sea surface temperature (SST) variations, and these in turn are strongly influenced by the motion of the atmosphere.

An understanding of tropical circulations thus requires consideration of equatorial wave dynamics, interactions of cumulus convection and mesoscale circulations with large-scale motions, and air–sea interactions. Detailed treatment of all these topics is beyond the scope of an introductory text. Nevertheless, because the tropics play a fundamental role in the general circulation of the atmosphere, and coupling between tropical and middle latitudes is an important consideration in extratropical extended-range forecasting, some discussion of the tropics is required even in a text with an extratropical emphasis.

Of course, it is not always possible to distinguish clearly between tropical and extratropical systems. In subtropical regions ($\sim 30°$ latitude), circulation systems characteristic of both tropical and extratropical regions may be observed depending on the season and geographical location. To keep the discussion in this chapter as simple as possible, we focus primarily on the zone well equatorward of 30° latitude where the influence of middle latitude systems should be minimum.

11.1 THE OBSERVED STRUCTURE OF LARGE-SCALE TROPICAL CIRCULATIONS

Due to the nature of their energy sources, as well as the smallness of the Coriolis parameter, large-scale equatorial motion systems have several distinctive characteristic structural features that are quite different from those of midlatitude systems. Many of these can be understood in terms of the equatorial wave modes discussed in Section 11.3. Before discussing equatorial wave theory, however, it is useful to review some of the major observed circulation features of the tropical atmosphere.

11.1.1 The Intertropical Convergence Zone

Traditionally, the tropical general circulation was thought to consist of a thermally direct Hadley circulation in which air in the lower troposphere in both hemispheres moved equatorward toward the *intertropical convergence zone* (ITCZ) where by continuity considerations it was forced to rise uniformly and move poleward, thus transporting heat away from the equator in the upper troposphere in both hemispheres. This simple model of large-scale overturning is not, however, consistent with the observed vertical profile of equivalent potential temperature (θ_e). As indicated in Fig. 11.1 the mean tropical atmosphere is conditionally stable above

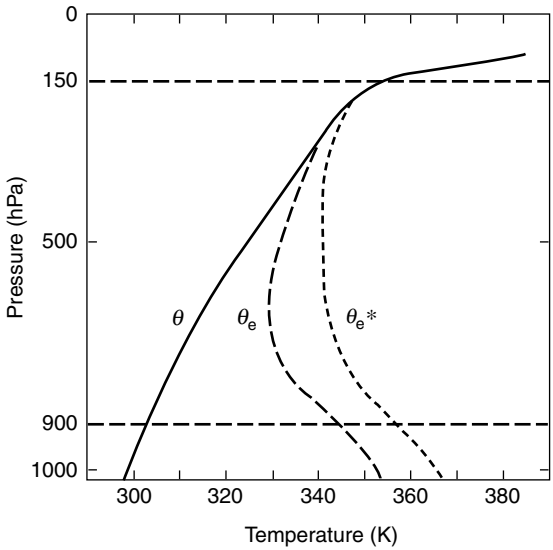


Fig. 11.1 Typical sounding in the tropical atmosphere showing the vertical profiles of potential temperature θ, equivalent potential temperature θ_e, and the equivalent potential temperature θ_ε^* of a hypothetically saturated atmosphere with the same temperature at each level. This figure should be compared with Fig. 9.10, which shows similar profiles for a midlatitude squall line sounding. (After Ooyama, 1969. Reproduced with permission of the American Meteorological Society.)

about 600 hPa. Thus, a large-scale upward mass flow, were it to exist, would be up the gradient of θ_e in the upper troposphere and would actually *cool* the upper troposphere in the region of the ITCZ. Such a circulation could not generate potential energy and would not, therefore, satisfy the heat balance in the equatorial zone.

It appears that the only way in which heat can effectively be brought from the surface to the upper troposphere in the ITCZ is through pseudo-adiabatic ascent in the cores of large cumulonimbus clouds (often referred to as "hot towers"). For such motions, cloud parcels approximately conserve θ_e. They can, therefore, arrive in the upper troposphere with moderate temperature excesses. Thus, the heat balance of the equatorial zone can be accounted for, at least qualitatively, provided that the vertical motion in the ITCZ is confined primarily to the updrafts of individual convective cells. Riehl and Malkus (1958) estimated that only 1500–5000 individual hot towers would need to exist simultaneously around the globe to account for the required vertical heat transport in the ITCZ.

This view of the ITCZ as a narrow zonal band of vigorous cumulus convection has been confirmed beyond a doubt by observations, particularly satellite cloud photos. An example is given in Fig. 11.2, which shows infrared cloud brightness temperatures in the tropics for the period August 14 to December 17, 1983. Low

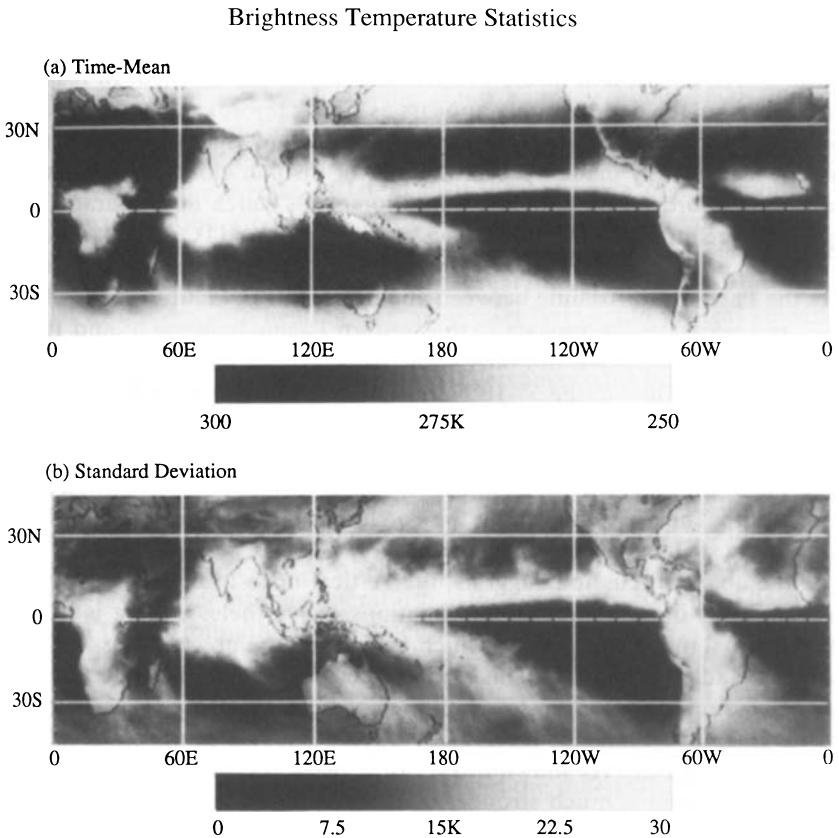


Fig. 11.2 (a) Time-mean IR brightness temperature and (b) three-hourly standard deviations about the time-mean IR brightness temperature for August 14 to December 17, 1983. Low values of brightness temperature indicate the presence of high cold cirrus anvil clouds. (After Salby et al., 1991.)

brightness temperatures signal the presence of deep anvil clouds characteristic of convective storms. The ITCZ appears as a line of deep convective cloud extending across the Atlantic and Pacific oceans between about 5° and 10°N.

Observations indicate that within the ITCZ precipitation greatly exceeds the moisture supplied by evaporation from the ocean surface below. Thus, much of the vapor necessary to maintain the convection in the ITCZ must be supplied by the converging trade wind flow in the lower troposphere. In this manner the large-scale flow provides the latent heat supply required to sustain the convection. The convective heating in turn produces large-scale midtropospheric temperature perturbations and (through hydrostatic adjustment) surface and upper level pressure perturbations, which maintain the low-level inflow.

The above description of the ITCZ is actually oversimplified. In reality, the ITCZ over the oceans rarely appears as a long unbroken band of heavy convective cloudiness, and it almost never is found right at the equator. Rather, it is usually made up of a number of distinct cloud clusters, with scales of the order of a few hundred kilometers, which are separated by regions of relatively clear skies. The strength of the ITCZ is also quite variable in both space and time. It is particularly persistent and well defined over the Pacific and Atlantic between about 5° and 10°N latitude (as in Fig. 11.2) and occasionally appears in the western Pacific between 5° and 10°S.

Figure 11.2 shows that not only is the mean deep convection associated with the ITCZ found in the 5°–10°N latitude belt, but during the period shown in Fig. 11.2 the standard deviation of deep convection is also a maximum there. This is consistent with the idea that the ITCZ is the locus of transient cloud clusters, rather than simply a region of steady-state precipitation and mean uplift. The dry zones along the equator in the oceanic regions are a particularly striking feature in Fig. 11.2.

As the above discussion suggests, the vertical mass flux associated with the ITCZ has important regional variations. Nevertheless, there is a significant zonal mean component, which constitutes the upward mass flux of the mean Hadley circulation. This Hadley circulation consists of overturning thermally direct cells in the low latitudes of both hemispheres, as shown in Fig. 10.7. The center of the Hadley circulation is located at the mean latitude of the ITCZ. As shown in Fig. 10.7, the winter hemisphere branch of the Hadley cell is much stronger than the summer hemisphere branch. Observations indicate that two Hadley cells, symmetric about the equator, are rarely observed even in the equinoctial seasons. Rather, the northern cell dominates from November to March, the southern cell dominates from May to September, and rapid transitions occur in April and October (see Oort, 1983).

11.1.2 Equatorial Wave Disturbances

The variance observed in the cloudiness associated with the ITCZ, as illustrated in Fig. 11.2b, is generally caused by transient precipitation zones associated with weak equatorial wave disturbances that propagate westward along the ITCZ. That such westward propagating disturbances exist and are responsible for a large part of the cloudiness in the ITCZ can be seen easily by viewing time-longitude sections constructed from daily satellite pictures cut into thin zonal strips. An example is shown in Fig. 11.3. The well-defined bands of cloudiness that slope from right to left down the page define the locations of the cloud clusters as functions of longitude and time. Clearly much of the cloudiness in the 5–10°N latitude zone of the Pacific is associated with westward moving disturbances. The slope of the cloud lines in Fig. 11.3 implies a westward propagation speed of about 8–10 m s^{-1}. The

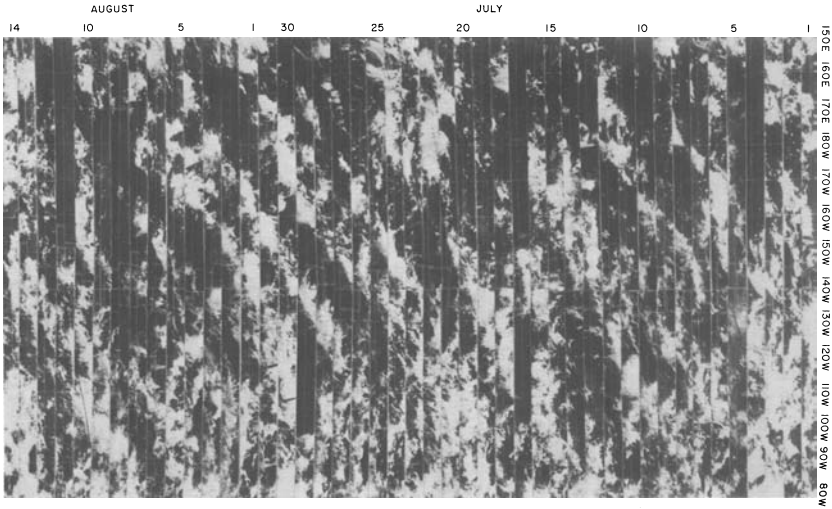


Fig. 11.3 Time-longitude sections of satellite photographs for the period July 1–August 14, 1967 in the 5–10°N latitude band of the Pacific. The westward progression of the cloud clusters is indicated by the bands of cloudiness sloping down the page from right to left. (After Chang, 1970. Reproduced with permission of the American Meteorological Society.)

longitudinal separation of the cloud bands is about 3000–4000 km, corresponding to a period range of about 4–5 days for this type of disturbance.

Diagnostic studies suggest that these westward-propagating wave disturbances are generally driven by the release of latent heat in their accompanying convective precipitation areas. The vertical structure of a typical equatorial wave disturbance is shown in schematic form in Fig. 11.4. Vertical motion in such disturbances is proportional to the diabatic heating rate; thus the maximum vertical velocities occur in the convective zone. By mass continuity there must be convergence at low levels in the convective zone and divergence in the upper levels. Hence, provided that the absolute vorticity has the sign of f, the divergence term in the vorticity equation will induce cyclonic vorticity tendencies in the lower troposphere and anticyclonic vorticity tendencies in the upper troposphere. The process of adjustment between mass and velocity fields will then generate a low-level trough and an upper-level ridge.[1] Thus, the thickness (or layer mean temperature) in the convective zone must be greater than in the surrounding environment.

[1] The terms "trough" and "ridge" as used by tropical meteorologists designate pressure minima and maxima, respectively, just as in midlatitudes. In the easterlies of the Northern Hemisphere tropics, however, the zonal-mean pressure increases with latitude so that the pattern of isobars depicting a tropical trough will resemble the pattern associated with a ridge in middle latitudes (i.e., there is a poleward deflection of the isobars).

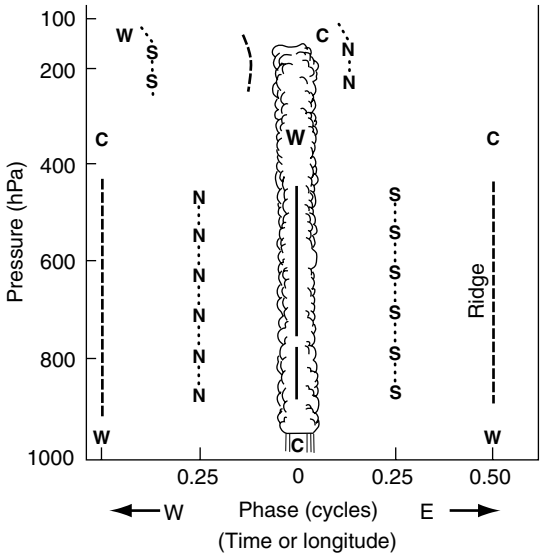


Fig. 11.4 Schematic model for equatorial wave disturbances showing trough axis (solid line), ridge axis (dashed lines), and axes of northerly and southerly wind components (dotted lines). Regions of warm and cold air are designated by W and C, respectively. (After Wallace, 1971.)

In the convectively active areas of equatorial waves there is net upward motion, and midtropospheric temperatures are warmer than average (although generally by less than 1°C). The correlations between temperature and vertical motion, and between temperature and diabatic heating, are thus both positive, and the potential energy generated by the diabatic heating is immediately converted to kinetic energy [i.e., the $[P' \bullet K']$ conversion balances R' in (10.62)]. There is, in this approximation, no storage in the form of available potential energy. The energy cycle of these disturbances, therefore, differs remarkably from that of midlatitude baroclinic systems in which the available potential energy greatly exceeds the kinetic energy.

For latent heat release by cumulonimbus clouds to be an effective energy source for large-scale disturbances there must be an interaction between the convective scale and the large scale, as mentioned in Section 9.7.2. In such interaction large-scale convergence at low levels moistens and destabilizes the environment so that small-scale thermals can easily reach the level of free convection and produce deep cumulus convection. The cumulus cells, in turn, act cooperatively to provide the large-scale heat source that drives the secondary circulation responsible for the low-level convergence.

A typical vertical profile of divergence in the precipitation zone of a synoptic-scale equatorial wave disturbance in the western Pacific is shown in Fig. 11.5. Convergence is not limited to low-level frictional inflow in the planetary boundary layer, but extends up to nearly 400 hPa, which is the height where the hot towers

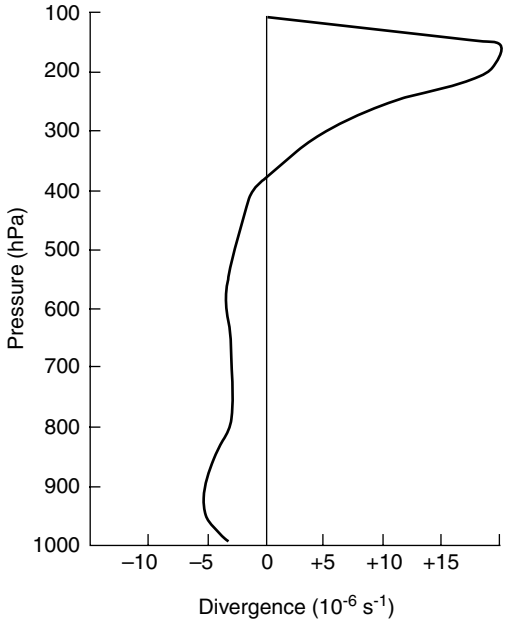


Fig. 11.5 Vertical profile of 4°-square area average divergence based on composites of many equatorial disturbances. (Adapted from Williams, 1971.)

achieve their maximum buoyancy. The deep convergence implies that there must be substantial entrainment of midtropospheric air into the convective cells. Because the midtropospheric air is relatively dry, this entrainment will require considerable evaporation of liquid water to bring the mixture of cloud and environment air to saturation. It will thus reduce the buoyancy of the cloud air and may in fact produce negatively buoyant convective downdrafts if there is sufficient evaporative cooling. However, in the large cumulonimbus clouds present in equatorial waves, the central core updrafts are protected from entrainment by the surrounding cloud air so that they can penetrate nearly to the tropopause without dilution by environmental air. These undiluted cores constitute the "hot towers" referred to in Section 11.1.1. Because the hot towers are responsible for much of the vertical heat and mass transport above the boundary layer in the ITCZ, and the wave disturbances contain most of the active convective precipitation areas along the ITCZ, it is obvious that equatorial waves play an essential role in the general circulation of the atmosphere.

11.1.3 African Wave Disturbances

The considerations of the previous subsection are valid for ITCZ disturbances over most regions of the tropical oceans. In the region of the North African

continent, however, local effects due to surface conditions create a unique situation that requires separate discussion. During the Northern Hemisphere summer, intense surface heating over the Sahara generates a strong positive meridional temperature gradient in the lower troposphere between the equator and about 25°N. The resulting easterly thermal wind is responsible for the existence of a strong easterly jet core near 650 hPa centered near 16°N as shown in Fig. 11.6. Synoptic-scale disturbances are observed to form and propagate westward in the cyclonic shear zone to the south of this jet core. Occasionally, such disturbances are progenitors of tropical storms and hurricanes in the western Atlantic. The average wavelength of observed African wave disturbances is about 2500 km, and the westward propagation speed is about 8 ms^{-1}, implying a period of about 3.5 days. The disturbances have horizontal velocity perturbations that reach maximum amplitude at the 650-hPa level, as indicated in Fig. 11.7. Although there is considerable organized convection associated with these waves, they do not appear to be driven primarily by latent heat release, but depend, rather, on barotropic and baroclinic conversions of energy from the easterly jet.

In Fig. 11.8 the absolute vorticity profile for the African easterly jet shown in Fig. 11.6 is plotted. The shaded region indicates the area in which the vorticity gradient is negative. Thus, it is clear that the African jet satisfies the necessary

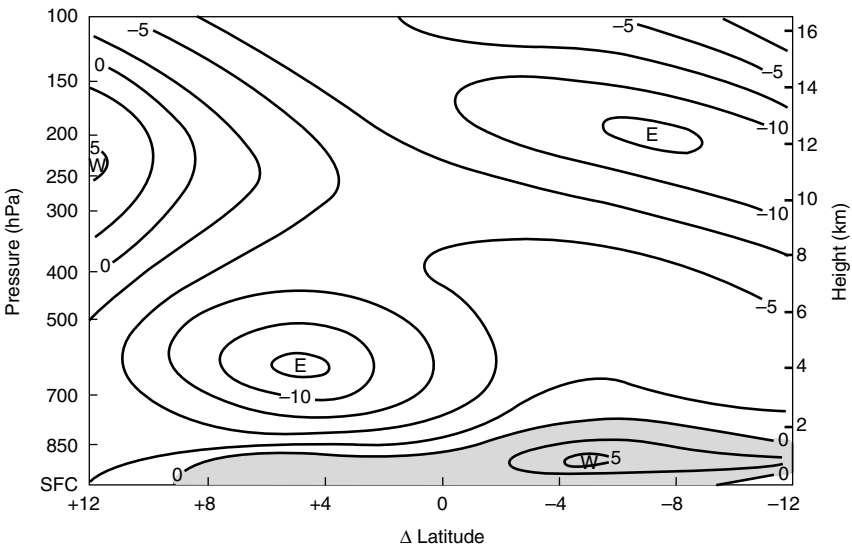


Fig. 11.6 Mean-zonal wind distribution in the North African region (30°W to 10°E longitude) for the period August 23, 1974 to September 19, 1974. Latitude is shown relative to latitude of maximum disturbance amplitude at 700 hPa (about 12°N). The contour interval is 2.5 m s^{-1}. (After Reed et al., 1977. Reproduced with permission of the American Meteorological Society.)

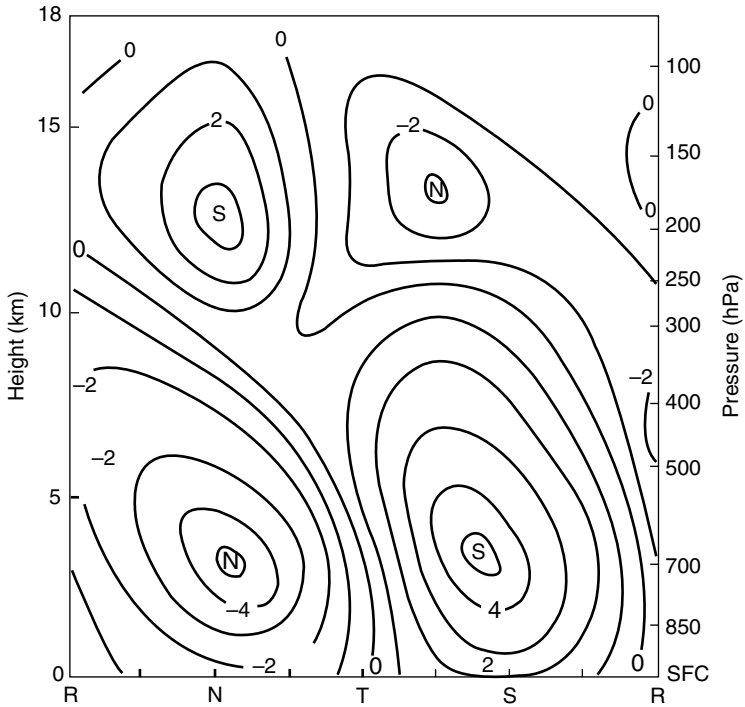


Fig. 11.7 Vertical cross section along the reference latitude of Fig. 11.6 showing perturbation meridional velocities in $m\,s^{-1}$. R, N, T, and S refer to ridge, northwind, trough, and southwind sectors of the wave, respectively. (After Reed et al., 1977. Reproduced with permission of the American Meteorological Society.)

condition for barotropic instability discussed in Section 8.4.2.[2] Baroclinic instability due to the strong easterly shear in the lower troposphere also appears to play a role in these disturbances. Thus, both barotropic and baroclinic conversions from the mean flow energy appear to be important for the generation of African wave disturbances.

Because such disturbances continue to exist in the absence of strong mean wind shears after they have propagated westward into the Atlantic, it is unlikely that either baroclinic or barotropic instability continues to be the primary energy source for their maintenance. Rather, diabatic heating through precipitating convective systems appears to be the main energy source for such waves over the ocean.

[2] It should be noted here that the profile shown in Fig. 11.6 is not a zonal mean. Rather, it is a time mean for a limited longitudinal domain. Provided that the longitudinal scale of variation of this time mean zonal flow is large compared to the scale of the disturbance, the time mean flow may be regarded as a locally valid basic state for linear stability calculations.

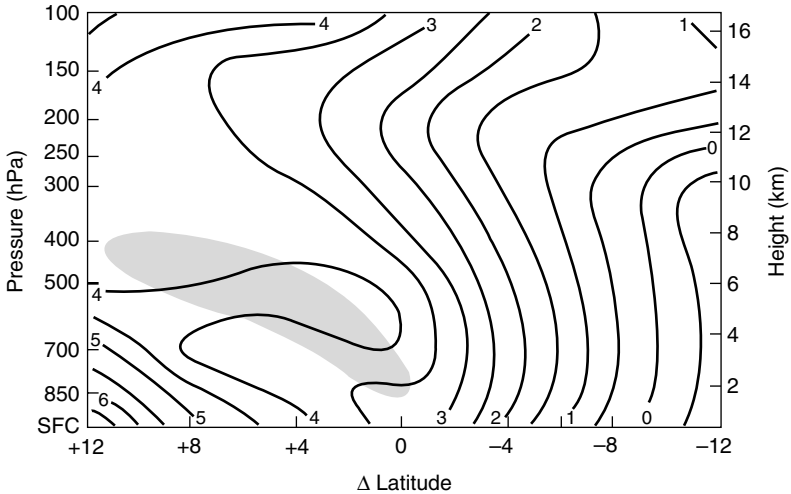


Fig. 11.8 Absolute vorticity (units 10^{-5}s^{-1}) corresponding to the mean wind field of Fig. 11.6. Shading shows region where $\beta - \partial^2\bar{u}/\partial y^2$ is negative. (After Reed et al., 1977. Reproduced with permission of the American Meteorological Society.)

11.1.4 Tropical Monsoons

The term "monsoon" is commonly used in a rather general sense to designate any seasonally reversing circulation system. The basic drive for a monsoon circulation is provided by the contrast in the thermal properties of the land and sea surfaces. Because the thin layer of soil that responds to the seasonal changes in surface temperature has a small heat capacity compared to the heat capacity of the upper layer of the ocean that responds on a similar time scale, the absorption of solar radiation raises the surface temperature over land much more rapidly than over the ocean. The warming of the land relative to the ocean leads to enhanced cumulus convection, and hence to latent heat release, which produces warm temperatures throughout the troposphere.

Much of the tropics is influenced by monsoons. The most extensive monsoon circulation by far is the Asian monsoon. This monsoon completely dominates the climate of the Indian subcontinent, producing warm wet summers and cool dry winters. An idealized model of the structure of the Asian summer monsoon is indicated in Fig. 11.9. As indicated in Fig. 11.9, the 1000-200-hPa thickness is larger over the land than over the ocean. As a result there is a pressure gradient force at the upper levels directed from the land to the ocean. The divergent wind that develops in response to this pressure gradient (shown by the arrows in Fig. 11.9) causes a net mass transport out of the air column above the continent and thereby generates a surface low over the continent (often called a *thermal low*). A compensating

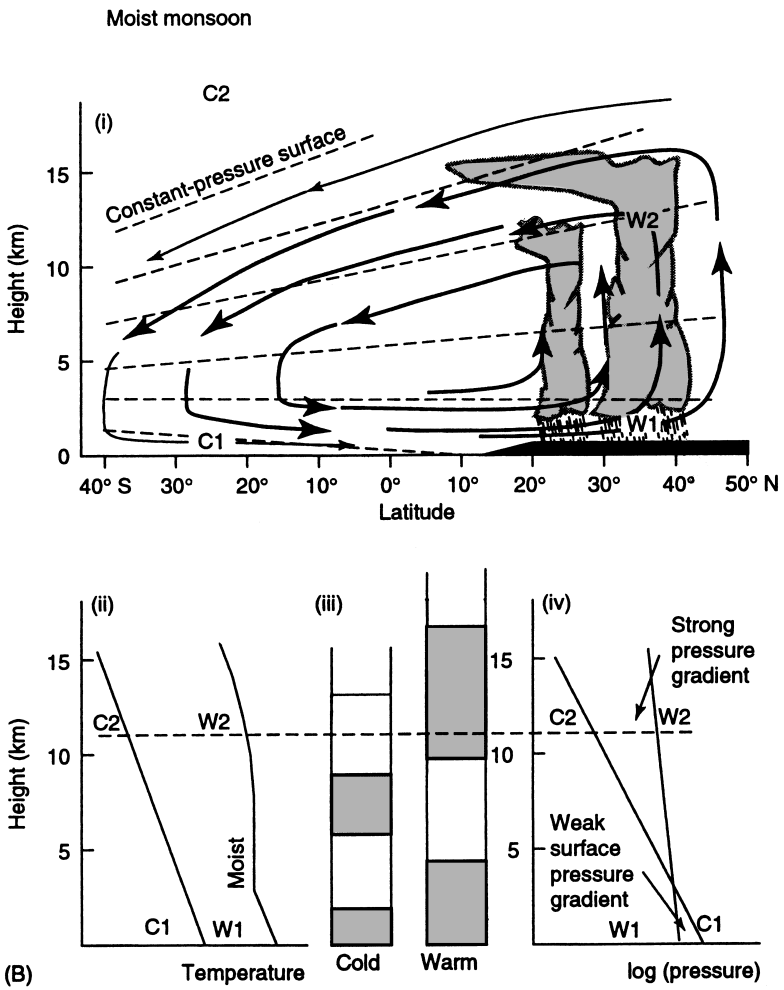


Fig. 11.9 Schematic representation of the moist phase of the Asian monsoon. (i) Arrows show the meridional circulation, dashed lines the isobars. (ii) Temperature profiles for columns C1–C2 and W1–W2, respectively. (iii) Schematic of the mass distribution in the cold and warm sectors. (iv) Horizontal pressure difference as a function of height. (After Webster and Fasullo, 2003.)

convergent wind then develops at low levels. This low-level flow produces a convergence of moisture, which by increasing the equivalent potential temperature in the boundary layer makes the environment more favorable for development of the cumulus convection, which is the primary energy source for the monsoon circulation.

The low-level convergence and upper-level divergence over the continent constitute a secondary circulation that concentrates cyclonic vorticity at the lower levels and anticyclonic vorticity at the upper levels. Thus, the vorticity adjusts toward geostrophic balance. From Fig. 11.9 it is clear that a positive correlation exists between the vertical motion and the temperature field. Therefore, the monsoon circulation converts eddy potential energy to eddy kinetic energy, just as midlatitude baroclinic eddies do.

Unlike the case of baroclinic eddies, however, the primary energy cycle of the monsoons does not involve the zonal mean potential or kinetic energy. Rather, eddy potential energy is generated directly by diabatic heating (latent and radiative heating); the eddy potential energy is converted to eddy kinetic energy by a thermally direct secondary circulation; and the eddy kinetic energy is frictionally dissipated. (A portion of the eddy kinetic energy may be converted to zonal kinetic energy.) In a dry atmosphere, monsoon circulations would still exist; however, because the diabatic heating would then be confined to a shallow layer near the surface, they would be much weaker than the observed monsoons. The presence of cumulus convection and its concomitant latent heat release greatly amplifies the eddy potential energy generation and makes the summer monsoons among the most important features of the global circulation.

In the winter season the thermal contrast between the land and the sea reverses so that the circulation is just opposite to that shown in Fig. 11.9. As a result the continents are cool and dry, and the precipitation is found over the relatively warm oceans.

11.1.5 The Walker Circulation

The pattern of diabatic heating in the equatorial regions exhibits strong departures from zonal symmetry. These are caused by longitudinal variations in sea surface temperature due mainly to the effects of wind-driven ocean currents. Such SST variations produce zonally asymmetric atmospheric circulations, which in some regions dominate over the Hadley circulation. Of particular significance is the east–west overturning along the equator, which is shown schematically in Fig. 11.10. Several overturning cells are indicated, which are associated with diabatic heating over equatorial Africa, Central and South America, and the Maritime Continent (i.e., the Indonesian region). The dominant cell in both zonal scale and amplitude, however, is that in the equatorial Pacific. This cell is referred to as *Walker circulation*, after G. T. Walker who first documented the surface pressure pattern associated with it.

As suggested by Fig. 11.10, this pressure pattern consists of low surface pressure in the western Pacific and high surface pressure in the eastern Pacific. The resulting westward-directed pressure gradient force drives mean surface easterlies in the equatorial Pacific, which are much stronger than the zonal-mean surface easterlies,

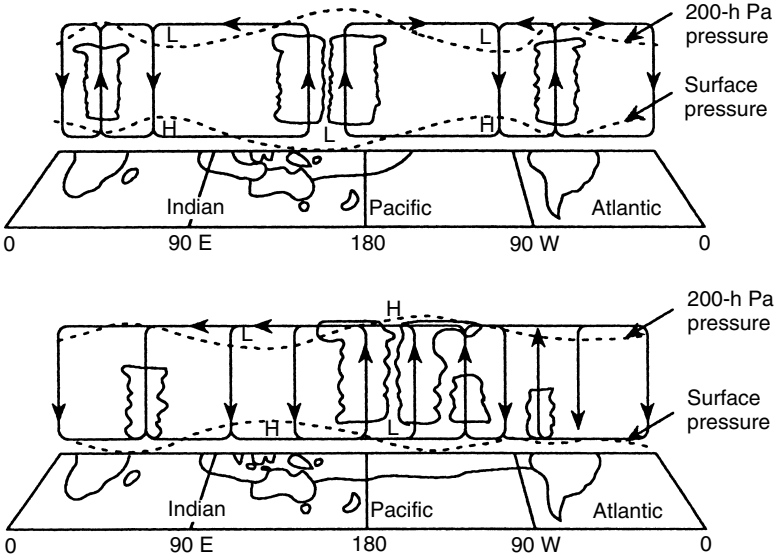


Fig. 11.10 Schematic diagrams of the Walker circulations along the equator for normal conditions (top) and El Niño conditions (bottom). (After Webster, 1983 and Webster and Chang, 1988.)

and by horizontal vapor transport provides a moisture source for convection in the western Pacific, in addition to that provided by the high evaporation rates caused by the high sea surface temperatures in that region.

The wind stress due to the time mean equatorial easterly surface winds over the Pacific has a strong influence on the heat balance of the ocean surface layer. It advects warm surface waters into the western Pacific and produces poleward drifts in the oceanic Ekman layer, which by continuity drive an equatorial upwelling. This upwelling accounts for the cold tongue of water along the equator, which in turn is a major reason for the equatorial dry zone exhibited in Fig. 11.2.

11.1.6 El Niño and the Southern Oscillation

The east–west pressure gradient associated with the Walker circulation undergoes an irregular interannual variation. This global scale "see-saw" in pressure, and its associated changes in patterns of wind, temperature, and precipitation, was named the *southern oscillation* by Walker. This oscillation can be clearly seen by comparing time series of surface pressure anomalies (i.e., departures from the long-term mean) for locations on the western and eastern sides of the equatorial Pacific. As shown in the upper portion of Fig. 11.11, surface pressure anomalies at Darwin Australia and Tahiti are negatively correlated and have strong variations in

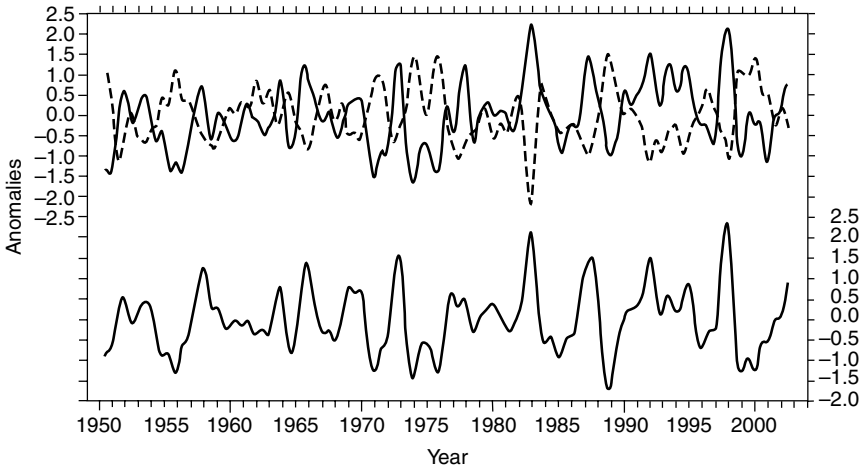


Fig. 11.11 Time series showing SST anomalies (°C) in the eastern Pacific (lower curve) and anomalies in sea level pressures (hPa) at Darwin (upper solid curve) and Tahiti (upper dashed curve). Data are smoothed to eliminate fluctuations with periods less than a year. Figure courtesy of Dr. Todd Mitchell, University of Washington.

the period range of 2–5 years. During periods of large pressure difference between these stations the Walker circulation is unusually strong and has the same general structure as its time mean pattern. In periods of weak pressure difference the Walker circulation weakens and the region of maximum rainfall shifts eastward (see Fig. 11.10, bottom).

The weakening of trade winds during periods of small Darwin–Tahiti pressure difference reduces wind-driven oceanic upwelling and deepens the oceanic thermocline in the eastern Pacific. This in turn, as shown in the lower portion of Fig. 11.11, leads to an increase in the sea surface temperature that is referred to as *El Niño* (Spanish for "the child"). This term was originally applied to a warming of the coastal waters of Peru and Ecuador that occurs annually near Christmas time (hence, El Niño refers to the Christ child), but is now used in a more general sense for the large-scale oceanic anomalies associated with the weak wind phase of the Southern Oscillation. In the opposite phase, strengthening of trade winds associated with large Darwin–Tahiti pressure differences enhances oceanic upwelling and deepens the oceanic thermocline. This causes a cooling of the sea surface temperatures in the equatorial Pacific that is now referred to as *La Niña*.

The ocean anomalies associated with weakening in the trade winds may begin near the coast, but over the course of several months spread westward along the equator to produce large-scale positive SST anomalies over much of the equatorial Pacific. These SST anomalies in turn cause a further weakening of the trade winds. This entire complex of atmospheric and oceanic variations is now referred to as

ENSO (for El Niño—Southern Oscillation). It is a dramatic example of interannual climate variability associated with atmosphere–ocean coupling.

The leading theoretical model for ENSO is the "delayed oscillator" model. In this model the sea surface temperature anomaly, T, in the eastern Pacific satisfies an equation of the form

$$\frac{dT}{dt} = bT(t) - cT(t - \tau)$$

where b and c are positive constants and τ is a time delay determined by the adjustment time for the equatorial ocean. The first term on the right represents a positive feedback associated with changes in the Darwin to Tahiti pressure difference. This term represents the atmosphere–ocean coupling through which an initial weakening of the wind causes an increase of the SST, which induces a further weakening of the wind, and hence a further increase in the SST. The second term provides a negative feedback due to the adjustment in the thermocline depth (and hence the ocean temperature) caused by propagating equatorial waves (see Section 11.4) in the ocean that are excited by the SST changes. The time delay in the negative feedback term is determined by the time that it takes for wave energy excited by air–sea interaction in the eastern Pacific to propagate to the western boundary of the ocean, undergo reflection, and propagate back to the region of origin. For realistic parameters the delayed oscillator model leads to ENSO oscillations in the period range of 3–4 years. This highly simplified model qualitatively accounts for the average characteristics of an ENSO cycle, but cannot, however, account for the observed irregularity of ENSO.

In addition to its profound effects in the equatorial region, ENSO is associated with a wide range of interannual climate anomalies in the extratropics. Thus, the development of models that show skill in predicting ENSO several months in advance is of considerable practical importance.

11.1.7 Equatorial Intraseasonal Oscillation

In addition to the interannual variability associated with El Niño, the equatorial circulation has an important intraseasonal oscillation, which occurs on a timescale of 30–60 days and is often referred to as the Madden–Julian oscillation (MJO) in honor of the meteorologists who first described it. The structure of the equatorial intraseasonal oscillation is shown schematically in Fig. 11.12, which shows the time development of the oscillation in the form of longitude–height sections along the equator, with time increasing at an interval of about 10 days for each panel from top to bottom. The circulations in Fig. 11.12 are intended to represent anomalies from the time-mean equatorial circulation.

The oscillation originates with development of a surface low-pressure anomaly over the Indian Ocean, accompanied by enhanced boundary layer moisture

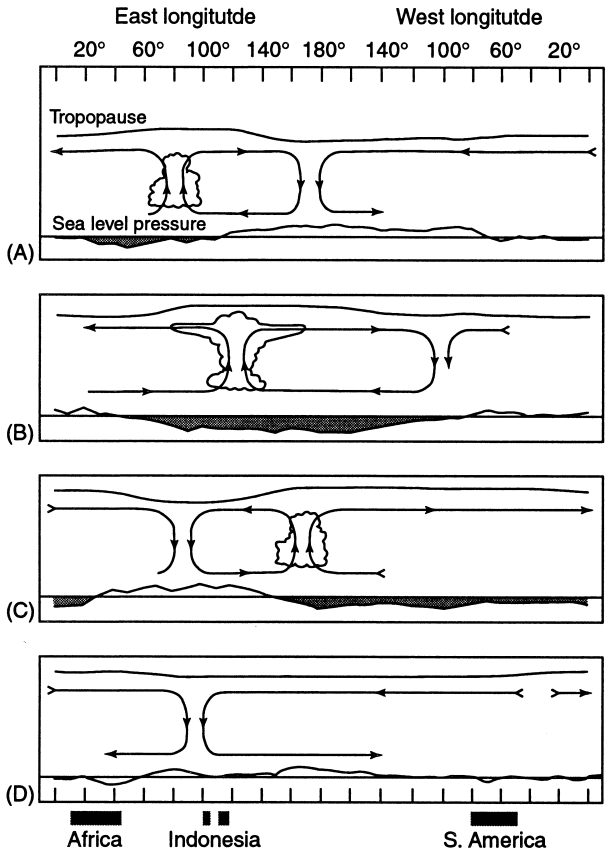


Fig. 11.12 Longitude-height section of the anomaly pattern associated with the tropical intraseasonal oscillation (MJO). Reading downward the panels represent a time sequence with intervals of about 10 days. Streamlines show the west–east circulation, wavy top line represents the tropopause height, and bottom line represents surface pressure (with shading showing below normal surface pressure). (After Madden, 2003; adapted from Madden and Julian, 1972.)

convergence, increased convection, warming of the troposphere, and raising of the tropopause height. The anomaly pattern gradually moves eastward at about $5\,\mathrm{m\,s^{-1}}$ and reaches maximum intensity over the western Pacific. As the anomaly moves over the cooler waters of the central Pacific, the anomalous convection gradually weakens, although a circulation disturbance continues eastward and can sometimes be traced completely around the globe. The observed intraseasonal oscillation is known to be associated with equatorial Rossby and Kelvin waves (see Section 11.4). However, a completely satisfactory theory for the oscillation has not yet been developed.

11.2 SCALE ANALYSIS OF LARGE-SCALE TROPICAL MOTIONS

Despite the uncertainties involved in the interaction between convective and synoptic scales, some information on the character of synoptic-scale motions in the tropics can be obtained through the methods of scale analysis. The scaling arguments can be carried out most conveniently if the governing equations are written in the log-pressure coordinate system introduced in Chapter 8:

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla + w^* \frac{\partial}{\partial z^*}\right)\mathbf{V} + f\mathbf{k} \times \mathbf{V} = -\nabla\Phi \tag{11.1}$$

$$\partial\Phi/\partial z^* = RT/H \tag{11.2}$$

$$\partial u/\partial x + \partial v/\partial y + \partial w^*/\partial z^* - w^*/H = 0 \tag{11.3}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla\right) T + w^* N^2 H/R = J/c_p \tag{11.4}$$

We wish to compare the magnitudes of the various terms in (11.1)–(11.4) for synoptic-scale motions in the tropics. We first note that an upper limit on the vertical velocity scale W is imposed by the continuity equation (11.3). Thus, following the discussion of Section 4.5,

$$\partial u/\partial x + \partial v/\partial y \leq U/L$$

However, for motions with vertical scales comparable to the density scale height H,

$$\partial w^*/\partial z^* - w^*/H \sim W/H$$

so that the vertical velocity scale must satisfy the constraint $W \leq HU/L$ if the horizontal divergence and vertical stretching terms are to balance in the continuity equation. We next define characteristic scales for the various field variables as follows:

$H \sim 10^4$ m	vertical length scale
$L \sim 10^6$ m	horizontal length scale
$U \sim 10\,\mathrm{m\,s^{-1}}$	horizontal velocity scale
$W \leq HU/L$	vertical velocity scale
$\delta\Phi$	geopotential fluctuation scale
$L/U \sim 10^5$ s	timescale for advection

The magnitudes chosen for the horizontal length and velocity scales are typical for observed values in synoptic-scale systems both in the tropics and at midlatitudes. We now wish to show how the corresponding characteristic scales for vertical velocity and geopotential fluctuations are limited by the dynamic constraints imposed by conservation of mass, momentum, and thermodynamic energy.

We can estimate the magnitude of the geopotential fluctuation $\delta\Phi$ by scaling the terms in the momentum equation (11.1). For this purpose it is convenient to compare the magnitude of the horizontal inertial acceleration,

$$(\mathbf{V} \bullet \boldsymbol{\nabla})\mathbf{V} \sim U^2/L$$

with each of the other terms in (11.1) as follows:

$$|\partial\mathbf{V}/\partial t| \,/\, |(\mathbf{V} \bullet \boldsymbol{\nabla})\mathbf{V}| \sim 1 \tag{11.5}$$

$$|w^*\partial\mathbf{V}/\partial z^*| / \,|(\mathbf{V} \bullet \boldsymbol{\nabla})\mathbf{V}| \sim WL/UH \leq 1 \tag{11.6}$$

$$|f\mathbf{k} \times \mathbf{V}| \,/\, |(\mathbf{V} \bullet \boldsymbol{\nabla})\mathbf{V}| \sim fL/U = \mathrm{Ro}^{-1} \leq 1 \tag{11.7}$$

$$|\boldsymbol{\nabla}\Phi| \,/\, |(\mathbf{V} \bullet \boldsymbol{\nabla})\mathbf{V}| \sim \delta\Phi/U^2 \tag{11.8}$$

We have shown previously that in middle latitudes where $f \sim 10^{-4}\,\mathrm{s}^{-1}$, the Rossby number Ro is small so that to first approximation the Coriolis force and pressure gradient force terms balance. In that case, $\delta\Phi \sim fUL$. In the equatorial region, however, $f \leq 10^{-5}\,\mathrm{s}^{-1}$ and the Rossby number is of order unity or greater. Therefore, it is not appropriate to assume that the Coriolis force balances the pressure gradient force. In fact, (11.5)–(11.8) show that if the pressure gradient force is to be balanced in (11.1), the geopotential perturbation must scale as $\delta\Phi \sim U^2 \sim 100\,\mathrm{m}^2\,\mathrm{s}^{-2}$, and geopotential perturbations associated with equatorial synoptic-scale disturbances will be an order of magnitude smaller than those for midlatitude systems of similar scale.

This constraint on the amplitude of geopotential fluctuations in the tropics has profound consequences for the structure of synoptic-scale tropical motion systems. These consequences can be understood easily by applying scaling arguments to the thermodynamic energy equation. It is first necessary to obtain an estimate of the temperature fluctuations. The hydrostatic approximation (11.2) implies that for systems whose vertical scale is comparable to the scale height

$$T = (H/R)\partial\Phi/\partial z^* \sim (\delta\Phi/R) \sim U^2/R \sim 0.3\,\mathrm{K} \tag{11.9}$$

Therefore, deep tropical systems are characterized by practically negligible synoptic-scale temperature fluctuations. Referring to the thermodynamic energy equation, we find that for such systems

$$\left(\frac{\partial}{\partial t} + \mathbf{V}\bullet\nabla\right) T \sim 0.3\,\mathrm{K\,d}^{-1}$$

In the absence of precipitation the diabatic heating is caused primarily by the emission of long wave radiation, which tends to cool the troposphere at a rate

of $J/c_p \sim 1\,\mathrm{K\,d^{-1}}$. Because the actual temperature fluctuations are small, this radiative cooling must be approximately balanced by adiabatic warming due to subsidence. Thus, to a first approximation, (11.4) becomes a diagnostic relationship for w^*:

$$w^*(N^2 H/R) = J/c_p \tag{11.10}$$

For the tropical troposphere, $N^2 H/R \sim 3\,\mathrm{K\,km^{-1}}$, thus the vertical motion scale must satisfy $W \sim 0.3\,\mathrm{cm\,s^{-1}}$, and $WL/UH \sim 0.03$ in (11.6). Therefore, in the absence of precipitation the vertical motion is constrained to be even smaller than in extratropical synoptic systems of a similar scale. Not only can the vertical advection term be neglected in (11.1), but from the continuity equation (11.3), the divergence of the horizontal wind is $\sim 3 \times 10^{-7}\,\mathrm{s^{-1}}$. Thus, the flow is nearly nondivergent.

The quasi-nondivergent nature of the flow in the absence of convective disturbances in the tropics makes it possible to simplify the governing equations for that situation. A theorem of Helmholtz[3] states that any velocity field can be divided into a *nondivergent* part $\mathbf{V}_\psi$ plus an i*rrotational* part $\mathbf{V}_e$ such that

$$\mathbf{V} = \mathbf{V}_\psi + \mathbf{V}_e$$

where $\boldsymbol{\nabla} \bullet \mathbf{V}_\psi = 0$ and $\boldsymbol{\nabla} \times \mathbf{V}_e = 0$.

For a two-dimensional velocity field the nondivergent part can be expressed in terms of a *streamfunction* ψ defined by letting

$$\mathbf{V}_\psi = \mathbf{k} \times \boldsymbol{\nabla}\psi \tag{11.11}$$

or in Cartesian components:

$$u_\psi = -\partial\psi/\partial y, \quad v_\psi = \partial\psi/\partial x$$

from which it is verified easily that $\boldsymbol{\nabla} \bullet \mathbf{V}_\psi = 0$ and $\zeta = \mathbf{k} \bullet \boldsymbol{\nabla} \times \mathbf{V}_\psi = \boldsymbol{\nabla}^2\psi$. Because the isolines of ψ correspond to streamlines for the nondivergent velocity and the distance separating the isolines of ψ is inversely proportional to the magnitude of the nondivergent velocity, the spatial distribution of $\mathbf{V}_\psi$ can be pictured easily by plotting lines of constant ψ on a synoptic chart.

We now approximate $\mathbf{V}$ by its nondivergent part $\mathbf{V}_\psi$ in (11.1) and neglect the small vertical advection term to obtain an approximate momentum equation valid for synoptic-scale motions in the tropics outside regions of precipitation:

$$\frac{\partial \mathbf{V}_\psi}{\partial t} + \left(\mathbf{V}_\psi \bullet \boldsymbol{\nabla}\right)\mathbf{V}_\psi + f\mathbf{k} \times \mathbf{V}_\psi = -\boldsymbol{\nabla}\Phi \tag{11.12}$$

[3] See, for example, Bourne and Kendall (1968, p. 190).

Using the vector identity

$$(\mathbf{V} \bullet \nabla)\mathbf{V} = \nabla\left(\frac{\mathbf{V} \bullet \mathbf{V}}{2}\right) + \mathbf{k} \times \mathbf{V}\zeta$$

we can rewrite (11.12) as

$$\frac{\partial \mathbf{V}_\psi}{\partial t} = -\nabla\left(\Phi + \frac{\mathbf{V}_\psi \bullet \mathbf{V}_\psi}{2}\right) - \mathbf{k} \times \mathbf{V}_\psi\,(\zeta + f) \tag{11.13}$$

We next take $\mathbf{k} \bullet \nabla \times$ (11.13) to obtain the vorticity equation

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_\psi \bullet \nabla\right)(\zeta + f) = 0 \tag{11.14}$$

valid for nondivergent flow. This equation shows that in the absence of condensation heating, synoptic-scale circulations in the tropics in which the vertical scale is comparable to the scale height of the atmosphere must be barotropic; absolute vorticity is conserved following the nondivergent horizontal wind. Such disturbances cannot convert potential energy to kinetic energy. They must be driven by barotropic conversion of mean-flow kinetic energy or by lateral coupling either to midlatitude systems or to precipitating tropical disturbances.

Because both the nondivergent velocity and the vorticity can be expressed in terms of the streamfunction, (11.14) requires only the field of ψ at any level in order to make a prediction. The pressure distribution is neither required nor predicted. Rather, it must be determined diagnostically. The relationship of the pressure and streamfunction fields can be obtained by taking $\nabla\bullet$(11.13). This yields a diagnostic relationship between the geopotential and streamfunction fields, which is usually referred to as the *nonlinear balance equation*:

$$\nabla^2\left[\Phi + \frac{1}{2}(\nabla\psi)^2\right] = \nabla\bullet\left[\left(f + \nabla^2\psi\right)\nabla\psi\right] \tag{11.15}$$

For the special case of stationary circularly symmetric flow, (11.15) is equivalent to the gradient wind approximation. Unlike the gradient wind, however, the balance in (11.15) does not require information on trajectory curvature, and thus can be solved for Φ from knowledge of the instantaneous distribution of ψ on an isobaric surface. Alternatively, if the Φ distribution is known, (11.15) can be solved for ψ. In this case the equation is quadratic so that there are generally two possible solutions, which correspond to the normal and anomalous gradient wind cases.

Such a balance condition is valid only when the above scaling arguments apply. These, however, have been based on the assumptions that the depth scale is comparable to the scale height of the atmosphere and that the horizontal scale is of

order 1000 km. There is a special class of planetary scale motions for which the divergence term in the vorticity equation is important even outside of regions of active precipitation (see Section 11.4). For such motions the pressure field cannot be diagnosed from a balance relationship. Rather, the pressure distribution must be predicted from the primitive equation form of the dynamics equations.

For precipitating synoptic-scale systems in the tropics, the above scaling considerations require considerable modification. Precipitation rates in such systems are typically of order 2 cm d^{-1}. This implies condensation of $m_w = 20$ kg water for an atmospheric column of 1 m^2 cross section. Because the latent heat of condensation is $L_c \approx 2.5 \times 10^6$ J kg^{-1}, this precipitation rate implies an addition of heat energy to the atmospheric column of

$$m_w L_c \sim 5 \times 10^7 \text{ J m}^{-2} \text{ d}^{-1}$$

If this heat is distributed uniformly over the entire atmospheric column of mass $p_0/g \approx 10^4$ kg m^{-2}, then the average heating rate per unit mass of air is

$$J/c_p \approx \left[L_c m_w / c_p \left(p_0/\text{g}\right)\right] \sim 5 \text{ K d}^{-1}$$

In reality the condensation heating due to deep convective clouds is not distributed evenly over the entire vertical column, but is a maximum between 300 and 400 hPa, where the heating rate can be as high as 10 K d^{-1}. In this case the approximate thermodynamic energy equation (11.10) implies that the vertical motion on the synoptic scale in precipitating systems must have a magnitude of order $W \sim 3$ cm s^{-1} in order that the adiabatic cooling can balance the condensation heating in the 300-400-hPa layer. Therefore, the average vertical motion in precipitating disturbances in the tropics is an order of magnitude larger than the vertical motion outside the disturbances. As a result the flow in these disturbances has a relatively large divergent component so that the barotropic vorticity equation (11.14) is no longer a reasonable approximation, and the full primitive equations must be used to analyze the flow.

11.3 CONDENSATION HEATING

The manner in which the atmosphere is heated by condensation of water vapor depends crucially on the nature of the condensation process. In particular, it is necessary to differentiate between latent heat release through large-scale vertical motion (i.e., synoptic-scale forced uplift) and the latent heat release due to deep cumulus convection. The former process, which is generally associated with midlatitude synoptic systems, can be incorporated easily into the thermodynamic energy equation in terms of the synoptic-scale field variables. The large-scale heating field resulting from the cooperative action of many cumulonimbus cells,

however, requires representation of this type of latent heating in terms of the synoptic-scale field variables, which is much more difficult.

Before considering the problem of condensation heating by cumulus convection, it is worth indicating briefly how the condensation heating by large-scale forced uplift can be included in a prediction model. The approximate thermodynamic energy equation for a pseudoadiabatic process (9.38) states that

$$\frac{D \ln \theta}{Dt} \approx -\left(\frac{L_c}{c_p T}\right) \frac{Dq_s}{Dt} \tag{11.16}$$

The change in q_s following the motion is caused primarily by ascent, so that

$$\frac{Dq_s}{Dt} \approx \begin{cases} w \partial q_s / \partial z & \text{for } w > 0 \\ 0 & \text{for } w < 0 \end{cases} \tag{11.17}$$

and (11.16) can be written in the form

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla\right) \ln \theta + w \left(\frac{\partial \ln \theta}{\partial z} + \frac{L_c}{c_p T} \frac{\partial q_s}{\partial z}\right) \approx 0 \tag{11.18}$$

for regions where $w > 0$. However, from (9.40)

$$\frac{\partial \ln \theta_e}{\partial z} \approx \frac{\partial \ln \theta}{\partial z} + \frac{L_c}{c_p T} \frac{\partial q_s}{\partial z}$$

so that (11.18) can be written in a form valid for both positive and negative vertical motion as

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla\right) \theta + w \Gamma_e \approx 0 \tag{11.19}$$

where Γ_e is an *equivalent static stability* defined by

$$\Gamma_e \approx \begin{cases} \theta \partial \ln \theta_e / \partial z & \text{for } q \geq q_s \quad \text{and } w > 0 \\ \partial \theta / \partial z & \text{for } q < q_s \quad \text{or } w < 0 \end{cases}$$

Thus, in the case of condensation due to large-scale forced ascent ($\Gamma_e > 0$), the thermodynamic energy equation has essentially the same form as for adiabatic motions except that the static stability is replaced by the equivalent static stability. As a consequence, the local temperature changes induced by the forced ascent will be smaller than for the case of forced dry ascent with the same lapse rate.

If, however, $\Gamma_e < 0$, the atmosphere is conditionally unstable and condensation will occur primarily through cumulus convection. In that case, (11.17) is still valid, but the vertical velocity must be that of the individual cumulus updrafts, not the synoptic-scale w. Thus, a simple formulation of the thermodynamic energy

equation in terms of only the synoptic-scale variables is not possible. We can still, however, simplify the thermodynamic energy equation to some extent. We recall from Section 11.2 [see Eq. (11.10)] that due to the smallness of temperature fluctuations in the tropics, the adiabatic cooling and diabatic heating terms must approximately balance. Thus, (11.16) becomes approximately

$$w\frac{\partial \ln\theta}{\partial z} \approx -\frac{L_c}{c_pT}\frac{Dq_s}{Dt} \tag{11.20}$$

The synoptic-scale vertical velocity w that appears in (11.20) is the average of very large vertical motions in the active convection cells and small vertical motions in the environment. Thus, if we let w' be the vertical velocity in the convective cells and the $\overline{w}$ vertical velocity in the environment, we have

$$w = aw' + (1-a)\overline{w} \tag{11.21}$$

where a is the fractional area occupied by the convection. With the aid of (11.17), we can than write (11.20) in the form

$$w\frac{\partial \ln\theta}{\partial z} \approx -\frac{L_c}{c_pT}aw'\frac{\partial q_s}{\partial z} \tag{11.22}$$

The problem is then to express the condensation heating term on the right in (11.22) in terms of synoptic-scale field variables.

This problem of *parameterizing* the cumulus convective heating is one of the most challenging areas in tropical meteorology. A simple approach that has been used successfully in some theoretical studies[4] is based on the fact that, because the storage of water in the clouds is rather small, the total vertically integrated heating rate due to condensation must be approximately proportional to the net precipitation rate:

$$-\int_{z_c}^{z_T}\left(\rho aw'\partial q_s/\partial z\right)dz = P \tag{11.23}$$

where z_c and z_T represent the cloud base and cloud top heights, respectively, and P is the precipitation rate (kg m^{-2} s^{-1}).

Because relatively little moisture goes into changing the atmospheric vapor mixing ratio, the net precipitation rate must approximately equal the moisture convergence into an atmospheric column plus surface evaporation

$$P = -\int_0^{z_m}\nabla\bullet\left(\rho q\mathbf{V}\right)dz + E \tag{11.24}$$

[4] See, for example, Stevens and Lindzen (1978).

where E is the evaporation rate (kg m^{-2} s^{-1}) and z_m is the top of the moist layer ($z_m \approx 2$ km over much of the equatorial oceans). Substituting into (11.24) from the approximate continuity equation for q,

$$\nabla \bullet (\rho q \mathbf{V}) + \partial (\rho q w) / \partial z \approx 0 \tag{11.25}$$

we obtain

$$P = (\rho w q)_{z_m} + E \tag{11.26}$$

Using (11.26) we can relate the vertically averaged heating rate to the synoptic-scale variables $w(z_m)$ and $q(z_m)$.

We still, however, need to determine distribution of the heating in the vertical. The most common approach is to use an empirically determined vertical distribution based on observations. In that case, (11.18) can be written as

$$\left(\frac{\partial}{\partial t} + \mathbf{V} \bullet \nabla\right) \ln \theta + w \frac{\partial \ln \theta}{\partial z} = \frac{L_c}{\rho c_p T} \eta(z) \left[(\rho w q)_{z_m} + E\right] \tag{11.27}$$

where $\eta(z) = 0$ for $z < z_c$ and $z > z_T$ and $\eta(z)$ for $z_c \leq z \leq z_T$ is a weighting function that must satisfy

$$\int_{z_c}^{z_T} \eta(z) dz = 1$$

Recalling that the diabatic heating must be approximately balanced by adiabatic cooling as indicated in (11.22), we see from (11.27) that $\eta(z)$ will have a vertical structure similar to that of the large-scale vertical mass flux, ρw . Observations indicate that for many tropical synoptic-scale disturbances, $\eta(z)$ reaches its maximum at about the 400-hPa level, consistent with the divergence pattern shown in Fig. 11.5.

The above formulation is designed to model average tropical conditions. In reality the vertical distribution of diabatic heating is determined by the local distribution of cloud heights. Thus, the cloud height distribution is apparently a key parameter in cumulus parameterization. A cumulus parameterization scheme in which this distribution is determined in terms of the large-scale variables was developed by Arakawa and Schubert (1974). A number of other schemes have been suggested in the past decade. Discussion of these is beyond the scope of this text.

11.4 EQUATORIAL WAVE THEORY

Equatorial waves are an important class of eastward and westward propagating disturbances in the atmosphere and in the ocean that are trapped about the equator (i.e., they decay away from the equatorial region). Diabatic heating by

organized tropical convection can excite atmospheric equatorial waves, whereas wind stresses can excite oceanic equatorial waves. Atmospheric equatorial wave propagation can cause the effects of convective storms to be communicated over large longitudinal distances, thus producing remote responses to localized heat sources. Furthermore, by influencing the pattern of low-level moisture convergence, atmospheric equatorial waves can partly control the spatial and temporal distribution of convective heating. Oceanic equatorial wave propagation, however, can cause local wind stress anomalies to remotely influence the thermocline depth and the SST, as was discussed in Section 11.1.6.

11.4.1 Equatorial Rossby and Rossby–Gravity Modes

A complete development of equatorial wave theory would be rather complicated. In order to introduce equatorial waves in the simplest possible context we here use a shallow water model, analogous to that introduced in Section 7.3.2, and concentrate on the *horizontal* structure. Vertical propagation in a stratified atmosphere is discussed in Chapter 12. For simplicity, we consider the linearized momentum and continuity equations for a fluid system of mean depth h_e in a motionless basic state. Because we are interested only in the tropics, we utilize Cartesian geometry on an *equatorial β-plane*. In this approximation, terms proportional to $\cos\phi$ are replaced by unity, and terms proportional to $\sin\phi$ are replaced by y/a, where y is the distance from the equator and a is the radius of the earth. The Coriolis parameter in this approximation is given by

$$f \approx \beta y \tag{11.28}$$

where $\beta \equiv 2\Omega/a$, and Ω is angular velocity of the earth. The resulting linearized shallow water equations for perturbations on a motionless basic state of mean depth h_e may be written as

$$\partial u'/\partial t - \beta y v' = -\partial \Phi' \partial x \tag{11.29}$$

$$\partial v'/\partial t + \beta y u' = -\partial \Phi' \partial y \tag{11.30}$$

$$\partial \Phi'/\partial t + g h_e \left(\partial u'/\partial x + \partial v'/\partial y\right) = 0 \tag{11.31}$$

where $\Phi' = gh'$ is the geopotential disturbance, and primed variables designate perturbation fields.

The x and t dependence may be separated by specifying solutions in the form of zonally propagating waves:

$$\begin{pmatrix} u' \\ v' \\ \Phi' \end{pmatrix} = \begin{bmatrix} \hat{u}\,(y) \\ \hat{v}\,(y) \\ \hat{\Phi}\,(y) \end{bmatrix} \exp\left[i\,(kx - \nu t)\right] \tag{11.32}$$

Substitution of (11.32) into (11.29)–(11.31) then yields a set of ordinary differential equations in y for the meridional structure functions $\hat{u}$, $\hat{v}$, $\hat{\Phi}$:

$$-i\nu\hat{u} - \beta y\hat{v} = -ik\hat{\Phi} \tag{11.33}$$

$$-i\nu\hat{v} + \beta y\hat{u} = -\partial\hat{\Phi}/\partial y \tag{11.34}$$

$$-i\nu\hat{\Phi} + gh_e\left(ik\hat{u} + \partial\hat{v}/\partial y\right) = 0 \tag{11.35}$$

If (11.33) is solved for $\hat{u}$ and the result substituted into (11.34) and (11.35), we obtain

$$\left(\beta^2 y^2 - \nu^2\right)\hat{v} = ik\beta y\hat{\Phi} + i\nu\partial\hat{\Phi}/\partial y \tag{11.36}$$

$$\left(\nu^2 - gh_e k^2\right)\hat{\Phi} + i\nu gh_e\left(\frac{\partial\hat{v}}{\partial y} - \frac{k}{\nu}\beta y\hat{v}\right) = 0 \tag{11.37}$$

Finally, (11.37) can be substituted into (11.36) to eliminate $\hat{\Phi}$, yielding a second-order differential equation in the single unknown, $\hat{v}$:

$$\frac{\partial^2\hat{v}}{\partial y^2} + \left[\left(\frac{\nu^2}{gh_e} - k^2 - \frac{k}{\nu}\beta\right) - \frac{\beta^2 y^2}{gh_e}\right]\hat{v} = 0 \tag{11.38}$$

Because (11.38) is homogeneous, we expect that nontrivial solutions satisfying the requirement of decay at large $|y|$ will exist only for certain values of ν, corresponding to frequencies of the normal mode disturbances.

Before discussing this equation in detail, it is worth considering the asymptotic limits that occur when either $h_e \to \infty$ or $\beta = 0$. In the former case, which is equivalent to assuming that the motion is nondivergent, (11.38) reduces to

$$\frac{\partial^2\hat{v}}{\partial y^2} + \left[-k^2 - \frac{k}{\nu}\beta\right]\hat{v} = 0$$

Solutions exist of the form $\hat{v} \sim \exp(ily)$, provided that ν satisfies the Rossby wave dispersion relationship, $\nu = -\beta k/(k^2 + l^2)$. This illustrates that for nondivergent barotropic flow, equatorial dynamics is in no way special. The rotation of the earth enters only in the form of the β effect; it is not dependent on f. However, if $\beta = 0$, all influence of rotation is eliminated and (11.38) reduces to the shallow water gravity wave model, which has nontrivial solutions for

$$\nu = \pm\left[gh_e\left(k^2 + l^2\right)\right]^{1/2}$$

Returning to (11.38), we seek solutions for the meridional distribution of $\hat{v}$, subject to the boundary condition that the disturbance fields vanish for $|y| \to \infty$.

This boundary condition is necessary because the approximation $f \approx \beta y$ is not valid for latitudes much beyond $\pm 30°$, so that solutions must be trapped equatorially if they are to be good approximations to the exact solutions on the sphere. Equation (11.38) differs from the classic equation for a harmonic oscillator in y because the coefficient in square brackets is not a constant, but is a function of y. For sufficiently small y this coefficient is positive and solutions oscillate in y, whereas for large y, solutions either grow or decay in y. Only the decaying solutions, however, can satisfy the boundary conditions.

It turns out[5] that solutions to (11.38), which satisfy the condition of decay far from the equator, exist only when the constant part of the coefficient in square brackets satisfies the relationship

$$\frac{\sqrt{gh_e}}{\beta}\left(-\frac{k}{\nu}\beta - k^2 + \frac{\nu^2}{gh_e}\right) = 2n + 1; \quad n = 0,1,2,\ldots \tag{11.39}$$

which is a cubic dispersion equation determining the frequencies of permitted equatorially trapped free oscillations for zonal wave number k and meridional mode number n. These solutions can be expressed most conveniently if y is replaced by the nondimensional meridional coordinate

$$\xi \equiv \left(\beta \Big/ \sqrt{gh_e}\right)^{1/2} y$$

Then the solution has the form

$$\hat{v}(\xi) = v_0 \mathrm{H_n}(\xi) \exp\left(-\xi^2/2\right) \tag{11.40}$$

where v_o is a constant with velocity units, and $\mathrm{H}_n(\xi)$ designates the nth *Hermite polynomial*. The first few of these polynomials have the values

$$\mathrm{H}_0 = 1, \quad \mathrm{H}_1(\xi) = 2\xi, \quad \mathrm{H}_2(\xi) = 4\xi^2 - 2$$

Thus, the index n corresponds to the number of nodes in the meridional velocity profile in the domain $|y| < \infty$.

In general the three solutions of (11.39) can be interpreted as eastward and westward moving equatorially trapped gravity waves, and westward moving equatorial Rossby waves. The case $n = 0$ (for which the meridional velocity perturbation has a Gaussian distribution centered at the equator) must be treated separately. In this case the dispersion relationship (11.39) factors as

$$\left(\frac{\nu}{\sqrt{gh_e}} - \frac{\beta}{\nu} - k\right)\left(\frac{\nu}{\sqrt{gh_e}} + k\right) = 0 \tag{11.41}$$

[5] See Matsuno (1966).

The root $\nu/k = -\sqrt{gh_e}$, corresponding to a westward propagating gravity wave, is not permitted, as the second term in parentheses in (11.41) was implicitly assumed not to vanish when (11.36) and (11.37) were combined to eliminate $\hat{\Phi}$. The roots given by the first term in parentheses in (11.41) are

$$\nu = k\sqrt{gh_e}\left[\frac{1}{2} \pm \frac{1}{2}\left(1 + \frac{4\beta}{k^2\sqrt{gh_e}}\right)^{1/2}\right] \tag{11.42}$$

The positive root corresponds to an eastward propagating equatorial inertia–gravity wave, whereas the negative root corresponds to a westward propagating wave, which resembles an inertia–gravity wave for long zonal scale ($k \to 0$) and resembles a Rossby wave for zonal scales characteristic of synoptic-scale disturbances. This mode is generally referred to as a *Rossby–gravity* wave. The horizontal structure of the westward propagating $n = 0$ solution is shown in Fig. 11.13, whereas the relationship between frequency and zonal wave number for this and several other equatorial wave modes is shown in Fig. 11.14.

11.4.2 Equatorial Kelvin Waves

In addition to the modes discussed in the previous section, there is another equatorial wave that is of great practical importance. For this mode, which is called the equatorial *Kelvin wave*, the meridional velocity perturbation vanishes and (11.33)–(11.35) are reduced to the simpler set

$$-i\nu\hat{u} = -ik\hat{\Phi} \tag{11.43}$$

$$\beta y\hat{u} = -\partial\hat{\Phi}/\partial y \tag{11.44}$$

$$-i\nu\hat{\Phi} + gh_e\left(ik\hat{u}\right) = 0 \tag{11.45}$$

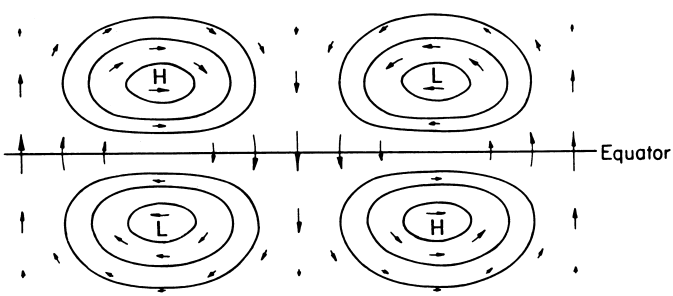


Fig. 11.13 Plan view of horizontal velocity and height perturbations associated with an equatorial Rossby–gravity wave. (Adapted from Matsuno, 1966.)

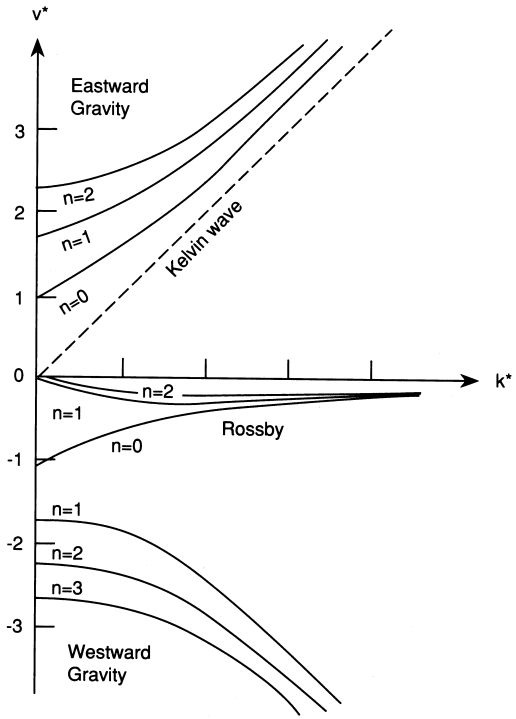


Fig. 11.14 Dispersion diagram for free equatorial waves. Frequency and zonal wave numbers have been nondimensionalized by defining

$$\nu^* \equiv \nu / \left(\beta\sqrt{gh_e}\right)^{1/2} \qquad k^* \equiv k\left(\sqrt{gh_e}/\beta\right)^{1/2}$$

Curves show dependence of frequency on zonal wave number for eastward and westward gravity modes and for Rossby and Kelvin modes. (k^* axis tic marks at unit interval with 0 on left.)

Combining (11.43) and (11.45) to eliminate $\hat{\Phi}$, we see that the Kelvin wave dispersion equation is identical to that for ordinary shallow water gravity waves:

$$c^2 \equiv (\nu/k)^2 = gh_e \tag{11.46}$$

According to (11.46), the phase speed c can be either positive or negative. However, if (11.43) and (11.44) are combined to eliminate $\hat{\Phi}$, we obtain a first-order equation for determining the meridional structure:

$$\beta y\hat{u} = -c\partial\hat{u}/\partial y \tag{11.47}$$

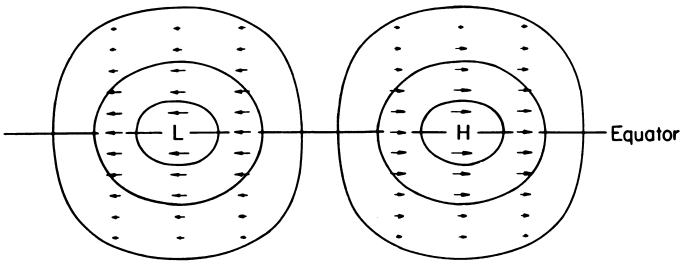


Fig. 11.15 Plan view of horizontal velocity and height perturbations associated with an equatorial Kelvin wave. (Adapted from Matsuno, 1966.)

which may be integrated immediately to yield

$$\hat{u} = u_o \exp\left(-\beta y^2/2c\right) \tag{11.48}$$

where u_0 is the amplitude of the perturbation zonal velocity at the equator. Equation (11.48) shows that if solutions decaying away from the equator are to exist, the phase speed must be positive ($c > 0$). Thus, Kelvin waves are eastward propagating and have zonal velocity and geopotential perturbations that vary in latitude as Gaussian functions centered on the equator. The e-folding decay width is given by

$$Y_K = |2c/\beta|^{1/2}$$

which for a phase speed $c = 30\,\mathrm{m\,s^{-1}}$ gives $Y_K \approx 1600$ km.

The perturbation wind and geopotential structure for the Kelvin wave are shown in plan view in Fig. 11.15. In the zonal direction the force balance is exactly that of an eastward propagating shallow water gravity wave. A vertical section along the equator would thus be the same as that shown in Fig. 7.9. The meridional force balance for the Kelvin mode is an exact geostrophic balance between the zonal velocity and the meridional pressure gradient. It is the change in sign of the Coriolis parameter at the equator that permits this special type of equatorial mode to exist.

11.5 STEADY FORCED EQUATORIAL MOTIONS

Not all zonally asymmetric circulations in the tropics can be explained on the basis of inviscid equatorial wave theory. For quasi-steady circulations the zonal pressure gradient force must be balanced by turbulent drag rather than by inertia. The Walker circulation may be regarded as a quasi-steady equatorially trapped circulation that is generated by diabatic heating. The simplest models of such circulations specify the diabatic heating and use the equations of equatorial wave

theory to compute the atmospheric response. These models, however, neglect the fact that the distribution of diabatic heating is highly dependent on the mean wind and equivalent potential temperature distributions in the boundary layer. These in turn depend on the surface pressure and the moisture distribution, which are themselves dependent on the motion field. Thus, in a consistent model the diabatic heating cannot be regarded as an externally specified quantity, but must be obtained as part of the solution by, for example, using the cumulus parameterization scheme illustrated in Section 11.3.

As indicated in (11.27) this scheme requires information on the vertical distribution of convective heating in order to solve for the temperature perturbation throughout the troposphere. There is, however, evidence that the essential features of stationary equatorial circulations can be partly explained on the basis of a model that involves only the boundary layer. This is perhaps not surprising since the maintenance of a convective system depends on evaporation and moisture convergence in the boundary layer. Over the tropical oceans the boundary layer can be approximated as a mixed layer of about 2 km depth, which is capped by an inversion, across which there is a density discontinuity $\delta\rho$ (see Fig. 5.2). The virtual temperature in the mixed layer is strongly correlated with the sea surface temperature. If we assume that the pressure field is uniform at the top of the mixed layer, the surface pressure will be determined by hydrostatic mass adjustment within the mixed layer. (The system is analogous to the two-layer model discussed in Section 7.3.2.) The resulting pressure perturbation in the layer depends on the density discontinuity at the top of the layer and the deviation, h, of the layer depth from its mean depth H_b. In addition, there will be a contribution from the perturbation virtual potential temperature, θ_v, within the mixed layer. Thus, the perturbation geopotential in the mixed layer can be expressed as

$$\Phi = g\,(\delta\rho/\rho_0)\,h - \Gamma\theta_v \tag{11.49}$$

where $\Gamma \equiv (gH_b/\theta_0)$ is a constant and ρ_0 and θ_0 are constant mixed layer reference values of density and potential temperature, respectively.

According to (11.49), positive sea surface temperature anomalies and negative boundary layer height anomalies will produce low surface pressures, and vice versa. If the boundary layer depth does not vary too much, the surface pressure gradient will thus tend to be proportional to the sea surface temperature gradient. The dynamics of steady circulations in such a mixed layer can be approximated by a set of linear equations analogous to the equatorial wave equations (11.29)–(11.31), but with the time derivative terms replaced by linear damping terms.

Thus, in the momentum equations the surface eddy stress is taken to be proportional to the mean velocity in the mixed layer. In the continuity equation the perturbation in the mixed layer height is proportional to the mass convergence in the layer, with a coefficient that is smaller in the presence of convection than in its absence, due to ventilation of the boundary layer by convection. The x and y

components of the vertically averaged momentum equation for steady motions in the mixed layer can then be written as

$$\alpha u - \beta y v + \partial \Phi / \partial x = 0 \tag{11.50}$$

$$\alpha v + \beta y u + \partial \Phi / \partial y = 0 \tag{11.51}$$

whereas the continuity equation can be expressed as

$$\alpha h + H_b (1 - \varepsilon) (\partial u / \partial x + \partial v / \partial y) = 0 \tag{11.52}$$

where ε is a coefficient that is zero in the absence of convection and of order 3/4 when convection occurs. Substituting from (11.49) into (11.52) yields

$$\alpha \Phi + c_b^2 (1 - \varepsilon) (\partial u / \partial x + \partial v / \partial y) = -\alpha \Gamma \theta_v \tag{11.53}$$

where $c_b^2 \equiv g (\delta \rho / \rho_0) H_b$ is the square of the phase speed for gravity waves propagating along the inversion at the top of the mixed layer. [See Battisti et al., (1999) for a more complete discussion.]

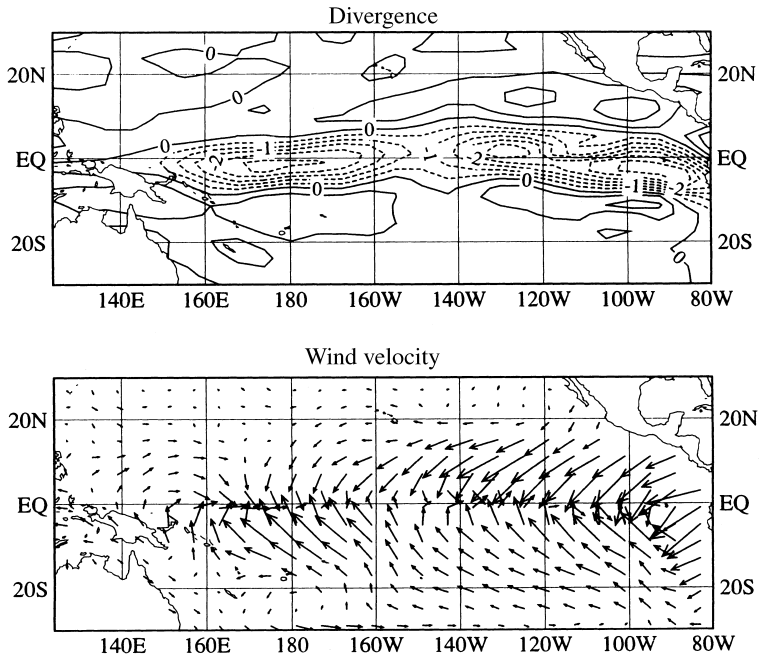


Fig. 11.16 Steady surface circulation forced by sea surface temperature anomalies characteristic of El Niño in the equatorial Pacific (courtesy of D. Battisti).

Equations (11.50), (11.51), and (11.53) are a closed set for prediction of the boundary layer variables u, v, Φ for a specified boundary layer perturbation virtual temperature θ_v. However, because the parameter ε depends on the presence or absence of convection, the system can only be solved by iteration through using (11.24) to test for the presence of convection. This model can be used to compute the steady surface circulation. A sample calculation of the anomalous circulation for temperature anomalies corresponding to a typical ENSO event is shown in Fig. 11.16. Note that the region of convergence is narrower than the region of warm SST anomalies due to the convective feedback in the boundary layer model.

PROBLEMS

11.1. Suppose that the relative vorticity at the top of an Ekman layer at 15°N is $\zeta = 2 \times 10^{-5}\,\mathrm{s}^{-1}$. Let the eddy viscosity coefficient be $K_m = 10\,\mathrm{m}^2\,\mathrm{s}^{-1}$ and the water vapor mixing ratio at the top of the Ekman layer be $12\,\mathrm{g\,kg}^{-1}$. Use the method of Section 11.3 to estimate the precipitation rate due to moisture convergence in the Ekman layer.

11.2. As mentioned in Section 11.1.3, barotropic instability is a possible energy source for some equatorial disturbances. Consider the following profile for an easterly jet near the equator:

$$\bar{u}\,(y) = -u_0 \sin^2 [l\,(y - y_0)]$$

where u_0, y_0, and l are constants and y is the distance from the equator. Determine the necessary conditions for this profile to be barotropically unstable.

11.3. Show that the nonlinear terms in the balance equation (11.15)

$$G\,(x, y) \equiv -\nabla^2 \left(\frac{1}{2}\nabla\psi \bullet \nabla\psi\right) + \nabla \bullet \left(\nabla\psi\nabla^2\psi\right)$$

may be written in Cartesian coordinates as

$$G\,(x, y) = 2\left[\left(\partial^2\psi/\partial x^2\right)\left(\partial^2\psi/\partial y^2\right) - \left(\partial^2\psi/\partial x\partial y\right)^2\right]$$

11.4. With the aid of the results of Problem 11.3, show that if f is assumed to be constant the balance equation (11.15) is equivalent to the gradient wind equation (3.15) for a circularly symmetric regular low with geopotential perturbation given by

$$\Phi = \Phi_0\left(x^2 + y^2\right)/L^2$$

where Φ_0 is a constant geopotential and L a constant length scale. *Hint:* Assume that $\psi(x, y)$ has the same functional dependence on (x, y) as does Φ.

11.5. Starting from the perturbation equations (11.29)–(11.31), show that the sum of kinetic plus available potential energy is conserved for equatorial waves. Hence, show that for the Kelvin wave there is an equipartition of energy between kinetic and available potential energy.

11.6. Solve for the meridional dependence of the zonal wind and geopotential perturbations for a Rossby–gravity mode in terms of the meridional velocity distribution (11.40).

11.7. Use the linearized model (11.50) and (11.51) to compute the meridional distribution of divergence in the mixed layer for a situation in which the geostrophic wind is given by $u_g = u_o \exp\left(-\beta y^2/2c\right)$, $v_g = 0$ where u_0 and c are constants.

11.8. Show that the frequency of the $n = 1$ equatorial Rossby mode is given approximately by $\nu = -k\beta\left(k^2 + 3\beta/\sqrt{gh_e}\right)^{-1}$ and use this result to solve for the $\hat{u}(y)$ and $\hat{\Phi}(y)$ fields in terms of $\hat{v}(y)$. *Hint:* Use the fact that the Rossby wave phase speed is much less than $\sqrt{gh_e}$.

MATLAB EXERCISES

M11.1. The MATLAB script **profile_2.m** and the function **Tmoist.m** use December–March seasonal mean pressure and temperature data for Samoa (14°S, 171°W) in the file **tropical_temp.dat**, together with an assumed constant relative humidity of 80%, to calculate and graph vertical profiles of temperature, dewpoint, and the temperature corresponding to pseudoadiabatic ascent from the lifting condensation level for a parcel lifted from the lowest level of the sounding. Modify this script to plot profiles of the potential temperature, the equivalent potential temperature, and the saturation equivalent potential temperature. (*Hint:* See Problem M9.2.)

M11.2. For the thermodynamic sounding of Exercise M11.1, compute the CAPE and the profile of vertical velocity for a parcel lifted from the lowest layer in the absence of entrainment. What is the maximum vertical velocity and what is the distance that the parcel overshoots its level of neutral buoyancy? What can you conclude about convection associated with this mean tropical sounding?

M11.3. Consider a simple nondivergent flow given by the streamfunction $\psi = A \sin kx \sin ly$. The MATLAB script **nonlinear_balance.m** solves the

nonlinear balance equation (11.15) for the corresponding geopotential Φ, assuming that the Coriolis parameter is a constant corresponding to its value at 30° N. Run this script for several values of the amplitude, A, in the range 0.4×10^7 to $4.0 \times 10^7\ \mathrm{m^2\,s^{-1}}$. Note how the geopotential field depends on A. For the case $A = 4.0 \times 10^7 \mathrm{m^2\,s^{-1}}$ use the **gradient** function in MATLAB to compute the geostrophic wind, and hence find the ageostrophic wind. Plot these as quiver plots and explain why the ageostrophic wind has the structure seen in your plot.

M11.4. A simple model for the horizontal flow in the mixed layer on an equatorial β plane forced by an equatorial wave mode is given by the following equations:

$$\frac{\partial u}{\partial t} = -\alpha u + \beta y v - \frac{\partial \Phi}{\partial x}$$

$$\frac{\partial v}{\partial t} = -\alpha v - \beta y u - \frac{\partial \Phi}{\partial y}$$

where $\Phi(x, y, t)$ is given by (11.32) for the appropriate mode. The MATLAB script **equatorial_mixed_layer.m** solves these equations numerically for the horizontal velocity components and divergence field in the mixed layer forced by a specified Rossby–gravity wave ($n = 0$ mode) geopotential perturbation corresponding to a westward propagating disturbance of zonal wavelength 4000 km and $\sqrt{gh_e} = 18\ \mathrm{m\,s^{-1}}$. Run the code for a long enough time so that the solution becomes periodic in time. Rerun the model for zonal wavelength 10,000 km and $\sqrt{gh_e} = 36\ \mathrm{m\,s^{-1}}$. In each case compare the frequency of the oscillation with the Coriolis parameter at the latitude of maximum convergence forced by the wave. What can you conclude from these results?

M11.5. Modify the MATLAB script of M11.4 to compute the mixed layer velocity and divergence pattern for the $n = 1$ Rossby mode for zonal wavelength 4000 km and $\sqrt{gh_e} = 18\ \mathrm{m\,s^{-1}}$ using the formulas for frequency and geopotential derived in Problem 11.8.

M11.6. The MATLAB script **forced_equatorial_mode2.m** shows the time development of the velocity and height perturbations produced by a transient localized mass source for shallow water waves on an equatorial β-plane. The model is based on equations (11.50), (11.51), and (11.53), but with time tendency terms included ($\varepsilon = 0$ is set here, but you can experiment with different values if you wish). The source is located at the equator at $x = 0$. It rises smoothly in amplitude for the first 2.5 days and then decreases to 0 at day 5. Modify the script to contour the divergence and the vorticity. Run the script for a 10-day period and interpret the results

in terms of equatorial wave theory by noting the structure and speeds of energy propagation for the disturbances to the east and west of the source.

Suggested References

Philander, *El Niño, La Niña, and the Southern Oscillation* is a well-written text that covers the dynamics of both atmospheric and oceanographic aspects of ENSO at an advanced level.

Trenberth (1991) is an excellent review article on ENSO.

Wallace (1971) is a detailed review of the structure of synoptic-scale tropospheric wave disturbances in the equatorial Pacific.

Webster and Fasullo (2003) present an excellent review of the structure and dynamics of tropical monsoons.

CHAPTER 12

Middle Atmosphere Dynamics

The *middle atmosphere* is generally regarded as the region extending from the tropopause (about 10–16 km altitude depending on latitude) to about 100 km. The bulk of the middle atmosphere consists of two main layers, the *stratosphere* and the *mesosphere*, which are distinguished on the basis of temperature stratification (Fig. 12.1). The stratosphere, which has large static stability associated with an overall increase of temperature with height, extends from the tropopause to the *stratopause* at about 50 km. The mesosphere, which has a lapse rate similar to that in the troposphere, extends from the stratopause to the *mesopause* at about 80–90 km.

Previous chapters of this text have focused almost exclusively on the dynamics of the troposphere. The troposphere accounts for about 85% of the total mass of the atmosphere and virtually all atmospheric water. There can be little doubt that processes occurring in the troposphere are primarily responsible for weather disturbances and climate variability. Nevertheless the middle atmosphere cannot be neglected. The troposphere and the middle atmosphere are linked through radiative and dynamical processes that must be represented in global forecast and climate models. They are also linked through the exchange of trace substances that are important in the photochemistry of the ozone layer. This chapter focuses primarily

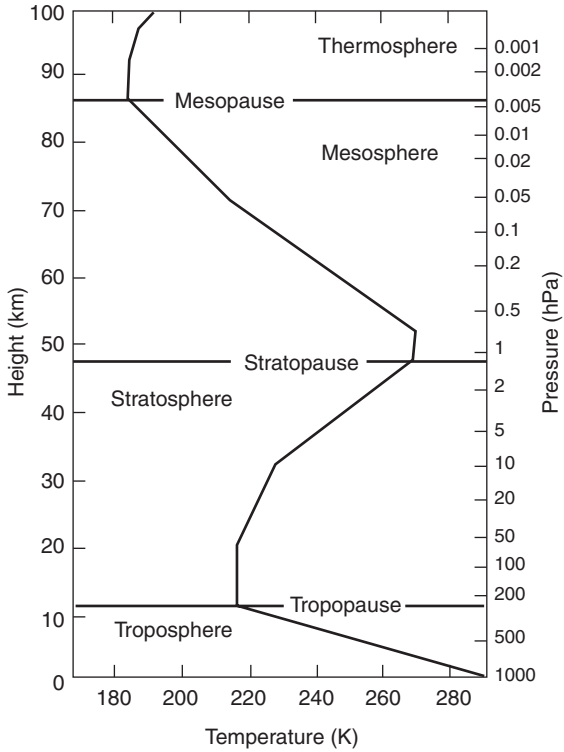


Fig. 12.1 Midlatitude mean temperature profile. Based on the *U.S. Standard Atmosphere* (1976).

on dynamical processes in the lower part of the middle atmosphere and their links to the troposphere.

12.1 STRUCTURE AND CIRCULATION OF THE MIDDLE ATMOSPHERE

Zonal mean temperature cross sections for January and July in the lower and middle atmosphere are shown in the upper panels of Figs. 12.2 and 12.3, respectively. Because very little solar radiation is absorbed in the troposphere, the thermal structure of the troposphere is maintained by an approximate balance among infrared radiative cooling, vertical transport of sensible and latent heat away from the surface by small-scale eddies, and large-scale heat transport by synoptic-scale eddies. The net result is a mean temperature structure in which the surface temperature has its maximum in the equatorial region and decreases toward both the winter and the summer poles. There is also a rapid decrease in altitude, with a lapse rate of about 6°C km^{-1}.

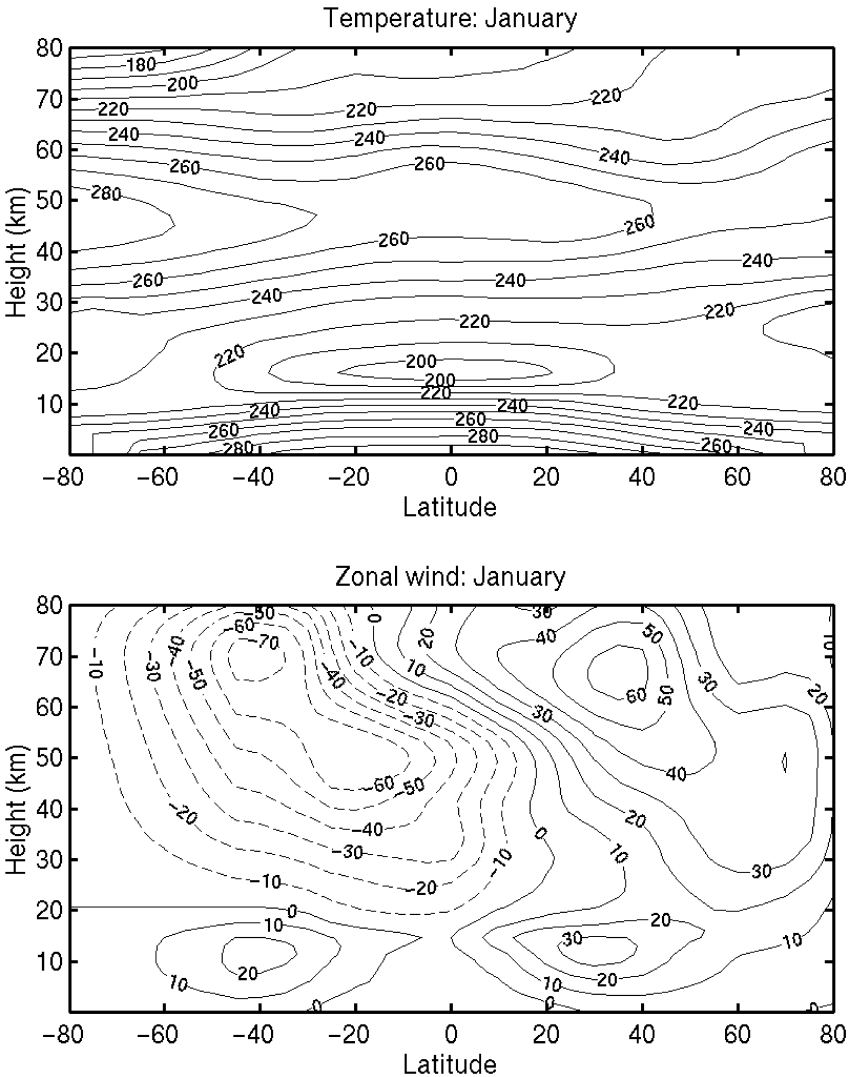


Fig. 12.2 Observed monthly and zonally averaged temperature (K) and zonal wind ($m\,s^{-1}$) for January. (Based on Fleming et al., 1990.)

In the stratosphere, however, infrared radiative cooling is in the mean balanced primarily by radiative heating due to the absorption of solar ultraviolet radiation by ozone. As a result of the solar heating in the ozone layer, the mean temperature in the stratosphere increases with height to a maximum at the stratopause near 50 km.

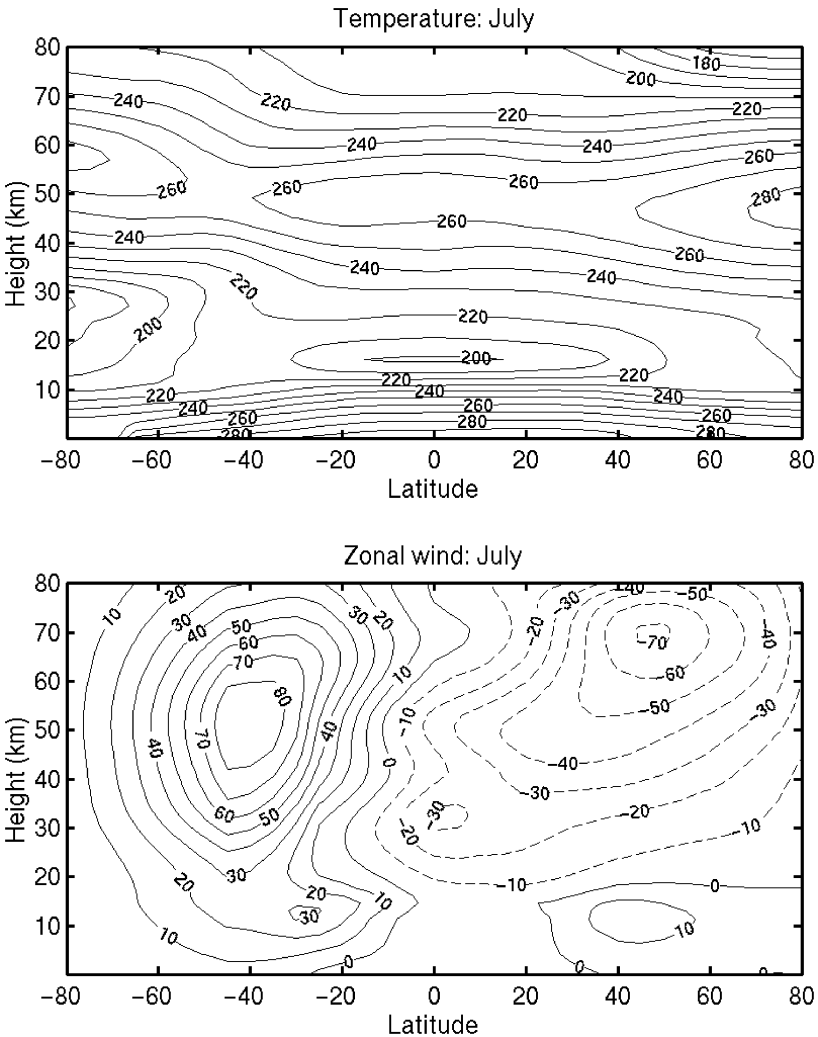


Fig. 12.3 Observed monthly and zonally averaged temperature (K) and zonal wind (m s^{-1}) for July. (Based on Fleming et al., 1990.)

Above the stratopause, temperature decreases with height due to the reduced solar heating of ozone.

The meridional temperature structure in the middle atmosphere is also quite different from that in the troposphere. In the lower stratosphere, where the temperature is strongly influenced by upper tropospheric processes, there is a temperature minimum at the equator and maxima at the summer pole and in midlatitudes of

the winter hemisphere. Above about 30 hPa the temperature decreases uniformly from summer pole to winter pole, in qualitative accord with radiative equilibrium conditions.

Mean zonal wind climatologies for the middle atmosphere are usually derived from the satellite observed temperature field. This is done using the geostrophic winds on an isobaric surface in the lower stratosphere (obtained from conventional meteorological analyses) as a lower boundary condition and integrating the thermal wind equation vertically. January and July mean zonal wind cross sections are shown in the lower panels of Figs. 12.2 and 12.3, respectively. The main features are an easterly jet in the summer hemisphere and a westerly jet in the winter hemisphere, with maxima in the wind speeds occurring near the 60-km level. Of particular significance are the high-latitude westerly jets in the winter hemispheres. These polar night jets provide wave guides for the vertical propagation of quasi-stationary planetary waves. In the Northern Hemisphere the EP flux convergence due to such waves occasionally leads to rapid deceleration of the mean zonal flow and an accompanying *sudden stratospheric warming* in the polar region, as discussed in Section 12.4.

The zonal mean flow in the equatorial middle atmosphere is strongly influenced by vertically propagating inertia–gravity waves and by equatorial wave modes, especially the Kelvin and Rossby–gravity modes. These waves interact with the mean flow to produce a long period oscillation called *quasi-biennial oscillation*. This oscillation produces large year-to-year variability in the mean zonal wind in the equatorial middle atmosphere, which is not shown in the long-term means of Fig. 12.3.

12.2 THE ZONAL-MEAN CIRCULATION OF THE MIDDLE ATMOSPHERE

As discussed in Section 10.1, the general circulation of the atmosphere considered as a whole can be regarded to a first approximation as the atmospheric response to the diabatic heating caused by absorption of solar radiation at the surface. Thus, it is reasonable to say that the atmosphere is driven by differential diabatic heating. For an open subregion of the atmosphere, such as the middle atmosphere, it is not correct, however, to assume that the circulation is driven by diabatic heating. It is, rather, necessary to consider the transfer of momentum and energy between that subregion and the rest of the atmosphere.

In the absence of eddy motions, the zonal-mean temperature of the middle atmosphere would relax to a radiatively determined state in which, except for a small lag due to thermal inertia, the temperature would correspond to an annually varying radiative equilibrium following the annual cycle in solar heating. The circulation in such a situation would consist only of a zonal-mean zonal flow in thermal wind balance with the meridional temperature gradient. Neglecting small

effects of the annual cycle, there would be no meridional or vertical circulation and no stratosphere–troposphere exchange for such a hypothetical state.

A cross section of the radiatively determined temperature during a Northern Hemisphere winter is shown in Fig. 12.4. It should be compared to the observed temperature profile for the same season that was shown in Fig. 12.2. Although the rather uniform increase of temperature from winter pole to summer pole in the region of 30- to 60-km altitude is qualitatively consistent with the radiatively determined distribution, the actual temperature difference between the two poles is much smaller than in the radiatively determined state. Above 60 km even the sign of the gradient is opposite to that in the radiative solution; the observed temperatures near the summer polar mesopause are much colder than those near the winter polar mesopause.

Departures from this radiatively determined state must be maintained by eddy transports. Thus, rather than *causing* the mean circulation, the radiative heating

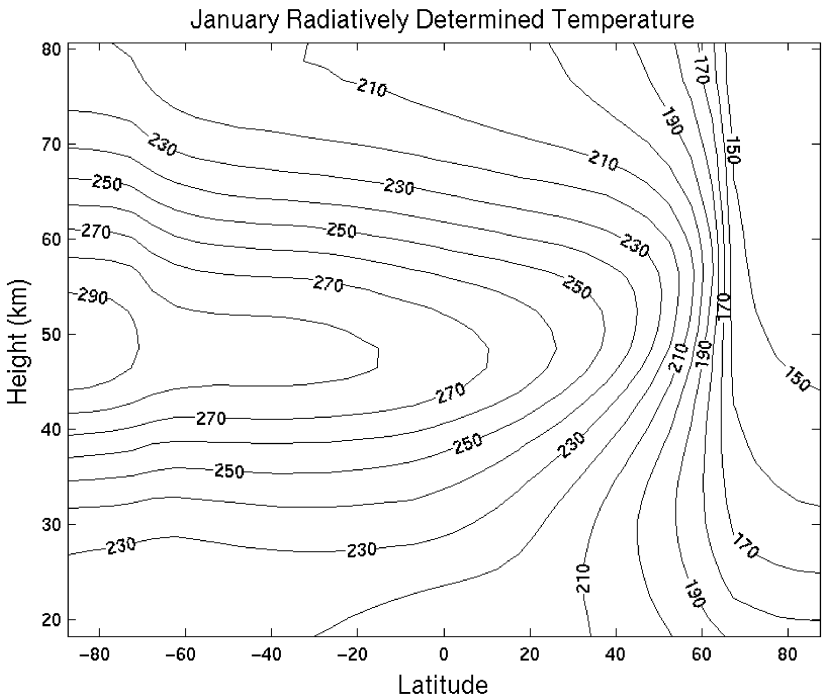


Fig. 12.4 Radiatively determined middle atmosphere temperature distribution (K) for Northern winter soltice from a radiative model that is time marched through an annual cycle. Realistic tropospheric temperatures and cloudiness are used to determine the upward radiative flux at the tropopause. (Based on Shine, 1987.)

and cooling patterns observed in the middle atmosphere are a *result* of the eddies driving the flow away from a state of radiative balance. This eddy-driven circulation has meridional and vertical wind components that induce substantial local departures from radiative equilibrium, especially in the winter stratosphere and in the mesosphere in both winter and summer.

12.2.1 Lagrangian Motion of Air Parcels

Viewed in the conventional Eulerian mean framework (Section 10.2.1), the time-averaged zonal-mean temperature distribution in the middle atmosphere is determined by the net balance among net radiative heating or cooling, eddy heat transport, and adiabatic heating or cooling by the mean meridional circulation $(\overline{v}, \overline{w})$. That this framework does not provide a useful model for transport in the meridional plane can be seen easily by considering steady-state adiabatic motion in the presence of large-scale waves. According to (10.12), in such a situation waves that have a positive poleward heat flux must drive a nonzero $\overline{w}$, with upwelling at high latitudes and downwelling at low latitudes corresponding to the regions of heat flux convergence and divergence, respectively. However, if the flow is adiabatic, there can be no motion across the isentropes, and hence in steady state there can be no net vertical transport even though $\overline{w}$ is nonzero. Thus, $\overline{w}$ clearly does not provide an approximation to the vertical motion of air parcels (i.e., the vertical transport) in such a situation.

But how can vertical transport vanish when the Eulerian mean vertical motion is finite? The resolution of this "nontransport" paradox can be illustrated by considering the kinematics of adiabatic flow in the presence of a stationary large-scale wave superposed on a background westerly flow as shown in Fig. 12.5a. Parcels moving along the streamlines labeled S_1, S_2, S_3 will oscillate about their mean latitudes, moving equatorward and upward, then poleward and downward following the up and down displacements of the isentropic surfaces associated with the wave. However, over a full wave period there can be no net vertical motion following the parcels, as the parcels must remain on the same isentropic surface. Hence there can be no vertical transport. If, however, an average vertical velocity is computed by taking the mean along a constant latitude circle, then as clearly shown in Fig. 12.5a, on the latitude circle ϕ_3 poleward of the maximum in wave amplitude the Eulerian mean is upward $[\overline{w}(\phi_3) > 0]$, as regions of upward motion dominate over regions of downward motion at that latitude. Conversely, on the latitude circle ϕ_1 equatorward of the maximum in wave amplitude the Eulerian mean is downward $[\overline{w}(\phi_1) < 0]$. The conventional Eulerian mean circulation thus suggests misleadingly that a trace constituent would be transported upward poleward of the latitude of maximum wave amplitude and downward equatorward of that latitude. In reality, air parcels in this idealized example of adiabatic wave motion are not undergoing net vertical displacements, but are simply oscillating back and forth

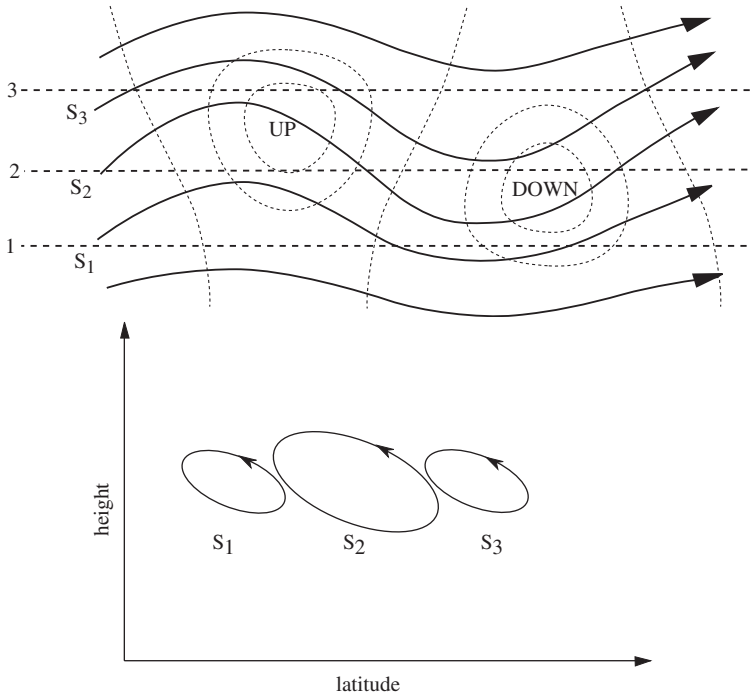


Fig. 12.5 Parcel motions for an adiabatic planetary wave in a westerly zonal flow. (a) Solid lines labeled S_1, S_2, S_3, are parcel trajectories, heavy dashed lines are latitude circles, and light dashed lines are contours of vertical velocity field. (b) Projection of parcel oscillations on the meridional plane.

along the same isentropic surface with trajectories whose meridional projections are closed ellipses as shown in Fig. 12.5b. In such a situation there is no net vertical parcel motion, and thus no net vertical transport.

Only if there is net diabatic heating or cooling can there be mean transport across the isentropes. In the winter stratosphere, where stationary planetary wave amplitudes tend to be large and there is diabatic heating at low latitudes and diabatic cooling at high latitudes, the meridional projection of actual parcel motions is a combination of the elliptical orbits of Fig. 12.5b and mean vertical drift as indicated in Fig. 12.6, which clearly shows that a vertically stratified long-lived trace constituent is transported upward at low latitudes and downward at high latitudes due to vertical transport of fluid parcels associated with diabatic heating and cooling. Thus, in order to represent net transport effects efficiently, a zonal averaging process should yield an average vertical circulation that mimics the net cross isentropic parcel motion.

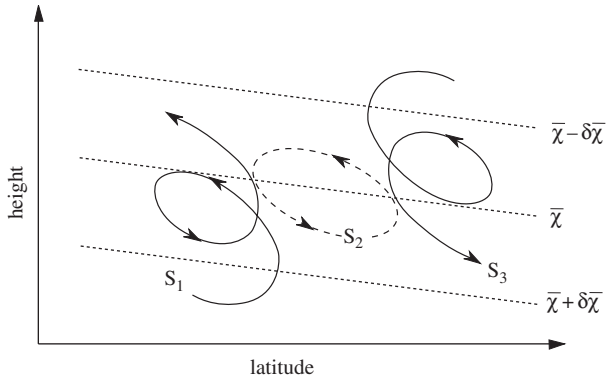


Fig. 12.6 Projections of parcel motions on the meridional plane for planetary waves in westerlies with diabatic heating at low latitudes and diabatic cooling at high latitudes. Thin dashed lines show tilting of iso-surfaces of tracer mixing ratio due to transport by the diabatic circulation.

12.2.2 The Transformed Eulerian Mean

In many circumstances the transformed Eulerian mean (TEM) equations introduced in Section 10.2.2 provide a useful model for the study of global-scale middle atmospheric transport. Recall that in this formulation the zonal-mean momentum, mass continuity, thermodynamic energy, and thermal wind equations have the form[1]

$$\partial\overline{u}/\partial t - f_0\overline{v}^* = \rho_0^{-1}\,\nabla\bullet\mathbf{F} + \overline{X} \equiv \overline{G} \tag{12.1}$$

$$\partial\overline{T}/\partial t + N^2 H R^{-1}\overline{w}^* = -\alpha_r\left[\overline{T} - \overline{T}_r\,(y,z,t)\right] \tag{12.2}$$

$$\partial\overline{v}^*/\partial y + \rho_0^{-1}\partial\left(\rho_0\overline{w}^*\right)/\partial z = 0 \tag{12.3}$$

$$f_0\partial\overline{u}/\partial z + RH^{-1}\partial\overline{T}/\partial y = 0 \tag{12.4}$$

Here the residual circulation $(\overline{v}^*, \overline{w}^*)$ is as defined in (10.16a,b), $\mathbf{F}$ designates the EP flux due to large-scale eddies, $\overline{X}$ is the zonal force due to small-scale eddies (e. g., gravity wave drag), and $\overline{G}$ designates the total *zonal force*. In (12.2) the diabatic heating is approximated in terms of a Newtonian relaxation proportional to the departure of the zonal mean temperature $\overline{T}(y,z,t)$ from its radiative equilibrium value $\overline{T_r}(y,z,t)$ where α_r is the Newtonian cooling rate.

To understand how eddies can lead to departures of the zonal-mean temperature distribution in the middle atmosphere from its radiatively determined state, we use the TEM system of equations to consider an idealized model of extratropical forcing. The dependence of the circulation on forcing frequency can be examined in an idealized model in which $\overline{T}_r$ is taken to be a function of height alone and the

[1] As in Chapter 10 we express the log-pressure coordinate simply as z rather than z^*.

forcing has a simple harmonic dependence in time with frequency σ. Then, $\overline{w}^*$ and $\overline{G}$ are of the form:

$$\begin{bmatrix} \overline{w}^* \\ \overline{G} \end{bmatrix} = Re \begin{bmatrix} \hat{w}(\phi, z) \\ \hat{G}(\phi, z) \end{bmatrix} e^{i\sigma t} \tag{12.5}$$

where $\hat{w}$ and $\hat{G}$ are complex latitude and height dependent amplitudes. Combining (12.1) and (12.3) to eliminate $\overline{v}^*$ and combining the resulting equation with (12.2) and (12.4) to eliminate $\overline{T}$ yields a partial differential equation for $\hat{w}$:

$$\frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial\left(\rho_0\hat{w}\right)}{\partial z}\right) + \left(\frac{i\sigma}{i\sigma + \alpha_r}\right)\left(\frac{N^2}{f_0^2}\right)\frac{\partial^2\hat{w}}{\partial y^2} = \frac{1}{f_0}\frac{\partial}{\partial y}\left(\frac{\partial\hat{G}}{\partial z}\right) \tag{12.6}$$

The structure of $\hat{w}$ in the meridional plane depends on the latitudinal and vertical distribution of the forcing $\hat{G}$. The differential equation (12.6) is elliptic in form. Such a form implies that a forcing applied locally in any region will produce a nonlocal response in the form of a large-scale circulation that acts to maintain the zonal flow in thermal wind balance with the meridional temperature gradient.

As suggested by the coefficients of (12.6), the character of the nonlocal response depends on the ratio of the frequency of the forcing to the Newtonian relaxation rate. We consider three cases:

(i) High-frequency variations $\sigma \gg \alpha_r$: The motions are approximately adiabatic (as is true in sudden stratospheric warmings); that is

$$\frac{i\sigma}{i\sigma + \alpha_r} \to 1$$

Away from the forcing region, the right-hand side of (12.6) vanishes so that the two terms on the left must balance each other. For this balance to hold, a simple scaling argument indicates that in midlatitudes the vertical scale δz of the response must be related to the horizontal scale as $\delta z \sim (f_0/N)\,\delta y \sim 10^{-2}\delta y$.

(ii) Low-frequency variations, such as the annual cycle: The Newtonian relaxation coefficient has a magnitude of order $\alpha_r \approx 1/(20\text{ days})$ in the lower stratosphere, increasing to $\alpha_r \approx 1/(5\text{ days})$ near the stratopause. Thus $\alpha_r > \sigma$ and the coefficient multiplying the second term on the left in (12.6) is reduced from its value in case (i) so that the vertical penetration scale is increased.

(iii) Steady-state regime $\sigma/\alpha_r \to 0$, In this limit $\partial\overline{u}//\partial t = 0$, and rather than using (12.6), it is simpler to note that (12.1) in this case reduces to a simple balance between the Coriolis force and the zonal drag force:

$$-f_0\overline{v}^* = \overline{G} \tag{12.7}$$

Combining this equation with the continuity equation (12.3) and requiring that

$$\rho_0 \overline{w}^* \to 0 \text{ as } z \to \infty$$

it follows that

$$\rho_o \overline{w}^* = -\frac{\partial}{\partial y}\left[\frac{1}{f_0}\int_z^\infty \rho_0 \overline{G} dz'\right] \tag{12.8}$$

According to (12.8), $\overline{w}^*$ is zero in the regions, *above* a localized forcing region, whereas it is constant for regions directly below it, hence the term "downward control" is sometimes used to refer to this steady-state situation.

Substitution of (12.8) into the first law of thermodynamics (12.2) and neglecting the time-dependent term yield an expression that explicitly shows the dependence on the zonal force distribution of the departure of the time and zonally averaged temperature from radiative equilibrium:

$$\left(\overline{T} - \overline{T}_r\right) = \frac{N^2 H}{\alpha_r \rho_0 R}\frac{\partial}{\partial y}\left[\frac{1}{f_0}\int_z^\infty \rho_0 \overline{G} dz'\right] \tag{12.9}$$

Thus, in steady state the departure of temperature from radiative equilibrium at a given level depends on the meridional gradient of the zonal force distribution in the column above that level.

For mathematical simplicity the variation of the Coriolis parameter with latitude and other effects of spherical geometry have been neglected in deriving the equations of this section. It is straightforward to extend this model to spherical coordinates. Figure 12.7 shows streamlines of the meridional mass circulation in response to isolated extratropical forcing for three cases corresponding to the differing frequency regimes discussed earlier, but with spherical geometry retained.

A comparison of Figs. 12.2 and 12.4 shows that the largest departures from radiative equilibrium occur in the summer and winter mesosphere and in the polar winter stratosphere. According to (12.9), these are the locations and seasons when zonal forcing must be strongest. The zonal force in the mesosphere is believed to be caused primarily by vertically propagating internal gravity waves. These transfer momentum from the troposphere into the mesosphere, where wave breaking produces strong zonal forcing. The zonal force in the winter stratosphere is due primarily to stationary planetary Rossby waves. These, as discussed in Section 12.3, can propagate vertically provided that the mean zonal wind is westerly and less than a critical value that depends strongly on the wavelength of the waves. Hence, in the extratropical stratosphere we expect a strong annual cycle in $\delta\overline{T}$, with large values (i.e., strong departure from radiative equilibrium) in winter and small in summer. This is indeed observed (see Figs. 12.2 and 12.4). Furthermore, because

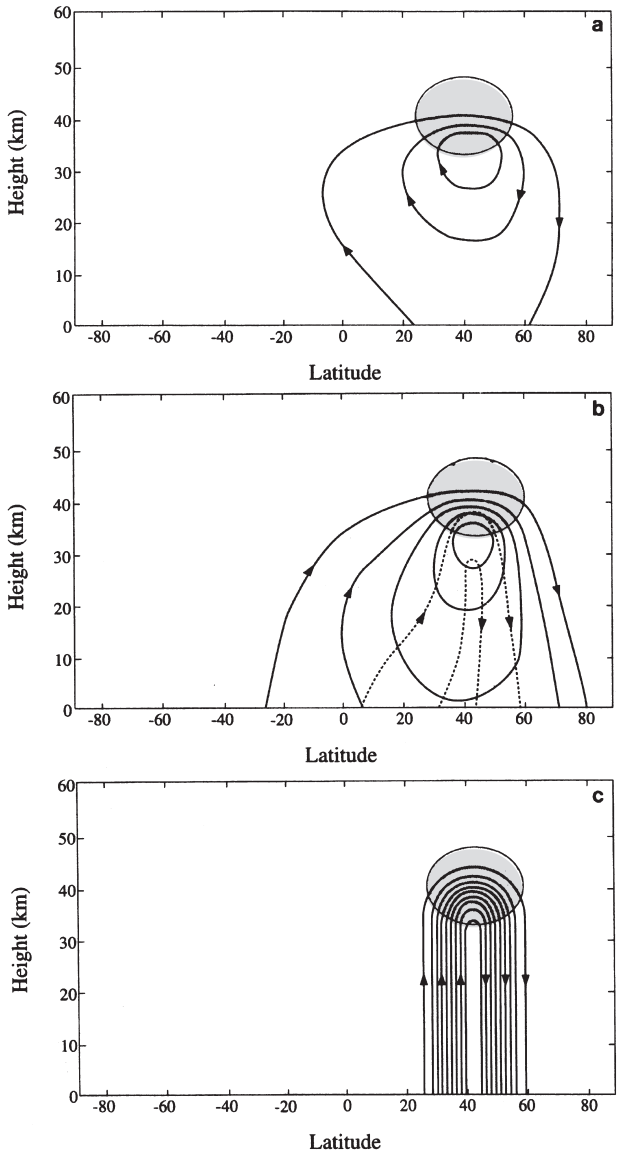


Fig. 12.7 Response of the zonally symmetric circulation to a westward force applied in the shaded region. Contours are streamlines of the TEM meridional circulation, with the same contour interval used in each panel. (a) Adiabatic response for high-frequency forcing. (b) Response for annual frequency forcing and 20-day radiative damping timescale; solid and dashed contours show the response that is in phase and 90° out of phase with the forcing, respectively. (c) Response to steady-state forcing. (After Holton et al., 1995.)

eddy forcing maintains the observed temperature above its radiative equilibrium in the extratropical stratosphere, there is radiative cooling and from (12.3) the residual vertical motion must be downward. By mass continuity it is then required that the residual vertical motion be upward in the tropics, implying that the temperature must be below radiative equilibrium in that region. Note that it is the dynamical driving by extratropical eddies, rather than local forcing, that is responsible for the upward residual motion and net radiative heating in the tropical stratosphere.

12.2.3 Zonal-Mean Transport

Viewed from the TEM perspective the overall meridional circulation in the winter stratosphere is qualitatively as pictured in Fig. 12.8. The residual circulation

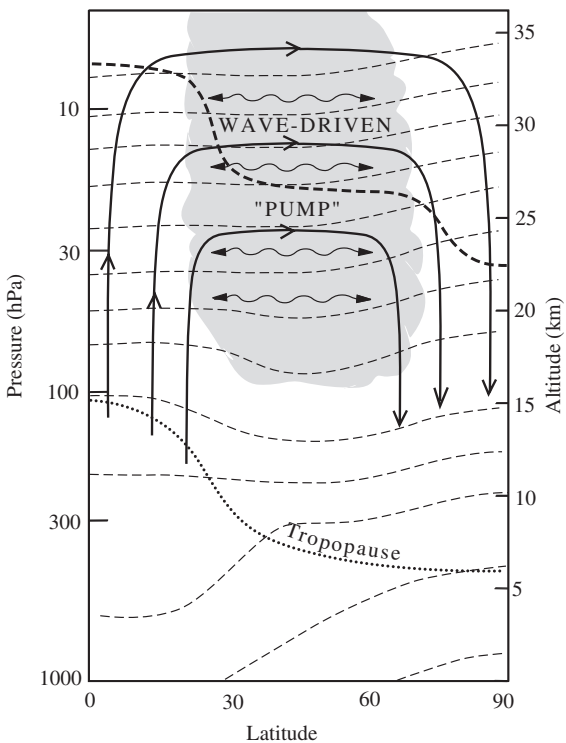


Fig. 12.8 Schematic cross section of the wave-driven circulation in the middle atmosphere and its role in transport. Thin dashed lines denote potential temperature surfaces. Dotted line is the tropopause. Solid lines are contours of the TEM meridional circulation driven by the wave-induced forcing (shaded region). Wavy double-headed arrows denote meridional transport and mixing by eddy motions. Heavy dashed line shows an isopleth of mixing ratio for a long-lived tracer.

transfers mass and trace chemicals upward across the tropopause in the tropics and downward in the extratropics. This vertical circulation is closed in the lower stratosphere by a poleward meridional drift balanced by EP flux convergence. That this schematic picture gives a qualitatively correct view can be ascertained by examination of the zonal mean mixing ratio distribution of any long-lived vertically stratified trace species. As an example, the distribution of N_2O is shown in Fig. 12.9. N_2O is produced at the surface, is mixed uniformly in the troposphere, but decays with height in the stratosphere due to photochemical dissociation. Thus, as shown in Fig. 12.9, the mixing ratio decreases upward in the stratosphere. Note, however, that surfaces of constant mixing ratio are displaced upward in the tropics and downward at higher latitudes, suggesting that the mean meridional mass transport is upward in the tropics and downward in the extratropics as suggested in Fig. 12.9.

In addition to the slow meridional drift by the residual meridional velocity shown in Fig. 12.9, tracers in the winter stratosphere are also subject to rapid quasi-

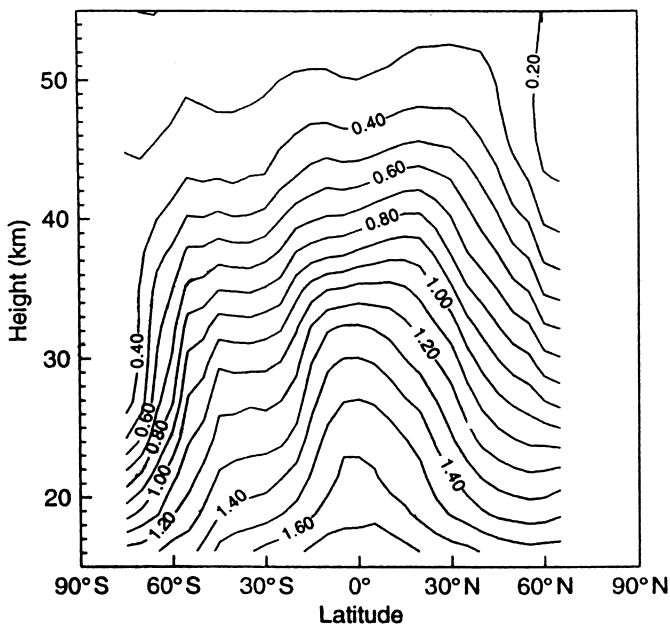


Fig. 12.9 October zonal mean cross section of methane (ppmv) from observations by the Halogen Occultation Experiment (HALOE) on the Upper Atmosphere Research Satellite (UARS). Note the strong vertical stratification due to photochemical destruction in the stratosphere. The upward bulging mixing ratio isopleths in the equatorial region are evidence of upward mass flow in the equatorial region, whereas the downward sloping isopleths in high latitudes are evidence of subsidence in the polar regions. The region of flattened isopleths in the midlatitudes of the Southern Hemisphere is evidence for quasi-adiabatic wave transport due to wintertime planetary wave activity. (After Norton, 2003.)

isentropic transport and mixing by breaking planetary waves. Such eddy transport, which is highly variable in space and time, must be included for quantitatively accurate modeling of transport within the stratosphere.

In the mesosphere the residual circulation is dominated by a single circulation cell with upward motion in the summer polar region, meridional drift from the summer to the winter hemisphere, and subsidence in the winter polar region. This circulation, like the residual circulation in the stratosphere, is eddy driven. However, in the mesosphere it appears that the dominant eddies are vertically propagating internal gravity waves, which have shorter scales in space and time than the planetary waves that dominate the eddy activity in the stratosphere.

12.3 VERTICALLY PROPAGATING PLANETARY WAVES

In Section 12.1 we pointed out that the predominant eddy motions in the stratosphere are vertically propagating quasi-stationary planetary waves (Rossby waves) and that these waves are confined to the winter hemisphere. In order to understand the absence of synoptic-scale motions and the confinement of stationary planetary waves to the winter hemisphere, it is necessary to examine the conditions under which planetary waves can propagate vertically.

12.3.1 Linear Rossby Waves

To analyze planetary wave propagation in the stratosphere, it is convenient to write the equations of motion in the log-pressure coordinate system introduced in Section 8.4.1. For analysis of extratropical planetary wave motions in the middle atmosphere, we may refer the motions to the midlatitude β-plane and use the quasi-geostrophic potential vorticity equation (6.24), which in log-pressure coordinates can be written as

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \bullet \boldsymbol{\nabla}\right) q = 0 \tag{12.10}$$

where

$$q \equiv \nabla^2 \psi + f + \frac{f_0^2}{\rho_0 N^2} \frac{\partial}{\partial z}\left(\rho_0 \frac{\partial \psi}{\partial z}\right)$$

[Compare with Eq. (6.25)]. Here $\psi = \Phi/f_0$ is the geostrophic streamfunction and f_0 is a constant midlatitude reference value of the Coriolis parameter. We now assume that the motion consists of a small-amplitude disturbance superposed on a constant zonal-mean flow. Thus, letting $\psi = -\overline{u}y + \psi'$, $q = \overline{q} + q'$, and linearizing (12.10) we find that the perturbation q' field must satisfy

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right) q' + \beta \frac{\partial \psi'}{\partial x} = 0 \tag{12.11}$$

where

$$q' \equiv \nabla^2 \psi' + \frac{f_0^2}{\rho_0 N^2} \frac{\partial}{\partial z} \left(\rho_0 \frac{\partial \psi'}{\partial z} \right) \tag{12.12}$$

Equation (12.11) has solutions in the form of harmonic waves with zonal and meridional wave numbers k and l and zonal phase speed c_x:

$$\psi'(x, y, z, t) = \Psi(z)\, e^{i(kx+ly-kc_x t)+z/2H} \tag{12.13}$$

Here the factor $e^{z/2H}$ (which is proportional to $\rho_0^{-1/2}$) is included to simplify the equation for the vertical dependence. Substituting (12.13) into (12.11) yields

$$d^2\Psi/dz^2 + m^2\Psi = 0 \tag{12.14}$$

where

$$m^2 \equiv \frac{N^2}{f_0^2} \left[\frac{\beta}{(\overline{u} - \mathrm{c_x})} - \left(k^2 + l^2\right) \right] - \frac{1}{4H^2} \tag{12.15}$$

Referring back to Section 7.4, we recall that $m^2 > 0$ is required for vertical propagation, and in that case m is the vertical wave number [i.e., solutions of (12.14) have the form $\Psi(z) = Ae^{imz}$, where A is a constant amplitude and the sign of m is determined by requiring that the vertical component of group velocity be positive]. For stationary waves ($c_x = 0$) we see from (12.15) that vertically propagating modes can exist only for mean zonal flows that satisfy the condition

$$0 < \overline{u} < \beta \left[\left(k^2 + l^2\right) + f_0^2 / \left(4N^2 H^2\right) \right]^{-1} \equiv U_c \tag{12.16}$$

where U_c is called the *Rossby critical velocity*. Thus, vertical propagation of stationary waves can occur only in the presence of westerly winds weaker than a critical value that depends on the horizontal scale of the waves. In the summer hemisphere the stratospheric mean zonal winds are easterly so that stationary planetary waves are all trapped vertically.

In the real atmosphere the mean zonal wind is not constant, but depends on latitude and height. However, both observational and theoretical studies suggest that (12.16) still provides a qualitative guide for estimating vertical propagation of planetary waves, even though the actual critical velocity may be larger than indicated by the β-plane theory.

12.3.2 Rossby Wavebreaking

The term *wavebreaking* for Rossby waves simply refers to a rapid, irreversible deformation of material contours. Since by (12.10) potential vorticity is approximately conserved in Rossby waves, isolines of potential vorticity on isentropic

surfaces approximate material contours, and wavebreaking can best be illustrated by considering the field of potential vorticity. Wavebreaking may occur when the disturbance fields reach amplitudes for which nonlinear effects can no longer be neglected in the dynamical equations. For example, if the flow is divided into mean and disturbance parts, and nonlinear terms are included, the equation for conservation of quasi-geostrophic potential vorticity (12.10) becomes

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right) q' + v'\frac{\partial \overline{q}}{\partial y} = -u'\frac{\partial q'}{\partial x} - v'\frac{\partial q'}{\partial y} \tag{12.17}$$

For steady waves propagating relative to the ground at zonal phase speed c_x, the variation of phase in time and space is given by $\phi = k(x - c_x t)$, where k is the zonal wave number, and it is readily verified that

$$\frac{\partial}{\partial t} = -c_x \frac{\partial}{\partial x}$$

so that in the linearized version of (12.17) there is a balance between advection of the disturbance potential vorticity q' by the Doppler-shifted mean wind and the advection of mean potential vorticity by the disturbance meridional wind:

$$(\overline{u} - c_x)\frac{\partial q'}{\partial x} = -v'\frac{\partial \overline{q}}{\partial y} \tag{12.18}$$

The validity of the linear approximation can thus be assessed by comparing the sizes of the two terms on the right side of (12.17) with either term in (12.18). Linearity holds provided that

$$\left|\overline{u} - c_x\right| \gg \left|u'\right| \tag{12.19a}$$

and

$$\partial \overline{q}/\partial y \gg \left|\partial q'/\partial y\right| \tag{12.19b}$$

Basically these criteria require that the slope of the material contours in the x, y plane must be small.

As indicated in (12.13), in an atmosphere with constant mean-zonal wind, vertically propagating linear Rossby waves have amplitudes increasing exponentially in height. Thus, at some altitude the disturbance amplitude will become sufficiently large so that wavebreaking must occur. In the real atmosphere, however, the mean zonal flow varies in both latitude and in height, and this variation is crucial for understanding the distribution and mean-flow forcing provided by Rossby wavebreaking. The simplest example of Rossby wavebreaking occurs in the presence of a *critical surface* along which the Doppler-shifted phase speed of the wave

vanishes ($\overline{u} - c_x = 0$). In that case (12.19a) cannot be satisfied even for small amplitude waves.

For understanding of wave behavior near critical lines it is helpful to generalize the Rossby wave analysis of Section 12.3.1 to a situation such as that shown in the seasonal mean cross sections of Fig. 12.2 in which the zonal wind depends on latitude and height, $\overline{u} = \overline{u}(y, z)$. The x and t dependence in (12.11) can then be separated by seeking solutions of the form

$$\psi' = e^{z/2H} \mathrm{Re}\left[\Psi(y, z)\, e^{ik(x - c_x t)}\right] \tag{12.20}$$

we obtain

$$\frac{\partial^2 \Psi}{\partial y^2} + \frac{f_0^2}{N^2}\frac{\partial^2 \Psi}{\partial z^2} + n_k^2 \Psi = 0 \tag{12.21}$$

where the small vertical variation of N^2 is neglected, and

$$n_k^2(y, z) = (\overline{u} - c_x)^{-1}\, \partial \overline{q}/\partial y - k^2 - f_0^2 / \left(4HN^2\right) \tag{12.22}$$

Equation (12.21) has a form similar to that of the equation governing the two-dimensional propagation of light waves in a medium with variable refractive index, n_k. The propagation of linear Rossby wave EP flux in that case can be shown to be along rays that behave somewhat like light rays. Thus, wave activity will tend to propagate along rays that bend toward regions of large positive n_k^2 and avoid regions of negative n_k^2. For stationary Rossby waves ($c_x = 0$) of low zonal wave number, n_k^2 is positive in a region with westerly winds that are not too strong and increases to infinity along a critical surface where the mean flow vanishes. Thus, the index of refraction for wave activity is positive in the winter hemisphere, but increases rapidly toward the equatorial zero wind line. As a result Rossby wave activity tends to propagate upward and equatorward, and wavebreaking occurs in the vicinity of the equatorial critical line.

12.4 SUDDEN STRATOSPHERIC WARMINGS

In the lower stratosphere the temperature is a minimum at the equator and has maxima at the summer pole and at about 45° latitude in the winter hemisphere (see Fig. 12.2). From thermal wind considerations the rapid decrease of temperature poleward of 45° in winter requires a zonal vortex with strong westerly shear with height.

In the Northern Hemisphere, every other year or so this normal winter pattern of a cold polar stratosphere with a westerly vortex is interrupted in a spectacular manner in midwinter. Within the space of a few days the polar vortex becomes highly distorted and breaks down (see Fig. 12.10) with an accompanying large-scale

warming of the polar stratosphere, which can quickly reverse the meridional temperature gradient and (through thermal wind balance) create a circumpolar easterly current. Warmings of 40 K in a few days have occurred at the 50-hPa level.

Numerous observational studies of sudden warmings confirm that enhanced propagation of planetary waves from the troposphere, primarily zonal wave numbers 1 and 2, is essential for the development of the warmings. Since major warmings are observed primarily in the Northern Hemisphere, it is logical to conclude that enhanced wave propagation into the stratosphere is due to topographically forced waves, which are much stronger in the Northern Hemisphere than in the Southern Hemisphere. Even in the Northern Hemisphere, however, it is apparently only in certain winters that conditions are right to produce sudden warmings.

It is generally accepted that sudden warmings are an example of transient mean-flow forcing due to planetary wave driving. Section 10.2.3 showed that in order for planetary waves to decelerate the zonal-mean circulation there must be a nonzero equatorward eddy potential vorticity flux (i.e., a net EP flux convergence). We further showed that for steady nondissipative waves, the divergence of the EP flux vanishes. For normal stationary planetary waves in the stratospheric polar night jet, this constraint should be at least approximately satisfied, as radiative and frictional dissipation are rather small. Thus, the strong interaction between the waves and the mean flow that occurs during the course of a sudden warming must be associated with wave transience (i.e., change of amplitude with respect to time) and wave dissipation. Most of the dramatic mean-flow deceleration that occurs during a sudden warming is caused by amplification of quasi-stationary planetary waves in the troposphere followed by propagation into the stratosphere.

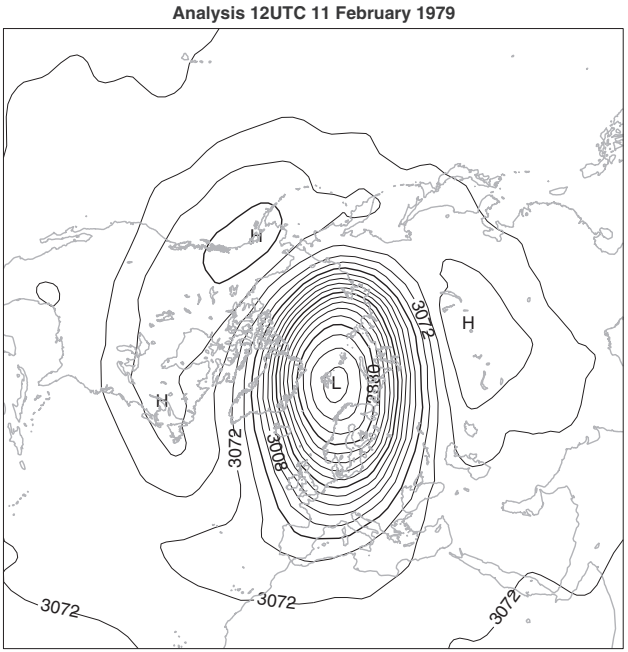
This process can be understood by considering the interaction between zonal mean and eddy potential vorticities. We assume that on the short time scales characteristic of sudden warmings, diabatic processes can be neglected. We also assume that the eddy motion is approximately governed by the linearized eddy potential vorticity equation and that the zonal mean is changed only by the eddy flux convergence. The eddies and mean flow are thus related by the quasi-linear system

$$\left(\frac{\partial}{\partial t} + \overline{u}\,\frac{\partial}{\partial x}\right) q' + v'\frac{\partial \overline{q}}{\partial y} = -S' \tag{12.23}$$

where again

$$q' \equiv \nabla^2 \psi' + \frac{f_0^2}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\frac{\partial \psi'}{\partial z}\right) \tag{12.24}$$

and an eddy damping term, $-S'$, has been added to represent the effects of both mechanical and thermal dissipation. (For constant and equal Rayleigh friction



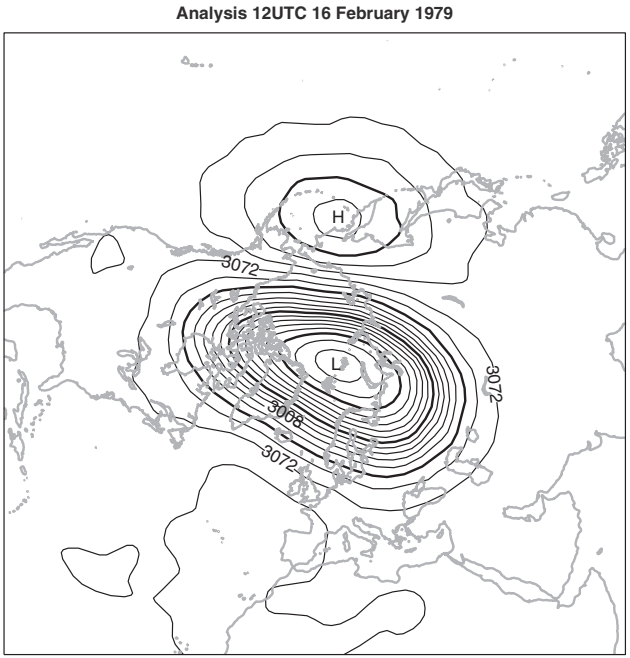


Fig. 12.10 *(continued)*

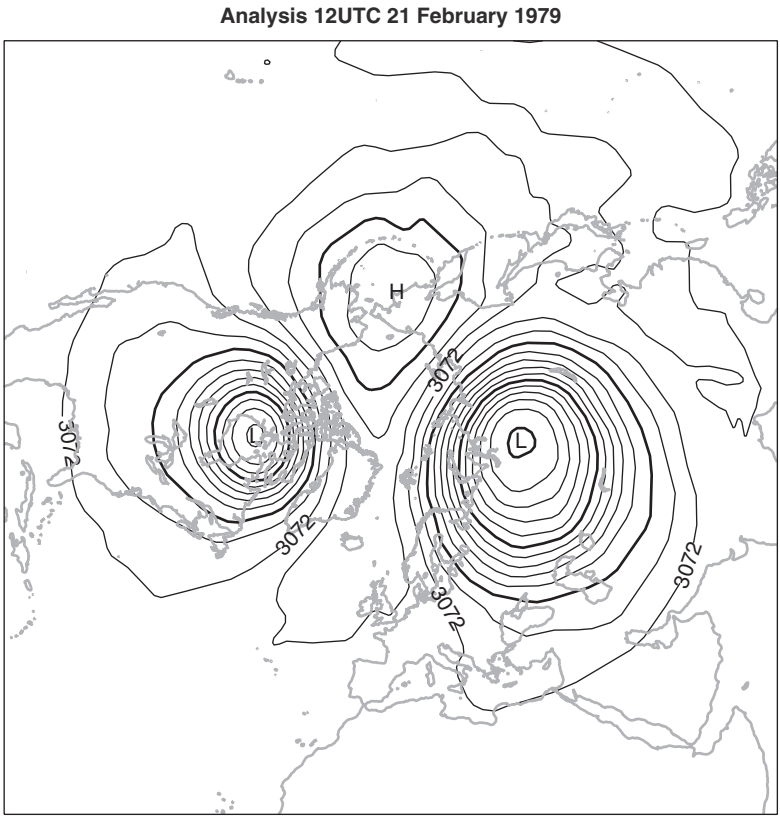


Fig. 12.10 10 hPa geopotential height analyses for February 11, 16, and 21, 1979 at 12UTC showing breakdown of the polar vortex associated with a wave number 2 sudden stratospheric warming. Contour interval: 16 dam. Analysis from ERA-40 reanalysis courtesy of the European Centre for Medium-Range Weather Forecasts (ECMWF).

and Newtonian cooling, $S' = \alpha q'$, where α is the dissipation rate coefficient.) Multiplying (12.23) by q' and averaging zonally after noting that

$$\overline{\overline{u}\, q' \frac{\partial q'}{\partial x}} = \frac{\overline{u}}{2} \frac{\partial \overline{q'^2}}{\partial x} = 0$$

yields with the aid of (10.26)

$$\frac{\partial A}{\partial t} + \nabla \bullet \mathbf{F} = D \tag{12.25}$$

where *wave activity*, A, and dissipation, D, are defined, respectively, as

$$A \equiv \frac{\rho_0 \overline{(q')^2}}{2\partial \overline{q}/\partial y}, \text{ and } \mathrm{D} \equiv -\frac{\rho_0 \overline{\mathrm{S'q'}}}{\partial \overline{\mathrm{q}}/\partial \mathrm{y}}. \tag{12.26}$$

For linear planetary waves satisfying the quasigeostrophic equations, it can be shown that

$$\mathbf{F} = \left(0, c_{gy}, c_{gz}\right) A \tag{12.27}$$

so that the EP flux is given by the wave activity times the group velocity in the meridional (y, z) plane. It is the flux of wave activity—not the flux of wave energy—that is fundamental for wave, mean-flow interaction.

Equation (12.25) shows that if dissipation and transience vanish, the EP flux divergence must vanish. Substitution into (12.1) then gives the *nonacceleration theorem*, which states that there is no wave-driven mean-flow acceleration for steady $(\partial A/\partial t = 0)$ and conservative $(D = 0)$ waves. The EP flux divergence, which must be nonzero for wave-induced forcing, is dependent on wave transience and dissipation.

In sudden stratospheric warmings the planetary wave amplitudes increase rapidly in time. During the Northern Hemisphere winter, quasi-stationary planetary waves of zonal wave numbers 1 and 2 are produced in the troposphere by orographic forcing. These waves propagate vertically into the stratosphere, implying a local increase in the wave activity. Thus, $\partial A/\partial t > 0$, implying from (12.25) that, for quasi-conservative flow, the flux of quasi-geostrophic potential vorticity is negative and that the Eliassen–Palm flux field is convergent:

$$\rho_0 \overline{v'q'} = \nabla \bullet \mathbf{F} < 0 \tag{12.28}$$

Usually $D < 0$ also, so dissipation and wave growth both produce a convergent EP flux (i.e., an equatorward potential vorticity flux).

According to the transformed Eulerian-mean zonal momentum equation (12.1), EP flux convergence leads to a deceleration of the zonal-mean zonal wind, partially offset by the Coriolis torque $f\overline{v}^*$. As a result, the polar night jet weakens, allowing even more waves to propagate into the stratosphere. At some point, the mean zonal wind changes sign so that a critical layer is formed. Stationary linear waves can no longer propagate beyond the level where this occurs, and strong EP flux convergence and even faster easterly acceleration then occurs below the critical layer.

Figure 12.11a illustrates the deceleration of the zonal-mean wind due to the Eliassen–Palm flux convergence caused by upward propagation of transient planetary wave activity. Note that the deceleration is spread out over a broader range in height than the eddy forcing. This reflects the elliptical nature of the equation for mean zonal-wind acceleration that can be derived from the transformed Eulerian mean equations. Fig. 12.11b shows the TEM residual meridional circulation and the pattern of temperature perturbation associated with the zonal wind deceleration. The thermal-wind relation implies that a deceleration of the polar night jet leads to a warming in the polar stratosphere and a cooling in the equatorial stratosphere, with a compensating warming in the polar mesosphere and cooling in the

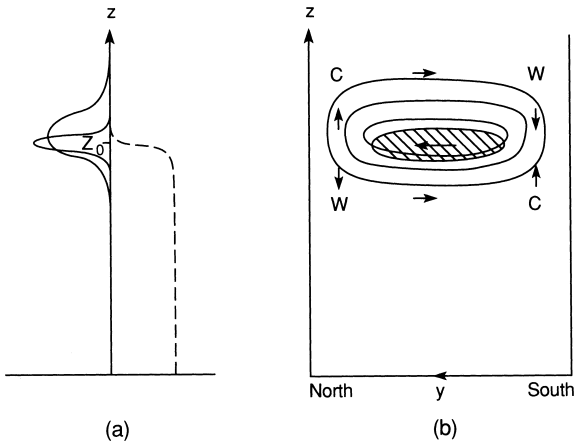


Fig. 12.11 Schematic of transient wave, mean–flow interactions occurring during a stratospheric warming. (a) Height profiles of EP flux (dashed), EP flux divergence (heavy line), and mean zonal flow acceleration (thin line); z_0 is the height reached by the leading edge of the wave packet at the time pictured. (b) Latitude–height cross section showing region where the EP flux is convergent (hatched), contours of induced zonal acceleration (thin lines), and induced residual circulation (arrows). Regions of warming (W) and cooling (C) are also shown. (After Andrews et al., 1987.)

equatorial mesosphere. Eventually, as more of the flow becomes easterly, waves can no longer propagate vertically. The wave-induced residual circulation then decreases, and radiative cooling processes are able to slowly reestablish the normal cold polar temperatures. Thermal-wind balance then implies that the normal westerly polar vortex is also reestablished.

In some cases the wave amplification may be large enough to produce a polar warming, but insufficient to lead to reversal of the mean zonal wind in the polar region. Such "minor warmings" occur every winter and are generally followed by a quick return to the normal winter circulation. A "major warming" in which the mean zonal flow reverses at least as low as the 30-hPa level in the polar region seems to occur only about once every couple of years. If the major warming occurs sufficiently late in the winter, the westerly polar vortex may not be restored at all before the normal seasonal circulation reversal.

12.5 WAVES IN THE EQUATORIAL STRATOSPHERE

Section 11.4 discussed equatorially trapped waves in the context of shallow water theory. Under some conditions, however, equatorial waves (both gravity and Rossby types) may propagate vertically, and the shallow water model must be replaced by a continuously stratified atmosphere in order to examine the vertical structure.

It turns out that vertically propagating equatorial waves share a number of physical properties with ordinary gravity modes. Section 7.5 discussed vertically propagating gravity waves in the presence of rotation for a simple situation in which the Coriolis parameter was assumed to be constant and the waves were assumed to be sinusoidal in both x and y. We found that such inertia–gravity waves can propagate vertically only when the wave frequency satisfies the inequality $f < \nu < N$. Thus, at middle latitudes, waves with periods in the range of several days are generally vertically trapped (i.e., they are not able to propagate significantly into the stratosphere). As the equator is approached, however, the decreasing Coriolis frequency should allow vertical propagation to occur for lower frequency waves. Thus, in the equatorial region there is the possibility for existence of long-period vertically propagating internal gravity waves.

As in Section 11.4 we consider linearized perturbations on an equatorial βplane. The linearized equations of motion, continuity equation, and first law of thermodynamics can then be expressed in log-pressure coordinates as

$$\partial u'/\partial t - \beta y v' = -\partial \Phi'/\partial x \tag{12.29}$$

$$\partial v'/\partial t + \beta y u' = -\partial \Phi'/\partial y \tag{12.30}$$

$$\partial u'/\partial x + \partial v'/\partial y + \rho_0^{-1} \partial \left(\rho_0 w'\right)/\partial z = 0 \tag{12.31}$$

$$\partial^2 \Phi'/\partial t \partial z + w' N^2 = 0 \tag{12.32}$$

We again assume that the perturbations are zonally propagating waves, but we now assume that they also propagate vertically with vertical wave number m. Due to the basic state density stratification, there will also be an amplitude growth in height proportional to $\rho_0^{-1/2}$. Thus, the x, y, z, and t dependencies can be separated as

$$\begin{pmatrix} u' \\ v' \\ w' \\ \Phi' \end{pmatrix} = e^{z/2H} \begin{bmatrix} \hat{u}(y) \\ \hat{v}(y) \\ \hat{w}(y) \\ \hat{\Phi}(y) \end{bmatrix} \exp\left[i\left(kx + mz - \nu t\right)\right] \tag{12.33}$$

Substituting from (12.33) into (12.29)–(12.32) yields a set of ordinary differential equations for the meridional structure:

$$-i\nu\hat{u} - \beta y \hat{v} = -ik\hat{\Phi} \tag{12.34}$$

$$-i\nu\hat{v} + \beta y \hat{u} = -\partial\hat{\Phi}/\partial y \tag{12.35}$$

$$\left(ik\hat{u} + \partial\hat{v}/\partial y\right) + i\left(m + i/2H\right)\hat{w} = 0 \tag{12.36}$$

$$\nu\left(m - i/2H\right)\hat{\Phi} + \hat{w}N^2 = 0 \tag{12.37}$$

12.5.1 Vertically Propagating Kelvin Waves

For Kelvin waves the aforementioned perturbation equations can be simplified considerably. Setting $\hat{v} = 0$ and eliminating $\hat{w}$ between (12.36) and (12.37), we obtain

$$-i\nu\hat{u} = -ik\hat{\Phi} \tag{12.38}$$

$$\beta y\hat{u} = -\partial\hat{\Phi}/\partial y \tag{12.39}$$

$$-\nu\left(m^2 + 1/4H^2\right)\hat{\Phi} + \hat{u}kN^2 = 0 \tag{12.40}$$

Equation (12.38) can be used to eliminate Ψ in (12.39) and (12.40). This yields two independent equations that the field of $\hat{u}$ must satisfy. The first of these determines the meridional distribution of $\hat{u}$ and is identical to (11.47). The second is simply the dispersion equation

$$c^2\left(m^2 + 1/4H^2\right) - N^2 = 0 \tag{12.41}$$

where, as in Section 11.4, $c^2 = (\nu^2/k^2)$.

If we assume that $m^2 \gg 1/4H^2$, as is true for most observed stratospheric Kelvin waves, (12.41) reduces to the dispersion relationship for internal gravity waves (7.44) in the hydrostatic limit ($|k| \ll |m|$). For waves in the stratosphere that are forced by disturbances in the troposphere, the energy propagation (i.e., the group velocity) must have an upward component. Therefore, according to the arguments of Section 7.4, the phase velocity must have a downward component. We showed in Section 11.4 that Kelvin waves must propagate eastward ($c > 0$) if they are to be trapped equatorially. However, eastward phase propagation requires $m < 0$ for downward phase propagation. Thus, the vertically propagating Kelvin wave has phase lines that tilt eastward with height as shown in Fig. 12.12.

12.5.2 Vertically Propagating Rossby–Gravity Waves

For all other equatorial modes, (12.34)–(12.37) can be combined in a manner exactly analogous to that described for the shallow water equations in Section 11.4.1. The resulting meridional structure equation is identical to (11.38) if we again assume that $m^2 \gg 1/4H^2$ and set

$$gh_e = N^2/m^2$$

For the $n = 0$ mode the dispersion relation (11.41) then implies that

$$|m| = N\nu^{-2}(\beta + \nu k) \tag{12.42}$$

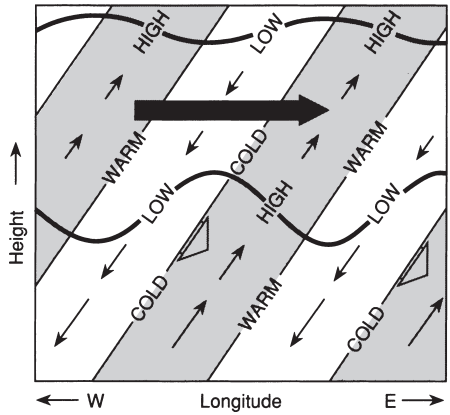


Fig. 12.12 Longitude–height section along the equator showing pressure, temperature, and wind perturbations for a thermally damped Kelvin wave. Heavy wavy lines indicate material lines; short blunt arrows show phase propagation. Areas of high pressure are shaded. Length of the small thin arrows is proportional to the wave amplitude, which decreases with height due to damping. The large shaded arrow indicates the net mean flow acceleration due to the wave stress divergence.

When $\beta = 0$ we again recover the dispersion relationship for hydrostatic internal gravity waves. The role of the β effect in (12.42) is to break the symmetry between eastward ($\nu > 0$) and westward ($\nu < 0$) propagating waves. Eastward propagating modes have shorter vertical wavelengths than westward propagating modes. Vertically propagating $n = 0$ modes can exist only for $c = \nu/k > -\beta/k^2$. Because $k = s/a$, where s is the number of wavelengths around a latitude circle, this condition implies that for $\nu < 0$ solutions exist only for frequencies satisfying the inequality

$$|\nu| < 2\Omega/s \tag{12.43}$$

For frequencies that do not satisfy (12.43), the wave amplitude will not decay away from the equator and it is not possible to satisfy boundary conditions at the pole.

After some algebraic manipulation, the meridional structure of the horizontal velocity and geopotential perturbations for the $n = 0$ mode can be expressed as

$$\begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{\Phi} \end{pmatrix} = v_0 \begin{pmatrix} i\,|m|\,N^{-1}\nu y \\ 1 \\ i\nu y \end{pmatrix} \exp\left(-\frac{\beta\,|m|\,y^2}{2N}\right) \tag{12.44}$$

The westward propagating $n = 0$ mode is generally referred to as the Rossby–gravity mode.[2] For upward energy propagation this mode must have downward

[2] Some authors use this term to describe both eastward and westward $n = 0$ waves.

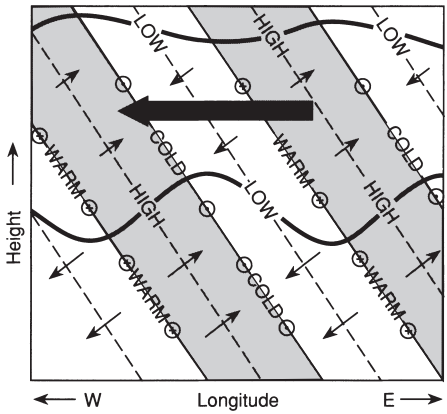


Fig. 12.13 Longitude–height section along a latitude circle north of the equator showing pressure, temperature, and wind perturbations for a thermally damped Rossby–gravity wave. Areas of high pressure are shaded. Small arrows indicate zonal and vertical wind perturbations with length proportional to the wave amplitude. Meridional wind perturbations are shown by arrows pointed into the page (northward) and out of the page (southward). The large shaded arrow indicates the net mean flow acceleration due to the wave stress divergence.

phase propagation ($m < 0$) just like an ordinary westward propagating internal gravity wave. The resulting wave structure in the x, z plane at a latitude north of the equator is shown in Fig. 12.13. Of particular interest is the fact that poleward moving air has positive temperature perturbations and vice versa so that the eddy heat flux contribution to the vertical EP flux is positive.

12.5.3 Observed Equatorial Waves

Both Kelvin and Rossby–gravity modes have been identified in observational data from the equatorial stratosphere. The observed stratospheric Kelvin waves are primarily of zonal wave number $s = 1$ and have periods in the range of 12–20 d. An example of zonal wind oscillations caused by the passage of Kelvin waves at a station near the equator is shown in the form of a time–height section in Fig. 12.14. During the observational period shown in Fig. 12.14, the westerly phase of the quasi-biennial oscillation (see Section 12.6) is descending so that at each level there is a general increase in the mean zonal wind with increasing time. However, superposed on this secular trend is a large fluctuating component with a period between speed maxima of about 12 days and a vertical wavelength (computed from the tilt of the oscillations with height) of about 10–12 km. Observations of the temperature field for the same period reveal that the temperature oscillation leads the zonal wind oscillation by 1/4 cycle (i.e., maximum temperature occurs prior to maximum westerlies), which is just the phase relationship required for upward propagating Kelvin waves (see Fig. 12.12). Additional observations from other

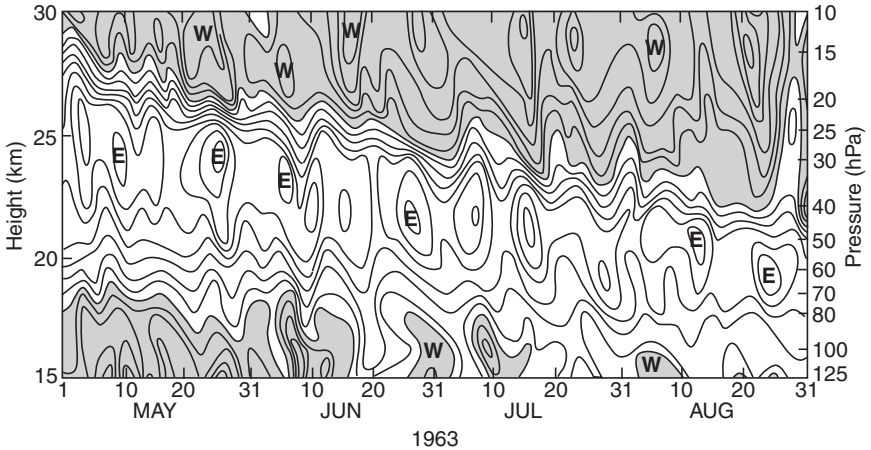


Fig. 12.14 Time–height section of zonal wind at Canton Island (3°S). Isotachs at intervals of 5 m s^{-1}. Westerlies are shaded. (Courtesy of J. M. Wallace and V. E. Kousky.)

stations indicate that these oscillations do propagate eastward at the theoretically predicted speed. Therefore, there can be little doubt that the observed oscillations are Kelvin waves.

The existence of the Rossby–gravity mode has been confirmed in observational data from the stratosphere in the equatorial Pacific. This mode is identified most easily in the meridional wind component, as v' is a maximum at the equator for the Rossby–gravity mode. The observed Rossby–gravity waves have $s = 4$, vertical wavelengths in the range of 6–12 km, and a period range of 4–5 d. Kelvin and Rossby–gravity waves each have significant amplitude only within about 20° latitude of the equator.

A more complete comparison of observed and theoretical properties of the Kelvin and Rossby–gravity modes is presented in Table 12.1. In comparing theory and observation, it must be recalled that it is the frequency relative to the mean flow, not relative to the ground, that is dynamically relevant.

It appears that Kelvin and Rossby–gravity waves are excited by oscillations in the large-scale convective heating pattern in the equatorial troposphere. Although these waves do not contain much energy compared to typical tropospheric weather disturbances, they are the predominant disturbances of the equatorial stratosphere, and through their vertical energy and momentum transport play a crucial role in the general circulation of the stratosphere. In addition to the stratospheric modes considered here, there are higher speed Kelvin and Rossby–gravity modes, which are important in the upper stratosphere and mesosphere. There is also a broad spectrum of equatorial gravity waves, which appears to be important for the momentum balance of the equatorial middle atmosphere.

Table 12.1 *Characteristics of the Dominant Observed Planetary-Scale Waves in the Equatorial Lower Stratosphere* [a]

Theoretical description	Kelvin wave	Rossby–gravity wave
Discovered by	Wallace and Kousky (1968)	Yanai and Maruyama (1966)
Period (ground-based) $2\pi\omega^{-1}$	15 days	4–5 days
Zonal wave number $s = ka\cos\phi$	1–2	4
Vertical wavelength $2\pi\,\mathrm{m}^{-1}$	6–10 km	4–8 km
Average phase speed relative to ground	+25 m s^{-1}	–23 m s^{-1}
Observed when mean zonal flow is	Easterly (maximum ≈ -25 m s^{-1})	Westerly (maximum $\approx +7$ m s^{-1})
Average phase speed relative to maximum zonal flow	+50 m s^{-1}	–30 m s^{-1}
Approximate observed amplitudes		
u'	8 m s^{-1}	2–3 m s^{-1}
v'	0	2–3 m s^{-1}
T'	2–3 K	1 K
Approximate inferred amplitudes		
Φ'/g	30 m	4 m
w'	1.5×10^{-3}m s^{-1}	1.5×10^{-3}m s^{-1}
Approximate meridional scales $(2N/\beta\lvert m\rvert)^{1/2}$	1300–1700 km	1000–1500 km

[a] (After Andrews et al., 1987).

12.6 THE QUASI-BIENNIAL OSCILLATION

The search for periodic oscillations in the atmosphere has a long history. Aside, however, from the externally forced diurnal and annual components and their harmonics, no compelling evidence exists for truly periodic atmospheric oscillations. The phenomenon that perhaps comes closest to exhibiting periodic behavior not associated with a periodic forcing function is the quasi-biennial oscillation (QBO) in the mean zonal winds of the equatorial stratosphere. This oscillation has the following observed features.

1. Zonally symmetric easterly and westerly wind regimes alternate regularly with periods varying from about 24 to 30 months.
2. Successive regimes first appear above 30 km, but propagate downward at a rate of 1 km month^{-1}.
3. The downward propagation occurs without loss of amplitude between 30 and 23 km, but there is rapid attenuation below 23 km.
4. The oscillation is symmetric about the equator with a maximum amplitude of about 20 m s^{-1}, and an approximately Gaussian distribution in latitude with a half-width of about 12°.

This oscillation is best depicted by means of a time–height section of the zonal wind speed at the equator as shown in Fig. 12.15. It is apparent from Fig. 12.15 that the vertical shear of the wind is quite strong at the level where one regime is replacing the other. Because the QBO is zonally symmetric and causes only very small mean meridional and vertical motions, the QBO mean zonal wind and temperature fields satisfy the thermal wind balance equation. For the equatorial β-plane this has the form [compare (10.13)]

$$\beta y \partial \overline{u}/\partial z = -RH^{-1} \partial \overline{T}/\partial y$$

For equatorial symmetry $\partial \overline{T}/\partial y = 0$ at $y = 0$, and by L'Hopital's rule thermal wind balance at the equator has the form

$$\partial \overline{u}/\partial z = -R\,(H\beta)^{-1}\,\partial^2 \overline{T}/\partial y^2 \tag{12.45}$$

Equation (12.45) can be used to estimate the magnitude of the QBO temperature perturbation at the equator. The observed magnitude of vertical shear of the mean zonal wind at the equator is ~ 5 m s^{-1} km^{-1}, and the meridional scale is $\sim$ 1200 km, from which (12.45) shows that the temperature perturbation has an amplitude ~ 3 K at the equator. Because the second derivative of temperature has the opposite sign to that of the temperature at the equator, the westerly and easterly shear zones have warm and cold equatorial temperature anomalies, respectively.

The main factors that a theoretical model of the QBO must explain are the approximate biennial periodicity, the downward propagation without loss of amplitude, and the occurrence of zonally symmetric westerlies at the equator. Because a zonal ring of air in westerly motion at the equator has an angular momentum per unit mass greater than that of the earth, no plausible zonally symmetric advection process could account for the westerly phase of the oscillation. Therefore, there must be a vertical transfer of momentum by eddies to produce the westerly accelerations in the downward-propagating shear zone of the QBO.

Observational and theoretical studies have confirmed that vertically propagating equatorial Kelvin and Rossby–gravity waves provide a significant fraction of the zonal momentum sources necessary to drive the QBO. From Fig. 12.12 it is clear that Kelvin waves with upward energy propagation transfer westerly momentum upward (i.e., u' and w' are positively correlated so that $u'w' > 0$). Thus, the Kelvin waves can provide a source of westerly momentum for the QBO.

Vertical momentum transfer by the Rossby–gravity mode requires special consideration. Examination of Fig. 12.13 shows that $\overline{u'w'} > 0$ also for the Rossby–gravity mode. The total effect of this wave on the mean flow cannot be ascertained from the vertical momentum flux alone, but rather the complete vertical EP flux must be considered. This mode has a strong poleward heat flux ($\overline{v'T'} > 0$), which

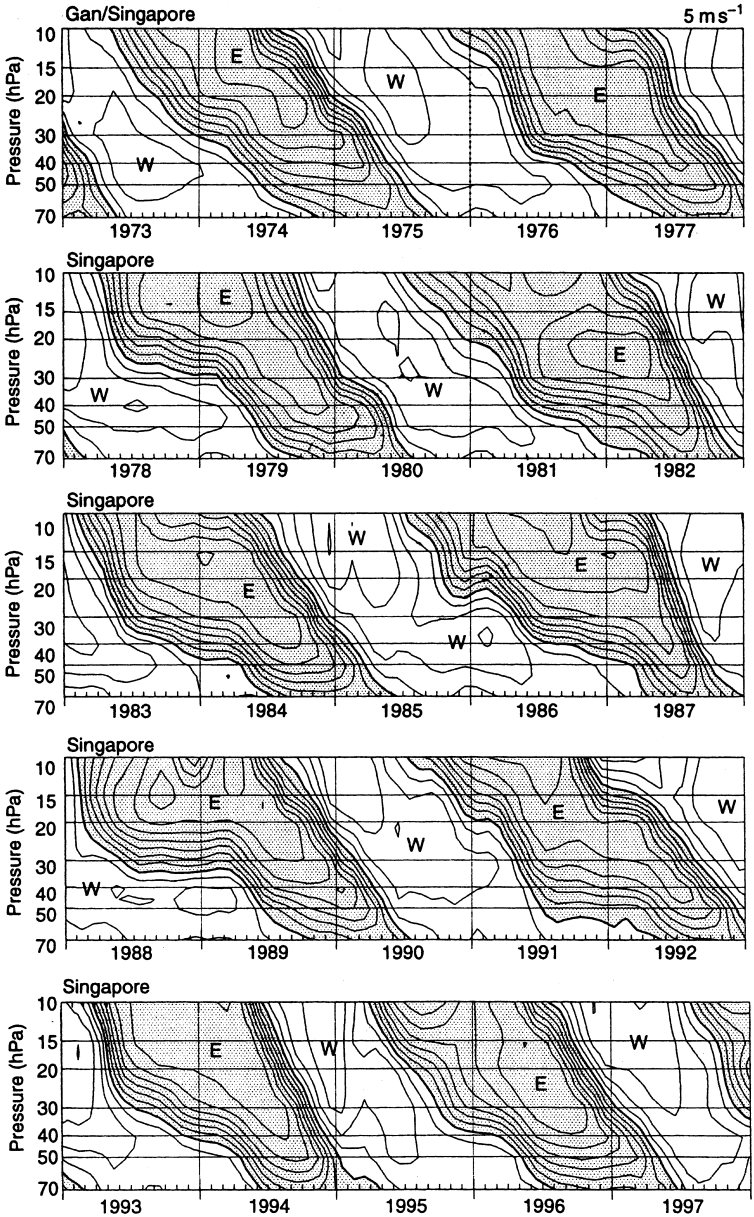


Fig. 12.15 Time–height section of departure of monthly mean zonal winds (m s^{-1}) for each month from the long-term average for that month at equatorial stations. Note the alternating downward propagating westerly (W) and easterly (E) regimes. (After Dunkerton, 2003. Data provided by B. Naujokat.)

provides an upward directed EP flux contribution. This dominates over the vertical momentum flux, and the net result is that the Rossby–gravity mode transfers easterly momentum upward and can provide a momentum source for the easterly phase of the QBO. Observed Kelvin and Rossby–gravity wave momentum fluxes are not, however, sufficient to account for the observed zonal accelerations of the QBO. Additional wave sources, such as gravity waves generated by convective storms, must also contribute to the forcing of the QBO.

It was pointed out in Section 12.4 that quasi-geostrophic wave modes do not produce any net mean flow acceleration unless the waves are transient or are damped mechanically or thermally. Similar considerations apply to gravity waves and to the equatorial Kelvin and Rossby–gravity modes. Stratospheric waves are subject to thermal damping by infrared radiation, and to both thermal and mechanical damping by small-scale turbulent motions. Such damping is strongly dependent on the Doppler-shifted frequency of the waves. As the Doppler-shifted frequency decreases, the vertical component of group velocity also decreases, and there is a longer time available for the wave energy to be damped as it propagates through a given vertical distance. Thus, eastward propagating gravity waves and Kelvin waves tend to be damped preferentially in westerly shear zones, where their Doppler-shifted frequencies decrease with height. The momentum flux convergence associated with this damping provides a westerly acceleration of the mean flow, and thus causes the westerly shear zone to descend. Similarly, the westward propagating gravity waves and Rossby–gravity waves are damped in easterly shear zones, thereby causing an easterly acceleration and descent of the easterly shear zone. We conclude that the QBO is indeed excited primarily by vertically propagating waves through wave transience and damping, which causes westerly accelerations in westerly shear zones and easterly accelerations in easterly shear zones.

This process of wave, mean-flow interaction can be elucidated by considering the heavy wavy lines in Figs. 12.12 and 12.13. These lines indicate the vertical displacement of horizontal surfaces of fluid parcels (material surfaces) by the velocity field associated with the waves. (For sufficiently weak thermal damping they are approximately the same as isentropic surfaces.) The wavy lines show that the maximum upward displacement occurs 1/4 cycle after the maximum upward perturbation velocity. In the Kelvin wave case (Fig. 12.12), positive pressure perturbations coincide with negative material surface slopes. Thus, the fluid below a wavy material line exerts an eastward directed pressure force on the fluid above. Because the wave amplitude decreases with height, this force is larger for the lower of the two material lines in Fig. 12.12. There will thus be a net westerly acceleration of the block of fluid contained between the two wavy material lines shown in Fig. 12.12. For the Rossby–gravity wave, however, positive pressure perturbations coincide with positive slopes of the material lines. There is thus a westward directed force exerted by the fluid below the lines on the fluid above as shown in Fig. 12.13. In this case the result is a net easterly acceleration of the fluid contained

between the two wavy material lines shown in Fig. 12.13. Thus, by considering the stresses acting across a material surface corrugated by the waves, it is possible to deduce the mean-flow acceleration caused by the waves without explicit reference to the EP fluxes of the waves.

How such a mechanism can cause a mean zonal flow oscillation when equal amounts of easterly and westerly momentum are transferred upward across the tropopause by the waves can be seen qualitatively by considering Fig. 12.16. If initially, as shown in Fig. 12.16a, the mean zonal wind is weak and westerly, the eastward propagating waves will be preferentially damped at lower altitude and produce a westerly acceleration, which will move downward in time as the mean westerlies intensify and the wave-driven acceleration is concentrated at ever lower altitudes. The westward propagating waves, however, initially penetrate to higher altitudes, where they produce easterlies, which also move downward in time. Eventually the westerlies are damped out as they approach the tropopause, and the stratosphere is dominated by easterlies so that the eastward propagating waves can penetrate to high altitude and produce a new westerly phase. In this manner the mean zonal wind is forced to oscillate back and forth between westerlies and easterlies with a period that depends primarily on the vertical momentum transport and other properties of the waves, not on an oscillating external forcing.

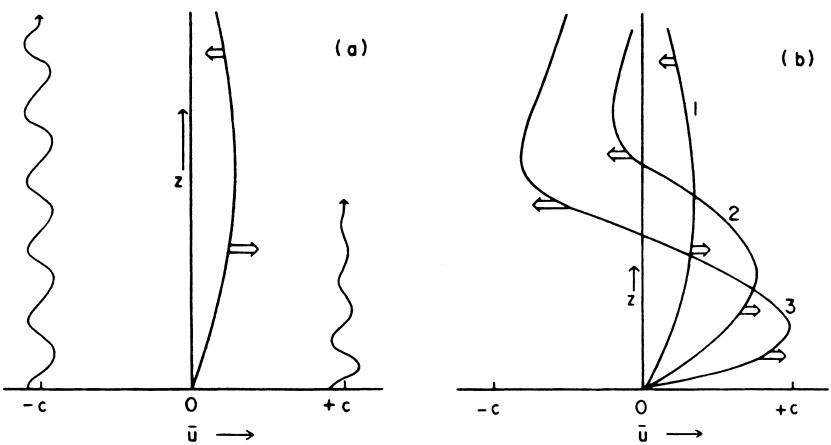


Fig. 12.16 Schematic representation of wave-driven accelerations that lead to the zonal wind QBO. Eastward and westward propagating gravity waves of phase speeds $+c$ and $-c$, respectively, propagate upward and are dissipated at rates dependent on the Doppler-shifted frequency. (a) Initial weak westerly current selectively damps the eastward propagating wave and leads to westerly acceleration at lower levels and easterly acceleration at higher levels. (b) Descending westerly shear zones block penetration of eastward propagating waves, whereas westward propagating waves produce descending easterlies aloft. Broad arrows show locations and direction of maxima in mean wind acceleration. Wavy lines indicate relative penetration of waves. (After Plumb, 1982.)

12.7 TRACE CONSTITUENT TRANSPORT

The study of global transport involves the motion of atmospheric tracers, defined as chemical or dynamical quantities that label fluid parcels. Chemical tracers consist of minor atmospheric species that have significant spatial variability in the atmosphere. Dynamical tracers (potential temperature and potential vorticity) are properties of the flow field that are conserved following the motion under certain conditions. These also can be useful for interpretation of transport.

12.7.1 Dynamical Tracers

Potential temperature, defined in (2.44), can be thought of as a label for the vertical position of an air parcel. Because the atmosphere is stably stratified, potential temperature increases monotonically with height (slowly in the troposphere and rapidly in the stratosphere as shown in Fig. 12.17) and thus can be used as an independent vertical coordinate, the isentropic coordinates introduced in Section 1.2.7. A parcel moving adiabatically remains on a surface of constant potential temperature and can be "tagged" by its value of potential temperature. Thus the motion of such a parcel is two dimensional when viewed in isentropic coordinates.

Potential temperature surfaces are quasi-horizontal, but they do move up and down in physical space with adiabatic changes in temperature. Thus, it is useful to reference trace constituent data to potential temperature, rather than pressure or altitude, as reversible variations at local altitudes or pressures caused by adiabatic vertical displacements associated with transient motions (e.g., gravity waves) are then accounted for.

The other commonly used dynamical tracer, potential vorticity (PV), is conserved for adiabatic, frictionless flow. As defined in (4.12), PV normally has a positive gradient in the meridional direction on an isentropic surface; PV is negative in the Southern Hemisphere, zero at (or near) the equator, and positive in the Northern Hemisphere. PV has strong gradients in both height and latitude. Because a parcel moving in adiabatic frictionless flow conserves PV as well as potential temperature, its motion must lie parallel to isocontours of PV on isentropic surfaces. Rapid meridional transport implies production of strong PV anomalies as high polar values of PV are advected equatorward or low equatorial values are advected poleward. As discussed earlier, the meridional gradient of the background PV field resists meridional displacements through the production of Rossby waves. Thus, regions of strong PV gradients on isentropic surfaces can act as semipermeable "barriers" to transport. This PV barrier effect is one of the reasons that eddy diffusion is often not a good model for meridional transport.

Because PV depends only on the distribution of horizontal winds and temperatures on isentropic surfaces, its distribution can be determined from conventional meteorological observations. Thus, the evolution of PV on an isentropic surface

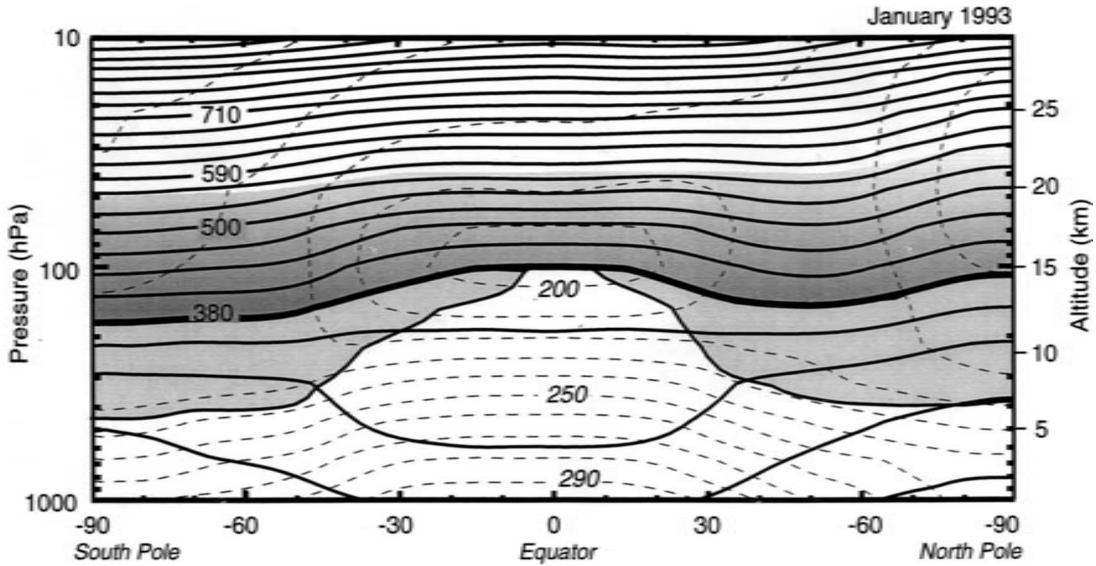


Fig. 12.17 *Latitude–height section showing January mean potential temperature surfaces (solid contours) and temperature (dashed contours). Heavy line shows the 380 K θ surface, above which all θ surfaces lie completely in the middle atmosphere. The light shaded region below the 380 K surface is the lowermost stratosphere, where Θ surfaces span the tropopause (shown by the line bounding the lower side of the shaded region). (Courtesy of C. Appenzeller.)*

can be used as a surrogate for the study of trace constituent transport on isentropic surfaces in circumstances in which there are inadequate observations of relevant chemical tracers. In addition, for dynamical studies, PV has the important property, not shared by other tracers, that it not only is advected by the flow, but actually determines the flow field. Thus, the distribution of PV on isentropic surfaces can be "inverted" to yield the wind and temperature fields. Changes in the PV distribution are then said to "induce" changes in the wind and temperature fields. In the quasi-geostrophic case the induced changes are required to maintain geostrophic and hydrostatic balance. More generally, such changes preserve a higher order balance among the dynamical fields.

12.7.2 Chemical Tracers

Much can be learned about the nature of large-scale transport by considering the climatological distribution of quasi-conservative chemical tracers in the atmosphere. The distribution of such trace substances is dependent on the competition between dynamical and chemical processes. This competition can be approximately measured by comparing the characteristic dynamical and chemical time scales. By *chemical timescale* is meant the average time for replacement or removal of a tracer due to chemical sources and sinks. By *dynamical timescale* is meant the average time for advective and diffusive processes to transport the tracer from equator to pole or across a scale height in the vertical. The role of transport in determining tracer climatology depends on the nature and distribution of tracer sources and sinks, and on the relative magnitudes of the timescales for dynamical and chemical processes.

If the chemical timescale is much shorter than the dynamical timescale, the tracer will be in photochemical equilibrium and transport does not directly influence its distribution. Transport can, however, play an important indirect role by partly determining the concentrations of other species that participate in the photochemical production or loss of the tracer in question.

If the chemical timescale is much longer than the dynamical timescale, the tracer will be passively advected by the flow field. In the absence of localized sources and sinks, such a tracer would eventually become well mixed due to the dispersive effects of transport. It is for this reason that species such as nitrous oxide (N_2O) tend to have uniform concentrations in the troposphere and are thus not useful for tropospheric transport studies.

When the chemical and dynamical timescales are comparable, the observed species concentration depends on the net effects of chemical sources and sinks and of transport. In many cases of interest, the ratio of chemical and dynamical timescales changes drastically between the troposphere and the stratosphere, and with altitude in the stratosphere. Tracers that have long lifetimes in the troposphere and the lower stratosphere are referred to as *long-lived tracers*, as the bulk of the

tracer mass is contained in the troposphere and lower stratosphere, and the lifetime is determined mainly by the (slow) flux of the tracer to altitudes where the chemical timescales become comparable to or shorter than dynamical timescales.

12.7.3 Transport in the Stratosphere

Transport processes are conveniently divided between those processes that involve mean motions of the atmosphere, or *advection,* and those processes that may be characterized as turbulent, or diffusive in nature. In the case of point sources, such as volcanic eruptions, the distinction is quite clear; advection moves the center of mass of the plume along the direction of the average wind, whereas turbulent diffusion disperses the plume in the plane orthogonal to the average wind. On a global scale, however, the distinction between advective and diffusive processes is not always clear. Because the atmosphere is characterized by spatially and temporally varying motions with a wide range of scales, there is no obvious physical separation between "mean" and "turbulent" motions. In practice, those transport processes that are explicitly resolved by the particular observational network or transport model being utilized are often regarded as the advective motions, whereas the remaining unresolved motions are assumed to be diffusive. The effects of the unresolved motions must then be parameterized in terms of the mean motions in some fashion. Usually this involves assuming that tracer fluxes by the unresolved motions are proportional to the gradient of the resolved tracer distribution. However, this approach is not always physically justified. A major problem in the modeling of global transport is accurate representation of the contribution of the unresolved eddy motions to the total transport.

As discussed in Section 12.2, the global-scale residual meridional circulation in the middle atmosphere is driven by wave-induced zonal forces associated with Rossby waves and gravity waves. Not surprisingly, the residual circulation plays an essential role in the meridional and vertical transport of trace chemical constituents in the middle atmosphere. Additionally, the waves responsible for the zonal force that drives the residual circulation are also responsible for the quasi-isentropic stirring and mixing that is associated with wavebreaking. Thus, understanding of transport involves both eddy and mean-flow transport effects.

In dynamical studies it is usual to characterize a chemical constituent by the volume mixing ratio (or mole fraction), defined as $\chi \equiv n_T/n_A$, where n_T and n_A designate the number densities (molecules m^{-3}) for the trace constituent and air, respectively. The mixing ratio is conserved following the motion in the absence of sources and sinks and hence satisfies the simple tracer continuity equation

$$\frac{D\chi}{Dt} = S \tag{12.46}$$

where S designates the sum of all chemical sources and sinks.

As in the case of the dynamical variables discussed in Section 12.2, it is useful to define a longitudinally averaged mixing ratio $\overline{\chi}$, and a disturbance or eddy ratio χ' such that $\chi = \overline{\chi} + \chi'$. Again, it proves useful to utilize the residual mean meridional circulation $\left(\overline{v}^*, \overline{w}^*\right)$ defined in (10.16ab). The zonal mean tracer continuity equation in the TEM framework can then be written as

$$\frac{\partial \overline{\chi}}{\partial t} + \overline{v}^* \frac{\partial \overline{\chi}}{\partial y} + \overline{w}^* \frac{\partial \overline{\chi}}{\partial z} = \overline{S} + \frac{1}{\rho_0} \boldsymbol{\nabla} \cdot \mathbf{M} \tag{12.47}$$

Here, **M** represents the diffusive effects of the eddies, plus advective effects not represented by the residual meridional circulation. In models the term involving **M** is often represented by meridional and vertical eddy diffusion, with empirically determined diffusion coefficients.

In order to appreciate the role of the wave-induced global circulation in determining the distribution of long-lived tracers in the middle atmosphere, it is useful to consider a hypothetical atmosphere in which there are no wave motions, and hence no wave-induced zonal force. In that case, as argued in Section 12.2, the middle atmosphere would relax to radiative equilibrium, the residual circulation would vanish, and the distribution of the tracer would be determined at each altitude by a balance between slow upward diffusion and photochemical destruction. Thus, tracer mixing ratio surfaces would, in an annual mean, tend to be close to horizontal. This is to be contrasted to observed distributions, which are characterized by mixing ratio surfaces that are bowed upward in the tropics and slope downward toward both poles (e.g., Fig. 12.9).

As discussed in Section 12.2, the wave-induced global-scale circulation consists of upward and poleward motion across the isentropes in low latitudes, accompanied by diabatic heating, and downward motion across the isentropes at high latitudes, accompanied by diabatic cooling. Actual parcel trajectories, of course, do not follow the zonally averaged motion, but are influenced by the three-dimensional wave motion. Nevertheless the diabatic circulation defined by the mean diabatic heating and cooling closely approximates the global transport circulation. For seasonal and longer timescales the TEM residual circulation generally provides a good approximation to the diabatic circulation and is generally simpler to compute from standard meteorological analyses. For shorter period phenomena in which the temperature tendency is large, the residual circulation is no longer a good approximation to the diabatic circulation.

The above conceptual model of global transport is clearly supported by long-lived tracer observations as shown in Figure 12.9. In middle latitudes there are regions in which tracer mixing ratio isopleths are nearly horizontal, reflecting the horizontal homogenizing role of meridional dispersion by planetary wave-breaking in the surf zone. The same processes also tend to homogenize the PV distribution on isentropes in the so-called surf zone in midlatitudes where Rossby wavebreaking tends to occur. The surf zone is bounded at both low and high

latitudes by regions of strong meridional gradients of long-lived tracers and of PV. The existence of such gradients is evidence that there is only weak mixing into and out of the tropics and into and out of the polar winter vortex. Thus these locations are sometimes referred to as "transport barriers." The strong PV gradients, strong winds, and strong wind shears that occur along the transport barriers at the subtropical and polar edges of the surf zone all act to suppress wave-breaking, and hence to minimize mixing and sustain the strong gradients at those locations.

PROBLEMS

12.1. Suppose that temperature increases linearly with height in the layer between 20 and 50 km at a rate of 2 K km^{-1}. If the temperature is 200 K at 20 km, find the value of the scale height H for which the log-pressure height z coincides with actual height at 50 km. (Assume that z coincides with the actual height at 20 km and let $g = 9.81$ m s^{-2} be a constant.)

12.2. Find the Rossby critical velocities for zonal wave numbers 1, 2, and 3 (i.e., for 1, 2, and 3 wavelengths around a latitude circle). Let the motion be referred to a β plane centered at 45°N, scale height $H = 7$ km, buoyancy frequency $N = 2 \times 10^{-2}$ s^{-1}, and infinite meridional scale ($l = 0$).

12.3. Suppose that a stationary linear Rossby wave is forced by flow over sinusoidal topography with height $h(x) = h_0 \cos(kx)$, where h_0 is a constant and k is the zonal wave number. Show that the lower boundary condition on the streamfunction ψ can be expressed in this case as

$$(\partial\psi/\partial z) = -hN^2/f_0$$

Using this boundary condition and an appropriate upper boundary condition, solve for $\psi(x, z)$ in the case $|m| \gg (1/2H)$ using the equations of Section 12.3.1. How does the position of the trough relative to the mountain ridge depend on the sign of m^2 for the limit $|m| \gg (1/2H)$.

12.4. Consider a very simple model of a steady-state mean meridional circulation for zonally symmetric flow in a midlatitude channel bounded by walls at $y = 0, \pi/l$ and $z = 0, \pi/m$. We assume that the zonal mean zonal flow $\overline{u}$ is in thermal wind balance and that the eddy momentum and heat fluxes vanish. For simplicity, we let $\rho_0 = 1$ (Boussinesq approximation) and let the zonal force due to small-scale motions be represented by a linear drag: $\overline{X} = -\gamma\overline{u}$. We assume that the diabatic heating has the form $J/c_p = (H/R)\, J_0 \cos ly \sin mz$, and we let N and f be constants.

Equations (12.1), (12.2), (12.3), and (12.4) then yield the following:

$$-f_0\overline{v}^* = -\gamma\overline{u}$$

$$+N^2HR^{-1}\overline{w}^* = +\overline{J}/c_p$$

$$\overline{v}^* = -\frac{\partial\overline{\chi}^*}{\partial z}; \quad \overline{w}^* = \frac{\partial\overline{\chi}^*}{\partial y}$$

$$f_0\partial\overline{u}/\partial z + RH^{-1}\partial\overline{T}/\partial y = 0$$

Assuming that there is no flow through the walls, solve for the residual circulation defined by $\overline{\chi}^*$, $\overline{v}^*$, and $\overline{w}^*$.

12.5. For the situation of Problem 12.4, solve for the steady-state zonal wind and temperature fields $\overline{u}$ and $\overline{T}$.

12.6. Find the geopotential and vertical velocity fluctuations for a Kelvin wave of zonal wave number 1, phase speed 40 m s^{-1}, and zonal velocity perturbation amplitude 5 m s^{-1}. Let $N^2 = 4 \times 10^{-4}$ s^{-2}.

12.7. For the situation of Problem 12.6 compute the vertical momentum flux $M \equiv \rho_0\overline{u'w'}$. Show that M is constant with height.

12.8. Determine the form for the vertical velocity perturbation for the Rossby–gravity wave corresponding to the u', v', and Φ' perturbations given in (12.44).

12.9. For a Rossby–gravity wave of zonal wave number 4 and phase speed -20 m s^{-1}, determine the latitude at which the vertical momentum flux $M \equiv \rho_0\overline{u'w'}$ is a maximum.

12.10. Suppose that the mean zonal wind shear in the descending westerlies of the equatorial QBO can be represented analytically on the equatorial β plane in the form $\partial u/\partial z = \Lambda \exp\left(-y^2/L^2\right)$ where $L = 1200$ km. Determine the approximate meridional dependence of the corresponding temperature anomaly for $|y| \ll L$.

12.11. Estimate the TEM residual vertical velocity in the westerly shear zone of the equatorial QBO assuming that radiative cooling can be approximated by Newtonian cooling with a 20-day relaxation time, that the vertical shear is 20 m s^{-1} per 5 km, and that the meridional half-width is 12° latitude.

MATLAB EXERCISES

M12.1. The MATLAB script **topo_Rossby_wave.m** plots solutions for various fields for a stationary linear Rossby wave forced by flow over an isolated ridge. A β–plane channel model is used following the discussion in

Section 12.3.1. Run the script for mean wind values of 5, 10, 15, 20, and 25 m s^{-1}. Describe how the geopotential and potential temperature perturbations change with changing mean wind speed. Explain these results based on the arguments given in Section 12.3.1.

M12.2. For the situation of Problem 12.4 let $J_0 = 10^{-6}$ s^{-3}, $N = 10^{-2}$ s^{-1}, $f = 10^{-4}$ s^{-1}, $l = 10^{-6}$ m^{-1}, $m = \pi/H$, where $H = 10^4$ m, and γ 10^{-5} s^{-1}. Make contour plots of the fields $\overline{u}$, $\overline{v}^*$, $\overline{w}^*$, $\overline{T}$ and discuss the relationship of these fields to the imposed diabatic heating.

M12.3. The MATLAB script **sudden_warming.m** provides an analogue to the sudden stratospheric warmings on an extratropical β plane centered at 60°N. The amplitude of a single planetary wave of zonal wave number $s = 1$ or $s = 2$ is specified at the bottom boundary (taken to be the 16-km level), and a time integration is carried out to determine the evolution of the zonal mean wind in the stratosphere forced by the EP flux of the specified wave. For the cases $s = 1$ and 2, run the model for geopotential heights between 100 and 400 m at a 50-m interval. For each case note whether the flow tends toward steady state or repeated sudden warmings. Modify the code to plot the time–height evolution of the EP flux divergence for the case $s = 2$ and forcing of 200 m.

M12.4. The MATLAB script **qbo_model.m** is a simplified one-dimensional model of the equatorial QBO originally introduced by Plumb (1977). In this model the mean zonal flow is forced by two waves of equal amplitude at the lower boundary and equal and opposite phase speeds. The initial mean wind profile has a weak westerly shear. If the model is run for sufficiently weak forcing the mean wind approaches a steady state, but for forcing greater than a critical amplitude a downward propagating oscillatory solution occurs, with period dependent on the amplitude of forcing. Run the script for a range of forcing between 0.04 and 0.40 and determine the approximate minimum forcing for a large-amplitude oscillatory motion to occur, and the dependence of period on the forcing amplitude. Modify the script to compute the time mean of the zonal wind over the course of several oscillations. Try a case in which the eastward forcing is 0.15 and the westward forcing is 0.075. How does the time mean wind change in this case?

Suggested References

Brasseur and Solomon, *Aeronomy of the Middle Atmosphere*, has an excellent discussion of the chemistry of the stratosphere at an advanced level.

Andrews, Holton, and Leovy, *Middle Atmosphere Dynamics*, present a graduate-level treatment of the dynamics of the stratosphere and mesosphere.

CHAPTER 13

Numerical Modeling and Prediction

Dynamical meteorology provides the theoretical basis and methodology for modern weather forecasting. Stated simply, the objective of dynamical forecasting is to predict the future state of the atmospheric circulation from knowledge of its present state by use of numerical approximations to the dynamical equations. To fulfill this objective requires observations of the initial state of the field variables, a closed set of prediction equations relating the field variables, and a method of integrating the equations in time to obtain the future distribution of the field variables.

Numerical prediction is a highly specialized field, which is continually evolving. Operational forecast centers utilize complex prediction models that require the largest available supercomputers for their solution. It is difficult to provide more than a superficial introduction to such models in an introductory text. Fortunately, however, many aspects of numerical prediction can be illustrated using a simple model, such as the barotropic vorticity equation. In fact, this equation was the basis of the earliest operational numerical prediction models.

13.1 HISTORICAL BACKGROUND

The British scientist L. F. Richardson made the first attempt to predict the weather numerically. His book, *Weather Prediction by Numerical Process*, published in 1922, is the classic treatise in this field. In his work Richardson showed how the differential equations governing atmospheric motions could be written approximately as a set of algebraic difference equations for values of the tendencies of various field variables at a finite number of points in space. Given the observed values of the field variables at these grid points, the tendencies could be calculated numerically by solving the algebraic difference equations. By extrapolating the computed tendencies ahead a small increment in time, an estimate of the fields at a short time in the future could be obtained. The new values of the field variables could then be used to recompute the tendencies, which could in turn be used to extrapolate further ahead in time, and so on. Even for short-range forecasting over a small area of the earth, this procedure requires an enormous number of arithmetic operations. Richardson did not foresee the development of high-speed digital computers. He estimated that a work force of 64,000 people would be required just to keep up with the weather on a global basis!

Despite the tedious labor involved, Richardson worked out one example forecast for surface pressure tendencies at two grid points. Unfortunately, the results were very poor. Predicted pressure changes were an order of magnitude larger than those observed. At the time this failure was thought to be due primarily to the poor initial data available, especially the absence of upper-air soundings. However, it is now known that there were other, even more serious, problems with Richardson's scheme.

After Richardson's failure to obtain a reasonable forecast, numerical prediction was not again attempted for many years. Finally, after World War II interest in numerical prediction revived due partly to the vast expansion of the meteorological observation network, which provided much improved initial data, but even more importantly to the development of digital computers, which made the enormous volume of arithmetic operations required in a numerical forecast feasible. At the same time it was realized that Richardson's scheme was not the simplest possible scheme for numerical prediction. His equations not only governed the slow-moving meteorologically important motions, but also included high-speed sound and gravity waves as solutions. Such waves are in nature usually very weak in amplitude. However, for reasons that will be explained later, if Richardson had carried his numerical calculation beyond the initial time step, these oscillations would have amplified spuriously, thereby introducing so much "noise" in the solution that the meteorologically relevant disturbances would have been obscured.

The American meteorologist J. G. Charney showed in 1948 how the dynamical equations could be simplified by systematic introduction of the geostrophic and hydrostatic approximations so that the sound and gravity oscillations were

filtered out. The equations that resulted from Charney's filtering approximations were essentially those of the quasi-geostrophic model. Thus, Charney's approach utilized the conservative properties of potential vorticity. A special case of this model, the *equivalent barotropic model*, was used in 1950 to make the first successful numerical forecast.

This model provided forecasts of the geopotential near 500 hPa. Thus, it did not forecast "weather" in the usual sense. It could, however, be used by forecasters as an aid in predicting the local weather associated with large-scale circulations. Later multilevel versions of the quasi-geostrophic model provided explicit predictions of the surface pressure and temperature distributions, but the accuracy of such predictions was limited due to the approximations inherent in the quasi-geostrophic model.

With the development of vastly more powerful computers and more sophisticated modeling techniques, numerical forecasting has now returned to models that are quite similar to Richardson's formulation, and are far more accurate than quasi-geostrophic models. Nevertheless, it is still worth considering the simplest filtered model, the barotropic vorticity equation, to illustrate some of the technical aspects of numerical prediction in a simple context.

13.2 FILTERING METEOROLOGICAL NOISE

One difficulty in directly applying the unsimplified equations of motion to the prediction problem is that meteorologically important motions are easily lost in the noise introduced by large-amplitude sound and gravity waves, which may arise as a result of errors in the initial data, and then can spuriously amplify to dominate the forecast fields. As an example of how this problem might arise, consider the pressure and density fields. On the synoptic scale these are in hydrostatic balance to a very good approximation. As a consequence, vertical accelerations are extremely small. However, if the pressure and density fields were determined independently by observations, as would be the case if the complete dynamical equations were used, small errors in the observed fields would lead to large errors in the computed vertical acceleration simply because the vertical acceleration is the very small difference between two large forces (the vertical pressure gradient and buoyancy). Such spurious accelerations would appear in the computed solutions as high-speed sound waves of very large amplitude. In a similar fashion, errors in the initial horizontal velocity and pressure fields would lead to spuriously large horizontal accelerations, as horizontal acceleration results from the small difference between Coriolis and pressure gradient forces. Such spurious horizontal accelerations would excite both sound and gravity waves.

In order to overcome this problem, either the "observed" fields must be modified systematically to remove unrealistic force imbalances or the prediction equations

must be modified to remove the physical mechanisms responsible for the occurrence of the unwanted oscillations, while still preserving the meteorologically important motions. In modern numerical forecasting, gravity waves are usually controlled by suitably adjusting the initial data; sound waves, however, are generally filtered from the dynamical equations.

To help understand how sound and gravity wave noise can be filtered from the prediction equations, it is useful to refer back to Section 7.3 in which the physical properties of such waves were discussed. If the pipe shown in Fig. 7.5 is tipped up in the vertical, it can be used to generate vertically traveling sound waves. If we now require that pressure be hydrostatic, the pressure at any point along the pipe is determined solely by the weight of the air above that point. Hence, the vertical pressure gradient cannot be influenced by adiabatic compression or expansion. Therefore, vertically propagating sound waves are not among the possible modes of oscillation in a hydrostatic system. Replacement of the vertical momentum equation by the hydrostatic approximation is thus sufficient to filter out ordinary sound waves. This approximation is used in virtually all models for forecasting medium- and large-scale motions.

A hydrostatically balanced atmosphere can still, however, support a special class of horizontally propagating acoustic waves. In this type of wave, the vertical velocity is zero (neglecting orographic effects and departures of the basic state temperature from isothermal conditions). The pressure, horizontal velocity, and density, however, oscillate with the horizontal structure of the simple acoustic waves described in Section 7.3.1. These oscillations have maximum amplitude at the lower boundary and decay away from the boundary with the pressure and density fields remaining in hydrostatic balance everywhere (see Problem 7.6). Because these oscillations, known as *Lamb waves*, have maximum pressure perturbation amplitude at the ground, they may be filtered simply by requiring that $\omega = Dp/Dt = 0$ at the lower boundary. This boundary condition is applied most easily by formulating the equations in isobaric coordinates. In that case the condition $\omega = 0$ at the lower boundary is a natural first approximation for geostrophically scaled motions over level terrain (see the discussion in Section 3.5).

The set of equations including the above approximations still has solutions in the form of internal gravity waves. Gravity waves may be important in accounting for mesoscale variability, and the zonal drag force due to gravity waves must be included in the overall zonal mean momentum balance. However, gravity waves are not important for short-range forecasting of synoptic and planetary scale circulations. Due, however, to their high frequency of oscillation, they can create serious errors in numerical forecasts. One possible solution to this problem is to filter gravity waves from the forecast equations. From the discussion of Section 7.4 it should be apparent that gravity waves can be solutions only if the dynamical equations are second order in time. In physical terms this turns out to require that the local rate of change of the divergence of the horizontal velocity field must be implicitly

included in the equations. One way to filter gravity waves is thus to divide the horizontal velocity into irrotational and nondivergent parts as discussed in Section 11.2 and to neglect the irrotational part of the motion in the velocity tendency term. The resulting prognostic equation is the barotropic vorticity equation (11.14). The stream function for the nondivergent flow can then be diagnostically related to the geopotential field by the nonlinear balance equation (11.15). For extratropical synoptic scale motions, (11.15) can be replaced approximately by geostrophic balance on the midlatitude β-plane, for which $\psi = \Phi/f_0$.

The barotropic vorticity equation can be used to compute the evolution of the flow at a single level in terms of the nondivergent flow at that level alone. No vertical coupling is involved. Thus, no explicit predictions for levels above or below the assumed nondivergent level are possible. The only "predictions" of surface weather possible with this model are those based on climatological relationships between the surface flow and flow at the nondivergent level. For extratropical flows it is possible to explicitly model the vertical structure and still filter gravity waves by utilizing the conservation of quasi-geostrophic potential vorticity (6.24), which is related to the geopotential field through the elliptic boundary value problem (6.25). Sections 13.4 and 13.5 utilize the barotropic vorticity equation to illustrate the basic methodology of numerical weather prediction.

13.3 NUMERICAL APPROXIMATION OF THE EQUATIONS OF MOTION

The equations of motion are an example of a general class of systems known as *initial value problems*. A system of differential equations is referred to as an initial value problem when the solution depends not only on boundary conditions, but also on the values of the unknown fields or their derivatives at some initial time. Clearly, weather forecasting is a primary example of a *nonlinear* initial value problem. Due to its nonlinearity even the simplest forecast equation, the barotropic vorticity equation, is rather complicated to analyze. Fortunately, general aspects of the numerical solution of initial value problems can be illustrated using linearized prototype equations that are much simpler than the barotropic vorticity equation.

13.3.1 Finite Differences

The equations of motions involve terms that are quadratic in the dependent variables (the advection terms). Such equations generally cannot be solved analytically. Rather, they must be approximated by some suitable discretization and solved numerically. The simplest form of discretization is the *finite difference* method.

To introduce the concept of finite differencing, we consider a field variable $\psi(x)$, which is a solution to some differential equation in the interval $0 \leq x \leq L$. If this

interval is divided into J subintervals of length δx, then $\psi(x)$ can be approximated by a set of $J+1$ values as $\Psi_j = \psi(j\delta x)$, which are the values of the field at the $J+1$ grid points given by $x = j\delta x$, $j = 0, 1, 2, \ldots, J$, where $\delta x = L/J$. Provided that δx is sufficiently small compared to the scale on which ψ varies, the $J+1$ grid point values should provide good approximations to $\psi(x)$ and its derivatives.

The limits of accuracy of a finite difference representation of a continuous variable can be assessed by noting that the field can also be approximated by a finite Fourier series expansion:

$$\psi(x) = \frac{a_0}{2} + \sum_{m=1}^{J/2} \left[a_m \cos \frac{2\pi m x}{L} + b_m \sin \frac{2\pi m x}{L} \right] \tag{13.1}$$

The available $J+1$ values of Ψ_j are just sufficient to determine the $J+1$ coefficients in (13.1). That is, it is possible to determine a_0 plus a_m and b_m for wave numbers $m = 1, 2, 3, \ldots, J/2$. The shortest wavelength component in (13.1) has wavelength $L/m = 2L/J = 2\delta x$. Thus, the shortest wave that can be resolved by finite differencing has a wavelength twice that of the grid increment. Accurate representation of derivatives is only possible, however, for wavelengths much greater than $2\delta x$.

We now consider how the grid point variable Ψ_j can be used to construct a finite difference approximation to a differential equation. That is, we wish to represent derivatives such as $d\psi/dx$ and $d^2\psi/dx^2$ in terms of the finite difference fields. We first consider the Taylor series expansions about the point x_0:

$$\psi(x_0 + \delta x) = \psi(x_0) + \psi'(x_0)\,\delta x + \psi''(x_0)\,\frac{(\delta x)^2}{2} + \psi'''(x_0)\,\frac{(\delta x)^3}{6} + \mathrm{O}\left[(\delta x)^4\right] \tag{13.2}$$

$$\psi(x_0 - \delta x) = \psi(x_0) - \psi'(x_0)\,\delta x + \psi''(x_0)\,\frac{(\delta x)^2}{2} - \psi'''(x_0)\,\frac{(\delta x)^3}{6} + \mathrm{O}\left[(\delta x)^4\right] \tag{13.3}$$

where the primes indicate differentiation with respect to x and $\mathrm{O}[(\delta x)^4]$ means that terms with order of magnitude $(\delta x)^4$ or less are neglected.

Subtracting (13.3) from (13.2) and solving for $\psi\prime(x)$ give a finite difference expression for the first derivative of the form

$$\psi'(x_0) = [\psi(x_0 + \delta x) - \psi(x_0 - \delta x)] / (2\delta x) + \mathrm{O}\left[(\delta x)^2\right] \tag{13.4}$$

while adding the same two expressions and solving for $\psi''(x)$ give a finite difference expression for the second derivative of the form

$$\psi''(x_0) = [\psi(x_0 + \delta x) - 2\psi(x_0) + \psi(x_0 - \delta x)] / (\delta x)^2 + \mathrm{O}\left[(\delta x)^2\right] \tag{13.5}$$

Because the difference approximations in (13.4) and (13.5) involve points at equal distances on either side of x_0, they are called *centered differences*. These approximations neglect terms of order $(\delta x)^2$. We thus say that the truncation error is order $(\delta x)^2$. Higher accuracy can be obtained by decreasing the grid interval, but at the cost of increasing the density of grid points. Alternatively, it is possible to obtain higher order accuracy without decreasing the grid spacing by writing formulas analogous to (13.2) and (13.3) for the interval $2\delta x$ and using these together with (13.2) and (13.3) to eliminate error terms less than order $(\delta x)^4$. This approach, however, has the disadvantage of producing more complicated expressions and can be difficult to implement near boundary points.

13.3.2 Centered Differences: Explicit Time Differencing

As a prototype model we consider the linear one-dimensional advection equation

$$\partial q/\partial t + c\partial q/\partial x = 0 \tag{13.6}$$

with c a specified speed and $q(x, 0)$ a known initial condition. This equation can be approximated to second-order accuracy in x and t by the centered difference equation

$$[q(x, t+\delta t) - q(x, t-\delta t)]/(2\delta t) = -c[q(x+\delta x, t) - q(x-\delta x, t)]/(2\delta x) \tag{13.7}$$

The original differential equation (13.6) is thus replaced by a set of algebraic equations (13.7), which can be solved to determine solutions for a finite set of points that define a grid mesh in x and t (see Fig. 13.1). For notational convenience it is useful to identify points on the grid mesh by indices m, s. These are defined by letting $x = m\delta x$; $m = 0, 1, 2, 3, \ldots, M$, and $t = s\delta t$; $s = 0, 1, 2, 3, \ldots, S$, and writing $\hat{q}_{m,s} \equiv q(m\delta x, s\delta t)$. The difference equation (13.7) can then be expressed as

$$\hat{q}_{m,s+1} - \hat{q}_{m,s-1} = -\sigma\left(\hat{q}_{m+1,s} - \hat{q}_{m-1,s}\right) \tag{13.8}$$

where $\sigma \equiv c\delta t/\delta x$ is the *Courant number*. This form of time differencing is referred to as the *leapfrog* method, as the tendency at time step s is given by the difference in values computed for time steps $s+1$ and $s-1$ (i.e., by leaping across point s).

Leapfrog differencing cannot be used for the initial time $t = 0$ ($s = 0$), as $\hat{q}_{m,-1}$ is not known. For the first time step, an alternative method is required such as the forward difference approximation

$$\hat{q}_{m,1} - \hat{q}_{m,0} \equiv -(\sigma/2)\left(\hat{q}_{m+1,0} - \hat{q}_{m-1,0}\right) \tag{13.9}$$

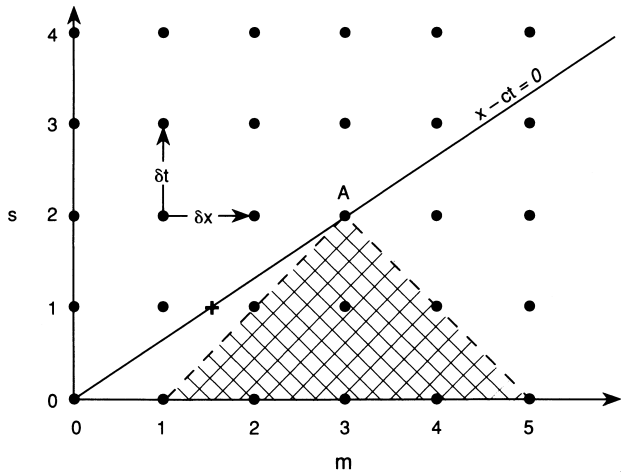


Fig. 13.1 Grid in $x - t$ space showing the domain of dependence of the explicit finite-difference solution at $m = 3$ and $s = 2$ for the one-dimensional linear advection equation. Solid circles show grid points. The sloping line is a characteristic curve along which $q(x, t) = q(0, 0)$, and the cross shows an interpolated point for the semi-Lagrangian differencing scheme. In this example the leapfrog scheme is unstable because the finite difference solution at point A does not depend on $q(0, 0)$.

The centered difference scheme for the advection equation (13.8) is an example of an *explicit* time differencing scheme. In an explicit difference scheme, the value of the predicted field at a given grid point for time step $s + 1$ depends only on the known values of the field at previous time steps. (In the case of the leapfrog method the fields at time step s and $s - 1$ are utilized.) The difference equation can then be solved simply by marching through the grid and obtaining the solution for each point in turn. The explicit leapfrog scheme is thus simple to solve. However, as shown in the next section, it has disadvantages in that it introduces a spurious "computational" mode and has stringent requirements on the maximum value permitted for the Courant number. There are a number of alternative explicit schemes that do not introduce a computational mode (e.g., see Problem 13.3), but these still require that the Courant number be sufficiently small.

13.3.3 Computational Stability

Experience shows that solutions to finite difference approximations such as (13.8) will not always resemble solutions to the original differential equations even when the finite difference increments in space and time are very small. It turns out that the character of the solutions depends critically on the *computational stability* of the difference equations. If the difference equations are not stable, numerical solutions will exhibit exponential growth even when, as in the case of the linear

advection equation, the original differential equation system has solutions whose amplitudes remain constant in time. In the example (13.8), stability considerations place stringent limitations on the value of the parameter σ, as we show later.

If the initial condition is specified as

$$q\,(x, 0) = \mathrm{Re}\left[\exp\,(ikx)\right] = \cos\,(kx)$$

The analytic solution of (13.6) that satisfies this initial condition is

$$q\,(x, t) = \mathrm{Re}\,\{\exp\,[ik\,(x - ct)]\} = \cos\,(kx - ct) \tag{13.10}$$

We now compare (13.10) with the solution of the finite difference system (13.8) and (13.9). In finite difference form the initial condition is

$$\hat{q}_{m,0} = \exp\,(ikm\delta x) = \exp\,(ipm) \tag{13.11}$$

where $p \equiv k\delta x$. Noting that the analytic solution (13.10) is separable in x and t, we consider solutions of (13.8) and (13.9) of the form

$$\hat{q}_{m,s} = B^s \exp\,(ipm) \tag{13.12}$$

where B is a complex constant. Substituting into (13.8) and dividing through by the common factor B^{s-1}, we obtain a quadratic equation in B:

$$B^2 + (2i\ sin\ \theta_P)\,B - 1 = 0 \tag{13.13}$$

where $\sin\theta_p \equiv \sigma \sin p$. Equation (13.3.3) has two roots, which may be expressed in the form

$$B_1 = \exp\left(-i\theta_p\right),\ \ B_2 = -\exp\left(+i\theta_p\right)$$

The general solution of the finite difference equation is thus

$$\hat{q}_{m,s} = \left[CB_1^s + DB_2^s\right]\exp\,(ipm) = Ce^{i\left(pm-\theta_p s\right)} + D\,(-1)^s\,e^{i\left(pm+\theta_p s\right)} \tag{13.14}$$

where C and D are constants to be determined by the initial conditions (13.11) and the first time step (13.9). The former gives $C + D = 1$, whereas the latter gives

$$Ce^{-i\theta_p} - De^{+i\theta_p} = 1 - i\sin\theta_p \tag{13.15}$$

Thus,

$$C = \frac{\left(1 + \cos\theta_p\right)}{2\cos\theta_p}, \quad D = -\frac{\left(1 - \cos\theta_p\right)}{2\cos\theta_p} \tag{13.16}$$

From inspection of (13.14) it is clear that the solution will remain finite for $s \to \infty$ provided that θ_p is real. If θ_p is imaginary, then one term in (13.14) will grow exponentially and the solution will be unbounded for $s \to \infty$. This sort of behavior is referred to as *computational instability*. Now because

$$\theta_p = \sin^{-1}(\sigma \sin p) \tag{13.17}$$

θ_p will be real only for $|\sigma \sin p| \leq 1$, which can only be valid for all waves (i.e., for all p) if $\sigma \leq 1$. Thus, computational stability of the difference equation (13.8), requires that

$$\sigma = c\delta t/\delta x \leq 1 \tag{13.18}$$

which is referred to as the Courant–Friedrichs–Levy (CFL) stability criterion.

The CFL criterion for this example states that for a given space increment δx the time step δt must be chosen so that the dependent field will be advected a distance less than one grid length per time step. The restriction on σ given by (13.18) can be understood physically by considering the character of the solution in the x, t plane as shown in Fig. 13.1. In the situation shown in Fig. 13.1, $\sigma = 1.5$. Examination of the centered difference system (13.8) shows that the numerical solution at point A in Fig. 13.1 depends only on grid points within the shaded region of the figure. However, because A lies on the *characteristic* line $x - ct = 0$, the true solution at point A depends only on the initial condition at $x = 0$ (i. e., a parcel which is initially at the origin will be advected to the point $3\delta x$ in time $2\delta t$). The point $x = 0$ is outside the *domain of influence* of the numerical solution. Hence, the numerical solution cannot possibly faithfully reproduce the solution to the original differential equation since, as shown in Fig. 13.1, the value at point A in the numerical solution has no dependence on the conditions at $x = 0$. Only when the CFL criterion is satisfied will the domain of influence of the numerical solution include the characteristic lines of the analytic solution.

Although the CFL condition (13.18) guarantees stability of the centered difference approximation to the one-dimensional advection equation, in general the CFL criterion is only a necessary and not sufficient condition for computational stability. Other forms of finite differencing the one-dimensional advection equation may lead to more stringent limits on σ than given in (13.18).

The existence of computational instability is one of the prime motivations for using filtered equations. In the quasi-geostrophic system, no gravity or sound waves occur. Thus, the speed c in (13.18) is just the maximum wind speed. Typically, $c < 100\,\mathrm{m\,s^{-1}}$ so that for a grid interval of 200 km a time increment of over 30 min is permissible. However, in the nonhydrostatic equations commonly used in cloud-resolving models, the solution would have characteristics corresponding to acoustic modes, and to assure that the domain of influence included such characteristics c would need to be set equal to the speed of sound, which is the fastest

wave described in that set of equations. In that case $c \approx 300\text{ms}^{-1}$ and for a 1-km vertical grid interval a time step of only a few seconds would be permitted.

13.3.4 Implicit Time Differencing

The spurious computational mode introduced by the leapfrog time differencing scheme does not occur in a number of other explicit schemes, such as the Euler backward scheme discussed in Problem 13.3. Such schemes do, however, retain the time step limitation imposed by the CFL condition. This restriction and the computational mode are eliminated by utilization of an alternative finite differencing scheme called the *trapezoidal implicit scheme*. For the linear advection equation (13.6) this scheme can be written in the form

$$\frac{\left(\hat{q}_{m,s+1} - \hat{q}_{m,s}\right)}{\delta t} = -\frac{c}{2}\left[\frac{\left(\hat{q}_{m+1,s+1} - \hat{q}_{m-1,s+1}\right)}{2\delta x} + \frac{\left(\hat{q}_{m+1,s} - \hat{q}_{m-1,s}\right)}{2\delta x}\right] \tag{13.19}$$

Substituting the trial solution (13.12) into (13.19) yields

$$B^{s+1} = \left[\frac{1 - i\,(\sigma/2)\sin p}{1 + i\,(\sigma/2)\sin p}\right] B^{s} \tag{13.20}$$

where as before $\sigma = c\delta t/\delta x$ and $p = k\delta x$. Defining

$$\tan\theta_p \equiv (\sigma/2)\sin p \tag{13.21}$$

and eliminating the common term B^s in (13.20), it can be shown that

$$B = \left(\frac{1 - i\tan\theta_p}{1 + i\tan\theta_p}\right) = \exp\left(-2i\theta_p\right) \tag{13.22}$$

so that the solution may be expressed simply as

$$\hat{q}_{m,s} = A\exp\left[ik\left(m\delta x - 2\theta_p s/k\right)\right] \tag{13.23}$$

Equation (13.19) involves only two time levels. Hence, unlike (13.14) the solution yields only a single mode, which has phase speed $c' = 2\theta_p/(k\delta t)$. According to (13.21), θ_p remains real for all values of δt. [This should be contrasted to the situation for the explicit scheme given by (13.17).] Thus, the implicit scheme is absolutely stable. The truncation errors, however, can become large if θ_p is not kept sufficiently small (see Problem 13.9). A disadvantage of the implicit scheme is that the integration cannot proceed by marching through the grid, as in the explicit

case. In (13.19) there are terms involving the $s+1$ time level on both sides of the equal sign, and these involve the values at a total of three grid points. Thus, the system (13.19) must be solved simultaneously for all points on the grid. If the grid is large, this may involve inverting a very large matrix, and so is computationally intensive. Furthermore, it is usually not feasible to utilize the implicit approach for nonlinear terms, such as the advection terms in the momentum equations. Semi-implicit schemes in which the linear terms are treated implicitly and the nonlinear terms explicitly have, however, become popular in modern forecasting models. These are discussed briefly in Section 13.6.

13.3.5 The Semi-Lagrangian Integration Method

The differencing schemes discussed above are Eulerian schemes in which the time integration is carried out by computing the tendencies of the predicted fields at a set of grid points fixed in space. Although it would be possible in theory to carry out predictions in a Lagrangian framework by following a set of marked fluid parcels, in practice this is not a viable alternative, as shear and stretching deformations tend to concentrate marked parcels in a few regions so that it is difficult to maintain uniform resolution over the forecast region. It is possible to take advantage of the conservative properties of Lagrangian schemes, while maintaining uniform resolution, by employing a semi-Lagrangian technique. This approach permits relatively long time steps while retaining numerical stability and high accuracy. The semi-Lagrangian method can be illustrated in a very simple fashion with the one-dimensional advection equation (13.6). According to this equation the field q is conserved following the zonal flow at speed c. Thus, for any grid point, $x_m = m\delta x$, and time $t_s = s\delta t$:

$$q\left(x_m, t_s + \delta t\right) = q\left(\tilde{x}_m^s, t_s\right) \tag{13.24}$$

Here, $\tilde{x}_m^s$ is the location at time t_s for the air parcel that is located at point x_m at time $t_s + \delta t$. This position in general does not lie on a grid point (see cross marked on Fig. 13.1), so that evaluation of the right-hand side of (13.24) requires interpolating from the grid point values at time t. For $c > 0$ the position $\tilde{x}_m^s$ lies between the grid points x_{m-p} and x_{m-p-1} where p is the integer part of the expression $c\delta t/\delta x$ (a measure of the number of grid points traversed in a timestep). If linear interpolation is used

$$q\left(\tilde{x}_m^s, t_s\right) = \alpha q\left(x_{m-p-1}, t_s\right) + (1-\alpha)\, q\left(x_{m-p}, t_s\right)$$

where $\alpha = \left(x_{m-p} - \tilde{x}_m^s\right)/\delta x$. Thus, in Fig. 13.1 $p = 1$, and to predict q at point A data are interpolated between the points $m = 1$ and $m = 2$ to the point shown by the cross.

In an actual prediction model the velocity field is predicted rather than known as in this simple example. Thus, for a two-dimensional field

$$q\,(x, y, t + \delta t) = q\,(x - u\delta t, y - v\delta t, t) \tag{13.25}$$

where the velocity components at time t can be used to estimate the fields at $t + \delta t$; once these are obtained they can be used to provide more accurate approximations to the velocities on the right in (13.25). The right side in (13.25) is again estimated by interpolation, which now must be carried out in two dimensions.

As shown in Fig. 13.1, the semi-Lagrangian scheme guarantees that the domain of influence in the numerical solution corresponds to that of the physical problem. Thus, the scheme is computational stability for time steps much longer than possible with an explicit Eulerian scheme. The semi-Lagrangian scheme also preserves the values of conservative properties quite accurately and is particularly useful for accurately advecting trace constituents such as water vapor.

13.3.6 Truncation Error

To be useful it is necessary not only that a numerical solution be stable, but that it also provide an accurate approximation to the true solution. The difference between the numerical solution to a finite difference equation and the solution to the corresponding differential equation is called the *discretization error*. If this error approaches zero with δt and δx, the solution is called *convergent*. The difference between a differential equation and the finite difference analog to it is referred to as a *truncation error* because it arises from truncating the Taylor series approximation to the derivatives. If this error approaches zero as δt and δx go to zero, the scheme is called *consistent*. According to the Lax equivalence theorem,[1] if the finite difference formulation satisfies the consistency condition, then stability is the necessary and sufficient condition for convergence. Thus, if a finite difference approximation is consistent and stable, one can be certain that the discretization error will decrease as the difference intervals are decreased, even if it is not possible to determine the error exactly.

Because numerical solutions are as a rule sought only when analytic solutions are unavailable, it is usually not possible to determine the accuracy of a solution directly. For the linear advection equation with constant advection speed considered in Section (13.3.3), it is possible, however, to compare the solutions of the finite difference equation (13.8) and the original differential equation (13.6). We can then use this example to investigate the accuracy of the difference method introduced above.

From the above discussion we already can conclude that the magnitude of the truncation error in the present case will be of order δx^2 and δt^2. It is possible

[1] See Richtmyer and Morton (1967).

to obtain more precise information on accuracy from examination of the solution (13.14). Note that for $\theta_p \to 0, C \to 1,$ and $D \to 0$. The part of the solution proportional to C is the *physical mode*. The part proportional to D is called the *computational mode* because it has no counterpart in the analytic solution to the original differential equation. It arises because centered time differencing has turned a differential equation that is first order in time into a second-order finite difference equation. The accuracy of the finite difference solution depends not only on the smallness of D and the closeness of C to unity, but on the match between the phase speed of the physical mode and the phase speed in the analytic solution. The phase of the physical mode is given in (13.14) by

$$pm - \theta_p s = (p/\delta x)\left(m\delta x - \theta_p s\delta x/p\right) = k\left(x - c't\right)$$

where $c' = \theta_p \delta x/(p\delta t)$ is the phase speed of the physical mode. Its ratio to the true phase speed is

$$c'/c = \theta_p \delta x/\left(pc\delta t\right) = \sin^{-1}\left(\sigma \sin p\right)/\left(\sigma p\right)$$

so that $c'/c \to 1$ as $\sigma p \to 0$. The dependence of c'/c and $|D|/|C|$ on wavelength is shown in Table 13.1 for the particular case where $\sigma = 0.75$.

It is clear from Table 13.1 that phase speed and amplitude errors both increase as the wavelength decreases. Short waves move slower than long waves in the finite difference solution, even though in the original equation all waves move at speed c. This dependence of phase speed on wavelength in the difference solution is called *numerical dispersion*. It is a serious problem for numerical modeling of any advected field that has sharp gradients (and hence large-amplitude short-wave components).

Short waves also suffer from having significant amplitude in the computational mode. This mode, which has no counterpart in the solution to the original differential equation, propagates opposite to the direction of the physical mode and changes sign from one time step to the next. This behavior makes it easy to recognize when the computational mode has significant amplitude.

Table 13.1 *Phase and Amplitude Accuracy of Centered Difference Solution of the Advection Equation as a Function of Resolution for* $\sigma = 0.75$

$L/\delta x$	p	θ_p	c'/c	$\lvert D\rvert/\lvert C\rvert$
2	π	π	—	∞
4	$\pi/2$	0.848	0.720	0.204
8	$\pi/4$	0.559	0.949	0.082
16	$\pi/8$	0.291	0.988	0.021
32	$\pi/16$	0.147	0.997	0.005

13.4 THE BAROTROPIC VORTICITY EQUATION IN FINITE DIFFERENCES

The simplest example of a dynamical forecast model is the barotropic vorticity equation (11.14), which for a Cartesian β plane can be written in the form

$$\frac{\partial \zeta}{\partial t} = -F(x, y, t) \tag{13.26}$$

where

$$F(x, y, t) = \mathbf{V}_\psi \bullet \nabla(\zeta + f) = \frac{\partial}{\partial x}\left(u_\psi \zeta\right) + \frac{\partial}{\partial y}\left(v_\psi \zeta\right) + \beta v_\psi \tag{13.27}$$

and $u_\psi = -\partial\psi/\partial y$, $v_\psi = \partial\psi/\partial x$, and $\zeta = \nabla^2\psi$. We have here used the fact that the horizontal velocity is nondivergent ($\partial u_\psi/\partial x + \partial v_\psi/\partial y = 0$) to write the advection term in flux form. The advection of absolute vorticity $F(x, y, t)$ may be calculated provided that we know the field of $\psi(x, y, t)$. Equation (13.26) can then be integrated forward in time to yield a prediction for ζ. It is then necessary to solve the Poisson equation $\zeta = \nabla^2\psi$ to predict the streamfunction.

A straightforward solution method is the leapfrog scheme discussed in Section 13.3.2. This requires writing (13.27) in finite difference form. Suppose that the horizontal x,y space is divided into a grid of $(M + 1) \times (N + 1)$ points separated by distance increments δx and δy. Then we can write the coordinate position of a given grid point as $x_m = m\delta x$, $y_n = n\delta y$, where $m = 0, 1, 2, \ldots, M$ and $n = 0, 1, 2, \ldots, N$. Thus any point on the grid is uniquely identified by the indices (m, n). A portion of such a grid space is shown in Fig. 13.2.

Centered difference formulas of the type (13.4) can then be used to approximate derivatives in the expression $F(x, y, t)$. For example, if we assume that $\delta x = \delta y \equiv d$

$$\begin{aligned} u_\psi &\approx u_{m,n} = -\left(\psi_{m,n+1} - \psi_{m,n-1}\right)/2d \\ v_\psi &\approx v_{m,n} = +\left(\psi_{m+1,n} - \psi_{m-1,n}\right)/2d \end{aligned} \tag{13.28}$$

Similarly, with the aid of (13.5) we find that the horizontal Laplacian can be approximated as

$$\nabla^2\psi \approx \left(\psi_{m+1,n} + \psi_{m-1,n} + \psi_{m,n+1} + \psi_{m,n-1} - 4\psi_{m,n}\right)/d^2 = \zeta_{m,n} \tag{13.29}$$

The finite difference form of the Laplacian is proportional to the difference between the value of the function at the central point and the average value at the four surrounding grid points. If there are $(M - 1) \times (N - 1)$ interior grid points, then (13.29) yields a set of $(M - 1) \times (N - 1)$ simultaneous equations, which together with suitable boundary conditions determine $\psi_{m,n}$ for a

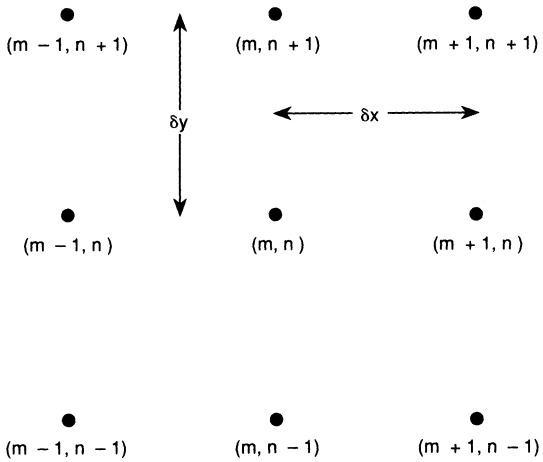


Fig. 13.2 A portion of a two-dimensional (x, y) grid mesh for solution of the barotropic vorticity equation.

given array $\zeta_{m,n}$. This set can be solved by standard methods of matrix inversion.

Before expressing the advection term $F(x, y, t)$ in finite difference form, it is worth noting that if ψ is taken to be constant on the northern and southern boundaries of the β-plane channel, it is easily shown by integration over the area of the channel that the average value of F is zero. This implies that the mean vorticity is conserved for the channel. It is also possible, with a little more algebra, to show that the mean kinetic energy and the mean square vorticity (called the *enstrophy*) are conserved.

For accuracy of long-term integrations, it is desirable that any finite difference approximation to F satisfy the same conservation constraints as the original differential form; otherwise the finite difference solution would not be conservative. The mean vorticity, for example, might then drift systematically in time purely due to the nature of the finite difference solution. Finite difference schemes that simultaneously conserve vorticity, kinetic energy, and enstrophy have been designed. They are, however, rather complex. For our purposes it is sufficient to note that by writing the advection in the flux form (13.27) and using centered space differences we can conserve both mean vorticity and mean kinetic energy:

$$F_{m,n} = \frac{1}{2d}\left[\left(u_{m+1,n}\zeta_{m+1,n} - u_{m-1,n}\zeta_{m-1,n}\right) + \left(v_{m,n+1}\zeta_{m,n+1} - v_{m,n-1}\zeta_{m,n-1}\right)\right] + \beta v_{m,n} \tag{13.30}$$

It is verified readily that if ψ satisfies periodic boundary conditions there is a cancellation of terms when (13.30) is summed over the domain. Thus,

$$\sum_{m=1}^{M}\sum_{n=1}^{N} F_{m,n} = 0 \tag{13.31}$$

Therefore, mean vorticity is conserved (except for errors introduced by time differencing) when (13.30) is used as the finite difference form of the advection term. This form also conserves mean kinetic energy (see Problem 2). Enstrophy is not conserved in this difference formulation, and it is conventional to add a small diffusion term in order to control any numerically generated increase in enstrophy.

The procedure for preparing a numerical forecast with the barotropic vorticity equation can now be summarized as follows:

(1) The observed geopotential field at the initial time is used to compute the initial stream function $\psi_{m,n}(t = 0)$ at all grid points.

(2) $F_{m,n}$ is evaluated at all grid points.

(3) $\zeta_{m,n}(t + \delta t)$ is determined using centered differencing except at the first time step when a forward difference must be used.

(4) The simultaneous set (13.29) is solved for $\psi_{m,n}(t + \delta t)$.

(5) The predicted array of $\psi_{m,n}$ is used as data, and steps 2–4 are repeated until the desired forecast period is reached. For example, a 24-h forecast with 30-min time increments would require 48 cycles.

13.5 THE SPECTRAL METHOD

In the finite difference method, the dependent variables are specified on a set of grid points in space and time, and derivatives are approximated using finite differences. An alternative approach, referred to as the spectral method, involves representing the spatial variations of the dependent variables in terms of finite series of orthogonal functions called *basis functions*. For the Cartesian geometry of a midlatitude β-plane channel, the appropriate set of basis functions is a double Fourier series in x and y. For the spherical earth, however, the appropriate basis functions are the spherical harmonics.

A finite difference approximation is *local* in the sense that the finite difference variable $\Psi_{m,n}$ represents the value of $\psi(x, y)$ at a particular point in space, and the finite difference equations determine the evolution of the $\Psi_{m,n}$ for all grid points. The spectral approach, however, is based on *global* functions, that is, the individual components of the appropriate series of basis functions. In the case of Cartesian geometry, for example, these components determine the amplitudes and phases of the sinusoidal waves that when summed determine the spatial distribution of the

dependent variable. The solution proceeds by determining the evolution of a finite number of Fourier coefficients. Because the distribution in wave number space of the Fourier coefficients for a given function is referred to as its *spectrum*, it is appropriate to call this approach the spectral method.

At low resolution the spectral method is generally more accurate than the grid point method, partly because for linear advection the numerical dispersion discussed in Section 13.3.4 can be severe in a grid point model, but does not occur in a properly formulated spectral model. For the range of resolutions used commonly in forecast models the two approaches are comparable in accuracy, and each has its advocates.

13.5.1 The Barotropic Vorticity Equation in Spherical Coordinates

The spectral method is particularly advantageous for solution of the vorticity equation. When the proper set of basis functions is chosen, it is trivial to solve the Poisson equation for the streamfunction. This property of the spectral method not only saves computer time, but eliminates the truncation error associated with finite differencing the Laplacian operator.

In practice the spectral method is applied most frequently to global models. This requires use of spherical harmonics, which are more complicated than Fourier series. In order to keep the discussion as simple as possible, it is again useful to consider the barotropic vorticity equation as a prototype forecast model in order to illustrate the spectral method on the sphere.

The barotropic vorticity equation in spherical coordinates may be expressed as

$$\frac{D}{Dt}(\zeta + 2\Omega \sin\phi) = 0 \tag{13.32}$$

where as usual $\zeta = \nabla^2\psi$, with ψ a streamfunction, and

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \frac{u}{a\cos\phi}\frac{\partial}{\partial\lambda} + \frac{v}{a}\frac{\partial}{\partial\phi} \tag{13.33}$$

It turns out to be convenient to use $\mu \equiv \sin\phi$ as the latitudinal coordinate, in which case the continuity equation can be written

$$\frac{1}{a}\frac{\partial}{\partial\lambda}\left(\frac{u}{\cos\phi}\right) + \frac{1}{a}\frac{\partial}{\partial\mu}(v\cos\phi) = 0 \tag{13.34}$$

so that the streamfunction is related to the zonal and meridional velocities according to

$$\frac{u}{\cos\phi} = -\frac{1}{a}\frac{\partial\psi}{\partial\mu}; \quad v\cos\phi = \frac{1}{a}\frac{\partial\psi}{\partial\lambda} \tag{13.35}$$

The vorticity equation can then be expressed as

$$\frac{\partial \nabla^2 \psi}{\partial t} = \frac{1}{a^2}\left[\frac{\partial \psi}{\partial \mu}\frac{\partial \nabla^2 \psi}{\partial \lambda} - \frac{\partial \psi}{\partial \lambda}\frac{\partial \nabla^2 \psi}{\partial \mu}\right] - \frac{2\Omega}{a^2}\frac{\partial \psi}{\partial \lambda} \tag{13.36}$$

where

$$\nabla^2 \psi = \frac{1}{a^2}\left\{\frac{\partial}{\partial \mu}\left[\left(1-\mu^2\right)\frac{\partial \psi}{\partial \mu}\right] + \frac{1}{1-\mu^2}\frac{\partial^2 \psi}{\partial \lambda^2}\right\} \tag{13.37}$$

The appropriate orthogonal basis functions are the *spherical harmonics*, which are defined as

$$Y_\gamma(\mu, \lambda) \equiv P_\gamma(\mu)\, e^{im\lambda} \tag{13.38}$$

where $\gamma \equiv (n, m)$ is a vector containing the integer indices for the spherical harmonics. These are given by $m = 0, \pm 1, \pm 2, \pm 3, \ldots, n = 1, 2, 3, \ldots$, where it is required that $|m| \leq n$. Here, P_γ designates an associated Legendre function of the first kind of degree n. From (13.38) it is clear that m designates the zonal wave number. It can be shown[2] that $n - |m|$ designates the number of nodes of P_γ in the interval $-1 < \mu < 1$ (i.e., between the poles), and thus measures the meridional scale of the spherical harmonic. The structures of a few spherical harmonics are shown in Fig. 13.3.

An important property of the spherical harmonics is that they satisfy the relationship

$$\nabla^2 Y_\gamma = -\frac{n(n+1)}{a^2} Y_\gamma \tag{13.39}$$

so that the Laplacian of a spherical harmonic is proportional to the function itself, which implies that the vorticity associated with a particular spherical harmonic component is simply proportional to the streamfunction for the same component.

In the spectral method on the sphere, the streamfunction is expanded in a finite series of spherical harmonics by letting

$$\psi(\lambda, \mu, t) = \sum_\gamma \psi_\gamma(t)\, Y_\gamma(\mu, \lambda) \tag{13.40}$$

where ψ_γ is the complex amplitude for the Y_γ spherical harmonic and the summation is over both n and m. The individual spherical harmonic coefficients ψ_γ are related to the streamfunction $\psi(\lambda, \mu)$ through the inverse transform

$$\psi_\gamma(t) = \frac{1}{4\pi}\int_S Y_\gamma^* \psi(\lambda, \mu, t)\, dS \tag{13.41}$$

where $dS = d\mu d\lambda$ and Y_γ^* designates the complex conjugate of Y_γ.

[2] See Washington and Parkinson (1986) for a discussion of the properties of the Legendre function.

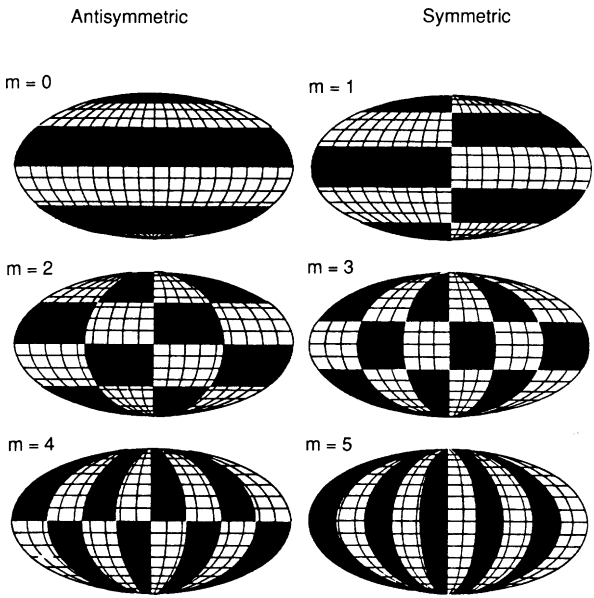


Fig. 13.3 Patterns of positive and negative regions for the spherical harmonic functions with $n = 5$ and $m = 0, 1, 2, 3, 4, 5$. (After Washington and Parkinson, 1986, adapted from Baer, 1972.)

13.5.2 Rossby–Haurwitz Waves

Before considering numerical solution of the barotropic vorticity equation, it is worth noting that an exact analytic solution of the nonlinear equation can be obtained in the special case where the streamfunction is equal to a single spherical harmonic mode. Thus, we let

$$\psi(\lambda, \mu, t) = \psi_\gamma(t)\, e^{im\lambda} P_\gamma(\mu) \tag{13.42}$$

Substituting from (13.42) into (13.36) and applying (13.39), we find that the nonlinear advection term is identically zero so that the amplitude coefficient satisfies the ordinary linear equation

$$-n(n+1)\, d\psi_\gamma/dt = -2\Omega i m \psi_\gamma \tag{13.43}$$

which has the solution $\psi_\gamma(t) = \psi_\gamma(0)\exp(i\nu_\gamma \mathrm{t})$, where

$$\nu_\gamma = 2\Omega m/[n(n+1)] \tag{13.44}$$

is the dispersion relationship for *Rossby–Haurwitz waves*, which is the name given to planetary waves on a sphere. [This should be compared to (7.91), which is the

equivalent expression for a midlatitude β-plane.] Because the horizontal scale of a spherical harmonic mode is proportional to n^{-1}, (13.44) shows that a single mode propagates westward on a sphere at a speed that is approximately proportional to the square of the horizontal scale. This solution also suggests why for some problems the spectral method is superior to the finite difference method at coarse resolution. A model containing even a single Fourier component can represent a realistic meteorological field (the Rossby wave), while many grid points are required for an equivalent representation in finite differences.

13.5.3 The Spectral Transform Method

When many spherical harmonic modes are present, the solution of (13.36) by a purely spectral method requires evaluation of the nonlinear interactions among various modes due to the advection term. It turns out that the number of interaction terms increases as the square of the number of modes retained in the series (13.40) so that this approach becomes computationally inefficient for models with the sort of spatial resolution required for prediction of synoptic-scale weather disturbances. The spectral transform method overcomes this problem by transforming between spherical harmonic wave number space and a latitude–longitude grid at every time step and carrying out the multiplications in the advection term in grid space so that it is never necessary to compute products of spectral functions.

To illustrate this method it is useful to rewrite the barotropic vorticity equation in the form

$$\frac{\partial \nabla^2 \psi}{\partial t} = -\frac{1}{a^2}\left[2\Omega \frac{\partial \psi}{\partial \lambda} + A(\lambda, \mu)\right] \tag{13.45}$$

where

$$A(\lambda, \mu) \equiv \left[-\frac{\partial \psi}{\partial \mu}\frac{\partial \nabla^2 \psi}{\partial \lambda} + \frac{\partial \psi}{\partial \lambda}\frac{\partial \nabla^2 \psi}{\partial \mu}\right] \tag{13.46}$$

Substituting from (13.40) into (13.45) then yields for the spherical harmonic coefficients

$$d\psi_\gamma/dt = i\nu_\gamma \psi_\gamma + A_\gamma\,[n\,(n+1)]^{-1} \tag{13.47}$$

where A_γ is the γ component of the transform of $A(\lambda, \mu)$:

$$A_\gamma = \frac{1}{4\pi}\int_0^{2\pi}\int_{-1}^{+1} A(\lambda, \mu) Y_\gamma^* d\lambda d\mu \tag{13.48}$$

The transform method utilizes the fact that if the sum $\gamma = (n, m)$ is taken over a finite number of modes, the integral in the transform (13.48) can be evaluated exactly by numerical quadrature [i.e., by summing appropriately weighted values of $A(\lambda, \mu)$ evaluated at the grid points (λ_j, μ_k) of a latitude–longitude grid mesh.]

Computation of the distribution of $A(\lambda, \mu)$ for all grid points can be carried out without need to introduce finite differences for derivatives by noting that we can express the advection term in the form

$$A\left(\lambda_j, \mu_k\right) = \left(1 - \mu^2\right)^{-1} \left[F_1 F_2 + F_3 F_4\right] \tag{13.49}$$

where

$$F_1 = -\left(1 - \mu^2\right) \partial\psi/\partial\mu, \quad F_2 = \partial\nabla^2\psi/\partial\lambda,$$
$$F_3 = \left(1 - \mu^2\right) \partial\nabla^2\psi/\partial\mu, \quad F_4 = \partial\psi/\partial\lambda$$

The quantities $F_1 - F_4$ can be computed exactly for each grid point using the spectral coefficients ψ_γ and the known differential properties of the spherical harmonics. For example,

$$F_4 = \partial\psi/\partial\lambda = \sum_\gamma im\psi_\gamma Y_\gamma\left(\lambda_j, \mu_k\right)$$

Once these terms have been computed for all grid points, $A(\lambda, \mu)$ can be computed at the grid points by forming the products $F_1 F_2$ and $F_3 F_4$; no finite difference approximations to the derivatives are required in this procedure. The transform (13.48) is then evaluated by numerical quadrature to compute the spherical harmonic components A_γ. Finally, (13.47) can be stepped ahead by a time increment δt in order to obtain new estimates of the spherical harmonic components of the streamfunction. The whole process is then repeated until the desired forecast period is reached. The steps in forecasting with the barotropic vorticity equation using the spectral transform method are summarized in schematic form in Fig. 13.4.

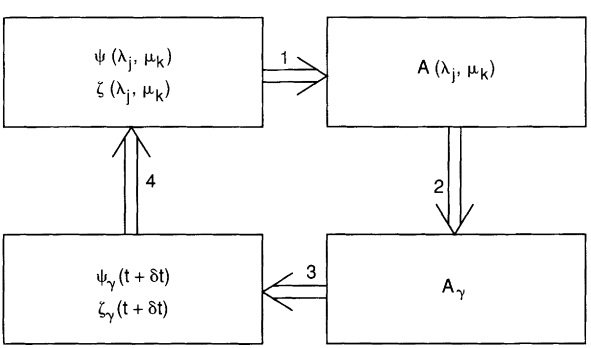


Fig. 13.4 Steps in the prediction cycle for the spectral transform method of solution of the barotropic vorticity equation.

13.6 PRIMITIVE EQUATION MODELS

Modern numerical forecast models are based on a formulation of the dynamical equations, referred to as the *primitive equations*, which is essentially the formulation proposed by Richardson. The primitive equations differ from the complete momentum equations (2.19)–(2.21) in that the vertical momentum equation is replaced by the hydrostatic approximation, and the small terms in the horizontal momentum equations given in columns C and D in Table 2.1 are neglected. In most models some version of the σ-coordinate system introduced in Section 10.3.1 is used, and the vertical dependence is represented by dividing the atmosphere into a number of levels and utilizing finite difference expressions for vertical derivatives. Both finite differencing and the spectral method have been used for horizontal discretization in operational primitive equation forecast models. Excellent examples of such models are the grid point and spectral models developed at the European Centre for Medium-Range Weather Forecasts (ECMWF). The numerical forecasting system at ECMWF, and several other centers, includes both a deterministic model and an ensemble prediction system. The deterministic model is typically run at very high resolution with the best estimate of the initial conditions. The ensemble prediction system consists of a number of parallel runs of a reduced resolution model in which the initializations (and sometimes the model physics) are slightly perturbed. The spread of forecasts given by members of the ensemble can then be used to assess confidence in the forecasts.

13.6.1 The Ecmwf Grid Point Model

The earliest version of the ECMWF operational model was a global grid point model with second-order horizontal differencing on a uniform grid with intervals of 1.875° in both latitude and longitude. For computational efficiency and noise control the grid was staggered in space, as shown in Fig. 13.5, so that not all variables were carried at the same points in λ and ϕ. Variables were also staggered vertically, as shown in Fig. 13.6, and there were a total of 15 unequally spaced σ layers, arranged to provide the finest resolution near the ground, the coarsest resolution in the stratosphere, and an uppermost layer near the 25-hPa level.

The difference equations on the staggered grid can be expressed most conveniently by defining a special operator notation as follows:

$$\begin{aligned} \overline{u}^{\lambda} &= [u(\lambda + \delta\lambda/2) + u(\lambda - \delta\lambda/2)]/2 \\ \delta_{\lambda} u &= [u(\lambda + \delta\lambda/2) - u(\lambda - \delta\lambda/2)]/\delta\lambda \end{aligned} \tag{13.50}$$

These represent, respectively, the averaging and differencing of adjacent grid point values of the field u on a latitude circle. Similar expressions can be defined for the

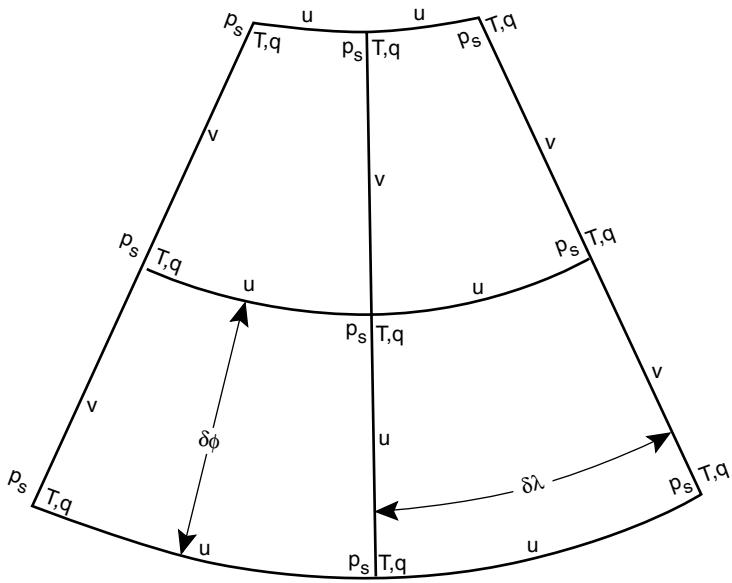


Fig. 13.5 Horizontal staggered grid arrangement in the ECMWF grid point model.

meridional and vertical directions. The continuity equation, for example, can then be expressed in the form

$$\frac{\partial p_s}{\partial t} + \frac{1}{a\cos\phi}\left[\delta_\lambda U + \delta_\phi\left(V\cos\phi\right)\right] + \delta_\sigma\left(p_s\dot{\sigma}\right) = 0 \tag{13.51}$$

where $U = \overline{p_s}^{\lambda} u$ and $V = \overline{p_s}^{\phi} v$ and are horizontal mass fluxes (see Section 10.3.1). Due to the arrangement of variables on the staggered grid of Fig. 13.5, the horizontal mass flux divergence evaluated by taking differences over the increments $\delta\lambda, \delta\phi$ in (13.51) is centered at the point where p_s is defined. This provides an improvement on the accuracy possible with a nonstaggered grid in which these differences would need to be evaluated over grid distances of $2\delta\lambda, 2\delta\phi$.

Grid staggering provides a similar gain in accuracy in computation of the pressure gradient terms in the momentum equations. Thus,

$$\frac{\partial u}{\partial t} = \frac{-1}{a\cos\phi}\left[\delta_\lambda\overline{\Phi}^{\sigma} + R\overline{T}^{\lambda}\delta_\lambda\left(\ln p_s\right) + \cdots\right] \tag{13.52}$$

$$\frac{\partial v}{\partial t} = -\frac{1}{a}\left[\delta_\phi\overline{\Phi}^{\sigma} + R\overline{T}^{\phi}\delta_\phi\left(\ln p_s\right) + \cdots\right] \tag{13.53}$$

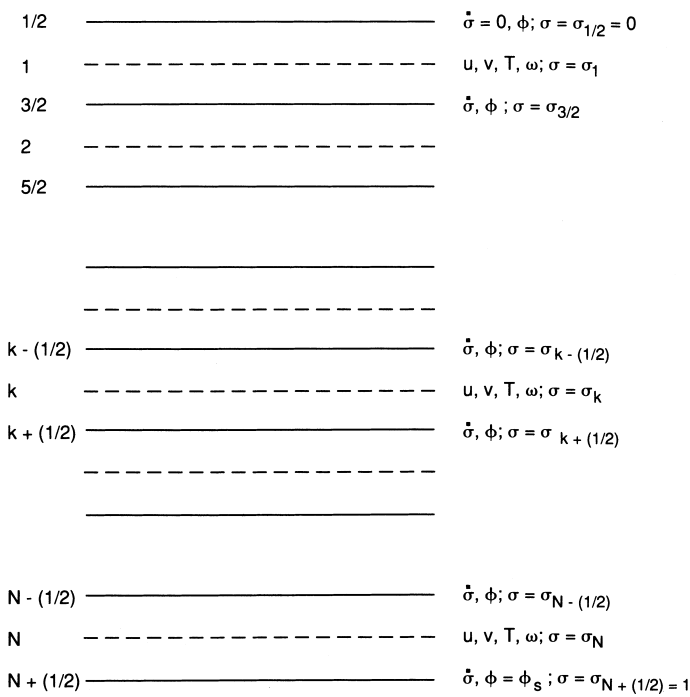


Fig. 13.6 Vertical disposition of variables in the ECMWF grid point model. (After Simmons et al., 1989.)

In this case, however, the vertical average of Φ is required since, as indicated in Fig. 13.6, Φ is staggered vertically with respect to u, v, T. The complete finite difference equations for this model, including the required special treatment of polar regions, are quite complex and will not be given here. They are discussed in Simmons et al. (1989).

The vertical finite differences in this model are defined so that the model can be run in either the σ coordinates discussed in Section 10.3.1 or in a *hybrid coordinate* system in which the coordinate varies from pure sigma coordinates in the troposphere to isobaric coordinates in the stratosphere.

13.6.2 Spectral Models

ECMWF and most other operational forecast centers utilize spectral models as their primary global forecast models. Grid point models are generally employed for fine scale limited area models and may become more common for very high resolution global models because the number of operations increases linearly with the number of grid points, but increases as $n \log n$ in a spectral model (where n is the number

of modes). Operational spectral models employ a primitive equation version of the spectral transform method described in Section 13.5. In this method the values of all meteorological fields are available in both spectral and grid point domains at each time step. Vorticity and divergence are employed as predictive variables rather than u and v. Physical computations involving such processes as radiative heating and cooling, condensation and precipitation, and convective overturning are calculated in physical space on the grid mesh, whereas differential dynamical quantities such as pressure gradients and velocity gradients are evaluated exactly in spectral space. This combination preserves the simplicity of the grid point representation for physical processes that are "local" in nature, while retaining the superior accuracy of the spectral method for dynamical calculations.

As of early 2003 the ECMWF deterministic global model employed a series of spherical harmonics truncated at $M = N = 511$, where (N, M) are the maximum retained values of (n, m). This truncation is referred to as a *triangular truncation* (T511) because on a plot of m versus n the retained modes occupy a triangular area. Another popular spectral truncation is *rhomboidal truncation*, which has $N = |m| + M$. Both of these truncations are shown schematically in Fig. 13.7. In triangular truncation, the horizontal resolution in the zonal and meridional directions is nearly equal. In rhomboidal truncation, however, the latitudinal resolution is the same for every zonal wave number. Rhomboidal truncation has some advantages for low-resolution models, but at high resolution the triangular truncation appears

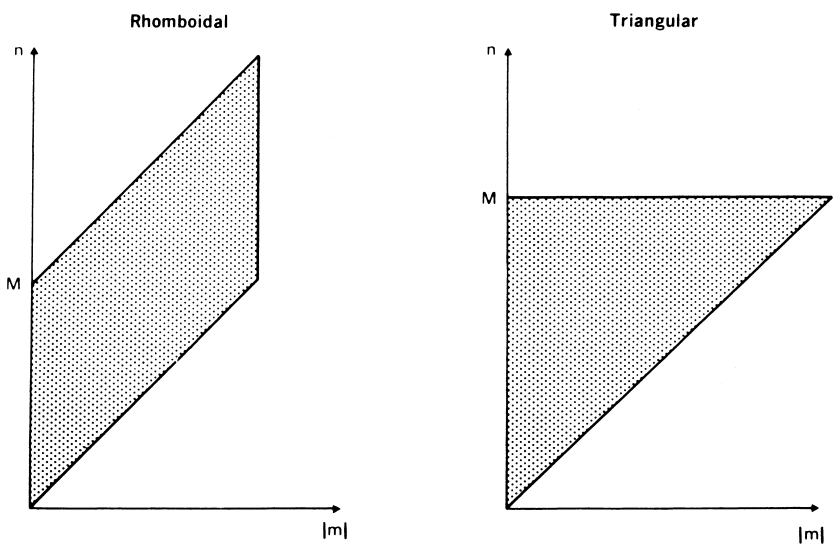


Fig. 13.7 Regions of wave number space (n, m) in which spectral components are retained for the rhomboidal truncation (left) and triangular truncation (right). (After Simmons and Bengtsson, 1984.)

to be superior. The deterministic ECMWF T511 spectral model has an equivalent horizontal grid spacing of about 40 km. The ECMWF ensemble prediction system in 2003 was based on a T255 spectral model (horizontal grid spacing of about 80 km). As of 2003 the deterministic ECMWF model had 60 vertical levels extending from the surface to about 64 km, while the ensemble prediction system had 40 levels extending from the surface to about 30 km.

Due to the high resolutions of the ECMWF T511 and T255 models, a time integration scheme based on the centered difference leapfrog method or other explicit scheme would require very small time steps in order to satisfy the CFL conditions (13.18) for the fastest propagating gravity wave modes and for the advection terms in areas of high wind speed. To avoid the computational overhead of such a model, the time integration is carried out using a semi-implicit technique in which terms associated with gravity wave propagation are treated implicitly; and a semi-Lagrangian technique is employed for computation of the nonlinear advection terms. A time step of 15 min or more can then be used.

13.6.3 Physical Parameterizations

The physical processes included in modern operational forecast models are generally the same as those included in general circulation models and were shown schematically in Fig. 10.22. Inclusion of such processes as boundary layer fluxes, vertical mixing by dry and moist convection, formation of clouds and precipitation, and the interaction of cloud and radiation fields requires that the relevant subgrid scale processes be represented in terms of model predicted fields. The approximation of unresolved processes in terms of resolved variables is referred to as *parameterization*; it is probably the most difficult and controversial area of weather and climate modeling.

Perhaps the most important physical process that must be parameterized is convection. Vertical heat transfer by convection is essential in maintaining the observed tropospheric lapse rate and moisture distribution. The simplest way to mimic this effect of unresolved convective motions is through *convective adjustment*. In a simple version of this method, the relative humidity and lapse rate in each grid column are examined at the end of each time step. If the lapse rate is super adiabatic, the temperature profile is adjusted to dry static neutrality in a manner that conserves energy; if the column is conditionally unstable, and the humidity exceeds a specified value, the column is adjusted to moist static neutrality. More sophisticated schemes utilize the fact that moist convection is dependent on low-level moisture convergence and include a moisture budget as part of the parameterization (see Section 11.3).

Various methods are used to relate clouds to the resolved humidity, temperature, and wind fields. None is completely satisfactory. In some models the model-predicted cloud fields are used only in the portion of the model concerned with

precipitation, and cloud–radiation interactions are parameterized using specified cloud climatologies.

13.7 DATA ASSIMILATION

The capability of a numerical forecasting model to produce useful forecasts depends not only on the resolution of the model and the accuracy with which dynamical and physical processes are represented; it is also critically dependent on the initial conditions employed for integrating the model. As Richardson's early attempt to forecast weather numerically showed, observations cannot be used directly to initialize a numerical forecast. It turns out that observational data must be modified in a dynamically consistent fashion in order to obtain a suitable set of data for model initialization. This process is usually referred to as *data assimilation*.

Traditionally, data assimilation was typically divided into two processes: *objective analysis* of the observations and *data initialization*. In the objective analysis step, all data acquired for a given time (generally 00Z or 12Z) from the irregularly spaced observational network of surface and upper air stations are checked for accuracy and interpolated to points on a regular latitude–longitude grid at standard pressure levels. A *background* or "first guess" estimate derived from a short-range forecast is used to fill in grid points in data-sparse regions such as the oceans. Such objectively analyzed data still contain noise that is likely to be interpreted as spuriously large gravity waves when data are used as initial data in a numerical model. In the initialization step, these objectively analyzed data are modified in order to minimize the gravity wave noise, and hence reduce the magnitude of initial velocity and pressure tendencies. More recently the objective analysis and initialization steps have been combined in an approach called *four-dimensional data assimilation*, which makes optimal use of all available observational and forecast information.

13.7.1 The Initialization Problem

The importance of the initialization step in preparing initial data for forecasting with primitive equation models can be illustrated by considering the relative magnitudes of the various terms in the horizontal momentum equation in pressure coordinates:

$$\frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V} \bullet \boldsymbol{\nabla})\,\mathbf{V} + \omega \frac{\partial \mathbf{V}}{\partial p} = -f\mathbf{k} \times \mathbf{V} - \boldsymbol{\nabla}\Phi \tag{13.54}$$

For synoptic-scale motions in extratropical latitudes the wind and pressure fields are in approximate geostrophic balance. Thus, the acceleration following

the motion is measured by the small difference between the two terms $-f\mathbf{k} \times \mathbf{V}$ and $-\nabla\Phi$. Although the Φ field can be determined observationally with quite good accuracy, observed winds are often 10–20% in error. Such observational errors inevitably affect the objectively analyzed fields, even though a suitable analysis procedure may be able to reduce the magnitude of the errors. Nevertheless, using the objectively analyzed winds in initial data can easily lead to an estimate of the Coriolis force that is 10% in error at the initial time. Because the acceleration is normally only about 10% of the Coriolis force in magnitude, an acceleration computed using the observed wind and geopotential fields will generally be 100% in error. Such spurious accelerations not only lead to poor estimates of the initial pressure and velocity tendencies, but also tend to produce large-amplitude gravity wave oscillations as the flow attempts to adjust from the initial unbalanced state back toward a state of quasi-geostrophic balance. These strong gravity waves are not present on the synoptic scale in nature, and their presence in the solution of the model equations quickly spoils any chance of a reasonable forecast.

One possible approach to avoid this problem might be to neglect the observed wind data and derive a wind field from the observed Φ field as part of the objective analysis process. The simplest scheme of this type would be to assume that the initial wind field was in geostrophic balance. However, the error in the computed initial local acceleration would still be ~100%. This can be seen by considering the example of zonally symmetric balanced flow about a circularly symmetric pressure system in a model with constant f. In that case $\partial\mathbf{V}/\partial t = 0$ so that if we neglect the effects of vertical advection the flow is in gradient wind balance (see Section 3.2.4). In vectorial form this balance is simply

$$(\mathbf{V}\bullet\nabla)\,\mathbf{V} + f\mathbf{k} \times \mathbf{V} = -\nabla\Phi \tag{13.55}$$

However, if $\mathbf{V}$ were replaced by $\mathbf{V}_g$ in (13.55), the Coriolis and pressure gradient forces would identically balance. Thus, the inertial force would be unbalanced and

$$(\partial\mathbf{V}/\partial t)_{t=0} \approx -\left(\mathbf{V}_g\bullet\nabla\right)\mathbf{V}_g \tag{13.56}$$

Therefore, by assuming that initially the wind is in geostrophic balance, rather than gradient balance, we compute a local acceleration that is completely erroneous!

It should now be clear that in order to avoid large errors in the initial acceleration, the initial wind field in the above example should be determined by the gradient wind balance, not by the geostrophic approximation or direct observations. The gradient wind formula (3.15) is not in itself a suitable balance condition to use because the radius of curvature must be computed for parcel trajectories. An appropriate balance condition (which is equivalent to the gradient wind in the special case of stationary circularly symmetric flow) can be obtained by taking

the divergence of the horizontal momentum equation, and assuming that at the initial time

$$(\partial \nabla \bullet \mathbf{V}/\partial t)_{t=0} = 0 \quad \text{and} \quad (\nabla \bullet \mathbf{V})_{t=0} = 0 \tag{13.57}$$

then the initial wind field is nondivergent and is related to the geopotential field by the balance equation (11.15).

The balance equation, however, is quadratic in the streamfunction ψ and hence does not provide a unique horizontal velocity field for a given Φ field. Furthermore, realistic baroclinic motions in the atmosphere are not exactly nondivergent, but rather have slowly evolving divergent wind components associated with the secondary circulation discussed in Chapter 6, which are required to maintain the delicate balance between the mass and velocity fields that is maintained as the flow evolves.

13.7.2 Nonlinear Normal Mode Initialization

It should be clear from the previous discussion that the process of model initialization should eliminate high-frequency gravity modes while retaining the relationships among the dynamical and thermodynamic fields that describe the slowly evolving large-scale flow. A method that has been used commonly to obtain these objectives is called *normal mode initialization*. Although some modern data assimilation systems no longer require this initialization step, it is worth reviewing to gain a better understanding of the initialization problem.

The normal modes of a dynamical system are the free modes of oscillation of the system. If the discretized form of the primitive equations used in a prediction model are linearized about a state of rest, the normal modes can be calculated. For each grid point and pressure level there are three normal modes: an eastward and westward propagating gravity mode and a westward propagating Rossby mode. The analyzed fields can be represented by a combination of these normal mode solutions in a manner analogous to the expansion of a field in terms of a finite series of Fourier components or spherical harmonics. Such a representation is referred to as a *projection* onto the normal modes.

This method can be illustrated in a simple fashion using the shallow water equations. We first must rewrite (7.69)–(7.71) to include the nonlinear terms neglected in Section 7.6:

$$\frac{\partial u}{\partial t} - fv + \frac{\partial \Phi}{\partial x} = -\varepsilon \left[u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial x} \right] \tag{13.58}$$

$$\frac{\partial v}{\partial t} + fu + \frac{\partial \Phi}{\partial y} = -\varepsilon \left[u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial x} \right] \tag{13.59}$$

$$\frac{\partial \Phi}{\partial t} + c^2 \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = -\varepsilon \left[\frac{\partial (\Phi u)}{\partial x} + \frac{\partial (\Phi v)}{\partial x} \right] \tag{13.60}$$

where $c^2 = gH$ with H the mean depth, $\Phi = gh$, with h the local deviation of depth from the mean, and ε is an indicator function with value 0 or 1.

If the nonlinear terms on the right in (13.58)–(13.60) are omitted by letting $\varepsilon = 0$, normal mode solutions can be obtained of the form

$$\begin{pmatrix} u \\ v \\ \Phi \end{pmatrix} = \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{\Phi} \end{pmatrix} \exp\left[i\left(kx + ly - \nu t\right)\right] \tag{13.61}$$

where for this simple f-plane model it can be shown (Problem 13.10) that the Rossby normal mode has $\nu_R = 0$, and the eastward and westward gravity modes satisfy $\nu_{G\pm} = \pm[f^2 + c^2(k^2 + l^2)]^{1/2}$. Substituting (13.61) into (13.58)–(13.60) and neglecting the nonlinear terms then yield the relationships of velocity to geopotential for the three normal mode solutions:

$$\hat{u}_R = -ilf^{-1}\hat{\Phi}_R, \hat{v}_R = +ikf^{-1}\hat{\Phi}_R \tag{13.62}$$

$$\begin{aligned} \hat{u}_{G\pm} &= \left(\nu_{G\pm}^2 - f^2\right)^{-1}\left[k\nu_{G\pm} + ilf\right]\hat{\Phi}_{G\pm} \\ \hat{v}_{G\pm} &= \left(\nu_{G\pm}^2 - f^2\right)^{-1}\left[l\nu_{G\pm} - ikf\right]\hat{\Phi}_{G\pm} \end{aligned} \tag{13.63}$$

Here the subscripts R and $G\pm$ indicate the Rossby (geostrophic) mode and the eastward and westward gravity modes, respectively.

The normal mode solutions thus can be expressed in terms of the three independent amplitude coefficients, $\hat{\Phi}_R$, $\hat{\Phi}_{G+}$, and $\hat{\Phi}_{G-}$, respectively.

Suppose that the observed fields at $t = 0$ are given by

$$\begin{pmatrix} u \\ v \\ \Phi \end{pmatrix} = \begin{pmatrix} \hat{u}_0 \\ \hat{v}_0 \\ \hat{\Phi}_0 \end{pmatrix} \exp\left[i\left(kx + ly\right)\right] \tag{13.64}$$

Thus, the observations involve three independent amplitude coefficients. These can be represented as sums over the normal modes:

$$\begin{aligned} \hat{u}_R + \hat{u}_{G+} + \hat{u}_{G-} &= \hat{u}_0 \\ \hat{v}_R + \hat{v}_{G+} + \hat{v}_{G-} &= \hat{v}_0 \\ \hat{\Phi}_R + \hat{\Phi}_{G+} + \hat{\Phi}_{G-} &= \hat{\Phi}_0 \end{aligned} \tag{13.65}$$

Upon substituting from (13.62) and (13.63) into (13.65) we obtain a set of three inhomogeneous linear equations, which can be solved for the relative geopotential amplitudes of the three normal modes when projected onto the observed fields.

The result is

$$\hat{\Phi}_R = \frac{c^2}{\nu_{G+}^2}\left[if\left(l\hat{u}_0 - k\hat{v}_0\right) + \frac{f^2}{c^2}\hat{\Phi}_0\right] \tag{13.66}$$

$$\hat{\Phi}_{G+} = \frac{c^2}{2\nu_{G+}^2}\left[\left(k\nu_{G+} - ifl\right)\hat{u}_0 + \left(l\nu_{G+} + ifk\right)\hat{v}_0 + \left(k^2 + l^2\right)\hat{\Phi}_0\right] \tag{13.67}$$

$$\hat{\Phi}_{G-} = \frac{c^2}{2\nu_{G-}^2}\left[\left(k\nu_{G-} - ifl\right)\hat{u}_0 + \left(l\nu_{G-} + ifk\right)\hat{v}_0 + \left(k^2 + l^2\right)\hat{\Phi}_0\right] \tag{13.68}$$

It is readily verified that if $\hat{u}_0$ and $\hat{v}_0$ are in exact geostrophic balance with $\hat{\Phi}_0$, then $\hat{\Phi}_R = \hat{\Phi}_0$ and $\hat{\Phi}_{G+} = \hat{\Phi}_{G-} = 0$. When the observed velocities are not exactly in geostrophic balance, (13.66)–(13.68) give the relative weightings for the projections onto the Rossby mode and the two gravity modes. For the linear system it is possible to set the gravity modes to zero and to initialize with the projection of the observed fields onto the Rossby mode. Note from (13.66) that the relative weighting of the velocity observations (which appear in the form of relative vorticity) increases as the scale of the disturbance decreases.

As explained in the previous section, merely setting the gravity modes to zero does not work for the nonlinear system. In that case when the normal mode projection defined by (13.62) and (13.66) is used to determine the amplitude coefficients in (13.61) and the result is substituted into (13.58)–(13.60), the unbalanced advective terms will lead to large initial accelerations. This problem can be solved by including the gravity modes, but with their amplitudes adjusted so that their initial tendencies vanish. Because the equations are nonlinear it is necessary to carry out the solution iteratively, using the linear solution as a first guess.

In practice the nonlinear normal mode method works very well when applied to primitive equation forecast models. The method can be modified to incorporate not only nonlinear advection, but also the effects of diabatic heating in the tropics in order to preserve the Hadley circulation in initial data.

13.7.3 Four-Dimensional Data Assimilation

The traditional approach to objective analysis works well as long as the prediction model is limited to a region of adequate data coverage (such as the North American continent) and all observations are made at the standard 00UTC and 12UTC times so that they can be directly incorporated into synoptic analyses. For global analysis the traditional approach is less successful, as in many oceanic regions there are very few conventional data sources available, and it is necessary to rely on asynoptic data (i.e., observations at nonstandard times, such as those from ships, aircraft, and satellites) and on estimates from earlier forecasts. Asynoptic data have proved invaluable, especially over the oceans and in the Southern

Hemisphere, but they are not incorporated easily into traditional objective analysis schemes.

Deficiencies of the traditional approach have led the major forecast centers to devise assimilation procedures in which the separate processes of objective analysis and data initialization are combined into a continuous cycle of data assimilation, which is referred to as *four-dimensional variational assimilation* (4DVAR). The objective of this approach is to utilize all available information during an assimilation window period of several hours to obtain a best estimate of the state of the atmosphere at the analysis time. The information utilized includes a first guess based on a short-term forecast from the previous analysis time and all observations acquired during the window period. These various pieces of information, weighted by their statistical errors, are used to derive a *cost function*, expressing the misfit between the observations and the analysis. The cost function is then minimized in order to achieve an analysis that is the most likely estimate of the true state of the atmosphere at the analysis time.

A simple example of the cost function approach was given by Kalnay (2003). Suppose that we have two independent observations of temperature, T_1 and T_2, which are unbiased but have error variances of σ_1^2 and σ_2^2, respectively. The cost function can then be defined as

$$J(T) = \frac{1}{2}\left[\frac{(T - T_1)^2}{\sigma_1^2} + \frac{(T - T_2)^2}{\sigma_2^2}\right] \tag{13.69}$$

where T is the analysis temperature that we wish to determine. The minimum of $J$is determined by evaluating $\partial J/\partial T = 0$, which yields the best estimate for T:

$$T = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right) T_1 + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right) T_2 \tag{13.70}$$

Thus, an observation that has small error variance is weighted more heavily than an observation with large error variance.

An example of a 4DVAR procedure, as used at ECMWF, is shown schematically in Fig. 13.8. The ECMWF scheme utilizes a 12-h assimilation window extending from 9 h prior to 3 h following the analysis times, which are 00Z and 12Z. An assimilation cycle is begun by using a forecast from the initial state at the previous analysis time as the first guess, or background state. All available data within the assimilation window are used to update the analysis defined by the first guess, with weightings determined by the error statistics for the particular type of data. The update process is subject to dynamical constraints to assure a balanced state at the analysis time. Forecasts are then carried out for a 10-day period once each day starting from the initialized state for 12Z.

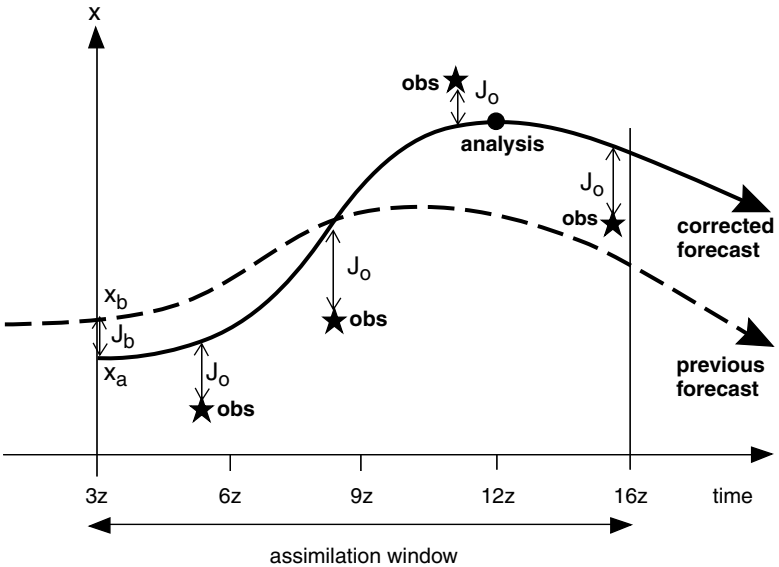


Fig. 13.8 The ECMWF data assimilation process. Vertical axis indicates the value of an atmospheric field variable. The curve labeled x_b is the first guess or "background" field, and x_a is the analysis field. J_b indicates the misfit between the background and analysis fields. Stars indicate the observations, and J_0 designates the "cost function," a measure of the misfit between the observations and the analysis, which along with J_b is minimized to produce the best estimate of the state of the atmosphere at the analysis time (Courtesy ECMWF.)

13.8 PREDICTABILITY AND ENSEMBLE PREDICTION SYSTEMS

For short-range forecasts (1 or 2 days) of the midlatitude 500-hPa flow, it is possible to neglect diabatic heating and frictional dissipation. It is important, however, to have good initial data in the region of interest because on this short time scale forecasting ability depends primarily on the proper advection of the initial vorticity field. As the length of the forecast period increases, the effects of propagation of influences from other regions and changes due to various sources and sinks of momentum and energy become increasingly important. Therefore, the flow at a point in the middle-latitude troposphere will depend on initial conditions for an increasing domain and on accurate representation of various physical processes as the period of the forecast is extended. In fact, according to the estimate of Smagorinsky shown in Fig. 13.9, for periods greater than 1 week it is necessary to know the initial state of the entire global atmosphere from the stratosphere to the surface, as well as the state of the upper layers of the oceans.

However, even if an ideal data network were available to specify the initial state on a global scale, there still would be a time limit beyond which a useful

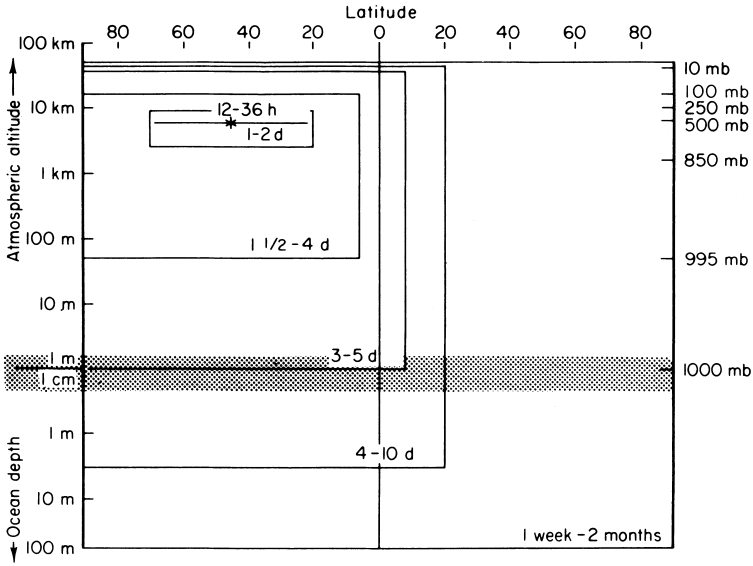


Fig. 13.9 A schematic diagram of the domain of initial dependence for a prediction point in midtroposphere at midlatitudes (denoted by a star) as a function of forecast time span. Atmospheric and ocean elevations are given on a logarithmic scale, increasing upward and downward, respectively. The stripped area is the interface zone. (After Smagorinsky, 1967. Reproduced with permission of the American Meteorological Society.)

forecast would not be possible. The atmosphere is a continuum with a continuous spectrum of scales of motion. No matter how fine the grid resolution is made, there will always be motions whose scales are too small to be properly observed and represented in the model. Thus, there is an unavoidable level of error in the determination of the initial state. The nonlinearity and instability of atmospheric flow will inevitably cause the small inherent errors in the initial data to grow and gradually affect the larger scales of motion so that the forecast flow field will eventually evolve differently from the actual flow.

A very simple example of this process of error growth was presented by Lorenz (1984). Lorenz illustrated the general problem of predictability by considering the first-order quadratic difference equation

$$Y_{s+1} = aY_s - Y_s^2 \tag{13.71}$$

which may be solved for Y_{s+1} by iteration if the constant a and the initial condition Y_0 are specified. The solid line in Fig. 13.10 shows a portion of the solution sequence with $a = 3.75$ and $Y_0 = 1.5$. This line represents a control run that may be regarded as defining the "observations." The dashed line is the solution for a case in which the initial value Y_0 is perturbed by 0.001, whereas the dash–dot line

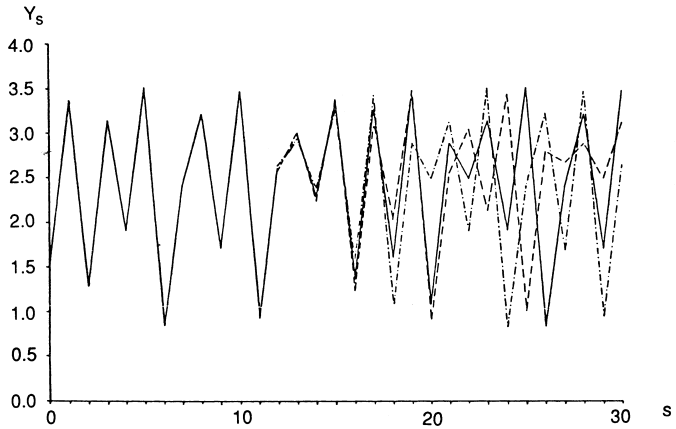


Fig. 13.10 Effects of small errors in the initial conditions (dashed line) and imperfect knowledge of the governing equation (dash–dotted line) on the prediction of a time series generated by the difference equation (13.69). (Courtesy of E. Källén.)

shows a solution in which the original initial data are used but the coefficient a is perturbed by 0.001.

For the first several steps the two "predictions" are very close to the observed values. However, after about 15 steps they begin to diverge. Thus, errors in either the initial conditions or the governing equation can produce comparable results—the departure of a predicted solution from the observed sequence increases rapidly. For the case shown here only about 20 steps are required for the three sequences to lose all resemblance to each other. Such behavior is characteristic of a wide variety of systems governed by deterministic equations, including atmospheric flows.

Estimates of how this sort of error growth limits the inherent predictability of the atmosphere can be made using primitive equation forecast models. In these predictability experiments a "control" run is made using initial data corresponding to the observed flow at a given time. The initial data are then perturbed by introducing small random errors, and the model is run again. The growth of inherent error can then be estimated by comparing the second "forecast" run with the control. Results from a number of such studies indicate that the doubling time for the root mean square geopotential height error is about 2 to 3 days for small errors and somewhat greater for large errors. Thus, the theoretical limit for predictability on the synoptic scale is probably about 2 weeks.

Actual predictive skill with present deterministic models is, however, less than the theoretical limit imposed by inherent error growth. A number of factors are probably responsible for the failure of current models to reach the theoretical forecast skill limits. These include observational and analysis errors in initial data, inadequate model resolution, and unsatisfactory representation of physical processes involving radiation, clouds, precipitation, and boundary layer fluxes.

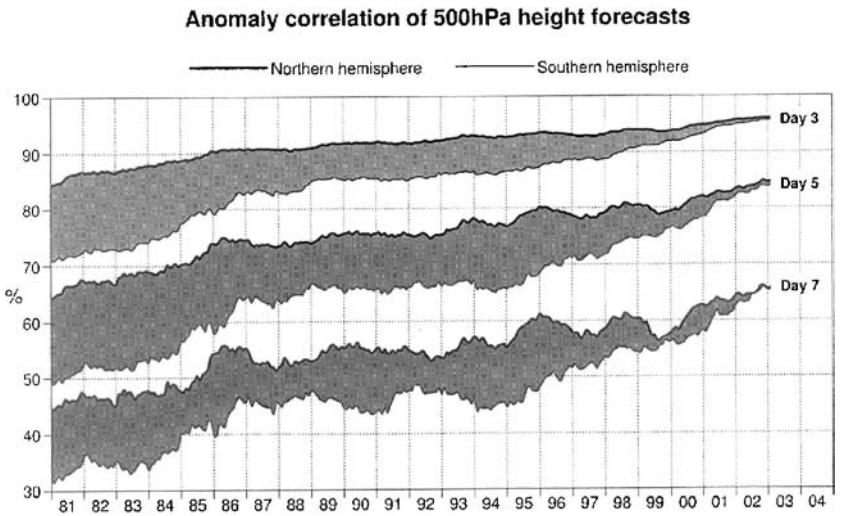


Fig. 13.11 Anomaly correlation of 500-hPa height for 3-, 5-, and 7-day forecasts for the ECMWF operational model as a function of year. Top and bottom of each band correspond to Northern and Southern Hemispheres, respectively. Note that the difference in skill between the two hemispheres has nearly disappeared in the past few years, due to the successful assimilation of satellite data. (Courtesy of Dr. Adrian Simmons, updated from Simmons and Hollingsworth, 2002.)

An indication of the skill of the present generation of global forecast models is given in Fig. 13.11, in which forecast skill is plotted in terms of the *anomaly correlation* of the 500-hPa geopotential height field, defined as the correlation between observed and predicted deviations from climatology. Subjective evaluations suggest that useful forecasts are obtained when the anomaly correlation is greater than 0.6. Thus, over the past 20 years the range for which ECMWF forecasts show useful skill has increased from 5 days to more than 7 days. There can, of course, be quite wide variations in skill from one situation to another reflecting variations in the degree of predictability from one atmospheric flow regime to another.

As mentioned earlier, it has become common for operational forecast centers to assess the likely skill of medium-range predictions by carrying out several parallel runs of a forecast model starting from slight perturbations of the initial conditions determined by the data assimilation process. For example, in the ECMWF ensemble prediction system as of 2003, 50 runs are made for each analysis time with a T255, 40 level model. In each of the runs, the analysis is perturbed by an amount that is within the uncertainty limits of the analysis. The perturbations are based on the method of *singular vectors*, which is designed to identify the regions of the atmosphere where small uncertainties are likely to magnify rapidly, and hence will tend to have a large impact on the forecast.

In general, the ensemble mean provides a better forecast than individual members of the ensemble, as the ensemble mean tends to filter out the most uncertain aspects of the forecast. The spread of ensemble members provides an important indication of the likely accuracy of the ensemble mean forecast. If there is a large spread among members of the ensemble, not all members can be correct, and the reliability of the forecast should be judged as low. However, when most members of the ensemble agree, the reliability of the forecast should be high (provided that the forecast model is unbiased). In summary, the ensemble prediction system not only can improve forecast reliability, but also can provide information on the degree of reliability of a specific forecast.

PROBLEMS

13.1. Show that for the barotropic vorticity equation on the Cartesian β plane (13.26) enstrophy and kinetic energy are conserved when averaged over the whole domain, that is, that the following integral constraints are satisfied:

$$\frac{d}{dt}\iint \frac{\zeta^2}{2}dxdy = 0 \qquad \frac{d}{dt}\iint \frac{\nabla\psi \bullet \nabla\psi}{2}dxdy = 0$$

Hint: To prove energy conservation, multiply (13.26) through by $-\psi$ and use the chain rule of differentiation.

13.2. Verify the expression (13.31) in the text. Use periodic boundary conditions in both x and y.

13.3. The Euler backward method of finite differencing the advection equation is a two-step method consisting of a forward prediction step, followed by a backward corrector step. In the notation of Section 13.3.2 the complete cycle is thus defined by

$$\hat{q}_m^* - \hat{q}_{m,s} = -\frac{\sigma}{2}\left(\hat{q}_{m+1,s} - \hat{q}_{m-1,s}\right)$$
$$\hat{q}_{m,s+1} - \hat{q}_{m,s} = -\frac{\sigma}{2}\left(\hat{q}_{m+1}^* - \hat{q}_{m-1}^*\right)$$

where $\hat{q}_m^*$ is the first guess for time step $s+1$. Use the method of Section 13.3.3 to determine the necessary condition for stability of this method.

13.4. Carry out truncation error analyses analogous to that of Table 13.1 for the centered difference approximation to the advection equation but for the cases $\sigma = 0.95$ and $\sigma = 0.25$.

13.5. Suppose that the streamfunction ψ is given by a single sinusoidal wave $\psi(x) = A\sin(kx)$. Find an expression for the error of the finite difference approximation

$$\frac{\partial^2 \psi}{\partial x^2} \approx \frac{\psi_{m+1} - 2\psi_m + \psi_{m-1}}{\delta x^2}$$

for $k\delta x = \pi/8, \pi/4, \pi/2,$ and π. Here $x = m\delta x$ with $m = 0, 1, 2, \ldots$.

13.6. Using the method given in Section 13.3.3, evaluate the computational stability of the following two finite difference approximations to the one-dimensional advection equation:

$$\text{(a) } \hat{\zeta}_{m,s+1} - \hat{\zeta}_{m,s} = -\sigma\left(\hat{\zeta}_{m,s} - \hat{\zeta}_{m-1,s}\right)$$

$$\text{(b) } \hat{\zeta}_{m,s+1} - \hat{\zeta}_{m,s} = -\sigma\left(\hat{\zeta}_{m+1,s} - \hat{\zeta}_{m,s}\right)$$

where $\sigma = c\delta t/\delta x > 0$. [The schemes labeled (a) and (b) are referred to as *upstream* and *downstream* differencing, respectively.] Show that scheme (a) damps the advected field and compute the fractional damping rate per time step for $\sigma = 0.25$ and $k\delta x = \pi/8$ for a field with the initial form $\zeta = exp(ikx)$.

13.7. Using a staggered horizontal grid analogous to that shown in Fig. 13.5 (but for an equatorial β-plane geometry), express the linearized shallow water equations (11.29)–(11.31) in finite difference form using the finite differencing and averaging operator notation introduced in Section 13.6.1.

13.8. Verify the equality

$$\left(\frac{1 - i\tan\theta_p}{1 + i\tan\theta_p}\right) = \exp\left(-2i\theta_p\right)$$

given in (13.22).

13.9. Compute the ratio of the numerical phase speed to the true phase speed, c'/c, for the implicit differencing scheme of (13.19) for $p = \pi, \pi/2, \pi/4, \pi/8$, and $\pi/16$. Let $\sigma = 0.75$ and $\sigma = 1.25$. Compare your results to those of Table 13.1.

13.10. Using the technique of Section 13.3.1, show that the following four-point difference formula for the first derivative is of fourth-order accuracy:

$$\psi'(x_0) \approx \frac{4}{3}\left(\frac{\psi(x_0 + \delta x) - \psi(x_0 - \delta x)}{2\delta x}\right) - \frac{1}{3}\left(\frac{\psi(x_0 + 2\delta x) - \psi(x_0 - 2\delta x)}{4\delta x}\right) \tag{13.72}$$

13.11. The Dufort–Frankel method for approximating the one-dimensional diffusion equation

$$\frac{\partial q}{\partial t} = K\frac{\partial^2 q}{\partial x^2}$$

can be expressed in the notation of Section 13.3.2 as

$$\hat{q}_{m,s+1} = \hat{q}_{m,s-1} + r\left[\hat{q}_{m+1,s} - \left(\hat{q}_{m,s+1} + \hat{q}_{m,s-1}\right) + \hat{q}_{m-1,s}\right]$$

where $r \equiv 2K\delta t/(\delta x)^2$. Show that this scheme is an explicit differencing scheme and that it is computationally stable for all values of δt.

13.12. Starting with the assumed solution (13.61), obtain the normal mode solutions to the linearized versions of (13.58)–(13.60) and hence verify (13.62) and (13.63).

13.13. Show that the projection of the observations onto the Rossby normal mode (13.66) is equivalent to requiring that the quasi-geostrophic potential vorticity of the Rossby mode be proportional to the sum of the observed relative vorticity minus the observed geopotential multiplied by the factor (f/c^2). *Hint*: Linearize (4.26) and assume that f is constant.

13.14. Derive the best estimate of T given in (13.70) based on minimization of the cost function of (13.69).

MATLAB EXERCISES

M13.1. The MATLAB script **finite_diff1.m** can be used to compare the centered difference first derivative formula (13.4) for the function $\psi(x) = \sin(\pi x/4)$ to the analytic expression $d\psi/dx = (\pi/4)\cos(\pi x/4)$ for various numbers of grid intervals in the domain $-4 \le x \le 4$. Graph the maximum error as a function of number of grid intervals, *ngrid*, in the range of 4 to 64. Carry out a similar analysis for the second derivative. (Note that first-order differencing is used next to the boundaries.)

M13.2. Modify the script of M13.1 to evaluate the error in the first derivative of the function $\tanh(x)$ using the same domain in x.

M13.3. The MATLAB script **advect_1.m** demonstrates the leapfrog differencing scheme for the one-dimensional advection equation as given in (13.8). By running the script for various grid intervals, find the dependence of phase error (degrees per wave period) on the number of grid intervals per wavelength for the range 4 to 64 and compare your results with Table 13.1.

M13.4. Modify the script **advect_1.m** by substituting the upstream differencing scheme described in Problem 13.6 and the Euler backward scheme described in Problem 13.3. Compare the accuracy in phase and amplitude of these two schemes with the leapfrog scheme of M13.3.

M13.5. The MATLAB script **advect_2.m** is similar to the script of M13.3 except that the initial tracer distribution is a localized positive definite pulse of width 0.25. Run the script for 100, 200, and 400 grid intervals. Explain why the pulse changes shape as it is advected. Using the results of Problem 13.10, modify the script to provide a fourth-order accurate approximation to the advection term and compare the accuracy of the fourth-order version to that of the second-order accurate system for the case with 400 grid intervals.

M13.6. The MATLAB script **advect_3.m** is a variant of the second-order accurate script of M13.5 in which the implicit differencing of (13.19) is utilized. Run this script with $\sigma = 1$ and 1.25 and 400 grid intervals and then try running **advect_2.m** with these values of σ. Give a qualitative explanation of these results referring to Problem 13.9.

M13.7. The MATLAB script **barotropic_model.m** can be used to solve a finite difference approximation to the barotropic vorticity equation using the flux form of the nonlinear terms given in (13.30) and leapfrog time differencing. In this example the initial flow consists of a localized vortex and a constant zonal mean flow. Run the model for a 10-day simulation using different values of the time step to determine the maximum time step permitted to maintain numerical stability. How does this compare with the CFL criterion (13.18)?

M13.8. The MATLAB script **predict.m** solves the quadratic difference equation (13.71) for a specified coefficient $a = 3.75$ and an initial condition $Y = 1.5$. (This will be called the control integration.) The user can specify additional iterations in which small positive and negative perturbations are added to the initial condition. Run this script and note how the solutions begin to diverge after about 15 iterations. Modify the script to compute an ensemble average of the perturbed predictions. Show that the ensemble mean provides a better prediction of the control run than a random member of the perturbed integrations. [*Hint*: Compute the standard deviation of the ensemble members at each iteration and compare to the magnitude of the difference between the ensemble mean and the control.]

M13.9. The MATLAB script **Lorenz_model.m** gives an accurate numerical solution to the famous three-component Lorenz equations, which are often used to demonstrate sensitive dependence on initial conditions in chaotic

systems (such as the atmosphere). The equations for the Lorenz model can be written as

$$dX/dt = -\sigma X + \sigma Y \tag{13.73}$$

$$dY/dt = -XZ + rX - Y \tag{13.74}$$

$$dZ/dt = XY - bZ \tag{13.75}$$

Here (X, Y, Z) may be regarded as a vector defining the "climate" and σ, r, and b are constants. Run the script letting the initial value of $X = 10$ and verify that in the $X - Z$ plane the resulting trajectory of solution points has the well known "butterfly-wings" shape. Modify the code to save and plot the time histories of the variables and compare solutions in which the initial condition on X is increased and decreased by 0.1%. How long (in nondimensional time units) do the three solutions remain within 10% of each other?

M13.10. Modify the code in **Lorenz_model.m** to include a constant forcing $F = 10$ on the right side of the equation governing dX/dt. Describe how the character of the solution changes in this case. [See Palmer (1993)].

M13.11. Use the results of Problem 13.7 to write a staggered grid version of the MATLAB script **forced_equatorial_mode2.m** (see Problem M11.6) for the shallow water model on the equatorial β-plane. Set the staggered grid such that u and Φ are defined at the equator and v is defined at points $\delta y/2$ north and south of the equator. Compare the results of the staggered grid model with those of **forced_equatorial_mode2.m** when the latter has a grid spacing half that of the former.

M13.12. The MATLAB script **nonlinear_advect_diffuse.m** provides a numerical approximation to the one-dimensional nonlinear advection–diffusion equation

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} + K\frac{\partial^2 u}{\partial x^2}$$

with initial condition $u(x, 0) = \sin\left(2\pi x \big/ Lx\right)$. The script uses leapfrog differencing for the advective term and forward differencing for the diffusion term. In the absence of diffusion the flow would quickly evolve to a shock, but diffusion prevents this from occurring. Run the script and by varying δt determine the maximum time step for a stable solution. Modify the code by expressing the diffusion term in the Dufort–Frankel differencing scheme given in Problem 13.11. Determine the maximum time step allowed for stability in this case. How does the accuracy of the solution change as the time step is increased?

Suggested References

Durran, *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics* is an excellent textbook covering the numerical methods used in the atmospheric sciences at the graduate level.

Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability* is an excellent graduate-level text on all aspects of modern numerical weather forecasting.

APPENDIX A

Useful Constants and Parameters

Gravitational constant	$G = 6.673 \times 10^{-11}\,\mathrm{N\,m^2\,kg^{-2}}$
Gravity at sea level	$g_0 = 9.81\,\mathrm{m\,s^{-2}}$
Mean radius of the earth	$a = 6.37 \times 10^6\,\mathrm{m}$
Earth's angular speed of rotation	$\Omega = 7.292 \times 10^{-5}\,\mathrm{rad\,s^{-1}}$
Universal gas constant	$R^* = 8.314 \times 10^3\,\mathrm{J\,K^{-1}\,kmol^{-1}}$
Gas constant for dry air	$R = 287\,\mathrm{J\,K^{-1}\,kg^{-1}}$
Specific heat of dry air at constant pressure	$c_p = 1004\,\mathrm{J\,K^{-1}\,kg^{-1}}$
Specific heat of dry air at constant volume	$c_v = 717\,\mathrm{J\,K^{-1}\,kg^{-1}}$
Ratio of specific heats	$\gamma = c_p/c_v = 1.4$
Molecular weight of water	$m_v = 18.016\,\mathrm{kg\,kmol^{-1}}$

Latent heat of condensation at 0°C	$L_c = 2.5 \times 10^6\,\mathrm{J\,kg^{-1}}$
Mass of the earth	$M = 5.988 \times 10^{24}\,\mathrm{kg}$
Standard sea level pressure	$p_0 = 1013.25\,\mathrm{hPa}$
Standard sea level temperature	$T_0 = 288.15\,\mathrm{K}$
Standard sea level density	$\rho_0 = 1.225\,\mathrm{kg\,m^{-3}}$

APPENDIX B

List of Symbols

Only the principal symbols are listed. Symbols formed by adding primes, overbars, or subscripted indices are not listed separately. Boldface type indicates vector quantities. Where symbols have more than one meaning, the section where the second meaning is first used is indicated in the list.

a	Radius of the earth; inner radius of a laboratory annulus (Section 10.5)
b	Buoyancy; outer radius of a laboratory annulus (Section 10.5)
c	Phase speed of a wave
c_p	Specific heat of dry air at constant pressure
c_{pv}	Specific heat of water vapor at constant pressure
c_v	Specific heat of dry air at constant volume
c_w	Specific heat of liquid water

d	Grid distance
e	Internal energy per unit mass
f	Coriolis parameter ($\equiv 2\Omega \sin \phi$)
g	Magnitude of gravity
$\mathbf{g}$	Gravity
$\mathbf{g}^*$	Gravitational acceleration
h	Depth of fluid layer; moist static energy (Section 9.5)
i	Square root of minus one
$\mathbf{i}$	Unit vector along the x axis
$\mathbf{j}$	Unit vector along the y axis
$\mathbf{k}$	Unit vector along the z axis
k	Zonal wave number
l	Mixing length; meridional wave number (Section 7.5)
m	Mass element; vertical wave number (Section 7.4); planetary wave number (Section 13.5)
m_v	Molecular weight of water
n	Distance in direction normal to a parcel trajectory; meridional index for equatorial waves (Section 11.4)
$\mathbf{n}$	Unit vector normal to a parcel trajectory
p	Pressure
p_s	Standard constant pressure; surface pressure in σ coordinates (Section 10.3)
q	Quasi-geostrophic potential vorticity; water vapor mixing ratio (Section 9.5)
q_s	Saturation mixing ratio
r	Radial distance in spherical coordinates
$\mathbf{r}$	A position vector
s	Generalized vertical coordinate; distance along a parcel trajectory (Section 3.2); entropy (Section 2.7); dry static energy (Section 9.5)

t	Time
$\mathbf{t}$	Unit vector parallel to a parcel trajectory
u^*	Friction velocity
u	x component of velocity (eastward)
v	y component of velocity (northward)
w	z component of velocity (upward)
w^*	Vertical motion in log-pressure system
x,y,z	Eastward, northward, and upward distance, respectively
z^*	Vertical coordinate in log-pressure system
$\mathbf{A}$	An arbitrary vector
A	Area
B	Convective available potential energy
C_d	Surface drag coefficient
D_e	Depth of Ekman layer
E	Evaporation rate
E_I	Internal energy
$\mathbf{F}$	A force; EP flux (Section 10.2)
$\mathbf{Fr}$	Frictional force
G	Universal gravitational constant; zonal force (Section 10.2)
H	Scale height
J	Diabatic heating rate
K	Total horizontal wavenumber; kinetic energy (Section 9.3)
K_m	Eddy viscosity coefficient
L	A length scale
L_c	Latent heat of condensation
M	Mass; mass convergence in Ekman layer (Section 5.4); absolute zonal momentum (Section 9.3); angular momentum (Section 10.3)
N	Buoyancy frequency

P	Ertel potential vorticity; available potential energy (Section 8.3); precipitation rate (Section 11.3)
$\mathbf{Q}$	Q vector
R	Gas constant for dry air; distance from the axis of rotation of the earth to a point on the surface of the earth (Section 1.5); diabatic energy generation rate (Section 10.4)
$\mathbf{R}$	Vector in the equatorial plane directed from the axis of rotation to a point on the surface of the earth
R^*	Universal gas constant
S_p	$\equiv -T\partial \ln\theta/\partial p$, stability parameter in pressure coordinates
T	Temperature
U	Horizontal velocity scale
V	Speed in natural coordinates
δV	Volume increment
$\mathbf{U}$	Three-dimensional velocity vector
$\mathbf{V}$	Horizontal velocity vector
W	Vertical motion scale
X	Zonal turbulent drag force
Z	Geopotential height
α	Specific volume
β	$\equiv df/dy$, variation of the Coriolis parameter with latitude; angular direction of the wind (Section 3.3)
γ	$\equiv c_p/c_v$, the ratio of specific heats
ε	Rate of frictional energy dissipation; thermal expansion coefficient of water (Section 10.7)
ζ	Vertical component of relative vorticity
η	Vertical component of absolute vorticity; weighting function for heating profile (Section 11.3)
θ	Potential temperature
$\dot{\theta}$	$\equiv D\theta/Dt$, vertical motion in isentropic coordinates

θ_e	Equivalent potential temperature
κ	$\equiv R/c_p$, ratio of gas constant to specific heat at constant pressure; Rayleigh friction coefficient (Section 10.6)
λ	Longitude, positive eastward
μ	Dynamic viscosity coefficient
ν	Angular frequency; kinematic viscosity (Section 1.4)
ρ	Density
σ	$\equiv -RT_0 p^{-1} d \ln \theta_0 / dp$, standard atmosphere static stability parameter in isobaric coordinates;
	$\equiv -p/p_s$, vertical coordinate in σ system (Section 10.3); "Density" in isentropic coordinates (Section 4.6)
τ_d	Diffusive time scale
τ_E	Eddy stress
ϕ	Latitude
χ	Geopotential tendency; meridional stream function; tracer mixing ratio
ψ	Horizontal stream function
ω	Vertical velocity ($\equiv dp/dt$) in isobaric coordinates
$\boldsymbol{\omega}$	Vorticity vector
Γ	$\equiv -dT/dz$, lapse rate of temperature
Γ_d	Dry adiabatic lapse rate
Φ	Geopotential
Π	Exner function
Θ	Potential temperature deviation
Ω	Angular speed of rotation of the earth; angular speed of rotation of laboratory annulus (Section 10.7)
$\boldsymbol{\Omega}$	Angular velocity of the earth

APPENDIX C

Vector Analysis

C.1 VECTOR IDENTITIES

The following formulas may be shown to hold where Φ is an arbitrary scalar and **A** and **B** are arbitrary vectors.

$$\nabla \times \nabla \Phi = 0$$

$$\nabla \bullet (\Phi \mathbf{A}) = \Phi \nabla \bullet (\mathbf{A}) + \mathbf{A} \bullet \nabla \Phi$$

$$\nabla \times (\Phi \mathbf{A}) = \nabla \Phi \times \mathbf{A} + \Phi (\nabla \times \mathbf{A})$$

$$\nabla \bullet (\nabla \times \mathbf{A}) = 0$$

$$(\mathbf{A} \bullet \nabla) \mathbf{A} = \frac{1}{2} \nabla (\mathbf{A} \bullet \mathbf{A}) - \mathbf{A} \times (\nabla \times \mathbf{A})$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A} (\nabla \bullet \mathbf{B}) - \mathbf{B} (\nabla \bullet \mathbf{A}) - (\mathbf{A} \bullet \nabla) \mathbf{B} + (\mathbf{B} \bullet \nabla) \mathbf{A}$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \bullet \mathbf{C}) \mathbf{B} - (\mathbf{A} \bullet \mathbf{B}) \mathbf{C}$$

C.2 INTEGRAL THEOREMS

(a) Divergence theorem:

$$\int_A \mathbf{B} \bullet \mathbf{n} dA = \int_V \mathbf{V} \bullet \mathbf{B} dV$$

where V is a volume enclosed by surface A and $\mathbf{n}$ is a unit normal on A.

(b) Stokes' theorem:

$$\oint \mathbf{B} \bullet d\mathbf{l} = \int_A (\mathbf{V} \times \mathbf{B}) \bullet \mathbf{n} dA$$

where A is a surface bounded by the line traced by the position vector $\mathbf{l}$ and $\mathbf{n}$ is a unit normal of A.

C.3 VECTOR OPERATIONS IN VARIOUS COORDINATE SYSTEMS

(a) Cartesian coordinates: (x, y, z)

Coordinate	Symbol	Velocity component	Unit vector
Eastward	x	u	$\mathbf{i}$
Northward	y	v	$\mathbf{j}$
Upward	z	w	$\mathbf{k}$

$$\nabla \Phi = \mathbf{i} \frac{\partial \Phi}{\partial x} + \mathbf{j} \frac{\partial \Phi}{\partial y} + \mathbf{k} \frac{\partial \Phi}{\partial z}$$

$$\nabla \bullet \mathbf{V} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$$

$$\mathbf{k} \bullet (\nabla \times \mathbf{V}) = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

$$\nabla_h^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2}$$

(b) Cylindrical coordinates: (r, λ, z)

Coordinate	Symbol	Velocity component	Unit vector
Radial	r	u	$\mathbf{i}$
Azimuthal	λ	v	$\mathbf{j}$
Upward	z	w	$\mathbf{k}$

$$\nabla\Phi = \mathbf{i}\frac{\partial\Phi}{\partial r} + \mathbf{j}\frac{1}{r}\frac{\partial\Phi}{\partial\lambda} + \mathbf{k}\frac{\partial\Phi}{\partial z}$$

$$\nabla\bullet\mathbf{V} = \frac{1}{r}\frac{\partial(ru)}{\partial r} + \frac{1}{r}\frac{\partial v}{\partial\lambda}$$

$$\mathbf{k}\bullet(\nabla\times\mathbf{V}) = \frac{1}{r}\frac{\partial(rv)}{\partial r} - \frac{1}{r}\frac{\partial u}{\partial\lambda}$$

$$\nabla_h^2\Phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Phi}{\partial\lambda^2}$$

(c) Spherical coordinates: (λ, ϕ, r)

Coordinate	Symbol	Velocity component	Unit vector
Longitude	λ	u	$\mathbf{i}$
Latitude	ϕ	v	$\mathbf{j}$
Height	r	w	$\mathbf{k}$

$$\nabla\Phi = \frac{\mathbf{i}}{r\cos\phi}\frac{\partial\Phi}{\partial\lambda} + \mathbf{j}\frac{1}{r}\frac{\partial\Phi}{\partial\phi} + \mathbf{k}\frac{\partial\Phi}{\partial r}$$

$$\nabla\bullet\mathbf{V} = \frac{1}{r\cos\phi}\left[\frac{\partial u}{\partial\lambda} + \frac{\partial(v\cos\phi)}{\partial\phi}\right]$$

$$\mathbf{k}\bullet(\nabla\times\mathbf{V}) = \frac{1}{r\cos\phi}\left[\frac{\partial v}{\partial\lambda} - \frac{\partial(u\cos\phi)}{\partial\phi}\right]$$

$$\nabla_h^2\Phi = \frac{1}{r^2\cos^2\phi}\left[\frac{\partial^2\Phi}{\partial\lambda^2} + \cos\phi\frac{\partial}{\partial\phi}\left(\cos\phi\frac{\partial\Phi}{\partial\phi}\right)\right]$$

APPENDIX D

Moisture Variables

D.1 EQUIVALENT POTENTIAL TEMPERATURE

A mathematical expression for θ_e can be derived by applying the first law of thermodynamics to a mixture of 1 kg dry air plus q kg of water vapor (q, called the *mixing ratio*, is usually expressed as grams of vapor per kilogram of dry air). If the parcel is not saturated, the dry air satisfies the energy equation

$$c_p dT - \frac{d\,(p-e)}{p-e} RT = 0 \tag{D.1}$$

and the water vapor satisfies

$$c_{pv} dT - \frac{de}{e}\frac{R^*}{m_v} T = 0 \tag{D.2}$$

where the motion is assumed to be adiabatic. Here, e is the partial pressure of the water vapor, c_{pv} the specific heat at constant pressure for the vapor, R^* the universal gas constant, and m_v the molecular weight of water. If the parcel is saturated, then consdensation of $-dq_s$ kg vapor per kilogram dry air will heat the

mixture of air and vapor by an amount of heat that goes into the liquid water, the saturated parcel must satisfy the energy equation

$$c_p dT + q_s c_{pv} dT - \frac{d\,(p - e_s)}{p - e_s} RT - q_s \frac{de_s}{e_s} \frac{R^*}{m_v} T = -L_c dq_s \tag{D.3}$$

where q_s and e_s are the saturation mixing ratio and vapor pressure, respectively. The quantity de_s/e_s may be expressed in terms of temperature using the Clausius–Clapeyron equation[1]

$$\frac{de_s}{dT} = \frac{m_v L_c e_s}{R^* T^2} \tag{D.4}$$

Substituting from (D.4) into (D.3) and rearranging terms we obtain

$$-L_c d\left(\frac{q_s}{T}\right) = c_p \frac{dT}{T} - \frac{Rd\,(p - e_s)}{p - e_s} + q_s c_{pv} \frac{dT}{T} \tag{D.5}$$

If we now define the potential temperature of the dry air θ_d, according to

$$c_p d \ln \theta_d = c_p d \ln T - Rd \ln (p - e_s)$$

we can rewrite (D.5) as

$$-L_c d\left(\frac{q_s}{T}\right) = c_p d \ln \theta_d + q_s c_{pv} d \ln T \tag{D.6}$$

However, it may be shown that

$$dL_c/dT = c_{pv} - c_w \tag{D.7}$$

where c_w is the specific heat of liquid water. Combining (D.7) and (D.6) to eliminate c_{pv} yields

$$-d\left(\frac{L_c q_s}{T}\right) = c_p d \ln \theta_d + q_s c_w d \ln T \tag{D.8}$$

Neglecting the last term in (D.8) we may integrate from the originial state (p, T, q_s, e_s , θ_d) to a state where $q_s \to 0$. Therefore, the equivalent potential temperature of a saturated parcel is given by

$$\theta_e = \theta_d \exp\left(L_c q_s / c_p T\right) \approx \theta \exp\left(L_c q_s / c_p T\right) \tag{D.9}$$

Equation (D. 9) may also be applied to an unsaturated parcel provided that the temperature used is the temperature that the parcel would have if brought to saturation by an adiabatic expansion.

[1] For a derivation, see Curry and Webster (1999, p. 108)

D.2 PSEUDOADIABATIC LAPSE RATE

Section 9.5.2 showed that from the first law of thermodynamics the lapse rate for a saturated parcel undergoing pseudoadiabatic ascent can be obtained from the formula

$$\frac{dT}{dz} + \frac{g}{c_p} = -\frac{L_c}{c_p}\left[\left(\frac{\partial q_s}{\partial T}\right)_p \frac{dT}{dz} - \left(\frac{\partial q_s}{\partial p}\right)_T \rho g\right] \tag{D.10}$$

Noting that $q_s \cong \varepsilon e_s/p$, where $\varepsilon = 0.622$ is the ratio of the molecular weight of water to that of dry air and utilizing (D.4), we can express the partial derivatives in (D.10) as

$$\left(\frac{\partial q_s}{\partial p}\right)_T \approx -\frac{q_s}{p} \quad \text{and} \quad \left(\frac{\partial q_s}{\partial T}\right)_p \approx \frac{\varepsilon}{p}\frac{\partial e_s}{\partial T} = \frac{\varepsilon^2 L_c e_s}{pRT^2} = \frac{\varepsilon L_c q_s}{RT^2}$$

Substitution into (D.10), and noting that $g/c_p = \Gamma_d$, then yields the desired result:

$$\Gamma_s \equiv -\frac{dT}{dz} = \Gamma_d \frac{[1 + L_c q_s/(RT)]}{\left[1 + \varepsilon L_c^2 q_s/\left(c_p RT^2\right)\right]}$$

APPENDIX E

Standard Atmosphere Data

Table E.1 *Geopotential Height versus Pressure*

Pressure (hPa)	Z (km)
1000	0.111
900	0.988
850	1.457
700	3.012
600	4.206
500	5.574
400	7.185
300	9.164
250	10.363
200	11.784
150	13.608
100	16.180
50	20.576
30	23.849
20	26.481
10	31.055

Table E.2 *Standard Atmosphere Temperature, Pressure, and Density as a Function of Geopotential Height*

Z (km)	Temperature (K)	Pressure (hPa)	Density (kg m^{-3})
0	288.15	1013,25	1.225
1	281.65	898.74	1.112
2	275.15	794.95	1.007
3	268.65	701.08	0.909
4	262.15	616.40	0.819
5	255.65	540.19	0.736
6	249.15	471.81	0.660
7	242.65	410.60	0.590
8	236.15	355.99	0.525
9	229.65	307.42	0.466
10	223.15	264.36	0.412
12	216.65	193.30	0.311
14	216.65	141.01	0.227
16	216.65	102.87	0.165
18	216.65	75.05	0.121
20	216.65	54.75	0.088
24	220.65	29.30	0.046
28	224.65	15.86	0.025
32	228.65	08.68	0.013

APPENDIX F

Symmetric Baroclinic Oscillations

A variant of the derivation of the Sawyer–Eliassen equation for forced transverse circulations in baroclinic zones (9.15) can be used to obtain an equation for free symmetric transverse oscillations, which can be used to derive an expression for the growth rate of unstable symmetric oscillations or the frequency of stable symmetric oscillations.

Suppose that the flow field is zonally symmetric so that $u_g = u_g(y, z)$ and $b = b(y, z)$. The ageostrophic (transverse) flow is given by the streamfunction $\psi(y, z)$ go that $v_a = -\partial\psi/\partial z;\ w_a = \partial\psi/\partial y$. Then from (9.12) $Q_2 = 0$ so that any transverse circulation is unforced and must arise from a departure from exact geostrophic balance. This can be simply represented by adding an acceleration term in the y-momentum equation so that (9.10) becomes

$$\frac{\partial}{\partial t}\left(-\frac{\partial^2\psi}{\partial z^2}\right) + f\frac{\partial u_g}{\partial z} + \frac{\partial b}{\partial y} = 0 \tag{F.1}$$

Then combining (9.11) and (9.13) as before and applying (F.1), we obtain

$$\frac{D}{Dt}\left[\frac{\partial}{\partial t}\left(\frac{\partial^2\psi}{\partial z^2}\right)\right] + N_s^2\frac{\partial^2\psi}{\partial y^2} + F^2\frac{\partial^2\psi}{\partial z^2} + 2S^2\frac{\partial^2\psi}{\partial y\partial z} = 0 \tag{F.2}$$

Neglecting terms quadratic in the streamfunction

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v_a \frac{\partial}{\partial y} + w_a \frac{\partial}{\partial z} \approx \frac{\partial}{\partial t}$$

then yields the desired result

$$\frac{\partial^2}{\partial t^2}\left(\frac{\partial^2 \psi}{\partial z^2}\right) + N_s^2 \frac{\partial^2 \psi}{\partial y^2} + F^2 \frac{\partial^2 \psi}{\partial z^2} + 2S^2 \frac{\partial^2 \psi}{\partial y \partial z} = 0 \tag{F.3}$$

ANSWERS TO SELECTED PROBLEMS

Chapter 1

1. $\alpha \approx (\Omega^2 a \sin 2\phi)/(2g)$ **2.** 35,830 km **3.** 0.35 N kg^{-1}
4. 556 m **5.** 1.46 cm to the right **6.** 6.25 m s^{-1}
7. Upward force $= 2 \times 10^3$ N; greater for westward travel **8.** 7.8 m eastward
10. 1 rad s^{-1}, 1.22 rad s^{-1}, 0.17 J **12.** 5.536 km, 5.070 km **13.** 3°C
14. 7.987 km **17.** 5.187 km

Chapter 2

1. −2 hPa/3 h **2.** −2°C/h **5.** 1.02 cm s^{-1} **7.** 0.725 kg m^{-3}
8. 135 J kg^{-1}, 16.4 m s^{-1} **11.** 287 s

Chapter 3

1. −1 m km^{-1} (−0.23 m s^{-1}) **2.** -10^{-3} m s^{-2} **3.** 940 hPa **5.** $V_{grad}/V_g = 2$
7. North, $R_t = 250$ km; west, $R_t = 500$ km; south, $R_t \to \infty$; east $R_t = 500$ km
8. North, 10.5 m s^{-1}; west and east, 12.1 m s^{-1}; south, 15 m s^{-1}
10. 25 m s^{-1}, 34° **11.** −1.5 °C h^{-1} **12.** 44 h
13. 238 km to the left
15. North, 9.4 m s^{-1}; west and east 11.5 m s^{-1}; south, 15 m s^{-1} **17.** 7.66 K
18. −852 km **20.** 2×10^{-5} s^{-1} **21.** 110% **23.** 0.96 cm s^{-1}

Chapter 4

1. -2×10^7 m^2 s^{-1}, -2×10^{-5} s^{-1} **2.** -5.5 m s^{-1} (antiyclonic) **3.** -2×10^{-5} s^{-1}
4. 9.4×10^{-5} s^{-1}, -8.4×10^{-6} s^{-1} **5.** 0 (annulus), 5×10^{-5} s^{-1} (inner cylinder)
7. -7.2 m^2 s^{-2} **9.** $h(r) = H + \Omega^2 r^2/(2g)$
12. 10^{-5} s^{-1}, -1.29×10^{-5} s^{-1}, -1546 km
14. 0.71° equatorward, 138.6 km upward

Chapter 5

4. 293 s, 1.7 cm **5.** 3.21×10^3 kg m^{-1} s^{-1} (mixed layer) **6.** $De = 14.05$ m
8. $w_{max} = 9.6 \times 10^{-3}$ mm s^{-1}, $v_{max} = 0.75$ mm s^{-1} **9.** 155 days
10. $K_m = 11$ m^2 s^{-1}

Chapter 6

3. 2.84×10^{-6} s^{-1} **4.** $\omega(p) = -(k^2c^2 p_0/f_0\pi)\sin(\pi p/p_0)\cos[k(x - ct)]$
5. $k^2 = \sigma[(f_0\pi)/(\sigma p_0)]^2$ **8.** $W_0 = -cf_0U_0\pi/(\sigma p_0)$
9. $\omega(p) = (Vp_0/f_0\pi)[k^2(c - U) + \beta]\sin(\pi p/p_0)\cos[k(x - ct)]$
10. $c = U - \beta[k^2 + (f_0^2\pi^2)/(\sigma p_0^2]^{-1}$ **11.** 0.047 Pa s^{-1}

Chapter 7

2. $B_r = A\cos kx_0$, $B_i = -A\sin kx_0$, $c_r = \nu/k$, $c_i = \alpha/k$
4. $u' = p'(c - \overline{u})\big/(\gamma\overline{p})$, $\rho' = p'\big/(c - \overline{u})^2$ **7.** $u' = h'(c - \overline{u})\big/H$
9. $\rho_0\overline{u'w'} = -\rho_0 m|A|^2\big/(2k)$
12. $u' = \left(0.66 \text{ m s}^{-1}\right)\sin(kx + mz)$, $w' = -\left(0.75 \text{ m s}^{-1}\right)\sin(kx + mz)$
16. -24.3 m s^{-1} **17.** -0.633 cm s^{-1}

Chapter 8

1. 6.04×10^4 s (about 17 hrs.)
2. $\psi_3' = -\psi_1' = \text{Acos}\left[k(x - c_r t)\right]$
$\omega_2' = -2kf_0\beta\text{Asin}\,[\text{k(x} - \text{c}_\text{r}\text{t)}]\,/\left[\sigma\delta p(k^2 + 2\lambda^2)\right]$
3. $\omega_2' = \delta p f_0^{-1}\lambda\beta\left(\sqrt{3} - 1\right)\text{Asin}\,[k\,(x - c_r t)]/3$
6. ψ_1' lags ψ_3' by 65.5° **7.** $|B| = \left|A\left(\sqrt{2} - 1\right)^{1/2}\right| = 6.44 \times 10^6$ m^2 s^{-1}
9. 2.38×10^9 J m^{-2} **13.** $\overline{v'T'}$ is independent of Λ.

Chapter 9

5. 344 K **6.** 1.42 m s^{-1} **8.** $r = 135.7$ km

Chapter 10

3. 134 m s^{-1} **8.** 6.25×10^5 kg s^{-2} **11.** 0.245 cm s^{-1}

Chapter 11

1. 9.6×10^{-8} m s^{-1} **2.** $2u_0 l^2 > \beta$

Chapter 12

1. 6697 m **2.** 87.6 m s^{-1}, 48.7 m s^{-1}, 28.0 m s^{-1}
6. $|\Phi'| = 200$ m^2 s^{-2}, $|w'| = 1.57$ mm s^{-1} **7.** 3.92×10^{-3} kg m^{-1} s^{-1}
8. $\hat{w} = \hat{v}(-imy)\left(\nu^2/N^2\right)\exp\left[-\beta|m|y^2/(2N^2)\right]$ **9.** $\pm 6.14^\circ$

Chapter 13

5. 1.3%, 5%, 19%, 59% **6.** 1.5% per time step

BIBLIOGRAPHY

Acheson, D. J., 1990: *Elementary Fluid Dynamics*. Oxford University Press, New York.

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. Academic Press, Orlando, FL.

Andrews, D. G., and M. E. McIntyre, 1976. Planetary waves in horizontal and vertical shear: The generalized Eliassen–Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–2048.

Arakawa, A., and W. Schubert, 1974. Interaction of a cumulus cloud ensemble with the large scale environment. Part I. *J. Atmos. Sci.*, **31**, 674–701.

Arya, S. P., 2001: *Introduction to Micrometeorology* (2nd Ed.). Academic Press, Orlando, FL.

Baer, F., 1972: An alternate scale representation of atmospheric energy spectra. *J. Atmos. Sci.*, **29**, 649–664.

Battisti, D. S., E. S. Sarachik, and A. C. Hirst, 1992: A consistent model for the large-scale steady surface atmospheric circulation in the tropics. *J. Climate*, **12**, 2956–2964.

Blackburn, M., 1985: Interpretation of ageostrophic winds and implications for jet stream maintenance. *J. Atmos. Sci*., **42**, 2604–2620.

Blackmon, M. L., J. M. Wallace, N.-C. Lau, and S. L. Mullen, 1977: An observational study of the northern hemisphere wintertime circulation. *J. Atmos. Sci.*, **34**, 1040–1053.

Bluestein, H., 1993: *Synoptic-Dynamic Meteorology in Midlatitudes,* Vol.II. Oxford University Press, New York.

Bourne, D. E., and P. C., Kendall, 1968: *Vector Analysis*. Allyn & Bacon, Boston.

Brasseur, G. and S. Solomon, 1986: *Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere* (2nd Ed.). Reidel.

Brown, R. A., 1970: A secondary flow model for the planetary boundary layer. *J. Atmos. Sci.*, **27**, 742–757.

Brown, R. A., 1991: *Fluid Mechanics of the Atmosphere*. Academic Press, Orlando, FL.

Chang, C. P., 1970: Westward propagating cloud patterns in the Tropical Pacific as seen from time-composite satellite photographs. *J. Atmos. Sci.*, **27**, 133–138.

Chapman, S. and R. S. Lindzen, 1970: *Atmospheric Tides: Thermal and Gravitational*. Reidel.

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, **4**, 135–163.

Charney, J. G., 1948: On the scale of atmospheric motions. *Geofys. Publ.* **17**(2), 1–17.

Charney, J. G., and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.*, **36**, 1205–1216.

Charney, J. G., and A. Eliassen, 1949: A numerical method for predicting the perturbations of the middle latitude westerlies. *Tellus,* **1**(2), 38–54.

Curry, J. A., and P. J. Webster, 1999: *Thermodynamics of Atmospheres and Oceans*. Academic Press, San Diego.

Dai, A., T. M. L. Wigley, B. A. Boville, J. T. Kiehl, and L. E. Buja, 2001: Climates of the Twentieth and Twenty-First Centuries Simulated by the NCAR Climate System Model. *Climate,* **14**, 485–519.

Dunkerton, T. J., 2003: Middle Atmosphere: Quasi-biennial Oscillation. In *Encyclopedia of Atmospheric Sciences,* ed. J. R. Holton, J. A. Curry and J. A. Pyle. Academic Press, London.

Durran, D. R., 1990: Mountain waves and downslope winds. In *Atmospheric Processes over Complex Terrain*, ed. W. Blumen. American Meteorological Society, 59–82.

Durran, D.R., 1993: Is the Coriolis force really responsible for the inertial oscillation? *Bull. Am. Meteorol. Soc.* **74**(11), 2179–2184.

Durran, D. R., 1999: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*. Springer, New York.

Durran, D. R., and L. W. Snellman, 1987: The diagnosis of synoptic-scale vertical motion in an operational environment. *Wea. Forecasting*, **1**, 17–31.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus* **1**(3), 33–52.

Eliassen, A., 1990: Transverse circulations in frontal Zones. In *Extratropical Cyclones, The Erik Palmén Memorial Volume,* ed. C. W. Newton and E. O. Holopainen. American Meteorological Society, Boston, 155–164.

Emanuel, K. A., 1988: Toward a general theory of hurricanes. *Am. Sci.*, **76**, 370–379.

Emanuel, K., 1994: *Atmospheric Convection*. Oxford University Press, New York.

Emanuel, K., 2000: Quasi-equilibrium thinking. In *General Circulation Model Development,* ed. D. A. Randall. Academic Press, New York, 225–255.

Fleming, E. L., S. Chandra, J. J. Barnett, and M. Corey, 1990: Zonal mean temperature, pressure, zonal wind and geopotential height as functions of latitude. *Adv. Space Res.*, **10**(12), 11–59.

Garratt, J. R., 1992: *The Atmospheric Boundary Layer*. Cambridge University Press, Cambridge.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press, New York.

Held, I. M., 1983: Stationary and quasi-stationary eddies in the extratropical troposphere: Theory. In *Large-Scale Dynamical Processes in the Atmosphere*, ed. B.J. Hoskins and R. Pearce. Academic Press, New York, 127–168.

Hess, S. L. 1959. *Introduction to Theoretical Meteorology*. Holt, New York.

Hide, R., 1966: On the dynamics of rotating fluids and related topics in geophysical fluid mechanics. *Bull. Am. Meteorol. Soc.*, **47**, 873–885.

Hildebrand, F. B., 1976: *Advanced Calculus for Applications* (2nd Ed.). Prentice Hall, New York.

Holton, J. R., 1986: Meridional distribution of stratospheric trace constituents. *J. Atmos. Sci.*, **43**, 1238–1242.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.,* **33**, 403–439.

Horel, J. D., and J. M. Wallace, 1981: Planetary scale atmospheric phenomena associated with the southern oscillation. *Mon. Wea. Rev*., **109**, 813–829.

Hoskins, B. J., 1975: The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.,* **32**, 233–242.

Hoskins, B. J., 1982: The mathematical theory of frontogenesis. *Annu. Rev. Fluid Mech*., **14**, 131–151.

Hoskins, B. J., 1983: Dynamical processes in the atmosphere and the use of models. *Quart. J. Roy. Meteor. Soc*., **109**, 1–21.

Hoskins, B. J., and F. P. Bretherton, 1972: Atmospheric frontogenesis models: Mathematical formulation and solution. *J. Atmos. Sci*., **29**, 11–37.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteorol. Soc*., **111**, 877–946.

Houze, R. A., Jr., 1993: *Cloud Dynamics*. Academic Press, San Diego.

James, I. N., 1994: *Introduction to Circulating Atmospheres*. Cambridge University Press, Cambridge.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge.

Klemp, J. B., 1987: Dynamics of tornadic thunderstorms. *Annu. Rev. Fluid Mech*., **19**, 369–402.

Lim, G. H., J. R. Holton, and J. M. Wallace, 1991: The structure of the ageostrophic wind field in baroclinic waves. *J. Atmos. Sci.,* **48**, 1733–1745.

Lindzen, R. S., E. S. Batten, and J. W. Kim, 1968: Oscillations in atmospheres with tops. *Mon. Wea. Rev.* **96**, 133–140.

Lorenz, E. N., 1960: Energy and numerical weather prediction. *Tellus*, **12**, 364–373.

Lorenz, E. N., 1967: *The Nature and Theory of the General Circulation of the Atmosphere*. World Meteorological Organization, Geneva.

Lorenz, E. N., 1984: Some aspects of atmospheric predictability. In *Problems and Prospects in Long and Medium Range Weather Forecasting*, ed. D. M. Burridge and E. Källén Springer-Verlag, New York, 1–20.

Madden, R. A., 2003: Intraseasonal oscillation (Madden–Julian oscillation). In *Encyclopedia of Atmospheric Sciences,* ed. J. R. Holton, J. A. Curry, and J. A. Pyle. Academic Press, London, 2334–2338.

Madden, R. A., and P. R. Julian, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. *Atmos. Sci.,* **29**, 1109–1123.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteorol. Soc. Japan*, **44**, 25–43.

Nappo, C. J., 2002: *An Introduction to Atmospheric Gravity Waves*. Academic Press, San Diego.

Naujokat, B., 1986: An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *J. Atmos. Sci.*, **43**, 1873–1877.

Norton, W. A., 2003: Middle atmosphere: Transport circulation. In *Encyclopedia of Atmospheric Sciences,* ed. J. R. Holton, J. A. Curry, and J. A. Pyle. Academic Press, London, 1353–1358.

Oort, A. H., 1983: Global atmospheric circulation statistics, 1958–1973. *NOAA Professional Paper 14*, U. S. Government Printing Office.

Oort, A. H., and J. P. Peixoto, 1974: The annual cycle of the energetics of the atmosphere on a planetary scale. *J. Geophys. Res*., **79**, 2705–2719.

Oort, A. H., and J. P. Peixoto, 1983: Global angular momentum and energy balance requirements from observations. *Adv. Geophys.,* **25**, 355–490.

Ooyama, K., 1969: Numerical simulation of the life cycle of tropical cyclones. *J. Atmos. Sci*., **26**, 3–40.

Palmén, E., and C. W. Newton, 1969: *Atmospheric Circulation Systems*. Academic Press, London.

Palmer, T. N., 1993: Extended-range atmospheric prediction and the Lorenz model. *Bull. Am. Meteor. Soc.,* **74**, 49–65.

Panofsky, H. A., and J. Dutton A., 1984: *Atmospheric Turbulence*. Wiley New York.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics* (2nd Ed.). Springer-Verlag, New York.

Philander, S. G.., 1990: *El Niño, La Niña, and the Southern Oscillation*. Academic Press, New York.

Phillips, N. A., 1956: The general circulation of the atmosphere: A numerical experiment. *Quart. J. Roy. Meteorol. Soc.*, **82**, 123–164.

Phillips, N. A. 1963: Geostrophic motion. *Rev. Geophys.*, **1**, 123–176.

Pierrehumbert, R. T., and K. L. Swanson, 1995: Baroclinic Instability. *Annu. Rev. Fluid Mech.,* **27**, 419–467.

Plumb, R. A., 1982: The circulation of the middle atmosphere. *Aust. Meteorol. Mag.,* **30**, 107–121.

Randall, D. A., 2000: *General Circulation Model Development*. Academic Press, San Diego.

Reed, R. J., D. C. Norquist, and E. E. Recker, 1977: The structure and properties of African wave disturbances as observed during Phase III of GATE. *Mon. Wea. Rev.*, **105**, 317–333.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press (reprt. Dover).

Richtmyer, R. D., and K. W. Morton, 1967: *Difference Methods for Initial Value Problems* (2nd Ed.) Wiley (Interscience), New York.

Riehl, H., and J. S. Malkus, 1958. On the heat balance of the equatorial trough zone. *Geophysica*, **6**, 503–538.

Salby, M. L., 1996: *Fundamentals of Atmospheric Physics*. Academic Press, San Diego.

Salby, M. L., H. H. Hendon, K. Woodberry, and K. Tanaka, 1991: Analysis of global cloud imagery from multiple satellites. *Bull. Am. Meteorol. Soc.*, **72**, 467–480.

Sanders, F., and B. J. Hoskins, 1990: An easy method for estimation of Q-vectors from weather maps. *Wea. Forecasting*, **5**, 346–353.

Sawyer, J. S., 1956: The vertical circulation at meteorological fronts and its relation to frontogenesis. *Proc. Roy. Soc. A*, **234**, 346–362.

Schubert, S., C.-K. Park, W. Higgins, S. Moorthi, and M. Suarez, 1990: *An Atlas of ECMWF Analyses (1980–1987) Part I — First Moment Quantities*. NASA Technical Memorandum 100747.

Scorer, R. S., 1958: *Natural Aerodynamics*, Pergamon Press, New York.

Shine, K. P., 1987: The middle atmosphere in the absence of dynamical heat fluxes. *Quart. J. Roy. Meteor. Soc.,* **113**, 603–633.

Simmons, A. J., and L. Bengtsson, 1984: Atmospheric general circulation models: Their design and use for climate studies. In *The Global Climate*, ed. J. T. Houghton. Cambridge University Press, Cambridge, 37–62.

Simmons, A. J., D. M. Burridge, M. Jarraud, C. Girard, and W. Wergen, 1989: The ECMWF medium-range prediction models: Development of the numerical formulations and the impact of increased resolution. *Meteorol. Atmos. Phys.*, **40**, 28–60.

Simmons, A. J., and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteorol. Soc.,* **128**, 647–678.

Sinclair, P. C., 1965: On the rotation of dust devils. *Bull. Am. Meteorol. Soc.,* **46**, 388–391.

Smagorinsky, J., 1967: The role of numerical modeling. *Bull. Am. Meteorol. Soc.*, **46**, 89–93.

Smith, R. B., 1979: The influence of mountains on the atmosphere. *Adv. Geophys.*, **21**, 87–230.

Stevens, D. E., and R. S. Lindzen, 1978: Tropical wave–CISK with a moisture budget and cumulus friction. *J. Atmos. Sci.*, **35**, 940–961.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers.

Thorpe, A. J., and C. H. Bishop, 1995: Potential vorticity and the electrostatics analogy: Ertel-Rossby formulation. *Quarti. J. Ref. Meteorol. Soc.,* **121**, 1477–1495.

Trenberth, K. E., 1991: General characteristics of El Niño–Southern Oscillation. *ENSO Teleconnections Linking Worldwide Climate Anomalies: Scientific Basis and Societal Impact*, ed. M. Glantz, R. Katz, and N. Nichols, Cambridge University Press, Cambridge, 13–42.

Turner, J. S., 1973: *Buoyancy Effects in Fluids*, Cambridge University Press, Cambridge.

U.S. Government Printing Office, 1976: *U. S. Standard Atmosphere, 1976*, U. S. Government Printing Office.

Wallace, J. M., 1971: Spectral studies of tropospheric wave disturbances in the tropical Western Pacific. *Rev. Geophys.*, **9**, 557–612.

Wallace, J. M., 2003: General circulation: Overview. In *Encyclopedia of Atmospheric Sciences,* ed. J. R. Holton, J. A. Curry, and J. A. Pyle. Academic Press, London, 821–829.

Wallace, J. M., and P. V. Hobbs, 1977: *Atmospheric Science: An Introductory Survey*. Academic Press, New York.

Wallace, J. M., and V. E. Kousky, 1968: Observational evidence of Kelvin waves in the tropical stratosphere. *J. Atmos. Sci*., **25**, 900–907.

Warsh, K. L., K. L. Echternacht, and M. Garstang, 1971: Structure of near-surface currents east of Barbados. *J. Phys. Oceanog*., **1**, 123–129.

Washington, W. M., and C. L. Parkinson, 1986: *An Introduction to Three-Dimensional Climate Modeling*. University Science Books, Mill Valley, CA.

Webster P. J., 1983: The large-scale structure of the tropical atmosphere. In *Large-Scale Dynamical Processes in the Atmosphere*, ed. B. J. Hoskins and R. Pearce. Academic Press, New York, 235–275.

Webster, P. J., and H.-R. Chang, 1988: Equatorial energy accumulation and emanation regions: Impacts of a zonally varying basic state. *J. Atmos. Sci.*, **45**, 803–829.

Webster, P. J., and J. Fasullo, 2003: Monsoon: Dynamical theory. In *Encyclopedia of Atmospheric Sciences,* ed. J. R. Holton, J. A. Curry, and J. A. Pyle. Academic Press, London, 1370–1386.

Williams, J., and S. A. Elder, 1989: *Fluid Physics for Oceanographers and Physicists*, Pergamon Press, New York.

Williams, K. T., 1971: A statistical analysis of satellite-observed trade wind cloud clusters in the western North Pacific. *Atmospheric Science Paper No. 161*, Dept. of Atmospheric Science, Colorado State Univ., Fort Collins, CO.

Yanai, M., and T. Maruyama, 1966: Stratospheric wave disturbances propagating over the equatorial Pacific. *J. Meteor. Soc. Jpn.,* **44**.

INDEX

B

C

D

E

F

G

H

I

J

K

T

V

W

Z

International Geophysics Series

EDITED BY

RENATA DMOWSKA

Division of Applied Science
Harvard University
Cambridge, Massachusetts

JAMES R. HOLTON

Department of Atmospheric Sciences
University of Washington
Seattle, Washington

H. THOMAS ROSSBY

Graduate School of Oceanography
University of Rhode Island
Narragansett, Rhode Island

Volume 1 BENO GUTENBERG. Physics of the Earth's Interior. 1959*

Volume 2 JOSEPH W. CHAMBERLAIN. Physics of the Aurora and Airglow. 1961*

Volume 3 S. K. RUNCORN (ed.). Continental Drift. 1962*

Volume 4 C. E. JUNGE. Air Chemistry and Radioactivity. 1963*

Volume 5 ROBERT G. FLEAGLE AND JOOST A. BUSINGER. An Introduction to Atmospheric Physics. 1963*

Volume 6 L. DEFOUR AND R. DEFAY. Thermodynamics of Clouds. 1963*

Volume 7 H. U. ROLL. Physics of the Marine Atmosphere. 1965*

Volume 8 RICHARD A. CRAIG. The Upper Atmosphere: Meteorology and Physics. 1965*

Volume 9 WILLIS L. WEBB. Structure of the Stratosphere and Mesophere. 1966*

* Out of print

Volume 10 MICHELE CAPUTO. The Gravity Field of the Earth from Classical and Modern Methods. 1967*

Volume 11 S. MATSUSHITA AND WALLACE H. CAMPBELL (eds.). Physics of Geomagnetic Phenomena (in two volumes). 1967*

Volume 12 K. YA KONDRATYEV. Radiation in the Atmosphere. 1969*

Volume 13 E. PALMÅN AND C. W. NEWTON. Atmospheric Circulation Systems: Their Structure and Physical Interpretation. 1969*

Volume 14 HENRY RISHBETH AND OWEN K. GARRIOTT. Introduction to Ionospheric Physics. 1969*

Volume 15 C. S. RAMAGE. Monsoon Meteorology. 1971*

Volume 16 JAMES R. HOLTON. An Introduction to Dynamic Meteorology. 1972*

Volume 17 K. C. YEH AND C. H. LIU. Theory of Ionospheric Waves. 1972*

Volume 18 M. I. BUDYKO. Climate and Life. 1974*

Volume 19 MELVIN E. STERN. Ocean Circulation Physics. 1975

Volume 20 J. A. JACOBS. The Earth's Core. 1975*

Volume 21 DAVID H. MILLER. Water at the Surface of the Earth: An Introduction to Ecosystem Hydrodynamics. 1977

Volume 22 JOSEPH W. CHAMBERLAIN. Theory of Planetary Atmospheres: An Introduction to Their Physics and Chemistry. 1978*

Volume 23 JAMES R. HOLTON. An Introduction to Dynamic Meteorology, Second Edition. 1979*

Volume 24 ARNETT S. DENNIS. Weather Modification by Cloud Seeding. 1980*

Volume 25 ROBERT G. FLEAGLE AND JOOST A. BUSINGER. An Introduction to Atmospheric Physics, Second Edition. 1980*

Volume 26 KUO-NAN LIOU. An Introduction to Atmospheric Radiation. 1980*

Volume 27 DAVID H. MILLER. Energy at the Surface of the Earth: An Introduction to the Energetics of Ecosystems. 1981

Volume 28 HELMUT G. LANDSBERG. The Urban Climate. 1991

Volume 29 M. I. BUDKYO. The Earth's Climate: Past and Future. 1982*

Volume 30 ADRIAN E. GILL. Atmosphere-Ocean Dynamics. 1982

Volume 31 PAOLO LANZANO. Deformations of an Elastic Earth. 1982*

Volume 32 RONALD T. MERRILL AND MICHAEL W. MCELHINNY. The Earth's Magnetic Field: Its History, Origin, and Planetary Perspective. 1983*

Volume 33 JOHN S. LEWIS AND RONALD G. PRINN. Planets and Their Atmospheres: Origin and Evolution. 1983

Volume 34 ROLF MEISSNER. The Continental Crust: A Geophysical Approach. 1986

Volume 35 M. U. SAGITOV, B. BODKI, V. S. NAZARENKO, AND KH. G. TADZHIDINOV. Lunar Gravimetry. 1986

Volume 36 JOSEPH W. CHAMBERLAIN AND DONALD M. HUNTEN. Theory of Planetary Atmospheres, Second Edition. 1987

Volume 37 J. A. JACOBS. The Earth's Core, Second Edition. 1987*

Volume 38 J. R. APEL. Principles of Ocean Physics. 1987

Volume 39 MARTIN A. UMAN. The Lightning Discharge. 1987*

Volume 40 DAVID G. ANDREWS, James R. HOLTON, and CONWAY B. LEOVY. Middle Atmosphere Dynamics. 1987

Volume 41 PETER WARNECK. Chemistry of the Natural Atmosphere. 1988*

Volume 42 S. PAL ARYA. Introduction to Micrometeorology. 1988*

Volume 43 MICHAEL C. KELLEY. The Earth's Ionosphere. 1989*

Volume 44 WILLIAM R. COTTON AND RICHARD A. ANTHES. Storm and Cloud Dynamics. 1989

Volume 45 WILLIAM MENKE. Geophysical Data Analysis: Discrete Inverse Theory, Revised Edition. 1989

Volume 46 S. GEORGE PHILANDER. El Niño, La Niña, and the Southern Oscillation. 1990

Volume 47 ROBERT A. BROWN. Fluid Mechanics of the Atmosphere. 1991

Volume 48 JAMES R. HOLTON. An Introduction to Dynamic Meterology, Third Edition. 1992

Volume 49 ALEXANDER A: KAUFMAN. Geophysical Field Theory and Method.
Part A: Gravitational, Electric, and Magnetic Fields. 1992*
Part B: Electromagnetic Fields I. 1994*
Part C: Electromagnetic Fields II. 1994*

Volume 50 SAMUEL S. BUTCHER, GORDON H. ORIANS, ROBERT J. CHARLSON, AND GORDON V. WOLFE. Global Biogeochemical Cycle. 1992

Volume 51 BRAIAN EVANS AND TENG-FONG WONG. Fault Mechanics and Transport Properties of Rocks. 1992

Volume 52 ROBERT E. HUFFMAN. Atmospheric Ultraviolet Remote Sensing. 1992

Volume 53 ROBERT A. HOUZE, JR. Cloud Dynamics. 1993

Volume 54 PETER V. HOBBS. Aerosol-Cloud-Climate Interactions. 1993

Volume 55 S. J. GIBOWICZ AND A. KIJKO. An Introduction to Mining Seismology. 1993

Volume 56 DENNIS L. HARTMANN. Global Physical Climatology. 1994

Volume 57 MICHAEL P. RYAN. Magmatic Systems. 1994

Volume 58 THORNE LAY AND TERRY C. WALLACE. Modern Global Seismology. 1995

Volume 59 DANIEL S. WILKS. Statistical Methods in the Atmospheric Sciences. 1995

Volume 60 FREDERIK NEBEKER. Calculating the Weather. 1995

Volume 61 MURRY L. SALBY. Fundamentals of Atmospheric Physics. 1996

Volume 62 JAMES P. MCCALPIN. Paleoseismology. 1996

Volume 63 RONALD T. MERRILL, MICHAEL W. MCELHINNY, AND PHILLIP L. MCFADDEN. The Magnetic Field of the Earth: Paleomagnetism, the Core, and the Deep Mantle. 1996

Volume 64 NEIL D. OPDYKE AND JAMES E. T. CHANNELL. Magnetic Stratigraphy. 1996

Volume 65 JUDITH A. CURRY AND PETER J. WEBSTER. Thermodynamics of Atmospheres and Oceans. 1998

Volume 66 LAKSHMI H. KANTHA AND CAROL ANNE CLAYSON. Numerical Models of Oceans and Oceanic Processes. 2000

Volume 67 LAKSHMI H. KANTHA AND CAROL ANNE CLAYSON. Small Scale Processes in Geophysical Fluid Flows. 2000

Volume 68 RAYMOND S. BRADLEY. Paleoclimatology, Second Edition. 1999

Volume 69 LEE-LUENG FU AND ANNY CAZANAVE. Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications. 2000

Volume 70 DAVID A. RANDALL. General Circulation Model Development: Past, Present, and Future. 2000

Volume 71 PETER WARNECK. Chemistry of the Natural Atmosphere, Second Edition. 2000

Volume 72 MICHAEL C. JACOBSON, ROBERT J. CHARLSON, HENNING RODHE, AND GORDON H. ORIANS. Earth System Science: From Biogeochemical Cycles to Global Change. 2000

Volume 73 MICHAEL W. MCELHINNY AND PHILLIP L. MCFADDEN. Paleomagnetism: Continents and Oceans. 2000

Volume 74 ANDREW E. DESSLER. The Chemistry and Physics of Stratospheric Ozone. 2000

Volume 75 BRUCE DOUGLAS, MICHAEL KEARNEY, AND STEPHEN LEATHERMAN. Sea Level Rise: History and Consequences. 2000

Volume 76 ROMAN TEISSEYRE AND EUGENISZ MAJEWSKI. Earthquake Thermodynamics and Phase Transformations in the Interior. 2001

Volume 77 GEROLD SIEDLER, JOHN CHURCH, AND JOHN GOULD. Ocean Circulation and Climate: Observing and Modelling the Global Ocean. 2001

Volume 78 ROGER A. PIELKE SR. Mesoscale Meteorological Modeling, Second Edition. 2001

Volume 79 S. PAL ARYA. Introduction to Micrometeorology. 2001

Volume 80 BARRY SALTZMAN. Dynamical Paleoclimatology: Generalized Theory of Global Climate Change. 2002

Volume 81A WILLIAM H. K. LEE, HIROO KANAMORI, PAUL JENNINGS, AND CARL KISSLINGER. International Handbook of Earthquake and Engineering Seismology, Part A. 2002

Volume 81B WILLIAM H. K. LEE, HIROO KANAMORI, PAUL JENNINGS, AND CARL KISSLINGER. International Handbook of Earthquake and Engineering Seismology, Part B. 2003

Volume 82 GORDON G. SHEPHERD. Spectral Imaging of the Atmosphere. 2002

Volume 83 ROBERT P. PEARCE. Meteorology at the Millennium. 2001

Volume 84 KUO-NAN LIOU. An Introduction to Atmospheric Radiation, Second Edition. 2002

Volume 85 CARMEN J. NAPPO. An Introduction to Atmospheric Gravity Waves. 2002

Volume 86 MICHAEL E. EVANS AND FRIEDRICH HELLER. Environmental Magnetism: Principles and Applications of Enviromagnetics. 2003

Volume 87 JOHN S. LEWIS. Physics and Chemistry of the Solar System, Second Edition. 2003

Volume 88 JAMES R. HOLTON. An Introduction to Dynamic Meteorology, Fourth Edition. 2004

ELSEVIER SCIENCE CD-ROM LICENSE AGREEMENT

PLEASE READ THE FOLLOWING AGREEMENT CAREFULLY BEFORE USING THIS CD-ROM PRODUCT. THIS CD-ROM PRODUCT IS LICENSED UNDER THE TERMS CONTAINED IN THIS CD-ROM LICENSE AGREEMENT ("Agreement"). BY USING THIS CD-ROM PRODUCT, YOU, AN INDIVIDUAL OR ENTITY INCLUDING EMPLOYEES, AGENTS AND REPRESENTATIVES ("You" or "Your"), ACKNOWLEDGE THAT YOU HAVE READ THIS AGREEMENT, THAT YOU UNDERSTAND IT, AND THAT YOU AGREE TO BE BOUND BY THE TERMS AND CONDITIONS OF THIS AGREEMENT. ELSEVIER SCIENCE INC. ("Elsevier Science") EXPRESSLY DOES NOT AGREE TO LICENSE THIS CD-ROM PRODUCT TO YOU UNLESS YOU ASSENT TO THIS AGREEMENT. IF YOU DO NOT AGREE WITH ANY OF THE FOLLOWING TERMS, YOU MAY, WITHIN THIRTY (30) DAYS AFTER YOUR RECEIPT OF THIS CD-ROM PRODUCT RETURN THE UNUSED CD-ROM PRODUCT AND ALL ACCOMPANYING DOCUMENTATION TO ELSEVIER SCIENCE FOR A FULL REFUND.

DEFINITIONS

As used in this Agreement, these terms shall have the following meanings:

"Proprietary Material" means the valuable and proprietary information content of this CD-ROM Product including all indexes and graphic materials and software used to access, index, search and retrieve the information content from this CD-ROM Product developed or licensed by Elsevier Science and/or its affiliates, suppliers and licensors.

"CD-ROM Product" means the copy of the Proprietary Material and any other material delivered on CD-ROM and any other human-readable or machine-readable materials enclosed with this Agreement, including without limitation documentation relating to the same.

OWNERSHIP

This CD-ROM Product has been supplied by and is proprietary to Elsevier Science and/or its affiliates, suppliers and licensors. The copyright in the CD-ROM Product belongs to Elsevier Science and/or its affiliates, suppliers and licensors and is protected by the national and state copyright, trademark, trade secret and other intellectual property laws of the United States and international treaty provisions, including without limitation the Universal Copyright Convention and the Berne Copyright Convention. You have no ownership rights in this CD-ROM Product.

TERM

This Agreement will remain in effect until terminated pursuant to the terms of this Agreement. You may terminate this Agreement at any time by removing from Your system and destroying the CD-ROM Product. Unauthorized copying of the CD-ROM Product, including without limitation, the Proprietary Material and documentation, or otherwise failing to comply with the terms and conditions of this Agreement shall result in automatic termination of this license and will make available to Elsevier Science legal remedies. Upon termination of this Agreement, the license granted herein will terminate and You must immediately destroy the CD-ROM Product and accompanying documentation. All provisions relating to proprietary rights shall survive termination of this Agreement.

LIMITED WARRANTY AND LIMITATION OF LIABILITY

NEITHER ELSEVIER SCIENCE NOR ITS LICENSORS REPRESENT OR WARRANT THAT THE INFORMATION CONTAINED IN THE PROPRIETARY MATERIALS IS COMPLETE OR FREE FROM ERROR, AND NEITHER ASSUMES, AND BOTH EXPRESSLY DISCLAIM, ANY LIABILITY TO ANY PERSON FOR ANY LOSS OR DAMAGE CAUSED BY ERRORS OR OMISSIONS IN THE PROPRIETARY MATERIAL, WHETHER SUCH ERRORS OR OMISSIONS RESULT FROM NEGLIGENCE, ACCIDENT, OR ANY OTHER CAUSE. IN ADDITION, NEITHER ELSEVIER SCIENCE NOR ITS LICENSORS MAKE ANY REPRESENTATIONS OR WARRANTIES, EITHER EXPRESS OR IMPLIED, REGARDING THE PERFORMANCE OF YOUR NETWORK OR COMPUTER SYSTEM WHEN USED IN CONJUNCTION WITH THE CD-ROM PRODUCT.

If this CD-ROM Product is defective, Elsevier Science will replace it at no charge if the defective CD-ROM Product is returned to Elsevier Science within sixty (60) days (or the greatest period allowable by applicable law) from the date of shipment.

Elsevier Science warrants that the software embodied in this CD-ROM Product will perform in substantial compliance with the documentation supplied in this CD-ROM Product. If You report significant defect in performance in writing to Elsevier Science, and Elsevier Science is not able to correct same within sixty (60) days after its receipt of Your notification, You may return this CD-ROM Product, including all copies and documentation, to Elsevier Science and Elsevier Science will refund Your money.

YOU UNDERSTAND THAT, EXCEPT FOR THE 60-DAY LIMITED WARRANTY RECITED ABOVE, ELSEVIER SCIENCE, ITS AFFILIATES,

LICENSORS, SUPPLIERS AND AGENTS, MAKE NO WARRANTIES, EXPRESSED OR IMPLIED, WITH RESPECT TO THE CD-ROM PRODUCT, INCLUDING, WITHOUT LIMITATION THE PROPRIETARY MATERIAL, AND SPECIFICALLY DISCLAIM ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

If the information provided on this CD-ROM contains medical or health sciences information, it is intended for professional use within the medical field. Information about medical treatment or drug dosages is intended strictly for professional use, and because of rapid advances in the medical sciences, independent verification of diagnosis and drug dosages should be made.

IN NO EVENT WILL ELSEVIER SCIENCE, ITS AFFILIATES, LICENSORS, SUPPLIERS OR AGENTS, BE LIABLE TO YOU FOR ANY DAMAGES, INCLUDING, WITHOUT LIMITATION, ANY LOST PROFITS, LOST SAVINGS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES, ARISING OUT OF YOUR USE OR INABILITY TO USE THE CD-ROM PRODUCT REGARDLESS OF WHETHER SUCH DAMAGES ARE FORESEEABLE OR WHETHER SUCH DAMAGES ARE DEEMED TO RESULT FROM THE FAILURE OR INADEQUACY OF ANY EXCLUSIVE OR OTHER REMEDY.

U.S. GOVERNMENT RESTRICTED RIGHTS

The CD-ROM Product and documentation are provided with restricted rights. Use, duplication or disclosure by the U.S. Government is subject to restrictions as set forth in subparagraphs (a) through (d) of the Commercial Computer Restricted Rights clause at FAR 52.22719 or in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.2277013, or at 252.2117015, as applicable. Contractor/Manufacturer is Elsevier Science Inc., 655 Avenue of the Americas, New York, NY 10010-5107, USA.

GOVERNING LAW

This Agreement shall be governed by the laws of the State of New York, USA. In any dispute arising out of this Agreement, you and Elsevier Science each consent to the exclusive personal jurisdiction and venue in the state and federal courts within New York County, New York, USA.