Fluid Dynamics of the Atmosphere and Ocean

Geoffrey K. Vallis
University of Exeter

ECMM719

Exeter, Autumn 2018

Contents

Preface

October 4, 2018

These are a set of lecture notes for ECMM719, Fluid Dynamics of the Atmosphere and Ocean, given at the University of Exeter. The notes are not self-contained – you will need to look in books for a full understanding, and this version of the notes is quite streamlined.

Chapter 1

Equations with Rotation and Stratification

Weeks 1 to 3

1.1 Review of Fluid Equations

First we just write down the equations without derivation. For dry air, or for a salt-free liquid, the equations of motion may be written as follows:

The *mass continuity equation:*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0. \tag{1.1}$$

If density is constant this reduces to $\nabla \cdot \boldsymbol{v} = 0$.

The *momentum equation:*

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} = -\frac{\nabla p}{\rho} + \nu \nabla^2 \boldsymbol{v} + \boldsymbol{F}, \tag{1.2}$$

where $\boldsymbol{F}$ represents the effects of body forces such as gravity and ν is the kinematic viscosity. If density is constant, or pressure, p, is given as a function of density alone (e.g., $p = C\rho^\gamma$ where γ is a constant), then (1.1) and (1.2) form a complete system.

The *thermodynamic equation:*

$$\frac{\mathrm{D}I}{\mathrm{D}t} + \frac{p}{\rho} \nabla \cdot \boldsymbol{v} = \dot{Q}, \tag{1.3}$$

where $\dot{Q}$ represents diabatic sources such as heating and diffusion, I is internal energy. In the ideal gas case the internal energy is given by $I = c_v T$ where T is temperature.

An *equation of state:*

$$p = f(I, \rho), \tag{1.4}$$

where f is some known function. For example, for an ideal gas, $p = \rho R I / c_v$ or, more simply, $p = \rho R T$, where R is the ideal gas constant for the gas at hand and T is temperature.

The above four equations have four unknowns: velocity (a vector), temperature, pressure and density. The equations are called the *Euler equations* if the viscous term is omitted, and the *Navier–Stokes equations* if viscosity is included.[1]

1.1.1 Ideal Gas

Let us look at the ideal gas case in a little more detail. For fluid dynamical purposes the ideal gas equation of state is usually written in the form

$$p = \rho R T \tag{1.5}$$

where R is the gas constant of the gas in question, related to the universal gas constant R_u by $R = R_u/m$, where m is the molecular weight of the gas.

The internal energy of an ideal gas is given by $I = c_v T$ where c_v is the heat capacity at constant volume. It is a function of temperature alone, and in fact is almost a constant. For an ideal gas we also have $c_p - c_v = R$, where c_p is the heat capacity at constant pressure.

For an ideal gas the first law of thermodynamics may be written in either of the two equivalent forms

$$đQ = c_v \, \mathrm{d}T + p \, \mathrm{d}\alpha \qquad \text{or} \qquad đQ = c_p \, \mathrm{d}T - \alpha \, \mathrm{d}p, \tag{1.6a,b}$$

where the second expression is derived using $\alpha = RT/p$. Forming the material derivative of the above gives two forms of the internal energy equation:

$$c_v \frac{\mathrm{D}T}{\mathrm{D}t} + p \frac{\mathrm{D}\alpha}{\mathrm{D}t} = \dot{Q} \qquad \text{or} \qquad c_p \frac{\mathrm{D}T}{\mathrm{D}t} - \frac{RT}{p} \frac{\mathrm{D}p}{\mathrm{D}t} = \dot{Q}. \tag{1.7a,b}$$

Using the mass continuity equation, (1.7a) is equivalent to

$$\frac{DT}{Dt} + \frac{p}{c_v \rho} \nabla \cdot \boldsymbol{v} = \frac{\dot{Q}}{c_v}. \tag{1.8}$$

Alternatively, again using the ideal gas equation, we may eliminate T in favour of p and α and obtain

$$\frac{Dp}{Dt} + \gamma p \nabla \cdot \boldsymbol{v} = \dot{Q} \frac{\rho R}{c_v}. \tag{1.9}$$

where $\gamma = c_p/c_v$.

Potential temperature

Using the ideal gas equation we can write (1.6b) as s

$$d\eta = \frac{đQ}{T} = c_p \, d \ln T - R \, d \ln p. \tag{1.10}$$

where η is the specific entropy, which is a function of state. Now, let us *define* the *potential temperature,* θ, by the expression

$$\theta \equiv T \left(\frac{p_0}{p} \right)^{\kappa}, \tag{1.11}$$

where $\kappa = R/c_p$. It straightforwardly follows that

$$c_p \, d \ln \theta = c_p \, d \ln T - R \, d \ln p, \tag{1.12}$$

and therefore the first law of thermodynamics can be written at

$$đQ = c_p \left(\frac{T}{\theta} \right) d\theta. \tag{1.13}$$

Taking the material derivative we have

$$c_p \frac{D\theta}{Dt} = \frac{\theta}{T} \dot{Q}. \tag{1.14}$$

This is a useful form because it just involves the material derivative of one quantity. The potential temperature is, in the absence of diabatic terms, a materially conserved quantity, unlike temperature. It is closely related to entropy, and in particular

$$d\eta = c_p \, d \ln \theta. \tag{1.15}$$

The potential temperature is the temperature that a fluid would have if moved adiabatically to the reference pressure p_0, but the explicit demonstration of this is left to the reader. Indeed, potential temperature may be defined this way, and for an ideal gas this is equivalent to (1.11).

1.2 The Equations of Motion in a Rotating Frame of Reference

Newton's second law of motion, that the acceleration on a body is proportional to the imposed force divided by the body's mass, applies in so-called inertial frames of reference; that is, frames that are stationary or moving only with a constant rectilinear velocity relative to the distant galaxies. Now the Earth spins round its own axis with a period of almost 24 hours (23h 56m) and so the surface of the Earth manifestly is not an inertial frame. Nevertheless, it is very convenient to describe the flow relative to the Earth's surface (which in fact is moving at speeds of up to a few hundreds of metres per second), rather than in some inertial frame. This necessitates recasting the equations into a form that is appropriate for a rotating frame of reference, and that is the subject of this section.

1.2.1 Rate of change of a vector

Consider first a vector $\boldsymbol{C}$ of constant length rotating relative to an inertial frame at a constant angular velocity $\boldsymbol{\Omega}$. Then, in a frame rotating with that same angular velocity it appears stationary and constant. If in a small interval of time δt the vector $\boldsymbol{C}$ rotates through a small angle $\delta\lambda$ then the change in $\boldsymbol{C}$, as perceived in the inertial frame, is given by (see Fig. 1.1)

$$\delta \boldsymbol{C} = |\boldsymbol{C}| \cos\vartheta \, \delta\lambda \, \boldsymbol{m}, \tag{1.16}$$

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

$$\delta \boldsymbol{C} = |\boldsymbol{C}||\boldsymbol{\Omega}| \sin\widehat{\vartheta} \, \boldsymbol{m} \, \delta t = \boldsymbol{\Omega} \times \boldsymbol{C} \, \delta t. \tag{1.17}$$

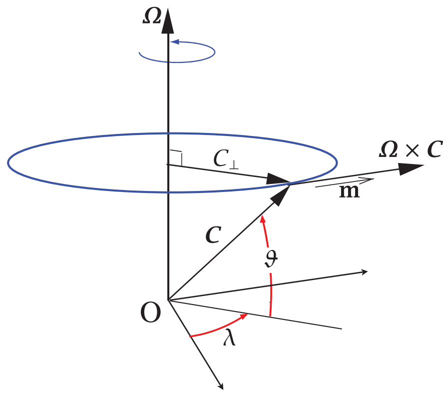

Figure 1.1 A vector $\boldsymbol{C}$ rotating at an angular velocity $\boldsymbol{\Omega}$. It appears to be a constant vector in the rotating frame, whereas in the inertial frame it evolves according to $(\mathrm{d}\boldsymbol{C}/\mathrm{d}t)_I = \boldsymbol{\Omega}\times\boldsymbol{C}$.

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

$$\left(\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}\right)_I = \boldsymbol{\Omega}\times\boldsymbol{C}, \tag{1.18}$$

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

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

$$(\delta\boldsymbol{B})_I = (\delta\boldsymbol{B})_R + (\delta\boldsymbol{B})_{rot}, \tag{1.19}$$

where the terms are, respectively, the change seen in the inertial frame, the change due to the vector itself changing as measured in the rotating frame, and the change due to the rotation. Using (1.17) $(\delta\boldsymbol{B})_{rot} = \boldsymbol{\Omega}\times\boldsymbol{B}\,\delta t$, and so the rates of change of the vector $\boldsymbol{B}$ in the inertial and rotating frames are related by

$$\left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega}\times\boldsymbol{B}. \tag{1.20}$$

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

1.2.2 Velocity and acceleration in a rotating frame

The velocity of a body is not measured to be the same in the inertial and rotating frames, so care must be taken when applying (1.20) to velocity. First apply (1.20) to $\boldsymbol{r}$, the position of a particle to obtain

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega}\times\boldsymbol{r} \tag{1.21}$$

or

$$\boldsymbol{v}_I = \boldsymbol{v}_R + \boldsymbol{\Omega}\times\boldsymbol{r}. \tag{1.22}$$

We refer to $\boldsymbol{v}_R$ and $\boldsymbol{v}_I$ as the relative and inertial velocity, respectively, and (1.22) relates the two. Apply (1.20) again, this time to the velocity $\boldsymbol{v}_R$ to give

$$\left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega}\times\boldsymbol{v}_R, \tag{1.23}$$

or, using (1.22)

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}_I - \boldsymbol{\Omega}\times\boldsymbol{r})\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega}\times\boldsymbol{v}_R, \tag{1.24}$$

or

$$\left(\frac{\mathrm{d}\boldsymbol{v}_I}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega}\times\boldsymbol{v}_R + \frac{\mathrm{d}\boldsymbol{\Omega}}{\mathrm{d}t}\times\boldsymbol{r} + \boldsymbol{\Omega}\times\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I. \tag{1.25}$$

Then, noting that

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega}\times\boldsymbol{r} = (\boldsymbol{v}_R + \boldsymbol{\Omega}\times\boldsymbol{r}), \tag{1.26}$$

and assuming that the rate of rotation is constant, (1.25) becomes

$$\left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R = \left(\frac{\mathrm{d}\boldsymbol{v}_I}{\mathrm{d}t}\right)_I - 2\boldsymbol{\Omega}\times\boldsymbol{v}_R - \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}). \tag{1.27}$$

This equation may be interpreted as follows. The term on the left-hand side is the rate of change of the relative velocity as measured in the rotating frame. The first term on the right-hand side is the rate of change of the inertial velocity as measured in the inertial frame (the inertial acceleration, which is, by Newton's second law, equal to the force on a fluid parcel divided by its mass). The second and third terms on the right-hand side (including the minus signs) are the *Coriolis force* and the *centrifugal force* per unit mass. Neither of these are true forces — they may be thought of as quasi-forces (i.e., 'as if' forces); that is, when a body is observed from a rotating frame it seems to behave as if unseen forces are present that affect its motion. If (1.27) is written, as is common, with the terms $+2\boldsymbol{\Omega}\times\boldsymbol{v}_r$ and $+\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r})$ on the left-hand side then these terms should be referred to as the Coriolis and centrifugal *accelerations.*

Centrifugal force

If $\boldsymbol{r}_\perp$ is the perpendicular distance from the axis of rotation (see Fig. 1.1 and substitute $\boldsymbol{r}$ for $\boldsymbol{C}$), then, because $\boldsymbol{\Omega}$ is perpendicular to $\boldsymbol{r}_\perp$, $\boldsymbol{\Omega}\times\boldsymbol{r} = \boldsymbol{\Omega}\times\boldsymbol{r}_\perp$. Then, using the vector identity $\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}_\perp) = (\boldsymbol{\Omega}\cdot\boldsymbol{r}_\perp)\boldsymbol{\Omega} - (\boldsymbol{\Omega}\cdot\boldsymbol{\Omega})\boldsymbol{r}_\perp$ and noting that the first term is zero, we see that the centrifugal force per unit mass is just given by

$$\boldsymbol{F}_{ce} = -\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}) = \Omega^2\boldsymbol{r}_\perp. \tag{1.28}$$

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

$$\boldsymbol{F}_{ce} = -\nabla\Phi_{ce}. \tag{1.29}$$

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

Coriolis force

The Coriolis force per unit mass is:

$$\boldsymbol{F}_{Co} = -2\boldsymbol{\Omega}\times\boldsymbol{v}_R. \tag{1.30}$$

It plays a central role in much of geophysical fluid dynamics and will be considered extensively later on. For now, we just note three basic properties.

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

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

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

1.2.3 Momentum equation in a rotating frame

Since (1.27) simply relates the accelerations of a particle in the inertial and rotating frames, then in the rotating frame of reference the momentum equation may be written

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v} = -\frac{1}{\rho}\nabla p - \nabla\Phi, \qquad (1.31)$$

incorporating the centrifugal term into the potential, Φ. We have dropped the subscript R; henceforth, unless we need to be explicit, all velocities without a subscript will be considered to be relative to the rotating frame.

1.2.4 Mass and tracer conservation in a rotating frame

Let ϕ be a scalar field that, in the inertial frame, obeys

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot \boldsymbol{v}_I = 0. \qquad (1.32)$$

Now, observers in both the rotating and inertial frame measure the same value of ϕ. Further, $\mathrm{D}\phi/\mathrm{D}t$ is simply the rate of change of ϕ associated with a material parcel, and therefore is reference frame invariant. Thus,

$$\left(\frac{\mathrm{D}\phi}{\mathrm{D}t}\right)_R = \left(\frac{\mathrm{D}\phi}{\mathrm{D}t}\right)_I, \qquad (1.33)$$

where $(\mathrm{D}\phi/\mathrm{D}t)_R = (\partial\phi/\partial t)_R + \boldsymbol{v}_R \cdot \nabla\phi$ and $(\mathrm{D}\phi/\mathrm{D}t)_I = (\partial\phi/\partial t)_I + \boldsymbol{v}_I\cdot\nabla\phi$ and the local temporal derivatives $(\partial\phi/\partial t)_R$ and $(\partial\phi/\partial t)_I$ are evaluated at fixed locations in the rotating and inertial frames, respectively.

Further, using (1.22), we have that we have that

$$\nabla\cdot \boldsymbol{v}_I = \nabla\cdot (\boldsymbol{v}_R + \boldsymbol{\Omega}\times\boldsymbol{r}) = \nabla\cdot \boldsymbol{v}_R \qquad (1.34)$$

since $\nabla\cdot (\boldsymbol{\Omega}\times\boldsymbol{r}) = 0$. Thus, using (1.33) and (1.34), (1.32) is equivalent to

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot \boldsymbol{v}_R = 0, \qquad (1.35)$$

where all observables are measured in the *rotating* frame. Thus, the equation for the evolution of a scalar whose measured value is the same in rotating and inertial frames is unaltered by the presence of rotation. In particular, the mass conservation equation is unaltered by the presence of rotation.

The individual components of the material derivative differ in the rotating and inertial frames. In particular

$$\left(\frac{\partial \phi}{\partial t}\right)_I = \left(\frac{\partial \phi}{\partial t}\right)_R - (\boldsymbol{\Omega} \times \boldsymbol{r}) \cdot \nabla \phi \tag{1.36}$$

because $\boldsymbol{\Omega} \times \boldsymbol{r}$ is the velocity, in the inertial frame, of a uniformly rotating body. Similarly,

$$\boldsymbol{v}_I \cdot \nabla \phi = (\boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}) \cdot \nabla \phi. \tag{1.37}$$

Adding the last two equations reprises (1.33).

1.3 ♦ Spherical Coordinates

We write these equations down for reference, but we won't derive them or use them in their spherical form.

1.3.1 Mass Conservation and Thermodynamic Equation

The mass conservation equation expanded in spherical co-ordinates, is

$$\frac{\partial \rho}{\partial t} + \frac{u}{r \cos \vartheta} \frac{\partial \rho}{\partial \lambda} + \frac{v}{r} \frac{\partial \rho}{\partial \vartheta} + w \frac{\partial \rho}{\partial r} + \frac{\rho}{r \cos \vartheta} \left[\frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \vartheta}(v \cos \vartheta) + \frac{1}{r} \frac{\partial}{\partial r}(w r^2 \cos \vartheta) \right] = 0. \tag{1.38}$$

Equivalently this is the same as

$$\frac{\partial \rho}{\partial t} + \frac{1}{r \cos \vartheta} \frac{\partial (u \rho)}{\partial \lambda} + \frac{1}{r \cos \vartheta} \frac{\partial}{\partial \vartheta}(v \rho \cos \vartheta) + \frac{1}{r^2} \frac{\partial}{\partial r}(r^2 w \rho) = 0. \tag{1.39}$$

The thermodynamic equation is a tracer advection equation. The (adiabatic) potential temperature in spherical coordinate form is

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \frac{\partial \theta}{\partial t} + \frac{u}{r \cos \vartheta} \frac{\partial \theta}{\partial \lambda} + \frac{v}{r} \frac{\partial \theta}{\partial \vartheta} + w \frac{\partial \theta}{\partial r} = 0, \tag{1.40}$$

and similarly for tracers such as water vapour or salt.

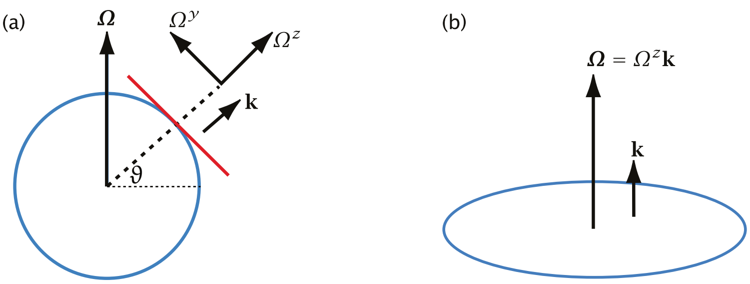

Figure 1.2 (a) On the sphere the rotation vector $\boldsymbol{\Omega}$ can be decomposed into two components, one in the local vertical and one in the local horizontal, pointing toward the pole. That is, $\boldsymbol{\Omega} = \Omega_y \mathbf{j} + \Omega_z \mathbf{k}$ where $\Omega_y = \Omega \cos\vartheta$ and $\Omega_z = \Omega \sin\vartheta$. In geophysical fluid dynamics, the rotation vector in the local vertical is often the more important component in the horizontal momentum equations. On a rotating disk, (b), the rotation vector $\boldsymbol{\Omega}$ is parallel to the local vertical $\mathbf{k}$.

1.3.2 Momentum Equation

The momentum equation is:

$$\frac{\mathrm{D}u}{\mathrm{D}t} - \left(2\Omega + \frac{u}{r\cos\vartheta}\right)(v\sin\vartheta - w\cos\vartheta) = -\frac{1}{\rho r\cos\vartheta}\frac{\partial p}{\partial\lambda}, \tag{1.41a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + \frac{wv}{r} + \left(2\Omega + \frac{u}{r\cos\vartheta}\right)u\sin\vartheta = -\frac{1}{\rho r}\frac{\partial p}{\partial\vartheta}, \tag{1.41b}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} - \frac{u^2 + v^2}{r} - 2\Omega u\cos\vartheta = -\frac{1}{\rho}\frac{\partial p}{\partial r} - g. \tag{1.41c}$$

The terms involving Ω are called Coriolis terms, and the quadratic terms on the left-hand sides involving $1/r$ are often called metric terms.

1.4 ♦ The primitive equations

The so-called *primitive equations* of motion are simplifications of the equations that make three related approximations:

(i) The hydrostatic approximation. In the vertical momentum equation the gravitational term is assumed

to be balanced by the pressure gradient term, so that

$$\frac{\partial p}{\partial z} = -\rho g. \tag{1.42}$$

The advection of vertical velocity, the Coriolis terms, and the metric term $(u^2 + v^2)/r$ are all neglected.

(ii) *The shallow-fluid approximation.* We write $r = a + z$ where the constant a is the radius of the Earth and z increases in the radial direction. The coordinate r is then replaced by a except where it is used as the differentiating argument. Thus, for example,

$$\frac{1}{r^2}\frac{\partial (r^2 w)}{\partial r} \rightarrow \frac{\partial w}{\partial z}. \tag{1.43}$$

(iii) *The traditional approximation.* Coriolis terms in the horizontal momentum equations involving the vertical velocity, and the still smaller metric terms uw/r and vw/r, are neglected.

The second and third of these approximations should be taken, or not, together, the underlying reason being that they both relate to the presumed small aspect ratio of the motion, so the approximations succeed or fail together.

Making these approximations, the momentum equations become

$$\frac{\mathrm{D}u}{\mathrm{D}t} - 2\Omega \sin\vartheta v - \frac{uv}{a}\tan\vartheta = -\frac{1}{a\rho\cos\vartheta}\frac{\partial p}{\partial \lambda}, \tag{1.44a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + 2\Omega \sin\vartheta u + \frac{u^2\tan\vartheta}{a} = -\frac{1}{\rho a}\frac{\partial p}{\partial \vartheta}, \tag{1.44b}$$

$$0 = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{1.44c}$$

where

$$\frac{\mathrm{D}}{\mathrm{D}t} = \left(\frac{\partial}{\partial t} + \frac{u}{a\cos\vartheta}\frac{\partial}{\partial\lambda} + \frac{v}{a}\frac{\partial}{\partial\vartheta} + w\frac{\partial}{\partial z}\right). \tag{1.45}$$

We note the ubiquity of the factor $2\Omega\sin\vartheta$, and take the opportunity to define the *Coriolis parameter*, $f \equiv 2\Omega\sin\vartheta$.

The corresponding mass conservation equation for a shallow fluid layer is:

$$\frac{\partial \rho}{\partial t} + \frac{u}{a\cos\vartheta}\frac{\partial \rho}{\partial \lambda} + \frac{v}{a}\frac{\partial \rho}{\partial \vartheta} + w\frac{\partial \rho}{\partial z} + \rho\left[\frac{1}{a\cos\vartheta}\frac{\partial u}{\partial \lambda} + \frac{1}{a\cos\vartheta}\frac{\partial}{\partial \vartheta}(v\cos\vartheta) + \frac{\partial w}{\partial z}\right] = 0, \tag{1.46}$$

or equivalently,

$$\frac{\partial \rho}{\partial t} + \frac{1}{a\cos\vartheta}\frac{\partial (u\rho)}{\partial \lambda} + \frac{1}{a\cos\vartheta}\frac{\partial}{\partial \vartheta}(v\rho\cos\vartheta) + \frac{\partial (w\rho)}{\partial z} = 0. \tag{1.47}$$

1.5 Cartesian Approximations: The Tangent Plane

1.5.1 The f-plane

Although the rotation of the Earth is central for many dynamical phenomena, the sphericity of the Earth is not always so. This is especially true for phenomena on a scale somewhat smaller than global where the use of spherical coordinates becomes awkward, and it is more convenient to use a locally Cartesian representation of the equations. Referring to Fig. 1.2 we will define a plane tangent to the surface of the Earth at a latitude ϑ_0, and then use a Cartesian coordinate system (x, y, z) to describe motion on that plane. For small excursions on the plane, $(x, y, z) \approx (a\lambda\cos\vartheta_0, a(\vartheta - \vartheta_0), z)$. Consistently, the velocity is $\boldsymbol{v} = (u, v, w)$, so that u,v and w are the components of the velocity *in the tangent plane,* in approximately in the east–west, north–south and vertical directions, respectively.

The momentum equations for flow in this plane are then

$$\frac{\partial u}{\partial t} + (\boldsymbol{v}\cdot\nabla)u + 2(\Omega^y w - \Omega^z v) = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{1.48a}$$

$$\frac{\partial v}{\partial t} + (\boldsymbol{v}\cdot\nabla)v + 2(\Omega^z u - \Omega^x w) = -\frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{1.48b}$$

$$\frac{\partial w}{\partial t} + (\boldsymbol{v}\cdot\nabla)w + 2(\Omega^x v - \Omega^y u) = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{1.48c}$$

where the rotation vector $\boldsymbol{\Omega} = \Omega^x \mathbf{i} + \Omega^y \mathbf{j} + \Omega^z \mathbf{k}$ and $\Omega^x = 0$, $\Omega^y = \Omega \cos \vartheta_0$ and $\Omega^z = \Omega \sin \vartheta_0$. If we make the traditional approximation, and so ignore the components of $\boldsymbol{\Omega}$ not in the direction of the local vertical, then

$$\frac{\mathrm{D}u}{\mathrm{D}t} - f_0 v = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{1.49a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + f_0 u = -\frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{1.49b}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g. \tag{1.49c}$$

where $f_0 = 2\Omega^z = 2\Omega \sin \vartheta_0$. Defining the horizontal velocity vector $\boldsymbol{u} = (u, v, 0)$, the first two equations may also be written as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}_0 \times \boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{1.50}$$

where $\mathrm{D}\boldsymbol{u}/\mathrm{D}t = \partial \boldsymbol{u}/\partial t + \boldsymbol{v}\cdot\nabla\boldsymbol{u}$, $\boldsymbol{f}_0 = 2\Omega \sin\vartheta_0 \mathbf{k} = f_0 \mathbf{k}$, and $\mathbf{k}$ is the direction perpendicular to the plane (it does not change its orientation with latitude). These equations are, evidently, exactly the same as the momentum equations in a system in which the rotation vector is aligned with the local vertical, as illustrated in the right-hand panel in Fig. 1.2 (on page 10). They will describe flow on the surface of a rotating sphere to a good approximation provided the flow is of limited latitudinal extent so that the effects of sphericity are unimportant; we have made what is known as the *f-plane* approximation since the Coriolis parameter is a constant. We may in addition make the hydrostatic approximation, in which case (1.49c) becomes the familiar $\partial p/\partial z = -\rho g$.

1.5.2 The beta-plane approximation

The magnitude of the vertical component of rotation varies with latitude, and this has important dynamical consequences. We can approximate this effect by allowing the effective rotation vector to vary. Thus, noting that, for small variations in latitude,

$$f = 2\Omega \sin\vartheta \approx 2\Omega \sin\vartheta_0 + 2\Omega(\vartheta - \vartheta_0)\cos\vartheta_0, \tag{1.51}$$

then on the tangent plane we may mimic this by allowing the Coriolis parameter to vary as

$$f = f_0 + \beta y \quad , \tag{1.52}$$

where $f_0 = 2\Omega \sin \vartheta_0$ and $\beta = \partial f/\partial y = (2\Omega \cos \vartheta_0)/a$. This important approximation is known as the *beta-plane*, or *β-plane,* approximation; it captures the the most important *dynamical* effects of sphericity, without the complicating *geometric* effects, which are not essential to describe many phenomena. The momentum equations (1.49) are unaltered except that f_0 is replaced by $f_0 + \beta y$ to represent a varying Coriolis parameter. Thus, sphericity combined with rotation is dynamically equivalent to a *differentially rotating* system. For future reference, we write down the β-plane horizontal momentum equations:

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{1.53}$$

where $\boldsymbol{f} = (f_0 + \beta y)\hat{\mathbf{k}}$. In component form this equation becomes

$$\frac{\mathrm{D}u}{\mathrm{D}t} - fv = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \qquad \frac{\mathrm{D}v}{\mathrm{D}t} + fu = -\frac{1}{\rho}\frac{\partial p}{\partial y}. \tag{1.54a,b}$$

The mass conservation, thermodynamic and hydrostatic equations in the β-plane approximation are the same as the usual Cartesian, f-plane, forms of those equations.

1.6 The Boussinesq Approximation

The density variations in the ocean are quite small compared to the mean density, and we may exploit this to derive somewhat simpler but still quite accurate equations of motion. Let us first examine how much density does vary in the ocean.

1.6.1 Variation of density in the ocean

The variations of density in the ocean are due to three effects: the compression of water by pressure (which we denote as $\varDelta_p\rho$), the thermal expansion of water if its temperature changes ($\varDelta_T\rho$), and the haline contraction if its salinity changes ($\varDelta_S\rho$). How big are these? An appropriate equation of state to approximately evaluate these effects is the linear one

$$\rho = \rho_0\left[1 - \beta_T(T - T_0) + \beta_S(S - S_0) + \frac{p}{\rho_0 c_s^2}\right], \qquad (1.55)$$

where $\beta_T \approx 2\times10^{-4}\,\mathrm{K}^{-1}$, $\beta_S \approx 10^{-3}\,\mathrm{psu}^{-1}$ and $c_s \approx 1500\,\mathrm{m\,s}^{-1}$. The three effects may then be evaluated as follows.

Pressure compressibility. We have $\varDelta_p\rho \approx \varDelta p/c_s^2 \approx \rho_0 gH/c_s^2$. where H is the depth and the pressure change is quite accurately evaluated using the hydrostatic approximation. Thus,

$$\frac{|\varDelta_p\rho|}{\rho_0} \ll 1 \qquad \text{if} \qquad \frac{gH}{c_s^2} \ll 1, \qquad (1.56)$$

or if $H \ll c_s^2/g$. The quantity $c_s^2/g \approx 200\,\mathrm{km}$ is the density scale height of the ocean. Thus, the pressure at the bottom of the ocean (say $H = 10\,\mathrm{km}$ in the deep trenches), enormous as it is, is insufficient to compress the water enough to make a significant change in its density. Changes in density due to dynamical variations of pressure are small if the Mach number is small, and this is also the case.

Thermal expansion. We have $\varDelta_T\rho \approx -\beta_T\rho_0\varDelta T$ and therefore

$$\frac{|\varDelta_T\rho|}{\rho_0} \ll 1 \qquad \text{if} \qquad \beta_T\varDelta T \ll 1. \qquad (1.57)$$

For $\varDelta T = 20\,\mathrm{K}$, $\beta_T\varDelta T \approx 4\times10^{-3}$, and evidently we would require temperature differences of order β_T^{-1}, or 5000 K to obtain order one variations in density.

Saline contraction. We have $\varDelta_S\rho \approx \beta_S\rho_0\varDelta S$ and therefore

$$\frac{|\varDelta_S\rho|}{\rho_0} \ll 1 \qquad \text{if} \qquad \beta_S\varDelta S \ll 1. \qquad (1.58)$$

As changes in salinity in the ocean rarely exceed 5 psu, for which $\beta_S\varDelta S = 5\times10^{-3}$, the fractional change in the density of seawater is correspondingly very small.

Evidently, fractional density changes in the ocean are very small.

1.6.2 The Boussinesq equations

The *Boussinesq equations* are a set of equations that exploit the smallness of density variations in many liquids.We write

$$\rho = \rho_0 + \delta\rho(x, y, z, t) \tag{1.59a}$$

where ρ_0 is a constant and we assume that

$$|\delta\rho| \ll \rho_0. \tag{1.60}$$

Associated with the reference density is a reference pressure that is defined to be in hydrostatic balance with it. That is,

$$p = p_0(z) + \delta p(x, y, z, t) \tag{1.61a}$$

where

$$\frac{\mathrm{d}p_0}{\mathrm{d}z} \equiv -g\rho_0 \tag{1.62}$$

Note that $\nabla_z p = \nabla_z \delta p$.

Momentum equations

To obtain the Boussinesq equations we use $\rho = \rho_0 + \delta\rho$, and assume $\delta\rho/\rho_0$ is small. Without approximation, the momentum equation can be written as

$$(\rho_0 + \delta\rho)\left(\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega}\times\boldsymbol{v}\right) = -\nabla\delta p - \frac{\partial p_0}{\partial z}\mathbf{k} - g(\rho_0 + \delta\rho)\mathbf{k}, \tag{1.63}$$

and using (1.62a) this becomes, again without approximation,

$$(\rho_0 + \delta\rho)\left(\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega}\times\boldsymbol{v}\right) = -\nabla\delta p - g\delta\rho\mathbf{k}. \tag{1.64}$$

If density variations are small this becomes

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega}\times\boldsymbol{v} = -\nabla\phi + b\mathbf{k}, \tag{1.65}$$

where $\phi = \delta p/\rho_0$ and $b = -g\,\delta\rho/\rho_0$ is the *buoyancy*. Note that we should not and do not neglect the term $g\,\delta\rho$, for there is no reason to believe it to be small ($\delta\rho$ may be small, but g is big). Equation (1.65) is the momentum equation in the Boussinesq approximation, and it is common to say that the Boussinesq approximation ignores all variations of density of a fluid in the momentum equation, except when associated with the gravitational term.

For most large-scale motions in the ocean the *deviation* pressure and density fields are also approximately in hydrostatic balance, and in that case the vertical component of (1.65) becomes

$$\frac{\partial\phi}{\partial z} = b. \tag{1.66}$$

A condition for (1.66) to hold is that vertical accelerations are small *compared to $g\,\delta\rho/\rho_0$, and not compared to the acceleration due to gravity itself.* For more discussion of this point, see section 1.7.

Mass Conservation

The unapproximated mass conservation equation is

$$\frac{\mathrm{D}\delta\rho}{\mathrm{D}t} + (\rho_0 + \delta\rho)\nabla\cdot\boldsymbol{v} = 0. \tag{1.67}$$

Provided that time scales advectively — that is to say that $\mathrm{D}/\mathrm{D}t$ scales in the same way as $\boldsymbol{v}\cdot\nabla$ — then we may approximate this equation by

$$\boxed{\nabla\cdot\boldsymbol{v} = 0}\,, \tag{1.68}$$

which is the same as that for a constant density fluid. This *absolutely does not* allow one to go back and use (1.67) to say that $D\delta\rho/Dt = 0$; the evolution of density is given by the thermodynamic equation in conjunction with an equation of state, and this should not be confused with the mass conservation equation. Note also that in eliminating the time-derivative of density we eliminate the possibility of sound waves.

Thermodynamic equation and equation of state

[This section is even more informal and non-rigorous than other sections.] We write the thermodynamic equation as

$$\frac{DI}{Dt} + \frac{p}{\rho}\nabla\cdot \boldsymbol{v} = \dot{Q} \tag{1.69}$$

We neglect the second term on the left-hand side (because the fluid is incompressible), and write the first term in terms of temperature

$$C\frac{DT}{Dt} = \dot{Q} \tag{1.70}$$

where c is the heat capacity of the fluid. We further suppose that the temperature is linearly related to the buoyancy, b. That is, we assume a linear equation of state, $b = b_0\left(1 + A(T - T_0)\right)$ where A is a constant coefficient of thermal expansion. The thermodynamic equation becomes

$$\frac{Db}{Dt} = Q_b, \tag{1.71}$$

where $Q_b = A\dot{Q}/C$. The momentum equation (1.65), mass continuity equation (1.68) and thermodynamic equation (1.71) then form a closed set, called the *Boussinesq equations*.

◆ *Mean stratification and the buoyancy frequency*

The processes that cause density to vary in the vertical often differ from those that cause it to vary in the horizontal. For this reason it is sometimes useful to write $\rho = \rho_0 + \widehat{\rho}(z) + \rho'(x, y, z, t)$ and define $\widetilde{b}(z) \equiv -g\widehat{\rho}/\rho_0$ and $b' \equiv -g\rho'/\rho_0$. The thermodynamic equation (1.69) becomes

$$\frac{Db'}{Dt} + N^2 w = 0, \tag{1.72}$$

where D/Dt remains a three-dimensional operator and

$$N^2(z) = \frac{d\widetilde{b}_\sigma}{dz}, \tag{1.73}$$

The quantity N^2 is a measure of the *mean stratification* of the fluid. N is known as the buoyancy frequency, something we return to later on.

Summary of Boussinesq Equations

The simple Boussinesq equations are, for an inviscid fluid:

momentum equations: $$\frac{D\boldsymbol{v}}{Dt} + \boldsymbol{f} \times \boldsymbol{v} = -\nabla\phi + b\mathbf{k}, \qquad \text{(B.1)}$$

mass conservation: $$\nabla\cdot \boldsymbol{v} = 0, \qquad \text{(B.2)}$$

buoyancy equation: $$\frac{Db}{Dt} = \dot{b}. \qquad \text{(B.3)}$$

A more general form replaces the buoyancy equation by:

thermodynamic equation: $$\frac{D\theta}{Dt} = \dot{\theta}, \qquad \text{(B.4)}$$

salinity equation: $$\frac{DS}{Dt} = \dot{S}, \qquad \text{(B.5)}$$

equation of state: $$b = b(\theta, S, z). \qquad \text{(B.6)}$$

1.6.3 Energetics of the Boussinesq system

In a uniform gravitational field but with no other forcing or dissipation, we write the simple Boussinesq equations as

$$\frac{D\boldsymbol{v}}{Dt} + 2\boldsymbol{\Omega} \times \boldsymbol{v} = b\mathbf{k} - \nabla\phi, \qquad \nabla\cdot \boldsymbol{v} = 0, \qquad \frac{Db}{Dt} = 0. \qquad \text{(1.74a,b,c)}$$

From (1.74a) and (1.74b) the kinetic energy density evolution is given by

$$\frac{1}{2}\frac{Dv^2}{Dt} = bw - \nabla\cdot(\phi\boldsymbol{v}), \qquad (1.75)$$

where the constant reference density ρ_0 is omitted. Let us now define the potential $\Phi \equiv -z$, so that $\nabla\Phi = -\mathbf{k}$ and

$$\frac{D\Phi}{Dt} = \nabla\cdot(\boldsymbol{v}\Phi) = -w, \qquad (1.76)$$

and using this and (1.74c) gives

$$\frac{D}{Dt}(b\Phi) = -wb. \qquad (1.77)$$

Adding (1.77) to (1.75) and expanding the material derivative gives

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\boldsymbol{v}^2 + b\Phi\right) + \nabla\cdot\left[\boldsymbol{v}\left(\frac{1}{2}\boldsymbol{v}^2 + b\Phi + \phi\right)\right] = 0. \qquad (1.78)$$

This constitutes an energy equation for the Boussinesq system. The energy density (divided by ρ_0) is just $\boldsymbol{v}^2/2 + b\Phi$. What does the term $b\Phi$ represent? Its integral, multiplied by ρ_0, is the potential energy of the flow minus that of the basic state, or $\int g(\rho - \rho_0)z\,\mathrm{d}z$. If there were a heating term on the right-hand side of (1.74c) this would directly provide a source of potential energy, rather than internal energy as in the compressible system. Because the fluid is incompressible, there is no conversion from kinetic and potential energy into internal energy.

1.7 Scaling for Hydrostatic Balance

We now look in more detail at the conditions required for hydrostatic balance to hold. Along with geostrophic balance, considered in the next section, it is one of the most fundamental balances in geophysical fluid dynamics. The corresponding states, hydrostasy and geostrophy, are not exactly realized, but their approximate satisfaction has profound consequences on the behaviour of the atmosphere and ocean.

1.7.1 Preliminaries

Consider the relative sizes of terms in (1.48c):

$$\frac{W}{T} + \frac{UW}{L} + \frac{W^2}{H} + \Omega U \sim \left|\frac{1}{\rho}\frac{\partial p}{\partial z}\right| + g. \qquad (1.79)$$

For most large-scale motion in the atmosphere and ocean the terms on the right-hand side are orders of magnitude larger than those on the left, and therefore must be approximately equal. Explicitly, suppose $W \sim 1\,\mathrm{cm\,s^{-1}}$, $L \sim 10^5\,\mathrm{m}$, $H \sim 10^3\,\mathrm{m}$, $U \sim 10\,\mathrm{m\,s^{-1}}$, $T = L/U$. Then by substituting into (1.79) it seems that the pressure term is the only one which could balance the gravitational term, and we are led to approximate (1.48c) by,

$$\frac{\partial p}{\partial z} = -\rho g. \tag{1.80}$$

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

However, (1.80) is not always a useful equation! Let us suppose that the density is a constant, ρ_0 . We can then write the pressure as

$$p(x, y, z, t) = p_0(z) + p'(x, y, z, t), \tag{1.81}$$

where

$$\frac{\partial p_0}{\partial z} \equiv -\rho_0 g. \tag{1.82}$$

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

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho_0}\frac{\partial p'}{\partial z}. \tag{1.83}$$

Thus, *for constant density fluids the gravitational term has no dynamical effect:* there is no buoyancy force, and the pressure term in the horizontal momentum equations can be replaced by p'. Hydrostatic balance, and in particular (1.82), is certainly not an appropriate vertical momentum equation in this case. If the fluid is stratified, we should therefore subtract off the hydrostatic pressure associated with the mean density before we can determine whether hydrostasy is a useful *dynamical* approximation, accurate enough to determine the horizontal pressure gradients. This is automatic in the Boussinesq equations, where the vertical momentum equation is

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial \phi}{\partial z} + b, \tag{1.84}$$

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

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{1.85}$$

we need to compare the advective term on the left-hand side with the pressure variations arising from horizontal flow in order to determine whether hydrostasy is an appropriate vertical momentum equation. Nevertheless, if we only need to determine the pressure for use in an equation of state then we simply need to compare the sizes of the dynamical terms in (1.48c) with g itself, in order to determine whether a hydrostatic approximation will suffice.

1.7.2 Scaling and the aspect ratio

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

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla_z \phi, \qquad \frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial \phi}{\partial z} + b. \tag{1.86a,b}$$

With $\boldsymbol{f} = 0$, (1.86a) implies the scaling

$$\phi \sim U^2. \tag{1.87}$$

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

$$w \sim W = \frac{H}{L} U = \alpha U, \tag{1.88}$$

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

$$\frac{\mathrm{D}w}{\mathrm{D}t} \sim \frac{UW}{L} = \frac{U^2 H}{L^2}. \tag{1.89}$$

Using (1.87) and (1.89) the ratio of the advective term to the pressure gradient term in the vertical momentum equations then scales as

$$\frac{|\mathrm{D}w/\mathrm{D}t|}{|\partial \phi/\partial z|} \sim \frac{U^2 H/L^2}{U^2/H} \sim \left(\frac{H}{L}\right)^2. \tag{1.90}$$

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

$$\boxed{\alpha^2 \equiv \left(\frac{H}{L}\right)^2 \ll 1}\,. \tag{1.91}$$

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

We can obtain the same result more formally by non-dimensionalizing the momentum equations. Using upper-case symbols to denote scaling values we write

$$
\begin{gathered}
(x, y) = L(\widehat{x}, \widehat{y}), \qquad z = H\widehat{z}, \\
\boldsymbol{u} = U\widehat{\boldsymbol{u}}, \qquad w = W\widehat{w} = \frac{HU}{L}\widehat{w}, \\
t = T\widehat{t} = \frac{L}{U}\widehat{t}, \qquad \phi = \Phi\widehat{\phi} = U^2\widehat{\phi}, \quad b = B\widehat{b} = \frac{U^2}{H}\widehat{b},
\end{gathered}
\tag{1.92}
$$

where the hatted variables are non-dimensional and the scaling for w is suggested by the mass conservation equation, $\nabla_z \cdot \boldsymbol{u} + \partial w/\partial z = 0$. Substituting (1.92) into (1.86) (with $\boldsymbol{f} = 0$) gives us the non-dimensional equations

$$
\frac{\mathrm{D}\widehat{\boldsymbol{u}}}{\mathrm{D}\widehat{t}} = -\nabla\widehat{\phi}, \qquad \alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}\widehat{t}} = -\frac{\partial\widehat{\phi}}{\partial\widehat{z}} + \widehat{b}, \tag{1.93a,b}
$$

where $\mathrm{D}/\mathrm{D}\widehat{t} = \partial/\partial\widehat{t} + \widehat{u}\partial/\partial\widehat{x} + \widehat{v}\partial/\partial\widehat{y} + \widehat{w}\partial/\partial\widehat{z}$ and we use the convention that when ∇ operates on non-dimensional quantities the operator itself is non-dimensional. From (1.93b) it is clear that hydrostatic balance pertains when $\alpha^2 \ll 1$.

1.8 Geostrophic and Thermal Wind Balance

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

1.8.1 The Rossby number

The *Rossby number* characterizes the importance of rotation in a fluid. It is, essentially, the ratio of the magnitude of the relative acceleration to the Coriolis acceleration, and

Variable	Scaling symbol	Meaning	Atmos. value	Ocean value
(x, y)	L	Horizontal length scale	10^6 m	10^5 m
t	T	Time scale	1 day (10^5 s)	10 days (10^6 s)
(u, v)	U	Horizontal velocity	$10\ \mathrm{m\,s^{-1}}$	$0.1\ \mathrm{m\,s^{-1}}$
	Ro	Rossby number, U/fL	0.1	0.01

Table 1.1 Scales of large-scale flow in atmosphere and ocean. The choices given are representative of large-scale mid-latitude eddying motion in both systems.

it is of fundamental importance in geophysical fluid dynamics. It arises from a simple scaling of the horizontal momentum equation, namely

$$\frac{\partial \boldsymbol{u}}{\partial t}+(\boldsymbol{v}\cdot\nabla)\boldsymbol{u}+\boldsymbol{f}\times\boldsymbol{u}=-\frac{1}{\rho}\nabla_z p, \tag{1.94a}$$

$$U^2/L \qquad fU \tag{1.94b}$$

where U is the approximate magnitude of the horizontal velocity and L is a typical length scale over which that velocity varies. (We assume that $W/H \lesssim U/L$, so that vertical advection does not dominate the advection.) The ratio of the sizes of the advective and Coriolis terms is defined to be the Rossby number,

$$Ro \equiv \frac{U}{fL}. \tag{1.95}$$

If the Rossby number is small then rotation effects are important, and as the values in Table 1.1 indicate this is the case for large-scale flow in both ocean and atmosphere.

Another intuitive way to think about the Rossby number is in terms of time scales. The Rossby number based on a time scale is

$$Ro_T \equiv \frac{1}{fT}, \tag{1.96}$$

where T is a time scale associated with the dynamics at hand. If the time scale is an advective one, meaning that

$T \sim L/U$, then this definition is equivalent to (1.95). Now, $f = 2\Omega \sin\vartheta$, where Ω is the angular velocity of the rotating frame and equal to $2\pi/T_p$ where T_p is the period of rotation (24 hours). Thus,

$$Ro_T = \frac{T_p}{4\pi T \sin\vartheta} = \frac{T_i}{T}, \tag{1.97}$$

where $T_i = 1/f$ is the 'inertial time scale', about three hours in mid-latitudes. Thus, for phenomena with time scales much longer than this, such as the motion of the Gulf Stream or a mid-latitude atmospheric weather system, the effects of the Earth's rotation can be expected to be important, whereas a short-lived phenomena, such as a cumulus cloud or tornado, may be oblivious to such rotation. The expressions (1.95) and (1.96) are, of course, just approximate measures of the importance of rotation.

1.8.2 Geostrophic balance

If the Rossby number is sufficiently small in (1.94a) then the rotation term will dominate the nonlinear advection term, and if the time period of the motion scales advectively then the rotation term also dominates the local time derivative. The only term that can then balance the rotation term is the pressure term, and therefore we must have

$$\boldsymbol{f} \times \boldsymbol{u} \approx -\frac{1}{\rho}\nabla_z p, \tag{1.98}$$

or, in Cartesian component form

$$fu \approx -\frac{1}{\rho}\frac{\partial p}{\partial y}, \qquad fv \approx \frac{1}{\rho}\frac{\partial p}{\partial x}. \tag{1.99}$$

This balance is known as *geostrophic balance,* and its consequences are profound, giving geophysical fluid dynamics a special place in the broader field of fluid dynamics. We *define* the geostrophic velocity by

$$fu_g \equiv -\frac{1}{\rho}\frac{\partial p}{\partial y}, \qquad fv_g \equiv \frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{1.100}$$

and for low Rossby number flow $u \approx u_g$ and $v \approx v_g$.

Geostrophic balance has a number of immediate ramifications:

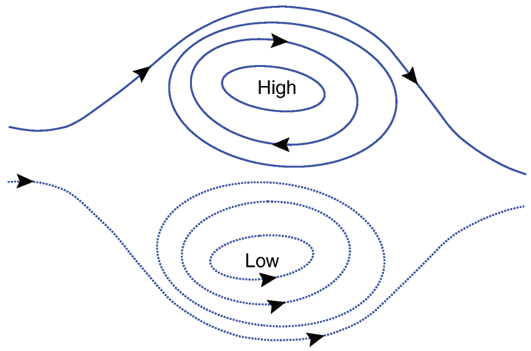

Figure 1.3 Schematic of geostrophic flow with a positive value of the Coriolis parameter f. Flow is parallel to the lines of constant pressure (isobars). Cyclonic flow is anticlockwise around a low pressure region and anticyclonic flow is clockwise around a high. If f were negative, as in the Southern Hemisphere, (anti)cyclonic flow would be (anti)clockwise.

- Geostrophic flow is parallel to lines of constant pressure (isobars). If $f > 0$ the flow is anticlockwise round a region of low pressure and clockwise around a region of high pressure (see Fig. 1.3).
- If the Coriolis force is constant and if the density does not vary in the horizontal the geostrophic flow is horizontally non-divergent and

$$\nabla_z \cdot \boldsymbol{u}_g = \frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} = 0. \tag{1.101}$$

We may define the *geostrophic streamfunction,* ψ, by

$$\psi \equiv \frac{p}{f_0 \rho_0}, \tag{1.102}$$

whence

$$u_g = -\frac{\partial \psi}{\partial y}, \qquad v_g = \frac{\partial \psi}{\partial x}. \tag{1.103}$$

The vertical component of vorticity, ζ, is then given by

$$\zeta = \mathbf{k} \cdot \nabla \times \boldsymbol{v} = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla_z^2 \psi. \tag{1.104}$$

- If the Coriolis parameter is not constant, then cross-differentiating (1.100) gives, for constant density geostrophic flow,

$$v_g \frac{\partial f}{\partial y} + f \nabla_z \cdot \boldsymbol{u}_g = 0, \qquad (1.105)$$

which, using the mass continuity equation $\nabla_z \cdot \boldsymbol{u}_g = -\partial w / \partial z$,

$$\beta v_g = f \frac{\partial w}{\partial z}. \qquad (1.106)$$

where $\beta \equiv \partial f / \partial y = 2\Omega \cos\vartheta / a$. This geostrophic vorticity balance is sometimes known as 'Sverdrup balance', although that expression is better restricted to the case when the vertical velocity from a wind stress.

1.8.3 Taylor–Proudman effect

If $\beta = 0$, then (1.106) implies that the vertical velocity is not a function of height. In fact, in that case none of the components of velocity vary with height if density is also constant. To show this, in the limit of zero Rossby number we first write the three-dimensional momentum equation as

$$\boldsymbol{f}_0 \times \boldsymbol{v} = -\nabla\phi - \nabla\chi, \qquad (1.107)$$

where $\boldsymbol{f}_0 = 2\boldsymbol{\Omega} = 2\Omega\mathbf{k}$, $\phi = p/\rho_0$, and $\nabla\chi$ represents other potential forces. If $\chi = gz$ then the vertical component of this equation represents hydrostatic balance, and the horizontal components represent geostrophic balance. On taking the curl of this equation, the terms on the right-hand side vanish and the left-hand side becomes

$$(\boldsymbol{f}_0 \cdot \nabla)\boldsymbol{v} - \boldsymbol{f}_0 \nabla\cdot \boldsymbol{v} - (\boldsymbol{v} \cdot \nabla)\boldsymbol{f}_0 + \boldsymbol{v}\nabla\cdot \boldsymbol{f}_0 = 0. \qquad (1.108)$$

But $\nabla\cdot \boldsymbol{v} = 0$ by mass conservation, and because $\boldsymbol{f}_0$ is constant both $\nabla\cdot \boldsymbol{f}_0$ and $(\boldsymbol{v} \cdot \nabla)\boldsymbol{f}_0$ vanish. Equation (1.108) thus reduces to

$$(\boldsymbol{f}_0 \cdot \nabla)\boldsymbol{v} = 0, \qquad (1.109)$$

which, since $\boldsymbol{f}_0 = f_0\mathbf{k}$, implies $f_0 \partial \boldsymbol{v} / \partial z = 0$, and in particular we have

$$\frac{\partial u}{\partial z} = 0, \quad \frac{\partial v}{\partial z} = 0, \quad \frac{\partial w}{\partial z} = 0. \qquad (1.110)$$

A different presentation of this argument proceeds as follows. If the flow is exactly in geostrophic and hydrostatic balance then

$$v = \frac{1}{f_0}\frac{\partial \phi}{\partial x}, \qquad u = -\frac{1}{f_0}\frac{\partial \phi}{\partial y}, \qquad \frac{\partial \phi}{\partial z} = -g. \quad \text{(1.111a,b,c)}$$

Differentiating (1.111a,b) with respect to z, and using (1.111c) yields

$$\frac{\partial v}{\partial z} = \frac{-1}{f_0}\frac{\partial g}{\partial x} = 0, \quad \frac{\partial u}{\partial z} = \frac{1}{f_0}\frac{\partial g}{\partial y} = 0. \qquad (1.112)$$

Noting that the geostrophic velocities are horizontally non-divergent ($\nabla_z \cdot \boldsymbol{u} = 0$), and using mass continuity then gives $\partial w/\partial z = 0$, as before.

If there is a solid horizontal boundary anywhere in the fluid, for example at the surface, then $w = 0$ at that surface and thus $w = 0$ everywhere. Hence the motion occurs in planes that lie perpendicular to the axis of rotation, and the flow is effectively two dimensional. This result is known as the *Taylor–Proudman effect,* namely that for constant density flow in geostrophic and hydrostatic balance the vertical derivatives of the horizontal and the vertical velocities are zero. At zero Rossby number, if the vertical velocity is zero somewhere in the flow, it is zero everywhere in that vertical column; furthermore, the horizontal flow has no vertical shear, and the fluid moves like a slab. The effects of rotation have provided a *stiffening* of the fluid in the vertical.

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

1.8.4 Thermal wind balance

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

$$-fv_g = -\frac{\partial \phi}{\partial x} = -\frac{1}{a\cos\vartheta}\frac{\partial \phi}{\partial \lambda}, \qquad fu_g = -\frac{\partial \phi}{\partial y} = -\frac{1}{a}\frac{\partial \phi}{\partial \vartheta}. \tag{1.113a,b}$$

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

$$-f\frac{\partial v_g}{\partial z} = -\frac{\partial b}{\partial x}, \tag{1.114a}$$

$$f\frac{\partial u_g}{\partial z} = -\frac{\partial b}{\partial y}. \tag{1.114b}$$

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

If the density or buoyancy is constant then there is no shear and (1.114b) gives the Taylor–Proudman result. But suppose that the temperature falls in the poleward direction. Then thermal wind balance implies that the (eastward) wind will increase with height — just as is observed in the atmosphere! In general, a vertical shear of the horizontal wind is associated with a horizontal temperature gradient, and this is one of the most simple and far-reaching effects in geophysical fluid dynamics. The underlying physical mechanism is illustrated in Fig. 1.4.

1.8.5 Vertical velocity and hydrostatic balance

Scaling for vertical velocity

If the Coriolis parameter is constant then the horizontal component of flows that are in geostrophic balance have zero horizontal divergence ($\nabla_x \cdot \boldsymbol{u} = 0$) and zero vertical velocity. We can therefore expect that any flow with small Rossby number will have a 'small' vertical velocity. Let us make this statement more precise using the rotating

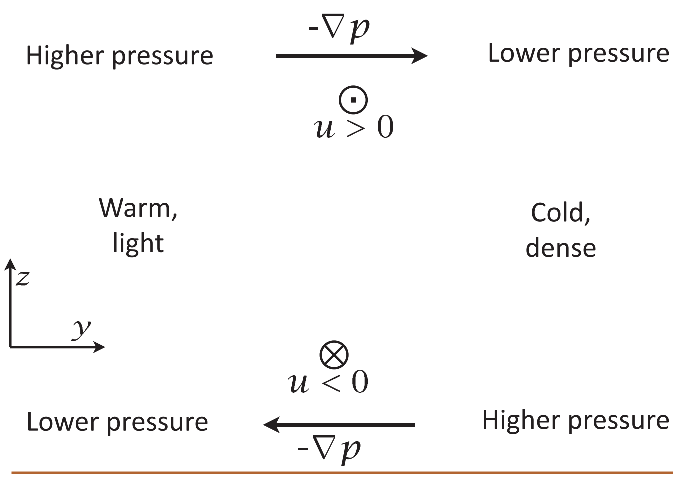

Figure 1.4 The mechanism of thermal wind. A cold fluid is denser than a warm fluid, so by hydrostasy the vertical pressure gradient is greater where the fluid is cold. Thus, the pressure gradients form as shown, where 'higher' and 'lower' mean relative to the average at that height. The horizontal pressure gradients are balanced by the Coriolis force, producing (for $f > 0$) the horizontal winds shown (⊗ into the paper, and ⊙ out of the paper). Only the wind *shear* is given by the thermal wind.

Boussinesq equations, (1.86) with constant Coriolis parameter. Let $\boldsymbol{u} = \boldsymbol{u}_g + \boldsymbol{u}_a$ where the geostrophic flow satisfies $\boldsymbol{f}_0 \times \boldsymbol{u}_g = -\nabla\phi$. The horizontal momentum equation, with corresponding scales for each term, then becomes

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + w\frac{\partial \boldsymbol{u}}{\partial z} + \boldsymbol{f}_0 \times \boldsymbol{u}_a = 0, \tag{1.115}$$

$$\frac{U^2}{L} \qquad \frac{U^2}{L} \qquad \frac{WU}{H} \qquad f_0 U_a. \tag{1.116}$$

This equation suggests a scaling for the ageostrophic flow of

$$U_a = \frac{U}{f_0 L} U = RoU. \tag{1.117}$$

That is, the ageostrophic flow is Rossby number smaller (at least) than the geostrophic flow. To obtain a scaling for the vertical velocity we look to the mass continuity equation written in the form

$$\frac{\partial w}{\partial z} = -\nabla\cdot\boldsymbol{u}_a, \tag{1.118}$$

since only the ageostrophic flow has a divergence. Equations (1.117) and (1.118) suggest the scaling

$$W = Ro\frac{HU}{L}. \tag{1.119}$$

That is, the vertical velocity is order Rossby number smaller than an estimate based purely on the mass continuity equation would suggest.

If the Coriolis parameter is not constant then the geostrophic flow itself is divergent and this induces a vertical velocity, as in (1.106). The scaling for vertical velocity is now

$$W = \frac{\beta}{f}HU = Ro_\beta \frac{HU}{L}. \tag{1.120}$$

where $Ro_\beta = \beta L/f$ is the *beta Rossby number*. It is less than one for all flows except those with a truly global scale.

Scaling for hydrostatic balance

Let us non-dimensionalize the rotating Boussinesq equations, (1.86), by writing

$$\begin{gathered}(x, y) = L(\widehat{x}, \widehat{y}), \qquad z = H\widehat{z}, \qquad \boldsymbol{u} = U\widehat{\boldsymbol{u}}, \\ t = T\widehat{t} = \frac{L}{U}\widehat{t}, \qquad \boldsymbol{f} = f_0\widehat{\boldsymbol{f}}, \\ w = \frac{\epsilon HU}{L}\widehat{w}, \qquad \phi = \Phi\widehat{\phi} = f_0 UL\widehat{\phi}, \\ b = B\widehat{b} = \frac{f_0 UL}{H}\widehat{b},\end{gathered} \tag{1.121}$$

These relations are almost the same as (1.92), except for the factor of ϵ in the scaling of w. If the Coriolis parameter is constant or nearly so then, from (1.119), $\epsilon = Ro$, whereas if the Coriolis parameter varies then $\epsilon = Ro_\beta$, as in (1.119). The scaling for ϕ and b' are suggested by geostrophic and thermal wind balance with f_0 a representative value of f. Substituting these values into (1.86) we obtain the following scaled momentum equations:

$$Ro\frac{\mathrm{D}\widehat{\boldsymbol{u}}}{\mathrm{D}\widehat{t}} + \widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}} = -\nabla\widehat{\phi}, \qquad Ro\epsilon\alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}\widehat{t}} = -\frac{\partial\widehat{\phi}}{\partial\widehat{z}} - \widehat{b}. \tag{1.122a,b}$$

where $D/D\hat{t} = \partial/\partial\hat{t} + \hat{\boldsymbol{u}} \cdot \nabla_z + \epsilon\hat{w}\partial/\partial\hat{z}$. There are two notable aspects to these equations. First and most obviously, when $Ro \ll 1$, (1.122a) reduces to geostrophic balance, $\boldsymbol{f} \times \boldsymbol{u} \approx -\nabla\hat{\phi}$. Second, the material derivative in (1.122b) is multiplied by three non-dimensional parameters, and we can understand the appearance of each as follows.

(i) The aspect ratio dependence (α^2) arises in the same way as for non-rotating flows — that is, because of the presence of w and z in the vertical momentum equation as opposed to (u, v) and (x, y) in the horizontal equations.

(ii) The Rossby number dependence (Ro) arises because in rotating flow the pressure gradient is balanced by the Coriolis force, which is Rossby number larger than the advective terms.

(iii) The factor ϵ arises because in rotating flow w is smaller than u by ϵ times the aspect ratio. The factor may be the Rossby number itself, or the beta Rossby number.

The factor $Ro\epsilon\alpha^2$ is very small for large-scale flow; the reader is invited to calculate representative values. Evidently, a rapidly rotating fluid is more likely to be in hydrostatic balance than a non-rotating fluid, other conditions being equal. The combined effects of rotation and stratification are, not surprisingly, quite subtle and we leave that topic for chapter 3.

Chapter 2

Shallow Water Systems

Weeks 3 to 5

2.1 Dynamics of a Single, Shallow Layer

Consider fluid in a container above which is another fluid of negligible density (and therefore negligible inertia) relative to the fluid of interest, as illustrated in Fig. 2.1. Our notation is that $\boldsymbol{v} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$ is the three-dimensional velocity and $\boldsymbol{u} = u\mathbf{i} + v\mathbf{j}$ is the horizontal velocity. $h(x, y)$ is the thickness of the liquid column, H is its mean height, and η is the height of the free surface. In a flat-bottomed container $\eta = h$, whereas in general $h = \eta - \eta_b$, where η_b is the height of the floor of the container.

2.1.1 Momentum equations

The vertical momentum equation is just the hydrostatic equation,

$$\frac{\partial p}{\partial z} = -\rho_0 g, \tag{2.1}$$

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

$$p(x, y, z, t) = -\rho_0 g z + p_o. \tag{2.2}$$

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

$$p(x, y, z, t) = \rho_0 g(\eta(x, y, t) - z). \tag{2.3}$$

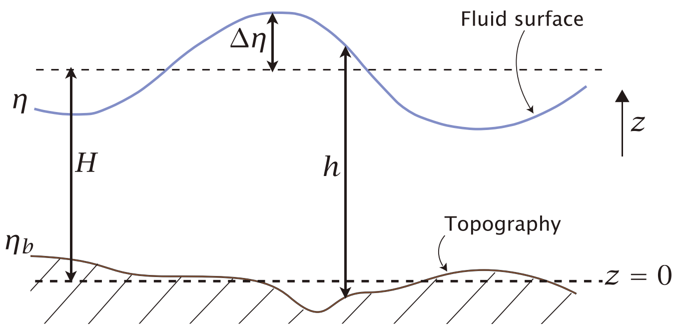

Figure 2.1 A shallow water system. h is the thickness of a water column, H its mean thickness, η the height of the free surface and η_b is the height of the lower, rigid, surface, above some arbitrary origin, typically chosen such that the average of η_b is zero. $\Delta\eta$ is the deviation free surface height, so we have $\eta = \eta_b + h = H + \Delta\eta$.

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

$$\nabla_z p = \rho_0 g \nabla_z \eta, \tag{2.4}$$

where

$$\nabla_z = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} \tag{2.5}$$

is the gradient operator at constant z. (In the rest of this chapter we will drop the subscript z unless that causes ambiguity. The three-dimensional gradient operator will be denoted by ∇_3. We will also mostly use Cartesian coordinates, but the shallow water equations may certainly be applied over a spherical planet — indeed, 'Laplace's tidal equations' are essentially the shallow water equations on a sphere.) The horizontal momentum equations therefore become

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\frac{1}{\rho_0}\nabla p = -g\nabla\eta. \tag{2.6}$$

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

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = \frac{\partial \boldsymbol{u}}{\partial t} + u\frac{\partial \boldsymbol{u}}{\partial x} + v\frac{\partial \boldsymbol{u}}{\partial y} = -g\nabla\eta. \tag{2.7}$$

That the horizontal velocity is independent of z is a consequence of the hydrostatic equation, which ensures that the horizontal pressure gradient is independent of height. (Another starting point would be to take this independence of the horizontal motion with height as the *definition* of shallow water flow. In real physical situations such independence does not hold exactly — for example, friction at the bottom may induce a vertical dependence of the flow in a boundary layer.) In the presence of rotation, (2.7) easily generalizes to

$$\boxed{\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u} = -g\nabla\eta}\ , \tag{2.8}$$

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

2.1.2 Mass continuity equation

The mass contained in a fluid column of height h and cross-sectional area A is given by $\int_A \rho_0 h \,\mathrm{d}\boldsymbol{A}$ (see Fig. 2.2). If there is a net flux of fluid across the column boundary (by advection) then this must be balanced by a net increase in the mass in A, and therefore a net increase in the height of the water column. The mass convergence into the column is given by

$$F_m = \text{mass flux in} = -\int_S \rho_0 \boldsymbol{u}\cdot\,\mathrm{d}\boldsymbol{S}, \tag{2.9}$$

where S is the area of the vertical boundary of the column. The surface area of the column is composed of elements of area $h\boldsymbol{n}\,\delta l$, where δl is a line element circumscribing the column and $\boldsymbol{n}$ is a unit vector perpendicular to the boundary, pointing outwards. Thus (2.9) becomes

$$F_m = -\oint \rho_0 h\boldsymbol{u}\cdot\boldsymbol{n}\,\mathrm{d}l. \tag{2.10}$$

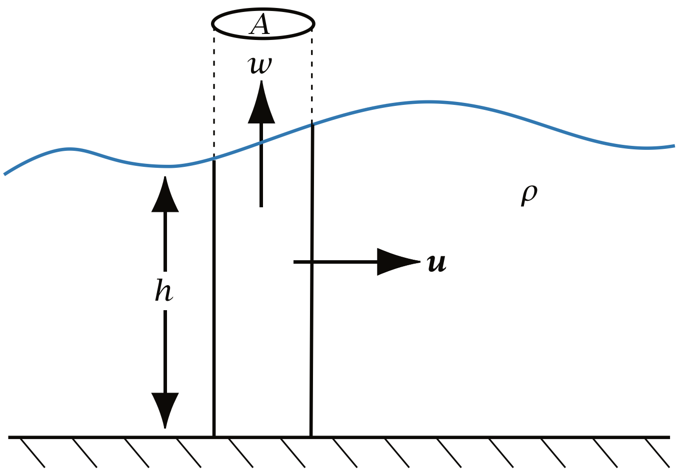

Figure 2.2 The mass budget for a column of area A in a shallow water system. The fluid leaving the column is $\oint \rho_0 h \boldsymbol{u} \cdot \boldsymbol{n} \, \mathrm{d}l$ where $\boldsymbol{n}$ is the unit vector normal to the boundary of the fluid column. There is a non-zero vertical velocity at the top of the column if the mass convergence into the column is non-zero.

Using the divergence theorem in two dimensions, (2.10) simplifies to

$$F_m = -\int_A \nabla \cdot (\rho_0 \boldsymbol{u} h) \, \mathrm{d}A, \tag{2.11}$$

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

$$F_m = \frac{\mathrm{d}}{\mathrm{d}t} \int \rho_0 \, \mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t} \int_A \rho_0 h \, \mathrm{d}A = \int_A \rho_0 \frac{\partial h}{\partial t} \, \mathrm{d}A. \tag{2.12}$$

Because ρ_0 is constant, the balance between (2.11) and (2.12) leads to

$$\int_A \left[\frac{\partial h}{\partial t} + \nabla \cdot (\boldsymbol{u} h) \right] \mathrm{d}A = 0, \tag{2.13}$$

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

$$\frac{\partial h}{\partial t} + \nabla \cdot (\boldsymbol{u} h) = 0, \tag{2.14}$$

The Shallow Water Equations

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

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

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

where $\boldsymbol{u}$ is the horizontal velocity, h is the total fluid thickness, η is the height of the upper free surface and η_b is the height of the lower surface (the bottom topography). Thus, $h(x, y, t) = \eta(x, y, t) - \eta_b(x, y)$. The material derivative is

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \boldsymbol{u}\cdot\nabla = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}, \tag{SW.3}$$

with the rightmost expression holding in Cartesian coordinates.

or equivalently

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla\cdot\boldsymbol{u} = 0. \tag{2.15}$$

This derivation holds whether or not the lower surface is flat. If it is, then $h = \eta$, and if not $h = \eta - \eta_b$.

From the 3D mass conservation equation

Since the fluid is incompressible, the three-dimensional mass continuity equation is just $\nabla\cdot\boldsymbol{v} = 0$. Writing this out in component form

$$\frac{\partial w}{\partial z} = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = -\nabla\cdot\boldsymbol{u}. \tag{2.16}$$

Integrating this from the bottom of the fluid to the top, and using the boundary conditions of w (express w in terms of h) gives (2.15). Details left to the reader.

2.2 Reduced Gravity Equations

Consider now a single shallow moving layer of fluid on top of a deep, quiescent fluid layer (Fig. 2.3), and beneath a fluid of negligible inertia. This configuration is often used as a model of the upper ocean: the upper layer represents flow in perhaps the upper few hundred metres of the ocean, the lower layer being the near-stagnant abyss. If we turn the model upside-down we have a perhaps slightly less realistic model of the atmosphere: the lower layer represents motion in the troposphere above which lies an inactive stratosphere. The equations of motion are virtually the same in both cases.

2.2.1 Pressure gradient in the active layer

We will derive the equations for the oceanic case (active layer on top) in two cases, which differ slightly in the assumption made about the upper surface.

I Free upper surface

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

$$p_1(z) = g\rho_1(\eta_0 - z), \tag{2.17}$$

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

$$\frac{1}{\rho_1}\nabla p_1 = g\nabla\eta_0, \tag{2.18}$$

and the momentum equation is

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u} = -g\nabla\eta_0. \tag{2.19}$$

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

$$p_2(z) = \rho_1 g(\eta_0 - \eta_1) + \rho_2 g(\eta_1 - z). \tag{2.20}$$

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

$$\rho_1 g\eta_0 = -\rho_1 g'\eta_1 + \text{constant}, \tag{2.21}$$

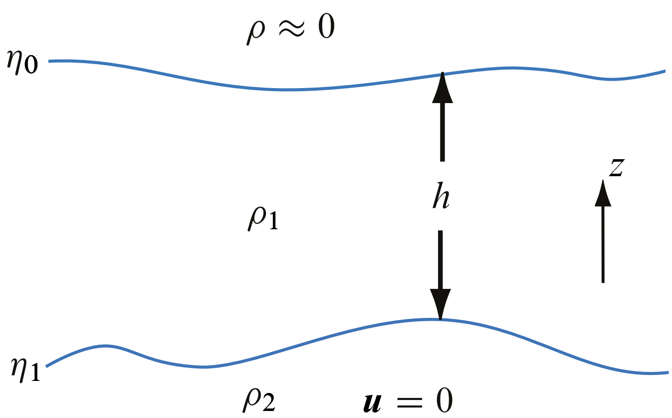

Figure 2.3 The reduced gravity shallow water system. An active layer lies over a deep, more dense, quiescent layer. In a common variation the upper surface is held flat by a rigid lid, and $\eta_0 = 0$.

where $g' = g(\rho_2 - \rho_1)/\rho_1$ is the *reduced gravity*. The momentum equation becomes

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = g' \nabla \eta_1. \tag{2.22}$$

The equations are completed by the usual mass conservation equation,

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h \nabla \cdot \boldsymbol{u} = 0, \tag{2.23}$$

where $h = \eta_0 - \eta_1$. Because $g \gg g'$, (2.21) shows that surface displacements are *much smaller* than the displacements at the interior interface. We see this in the real ocean where the mean interior isopycnal displacements may be several tens of metres but variations in the mean height of ocean surface are of the order of centimetres.

II The rigid lid approximation

The smallness of the upper surface displacement suggests that we will make little error is we impose a *rigid lid* at the top of the fluid. Displacements are no longer allowed, but the lid will in general impart a pressure force to the fluid. Suppose that this is $P(x, y, t)$, then the horizontal pressure gradient in the upper layer is simply

$$\nabla p_1 = \nabla P. \tag{2.24}$$

The pressure in the lower layer is again given by hydrostasy, and is

$$p_2 = -\rho_1 g\eta_1 + \rho_2 g(\eta_1 - z) + P = \rho_1 gh - \rho_2 g(h+z) + P, \quad (2.25)$$

so that

$$\nabla p_2 = -g(\rho_2 - \rho_1)\nabla h + \nabla P. \quad (2.26)$$

Then if $\nabla p_2 = 0$ we have

$$g(\rho_2 - \rho_1)\nabla h = \nabla P, \quad (2.27)$$

and the momentum equation for the upper layer is just

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -g_1'\nabla h. \quad (2.28)$$

where $g' = g(\rho_2 - \rho_1)/\rho_1$. These equations differ from the usual shallow water equations only in the use of a reduced gravity g' in place of g itself. It is the density *difference* between the two layers that is important. Similarly, if we take a shallow water system, with the moving layer on the bottom, and we suppose that overlying it is a stationary fluid of finite density, then we would easily find that the fluid equations for the moving layer are the same as if the fluid on top had zero inertia, except that g would be replaced by an appropriate reduced gravity (problem 2.**??**).

2.3 Geostrophic Balance

Geostrophic balance occurs in the shallow water equations, just as in the continuously stratified equations, when the Rossby number U/fL is small and the Coriolis term dominates the advective terms in the momentum equation. In the single-layer shallow water equations the geostrophic flow is:

$$\boldsymbol{f} \times \boldsymbol{u}_g = -g\nabla\eta. \quad (2.29)$$

Thus, the geostrophic velocity is proportional to the slope of the surface, as sketched in Fig. 2.4. (For the rest of this section, we will drop the subscript g, and take all velocities to be geostrophic.)

In both the single-layer and multi-layer cases, the slope of an interfacial surface is directly related to the difference

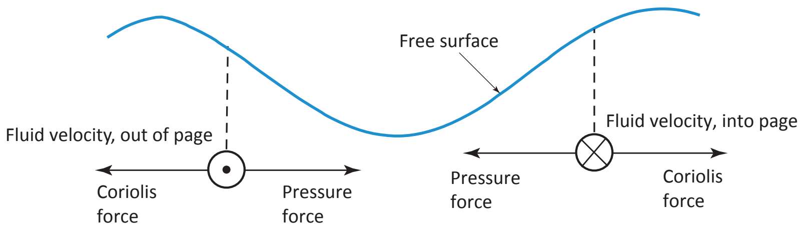

Figure 2.4 Geostrophic flow in a shallow water system, with a positive value of the Coriolis parameter f, as in the Northern Hemisphere. The pressure force is directed down the gradient of the height field, and this can be balanced by the Coriolis force if the fluid velocity is at right angles to it. If f were negative, the geostrophic flow would be reversed.

in pressure gradient on either side and so, by geostrophic balance, to the shear of the flow. This is the shallow water analogue of the thermal wind relation. To obtain an expression for this, consider the interface, η, between two layers labelled 1 and 2. The pressure in two layers is given by the hydrostatic relation and so,

$$p_1 = A(x, y) - \rho_1 g z \qquad \text{(at some } z \text{ in layer 1)} \tag{2.30a}$$

$$\begin{aligned} p_2 &= A(x, y) - \rho_1 g\eta + \rho_2 g(\eta - z) \\ &= A(x, y) + \rho_1 g_1'\eta - \rho_2 g z \qquad \text{(at some } z \text{ in layer 2)} \end{aligned} \tag{2.30b}$$

where $A(x, y)$ is a function of integration. Thus we find

$$\frac{1}{\rho_1}\nabla(p_1 - p_2) = -g_1'\nabla\eta. \tag{2.31}$$

If the flow is geostrophically balanced and Boussinesq then, in each layer, the velocity obeys

$$f\boldsymbol{u}_i = \frac{1}{\rho_1}\mathbf{k}\times\nabla p_i. \tag{2.32}$$

Using (2.31) then gives

$$f(\boldsymbol{u}_1 - \boldsymbol{u}_2) = -\mathbf{k}\times g_1'\nabla\eta, \tag{2.33}$$

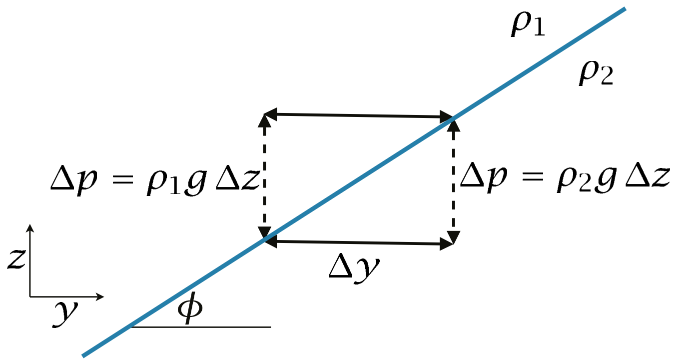

Figure 2.5 Margules' relation: using hydrostasy, the difference in the horizontal pressure gradient between the upper and the lower layer is given by $-g'\rho_1 s$, where $s = \tan\phi = \Delta z/\Delta y$ is the interface slope and $g' = g(\rho_2 - \rho_1)/\rho_1$. Geostrophic balance then gives $f(u_1 - u_2) = g's$, which is a special case of (2.34).

or in general

$$f(\boldsymbol{u}_n - \boldsymbol{u}_{n+1}) = -\mathbf{k} \times g_n' \nabla \eta. \tag{2.34}$$

This is the thermal wind equation for the shallow water system. It applies at any interface, and it implies *the shear is proportional to the interface slope*, a result known as the 'Margules relation' (Fig. 2.5).

Suppose that we represent the atmosphere by two layers of fluid; a meridionally decreasing temperature may then be represented by an interface that slopes upwards toward the pole. Then, in either hemisphere, we have

$$u_1 - u_2 = \frac{g_1'}{f}\frac{\partial \eta}{\partial y} > 0, \tag{2.35}$$

and the temperature gradient is associated with a positive shear (see problem 2.??).

2.4 Conservation Properties of Shallow Water Systems

There are two common types of conservation property in fluids: (i) material invariants and (ii) integral invariants. Material invariance occurs when a property (ϕ say) is conserved on each fluid element, and so obeys the equation

$D\phi/Dt = 0$. An integral invariant is one that is conserved following an integration over some, usually closed, volume; energy is an example.

2.4.1 Potential vorticity: a material invariant

The vorticity of a fluid, denoted $\boldsymbol{\omega}$, is defined to be the curl of the velocity field. Let us also define the shallow water vorticity, $\boldsymbol{\omega}^*$, as the curl of the horizontal velocity. We therefore have:

$$\boldsymbol{\omega} \equiv \nabla \times \boldsymbol{v}, \qquad \boldsymbol{\omega}^* \equiv \nabla \times \boldsymbol{u}. \tag{2.36}$$

Because $\partial u/\partial z = \partial v/\partial z = 0$, only the vertical component of $\boldsymbol{\omega}^*$ is non-zero and

$$\boldsymbol{\omega}^* = \mathbf{k}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) \equiv \mathbf{k}\,\zeta. \tag{2.37}$$

Considering first the non-rotating case, we use the vector identity

$$(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = \frac{1}{2}\nabla(\boldsymbol{u}\cdot\boldsymbol{u}) - \boldsymbol{u}\times(\nabla\times\boldsymbol{u}), \tag{2.38}$$

to write the momentum equation, (2.8) with $f = 0$, as

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{\omega}^* \times \boldsymbol{u} = -\nabla\left(g\eta + \frac{1}{2}\boldsymbol{u}^2\right). \tag{2.39}$$

To obtain an evolution equation for the vorticity we take the curl of (2.39), and make use of the vector identity

$$\begin{aligned}\nabla\times(\boldsymbol{\omega}^*\times\boldsymbol{u}) &= (\boldsymbol{u}\cdot\nabla)\boldsymbol{\omega}^* - (\boldsymbol{\omega}^*\cdot\nabla)\boldsymbol{u} + \boldsymbol{\omega}^*\nabla\cdot\boldsymbol{u} - \boldsymbol{u}\nabla\cdot\boldsymbol{\omega}^* \\ &= (\boldsymbol{u}\cdot\nabla)\boldsymbol{\omega}^* + \boldsymbol{\omega}^*\nabla\cdot\boldsymbol{u},\end{aligned} \tag{2.40}$$

using the fact that $\nabla\cdot\boldsymbol{\omega}^*$ is the divergence of a curl and therefore zero, and $(\boldsymbol{\omega}^*\cdot\nabla)\boldsymbol{u} = 0$ because $\boldsymbol{\omega}^*$ is perpendicular to the surface in which $\boldsymbol{u}$ varies. Taking the curl of (2.39) gives

$$\frac{\partial \zeta}{\partial t} + (\boldsymbol{u}\cdot\nabla)\zeta = -\zeta\nabla\cdot\boldsymbol{u}, \tag{2.41}$$

where $\zeta = \mathbf{k}\cdot\boldsymbol{\omega}^*$. Now, the mass conservation equation may be written as

$$-\zeta\nabla\cdot\boldsymbol{u} = \frac{\zeta}{h}\frac{Dh}{Dt}, \tag{2.42}$$

and using this (2.41) becomes

$$\frac{D\zeta}{Dt} = \frac{\zeta}{h}\frac{Dh}{Dt}, \tag{2.43}$$

which simplifies to

$$\frac{D}{Dt}\left(\frac{\zeta}{h}\right) = 0. \tag{2.44}$$

The important quantity ζ/h, often denoted by Q, is known as the *potential vorticity*, and (2.44) is the potential vorticity equation. We re-derive this conservation law in a more general way in section **??**

Because Q is conserved on parcels, then so is any function of Q; that is, $F(Q)$ is a material invariant, where F is any function. To see this algebraically, multiply (2.44) by $F'(Q)$, the derivative of F with respect to Q, giving

$$F'(Q)\frac{DQ}{Dt} = \frac{D}{Dt}F(Q) = 0. \tag{2.45}$$

Since F is arbitrary there are an infinite number of material invariants corresponding to different choices of F.

Effects of rotation

In a rotating frame of reference, the shallow water momentum equation is

$$\frac{D\boldsymbol{u}}{Dt} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \tag{2.46}$$

where (as before) $\boldsymbol{f} = f\mathbf{k}$. This may be written in vector invariant form as

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{\omega}^* + \boldsymbol{f}) \times \boldsymbol{u} = -\nabla\left(g\eta + \frac{1}{2}\boldsymbol{u}^2\right), \tag{2.47}$$

and taking the curl of this gives the vorticity equation

$$\frac{\partial \zeta}{\partial t} + (\boldsymbol{u}\cdot\nabla)(\zeta + f) = -(f+\zeta)\nabla\cdot\boldsymbol{u}. \tag{2.48}$$

This is the same as the shallow water vorticity equation in a non-rotating frame, save that ζ is replaced by $\zeta + f$, the reason for this being that f is the vorticity that the fluid has by virtue of the background rotation. Thus, (2.48) is simply the equation of motion for the total or absolute vorticity, $\boldsymbol{\omega}_a = \boldsymbol{\omega}^* + \boldsymbol{f} = (\zeta + f)\mathbf{k}$.

The potential vorticity equation in the rotating case follows, much as in the non-rotating case, by combining (2.48) with the mass conservation equation, giving

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{\zeta + f}{h}\right) = 0 \quad . \tag{2.49}$$

That is, $Q \equiv (\zeta + f)/h$, the potential vorticity in a rotating shallow system, is a material invariant.

Vorticity and circulation

Although vorticity itself is not a material invariant, its integral over a horizontal material area is invariant. To demonstrate this in the non-rotating case, consider the integral

$$C = \int_A \zeta \,\mathrm{d}A = \int_A Qh \,\mathrm{d}A, \tag{2.50}$$

over a surface A, the cross-sectional area of a column of height h (as in Fig. 2.2). Taking the material derivative of this gives

$$\frac{\mathrm{D}C}{\mathrm{D}t} = \int_A \frac{\mathrm{D}Q}{\mathrm{D}t} h \,\mathrm{d}A + \int_A Q \frac{\mathrm{D}}{\mathrm{D}t}(h \,\mathrm{d}A). \tag{2.51}$$

The first term is zero, by (2.43); the second term is just the derivative of the volume of a column of fluid and it too is zero, by mass conservation. Thus,

$$\frac{\mathrm{D}C}{\mathrm{D}t} = \frac{\mathrm{D}}{\mathrm{D}t}\int_A \zeta \,\mathrm{d}A = 0. \tag{2.52}$$

Thus, the integral of the vorticity over a some cross-sectional area of the fluid is unchanging, although both the vorticity and area of the fluid may individually change. Using Stokes' theorem, it may be written as

$$\frac{\mathrm{D}C}{\mathrm{D}t} = \frac{\mathrm{D}}{\mathrm{D}t}\oint \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{l}, \tag{2.53}$$

where the line integral is around the boundary of A. This is an example of Kelvin's circulation theorem, which we shall meet again in a more general form in chapter **??**, where we also consider the rotating case.

A slight generalization of (2.52) is possible. Consider the integral $I = \int F(Q)h\,\mathrm{d}A$ where again F is any differentiable function of its argument. It is clear that

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_A F(Q)h\,\mathrm{d}A = 0. \tag{2.54}$$

If the area of integration in (2.39) or (2.54) is the whole domain (enclosed by frictionless walls, for example) then it is clear that the integral of $hF(Q)$ is a constant, including as a special case the integral of ζ.

2.4.2 Energy conservation: an integral invariant

Since we have made various simplifications in deriving the shallow water system, it is not self-evident that energy should be conserved, or indeed what form the energy takes. The kinetic energy density (KE), that is the kinetic energy per unit area, is $\rho_0 h\boldsymbol{u}^2/2$. The potential energy density of the fluid is

$$PE = \int_0^h \rho_0 g z\,\mathrm{d}z = \frac{1}{2}\rho_0 g h^2. \tag{2.55}$$

The factor ρ_0 appears in both kinetic and potential energies and, because it is a constant, we will omit it. For algebraic simplicity we also assume the bottom is flat, at $z = 0$.

Using the mass conservation equation (2.15) we obtain an equation for the evolution of potential energy density, namely

$$\frac{\mathrm{D}}{\mathrm{D}t}\frac{gh^2}{2} + gh^2\nabla\cdot\boldsymbol{u} = 0 \tag{2.56a}$$

or

$$\frac{\partial}{\partial t}\frac{gh^2}{2} + \nabla\cdot\left(\boldsymbol{u}\frac{gh^2}{2}\right) + \frac{gh^2}{2}\nabla\cdot\boldsymbol{u} = 0. \tag{2.56b}$$

From the momentum and mass continuity equations we obtain an equation for the evolution of kinetic energy density, namely

$$\frac{\mathrm{D}}{\mathrm{D}t}\frac{h\boldsymbol{u}^2}{2} + \frac{\boldsymbol{u}^2 h}{2}\nabla\cdot\boldsymbol{u} = -g\boldsymbol{u}\cdot\nabla\frac{h^2}{2} \tag{2.57a}$$

or

$$\frac{\partial}{\partial t}\frac{h\boldsymbol{u}^2}{2} + \nabla\cdot\left(\boldsymbol{u}\frac{h\boldsymbol{u}^2}{2}\right) + g\boldsymbol{u}\cdot\nabla\frac{h^2}{2} = 0. \tag{2.57b}$$

Adding (2.56b) and (2.57b) we obtain

$$\frac{\partial}{\partial t}\frac{1}{2}\left(h\boldsymbol{u}^2 + gh^2\right)+\nabla\cdot\left[\frac{1}{2}\boldsymbol{u}\left(gh^2 + h\boldsymbol{u}^2 + gh^2\right)\right] = 0, \tag{2.58}$$

or

$$\frac{\partial E}{\partial t} + \nabla\cdot\boldsymbol{F} = 0, \tag{2.59}$$

where $E = KE + PE = (h\boldsymbol{u}^2 + gh^2)/2$ is the density of the total energy and $\boldsymbol{F} = \boldsymbol{u}(h\boldsymbol{u}^2/2 + gh^2)$ is the energy flux. If the fluid is confined to a domain bounded by rigid walls, on which the normal component of velocity vanishes, then on integrating (2.58) over that area and using Gauss's theorem, the total energy is seen to be conserved; that is

$$\frac{\mathrm{d}\widehat{E}}{\mathrm{d}t} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_A (h\boldsymbol{u}^2 + gh^2)\,\mathrm{d}A = 0. \tag{2.60}$$

Such an energy principle also holds in the case with bottom topography. Note that, as we found in the case for a compressible fluid in chapter **??**, the energy flux in (2.59) is not just the energy density multiplied by the velocity; it contains an additional term $guh^2/2$, and this represents the energy transfer occurring when the fluid does work against the pressure force (see problem 2.**??**).

2.5 Shallow Water Waves

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

2.5.1 Non-rotating shallow water waves

Given a flat bottom the fluid thickness is equal to the free surface displacement (Fig. 2.1), and taking the basic state of the fluid to be at rest we let

$$h(x, y, t) = H + h'(x, y, t) = H + \eta'(x, y, t), \tag{2.61a}$$

$$\boldsymbol{u}(x, y, t) = \boldsymbol{u}'(x, y, t). \tag{2.61b}$$

The mass conservation equation, (2.15), then becomes

$$\frac{\partial \eta'}{\partial t} + (H + \eta')\nabla \cdot \boldsymbol{u}' + \boldsymbol{u}' \cdot \nabla \eta' = 0, \tag{2.62}$$

and neglecting squares of small quantities this yields the linear equation

$$\frac{\partial \eta'}{\partial t} + H\nabla \cdot \boldsymbol{u}' = 0. \tag{2.63}$$

Similarly, linearizing the momentum equation, (2.8) with $\boldsymbol{f} = 0$, yields

$$\frac{\partial \boldsymbol{u}'}{\partial t} = -g\nabla \eta'. \tag{2.64}$$

Eliminating velocity by differentiating (2.63) with respect to time and taking the divergence of (2.64) leads to

$$\frac{\partial^2 \eta'}{\partial t^2} - gH\nabla^2 \eta' = 0, \tag{2.65}$$

which may be recognized as a wave equation. We can find the dispersion relationship for this by substituting the trial solution

$$\eta' = \operatorname{Re} \tilde{\eta} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.66}$$

where $\tilde{\eta}$ is a complex constant, $\boldsymbol{k} = \mathbf{i}k + \mathbf{j}l$ is the horizontal wavenumber and Re indicates that the real part of the solution should be taken. If, for simplicity, we restrict attention to the one-dimensional problem, with no variation in the y-direction, then substituting into (2.65) leads to the dispersion relationship

$$\omega = \pm ck, \tag{2.67}$$

where $c = \sqrt{gH}$; that is, the wave speed is proportional to the square root of the mean fluid depth and is independent of the wavenumber — the waves are dispersionless. The general solution is a superposition of all such waves, with the amplitudes of each wave (or Fourier component) being determined by the Fourier decomposition of the initial conditions.

Because the waves are dispersionless, the general solution can be written as

$$\eta'(x,t) = \frac{1}{2}\left[F(x-ct) + F(x+ct)\right], \tag{2.68}$$

where $F(x)$ is the height field at $t = 0$. From this, it is easy to see that the shape of an initial disturbance is preserved as it propagates both to the right and to the left at speed c, (see also problem 2.**??**).

2.5.2 Rotating shallow water (Poincaré) waves

We now consider the effects of rotation on shallow water waves. Linearizing the rotating, flat-bottomed f-plane shallow water equations [i.e., (SW.1) and (SW.2) on page 37] about a state of rest we obtain

$$\frac{\partial u'}{\partial t} - f_0 v' = -g\frac{\partial \eta'}{\partial x}, \qquad \frac{\partial v'}{\partial t} + f_0 u' = -g\frac{\partial \eta'}{\partial y},$$
$$\frac{\partial \eta'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0. \tag{2.69a,b,c}$$

To obtain a dispersion relationship we let

$$(u, v, \eta) = (\widetilde{u}, \widetilde{v}, \widetilde{\eta})\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{2.70}$$

and substitute into (2.69), giving

$$\begin{pmatrix} -\mathrm{i}\,\omega & -f_0 & \mathrm{i}\,gk \\ f_0 & -\mathrm{i}\,\omega & \mathrm{i}\,gl \\ \mathrm{i}\,Hk & \mathrm{i}\,Hl & -\mathrm{i}\,\omega \end{pmatrix}\begin{pmatrix} \widetilde{u} \\ \widetilde{v} \\ \widetilde{\eta} \end{pmatrix} = 0. \tag{2.71}$$

This homogeneous equation has non-trivial solutions only if the determinant of the matrix vanishes. This condition gives

$$\omega(\omega^2 - f_0^2 - c^2K^2) = 0. \tag{2.72}$$

where $K^2 = k^2 + l^2$ and $c^2 = gH$. There are two classes of solution to (2.72). The first is simply $\omega = 0$, i.e., time-independent flow corresponding to geostrophic balance in (2.69). Because geostrophic balance gives a divergence-free velocity field for a constant Coriolis parameter the

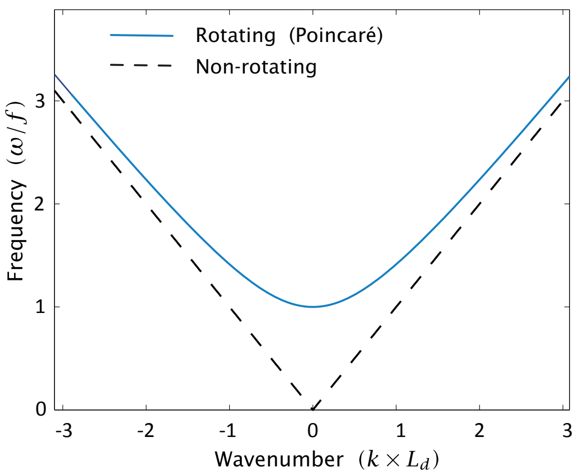

Figure 2.6 Dispersion relation for Poincaré waves (solid) and non-rotating shallow water waves (dashed). Frequency is scaled by the Coriolis frequency f, and wavenumber by the inverse deformation radius $\sqrt{gH}/f$. For small wavenumbers the frequency is approximately f; for high wavenumbers it asymptotes to that of non-rotating waves.

equations are satisfied by a time-independent solution. The second set of solutions gives the dispersion relation

$$\omega^2 = f_0^2 + c^2(k^2 + l^2), \tag{2.73}$$

or

$$\omega^2 = f_0^2 + gH(k^2 + l^2). \tag{2.74}$$

The corresponding waves are known as *Poincaré* waves,[2] and the dispersion relationship is illustrated in Fig. 2.6. Note that the frequency is always greater than the Coriolis frequency f_0. There are two interesting limits.

(i) The short wave limit. If

$$K^2 \gg \frac{f_0^2}{gH}, \tag{2.75}$$

where $K^2 = k^2 + l^2$, then the dispersion relationship reduces to that of the non-rotating case (2.67). This condition is equivalent to requiring that the wavelength be much shorter than the *deformation radius*,

$L_d \equiv \sqrt{gH}/f$. Specifically, if $l = 0$ and $\lambda = 2\pi/k$ is the wavelength, the condition is

$$\lambda^2 \ll L_d^2(2\pi)^2. \tag{2.76}$$

The numerical factor of $(2\pi)^2$ is more than an order of magnitude, so care must be taken when deciding if the condition is satisfied in particular cases. Furthermore, the wavelength must still be longer than the depth of the fluid, otherwise the shallow water condition is not met.

(ii) *The long wave limit.* If

$$K^2 \ll \frac{f_0^2}{gH}, \tag{2.77}$$

that is if the wavelength is much longer than the deformation radius L_d, then the dispersion relationship is

$$\omega = f_0. \tag{2.78}$$

These are known as *inertial oscillations*. The equations of motion giving rise to them are

$$\frac{\partial u'}{\partial t} - f_0 v' = 0, \qquad \frac{\partial v'}{\partial t} + f_0 u' = 0, \tag{2.79}$$

which are equivalent to material equations for free particles in a rotating frame, unconstrained by pressure forces, namely

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} - f_0 v = 0, \qquad \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + f_0 u = 0. \tag{2.80}$$

2.5.3 Kelvin waves

The Kelvin wave is a particular type of gravity wave that exists in the presence of both rotation and a lateral boundary. Suppose there is a solid boundary at $y = 0$; clearly harmonic solutions in the y-direction are not allowable, as these would not satisfy the condition of no normal flow at the boundary. Do any wave-like solutions exist? The affirmative answer to this question was provided by Kelvin and the associated waves are now eponymously known as

Kelvin waves.[3] We begin with the linearized shallow water equations, namely

$$\frac{\partial u'}{\partial t} - f_0 v' = -g\frac{\partial \eta'}{\partial x}, \qquad \frac{\partial v'}{\partial t} + f_0 u' = -g\frac{\partial \eta'}{\partial y},$$
$$\frac{\partial \eta'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0. \tag{2.81a,b,c}$$

The fact that $v' = 0$ at $y = 0$ suggests that we look for a solution with $v' = 0$ everywhere, whence these equations become

$$\frac{\partial u'}{\partial t} = -g\frac{\partial \eta'}{\partial x}, \qquad f_0 u' = -g\frac{\partial \eta'}{\partial y}, \qquad \frac{\partial \eta'}{\partial t} + H\frac{\partial u'}{\partial x} = 0. \tag{2.82a,b,c}$$

Equations (2.82a) and (2.82c) lead to the standard wave equation

$$\frac{\partial^2 u'}{\partial t^2} = c^2\frac{\partial^2 u'}{\partial x^2}, \tag{2.83}$$

where $c = \sqrt{gH}$, the usual wave speed of shallow water waves. The solution of (2.83) is

$$u' = F_1(x + ct, y) + F_2(x - ct, y), \tag{2.84}$$

with corresponding surface displacement

$$\eta' = \sqrt{H/g}\left[-F_1(x + ct, y) + F_2(x - ct, y)\right]. \tag{2.85}$$

The solution represents the superposition of two waves, one (F_1) travelling in the negative x-direction, and the other in the positive x-direction. To obtain the y dependence of these functions we use (2.82b) which gives

$$\frac{\partial F_1}{\partial y} = \frac{f_0}{\sqrt{gH}}F_1, \qquad \frac{\partial F_2}{\partial y} = -\frac{f_0}{\sqrt{gH}}F_2, \tag{2.86}$$

with solutions

$$F_1 = F(x + ct)\mathrm{e}^{y/L_d} \qquad F_2 = G(x - ct)\mathrm{e}^{-y/L_d}, \tag{2.87}$$

where $L_d = \sqrt{gH}/f_0$ is the radius of deformation. If we consider flow in the half-plane in which $y > 0$, then for positive f_0 the solution F_1 grows exponentially away from the wall, and so fails to satisfy the condition of boundedness at infinity. It thus must be eliminated, leaving the general solution

$$u' = e^{-y/L_d} G(x - ct), \qquad v' = 0, \qquad \eta' = \sqrt{H/g}\, e^{-y/L_d} G(x - ct). \tag{2.88a,b,c}$$

These are Kelvin waves, and they decay exponentially away from the boundary. In general, for f_0 positive the boundary is to the right of an observer moving with the wave. Given a constant Coriolis parameter, we could equally well have obtained a solution on a meridional wall, in which case we would find that the wave again moves such that the wall is to the right of the wave direction. (This is obvious once it is realized that f-plane dynamics are isotropic in x and y.) Thus, in the Northern Hemisphere the wave moves anticlockwise round a basin, and conversely in the Southern Hemisphere, and in both hemispheres the direction is cyclonic.

2.6 Geostrophic Adjustment

Large-scale, extratropical circulation of the atmosphere is in near-geostrophic balance. Why is this? Why should the Rossby number be small? It turs out there is in fact a powerful and ubiquitous process whereby a fluid in an initially unbalanced state naturally evolves toward a state of geostrophic balance, namely *geostrophic adjustment*. This process occurs quite generally in rotating fluids, whether stratified or not. We consider the free evolution of a single shallow layer of fluid whose initial state is manifestly unbalanced, and we will suppose that surface displacements are small so that the evolution of the system is described by the linearized shallow equations of motion. These are

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \qquad \frac{\partial \eta}{\partial t} + H\nabla\cdot \boldsymbol{u} = 0, \tag{2.89a,b}$$

where η is the free surface displacement and H is the mean fluid depth, and we omit the primes on the linearized variables.

2.6.1 Non-rotating flow

We consider first the non-rotating problem set, with little loss of generality, in one dimension. We suppose that initially the fluid is at rest but with a simple discontinuity in the height field so that

$$\eta(x, t=0) = \begin{cases} +\eta_0 & x < 0 \\ -\eta_0 & x > 0 \end{cases} \tag{2.90}$$

and $u(x, t = 0) = 0$ everywhere. We can realize these initial conditions physically by separating two fluid masses of different depths by a thin dividing wall, and then quickly removing the wall. What is the subsequent evolution of the fluid? The general solution to the linear problem is given by (2.68) where the functional form is determined by the initial conditions so that here

$$F(x) = \eta(x, t=0) = -\eta_0 \operatorname{sgn}(x). \tag{2.91}$$

Equation (2.68) states that this initial pattern is propagated to the right and to the left. That is, two discontinuities in fluid height move to the right and left at a speed $c = \sqrt{gH}$. Specifically, the solution is

$$\eta(x, t) = -\frac{1}{2}\eta_0[\operatorname{sgn}(x + ct) + \operatorname{sgn}(x - ct)]. \tag{2.92}$$

The initial conditions may be much more complex than a simple front, but, because the waves are dispersionless, the solution is still simply a sum of the translation of those initial conditions to the right and to the left at speed c. The velocity field in this class of problem is obtained from

$$\frac{\partial u}{\partial t} = -g\frac{\partial \eta}{\partial x}, \tag{2.93}$$

which gives, using (2.68),

$$u = -\frac{g}{2c}[F(x + ct) - F(x - ct)]. \tag{2.94}$$

Consider the case with initial conditions given by (2.90). At a given location, away from the initial disturbance, the fluid remains at rest and undisturbed until the front arrives.

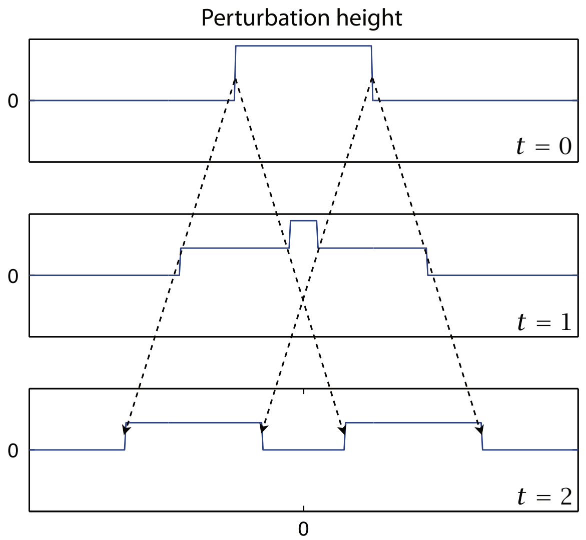

Figure 2.7 The time development of an initial 'top hat' height disturbance, with zero initial velocity, in non-rotating flow. Fronts propagate in both directions, and the velocity is non-zero between fronts, but ultimately the disturbance are radiated away to infinity, and the fluid is left at rest with zero perturbation height.

After the front has passed, the fluid surface is again undisturbed and the velocity is uniform and non zero. Specifically:

$$\eta = \begin{cases} -\eta_0 \mathrm{sgn}(x) \\ 0 \end{cases} \qquad u = \begin{cases} 0 & |x| > ct \\ (\eta_0 g/c) & |x| < ct. \end{cases} \tag{2.95}$$

The solution with 'top-hat' initial conditions in the height field, and zero initial velocity, is a superposition two discontinuities similar to (2.95) and is illustrated in Fig. 2.7. Two fronts propagate in either direction from each discontinuity and, in this case, the final velocity, as well as the fluid displacement, is zero after all the fronts have passed. That is, the disturbance is radiated completely away.

2.6.2 Rotating flow

Rotation makes a profound difference to the adjustment problem of the shallow water system, because a steady, adjusted, solution can exist with non-zero gradients in the height field — the associated pressure gradients being balanced by the Coriolis force — and potential vorticity conservation provides a powerful constraint on the fluid evolution.[4] In a rotating shallow fluid that conservation is represented by

$$\frac{\partial Q}{\partial t} + \boldsymbol{u} \cdot \nabla Q = 0, \tag{2.96}$$

where $Q = (\zeta + f)/h$. In the linear case with constant Coriolis parameter, (2.96) becomes

$$\frac{\partial q}{\partial t} = 0, \qquad q = \left(\zeta - f_0 \frac{\eta}{H}\right). \tag{2.97}$$

This equation may be obtained either from the linearized velocity and mass conservation equations, (2.89), or from (2.96) directly. In the latter case, we write

$$Q = \frac{\zeta + f_0}{H + \eta} \approx \frac{1}{H}(\zeta + f_0)\left(1 - \frac{\eta}{H}\right) \approx \frac{1}{H}\left(f_0 + \zeta - f_0\frac{\eta}{H}\right) = \frac{f_0}{H} + \frac{q}{H} \tag{2.98}$$

having used $f_0 \gg |\zeta|$ and $H \gg |\eta|$. The term f_0/H is a constant and so dynamically unimportant, as is the H^{-1} factor multiplying q. Further, the advective term $\boldsymbol{u} \cdot \nabla Q$ becomes $\boldsymbol{u} \cdot \nabla q$, and this is second order in perturbed quantities and so is neglected. Thus, making these approximations, (2.96) reduces to (2.97). The potential vorticity field is therefore fixed in space! Of course, this was also true in the non-rotating case where the fluid is initially at rest. Then $q = \zeta = 0$ and the fluid remains irrotational throughout the subsequent evolution of the flow. However, this is rather a weak constraint on the subsequent evolution of the fluid; it does nothing, for example, to prevent the conversion of all the potential energy to kinetic energy. In the rotating case the potential vorticity is non-zero, and potential vorticity conservation and geostrophic balance are all we need to infer the final steady state, assuming it exists, without solving for the details of the flow evolution, as we now see.

With an initial condition for the height field given by (2.90), the initial potential vorticity is given by

$$q(x,y) = \begin{cases} -f_0\eta_0/H & x < 0 \\ f_0\eta_0/H & x > 0, \end{cases} \tag{2.99}$$

and this remains unchanged throughout the adjustment process. The final steady state is then the solution of the equations

$$\zeta - f_0\frac{\eta}{H} = q(x,y), \qquad f_0 u = -g\frac{\partial \eta}{\partial y}, \qquad f_0 v = g\frac{\partial \eta}{\partial x}, \tag{2.100a,b,c}$$

where $\zeta = \partial v/\partial x - \partial u/\partial y$. Because the Coriolis parameter is constant, the velocity field is horizontally non-divergent and we may define a streamfunction $\psi = g\eta/f_0$. Equations (2.100) then reduce to

$$\left(\nabla^2 - \frac{1}{L_d^2}\right)\psi = q(x,y), \tag{2.101}$$

where $L_d = \sqrt{gH}/f_0$ is known as the *Rossby radius of deformation* or often just the 'deformation radius' or the 'Rossby radius'. It is a naturally occurring length scale in problems involving both rotation and gravity, and arises in a slightly different form in stratified fluids.

The initial conditions (2.99) admit of a nice analytic solution, for the flow will remain uniform in y, and (2.101) reduces to

$$\frac{\partial^2\psi}{\partial x^2} - \frac{1}{L_d^2}\psi = \frac{f_0\eta_0}{H}\operatorname{sgn}(x). \tag{2.102}$$

We solve this separately for $x > 0$ and $x < 0$ and then match the solutions and their first derivatives at $x = 0$, also imposing the condition that the velocity decays to zero as $x \to \pm\infty$. The solution is

$$\psi = \begin{cases} -(g\eta_0/f_0)(1 - e^{-x/L_d}) & x > 0 \\ +(g\eta_0/f_0)(1 - e^{x/L_d}) & x < 0. \end{cases} \tag{2.103}$$

The velocity field associated with this is obtained from (2.100b,c), and is

$$u = 0, \qquad v = -\frac{g\eta_0}{f_0 L_d} e^{-|x|/L_d}. \tag{2.104}$$

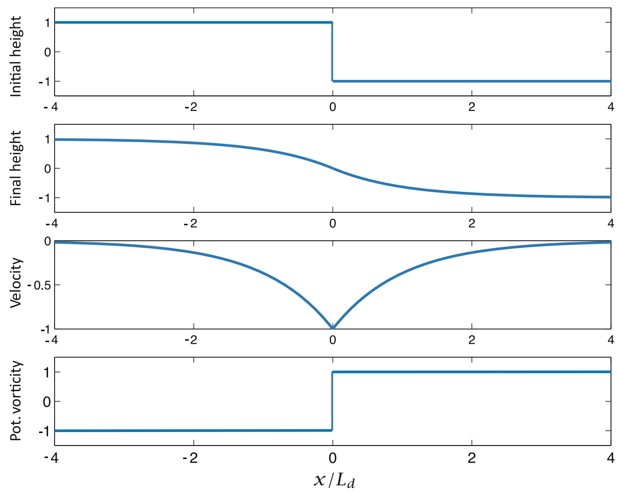

Figure 2.8 Solutions of a linear geostrophic adjustment problem. Top panel: the initial height field, given by (2.90) with $\eta_0 = 1$. Second panel: equilibrium (final) height field, η given by (2.103) and $\eta = f_0\psi/g$. Third panel: equilibrium geostrophic velocity (normal to the gradient of height field), given by (2.104). Bottom panel: potential vorticity, given by (2.99), and this does note evolve. The distance, x is non-dimensionalized by the deformation radius $L_d = \sqrt{gH}/f_0$, and the velocity by $\eta_0(g/f_0 L_d)$. Changes to the initial state occur only within $\mathcal{O}(L_d)$ of the initial discontinuity; and as $x \to \pm\infty$ the initial state is unaltered.

The velocity is perpendicular to the slope of the free surface, and a jet forms along the initial discontinuity, as illustrated in Fig. 2.8.

The important point of this problem is that the variations in the height and field are not radiated away to infinity, as in the non-rotating problem. Rather, potential vorticity conservation constrains the influence of the adjustment to within a deformation radius (we see now why this name is appropriate) of the initial disturbance. This property is a general one in geostrophic adjustment — it also arises if the initial condition consists of a velocity jump, as considered in problem 2.**??**.

2.6.3 ♦ Energetics of adjustment

How much of the initial potential energy of the flow is lost to infinity by gravity wave radiation, and how much is converted to kinetic energy? The linear equations (2.89) lead to

$$\frac{1}{2}\frac{\partial}{\partial t}(H\boldsymbol{u}^2 + g\eta^2) + gH\nabla\cdot(\boldsymbol{u}\eta) = 0, \qquad (2.105)$$

so that energy conservation holds in the form

$$E = \frac{1}{2}\int (H\boldsymbol{u}^2 + g\eta^2)\,\mathrm{d}\boldsymbol{x}, \qquad \frac{\mathrm{d}E}{\mathrm{d}t} = 0, \qquad (2.106)$$

provided the integral of the divergence term vanishes, as it normally will in a closed domain. The fluid has a non-zero potential energy, $(1/2)\int_{-\infty}^{\infty} g\eta^2\,\mathrm{d}x$, if there are variations in fluid height, and with the initial conditions (2.90) the initial potential energy is

$$PE_I = \int_0^{\infty} g\eta_0^2\,\mathrm{d}x. \qquad (2.107)$$

This is nominally infinite if the fluid has no boundaries, and the initial potential energy density is $g\eta_0^2/2$ everywhere.

In the non-rotating case, and with initial conditions (2.90), after the front has passed, the potential energy density is zero and the kinetic energy density is $Hu^2/2 = g\eta_0^2/2$, using (2.95) and $c^2 = gH$. Thus, all the potential energy is locally converted to kinetic energy as the front passes, and eventually the kinetic energy is distributed uniformly along the line. In the case illustrated in Fig. 2.7, the potential energy and kinetic energy are both radiated away from the initial disturbance. (Note that although we can superpose the solutions from different initial conditions, we cannot superpose their potential and kinetic energies.) The general point is that the evolution of the disturbance is not confined to its initial location.

In contrast, in the rotating case the conversion from potential to kinetic energy *is largely confined to within a deformation radius of the initial disturbance,* and at locations far from the initial disturbance the initial state is essentially unaltered. The conservation of potential vorticity has prevented the complete conversion of potential energy to kinetic energy, a result that is not sensitive to the precise form of the initial conditions (see also problem 2.??).

In fact, in the rotating case, some of the initial potential energy is converted to kinetic energy, some remains as potential energy and some is lost to infinity; let us calculate these amounts. The final potential energy, after adjustment, is, using (2.103),

$$PE_F = \frac{1}{2} g\eta_0^2 \left[\int_0^\infty \left(1 - \mathrm{e}^{-x/L_d}\right)^2 \, \mathrm{d}x + \int_{-\infty}^0 \left(1 - \mathrm{e}^{x/L_d}\right)^2 \, \mathrm{d}x \right]. \tag{2.108}$$

This is nominally infinite, but the change in potential energy is finite and is given by

$$PE_I - PE_F = g\eta_0^2 \int_0^\infty (2\mathrm{e}^{-x/L_d} - \mathrm{e}^{-2x/L_d}) \, \mathrm{d}x = \frac{3}{2} g\eta_0^2 L_d. \tag{2.109}$$

The initial kinetic energy is zero, because the fluid is at rest, and its final value is, using (2.104),

$$KE_F = \frac{1}{2} H \int \boldsymbol{u}^2 \, \mathrm{d}x = H \left(\frac{g\eta_0}{fL_d} \right)^2 \int_0^\infty \mathrm{e}^{-2x/L_d} \, \mathrm{d}x = \frac{g\eta_0^2 L_d}{2}. \tag{2.110}$$

Thus one-third of the difference between the initial and final potential energies is converted to kinetic energy, and this is trapped within a distance of the order of a deformation radius of the disturbance; the remainder, an amount $gL_d\eta_0^2$ is radiated away and lost to infinity. In any finite region surrounding the initial discontinuity the final energy is less than the initial energy.

2.6.4 A variational perspective

In the non-rotating problem, all of the initial potential energy is eventually radiated away to infinity. In the rotating problem, the final state contains both potential and kinetic energy. Why is the energy not all radiated away to infinity? It is because potential vorticity conservation on parcels prevents all of the energy being dispersed. This suggests that it may be informative to think of the geostrophic adjustment problem as a *variational problem:* we seek to minimize the energy consistent with the conservation of potential vorticity. We stay in the linear approximation in which, because the advection of potential vorticity is neglected, potential vorticity remains constant at each point.

The energy of the flow is given by the sum of potential and kinetic energies, namely

$$\text{energy} = \int (H\boldsymbol{u}^2 + g\eta^2)\, \mathrm{d}A, \tag{2.111}$$

(where $\mathrm{d}A \equiv \mathrm{d}x\,\mathrm{d}y$) and the potential vorticity field is

$$q = \zeta - f_0 \frac{\eta}{H} = (v_x - u_y) - f_0 \frac{\eta}{H}, \tag{2.112}$$

where the subscripts x, y denote derivatives. The problem is then to extremize the energy subject to potential vorticity conservation. This is a constrained problem in the calculus of variations, sometimes called an *isoperimetric* problem because of its origins in maximizing the area of a surface for a given perimeter.[5] The mathematical problem is to extremize the integral

$$I = \int \left\{ H(u^2 + v^2) + g\eta^2 + \lambda(x, y)\big[(v_x - u_y) - f_0\eta/H\big] \right\} \mathrm{d}A, \tag{2.113}$$

where $\lambda(x, y)$ is a Lagrange multiplier, undetermined at this stage. It is a function of space: if it were a constant, the integral would merely extremize energy subject to a given integral of potential vorticity, and rearrangements of potential vorticity (which here we wish to disallow) would leave the integral unaltered.

As there are three independent variables there are three Euler–Lagrange equations that must be solved in order to minimize I. These are

$$\frac{\partial L}{\partial \eta} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \eta_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial \eta_y} = 0,$$

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial u_y} = 0, \qquad \frac{\partial L}{\partial v} - \frac{\partial}{\partial x}\frac{\partial L}{\partial v_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial v_y} = 0, \tag{2.114}$$

where L is the integrand on the right-hand side of (2.113). Substituting the expression for L into (2.114) gives, after a little algebra,

$$2g\eta - \frac{\lambda f_0}{H} = 0, \qquad 2Hu + \frac{\partial \lambda}{\partial y} = 0, \qquad 2Hv - \frac{\partial \lambda}{\partial x} = 0, \tag{2.115}$$

and then eliminating λ gives the simple relationships

$$u = -\frac{g}{f_0}\frac{\partial \eta}{\partial y}, \qquad v = \frac{g}{f_0}\frac{\partial \eta}{\partial x}, \tag{2.116}$$

which are the equations of geostrophic balance. Thus, in the linear approximation, *geostrophic balance is the minimum energy state for a given field of potential vorticity.*

Chapter 3

Geostrophic Theory

Weeks 5 to 7

This chapter is concerned with flows that are close to geostrophic balance, with the specific goal of deriving equation sets that exploit this closeness and that are simpler than the original, 'primitive' equations. We will in particular derive the quasi-geostrophic and planetary-geostrophic sets of equations. We do this first for shallow water and then for the stratified three-dimensional equations.

3.1 Geostrophic Scaling in the Shallow Water Equations

For simplicity we will assume a flat bottom, so that $\eta = h$. With the odd exception, we will denote the scales of variables by capital letters; thus, if L is a typical length scale of the motion we wish to describe, and U is a typical velocity scale, and assuming the scales are horizontally isotropic, we write

$$\begin{aligned} (x, y) &\sim L \qquad \text{or} \qquad (x, y) = \mathcal{O}(L) \\ (u, v) &\sim U \qquad \text{or} \qquad (u, v) = \mathcal{O}(U). \end{aligned} \tag{3.1}$$

and similarly for other variables. We may then non-dimensionalize the variables by writing

$$(x, y) = L(\widehat{x}, \widehat{y}), \qquad (u, v) = U(\widehat{u}, \widehat{v}), \tag{3.2}$$

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

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} = -g \nabla \eta, \tag{3.3a}$$

$$\frac{U}{T} \qquad \frac{U^2}{L} \qquad fU \qquad g\frac{\mathcal{H}}{L}, \tag{3.3b}$$

where the ∇ operator acts in the x, y plane and $\mathcal{H}$ is the amplitude of the variations in the surface displacement. (We use η to denote the height of the free surface above some arbitrary reference level, as in Fig. 2.1. Thus, $\eta = H + \Delta\eta$, where $\Delta\eta$ denotes the variation of η about its mean position.)

The ratio of the advective term to the rotational term in the momentum equation (3.3) is $(U^2/L)/(fU) = U/fL$; this is the Rossby number, first encountered in chapter **??**.[6] Using values typical of the large-scale circulation (e.g., from Table 1.1) we find that $Ro \approx 0.1$ for the atmosphere and $Ro \approx 0.01$ for the ocean: small in both cases. If we are interested in motion that has the advective time scale $T = L/U$ then we scale time by L/U so that

$$t = \frac{L}{U}\hat{t}, \tag{3.4}$$

and the local time derivative and the advective term then both scale as U^2/L, and both are smaller than the rotation term by a factor of the order of the Rossby number. Then, either the Coriolis term is the dominant term in the equation, in which case we have a state of no motion with $-fv = 0$, or else the Coriolis force is balanced by the pressure force, and the dominant balance is

$$-fv = -g\frac{\partial \eta}{\partial x}, \tag{3.5}$$

namely *geostrophic balance*. If we make this non-trivial choice, then variations in η (i.e., $\Delta\eta$) scale according to

$$\Delta\eta \sim \mathcal{H} = \frac{fUL}{g}. \tag{3.6}$$

We can also write $\mathcal{H}$ as

$$\mathcal{H} = Ro\frac{f^2L^2}{g} = Ro\, H\frac{L^2}{L_d^2}, \tag{3.7}$$

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

$$\frac{\Delta\eta}{H} \sim Ro\frac{L^2}{L_d^2}, \tag{3.8}$$

and the height of the fluid may be written as

$$\eta = H\left(1 + Ro\frac{L^2}{L_d^2}\widehat{\eta}\right) \qquad \text{and} \qquad \Delta\eta = Ro\frac{L^2}{L_d^2}H\widehat{\eta}, \tag{3.9}$$

where $\widehat{\eta}$ is the $\mathcal{O}(1)$ non-dimensional value of the surface height deviation.

Non-dimensional momentum equation

If we use (3.9) to scale height variations, (3.2) to scale lengths and velocities, and (3.4) to scale time, then the momentum equation (3.3) becomes

$$Ro\left[\frac{\partial\widehat{\boldsymbol{u}}}{\partial\widehat{t}} + (\widehat{\boldsymbol{u}}\cdot\nabla)\widehat{\boldsymbol{u}}\right] + \widehat{\boldsymbol{f}}\times\widehat{\boldsymbol{u}} = -\nabla\widehat{\eta}\ , \tag{3.10}$$

where $\widehat{\boldsymbol{f}} = \mathbf{k}\widehat{f} = \mathbf{k}f/f_0$, where f_0 is a representative value of the Coriolis parameter. (If f is a constant, then $\widehat{f} = 1$, but it is informative to explicitly write $\widehat{f}$ in the equations. Also, where the operator ∇ operates on a non-dimensional variable then the differentials are taken with respect to the non-dimensional variables $\widehat{x}$, $\widehat{y}$.) All the variables in (3.10) will be assumed to be of order unity, and the Rossby number multiplying the local time derivative and the advective terms indicates the smallness of those terms. By construction, the dominant balance in this equation is the geostrophic balance between the last two terms.

Non-dimensional mass continuity (height) equation

The (dimensional) mass continuity equation can be written as

$$\frac{1}{H}\frac{\mathrm{D}\eta}{\mathrm{D}t} + \left(1 + \frac{\Delta\eta}{H}\right)\nabla\cdot\boldsymbol{u} = 0, \tag{3.11}$$

Using (3.2), (3.4) and (3.9) this equation may be written

$$Ro\left(\frac{L}{L_d}\right)^2\frac{\mathrm{D}\widehat{\eta}}{\mathrm{D}\widehat{t}} + \left[1 + Ro\left(\frac{L}{L_d}\right)^2\widehat{\eta}\right]\nabla\cdot\widehat{\boldsymbol{u}} = 0\ . \tag{3.12}$$

Equations (3.10) and (3.12) are the non-dimensional versions of the full shallow water equations of motion. Evidently, some terms in the equations of motion are small and may be eliminated with little loss of accuracy, and the way this is done will depend on the size of the second non-dimensional parameter, $(L/L_d)^2$. We explore this in sections 3.2 and 3.3.

Froude and Burger numbers

The Froude number may be generally defined as the ratio of a fluid particle speed to a wave speed. In a shallow water system this gives

$$Fr \equiv \frac{U}{\sqrt{gH}} = \frac{U}{f_0 L_d} = Ro\frac{L}{L_d}. \tag{3.13}$$

The Burger number is a useful measure of the scale of motion of the fluid, relative to the deformation radius, and may be defined by

$$Bu \equiv \left(\frac{L_d}{L}\right)^2 = \frac{gH}{f_0^2 L^2} = \left(\frac{Ro}{Fr}\right)^2. \tag{3.14}$$

It is also useful to define the parameter $F \equiv Bu^{-1}$, which is like the square of a Froude number but uses the rotational speed fL instead of U in the numerator.

3.2 The Shallow Water Planetary-Geostrophic Equations

3.2.1 Informal derivation

The advection and time derivative terms in the momentum equation (3.10) are order Rossby number smaller than the Coriolis and pressure terms (the term in square brackets is multiplied by Ro), and therefore let us neglect them. The momentum equation straightforwardly becomes

$$\widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}} = -\nabla\widehat{\eta}. \tag{3.15}$$

The mass conservation equation (3.12), contains two non-dimensional parameters, $Ro = U/(f_0 L)$ (the Rossby number), and $F = (L/L_d)^2$ (the ratio of the length scale of the motion to the deformation scale; $F = Bu^{-1}$) and we must make a choice as to the relationship between these two numbers. We will choose

$$FRo = \mathcal{O}(1), \tag{3.16}$$

which implies

$$L^2 \gg L_d^2 \qquad \text{or equivalently} \qquad F \gg 1, \quad Bu \ll 1. \tag{3.17}$$

That is to say, we suppose that the scales of motion are much larger than the deformation scale. Given this choice, all the terms in the mass conservation equation, (3.12), are of roughly the same size, and we retain them all. Thus, the shallow water planetary geostrophic equations are the full mass continuity equation along with geostrophic balance and a geometric relationship between the height field and the fluid thickness, and in dimensional form these are:

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla\cdot\boldsymbol{u} = 0$$
$$\boldsymbol{f}\times\boldsymbol{u} = -g\nabla\eta, \qquad \eta = h + \eta_b. \tag{3.18a,b}$$

We emphasize that *the planetary-geostrophic equations are only valid for scales of motion much larger than the deformation radius.* The height variations are then as large as the mean height field itself; that is, using (3.8), $\Delta\eta/H = \mathcal{O}(1)$.

♦ *Formal derivation*

We make the following assumptions.

(i) The Rossby number is small. $Ro = U/f_0L \ll 1$.

(ii) The scale of the motion is significantly larger than the deformation scale. That is, (3.16) holds or

$$F = Bu^{-1} = \left(\frac{L}{L_d}\right)^2 \gg 1 \tag{3.19}$$

and in particular

$$FRo = \mathcal{O}(1). \tag{3.20}$$

(iii) Time scales advectively, so that $T = L/U$.

We expand the non-dimensional variables velocity and height fields in an asymptotic series with the Rossby number as the small parameter, substitute into the equations of motion and derive a simpler set of equations. It is a nearly trivial exercise in this instance, and so it illustrates the methodology well. The expansions are

$$\widehat{\boldsymbol{u}} = \widehat{\boldsymbol{u}}_0 + Ro\,\widehat{\boldsymbol{u}}_1 + Ro^2\widehat{\boldsymbol{u}}_2 + \cdots \tag{3.21a}$$

and

$$\widehat{\eta} = \widehat{\eta}_0 + Ro\,\widehat{\eta}_1 + Ro^2\widehat{\eta}_2 + \cdots . \tag{3.21b}$$

Then substituting (3.21a) and (3.21b) into the momentum equation gives

$$Ro\left[\frac{\partial \widehat{\boldsymbol{u}}_0}{\partial \widehat{t}} + \widehat{\boldsymbol{u}}_0 \cdot \nabla \widehat{\boldsymbol{u}}_0 + \widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}}_1\right] + \widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}}_0 = -\nabla\widehat{\eta}_0 - Ro\left[\nabla\widehat{\eta}_1\right] + \mathcal{O}(Ro^2) \tag{3.22}$$

The Rossby number is an asymptotic ordering parameter; thus, the sum of all the terms at any particular order in Rossby number must vanish. At lowest order we obtain the simple expression

$$\widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}}_0 = -\nabla\widehat{\eta}_0. \tag{3.23}$$

Note that although f_0 is a representative value of f, we have made no assumptions about the constancy of f. In particular, f is allowed to vary by an order one amount, provided that it does not become so small that the Rossby number $U/(f_0 L)$ is not small.

The appropriate height (mass conservation) equation is similarly obtained by substituting (3.21a) and (3.21b) into the shallow water mass conservation equation. Because $FRo = \mathcal{O}(1)$ at lowest order we simply retain all the terms in the equation to give

$$FRo\left[\frac{\partial \widehat{\eta}_0}{\partial t} + \widehat{\boldsymbol{u}}_0 \cdot \nabla\widehat{\eta}_0\right] + \left[1 + FRo\,\widehat{\eta}\right]\nabla\cdot\widehat{\boldsymbol{u}}_0 = 0. \tag{3.24}$$

Equations (3.23) and (3.24) are a closed set, and constitute the non-dimensional planetary-geostrophic equations. The dimensional forms of these equations are just (3.18).

Variation of the Coriolis parameter

Suppose then that f is a constant (f_0). Then, from the curl of (3.23), $\nabla \cdot \boldsymbol{u}_0 = 0$. This means that we can define a streamfunction for the flow and, from geostrophic balance, the height field is just that streamfunction. That is, in dimensional form,

$$\psi = \frac{g}{f_0}\eta, \qquad \boldsymbol{u} = \mathbf{k} \times \nabla\psi, \tag{3.25a,b}$$

and (3.24) becomes, in dimensional form,

$$\frac{\partial \eta}{\partial t} + \boldsymbol{u} \cdot \nabla\eta = 0 \qquad \text{or} \qquad \frac{\partial \eta}{\partial t} + J(\psi, \eta) = 0, \tag{3.26}$$

where $J(a,b) \equiv a_x b_y - a_y b_x$. But since $\eta \propto \psi$ the advective term is proportional to $J(\psi, \psi)$, which is zero. Thus, the flow does not evolve at this order. The planetary-geostrophic equations are *uninteresting* if the scale of the motion is such that the Coriolis parameter is not variable. On Earth, the scale of motion on which this parameter regime exists is rather limited, since the planetary-geostrophic equations require that the scale of motion also be larger than the deformation radius. In the Earth's atmosphere, any scale that is larger than the deformation radius will be such that the Coriolis parameter varies significantly over it, and we do not encounter this parameter regime. On the other hand, in the Earth's ocean the deformation radius is relatively small and there exists a small parameter regime that has scales larger than the deformation radius but smaller than that on which the Coriolis parameter varies.

Potential vorticity

The shallow water PG equations may be written as an evolution equation for an approximated potential vorticity. A little manipulation reveals that (3.18) are equivalent to:

$$\begin{gathered} \frac{\mathrm{D}Q}{\mathrm{D}t} = 0 \\ Q = \frac{f}{h}, \quad \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \quad \eta = h + \eta_b. \end{gathered} \tag{3.27}$$

Thus, potential vorticity is a material invariant in the approximate equation set, just as it is in the full equations. The other variables — the free surface height and the velocity — are diagnosed from it, a process known as *potential vorticity inversion*. In the planetary geostrophic approximation, the inversion proceeds using the approximate form f/h rather than the full potential vorticity, $(f+\zeta)/h$. (Strictly speaking, we do not approximate potential vorticity, because this is the evolving variable. Rather, we approximate the inversion relations from which we derive the height and velocity fields.) The simplest way of all to derive the shallow water PG equations is to *begin* with the conservation of potential vorticity, and to note that at small Rossby number the expression $(\zeta + f)/h$ may be approximated by f/h. Then, noting in addition that the flow is geostrophic, (3.27) immediately emerges. *Every* approximate set of equations that we derive in this chapter may be expressed as the evolution of potential vorticity, with the other fields being obtained diagnostically from it.

3.3 The Shallow Water Quasi-Geostrophic Equations

We now derive a set of geostrophic equations that is valid (unlike the PG equations) when the horizontal scale of motion is similar to that of the deformation radius. These equations are called the *quasi-geostrophic* equations, and are perhaps the most widely used set of equations for theoretical studies of the atmosphere and ocean. The specific assumptions we make are as follows.

(i) The Rossby number is small, so that the flow is in near-geostrophic balance.

(ii) The scale of the motion is not significantly larger than the deformation scale. Specifically, we shall require that

$$Ro\left(\frac{L}{L_d}\right)^2 = \mathcal{O}(Ro). \tag{3.28}$$

For the shallow water equations, this assumption implies, using (3.9), that the variations in fluid depth are small compared to its total depth. For the continuously stratified system it implies, using (3.53), that the variations in stratification are small compared to the background stratification.

(iii) Variations in the Coriolis parameter are small; that is, $|\beta L| \ll |f_0|$ where L is the length scale of the motion.

(iv) Time scales advectively; that is, the scaling for time is given by $T = L/U$.

The second and third of these differ from the planetary-geostrophic counterparts: we make the second assumption because we wish to explore a different parameter regime, and we then find that the third assumption is necessary to avoid a rather trivial state [i.e., a leading order balance of $\beta v = 0$, see the discussion surrounding (3.44)]. All of the assumptions are the same whether we consider the shallow water equations or a continuously stratified flow, and in this section we consider the former.

3.3.1 Shallow water quasi-geostrophic equations

let us set the velocity equal to a geostrophic component, $\boldsymbol{u}_g$ plus an ageostrophic component, $\boldsymbol{u}_a$. We will suppose that $f = f_0 + \beta y$, where $|f_0| \gg |\beta y|$, and we will define the geostrophic flow to be the flow that satisfies

$$f_0 \times \boldsymbol{u}_g = -g\nabla\eta, \tag{3.29}$$

which in turn implies $\nabla\cdot \boldsymbol{u}_g = 0$. Rather than make approximations to the momentum approximation let us begin with the shallow water vorticity equation which, reprising 2.41, is

$$\frac{\partial \zeta}{\partial t} + (\boldsymbol{u}\cdot\nabla)(\zeta + f) = -(\zeta + f)\nabla\cdot \boldsymbol{u}. \tag{3.30}$$

The right-hand side contains only the ageostrophic velocity, which is small, and since ζ is smaller than f by a factor of the Rossby number we can ignore $\zeta\nabla\cdot \boldsymbol{u}$ and take f to be equal to f_0. The left-hand side may be well-approximated by using the geostrophic flow, (3.29), so that we have

$$\frac{\partial \zeta_g}{\partial t} + (\boldsymbol{u}_g\cdot\nabla)(\zeta_g + f) = -f_0\nabla\cdot \boldsymbol{u}_a. \tag{3.31}$$

Note that on the left-hand side f can be replaced by βy.

We now use the mass continuity equation to obtain an expression for the divergence. From (**??**) the mass continuity equation may be written as

$$\frac{\mathrm{D}\eta}{\mathrm{D}t} + (H + \Delta\zeta)\nabla\cdot \boldsymbol{u} = 0, \tag{3.32}$$

and since $H \gg \Delta\zeta$ (using (3.12), H is bigger by a factor $(L_d/L)^2 Ro^{-1}$), the equation becomes

$$\frac{\mathrm{D}\eta}{\mathrm{D}t} + H\nabla\cdot \boldsymbol{u}_a = 0. \tag{3.33}$$

Combining (3.31) and (3.33) gives

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\zeta_g + f - \frac{f_0\eta}{H}\right) = 0. \tag{3.34}$$

It appears that we have two variables here, ζ_g and η. However, they are related through geostrophic balance, and the fact that the geostrophic flow is non-divergent. Thus, we may define a streamfunction ψ such that $u_g = -\partial\psi/\partial y, v_g = \partial\psi/\partial x$, whence $\partial u/\partial x + \partial v/\partial y = 0$, The vorticity and height field are related to the streamfunction by

$$\zeta_g = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla^2\psi, \qquad \text{and} \qquad \eta = \frac{f_0\psi}{g}, \tag{3.35a,b}$$

where the second relation comes from geostrophic balance. Eqrefgs:pg10 may then be written as

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \qquad q = \nabla^2\psi + \beta y - \frac{\psi}{L_d^2}, \tag{3.36}$$

where $L_d = \sqrt{gH}/f_0$. The variable q is the *quasi-geostrophic potential vorticity.*

Connection to shallow water potential vorticity

The quantity q given by (3.36) is an approximation (except for dynamically unimportant constant additive and multiplicative factors) to the shallow water potential vorticity. To see the truth of this statement, begin with the expression for the shallow water potential vorticity,

$$Q = \frac{f + \zeta}{h}. \tag{3.37}$$

Now let $h = H(1 + \eta'/H)$, where η' is the perturbation of the free-surface height, and assume that η'/H is small to obtain

$$
\begin{aligned}
Q = \frac{f+\zeta}{H(1+\eta'/H)} &\approx \frac{1}{H}(f+\zeta)\left(1 - \frac{\eta'}{H}\right) \\
&\approx \frac{1}{H}\left(f_0 + \beta y + \zeta - f_0\frac{\eta'}{H}\right).
\end{aligned}
\tag{3.38}
$$

Because f_0/H is a constant it has no effect in the evolution equation, and the quantity given by

$$
q = \beta y + \zeta - f_0 \frac{\eta'}{H} \tag{3.39}
$$

is materially conserved. Using geostrophic balance we have $\zeta = \nabla^2\psi$ and $\eta' = f_0\psi/g$ so that (3.39) is identical to the q given in (3.36).

The approximations needed to go from (3.37) to (3.39) are the same as those used in our earlier, more long-winded, derivation of the quasi-geostrophic equations. That is, we assumed that f itself is nearly constant, and that f_0 is much larger than ζ, equivalent to a low Rossby number assumption. It was also necessary to assume that $H \gg \eta'$ to enable the expansion of the height field which, using assumption (*(ii)*) on page 69, is equivalent to requiring that the scale of motion not be significantly larger than the deformation scale. The derivation is completed by noting that the advection of the potential vorticity should be by the geostrophic velocity alone, and we recover (**??**) or (**??**).

Two interesting limits

There are two interesting limits to the quasi-geostrophic potential vorticity equation which, taking $\beta = 0$ for simplicity, are as follows.

(i) Motion on scales much smaller than the deformation radius. That is, $L \ll L_d$ and thus $Bu \gg 1$ or $F \ll 1$. Then (**??**) becomes

$$
\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) = 0, \tag{3.40}
$$

where $\zeta = \nabla^2\psi$ and $J(\psi, \zeta) = \psi_x\zeta_y - \psi_y\zeta_x$. Thus, the motion obeys the two-dimensional vorticity equation. Physically, on small length scales the deviations in the height field are very small and may be neglected.

(ii) Motion on scales much larger than the deformation radius. Although scales are not allowed to become so large that $Ro(L/L_d)^2$ is of order unity, we may, a posteriori, still have $L \gg L_d$, whence the potential vorticity equation, (**??**), becomes

$$\frac{\partial \psi}{\partial t} + J(\psi, \psi) = 0 \qquad \text{or} \qquad \frac{\partial \eta}{\partial t} + J(\psi, \eta) = 0, \tag{3.41}$$

because $\psi = g\eta/f_0$. The Jacobian term evidently vanishes. Thus, one is left with a trivial equation that implies there is no advective evolution of the height field. There is nothing wrong with our reasoning; the mathematics has indeed pointed out a limit interesting in its uninterestingness. From a physical point of view, however, such a lack of motion is likely to be rare, because on such large scales the Coriolis parameter varies considerably, and we are led to the planetary-geostrophic equations.

In practice, often the most severe restriction of quasi-geostrophy is that variations in layer thickness are small: what does this have to do with geostrophy? If we scale η assuming geostrophic balance then $\eta \sim fUL/g$ and $\eta/H \sim Ro(L/L_d)^2$. Thus, if Ro is to remain small, η/H can only be of order one if $(L/L_d)^2 \gg 1$. That is, the height variations must occur on a large scale, or we are led to a scaling inconsistency. Put another way, *if there are order-one height variations over a length scale of less than or of the order of the deformation scale, the Rossby number will not be small.* Large height variations are allowed if the scale of motion is large, but this contingency is described by the planetary-geostrophic equations.

Another flow regime

Although perhaps of little terrestrial interest, we can imagine a regime in which the Coriolis parameter varies fully, but the scale of motion remains no larger than the deformation radius. This parameter regime is not quasi-geostrophic, but it gives an interesting result. Because $\eta'/H \sim Ro(L/L_d)^2$ deviations of the height field are at least of order Rossby number smaller than the reference height and $|\eta'| \ll H$. The dominant balance in the height equation is then

$$H\nabla\cdot \boldsymbol{u} = 0, \tag{3.42}$$

presuming that time still scales advectively. This zero horizontal divergence must remain consistent with geostrophic balance

$$\boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \tag{3.43}$$

where now f is a fully variable Coriolis parameter. Taking the curl of (that is, cross-differentiating) (3.43) gives

$$\beta v + f\nabla\cdot \boldsymbol{u} = 0, \tag{3.44}$$

whence, using (3.42), $v = 0$, and the flow is purely zonal. Although not at all useful as an evolution equation, this illustrates the constraining effect that differential rotation has on meridional velocity. This effect may be the cause of the banded, highly zonal flow on some of the giant planets.

3.4 Geostrophic scaling in the stratified equations

We use the hydrostatic Boussinesq equations, which we write as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla_z\phi, \tag{3.45a}$$

$$\frac{\partial\phi}{\partial z} = b, \tag{3.45b}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} = 0, \tag{3.45c}$$

$$\nabla\cdot \boldsymbol{v} = 0. \tag{3.45d}$$

where b is the buoyancy. Anticipating that the average stratification may not scale in the same way as the deviation from it, let us separate out the contribution of the advection of a reference stratification in (3.45c) by writing

$$b = \widetilde{b}(z) + b'(x, y, z, t). \tag{3.46}$$

Then the thermodynamic equation becomes

$$\frac{\mathrm{D}b'}{\mathrm{D}t} + N^2 w = 0, \tag{3.47}$$

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

$$\frac{\partial \phi'}{\partial z} = b'. \tag{3.48}$$

Equations (3.47) and (3.48) replace (3.45c) and (3.45b), and ϕ' is used in (3.45a).

3.4.1 Non-dimensional equations

We scale the basic variables by supposing that

$$\begin{gathered}(x, y) \sim L, \quad (u, v) \sim U, \quad t \sim \frac{L}{U}, \quad z \sim H, \\ f \sim f_0, \qquad N \sim N_0,\end{gathered} \tag{3.49}$$

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

$$|\boldsymbol{f} \times \boldsymbol{u}| \sim |\nabla \phi'| \tag{3.50}$$

and so the pressure scales as

$$\phi' \sim \Phi = f_o U L. \tag{3.51}$$

Using the hydrostatic relation, (3.51) implies that the buoyancy scales as

$$b' \sim B = \frac{f_0 U L}{H}, \tag{3.52}$$

and from this we obtain

$$\frac{(\partial b'/\partial z)}{N_0^2} \sim Ro \frac{L^2}{L_d^2}, \tag{3.53}$$

where $L_d = N_0 H/f_0$ is the deformation radius in the continuously stratified fluid, analogous to the quantity $\sqrt{gH}/f_0$ in the shallow water system, and we use the same symbol, L_d, for both. In the continuously stratified system, *if the scale of motion is the same as or smaller than the deformation radius, and the Rossby number is small, then the variations in stratification are small.* The choice of scale is the key difference between the planetary-geostrophic and quasi-geostrophic equations.

Finally, we will non-dimensionalize the vertical velocity by using the mass conservation equation,

$$\frac{\partial w}{\partial z} = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \tag{3.54}$$

and we suppose that this implies

$$w \sim W = \frac{UH}{L}. \tag{3.55}$$

This is a naïve scaling for rotating flow: if the Coriolis parameter is nearly constant the geostrophic velocity is nearly horizontally non-divergent and the right-hand side of (3.54) is small, and $W \ll UH/L$. We might then estimate w by cross-differentiating geostrophic balance to obtain the linear geostrophic vorticity equation and corresponding scaling:

$$\beta v \approx f\frac{\partial w}{\partial z}, \qquad w \sim W = \frac{\beta UH}{f_0}. \tag{3.56a,b}$$

However, rather than using (3.56b) from the outset, we will use (3.55) and let the proper scaling emerge in the fullness of time. Note that if variations in the Coriolis parameter are large and $\beta \sim f_0/L$, then (3.56b) is the same as (3.55).

Given the scalings above [using (3.55) for w] we non-dimensionalize by setting

$$\begin{gathered}(\widehat{x}, \widehat{y}) = L^{-1}(x, y), \qquad \widehat{z} = H^{-1}z, \\ (\widehat{u}, \widehat{v}) = U^{-1}(u, v), \quad \widehat{w} = \frac{L}{UH}w, \quad \widehat{t} = \frac{U}{L}t, \\ \widehat{f} = f_0^{-1}f, \quad \widehat{\phi} = \frac{\phi'}{f_0 UL}, \quad \widehat{b} = \frac{H}{f_0 UL}b',\end{gathered} \tag{3.57}$$

Non-dimensional Primitive Equations

The non-dimebnsional, hydrostatic, Boussinesq equations in a rotating frame of reference are

Horizontal momentum: $$Ro\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} + \hat{\boldsymbol{f}} \times \hat{\boldsymbol{u}} = -\nabla\hat{\phi} \qquad \text{(PE.1)}$$

Hydrostatic: $$\frac{\partial\hat{\phi}}{\partial\hat{z}} = \hat{b} \qquad \text{(PE.2)}$$

Mass continuity: $$\left(\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}} + \frac{\partial\hat{w}}{\partial\hat{z}}\right) = 0 \qquad \text{(PE.3)}$$

Thermodynamic: $$Ro\frac{\mathrm{D}\hat{b}}{\mathrm{D}\hat{t}} + \left(\frac{L_d}{L}\right)^2 \hat{N}^2\hat{w} = 0. \qquad \text{(PE.4)}$$

where the hatted variables are non-dimensional. The horizontal momentum and hydrostatic equations then become

$$Ro\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} + \hat{\boldsymbol{f}} \times \hat{\boldsymbol{u}} = -\nabla\hat{\phi}, \qquad (3.58)$$

and

$$\frac{\partial\hat{\phi}}{\partial\hat{z}} = \hat{b}. \qquad (3.59)$$

The non-dimensional mass conservation equation is simply

$$\nabla\cdot\hat{\boldsymbol{v}} = \left(\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}} + \frac{\partial\hat{w}}{\partial\hat{z}}\right) = 0, \qquad (3.60)$$

and the nondimensional thermodynamic equation is

$$\frac{f_0 UL}{H}\frac{U}{L}\frac{\mathrm{D}\hat{b}}{\mathrm{D}\hat{t}} + \hat{N}^2 N_0^2\frac{HU}{L}\hat{w} = 0, \qquad (3.61)$$

or, re-arranging,

$$Ro\frac{\mathrm{D}\hat{b}}{\mathrm{D}\hat{t}} + \left(\frac{L_d}{L}\right)^2 \hat{N}^2\hat{w} = 0. \qquad (3.62)$$

The nondimensional primitive equations are summarized in the box above.

3.5 PLANETARY-GEOSTROPHIC EQUATIONS FOR STRATIFIED FLOW

We use the inviscid and adiabatic Boussinesq equations of motion with the hydrostatic approximation. The essential assumptions in deriving the PG equations are:

1. $Ro \ll 1$
2. $(L_d/L)^2 \ll 1$. And specifically $(L_d/L)^2 = \mathcal{O}(Ro)$.

We are also assuming that time scales advectively and we allow f to have a full variation.

Given these assumptions the only simplification we make to the equations in the shaded box on the preceding page is that the momentum equation is replaced by geostrophic balance. Then, in dimensional form, the *planetary-geostrophic* equations of motion are:

$$\begin{gathered} \frac{\mathrm{D}b'}{\mathrm{D}t} + wN^2 = 0. \\ \boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi', \qquad \frac{\partial \phi'}{\partial z} = b', \qquad \nabla \cdot \boldsymbol{v} = 0 \end{gathered} \quad . \tag{3.63}$$

The thermodynamic equation may also be written simply as

$$\frac{\mathrm{D}b}{\mathrm{D}t} = \dot{b}, \tag{3.64}$$

where b now represents the total stratification. The relevant pressure, ϕ, is then the pressure that is in hydrostatic balance with b, so that geostrophic and hydrostatic balance are most usefully written as

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi, \qquad \frac{\partial \phi}{\partial z} = b. \tag{3.65a,b}$$

3.5.1 Potential vorticity

Manipulation of (3.63) reveals that we can equivalently write the equations as an evolution equation for potential vorticity. Thus, the evolution equations may be written as

$$\begin{aligned} \frac{DQ}{Dt} &= \dot{Q} \\ Q &= f\frac{\partial b}{\partial z} \end{aligned}, \qquad (3.66)$$

where $\dot{Q} = f\partial\dot{b}/\partial z$, and the inversion — i.e., the diagnosis of velocity, pressure and buoyancy — is carried out using the hydrostatic, geostrophic and mass conservation equations.

3.5.2 Applicability to the ocean and atmosphere

In the atmosphere a typical deformation radius NH/f is about 1 000 km. The constraint that the scale of motion be much larger than the deformation radius is thus quite hard to satisfy, since one quickly runs out of room on a planet whose equator-to-pole distance is 10 000 km. Thus, only the largest planetary waves can satisfy the planetary-geostrophic scaling in the atmosphere and we should then also write the equations in spherical coordinates. In the ocean the deformation radius is about 100 km, so there is lots of room for the planetary-geostrophic equations to hold, and indeed much of the theory of the large-scale structure of the ocean involves the planetary-geostrophic equations.

3.6 The Continuously Stratified Quasi-Geostrophic System

We now consider the quasi-geostrophic equations for the continuously stratified hydrostatic system. The primitive equations of motion are given by (3.45), and we extract the mean stratification so that the thermodynamic equation is given by (3.47). We stay on the β-plane for simplicity.

3.6.1 Scaling and assumptions

The non-dimensionalization and scaling are initially precisely that of section 3.4 and the nondimensional equations are just those in the shaded box on page 77. The Coriolis parameter is given

$$\boldsymbol{f} = (f_0 + \beta y)\,\widehat{\mathbf{k}} \tag{3.67}$$

The variation of the Coriolis parameter is assumed to be small (this is a key difference between the quasi-geostrophic system and the planetary-geostrophic system), and in particular we shall assume that βy is approximately the size of the relative vorticity, and so is much smaller than f_0 itself. The assumptions needed to derive the QG system are:

1. The Rossby number is small, $Ro \ll 1$.
2. Length scales are of the same order as the deformation radius, $L \sim L_d$ or $L/L_d = \mathcal{O}(1)$.
3. Variations in Coriolis parameter are small, and specifically $|\beta y| \sim Ro\, f_0$.

Given these assumptions, we can write the horizontal velocity as the sum of a geostrophic component and an ageostrophic one:

$$\begin{gathered} \boldsymbol{u} = \boldsymbol{u}_g + \boldsymbol{u}_a, \\ \text{where}\quad f_0\widehat{\mathbf{k}} \times \boldsymbol{u}_g = -\nabla\phi \quad \text{and} \quad |\boldsymbol{u}_g| \gg |\boldsymbol{u}_a|. \end{gathered} \tag{3.68}$$

I follows from the definition of the geostrophic velocity that its divergence is zero; that is

$$\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} = 0. \tag{3.69}$$

The vertical velocity is thus given by the divergence of the *ageostrophic* velocity,

$$\frac{\partial w}{\partial z} = -\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y}. \tag{3.70}$$

Since the ageostrophic velocity is small, the actual vertical velocity is smaller than the scaling suggested by the mass conservation equation in its original form. That is,

$$W \ll \frac{UH}{L}. \tag{3.71}$$

3.6.2 Derivation of Stratified QG Equations

For reference we write down the primitive equations of motion again. These are

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla_z \phi, \tag{3.72a}$$

$$\frac{\partial \phi}{\partial z} = b, \tag{3.72b}$$

$$\frac{\mathrm{D}b'}{\mathrm{D}t} + N^2 w = 0, \tag{3.72c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \tag{3.72d}$$

These are the horizontal momentum equation, the hydrostatic equation, the thermodynamic equation and the mass continuity equation. The material derivative is three dimensional.

We being by cross differentiating the horizontal momentum equation to give, after a few lines of algebra, the vorticity equation:

$$\frac{\mathrm{D}}{\mathrm{D}t}(\zeta + f) = -(\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \left(\frac{\partial u}{\partial z}\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\frac{\partial w}{\partial x}\right). \tag{3.73}$$

We now apply the above quasi-geostrophic assumptions, so that:

1. The geostrophic velocity and vorticity are much larger than their ageostrophic counterparts, so we use geostrophic values for the terms on the left-hand side.
2. On the right hand side we keep the horizontal divergence (which is small) on the right-hand side where it is multiplied by the big term f. Furthermore, because f is nearly constant we replace it with f_0 except where it is differentiated.
3. The second term (tilting) on the right-hand side is smaller than the advection terms on the left-hand side by the ratio $[UW/(HL)]/[U^2/L^2] = [W/H]/[U/L] \ll 1$, because w is small, as noted above

Given the above, (3.73) becomes

$$\frac{\mathrm{D}_g}{\mathrm{D}t}(\zeta_g + f) = -f_0\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = f_0\frac{\partial w}{\partial z}, \tag{3.74}$$

where the second equality uses mass continuity and $\mathrm{D}_g/\mathrm{D}t = \partial/\partial t + \boldsymbol{u}_g \cdot \nabla$ — note that only the (horizontal) geostrophic velocity does any advecting.

Now consider the three-dimensional thermodynamic equation. Since w is small it only advects the basic state, and the perturbation buoyancy is advected only by the geostrophic velocity. Thus, (3.72c) becomes

$$\frac{\mathrm{D}_g b'}{\mathrm{D}t} + wN^2 = 0. \tag{3.75}$$

We now eliminate w between (3.74) and (3.75), and (with some algebra) gives

$$\frac{\mathrm{D}_g q}{\mathrm{D}t} = 0, \qquad q = \zeta_g + f + \frac{\partial}{\partial z}\left(\frac{f_0 b'}{N^2}\right). \tag{3.76}$$

Hydrostatic and geostrophic wind balance enable us to write the geostrophic velocity, vorticity, and buoyancy in terms of streamfunction ψ [$= p/(f_0 \rho_0)$]:

$$\boldsymbol{u}_g = \mathbf{k} \times \nabla\psi, \qquad \zeta_g = \nabla^2\psi, \qquad b' = f_0 \partial\psi/\partial z. \tag{3.77}$$

Thus, we have, now omitting the subscript g,

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \qquad q = \nabla^2\psi + f + f_0^2 \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial\psi}{\partial z}\right), \tag{3.78a,b}$$

Only the variable part of f (e.g., βy) is relevant in the second term on the right-hand side of the expression for q. The material derivative may be expressed as

$$\frac{\mathrm{D}w}{\mathrm{D}t} = \partial q/\partial t + J(\psi, q). \tag{3.79}$$

The quantity q is known as the *quasi-geostrophic potential vorticity.* It is analogous to the exact (Ertel) potential vorticity (see section **??** for more about this), and it is conserved when advected by the *horizontal* geostrophic flow. All the other dynamical variables may be obtained from potential vorticity as follows.

(i) Streamfunction, using (3.78b).

(ii) Velocity: $\boldsymbol{u} = \mathbf{k} \times \nabla\psi$ [$\equiv \nabla^{\perp}\psi = -\nabla \times (\mathbf{k}\psi)$].

(iii) Relative vorticity: $\zeta = \nabla^2\psi$.

(iv) Perturbation pressure: $\phi = f_0\psi$.

(v) Perturbation buoyancy: $b' = f_0\partial\psi/\partial z$.

The length scale $L_d = NH/f_0$, emerges naturally from the QG dynamics. It is the scale at which buoyancy and relative vorticity effects contribute equally to the potential vorticity, and is called the *deformation radius;* it is analogous to the quantity $\sqrt{gH}/f_0$ arising in shallow water theory. In the upper ocean, with $N \approx 10^{-2}\,\mathrm{s}^{-1}$, $H \approx 10^3$ m and $f_0 \approx 10^{-4}\,\mathrm{s}^{-1}$, then $L_d \approx 100$ km. At high latitudes the ocean is much less stratified and f is somewhat larger, and the deformation radius may be as little as 30 km. In the atmosphere, with $N \approx 10^{-2}\,\mathrm{s}^{-1}$, $H \approx 10^4$ m, then $L_d \approx 1000$ km. It is this order of magnitude difference in the deformation scales that accounts for a great deal of the quantitative difference in the dynamics of the ocean and the atmosphere. If we take the limit $L_d \to \infty$ then the stratified quasi-geostrophic equations reduce to

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \qquad q = \nabla^2\psi + f. \tag{3.80}$$

This is the two-dimensional vorticity equation, identical to (**??**). The high stratification of this limit has suppressed all vertical motion, and variations in the flow become confined to the horizontal plane. Finally, we note that it is typical in quasi-geostrophic applications to omit the prime on the buoyancy perturbations, and write $b = f_0\partial\psi/\partial z$; however, we will keep the prime in this chapter.

3.6.3 Buoyancy advection at the surface

The solution of the elliptic equation in (3.78) requires vertical boundary conditions on ψ at the ground and at the top of the atmosphere, and these are given by use of the thermodynamic equation. For a flat, slippery, rigid surface the vertical velocity is zero so that the thermodynamic equation may be written as

$$\frac{\mathrm{D}b'}{\mathrm{D}t} = 0, \qquad b' = f_0\frac{\partial\psi}{\partial z}. \tag{3.81}$$

We apply this at the ground and at the tropopause, treating the latter as a lid on the lower atmosphere. In the presence of friction and topography the vertical velocity is not zero, but is given by

$$w = r\nabla^2\psi + \boldsymbol{u}\cdot\nabla\eta_b \qquad (3.82)$$

where the first term represents Ekman friction (with the constant r proportional to the thickness of the Ekman layer) and the second term represents topographic forcing. The boundary condition becomes

$$\frac{\partial}{\partial t}\left(f_0\frac{\partial\psi}{\partial z}\right)+\boldsymbol{u}\cdot\nabla\left(f_0\frac{\partial\psi}{\partial z}+N^2\eta_b\right)+N^2r\nabla^2\psi = 0, \quad (3.83)$$

where all the fields are evaluated at $z = 0$ or $z = H$, the height of the lid. Thus, the quasi-geostrophic system is characterized by the horizontal advection of potential vorticity in the interior and the advection of buoyancy at the boundary. Instead of a lid at the top, then in a compressible fluid such as the atmosphere we may suppose that all disturbances tend to zero as $z \to \infty$.

3.7 Energetics of Quasi-Geostrophy

If the quasi-geostrophic set of equations is to represent a real fluid system in a physically meaningful way, then it should have a consistent set of energetics. In particular, the total energy should be conserved, and there should be analogs of kinetic and potential energy and conversion between the two. We now show that such energetic properties do hold, using the Boussinesq set as an example.

Let us write the governing equations as a potential vorticity equation in the interior,

$$\frac{\mathrm{D}}{\mathrm{D}t}\left[\nabla^2\psi + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial\psi}{\partial z}\right)\right] + \beta\frac{\partial\psi}{\partial x} = 0, \quad 0 < z < 1, \qquad (3.84)$$

and buoyancy advection at the boundary,

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{\partial\psi}{\partial z}\right) = 0, \quad z = 0, 1. \quad (3.85)$$

For lateral boundary conditions we may assume that ψ = constant, or impose periodic conditions. If we multiply (3.84) by $-\psi$ and integrate over the domain, using the boundary conditions, we easily find

$$\frac{\mathrm{d}\widehat{E}}{\mathrm{d}t} = 0, \qquad \widehat{E} = \frac{1}{2}\int_V \left[(\nabla\psi)^2 + \frac{f_0^2}{N^2}\left(\frac{\partial\psi}{\partial z}\right)^2 \right] \mathrm{d}V. \tag{3.86a,b}$$

The term involving β makes no direct contribution to the energy budget. Equation (3.86) is the fundamental energy equation for quasi-geostrophic motion, and it states that in the absence of viscous or diabatic terms the total energy is conserved. The two terms in (3.86b) can be identified as the kinetic energy (KE) and available potential energy (APE) of the flow, where

$$KE = \frac{1}{2}\int_V (\nabla\psi)^2 \,\mathrm{d}V, \qquad APE = \frac{1}{2}\int_V \frac{f_0^2}{N^2}\left(\frac{\partial\psi}{\partial z}\right)^2 \mathrm{d}V. \tag{3.87a,b}$$

The available potential energy may also be written as

$$APE = \frac{1}{2}\int_V \frac{H^2}{L_d^2}\left(\frac{\partial\psi}{\partial z}\right)^2 \mathrm{d}V, \tag{3.88}$$

where L_d is the deformation radius NH/f_0 and we may choose H such that $z \sim H$. At some scale L the ratio of the kinetic energy to the potential energy is thus, roughly,

$$\frac{KE}{APE} \sim \frac{L_d^2}{L^2}. \tag{3.89}$$

For scales much larger than L_d the potential energy dominates the kinetic energy, and contrariwise.

3.7.1 Conversion between APE and KE

Let us return to the vorticity and thermodynamic equations,

$$\frac{\mathrm{D}\zeta}{\mathrm{D}t} = f\frac{\partial w}{\partial z} \tag{3.90}$$

where $\zeta = \nabla^2\psi$, and

$$\frac{\mathrm{D}b'}{\mathrm{D}t} + N^2 w = 0 \tag{3.91}$$

where $b' = f_0 \partial\psi/\partial z$. From (3.90) we form a kinetic energy equation namely

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_V (\nabla\psi)^2 \,\mathrm{d}V = -\int_V f_0 \frac{\partial w}{\partial z}\psi \,\mathrm{d}V = \int_V f_0 w \frac{\partial \psi}{\partial z}\,\mathrm{d}V. \tag{3.92}$$

From (3.91) we form a potential energy equation, namely

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int_V \frac{f_0^2}{N^2}\left(\frac{\partial\psi}{\partial z}\right)^2 \mathrm{d}V = -\int_V f_0 w \frac{\partial\psi}{\partial z}\,\mathrm{d}V. \tag{3.93}$$

Thus, the *conversion* from APE to KE is represented by

$$\frac{\mathrm{d}}{\mathrm{d}t}KE = -\frac{\mathrm{d}}{\mathrm{d}t}APE = \int_v f_0 w \frac{\partial\psi}{\partial z}\,\mathrm{d}V. \tag{3.94}$$

Because the buoyancy is proportional to $\partial\psi/\partial z$, when warm fluid rises there is a correlation between w and $\partial\psi/\partial z$ and APE is converted to KE. Whether such a phenomenon occurs depends of course on the dynamics of the flow; however, such a conversion *is,* in fact, a common feature of geophysical flows.

Chapter 4

Rossby Waves

Weeks 7 to 9

4.1 Fundamentals and Formalities

4.1.1 Wave propagation and phase speed

Consider the propagation of monochromatic plane waves satisfying

$$\psi = \operatorname{Re} \widetilde{\psi} e^{i\theta(\boldsymbol{x},t)} = \operatorname{Re} \widetilde{\psi} e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{4.1}$$

where $\widetilde{\psi}$ is a complex constant, θ is the phase, ω is the wave frequency and $\boldsymbol{k}$ is the vector wavenumber (k, l, m) (also written as (k^x, k^y, k^z) or, in subscript notation, k_i). The prefix Re denotes the real part of the expression, but we will drop it if there is no ambiguity. Given (4.1) a wave will propagate in the direction of $\boldsymbol{k}$ (Fig. 4.1). At a given instant and location we can align our coordinate axis along this direction, and we write $\boldsymbol{k} \cdot \boldsymbol{x} = Kx^*$, where x^* increases in the direction of $\boldsymbol{k}$ and $K^2 = |\boldsymbol{k}|^2$ is the magnitude of the wavenumber. With this, we can write (4.1) as

$$\psi = \operatorname{Re} \widetilde{\psi} e^{i(Kx^*-\omega t)} = \operatorname{Re} \widetilde{\psi} e^{iK(x^*-ct)}, \tag{4.2}$$

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

$$c_p \equiv \frac{\omega}{K}. \tag{4.3}$$

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

$$\lambda = \frac{2\pi}{K}. \tag{4.4}$$

In (for simplicity) a two-dimensional wave, and referring to Fig. 4.1, the wavelength and wave vectors in the x- and y-directions are given by,

$$\lambda^x = \frac{\lambda}{\cos\phi}, \quad \lambda^y = \frac{\lambda}{\sin\phi}, \qquad k^x = K\cos\phi, \quad k^y = K\sin\phi. \tag{4.5}$$

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

$$c_p^x = c_p\frac{l^x}{l} = \frac{c_p}{\cos\phi} = c_p\frac{K}{k^x} = \frac{\omega}{k^x}, \qquad c_p^y = c_p\frac{l^y}{l} = \frac{c_p}{\sin\phi} = c_p\frac{K}{k^y} = \frac{\omega}{k^y}, \tag{4.6}$$

using (4.3) and (4.5), and again referring to Fig. 4.1 for notation. The speed of phase propagation along any one of the axis is in general *larger* than the phase speed in the primary direction of the wave. The phase speeds are clearly *not* components of a vector: for example, $c_p^x \neq c_p\cos\phi$. Analogously, the wavevector $\boldsymbol{k}$ is a true vector, whereas the wavelength λ is not.

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

$$c_p = \frac{\omega}{K}, \qquad c_p^x = \frac{\omega}{k^x}, \quad c_p^y = \frac{\omega}{k^y}. \tag{4.7}$$

4.1.2 The dispersion relation

The above description is mostly kinematic and a little abstract, applying to almost any disturbance that has a wavevector and a frequency. The particular *dynamics* of a wave are determined by the relationship between the wavevector and the frequency; that is, by the *dispersion relation*. Once

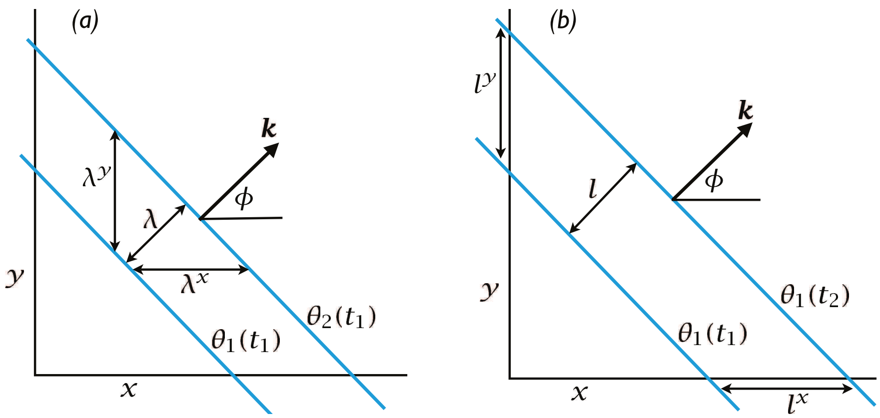

Figure 4.1 The propagation of a two-dimensional wave. (a) Two lines of constant phase (e.g., two wavecrests) at a time t_1. The wave is propagating in the direction $\boldsymbol{k}$ with wavelength λ. (b) The same line of constant phase at two successive times. The phase speed is the speed of advancement of the wavecrest in the direction of travel, and so $c_p = l/(t_2 - t_1)$. The phase speed in the x-direction is the speed of propagation of the wavecrest along the x-axis, and $c_p^x = l^x/(t_2 - t_1) = c_p/\cos\phi$.

the dispersion relation is known a great many of the properties of the wave follow in a more-or-less straightforward manner, as we will see. Picking up from (**??**), the dispersion relation is a functional relationship between the frequency and the wavevector of the general form

$$\omega = \Omega(\boldsymbol{k}). \tag{4.8}$$

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

$$\frac{\partial \psi}{\partial t} + c\frac{\partial \psi}{\partial x} = 0. \tag{4.9}$$

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

$$\omega = ck. \tag{4.10}$$

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

$$\frac{\partial \psi}{\partial t} + \boldsymbol{c} \cdot \nabla \psi = 0, \quad \omega = \boldsymbol{c} \cdot \boldsymbol{k}, \quad c_p = |\boldsymbol{c}| \cos\theta, \quad c_p^x = \frac{\boldsymbol{c} \cdot \boldsymbol{k}}{k}, \quad c_p^y = \frac{\boldsymbol{c} \cdot \boldsymbol{k}}{l} \tag{4.11a}$$

$$\frac{\partial^2 \psi}{\partial t^2} - c^2 \nabla^2 \psi = 0, \quad \omega^2 = c^2 K^2, \quad c_p = \pm c, \quad c_p^x = \pm \frac{cK}{k}, \quad c_p^y = \pm \frac{cK}{l}, \tag{4.11b}$$

$$\frac{\partial}{\partial t} \nabla^2 \psi + \beta \frac{\partial \psi}{\partial x} = 0, \quad \omega = \frac{-\beta k}{K^2}, \quad c_p = \frac{\omega}{K}, \quad c_p^x = -\frac{\beta}{K^2}, \quad c_p^y = -\frac{\beta k/l}{K^2}. \tag{4.11c}$$

where $K^2 = k^2 + l^2$ and θ is the angle between $\boldsymbol{c}$ and $\boldsymbol{k}$, and the examples are all two-dimensional, with variation in x and y only.

A wave is said to be *nondispersive* or *dispersionless* if the phase speed is independent of the wavelength. This condition is clearly satisfied for the simple example (4.9) but is manifestly not satisfied for (4.11c), and these waves (Rossby waves, in fact) are *dispersive*. Waves of different wavelengths then travel at different speeds so that a group of waves will spread out — disperse — even if the medium is homogeneous. When a wave is dispersive there is another characteristic speed at which the waves propagate, known as the group velocity, and we come to this in the next section.

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

$$\omega = \Omega(\boldsymbol{k}; \boldsymbol{x}, t). \tag{4.12}$$

Although Ω is a function of $\boldsymbol{k}$, $\boldsymbol{x}$ and t the semi-colon in (4.12) is used to suggest that x and t are slowly varying parameters of a somewhat different nature than $\boldsymbol{k}$. We'll resume our discussion of this topic in section **??**, but before that we must introduce the group velocity.

4.2 Group Velocity

Information and energy clearly cannot travel at the phase speed, for as the direction of propagation of the phase line

Wave Fundamentals

- A wave is a propagating disturbance that has a characteristic relationship between its frequency and size, known as the dispersion relation. Waves typically arise as solutions to a linear problem of the form
$$L(\psi) = 0, \tag{WF.1}$$
where L is, commonly, a linear operator in space and time. Two examples are
$$\frac{\partial^2 \psi}{\partial t^2} - c^2 \nabla^2 \psi = 0 \qquad \text{and} \qquad \frac{\partial}{\partial t}\nabla^2 \psi + \beta \frac{\partial \psi}{\partial x} = 0. \tag{WF.2}$$
The first example is s
- Solutions to the governing equation are often sought in the form of plane waves that have the form
$$\psi = \mathrm{Re}\, A \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{WF.3}$$
where A is the wave amplitude, $\boldsymbol{k} = (k, l, m)$ is the wavevector, and ω is the frequency.
- The dispersiono common in all areas of physics it is sometimes called 'the' wave equation. The second example gives rise to Rossby waves. relation connects the frequency and wavevector through an equation of the form $\omega = \Omega(\boldsymbol{k})$ where Ω is some function. The relation is normally derived by substituting a trial solution like (WF.3) into the governing equation (WF.1). For the examples of (WF.2) we obtain $\omega = c^2 K^2$ and $\omega = -\beta k / K^2$ where $K^2 = k^2 + l^2 + m^2$ or, in two dimensions, $K^2 = k^2 + l^2$.
- The phase speed is the speed at which the wave crests move. In the direction of propagation and in the x, y and z directions the phase speed is given by, respectively,
$$c_p = \frac{\omega}{K}, \qquad c_p^x = \frac{\omega}{k}, \quad c_p^y = \frac{\omega}{l}, \quad c_p^z = \frac{\omega}{m}. \tag{WF.4}$$
where $K = 2\pi/\lambda$ where λ is the wavelength. The wave crests have both a speed (c_p) and a direction of propagation (the direction of $\boldsymbol{k}$), like a vector, but the components defined in (WF.4) are not the components of that vector.
- The group velocity is the velocity at which a wave packet or wave group moves. It is a vector and is given by
$$\boldsymbol{c}_g = \frac{\partial \omega}{\partial \boldsymbol{k}} \quad \text{with components} \quad c_g^x = \frac{\partial \omega}{\partial k}, \quad c_g^y = \frac{\partial \omega}{\partial l}, \quad c_g^z = \frac{\partial \omega}{\partial m}. \tag{WF.5}$$
Most physical quantities of interest are transported at the group velocity.
- If the coefficients of the wave equation are not constant (for example if the medium is inhomogeneous) then, if the coefficients are only slowly varying, approximate solutions may sometimes be found in the form
$$\psi = \mathrm{Re}\, A(\boldsymbol{x}, t) \mathrm{e}^{\mathrm{i}\theta(\boldsymbol{x}, t)}, \tag{WF.6}$$
where the amplitude A is also slowly varying and the local wavenumber and frequency are related to the phase, θ, by $\boldsymbol{k} = \nabla\theta$ and $\omega = -\partial\theta/\partial t$. The dispersion relation is then a *local* one of the form $\omega = \Omega(\boldsymbol{k}; x, t)$.

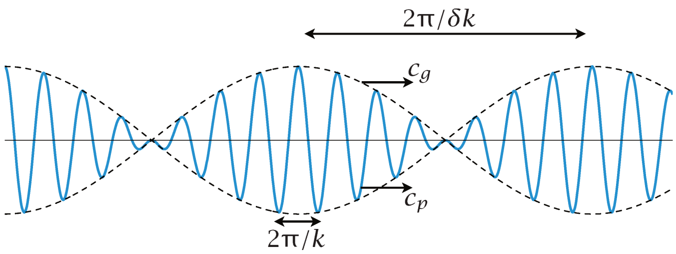

Figure 4.2 Superposition of two sinusoidal waves with wavenumbers k and $k+\delta k$, producing a wave (solid line) that is modulated by a slowly varying wave envelope or wave packet (dashed line). The envelope moves at the group velocity, $c_g = \partial\omega/\partial k$ and the phase of the wave moves at the group speed $c_p = \omega/k$.

tends to a direction parallel to the y-axis, the phase speed in the x-direction tends to infinity! Rather, it turns out that most quantities of interest, including energy, propagate at the *group velocity,* a quantity of enormous importance in wave theory.[7] Roughly speaking, group velocity is the velocity at which a packet or a group of waves will travel, whereas the individual wave crests travel at the phase speed. To introduce the idea we will consider the superposition of plane waves, noting that a monochromatic plane wave already fills space uniformly so that there can be no propagation of energy from place to place. We will restrict attention to waves propagating in one direction, but the argument may be extended to two or three dimensions.

4.2.1 Superposition of two waves

Consider the linear superposition of two waves. Limiting attention to the one-dimensional case for simplicity, consider a disturbance represented by

$$\psi = \operatorname{Re}\widetilde{\psi}(\mathrm{e}^{\mathrm{i}(k_1 x-\omega_1 t)} + \mathrm{e}^{\mathrm{i}(k_2 x-\omega_2 t)}). \tag{4.13}$$

Let us further suppose that the two waves have similar wavenumbers and frequency, and, in particular, that $k_1 = k + \Delta k$ and $k_2 = k - \Delta k$, and $\omega_1 = \omega + \Delta\omega$ and $\omega_2 = \omega - \Delta\omega$. With this, (4.13) becomes

$$\begin{aligned}\psi &= \operatorname{Re}\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx-\omega t)}[\mathrm{e}^{\mathrm{i}(\Delta k\, x-\Delta\omega\, t)} + \mathrm{e}^{-\mathrm{i}(\Delta k\, x-\Delta\omega\, t)}] \\ &= 2\operatorname{Re}\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx-\omega t)}\cos(\Delta k\, x - \Delta\omega\, t).\end{aligned} \tag{4.14}$$

The resulting disturbance, illustrated in Fig. 4.2 has two aspects: a rapidly varying component, with wavenumber k and frequency ω, and a more slowly varying envelope, with wavenumber Δk and frequency $\Delta\omega$. The envelope modulates the fast oscillation, and moves with velocity $\Delta\omega/\Delta k$; in the limit $\Delta k \to 0$ and $\Delta\omega \to 0$ this is the *group velocity*, $c_g = \partial\omega/\partial k$. Group velocity is equal to the phase speed, ω/k, only when the frequency is a linear function of wavenumber. The energy in the disturbance must move at the group velocity — note that the node of the envelope moves at the speed of the envelope and no energy can cross the node. These concepts generalize to more than one dimension, and if the wavenumber is the three-dimensional vector $\boldsymbol{k} = (k, l, m)$ then the three-dimensional envelope propagates at the group velocity given by

$$\boldsymbol{c}_g = \frac{\partial\omega}{\partial\mathbf{k}} \equiv \left(\frac{\partial\omega}{\partial k}, \frac{\partial\omega}{\partial l}, \frac{\partial\omega}{\partial m}\right) . \tag{4.15}$$

The group velocity is also written as $\boldsymbol{c}_g = \nabla_{\boldsymbol{k}}\omega$ or, in subscript notation, $c_{gi} = \partial\Omega/\partial k_i$, with the subscript i denoting the component of a vector.

4.3 Rossby Waves

4.3.1 The linear equation of motion

For most of the rest of this chapter we will be concerned with the quasi-geostrophic equations of motion for which (as discussed in chapter 3) the inviscid, adiabatic potential vorticity equation is

$$\frac{\partial q}{\partial t} + \boldsymbol{u}\cdot\nabla q = 0, \tag{4.16}$$

where $q(x, y, z, t)$ is the potential vorticity and $\boldsymbol{u}(x, y, z, t)$ is the horizontal velocity. The velocity is related to a streamfunction by $u = -\partial\psi/\partial y$, $v = \partial\psi/\partial x$ and the potential vorticity is some function of the streamfunction, which might differ from system to system. Two examples, one applying to a continuously stratified system and the second to a single layer system, are

$$q = f + \zeta + \frac{\partial}{\partial z}\left(S(z)\frac{\partial\psi}{\partial z}\right), \qquad q = \zeta + f - k_d^2\psi. \tag{4.17a,b}$$

where $S(z) = f_0^2/N^2$, $\zeta = \nabla^2\psi$ is the relative vorticity and $k_d = 1/L_d$ is the inverse radius of deformation for a shallow water system. (Note that definitions of k_d and L_d can vary, typically by factors of 2, π, etc.) Boundary conditions may be needed to form a complete system.

We now *linearize* (4.16); that is, we suppose that the flow consists of a time-independent component (the 'basic state') plus a perturbation, with the perturbation being small compared with the mean flow. The basic state must satisfy the time-independent equation of motion, and it is common and useful to linearize about a zonal flow, $\overline{u}(y,z)$. The basic state is then purely a function of y and so we write

$$q = \overline{q}(y,z) + q'(x,y,t), \qquad \psi = \overline{\psi}(y,z) + \psi'(x,y,z,t) \tag{4.18}$$

with a similar notation for the other variables. Note that $\overline{u} = -\partial\overline{\psi}/\partial y$ and $\overline{v} = 0$. Substituting into (4.16) gives, without approximation,

$$\frac{\partial q'}{\partial t} + \overline{\boldsymbol{u}} \cdot \nabla\overline{q} + \overline{\boldsymbol{u}} \cdot \nabla q' + \boldsymbol{u}' \cdot \nabla\overline{q} + \boldsymbol{u}' \cdot \nabla q' = 0. \tag{4.19}$$

The primed quantities are presumptively small so we neglect terms involving their products. Further, we are assuming that we are linearizing about a state that is a solution of the equations of motion, so that $\overline{\boldsymbol{u}} \cdot \nabla\overline{q} = 0$. Finally, since $\overline{v} = 0$ and $\partial\overline{q}/\partial x = 0$ we obtain

$$\frac{\partial q'}{\partial t} + \overline{u}\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = 0. \tag{4.20}$$

This equation or one very similar appears very commonly in studies of Rossby waves. To proceed, let us consider the simple example of waves in a single layer.

4.3.2 Waves in a single layer

Consider a system obeying (4.16) and (4.17b). The equation could be written in spherical coordinates with $f = 2\Omega\sin\vartheta$, but the dynamics are more easily illustrated on Cartesian β-plane for which $f = f_0 + \beta y$, and since f_0 is a constant it does not appear in our subsequent derivations.

Infinite deformation radius

If the scale of motion is much less than the deformation scale then we make the approximation that $k_d = 0$ and the equation of motion may be written as

$$\frac{\partial \zeta}{\partial t} + \boldsymbol{u} \cdot \nabla \zeta + \beta v = 0 \tag{4.21}$$

We linearize about a constant zonal flow, $\overline{u} = U$, by writing

$$\psi = \overline{\psi}(y) + \psi'(x, y, t), \tag{4.22}$$

where $\overline{\psi} = -Uy$. Substituting(4.22) into (4.21) and neglecting the nonlinear terms involving products of ψ' to give

$$\frac{\partial}{\partial t}\nabla^2\psi' + U\frac{\partial \nabla^2 \psi'}{\partial x} + \beta\frac{\partial \psi'}{\partial x} = 0. \tag{4.23}$$

This equation is just a single-layer version of (4.20), with $\partial\overline{q}/\partial y = \beta$, $q' = \nabla^2\psi'$ and $v' = \partial\psi'/\partial x$.

The coefficients in (4.23) are not functions of y or z; this is not a requirement for wave motion to exist but it does enable solutions to be found more easily. Let us seek solutions in the form of a plane wave, namely

$$\psi' = \mathrm{Re}\, \widetilde{\psi} \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \tag{4.24}$$

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

Substituting (4.24) into (4.23) yields

$$[(-\omega + Uk)(-K^2) + \beta k]\widetilde{\psi} = 0, \tag{4.25}$$

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

$$\omega = Uk - \frac{\beta k}{K^2}. \tag{4.26}$$

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

$$c_p^x \equiv \frac{\omega}{k} = U - \frac{\beta}{K^2}, \qquad c_p^y \equiv \frac{\omega}{l} = U\frac{k}{l} - \frac{\beta k}{K^2 l}, \quad \text{(4.27a,b)}$$

and

$$c_g^x \equiv \frac{\partial\omega}{\partial k} = U + \frac{\beta(k^2 - l^2)}{(k^2 + l^2)^2}, \qquad c_g^y \equiv \frac{\partial\omega}{\partial l} = \frac{2\beta kl}{(k^2 + l^2)^2}. \quad \text{(4.28a,b)}$$

The phase speed in the absence of a mean flow is *westwards,* with waves of longer wavelengths travelling more quickly, and the eastward current speed required to hold the waves of a particular wavenumber stationary (i.e., $c_p^x = 0$) is $U = \beta/K^2$. The background flow U evidently just provides a uniform shift to the phase speed, and could be transformed away by a change of coordinate.

Finite deformation radius

For a finite deformation radius the basic state $\Psi = -Uy$ is still a solution of the original equations of motion, but the potential vorticity corresponding to this state is $q = Uyk_d^2 + \beta y$ and its gradient is $\nabla q = (\beta + Uk_d^2)\mathbf{j}$. The linearized equation of motion is thus

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)(\nabla^2\psi' - \psi' k_d^2) + (\beta + Uk_d^2)\frac{\partial\psi'}{\partial x} = 0. \quad \text{(4.29)}$$

Substituting $\psi' = \widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}$ we obtain the dispersion relation,

$$\omega = \frac{k(UK^2 - \beta)}{K^2 + k_d^2} = Uk - k\frac{\beta + Uk_d^2}{K^2 + k_d^2}. \quad \text{(4.30)}$$

The corresponding components of phase speed and group velocity are

$$c_p^x = U - \frac{\beta + Uk_d^2}{K^2 + k_d^2} = \frac{UK^2 - \beta}{K^2 + k_d^2}, \qquad c_p^y = U\frac{k}{l} - \frac{k}{l}\left(\frac{UK^2 - \beta}{K^2 + k_d^2}\right) \quad \text{(4.31a,b)}$$

and

$$c_g^x = U + \frac{(\beta + Uk_d^2)(k^2 - l^2 - k_d^2)}{(k^2 + l^2 + k_d^2)^2}, \quad c_g^y = \frac{2kl(\beta + Uk_d^2)}{(k^2 + l^2 + k_d^2)^2}. \tag{4.32a,b}$$

The uniform velocity field now no longer provides just a simple Doppler shift of the frequency, nor a uniform addition to the phase speed. From (4.31a) the waves are stationary when $K^2 = \beta/U \equiv K_s^2$; that is, the current speed required to hold waves of a particular wavenumber stationary is $U = \beta/K^2$. However, this is *not* simply the magnitude of the phase speed of waves of that wavenumber in the absence of a current — this is given by

$$c_p^x = \frac{-\beta}{K_s^2 + k_d^2} = \frac{-U}{1 + k_d^2/K_s^2} \neq -U. \tag{4.33}$$

Why is there a difference? It is because the current does not just provide a uniform translation, but, if k_d is non-zero, it also modifies the basic potential vorticity gradient. The basic state height field η_0 is sloping; that is $\eta_0 = -(f_0/g)Uy$, and the ambient potential vorticity field increases with y and $q = (\beta + Uk_d^2)y$. Thus, the basic state defines a preferred frame of reference, and the problem is not Galilean invariant.[8]

We also note that, from (4.31b), the group velocity is negative (westward) if the x-wavenumber is sufficiently small, compared to the y-wavenumber or the deformation wavenumber. That is, said a little loosely, *long waves move information westward and short waves move information eastward,* and this is a common property of Rossby waves. The x-component of the phase speed, on the other hand, is always westward relative to the mean flow.

4.3.3 The mechanism of Rossby waves

The fundamental mechanism underlying Rossby waves is easily understood. Consider a material line of stationary fluid parcels along a line of constant latitude, and suppose that some disturbance causes their displacement to the line marked $\eta(t = 0)$ in Fig. 4.3. In the displacement, the potential vorticity of the fluid parcels is conserved, and

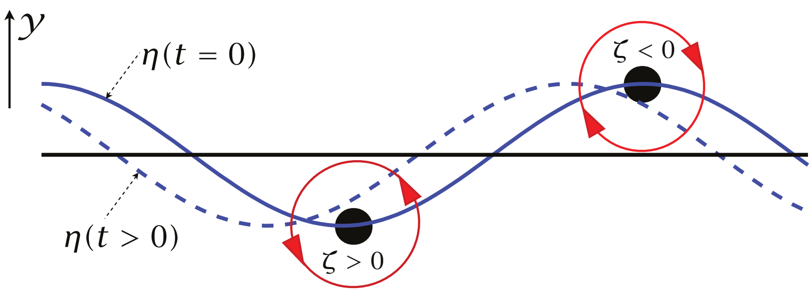

Figure 4.3 The mechanism of a two-dimensional (x–y) Rossby wave. An initial disturbance displaces a material line at constant latitude (the straight horizontal line) to the solid line marked $\eta(t = 0)$. Conservation of potential vorticity, $\beta y + \zeta$, leads to the production of relative vorticity, as shown for two parcels. The associated velocity field (arrows on the circles) then advects the fluid parcels, and the material line evolves into the dashed line. The phase of the wave has propagated westwards.

in the simplest case of barotropic flow on the β-plane the potential vorticity is the absolute vorticity, $\beta y + \zeta$. Thus, in either hemisphere, a northward displacement leads to the production of negative relative vorticity and a southward displacement leads to the production of positive relative vorticity. The relative vorticity gives rise to a velocity field which, in turn, advects the parcels in material line in the manner shown, and the wave propagates westwards.

In more complicated situations, such as flow in two layers, considered below, or in a continuously stratified fluid, the mechanism is essentially the same. A displaced fluid parcel carries with it its potential vorticity and, in the presence of a potential vorticity gradient in the basic state, a potential vorticity anomaly is produced. The potential vorticity anomaly produces a velocity field (an example of potential vorticity inversion) which further displaces the fluid parcels, leading to the formation of a Rossby wave. The vital ingredient is a basic state potential vorticity gradient, such as that provided by the change of the Coriolis parameter with latitude.

4.4 Rossby Waves in Stratified Quasi-Geostrophic Flow

4.4.1 Setting up the problem

Let us now consider the dynamics of linear waves in stratified quasi-geostrophic flow on a β-plane, with a resting basic state.

The interior flow is governed by the potential vorticity equation, (3.78), and linearizing this about a uniform E–W flow gives rest gives

$$\left[\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right]\left[\nabla^2\psi' + \frac{\partial}{\partial z}\left(F(z)\frac{\partial \psi'}{\partial z}\right)\right] + \beta\frac{\partial \psi'}{\partial x} = 0, \tag{4.34}$$

where $F(z) = f_0^2/N^2$. (F is the square of the inverse Prandtl ratio, N/f_0.) The vertical boundary conditions are determined by the thermodynamic equation, (3.81). If the boundaries are flat, rigid, slippery surfaces then $w = 0$ at the boundaries and if there is no surface buoyancy gradient the linearized thermodynamic equation is

$$\frac{\partial}{\partial t}\left(\frac{\partial \psi'}{\partial z}\right) = 0. \tag{4.35}$$

We apply this at the ground and at the tropopause, so at $z = 0$ and at $z = H$.

4.4.2 Wave motion

We may seek solutions of the form

$$\psi' = \mathrm{Re}\,\widetilde{\psi}(z)\mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \tag{4.36}$$

where $\widetilde{\psi}(z)$ will determine the vertical structure of the waves. In the zonal direction (the x-direction) the flow is periodic, and if the domain is of horizontal length L_x then we have $k = 2\pi n_x/L_x$ where $n_x = 1, 2, 3\ldots$. In there are 'walls' at $y = 0$ and $y = L_y$ where $\psi = 0$ then the y variation should be of the form $\psi' \sim \sin ly$ where $l = \pi n_y/L_y$ where n_y is an integer. However, we will keep the exponential form (4.36) for the y variation for simplicity. Finally, if $F(z)$ is a constant then the problem further simplifies and we can seek solutions of the form

$$\psi' = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx+ly+mz-\omega t)}, \tag{4.37}$$

and this is what we shall do. This solution does not of itself satisfy (4.35), and we can make it do so by restricting the vertical variations to be of the form:

$$\psi' = A\cos mz \qquad \text{where} \qquad m = n_z\pi/H, \tag{4.38}$$

where n_z is an integer. These solutions then satisfy $\partial\psi/\partial z$ at $z = 0$ and $z = H$. Having said this, we will stick with eqref[qg:sepwave2] for our manipulations, just because that is simpler, bearing in mind that the y and z variations should just be sines and cosines, respectively.

The dispersion relation is obtained by substituting (**??**)qg;sepwave2] into (4.34) giving

$$\omega = Uk - \frac{\beta k}{k^2 + l^2 + (f_0^2/N^2)m^2}. \tag{4.39}$$

It is interesting to re-write this as an equation for m, and we obtain

$$\frac{f_0^2}{N^2}m^2 = \frac{\beta}{U - c} - K^2 \tag{4.40}$$

where $K^2 = k^2 + l^2$ and $c = \omega/k$. We'll come back to this in section 4.5, and the next subsection may be skipped if you wish.

4.4.3 ♦ The case with non-constant N^2

For simplicity let $U = 0$, and then substituting (4.36) into (4.34) gives

$$\omega\left[-K^2\widetilde{\psi}(z) + \frac{1}{\widetilde{\rho}}\frac{\mathrm{d}}{\mathrm{d}z}\left(\widetilde{\rho}F(z)\frac{\mathrm{d}\widetilde{\psi}}{\mathrm{d}z}\right)\right] - \beta k\widetilde{\psi}(z) = 0. \tag{4.41}$$

Now, if $\widetilde{\psi}$ satisfies

$$\frac{1}{\widetilde{\rho}}\frac{\mathrm{d}}{\mathrm{d}z}\left(\widetilde{\rho}F(z)\frac{\mathrm{d}\widetilde{\psi}}{\mathrm{d}z}\right) = -\Gamma\widetilde{\psi}, \tag{4.42}$$

where Γ is a constant, then the equation of motion becomes

$$-\,\omega\left[K^2 + \Gamma\right]\widetilde{\psi} - \beta k\widetilde{\psi} = 0, \tag{4.43}$$

and the dispersion relation follows, namely

$$\omega = -\frac{\beta k}{K^2 + \Gamma}. \tag{4.44}$$

Equation (4.42) constitutes an eigenvalue problem for the vertical structure; the boundary conditions, derived from (4.35), are $\partial\tilde{\psi}/\partial z = 0$ at $z = 0$ and $z = H$. The resulting eigenvalues, Γ are proportional to the inverse of the squares of the deformation radii for the problem and the eigenfunctions are the vertical structure functions.

Consider the case in which $F(z)$ is constant, and in which the domain is confined between two rigid surfaces at $z = 0$ and $z = H$. Then the eigenvalue problem for the vertical structure is

$$F\frac{\mathrm{d}^2\tilde{\psi}}{\mathrm{d}z^2} = -\Gamma\tilde{\psi} \tag{4.45a}$$

with boundary conditions of

$$\frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}z} = 0, \qquad \text{at } z = 0, H. \tag{4.45b}$$

There is a sequence of solutions to this, namely

$$\tilde{\psi}_n(z) = \cos(n\pi z/H), \qquad n = 1, 2 \ldots \tag{4.46}$$

with corresponding eigenvalues

$$\Gamma_n = n^2\frac{F\pi^2}{H^2} = (n\pi)^2\left(\frac{f_0}{NH}\right)^2, \qquad n = 1, 2 \ldots. \tag{4.47}$$

Equation (4.47) may be used to define the deformation radii for this problem, namely

$$L_n \equiv \frac{1}{\sqrt{\Gamma}_n} = \frac{NH}{n\pi f_0}. \tag{4.48}$$

The first deformation radius is the same as the expression obtained by dimensional analysis, namely NH/f, except for a factor of π. (Definitions of the deformation radii both with and without the factor of π are common in the literature, and neither is obviously more correct. In the latter case, the first deformation radius in a problem with uniform stratification is given by NH/f, equal to $\pi/\sqrt{\Gamma_1}$.)

In addition to these baroclinic modes, the case with $n = 0$, that is with $\widetilde{\psi} = 1$, is also a solution of (4.45) for any $F(z)$.

Using (4.44) and (4.47) the dispersion relation becomes

$$\omega = -\frac{\beta k}{K^2 + (n\pi)^2 (f_0/NH)^2}, \qquad n = 0, 1, 2 \ldots \tag{4.49}$$

and, of course, the horizontal wavenumbers k and l are also quantized in a finite domain. This equation is the same as (4.39)

The dynamics of the barotropic mode ($n = 0$) are independent of height and independent of the stratification of the basic state, and so these Rossby waves are *identical* with the Rossby waves in a two-dimensional fluid.

4.5 Vertical Propagation of Rossby waves

4.5.1 Conditions for wave propagation

The dispersion relation is

$$m^2 = \frac{N^2}{f_0^2} \left(\frac{\beta}{U - c} - (k^2 + l^2) \right). \tag{4.50}$$

For waves to propagate upwards we require that $m^2 > 0$ and, from (4.50), that

$$0 < U - c < \frac{\beta}{k^2 + l^2}, \tag{4.51}$$

where $u_c = \beta/(k^2 + l^2)$ is the Rossby critical velocity. For waves of some given frequency ($\omega = kc$) the above expression provides a condition on U for the vertical propagation of planetary waves. For stationary waves $c = 0$ and the criterion is

$$\boxed{0 < U < \frac{\beta}{k^2 + l^2}}, \tag{4.52}$$

and this is illustrated in Fig. 4.4. That is to say, the vertical propagation of stationary Rossby waves occurs only in westerly winds, and winds that are weaker than some critical value, $u_c = \beta/(k^2 + l^2)$ that depends on the scale of the wave. If the waves can take any frequency there is no such condition on U, for (4.50) is just a form of the dispersion relation and (4.51) is naturally satisfied.

Essentials of Rossby Waves

- Rossby waves owe their existence to a gradient of potential vorticity in the fluid. If a fluid parcel is displaced, it conserves its potential vorticity and so its relative vorticity will in general change. The relative vorticity creates a velocity field that displaces neighbouring parcels, whose relative vorticity changes and so on.

- A common source of a potential vorticity gradient is differential rotation, or the β-effect, and *planetary waves* is the name given to this type of Rossby wave. In the presence of non-zero β the ambient potential vorticity increases northward and the phase of the Rossby waves propagates westward. In general, Rossby waves propagate pseudo-westwards, meaning to the left of the direction of the potential vorticity gradient.

- A common equation of motion for Rossby waves is

$$\frac{\partial q'}{\partial t} + \overline{u}\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = 0, \tag{RW.1}$$

  with an overbar denoting the basic state and a prime a perturbation. In the case of a single layer of fluid with no mean flow this equation becomes

$$\frac{\partial}{\partial t}(\nabla^2 + k_d^2)\psi' + \beta\frac{\partial \psi'}{\partial x} = 0 \tag{RW.2}$$

  with dispersion relation

$$\omega = \frac{-\beta k}{k^2 + l^2 + k_d^2}. \tag{RW.3}$$

- The phase speed in the zonal direction ($c_p^x = \omega/k$) is always negative, or westward, and is larger for large waves. For (RW.2) components of the group velocity are given by

$$c_g^x = \frac{\beta(k^2 - l^2 - k_d^2)}{(k^2 + l^2 + k_d^2)^2}, \qquad c_g^y = \frac{2\beta k l}{(k^2 + l^2 + k_d^2)^2}. \tag{RW.4}$$

  The group velocity is westward if the zonal wavenumber is sufficiently small, and eastward if the zonal wavenumber is sufficiently large.

- Rossby waves exist in stratified fluids, and have a similar dispersion relation to (RW.3) with an appropriate vertical wavenumber appearing in place of the inverse deformation radius, k_d.

- The reflection of such Rossby waves at a wall is specular, meaning that the group velocity of the reflected wave makes the same angle with the wall as the group velocity of the incident wave. The energy flux of the reflected wave is equal and opposite to that of the incoming wave in the direction normal to the wall.

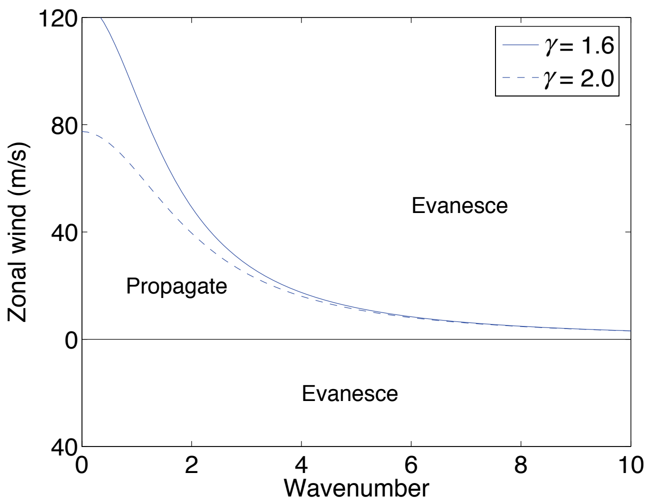

Figure 4.4 The boundary between propagating waves and evanescent waves as a function of zonal wind and wavenumber, using (4.52), for a couple of values of l (labelled γ here).

Stationary, vertically oscillatory modes can exist only for zonal flows that are eastwards and that are less than the critical velocity $U_c = \beta/(k^2 + l^2)$. One way to interpret this condition is note that in a resting medium the Rossby wave frequency has a minimum value (and maximum absolute value), when $m = 0$, of

$$\omega = -\frac{\beta k}{k^2 + l^2}. \tag{4.53}$$

Note too that in a frame moving with speed U our Rossby waves (stationary in the Earth's frame) have frequency $-Uk$, and this is the forcing frequency arising from the now-moving bottom topography. Thus, (4.52) is equivalent to saying that for oscillatory waves to exist *the forcing frequency must lie within the frequency range of vertically propagating Rossby waves.*

For westward flow, or for sufficiently strong eastward flow, the waves decay exponentially as $\Phi = \Phi_0 \exp(-\alpha z)$ where

$$\alpha = \frac{N}{f_0}\left(k^2 + l^2 - \frac{\beta}{U}\right)^{1/2}. \tag{4.54}$$

Note that the critical velocity $u_c = (\beta/k^2 + l^2)$ is a function of wavenumber, and that it increases with horizontal wavelength. Thus, for a given eastward flow long waves may penetrate vertically when short waves are trapped, an effect sometimes referred to as 'Charney–Drazin filtering'.[9] One important consequence of this is that the stratospheric motion is typically of larger scales than that of the troposphere, because Rossby waves tend to be excited first in the troposphere (by baroclinic instability and by flow over topography, among other things), but the shorter waves are trapped and only the longer ones reach the stratosphere. In the summer, the stratospheric winds are often westwards (because the pole is warmer than the equator) and all waves are trapped in the troposphere; the eastward stratospheric winds that favour vertical penetration occur in the other three seasons, although very strong eastward winds can suppress penetration in mid-winter.

4.5.2 Dispersion relation and group velocity

The dispersion relation for three-dimensional Rossby waves is again

$$\omega = Uk - \frac{\beta k}{K^2 + \gamma^2 + m^2 f_0^2/N^2}. \tag{4.55}$$

The three components of the group velocity for these waves are then:

$$c_g^x = U + \frac{\beta[k^2 - (l^2 + m^2 f_0^2/N^2)]}{\left(K^2 + m^2 f_0^2/N^2\right)^2}, \tag{4.56a}$$

$$c_g^y = \frac{2\beta kl}{\left(K^2 + m^2 f_0^2/N^2\right)^2}, \quad c_g^z = \frac{2\beta km f_0^2/N^2}{\left(K^2 + m^2 f_0^2/N^2\right)^2}. \tag{4.56b,c}$$

The propagation in the horizontal is analogous to the propagation in a shallow water model [c.f. (4.31b)]; note also that higher baroclinic modes (bigger m) will have a more westward group velocity. The vertical group velocity is proportional to m, and for waves that propagate signals upward we must choose m to have the same sign as k so that c_g^z is positive. If there is no mean flow then the zonal wavenumber k is negative (in order that frequency is positive) and m must then also be negative. Energy then propagates upward but the phase propagates downward.

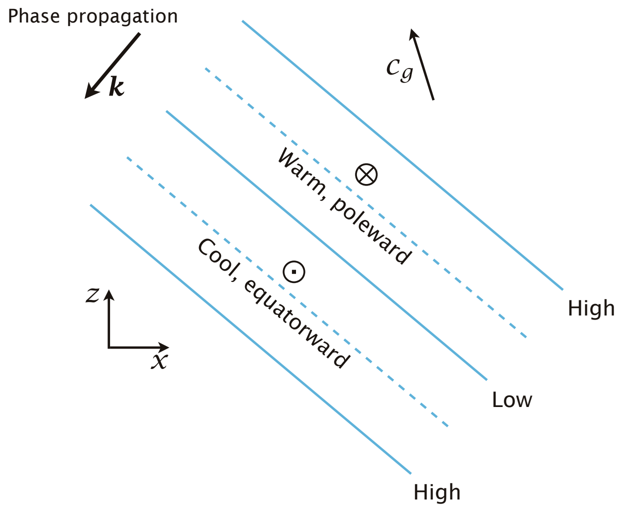

Figure 4.5 A schematic east-west section of an upwardly propagating Rossby wave. The slanting lines are lines of constant phase and 'high' and 'low' refer to the pressure or streamfunction values. Both k and m are negative so the phase lines are oriented up and to the west. The phase propagates westward and downward, but the group velocity is upward.

4.5.3 Vertical wave propagation and heat transport

If the group velocity in the z-direction, given by (4.56) is to be positive, then we require the product $km > 0$. This has consequences for the heat transport.

Remember that the buoyancy b, which is a proxy for temperature, is given by $f_0 \partial\psi/\partial z$. And the northward velocity is $v = \partial\psi/\partial x$. Thus, the northward flux of heat, H say, is given by

$$H = \overline{vb} = f_0 \overline{\frac{\partial\psi}{\partial z}\frac{\partial\psi}{\partial x}}, \tag{4.57}$$

where an overbar denotes a zonal average. Thus

$$H = \overline{vb} = f_0 \overline{\frac{\partial\psi}{\partial z}\frac{\partial\psi}{\partial x}} = f_0 \overline{\mathrm{Re}\,\widetilde{\psi} \mathrm{i} m \exp(\mathrm{i}\theta)\,\mathrm{Re}\,\widetilde{\psi} \mathrm{i} k \exp(\mathrm{i}\theta)} \tag{4.58}$$

where $\theta = (kx + ly + mz)$. Following manipulations exactly analogous to those given in the appendix, we find

$$H = \frac{1}{2} f_0 |\widetilde{\psi}|^2 km. \tag{4.59}$$

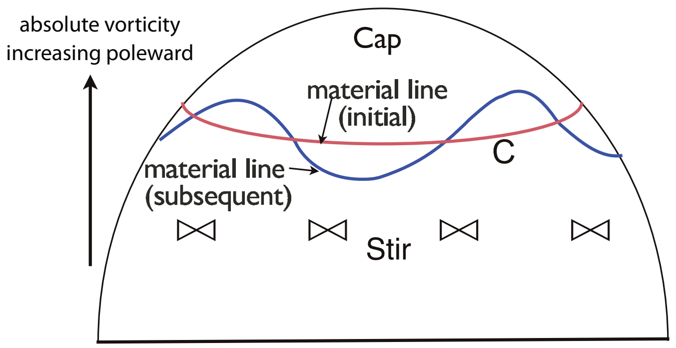

Figure 4.6 The effects of a mid-latitude disturbance on the circulation around the latitude line C. If initially the absolute vorticity increases monotonically polewards, then the disturbance will bring fluid with lower absolute vorticity into the cap region. Then, using Stokes theorem, the velocity around the latitude line C will become more westwards.

The conclusion is that vertical propagation of Rossby waves is associated with a polewards heat flux.

4.6 Rossby Waves and Jets

4.6.1 I. The vorticity budget

Suppose that the absolute vorticity normal to the surface (i.e., $\zeta + 2\Omega \sin \vartheta$) increases monotonically polewards. (A sufficient condition for this is that the fluid is at rest.) By Stokes' theorem, the circulation around a line of latitude circumscribing the polar cap, I, is equal to the integral of the absolute vorticity over the cap. That is,

$$I_i = \int_{\text{cap}} \boldsymbol{\omega}_{ia} \cdot \mathrm{d}\boldsymbol{A} = \oint_C u_{ia}\, \mathrm{d}l = \oint_C (u_i + \Omega a \cos \vartheta)\, \mathrm{d}l, \quad (4.60)$$

where $\boldsymbol{\omega}_{ia}$ and u_{ia} are the initial absolute vorticity and velocity, respectively, u_i is the initial zonal velocity in the Earth's frame of reference, and the line integrals are around the line of latitude. For simplicity let us take $u_i = 0$ and suppose there is a disturbance equatorwards of the polar cap, and that this results in a distortion of the material line around the latitude circle C (Fig. 4.6). Since we are supposing the source of the disturbance to be distant from the latitude of interest, then if we neglect viscosity the circulation along the material line is conserved, by Kelvin's circulation theorem. Thus, vorticity with a lower value is brought into the region of the polar cap — that is, the region polewards of the latitude line C. Using Stokes' theorem again the circulation around the latitude circle C must therefore fall; that is, denoting values after the disturbance with a subscript f,

$$I_f = \int_{\text{cap}} \boldsymbol{\omega}_{fa} \cdot \mathrm{d}\boldsymbol{A} < I_i \tag{4.61}$$

so that

$$\oint_C (u_f + \Omega a \cos\vartheta)\,\mathrm{d}l < \oint_C (u_i + \Omega a \cos\vartheta)\,\mathrm{d}l \tag{4.62}$$

and

$$\overline{u}_f < \overline{u}_i \tag{4.63}$$

with the overbar indicating a zonal average. Thus, there is a tendency to produce *westward* flow polewards of the disturbance. By a similar argument westward flow is also produced equatorwards of the disturbance — to see this one might apply Kelvin's theorem over all of the globe south of the source of the disturbance (taking care to take the dot-product correctly between the direction of the vorticity vector and the direction normal to the surface). Finally, note that the overall situation is the same in the Southern Hemisphere. Thus, on the surface of a rotating sphere, external stirring will produce westward flow *away* from the region of the stirring.

Now suppose, furthermore, that the disturbance imparts no net angular momentum to the fluid. Then the integral of $ua\cos\vartheta$ over the entire hemisphere must be constant. But the fluid is accelerating westwards away from the disturbance. Therefore, the fluid in the region of the disturbance must accelerate *eastwards;* that is, angular momentum must converge into the stirred region, producing an eastward flow. This simple mechanism is the essence of the production of eastward eddy-driven jets in the atmosphere, and of the eastward surface winds in mid-latitudes. The stirring that here we have externally imposed comes, in reality, from baroclinic instability.

If the stirring subsides then the flow may reversibly go back to its initial condition, with a concomitant reversal of the momentum convergence that caused the zonal

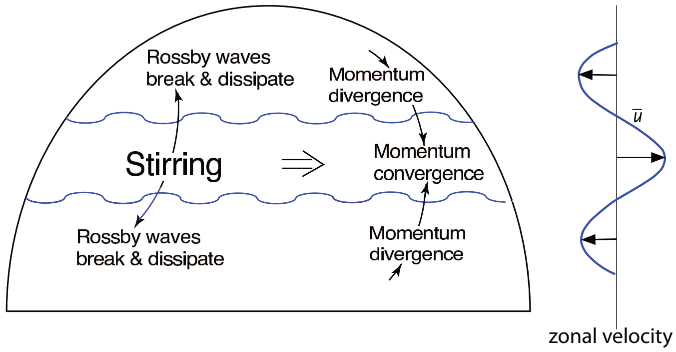

Figure 4.7 Generation of zonal flow on a β-plane or on a rotating sphere. Stirring in mid-latitudes (by baroclinic eddies) generates Rossby waves that propagate away from the disturbance. Momentum converges in the region of stirring, producing eastward flow there and weaker westward flow on its flanks.

flow. Thus, we must have some form of dissipation and irreversibility in order to produce permanent changes, and in particular we need to irreversibly mix vorticity. If the fluid is continuously mixed, then of course we also need a source that restores the absolute vorticity gradient, otherwise we will completely homogenize the vorticity over the hemisphere, so let us now set up a simple model that shows how a permanent jet structure can be maintained.

4.6.2 II. Rossby waves and momentum flux

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

If a stirring is present in mid-latitudes then we expect that Rossby waves will be generated there, propagate away and break and dissipate. To the extent that the waves are quasi-linear and do not interact, then just away from the source region each wave has the form

$$\psi = \mathrm{Re}\, C \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)} = \mathrm{Re}\, C \mathrm{e}^{\mathrm{i}(kx+ly-kct)}, \qquad (4.64)$$

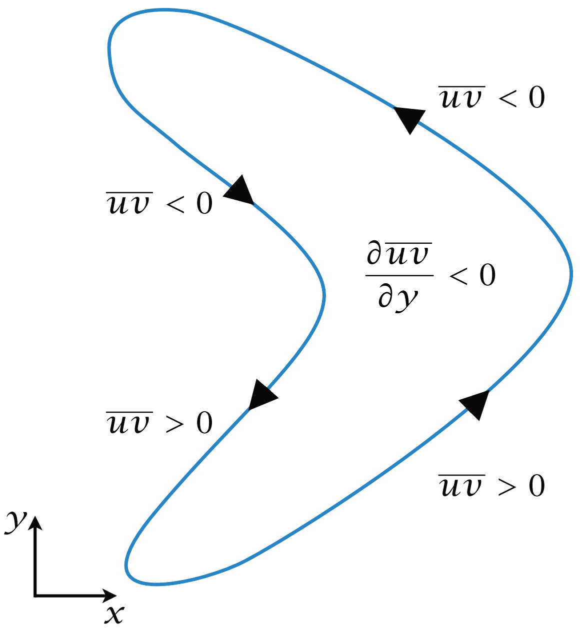

Figure 4.8 The momentum transport in physical space, caused by the propagation of Rossby waves away from a source in mid-latitudes. The ensuing bow-shaped eddies are responsible for a convergence of momentum, as indicated in the idealization pictured.

where C is a constant, with dispersion relation

$$\omega = ck = Uk - \frac{\beta k}{k^2 + l^2} \equiv \omega_R, \tag{4.65}$$

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

$$c_g^y = \frac{\partial \omega}{\partial l} = \frac{2\beta k l}{(k^2 + l^2)^2}. \tag{4.66}$$

Now, the direction of the group velocity must be *away* from the source region; this is a radiation condition (discussed more in the next subsection), demanded by the requirement that Rossby waves transport energy *away* from the disturbance. Thus, northwards of the source kl is positive and southwards of the source kl is negative. That the product kl can be positive or negative arises because for each k there are two possible values of l that satisfy the dispersion relation (4.65), namely

$$l = \pm \left(\frac{\beta}{U - c} - k^2 \right)^{1/2}, \tag{4.67}$$

assuming that the quantity in parentheses is positive.

The velocity variations associated with the Rossby waves are

$$u' = -\mathrm{Re}\, C\, \mathrm{i} l \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \qquad v' = \mathrm{Re}\, C\, \mathrm{i} k \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \tag{4.68a,b}$$

and the associated momentum flux is (see appendix for algebraic details)

$$\overline{u'v'} = -\frac{1}{2} C^2 kl. \tag{4.69}$$

Thus, given that the sign of kl is determined by the group velocity, northwards of the source the momentum flux associated with the Rossby waves is southward (i.e., $\overline{u'v'}$ is negative), and southwards of the source the momentum flux is northward (i.e., $\overline{u'v'}$ is positive). That is, the momentum flux associated with the Rossby waves is *toward* the source region. Momentum converges in the region of the stirring, producing net eastward flow there and westward flow to either side (Fig. 4.7).

Another way of describing the same effect is to note that if kl is positive then lines of constant phase ($kx + ly =$ constant) are tilted north-west/south-east, and the momentum flux associated with such a disturbance is negative ($\overline{u'v'} < 0$). Similarly, if kl is negative then the constant-phase lines are tilted north-east/south-west and the associated momentum flux is positive ($\overline{u'v'} > 0$). The net result is a convergence of momentum flux into the source region. In physical space this is reflected by having eddies that are 'bow-shaped', as in Fig. 4.8.

Appendix: Calculation of Fluxes

In two places in this chapter we had to calculate the average flux of a quantity and in this appendix we do that explicitly in the case of the northward flux of momentum in a Rossby wave. The same method can be used to calculate the vertical flux of buoyancy in a Rossby wave. It is important to take the real part of each expression before taking the average. To proceed, let

$$\psi = \mathrm{Re}\, A e^{\mathrm{i}(kx+ly-\omega t)} \tag{4.70}$$

where $A = a + \mathrm{i}b$. The velocities are given by

$$u = -\frac{\partial \psi}{\partial y}, \qquad v = \frac{\partial \psi}{\partial x}. \tag{4.71}$$

Thus,

$$u = -\mathrm{Re}\, \mathrm{i}lA e^{\mathrm{i}\theta} = al \sin\theta + bl \cos\theta \tag{4.72}$$

and

$$v = +\mathrm{Re}\, \mathrm{i}kA e^{\mathrm{i}\theta} = -ak \sin\theta - bk \cos\theta \tag{4.73}$$

where $\theta = kx + ly - \omega t$. The northwards momentum flux is then

$$\overline{uv} = \frac{1}{L}\int_0^L uv \, \mathrm{d}x \tag{4.74}$$

where L is a wavelength or a multiple of wavelengths. Now, a standard result is that

$$\frac{1}{L}\int_0^L \sin^2 kx \, \mathrm{d}x = \frac{1}{L}\int_0^L \cos^2 kx \, \mathrm{d}x = \frac{1}{2}, \tag{4.75}$$

and

$$\frac{1}{L}\int_0^L \sin kx \cos kx \, \mathrm{d}x = 0. \tag{4.76}$$

Thus,

$$\begin{aligned} \overline{uv} &= \frac{1}{L}\int_0^L (al \sin\theta + bl \cos\theta) \times (-ak \sin\theta - bk \cos\theta) \\ &= -\frac{kl}{2}(a^2 + b^2) = -\frac{1}{2}|A|^2 kl \end{aligned} \tag{4.77}$$

Thus, the poleward flux of momentum is proportional to $-kl$.

A similar methodology applies when calculating the poleward flux of buoyancy, $\overline{vb}$. Since $v = \partial\psi/\partial x = \mathrm{Re}\, \mathrm{i}kA \exp(\mathrm{i}\theta)$

and $b = f_0 \partial\psi/\partial z = \mathrm{Re}\, \mathrm{i} f_0 m A \exp(\mathrm{i}\theta)$ then by the same technique we find, skipping some algebra,

$$\begin{aligned}\overline{vb} &= \frac{f_0}{L}\int_0^L (-ak\sin\theta - bk\cos\theta)\times(-am\sin\theta - bm\cos\theta) \\ &= \frac{f_0 km}{2}(a^2+b^2) = \frac{f_0}{2}|A|^2 km\end{aligned} \tag{4.78}$$

and is proportional to $+km$.

Chapter 5

Ekman Layers and Ocean Gyre

Weeks 9 to 11

5.1 Ekman Layers

The fluid fields in the interior of a domain are often set by different physical processes than those occurring at a boundary, and consequently often change rapidly in a thin *boundary layer*, as in Fig. 5.1. Such boundary layers nearly always involve one or both of viscosity and diffusion, because these appear in the terms of highest differential order in the equations of motion, and so are responsible for the number and type of boundary conditions that the equations must satisfy — for example, the presence of molecular viscosity leads to the condition that the tangential flow (as well as the normal flow) must vanish at a rigid surface. In many boundary layers in non-rotating flow the dominant balance in the momentum equation is between the advective and viscous terms. In large-scale atmospheric and oceanic flow the effects of rotation are large and the dominant balance is between Coriolis and frictional or stress terms.

The atmospheric Ekman layer occurs near the ground, and the stress at the ground itself is due to the surface wind (and its vertical variation). In the ocean the main Ekman layer is near the surface, and the stress at ocean surface is largely due to the presence of the overlying wind. There

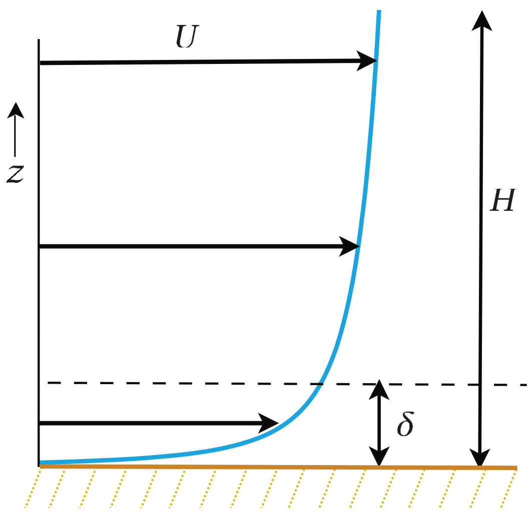

Figure 5.1 An idealized boundary layer. The values of a field, such as velocity, U, may vary rapidly in a boundary in order to satisfy the boundary conditions at a rigid surface. The parameter δ is a measure of the boundary layer thickness, H is a typical scale of variation away from the boundary, and typically a boundary layer has $\delta \ll H$.

is also a weak Ekman layer at the bottom of the ocean, analogous to the atmospheric Ekman layer. To analyze all these layers we assume:

- The Ekman layer is Boussinesq.
- The Ekman layer has a finite depth that is less than the total depth of the fluid, this depth being given by the level at which the frictional stresses essentially vanish. Within the Ekman layer, frictional terms are important, whereas geostrophic balance holds beyond it.
- The nonlinear and time-dependent terms in the equations of motion are negligible, hydrostatic balance holds in the vertical, and buoyancy is constant, not varying in the horizontal.
- Friction can be parameterized by a viscous term of the form $\rho_0^{-1}\partial\boldsymbol{\tau}/\partial z = A\,\partial^2\boldsymbol{u}/\partial z^2$, where A is constant and $\boldsymbol{\tau}$ is the stress. [In general, stress is a tensor, τ_{ij}, with an associated force given by $F_i = \partial\tau_{ij}/\partial x_j$, sum-

ming over the repeated index. It is common in geophysical fluid dynamics that the vertical derivative dominates, and in this case the force is $\boldsymbol{F} = \partial\boldsymbol{\tau}/\partial z$. We still use the word stress for $\boldsymbol{\tau}$, but it now refers to a vector whose derivative in a particular direction (z in this case) is the force on a fluid.] In laboratory settings A may be the molecular viscosity, whereas in the atmosphere and ocean it is a so-called *eddy viscosity.*

5.1.1 Equations of motion and scaling

Frictional–geostrophic balance in the horizontal momentum equation is:

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla_z \phi + \frac{\partial \widetilde{\boldsymbol{\tau}}}{\partial z}, \tag{5.1}$$

where $\widetilde{\boldsymbol{\tau}} \equiv \boldsymbol{\tau}/\rho_0$ is the kinematic stress and $\boldsymbol{f} = f\mathbf{k}$, where the Coriolis parameter f is allowed to vary with latitude. If we model the stress with an eddy viscosity, (5.1) becomes

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla_z \phi + A\frac{\partial^2 \boldsymbol{u}}{\partial z^2}. \tag{5.2}$$

The vertical momentum equation is $\partial\phi/\partial z = b$, i.e., hydrostatic balance, and, because buoyancy is constant, we may without loss of generality write this as

$$\frac{\partial \phi}{\partial z} = 0. \tag{5.3}$$

The equation set is completed by the mass continuity equation, $\nabla \cdot \boldsymbol{v} = 0$.

The Ekman number

We non-dimensionalize the equations by setting

$$(u, v) = U(\widehat{u}, \widehat{v}), \quad (x, y) = L(\widehat{x}, \widehat{y}), \quad f = f_0 \widehat{f}, \quad z = H\widehat{z}, \quad \phi = \Phi\widehat{\phi}, \tag{5.4}$$

where hatted variables are non-dimensional. H is a scaling for the height, and at this stage we will suppose it to be some height scale in the free atmosphere or ocean, not the height of the Ekman layer itself. Geostrophic balance suggests that $\Phi = f_0 UL$. Substituting (5.4) into (5.2) we obtain

$$\widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}} = -\widehat{\nabla}\widehat{\phi} + Ek\frac{\partial^2 \widehat{\boldsymbol{u}}}{\partial \widehat{z}^2}, \tag{5.5}$$

where the parameter

$$Ek \equiv \left(\frac{A}{f_0 H^2}\right), \tag{5.6}$$

is the *Ekman number*, and it determines the importance of frictional terms in the horizontal momentum equation. If $Ek \ll 1$ then the friction is small in the flow interior where $\widehat{z} = \mathcal{O}(1)$. However, the friction term cannot necessarily be neglected in the boundary layer because it is of the highest differential order in the equation, and so determines the boundary conditions; if Ek is small the vertical scales become small and the second term on the right-hand side of (5.5) remains finite. The case when this term is simply omitted from the equation is therefore a *singular limit,* meaning that it differs from the case with $Ek \to 0$. If $Ek \geq 1$ friction is important everywhere, but it is usually the case that Ek is small for atmospheric and oceanic large-scale flow, and the interior flow is very nearly geostrophic. (In part this is because A itself is only large near a rigid surface where the presence of a shear creates turbulence and a significant eddy viscosity.)

Momentum balance in the Ekman layer

For definiteness, suppose the fluid lies above a rigid surface at $z = 0$. Sufficiently far away from the boundary the velocity field is known, and we suppose this flow to be in geostrophic balance. We then write the velocity field and the pressure field as the sum of the interior geostrophic part, plus a boundary layer correction:

$$\widehat{\boldsymbol{u}} = \widehat{\boldsymbol{u}}_g + \widehat{\boldsymbol{u}}_E, \qquad \widehat{\phi} = \widehat{\phi}_g + \widehat{\phi}_E, \tag{5.7}$$

where the Ekman layer corrections, denoted with a subscript E, are negligible away from the boundary layer. Now, in the fluid interior we have, by hydrostatic balance, $\partial\widehat{\phi}_g/\partial\widehat{z} = 0$. In the boundary layer we still have $\partial\widehat{\phi}_g/\partial\widehat{z} = 0$ so that, to satisfy hydrostasy, $\partial\widehat{\phi}_E/\partial\widehat{z} = 0$. But because $\widehat{\phi}_E$ vanishes away from the boundary we have $\widehat{\phi}_E = 0$ everywhere. Thus, *there is no boundary layer in the pressure field.* Note that this is a much stronger result than saying that pressure is continuous, which is nearly always true in fluids; rather, it is a special result for Ekman layers.

Using (5.7) with $\widehat{\phi}_E = 0$, the dimensional horizontal momentum equation (5.1) becomes, in the Ekman layer,

$$\boldsymbol{f} \times \boldsymbol{u}_E = \frac{\partial \widetilde{\boldsymbol{\tau}}}{\partial z}. \tag{5.8}$$

The dominant force balance in the Ekman layer is thus between the Coriolis force and the friction. We can determine the thickness of the Ekman layer if we model the stress with an eddy viscosity so that

$$\boldsymbol{f} \times \boldsymbol{u}_E = A \frac{\partial^2 \boldsymbol{u}_E}{\partial z^2}, \tag{5.9}$$

or, non-dimensionally,

$$\widehat{\boldsymbol{f}} \times \widehat{\boldsymbol{u}}_E = Ek \frac{\partial^2 \widehat{\boldsymbol{u}}_E}{\partial \widehat{z}^2}. \tag{5.10}$$

It is evident this equation can only be satisfied if $\widehat{z} \neq \mathcal{O}(1)$, implying that H is not a proper scaling for z in the boundary layer. Rather, if the vertical scale in the Ekman layer is $\widehat{\delta}$ (meaning $\widehat{z} \sim \widehat{\delta}$) we must have $\widehat{\delta} \sim Ek^{1/2}$. In dimensional terms this means the thickness of the Ekman layer is

$$\delta = H\widehat{\delta} = HEk^{1/2} \tag{5.11}$$

or

$$\delta = \left(\frac{A}{f_0}\right)^{1/2}. \tag{5.12}$$

[This estimate also emerges directly from (5.9).] Note that (5.11) can be written as

$$Ek = \left(\frac{\delta}{H}\right)^2. \tag{5.13}$$

That is, the Ekman number is equal to the square of the ratio of the depth of the Ekman layer to an interior depth scale of the fluid motion. In laboratory flows where A is the molecular viscosity we can thus estimate the Ekman layer thickness, and if we know the eddy viscosity of the ocean or atmosphere we can estimate their respective Ekman layer thicknesses. We can invert this argument and obtain an estimate of A if we know the Ekman layer depth. In the atmosphere, deviations from geostrophic balance are very small in the atmosphere above 1 km, and using this gives $A \approx 10^2\,\mathrm{m^2\,s^{-1}}$. In the ocean Ekman depths are often 50 m or less, and eddy viscosities are about $0.1\,\mathrm{m^2\,s^{-1}}$.

5.1.2 Integral properties of the Ekman layer

What can we deduce about the Ekman layer without specifying the detailed form of the frictional term? Using dimensional notation we recall frictional–geostrophic balance,

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi + \frac{1}{\rho_0}\frac{\partial \boldsymbol{\tau}}{\partial z}, \tag{5.14}$$

where $\boldsymbol{\tau}$ is zero at the edge of the Ekman layer. In the Ekman layer itself we have

$$\boldsymbol{f} \times \boldsymbol{u}_E = \frac{1}{\rho_0}\frac{\partial \boldsymbol{\tau}}{\partial z}. \tag{5.15}$$

Consider either a top or bottom Ekman layer, and integrate over its thickness. From (5.15) we obtain

$$\boldsymbol{f} \times \boldsymbol{M}_E = \boldsymbol{\tau}_T - \boldsymbol{\tau}_B, \tag{5.16}$$

where

$$\boldsymbol{M}_E = \int_{Ek} \rho_0 \boldsymbol{u}_E \,\mathrm{d}z \tag{5.17}$$

is the ageostrophic mass transport in the Ekman layer, and $\boldsymbol{\tau}_T$ and $\boldsymbol{\tau}_B$ are the respective stresses at the top and the bottom of the Ekman layer at hand. The stress at the top (bottom) will be zero in a bottom (top) Ekman layer and therefore, from (5.16),

$$\begin{aligned} &\text{top Ekman layer:} && \boldsymbol{M}_E = -\frac{1}{f}\mathbf{k} \times \boldsymbol{\tau}_T \\ &\text{bottom Ekman layer:} && \boldsymbol{M}_E = \frac{1}{f}\mathbf{k} \times \boldsymbol{\tau}_B \end{aligned} . \tag{5.18a,b}$$

The transport is thus at right angles to the stress at the surface, and proportional to the magnitude of the stress. These properties have a simple physical explanation: integrated over the depth of the Ekman layer the surface stress must be balanced by the Coriolis force, which in turn acts at right angles to the mass transport. A consequence of (5.18) is that the mass transports in adjacent oceanic and atmospheric Ekman layers are equal and opposite, because the stress is continuous across the ocean–atmosphere interface. Equation (5.18a) is particularly useful in the ocean, where the stress at the surface is primarily due to the wind, and is largely independent of the interior oceanic flow. In the atmosphere, the surface stress mainly arises as a result of the interior atmospheric flow, and to calculate it we need to parameterize the stress in terms of the flow.

Finally, we obtain an expression for the vertical velocity induced by an Ekman layer. The mass conservation equation is

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \tag{5.19}$$

Integrating this over an Ekman layer gives

$$\frac{1}{\rho_0}\nabla\cdot \boldsymbol{M}_T = -(w_T - w_B), \tag{5.20}$$

where $\boldsymbol{M}_T$ is the total (Ekman plus geostrophic) mass transport in the Ekman layer,

$$\boldsymbol{M}_T = \int_{Ek} \rho_0 \boldsymbol{u}\,\mathrm{d}z = \int_{Ek} \rho_0(\boldsymbol{u}_g + \boldsymbol{u}_E)\,\mathrm{d}z \equiv \boldsymbol{M}_g + \boldsymbol{M}_E, \tag{5.21}$$

and w_T and w_B are the vertical velocities at the top and bottom of the Ekman layer; the former (latter) is zero in a top (bottom) Ekman layer. Equations (5.21) and (5.16) give

$$\mathbf{k}\times(\boldsymbol{M}_T - \boldsymbol{M}_g) = \frac{1}{f}(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B). \tag{5.22}$$

Taking the curl of this (i.e., cross-differentiating) gives

$$\nabla\cdot(\boldsymbol{M}_T - \boldsymbol{M}_g) = \mathrm{curl}_z[(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B)/f], \tag{5.23}$$

where the curl_z operator on a vector $\boldsymbol{A}$ is defined by $\mathrm{curl}_z\boldsymbol{A} \equiv \partial_x A_y - \partial_y A_x$. Using (5.20) we obtain, for top and bottom

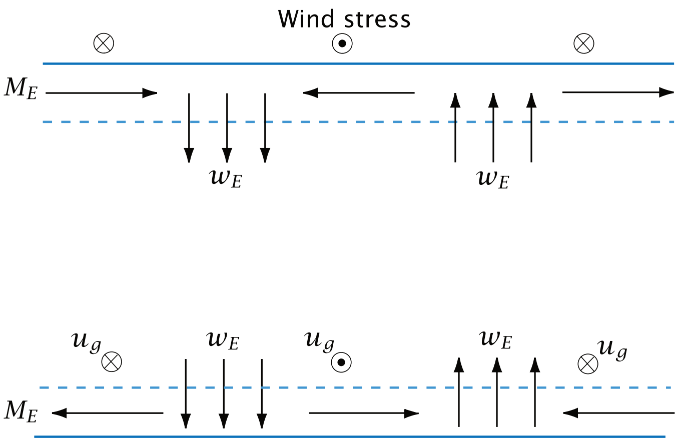

Figure 5.2 Upper and lower Ekman layers. The upper Ekman layer in the ocean is primarily driven by an imposed wind stress, whereas the lower Ekman layer in the atmosphere or ocean largely results from the interaction of interior geostrophic velocity and a rigid lower surface. The upper part of figure shows the vertical Ekman 'pumping' velocities that result from the given wind stress, and the lower part of the figure shows the Ekman pumping velocities given the interior geostrophic flow.

Ekman layers respectively,

$$w_B = \frac{1}{\rho_0}\left(\mathrm{curl}_z \frac{\boldsymbol{\tau}_T}{f} + \nabla\cdot \boldsymbol{M}_g\right), \quad w_T = \frac{1}{\rho_0}\left(\mathrm{curl}_z \frac{\boldsymbol{\tau}_B}{f} - \nabla\cdot \boldsymbol{M}_g\right), \tag{5.24a,b}$$

where $\nabla\cdot \boldsymbol{M}_g = -(\beta/f)\boldsymbol{M}_g \cdot \mathbf{j}$ is the divergence of the geostrophic transport in the Ekman layer, and this is often small compared to the other terms in these equations. Thus, friction induces a vertical velocity at the edge of the Ekman layer, proportional to the curl of the stress at the surface, and this is perhaps the most used result in Ekman layer theory. Numerical models sometimes do not have the vertical resolution to explicitly resolve an Ekman layer, and (5.24) provides a means of *parameterizing* the Ekman layer in terms of resolved or known fields. It is particularly useful for the top Ekman layer in the ocean, where the stress can be regarded as a given function of the overlying wind.

5.1.3 Sverdrup Balance

In this section we rederive the above results in a slightly more direct way, and also obtain a result for the total transport induced by a windstress. To this end, consider an ocean forced by a windstress at the top that satisfies the Ekman-layer equations

$$-fv = -\frac{\partial \phi}{\partial x} + \frac{\partial \tilde{\tau}_x}{\partial z}, \qquad fu = -\frac{\partial \phi}{\partial y} + \frac{\partial \tilde{\tau}_y}{\partial z}. \tag{5.25}$$

where $\tilde{\tau} = \tau/\rho_0$. Equivalently we have

$$f(v_g - v) = \frac{\partial \tilde{\tau}_x}{\partial z}, \qquad f(u - u_g) = \frac{\partial \tilde{\tau}_y}{\partial z}. \tag{5.26}$$

We note that the geostrophic velocity field satisfies,

$$f\left(\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y}\right) = -\beta v_g. \tag{5.27}$$

If we integrate the mass continuity equation over the depth of the Ekman layer, the vertical velocity at its base is given by

$$w_E = \int_{-H_E}^{0} \left(\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y} + \frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y}\right). \tag{5.28}$$

The divergence of the geostrophic velocity is given by (5.27), and that of the ageostrophic velocity is obtained from (5.26). We thus obtain

$$w_E = \left[\frac{\partial}{\partial x}\left(\frac{\tilde{\tau}_{y0}}{f}\right) - \frac{\partial}{\partial y}\left(\frac{\tilde{\tau}_{x0}}{f}\right)\right] - \int_{-H_E}^{0} \frac{\beta}{f} v_g \, dz, \tag{5.29}$$

where $\tilde{\tau}_{x0}, \tilde{\tau}_{y0}$ are the components of the stress at the surface. This equation is essentially the same as (5.24a).

If we go back to (5.25), cross differentiate and integrate from the top down we obtain an expression for the vertical velocity at the base of the Ekman layer in terms of the stress and the *total* velocity,

$$w_E = \frac{1}{f}\left[\frac{\partial \tilde{\tau}_{y0}}{\partial x} - \frac{\partial \tilde{\tau}_{x0}}{\partial y}\right] - \int_{-H_E}^{0} \frac{\beta}{f} v \, dz. \tag{5.30}$$

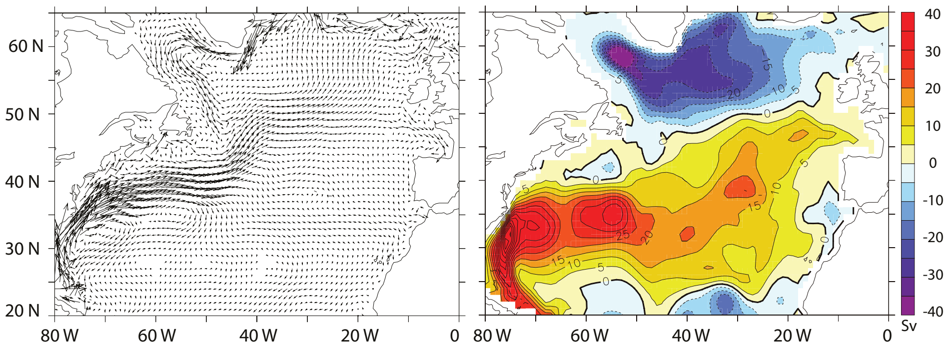

Figure 5.3 Left: the time averaged velocity field at a depth of 75 m in the North Atlantic. Right: the streamfunction of the vertically integrated flow, in Sverdrups (1 Sv = 10^9 kg s^{-1}). Note the presence of an anticyclonic subtropical gyre (clockwise circulation, shaded red), a cyclonic subpolar gyre (anticlockwise, blue), and intense western boundary currents.

If we let the integral go over the entire depth of the ocean, and assume that the vertical velocity is zero at the bottom, we obtain

$$\int \beta v \, \mathrm{d}z = \frac{\partial \widetilde{\tau}_{y0}}{\partial x} - \frac{\partial \widetilde{\tau}_{x0}}{\partial y}. \tag{5.31}$$

This is known as the Sverdrup relation, and is a relation between the stress at the surface and the total meridional transport in the ocean.

5.2 Ocean Gyres

5.3 The Depth Integrated Wind-Driven Circulation

The large-scale mean currents shown in Fig. 5.3 and in Fig. 5.4, where we see subtropical and subpolar gyres, all of them intensified in the west. Our goal is to explain the main features seen in these figures in as simple and straightforward a manner as is possible.

The equations that govern the large-scale flow in the oceans are the planetary-geostrophic equations, but these equations are still quite daunting: a prognostic equation for buoyancy is coupled to the advecting velocity via hydrostatic and geostrophic balance, and the resulting problem is formidably nonlinear. However, it turns out that thermodynamic effects can effectively be eliminated by the simple device of vertical integration; the resulting equations are linear, and the only external forcing is that due to the wind stress.

5.3.1 The Stommel Model

The planetary-geostrophic equations for a Boussinesq fluid are:

$$\frac{\mathrm{D}b}{\mathrm{D}t} = \dot{b}, \qquad \nabla_3 \cdot \boldsymbol{v} = 0, \tag{5.32a,b}$$

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi + \frac{1}{\rho_0}\frac{\partial \boldsymbol{\tau}}{\partial z}, \qquad \frac{\partial \phi}{\partial z} = b. \tag{5.33a,b}$$

These equations are, respectively, the thermodynamic equation (5.32a), the mass continuity equation (5.32b), the horizontal momentum equation (5.33a), (i.e., geostrophic balance, plus a stress term), and the vertical momentum equation (5.33b) — that is, hydrostatic balance. These equations are derived more fully in Chapter 3, but they are essentially the Boussinesq primitive equations with the advection terms omitted from the horizontal momentum equation, on the basis of small Rossby number. In this chapter we will henceforth absorb the factor of ρ_0 into the $\boldsymbol{\tau}$, so that $\boldsymbol{\tau}$ denotes the kinematic stress, and the gradient operator will be two dimensional, in the x-y plane, unless noted.

Take the curl of (5.33a) (that is, cross differentiate its x and y components) and integrate over the depth of the ocean to give

$$\int \boldsymbol{f}\nabla\cdot \boldsymbol{u}\,\mathrm{d}z + \frac{\partial f}{\partial y}\int v\,\mathrm{d}z = \mathrm{curl}_z(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B), \tag{5.34}$$

where the operator curl_z is defined by $\mathrm{curl}_z\boldsymbol{A} \equiv \partial A^y/\partial x - \partial A^x/\partial y = \mathbf{k}\cdot\nabla\times\boldsymbol{A}$, and the subscripts T and B are for top and bottom. The divergence term vanishes if the vertical velocity is zero at the top and bottom of the ocean. Strictly, at the top of the ocean the vertical velocity is given by the material derivative of height of the ocean's surface, $\mathrm{D}h/\mathrm{D}t$, but on the large-scales this has a negligible effect and we may make the rigid-lid approximation and set it to zero.

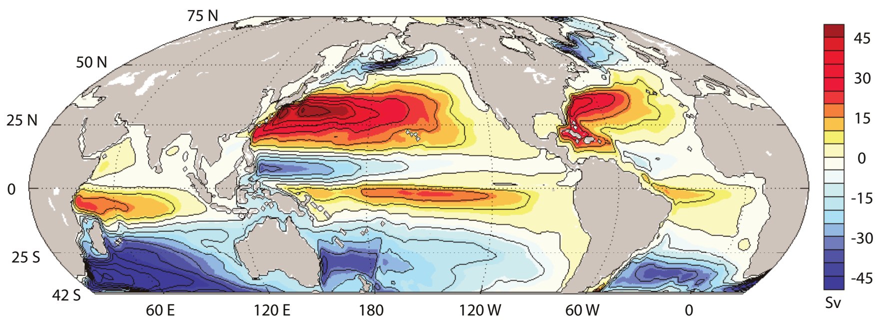

Figure 5.4 The streamfunction of the vertically integrated flow for the near global ocean. Red shading indicates clockwise flow, and blue shading anticlockwise, but in both hemispheres the subtropical (subpolar) gyres are anticyclonic (cyclonic).

At the bottom of the ocean the vertical velocity is only zero if the ocean is flat-bottomed; otherwise it is $\boldsymbol{u} \cdot \nabla \eta_B$, where η_B is the orographic height at the ocean floor. The neglect of this topographic term is probably the most restrictive single approximation in the model. Given this neglect, (5.34) becomes

$$\beta \overline{v} = \operatorname{curl}_z(\boldsymbol{\tau}_T - \boldsymbol{\tau}_B), \tag{5.35}$$

where henceforth, in this section, quantities with an overbar are understood to be the vertical integral over the depth of the ocean. If the stresses depend only on the velocity fields then thermodynamic fields do not affect the vertically integrated flow.

At the top of the ocean, the stress is given by the wind. At the bottom, in the absence of topography we assume that the stress may be parameterized by a linear drag, or Rayleigh friction, as might be generated by an Ekman layer; it is this assumption that particularly characterizes this model as being due to Stommel. Equation (5.35) then becomes

$$\beta \overline{v} = -r\overline{\zeta} + F_\tau(x, y), \tag{5.36}$$

where $F_\tau = \operatorname{curl}_z \boldsymbol{\tau}_T$ is the wind-stress curl at the top of the ocean and is a known function. Because the velocity is divergence-free, we can define a streamfunction ψ such that $\overline{u} = -\partial\psi/\partial y$ and $\overline{v} = \partial\psi/\partial x$. Equation (5.36) then becomes

$$r\nabla^2\psi + \beta\frac{\partial\psi}{\partial x} = F_\tau(x, y). \tag{5.37}$$

This equation is often referred to as the *Stommel problem* or the *Stommel model,* and may be posed in a variety of two dimensional domains.

5.3.2 Approximate Solution of Stommel Model

Sverdrup balance

Equation (5.37) is linear and it is possible to obtain an exact, analytic solution. However, it is more insightful to approach the problem perturbatively, by supposing that the frictional term is small, meaning there is an approximate balance between wind stress and the β-effect.[10] Friction is small if $|r\zeta| \ll |\beta v|$ or

$$\frac{r}{L} = \frac{f\delta_B}{HL} \ll \beta \qquad (5.38)$$

using $r = f\delta_B/H$, and where L is the horizontal scale of the motion, and generally speaking this inequality is well satisfied for large-scale flow. The vorticity equation becomes

$$\beta\overline{v} \approx \mathrm{curl}_z \boldsymbol{\tau}_T, \qquad (5.39)$$

which is known as *Sverdrup balance.*[11] (Sometimes Sverdrup balance is taken to mean the linear geostrophic vorticity balance $\beta v = f\partial w/\partial z$, but we will restrict its use to mean a balance between the beta effect and wind stress curl.) The observational support for Sverdrup balance is rather mixed, discrepancies arising not so much from the failure of (5.38), but from the presence of small-scale eddying motion with concomitantly large nonlinear terms, and the presence of non-negligible vertical velocities induced by the interaction with bottom topography.[12] Nevertheless, Sverdrup balance provides a useful, if not impregnable, foundation on which to build.

Boundary-layer solution

For simplicity, consider a square domain of side a and rescale the variables by setting

$$x = a\widehat{x}, \qquad y = a\widehat{y}, \qquad \tau = \tau_0\widehat{\tau}, \qquad \psi = \widehat{\psi}\frac{\tau_0}{\beta}, \quad (5.40)$$

where τ_0 is the amplitude of the wind stress. The hatted variables are nondimensional and, assuming our scaling to be sensible, these are $\mathcal{O}(1)$ quantities in the interior. Equation (??) becomes

$$\frac{\partial \widehat{\psi}}{\partial \widehat{x}} + \epsilon_S \nabla^2 \widehat{\psi} = \mathrm{curl}_z \widehat{\boldsymbol{\tau}}_T, \tag{5.41}$$

where $\epsilon_S = (r/a\beta) \ll 1$, in accord with (5.38). For the rest of this section we will drop the hats over nondimensional quantities. Over the interior of the domain, away from boundaries, the frictional term in (5.41) is small. We can take advantage of this by writing

$$\psi(x, y) = \psi_I(x, y) + \phi(x, y), \tag{5.42}$$

where ψ_I is the interior streamfunction and ϕ is a boundary layer correction. Away from boundaries ψ_I is presumed to dominate the flow, and this satisfies

$$\frac{\partial \psi_I}{\partial x} = \mathrm{curl}_z \boldsymbol{\tau}_T. \tag{5.43}$$

The solution of this equation (called the 'Sverdrup interior') is

$$\psi_I(x, y) = \int_0^x \mathrm{curl}_z \boldsymbol{\tau}(x', y)\, \mathrm{d}x' + g(y), \tag{5.44}$$

where $g(y)$ is an arbitrary function of integration that gives rise to an arbitrary zonal flow. The corresponding velocities are

$$v_I = \mathrm{curl}_z \boldsymbol{\tau}, \qquad u_I = -\frac{\partial}{\partial y} \int_0^x \mathrm{curl}_z \boldsymbol{\tau}(x', y)\, \mathrm{d}x' - \frac{\mathrm{d}g(y)}{\mathrm{d}y}. \tag{5.45}$$

The dynamics is most clearly illustrated if we now restrict our attention to a wind-stress curl that is zonally uniform, and that vanishes at two latitudes, $y = 0$ and $y = 1$. An example is

$$\tau_T^y = 0, \qquad \tau_T^x = -\cos(\pi y), \tag{5.46}$$

for which $\mathrm{curl}_z \boldsymbol{\tau}_T = -\pi \sin(\pi y)$. The Sverdrup (interior) flow may then be written as

$$\psi_I(x, y) = [x - C(y)]\mathrm{curl}_z \boldsymbol{\tau}_T = \pi[C(y) - x] \sin \pi y, \tag{5.47}$$

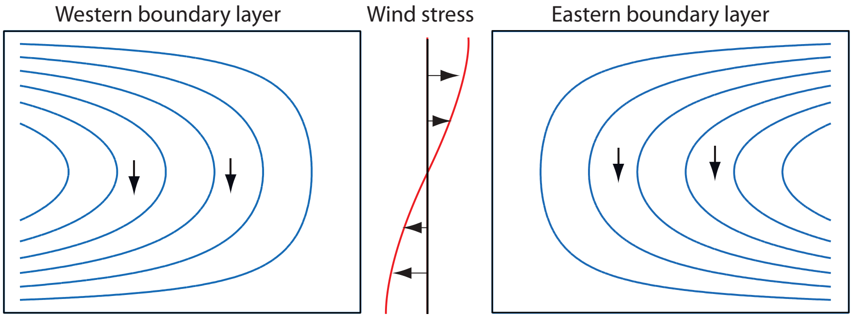

Figure 5.5 Two possible Sverdrup flows, ψ_I, for the wind stress shown in the centre. Each solution satisfies the no-flow condition at either the eastern or western boundary, and a boundary layer is therefore required at the other boundary. Both flows have the same, equatorward, meridional flow in the interior. Only the flow with the western boundary current is physically realizable, however, because only then can friction produce a curl that opposes that of the wind stress, so allowing the flow to equilibrate.

where $C(y)$ is the arbitrary function of integration $[C(y) = -g(y)/\mathrm{curl}_z\boldsymbol{\tau}]$. If we choose C to be a constant, the zonal flow associated with it is $C\,\mathrm{curl}_z\boldsymbol{\tau}_T$. We can then satisfy $\psi = 0$ at *either* $x = 0$ (if $C = 0$) *or* $x = 1$ (if $C = 1$). These solutions are illustrated in Fig. 5.5 for the particular stress (5.46).

Regardless of our choice of C we cannot satisfy $\psi = 0$ at both zonal boundaries. We must choose one, and then construct a *boundary layer* solution (i.e., we determine ϕ) to satisfy the other condition. Which choice do we make? On intuitive grounds it seems that we should choose the solution that satisfies $\psi = 0$ at $x = 1$ (the solution on the left in Fig. 5.5), for the interior flow then goes round in the same direction as the wind: the wind is supplying a clockwise torque, and to achieve an angular momentum balance anticlockwise angular momentum must be supplied by friction. We can imagine that this would be provided by the frictional forces at the western boundary layer if the interior flow is clockwise, but not by friction at an eastern boundary layer when the interior flow is anticlockwise. Note that this argument is not dependent on the sign of the wind-stress curl: if the wind blew the other way a similar argument still implies that a western boundary layer is needed. We will now see if and how the mathematics reflects this intuitive but non-rigorous argument.

Asymptotic matching

Near the walls of the domain the boundary layer correction $\phi(x, y)$ must become important in order that the boundary conditions may be satisfied, and the flow, and in particular $\phi(x, y)$, will vary rapidly with x. To reflect this, let us *stretch* the x-coordinate near this point of failure (i.e., at either $x = 0$ or $x = 1$, but we do not know at which yet) and let

$$x = \epsilon\alpha \qquad \text{or} \qquad x - 1 = \epsilon\alpha. \tag{5.48a,b}$$

Here, α is the stretched coordinate, which has values $\mathcal{O}(1)$ in the boundary layer, and ϵ is a small parameter, as yet undetermined. We then suppose that $\phi = \phi(\alpha, y)$, and using (5.42) in (5.41), we obtain

$$\epsilon_S(\nabla^2\psi_I + \nabla^2\phi) + \frac{\partial\psi_I}{\partial x} + \frac{1}{\epsilon}\frac{\partial\phi}{\partial\alpha} = \mathrm{curl}_z\boldsymbol{\tau}_T, \tag{5.49}$$

where $\phi = \phi(\alpha, y)$ and $\nabla^2\phi = \epsilon^{-2}\partial^2\phi/\partial\alpha^2 + \partial^2\phi/\partial y^2$. Now, by choice, ψ_I exactly satisfies Sverdrup balance, and so (5.49) becomes

$$\epsilon_S\left(\nabla^2\psi_I + \frac{1}{\epsilon^2}\frac{\partial^2\phi}{\partial\alpha^2} + \frac{\partial^2\phi}{\partial y^2}\right) + \frac{1}{\epsilon}\frac{\partial\phi}{\partial\alpha} = 0. \tag{5.50}$$

We now choose ϵ to obtain a physically meaningful solution. An obvious choice is $\epsilon = \epsilon_S$, for then the leading-order balance in (5.50) is

$$\frac{\partial^2\phi}{\partial\alpha^2} + \frac{\partial\phi}{\partial\alpha} = 0, \tag{5.51}$$

the solution of which is

$$\phi = A(y) + B(y)\mathrm{e}^{-\alpha}. \tag{5.52}$$

Evidently, ϕ grows exponentially in the negative α direction. If this were allowed, it would violate our assumption that solutions are small in the interior, and we must eliminate this possibility by allowing α to take only positive values in the interior of the domain, and by setting

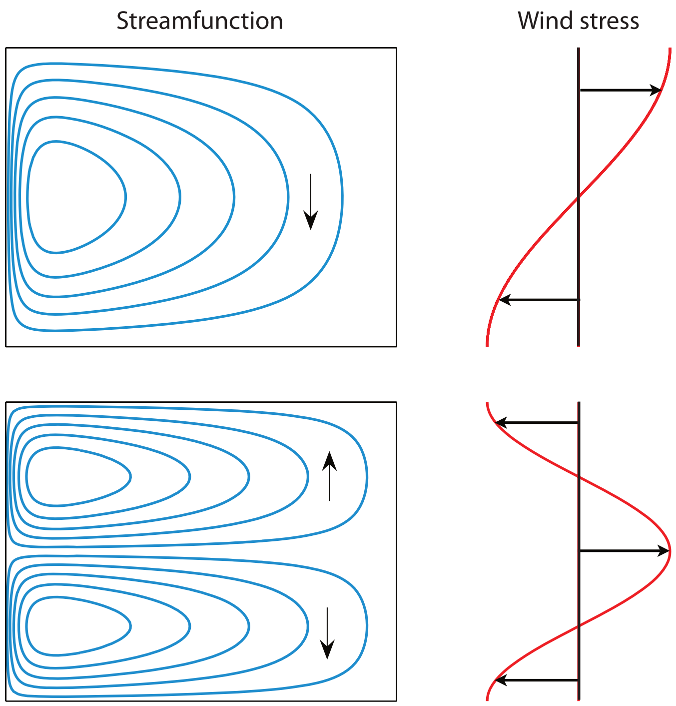

Figure 5.6 Two solutions of the Stommel model. Upper panel shows the streamfunction of a single-gyre solution, with a wind stress proportional to $-\cos(\pi y/a)$ (in a domain of side a), and the lower panel shows a two-gyre solution, with wind stress proportional to $\cos(2\pi y/a)$. In both cases $\epsilon_S = 0.04$.

$A(y) = 0$. We therefore choose $x = \epsilon\alpha$ so that $\alpha > 0$ for $x > 0$; the boundary layer is then at $x = 0$, that is, it is a *western boundary*, and it decays eastwards in the direction of increasing α — that is, into the ocean interior. We now choose $C = 1$ in (5.47) to make $\psi_I = 0$ at $x = 1$ in (5.47) and then, for the wind stress (5.46), the interior solution is given by

$$\psi_I = \pi(1 - x)\sin\pi y. \tag{5.53}$$

This alone satisfies the boundary condition at the eastern boundary. The function $B(y)$ is chosen to satisfy the additional condition that

$$\psi = \psi_I + \phi = 0 \qquad \text{at} \quad x = 0, \tag{5.54}$$

and using (5.53) this gives

$$\pi\sin\pi y + B(y) = 0. \tag{5.55}$$

Using this in (5.52), with $A(y) = 0$, then gives the boundary layer solution

$$\phi = -\pi \sin \pi y e^{-x/\epsilon_S}. \tag{5.56}$$

The composite (boundary layer plus interior) solution is the sum of (5.53) and (5.56), namely

$$\psi = (1 - x - e^{-x/\epsilon_S})\pi \sin \pi y. \tag{5.57}$$

With dimensional variables this is

$$\psi = \frac{\tau_0 \pi}{\beta}\left(1 - \frac{x}{a} - e^{-x/(a\epsilon_S)}\right) \sin \frac{\pi y}{a}. \tag{5.58}$$

This is a 'single gyre' solution. Two or more gyres can be obtained with a different wind forcing, such as $\tau^x = -\tau_0 \cos(2\pi y)$, as in Fig. 5.6.

It is a relatively straightforward matter to generalize to other wind stresses, provided these also vanish at the two latitudes between which the solution is desired. It is left as a problem to show that in general

$$\psi_I = \int_{x_E}^{x} \text{curl}_z \boldsymbol{\tau}(x', y)\, dx', \tag{5.59}$$

and that the composite solution is

$$\psi = \psi_I - \psi_I(0, y) e^{-x/(x_E \epsilon_S)}. \tag{5.60}$$