Chapter

4

Shallow Water Equations

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

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

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

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

4.1 Shallow Water Equations of Motion

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

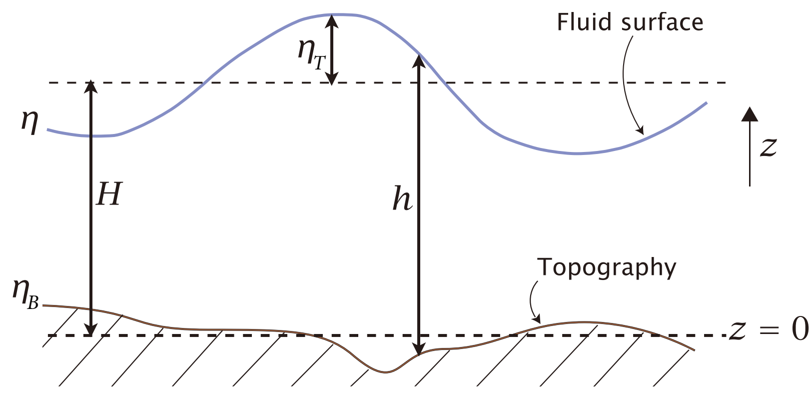

Fig. 4.1: A shallow water system where h is the thickness of a water column, H its mean thickness, η the height of the free surface and η_B is the height of the lower, rigid, surface above some arbitrary origin, typically chosen such that the average of η_B is zero. The quantity η_B is the deviation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

4.1.1 Momentum Equations

The vertical momentum equation is just the hydrostatic equation,

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

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

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

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

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

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

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

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

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

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

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

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

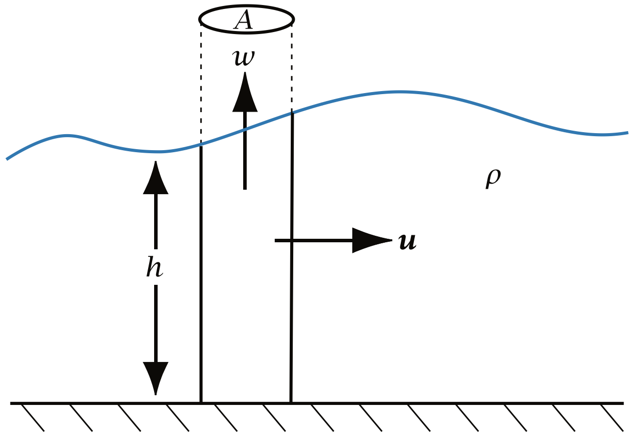

Fig. 4.2: The mass budget for a column of area A in a flat-bottomed shallow water system. The fluid leaving the column is $\oint \rho h \boldsymbol{u} \cdot \boldsymbol{n}\, dl$ where $\boldsymbol{n}$ is the unit vector normal to the boundary of the fluid column. There is a non-zero vertical velocity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

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

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

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

4.1.2 Mass Continuity Equation

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

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

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

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

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

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

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

$$F_m = \frac{d}{dt}\int \rho_0\, dV = \frac{d}{dt}\int_A \rho_0 h\, dA = \int_A \rho_0 \frac{\partial h}{\partial t}\, dA. \tag{4.11}$$

The Shallow Water Equations

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

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

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

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

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

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

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

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

with the rightmost expression holding in Cartesian coordinates.

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

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

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

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

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

4.1.3 Reduced Gravity Equations

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

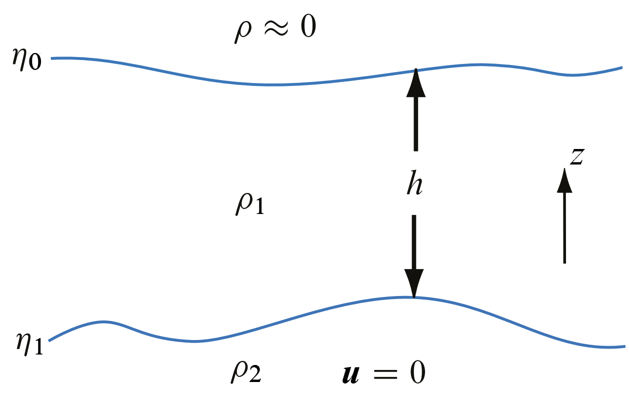

Fig. 4.3: The reduced gravity shallow water system. An active layer lies over a deep, denser, quiescent layer. In a common variation the upper surface is held flat by a rigid lid, and $\eta_0 = 0$.

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

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

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

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

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

and the momentum equation is

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

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

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

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

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

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

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

The equations are completed by the usual mass conservation equation,

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

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

4.2 Conservation Properties

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 $\mathrm{D}\varphi/\mathrm{D}t = 0$. An integral invariant is one that is conserved after an integration over some, usually closed, volume; energy is an example. The simplicity of the shallow water equations allows us to transparently see how these arise.

4.2.1 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*), meaning the kinetic energy per unit area, is $\rho_0 h\boldsymbol{u}^2/2$. The potential energy density (*PE*) of the fluid is

$$PE = \int_0^h \rho_0 g z \,\mathrm{d}z = \frac{1}{2}\rho_0 g h^2. \tag{4.21}$$

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 $z = 0$.

Energy is moved from place to place by pressure forces as well as advection, and is not a material invariant. However, in total energy is conserved and it is an integral invariant.

Using the mass conservation equation (4.13b) we obtain an equation for the evolution of potential energy density, namely

$$\frac{\mathrm{D}}{\mathrm{D}t}\frac{gh^2}{2} + gh^2\nabla\cdot\boldsymbol{u} = 0 \tag{4.22a}$$

or

$$\frac{\partial}{\partial t}\frac{gh^2}{2} + \nabla\cdot\left(\boldsymbol{u}\frac{gh^2}{2}\right) + \frac{gh^2}{2}\nabla\cdot\boldsymbol{u} = 0. \tag{4.22b}$$

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

$$\frac{\mathrm{D}}{\mathrm{D}t}\frac{h\boldsymbol{u}^2}{2} + \frac{\boldsymbol{u}^2 h}{2}\nabla\cdot\boldsymbol{u} = -g\boldsymbol{u}\cdot\nabla\frac{h^2}{2} \tag{4.23a}$$

or

$$\frac{\partial}{\partial t}\frac{h\boldsymbol{u}^2}{2} + \nabla\cdot\left(\boldsymbol{u}\frac{h\boldsymbol{u}^2}{2}\right) + g\boldsymbol{u}\cdot\nabla\frac{h^2}{2} = 0. \tag{4.23b}$$

Adding (4.22b) and (4.23b) we obtain

$$\frac{\partial}{\partial t}\frac{1}{2}\left(h\boldsymbol{u}^2 + gh^2\right) + \nabla\cdot\left[\frac{1}{2}\boldsymbol{u}\left(gh^2 + h\boldsymbol{u}^2 + gh^2\right)\right] = 0 \tag{4.24}$$

or

$$\frac{\partial E}{\partial t} + \nabla\cdot\boldsymbol{F} = 0, \tag{4.25}$$

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 (4.24) over that area and using Gauss's theorem, the total energy is seen to be conserved; that is

$$\frac{\mathrm{d}\widehat{E}}{\mathrm{d}t} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_A (h\boldsymbol{u}^2 + gh^2)\,\mathrm{d}A = 0. \tag{4.26}$$

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 2, the energy flux in (4.25) is not just the energy density multiplied by the velocity; it contains an additional term $g\boldsymbol{u}h^2/2$, and this represents the energy transfer occurring when the fluid does work against the pressure force.

4.2.2 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:

$$\boldsymbol{\omega} \equiv \nabla \times \boldsymbol{v}, \qquad \boldsymbol{\omega}^* \equiv \nabla \times \boldsymbol{u}. \tag{4.27}$$

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

$$\boldsymbol{\omega}^* = \hat{\mathbf{k}}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) = \hat{\mathbf{k}}\zeta. \tag{4.28}$$

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

$$(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = \frac{1}{2}\nabla(\boldsymbol{u}\cdot\boldsymbol{u}) - \boldsymbol{u}\times(\nabla\times\boldsymbol{u}), \tag{4.29}$$

to write the momentum equation, (4.7) as

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{\omega}^* \times \boldsymbol{u} + \boldsymbol{f}\times\boldsymbol{u} = -\nabla\left(g\eta + \frac{1}{2}\boldsymbol{u}^2\right). \tag{4.30}$$

To obtain an evolution equation for the vorticity we take the curl of (4.30), and make use of the vector identity

$$\begin{aligned}\nabla\times(\boldsymbol{\omega}^*\times\boldsymbol{u}) &= (\boldsymbol{u}\cdot\nabla)\boldsymbol{\omega}^* - (\boldsymbol{\omega}^*\cdot\nabla)\boldsymbol{u} + \boldsymbol{\omega}^*\nabla\cdot\boldsymbol{u} - \boldsymbol{u}\nabla\cdot\boldsymbol{\omega}^* \\ &= (\boldsymbol{u}\cdot\nabla)\boldsymbol{\omega}^* + \boldsymbol{\omega}^*\nabla\cdot\boldsymbol{u},\end{aligned} \tag{4.31}$$

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. A similar expression holds for $\boldsymbol{f}$ so that taking the curl of (4.30) gives

$$\frac{\partial\zeta}{\partial t} + (\boldsymbol{u}\cdot\nabla)(\zeta + f) = -(\zeta + f)\nabla\cdot\boldsymbol{u}, \tag{4.32}$$

where $\zeta = \hat{\mathbf{k}}\cdot\boldsymbol{\omega}^*$ and f varies with latitude (and on the beta-plane $f = f_0 + \beta y$). Since f does not vary with time we can write (4.32) as

$$\frac{\mathrm{D}}{\mathrm{D}t}(\zeta + f) = -(\zeta + f)\nabla\cdot\boldsymbol{u}. \tag{4.33}$$

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

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

and using this equation and (4.32) we obtain

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

which is equivalent to

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

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

4.3 Shallow Water Waves

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

4.3.1 Non-Rotating Shallow Water Waves

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

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

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

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

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

and neglecting squares of small quantities this yields the linear equation

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

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

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

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

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

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

$$\eta' = \operatorname{Re} \tilde{\eta} e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{4.42}$$

where $\tilde{\eta}$ is a complex constant, $\boldsymbol{k} = k\hat{\mathbf{i}} + l\hat{\mathbf{j}}$ 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 (4.41) leads to the dispersion relationship

$$\omega = \pm ck, \tag{4.43}$$

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

$$\eta'(x,t) = \frac{1}{2}\left[F(x-ct) + F(x+ct)\right], \tag{4.44}$$

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.

4.3.2 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, (SW.1) and (SW.2) on page 66, about a state of rest we obtain

$$\frac{\partial u'}{\partial t} - f_0 v' = -g\frac{\partial \eta'}{\partial x}, \qquad \frac{\partial v'}{\partial t} + f_0 u' = -g\frac{\partial \eta'}{\partial y}, \tag{4.45a,b}$$

$$\frac{\partial \eta'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0. \tag{4.45c}$$

To obtain a dispersion relationship we let

$$(u, v, \eta) = (\tilde{u}, \tilde{v}, \tilde{\eta}) e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{4.46}$$

and substitute into (4.45), giving

$$\begin{pmatrix} -i\omega & -f_0 & i\,gk \\ f_0 & -i\omega & i\,gl \\ i\,Hk & i\,Hl & -i\omega \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\eta} \end{pmatrix} = 0. \tag{4.47}$$

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

$$\omega(\omega^2 - f_0^2 - c^2K^2) = 0, \tag{4.48}$$

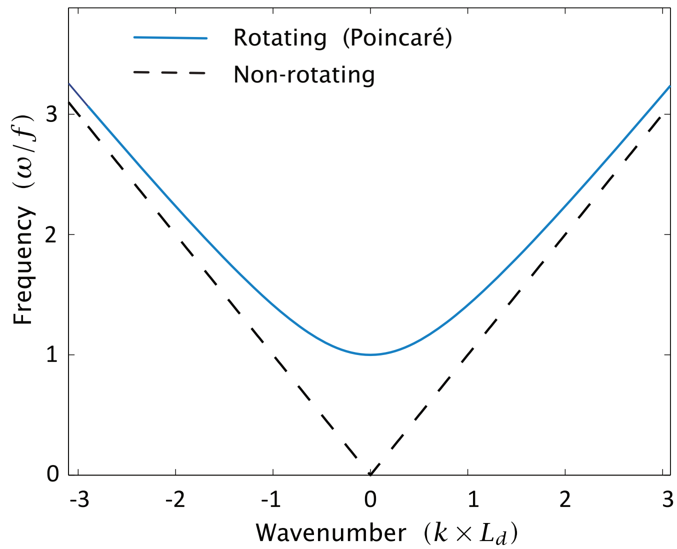

Fig. 4.4: Dispersion relation for Poincaré waves and non-rotating shallow water waves. Frequency is scaled by the Coriolis frequency f, and wavenumber by the inverse deformation radius $\sqrt{gH}/f$. For small wavenumbers the frequency of the Poincaré waves is approximately f, and for high wavenumbers it asymptotes to that of non-rotating waves.

where $K^2 = k^2 + l^2$ and $c^2 = gH$. There are two classes of solution to (4.48). The first is simply $\omega = 0$, i.e., time-independent flow corresponding to geostrophic balance in (4.45). Because geostrophic balance gives a divergence-free velocity field for a constant Coriolis parameter the equations are satisfied by a time-independent solution. The second set of solutions gives the dispersion relation

$$\omega^2 = f_0^2 + c^2(k^2 + l^2), \tag{4.49}$$

or

$$\omega^2 = f_0^2 + gH(k^2 + l^2). \tag{4.50}$$

The corresponding waves are known as *Poincaré* waves, and the dispersion relationship is illustrated in Fig. 4.4. Note that the frequency is always greater than the Coriolis frequency f_0. There are two interesting limits:

(Jules) Henri Poincaré (1854–1912) was a remarkable French mathematician, physicist and philosopher. As well as the Poincaré waves of this chapter he is also known as one of the founders of what is now known as chaos theory — he knew that the weather was inherently, and not just in practice, unpredictable.

(i) The short wave limit. If

$$K^2 \gg \frac{f_0^2}{gH}, \tag{4.51}$$

where $K^2 = k^2 + l^2$, then the dispersion relationship reduces to that of the non-rotating case (4.43). This condition is equivalent to requiring that the wavelength be much shorter than the *deformation radius*, defined by

$$L_d = \frac{\sqrt{gH}}{f}. \tag{4.52}$$

Specifically, if $l = 0$ and $\lambda = 2\pi/k$ is the wavelength, the condition is $\lambda^2 \ll L_d^2(2\pi)^2$. 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

$$K^2 \ll \frac{f_0^2}{gH}, \tag{4.53}$$

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

$$\omega = f_0. \tag{4.54}$$

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

$$\frac{\partial u'}{\partial t} - f_0 v' = 0, \qquad \frac{\partial v'}{\partial t} + f_0 u' = 0, \tag{4.55}$$

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

4.3.3 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 $y = 0$; general 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 answer is yes, and to show that we begin with the linearized shallow water equations, namely

The affirmative answer to the question in the text was provided by William Thomson in the nineteenth century. Thomson later became Lord Kelvin, taking the name of the river flowing near his university in Glasgow, and the waves are now eponymously known as Kelvin waves.

$$\frac{\partial u'}{\partial t} - f_0 v' = -g\frac{\partial \eta'}{\partial x}, \qquad \frac{\partial v'}{\partial t} + f_0 u' = -g\frac{\partial \eta'}{\partial y}, \tag{4.56a,b}$$

$$\frac{\partial \eta'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0. \tag{4.56c}$$

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

$$\frac{\partial u'}{\partial t} = -g\frac{\partial \eta'}{\partial x}, \qquad f_0 u' = -g\frac{\partial \eta'}{\partial y}, \qquad \frac{\partial \eta'}{\partial t} + H\frac{\partial u'}{\partial x} = 0. \tag{4.57a,b,c}$$

Equations (4.57a) and (4.57c) lead to the standard wave equation

$$\frac{\partial^2 u'}{\partial t^2} = c^2\frac{\partial^2 u'}{\partial x^2}, \tag{4.58}$$

where $c = \sqrt{gH}$, the usual wave speed of shallow water waves. The solution of (4.58) is

$$u' = F_1(x + ct, y) + F_2(x - ct, y), \tag{4.59}$$

with corresponding surface displacement

$$\eta' = \sqrt{H/g}\left[-F_1(x + ct, y) + F_2(x - ct, y)\right]. \tag{4.60}$$

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 (4.57b) which gives

$$\frac{\partial F_1}{\partial y} = \frac{f_0}{\sqrt{gH}} F_1, \qquad \frac{\partial F_2}{\partial y} = -\frac{f_0}{\sqrt{gH}} F_2, \tag{4.61}$$

with solutions

$$F_1 = F(x + ct)\mathrm{e}^{y/L_d}, \qquad F_2 = G(x - ct)\mathrm{e}^{-y/L_d}, \tag{4.62}$$

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

$$u' = \mathrm{e}^{-y/L_d} G(x - ct), \qquad v' = 0, \qquad \eta' = \sqrt{H/g}\,\mathrm{e}^{-y/L_d} G(x - ct). \tag{4.63a,b,c}$$

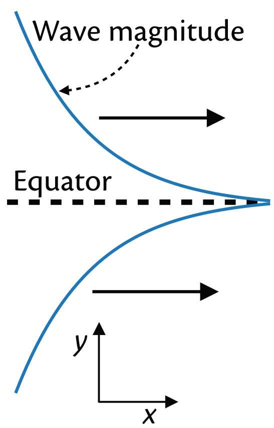

Fig. 4.5: Kelvin waves propagating eastward at the equator and decaying rapidly away to either side.

These are Kelvin waves and they decay exponentially away from the boundary. In general, for f_0 positive (negative) the boundary is to the right (left) of the wave direction. Given a constant Coriolis parameter, we could also have obtained a solution on a meridional wall, again with the wave moving with the wall on its the right if f_0 is positive. The equator also acts like a wall, with Kelvin waves propagating eastward along it, decaying to either side, as in Fig. 4.5. (The details of the solution for equatorial Kelvin waves differ from the case with $f = f_0$, because now $f = 0$ at the equator itself, but the structure is similar to the solution above.)

4.4 Geostrophic Adjustment

Geostrophic balance occurs in the shallow water 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:

$$\boldsymbol{f} \times \boldsymbol{u}_g = -g\nabla\eta. \tag{4.64}$$

Thus, the geostrophic velocity is proportional to the slope of the surface, as sketched in Fig. 4.6. The result arises because by hydrostatic balance the slope of an interfacial surface is directly related to the difference in pressure gradient on either side, so that in geostrophic balance the velocity is related to the slope.

But consider the more general question, *why* are the atmosphere and ocean close to geostrophic balance? Suppose that the initial state were not balanced, that it had small scale motion and sharp pressure gradients. How would the system evolve? It turns out that all of these sharp gradients would radiate away, leaving behind a geostrophically balanced state. The process is called *geostrophic adjustment*, and it occurs quite generally in rotating fluids, whether stratified or not. The shallow water equations are an ideal set of equations to explore this process, as we now see.

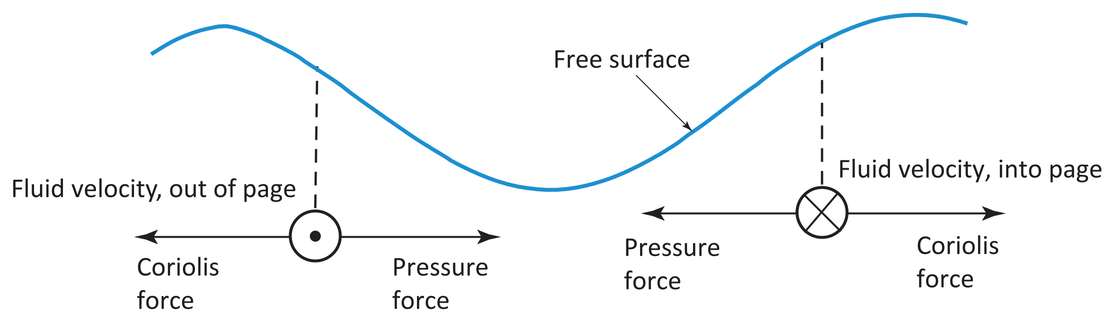

Fig. 4.6: Geostrophic flow in a shallow water system, with a positive value of the Coriolis parameter f, as in the Northern Hemisphere. The pressure force is directed down the gradient of the height field, and this can be balanced by the Coriolis force if the fluid velocity is at right angles to it. If f were negative, the geostrophic flow would be reversed.

4.4.1 Posing the Problem

Let us consider the free evolution of a single shallow layer of fluid whose initial state is manifestly unbalanced, and we suppose that surface displacements are small so that the evolution of the system is described by the linearized shallow equations of motion. These are

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{f} \times \boldsymbol{u} = -g\nabla\eta, \qquad \frac{\partial \eta}{\partial t} + H\nabla \cdot \boldsymbol{u} = 0, \tag{4.65a,b}$$

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

4.4.2 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

$$\eta(x, t = 0) = \begin{cases} +\eta_0 & x < 0 \\ -\eta_0 & x > 0, \end{cases} \tag{4.66}$$

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 (4.44) where the functional form is determined by the initial conditions so that here

$$F(x) = \eta(x, t = 0) = -\eta_0 \operatorname{sgn}(x). \tag{4.67}$$

Equation (4.44) 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

$$\eta(x, t) = -\frac{1}{2}\eta_0[\operatorname{sgn}(x + ct) + \operatorname{sgn}(x - ct)]. \tag{4.68}$$

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

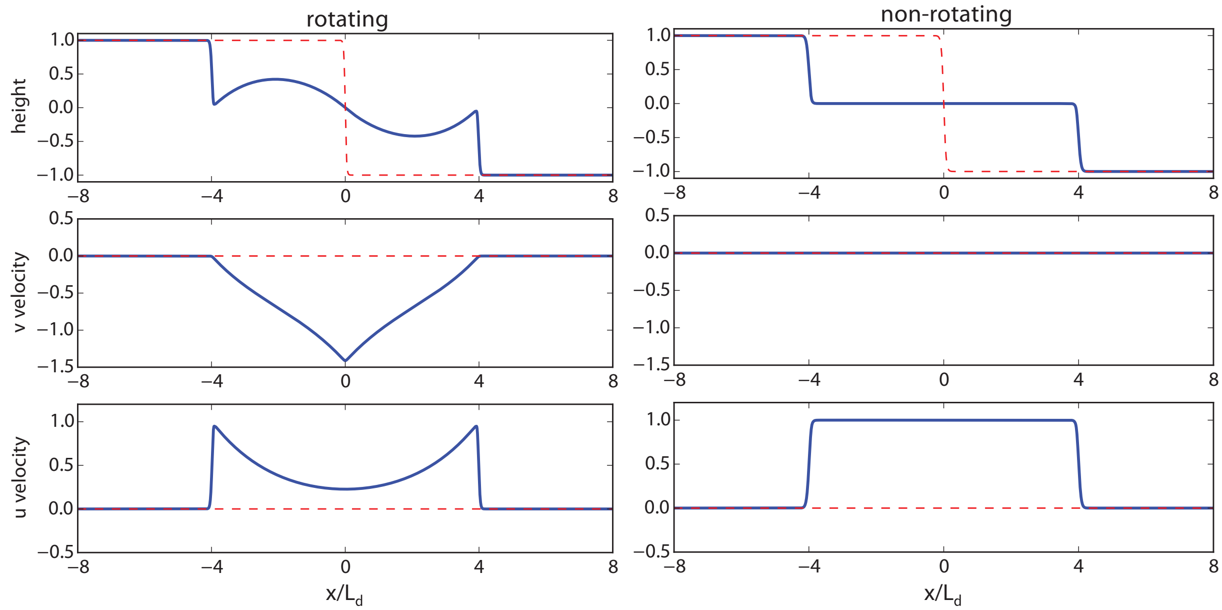

Fig. 4.7: The solutions of the shallow water equations obtained by numerically integrating the equations of motion with and without rotation. The panels show snapshots of the state of the fluid (solid lines) soon after being released from a stationary initial state (red dashed lines) with a height discontinuity. The rotating flow is evolving toward an end state similar to Fig. 4.8 whereas the non-rotating flow will eventually become stationary. In the non-rotating case L_d is defined using the rotating parameters.

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

$$\frac{\partial u}{\partial t} = -g\frac{\partial \eta}{\partial x}, \tag{4.69}$$

which gives, using (4.44),

$$u = -\frac{g}{2c}[F(x+ct) - F(x-ct)]. \tag{4.70}$$

Consider the case with initial conditions given by (4.66). At a given location, away from the initial disturbance, the fluid remains at rest and undisturbed until the front arrives. After the front has passed, the fluid surface is again undisturbed and the velocity is uniform and non-zero. Specifically:

$$\eta = \begin{cases} -\eta_0 \operatorname{sgn}(x) \\ 0 \end{cases} \qquad u = \begin{cases} 0 & |x| > ct \\ (\eta_0 g/c) & |x| < ct. \end{cases} \tag{4.71}$$

The solution with a discontinuity in the height field, and zero initial velocity, is illustrated in the right-hand panels in Fig. 4.7. The front propagates in either direction from the discontinuity and, in this case, the final velocity, as well as the fluid displacement, is zero. That is, the disturbance is radiated completely away.

4.4.3 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. In a rotating shallow fluid that conservation is represented by

$$\frac{\partial Q}{\partial t} + \boldsymbol{u} \cdot \nabla Q = 0, \tag{4.72}$$

where $Q = (\zeta + f)/h$. In the linear case with constant Coriolis parameter, (4.72) becomes

$$\frac{\partial q}{\partial t} = 0, \qquad q = \left(\zeta - f_0 \frac{\eta}{H}\right). \tag{4.73}$$

This equation may be obtained either from the linearized velocity and mass conservation equations, (4.65), or from (4.72) directly. In the latter case we can write

$$Q = \frac{\zeta + f_0}{H + \eta} \approx \frac{1}{H}(\zeta + f_0)\left(1 - \frac{\eta}{H}\right) \approx \frac{1}{H}\left(f_0 + \zeta - f_0\frac{\eta}{H}\right) = \frac{f_0}{H} + \frac{q}{H}, \tag{4.74}$$

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, (4.72) reduces to (4.73). 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 (4.66), the initial potential vorticity is given by

$$q(x, y) = \begin{cases} -f_0\eta_0/H & x < 0 \\ f_0\eta_0/H & x > 0, \end{cases} \tag{4.75}$$

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

$$\zeta - f_0\frac{\eta}{H} = q(x, y), \qquad f_0 u = -g\frac{\partial \eta}{\partial y}, \qquad f_0 v = g\frac{\partial \eta}{\partial x}, \tag{4.76a,b,c}$$

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 (4.76) then reduce to

$$\left(\nabla^2 - \frac{1}{L_d^2}\right)\psi = q(x, y), \tag{4.77}$$

where $L_d = \sqrt{gH}/f_0$ is the radius of deformation, as in (4.52), sometimes called 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.

If an unbalanced flow is left to freely evolve, gravity waves will propagate quickly away from the disturbance. However, the balanced component of the flow is constrained by potential vorticity conservation and it can only evolve advectively, and so much more slowly. Thus, after some time, only the balanced part of the flow remains, with the unbalanced flow having been radiated away and dissipated. This process is *geostrophic adjustment*.

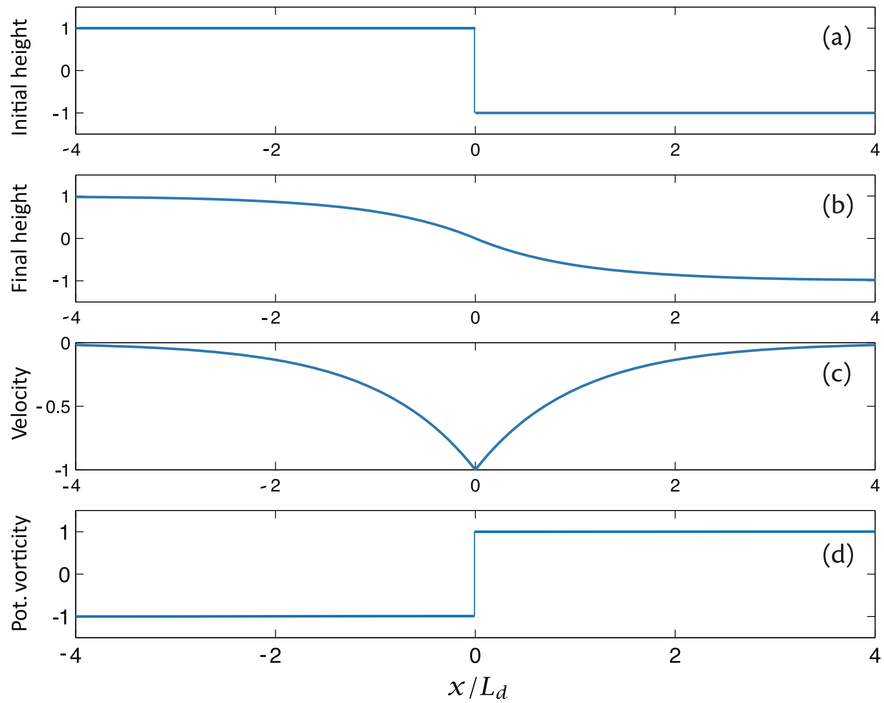

Fig. 4.8: Solutions of a linear geostrophic adjustment problem. (a) Initial height field, given by (4.66) with $\eta_0 = 1$. (b) Equilibrium (final) height field, η given by (4.79) and $\eta = f_0\psi/g$. (c) Equilibrium geostrophic velocity, normal to the gradient of the height field, given by (4.80). (d) Potential vorticity, given by (4.75), and this does not evolve. The distance, x is nondimensionalized by the deformation radius L_d and the velocity by $\eta_0(g/f_0L_d)$. Changes to the initial state occur within $\mathcal{O}(L_d)$ of the initial discontinuity.

The initial conditions (4.75) admit of a nice analytic solution, for the flow will remain uniform in y, and (4.77) reduces to

$$\frac{\partial^2\psi}{\partial x^2} - \frac{1}{L_d^2}\psi = \frac{f_0\eta_0}{H}\,\mathrm{sgn}(x). \tag{4.78}$$

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

$$\psi = \begin{cases} -(g\eta_0/f_0)(1 - \mathrm{e}^{-x/L_d}) & x > 0 \\ +(g\eta_0/f_0)(1 - \mathrm{e}^{x/L_d}) & x < 0. \end{cases} \tag{4.79}$$

The velocity field associated with this is obtained from (4.76b,c), and is

$$u = 0, \qquad v = -\frac{g\eta_0}{f_0 L_d}\,\mathrm{e}^{-|x|/L_d}. \tag{4.80}$$

The velocity is perpendicular to the slope of the free surface, and a jet forms along the initial discontinuity, as illustrated in Fig. 4.8.

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. The time evolution of the rotating flow, obtained by a numerical integration, is illustrated in the left-hand panels of Fig. 4.7.

Fronts propagate away at a speed $\sqrt{gH} = 1$, just as in the non-rotating case, but in the rotating flow they leave behind a geostrophically balanced state with a non-zero meridional velocity.

4.5 ♦ A Variational Perspective on Adjustment

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, mostly trapped within a deformation radius of the initial disturbance, 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, $\widehat{E}$, is given by the sum of potential and kinetic energies, namely

$$\widehat{E} = \int (H\boldsymbol{u}^2 + g\eta^2)\,\mathrm{d}A, \tag{4.81}$$

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

$$q = \zeta - f_0\frac{\eta}{H} = (v_x - u_y) - f_0\frac{\eta}{H}, \tag{4.82}$$

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.

The mathematical problem is to extremize the integral

$$I = \int \left\{H(u^2 + v^2) + g\eta^2 + \lambda(x, y)\left[(v_x - u_y) - f_0\eta/H\right]\right\}\,\mathrm{d}A, \tag{4.83}$$

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

$$\frac{\partial L}{\partial \eta} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \eta_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial \eta_y} = 0,$$
$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial u_y} = 0, \qquad \frac{\partial L}{\partial v} - \frac{\partial}{\partial x}\frac{\partial L}{\partial v_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial v_y} = 0, \tag{4.84}$$

where L is the integrand on the right-hand side of (4.83). Substituting the expression for L into (4.84) gives, after a little algebra,

$$2g\eta - \frac{\lambda f_0}{H} = 0, \qquad 2Hu + \frac{\partial \lambda}{\partial y} = 0, \qquad 2Hv - \frac{\partial \lambda}{\partial x} = 0, \tag{4.85}$$

and then eliminating λ gives the simple relationships

$$u = -\frac{g}{f_0}\frac{\partial \eta}{\partial y}, \qquad v = \frac{g}{f_0}\frac{\partial \eta}{\partial x}. \tag{4.86}$$

These 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.*

Notes and References

The shallow water equations are sometimes called the Saint-Venant equations after Adhémar Jean Claude Barré de Saint-Venant (1797–1886) who wrote down a form of the equations in 1871. Pierre-Simon Laplace (1749–1825) previously wrote down a linear version of the equations on the sphere, now known as Laplace's tidal equations, in 1776, and one would not be surprised if Euler knew about the equations before that.

Problems

4.1 Using the shallow water equations:

(*a*) A cylindrical column of air at 30° latitude with radius 100 km expands horizontally (shrinking in depth accordingly) to twice its original radius. If the air is initially at rest, what is the mean tangential velocity at the perimeter after the expansion?

(*b*) An air column at 60°N with zero relative vorticity ($\zeta = 0$) stretches from the surface to the tropopause, which we assume is a rigid lid, at 10 km. The air column moves zonally on to a plateau 2.5 km high. What is its relative vorticity? Suppose it then moves southwards to 30°N, staying on the plateau. What is its vorticity?

4.2 Show that the vertical velocity within a shallow water system is given by

$$w = \frac{z - \eta_b}{h}\frac{\mathrm{D}h}{\mathrm{D}t} + \frac{\mathrm{D}\eta_b}{\mathrm{D}t}. \tag{P4.1}$$

Interpret this result, showing that it gives sensible answers at the top and bottom of the fluid layer.

4.3 In the long-wave limit of Poincaré waves, fluid parcels behave as free-agents; that is, like free solid particles moving in a rotating frame unencumbered by pressure forces. Why then, is their frequency given by $\omega = f = 2\Omega$ where Ω is the rotation rate of the coordinate system, and not by Ω itself? Do particles that are stationary or move in a straight line in the inertial frame of reference satisfy the dispersion relationship for Poincaré waves in this limit? Explain. [See also Durran (1993) and Egger (1999).]

4.4 Linearize the f-plane shallow water system about a state of rest. Suppose that there is an initial disturbance given in the general form

$$\eta = \iint \tilde{\eta}_{k,l} e^{i(kx+ly)} \, dk \, dl, \tag{P4.2}$$

where η is the deviation surface height and the Fourier coefficients $\tilde{\eta}_{k,l}$ are given, and that the initial velocity is zero.

(a) Obtain the geopotential field at the completion of geostrophic adjustment, and show that the deformation scale is a natural length scale in the problem.

(b) Show that the change in total energy during the adjustment is always less than or equal to zero. Neglect any initial divergence.
N.B. Because the problem is linear, the Fourier modes do not interact.

4.5 If energy conservation is one of the most basic physical laws, how can energy be lost in geostrophic adjustment?

4.6 *Geostrophic adjustment of a velocity jump.* In which we consider the evolution of the linearized f-plane shallow water equations in an infinite domain.

(a) Show that the *linearized* potential vorticity, q', for the shallow water system is given by

$$q' = \zeta' - f_0 \frac{h'}{H}, \tag{P4.3}$$

using standard notation.

(b) If the flow is in geostrophic balance show that the relative vorticity is given by

$$\zeta' = \nabla^2 \psi,$$

where $\psi = gh'/f_0$. Hence show that the potential vorticity is then given by

$$q' = \nabla^2 \psi - \frac{1}{L_d^2} \psi,$$

and write down an expression for L_d.

(c) Suppose that initially the fluid surface is flat, the zonal velocity is zero and the meridional velocity is given by

$$v(x) = v_0 \, \mathrm{sgn}(x). \tag{P4.4}$$

Find the equilibrium height and velocity fields at $t = \infty$.

(d) What are the initial and final kinetic and potential energies?

Partial solution:
The potential vorticity is $q = \zeta - f_0 \eta / H$, so that the initial state and final state are both given by

$$q = 2v_0 \, \delta(x). \tag{P4.5}$$

(Why?) The final state streamfunction is thus given by $\left(\partial^2/\partial x^2 - L_d^{-2}\right)\psi = q$, with solution $\psi = \psi_0 \exp(x/L_d)$ and $\psi = \psi_0 \exp(-x/L_d)$ for $x < 0$ and $x > 0$, where $\psi_0 = -L_d v_0$ (why?), and $\eta = f_0 \psi / g$. The energy is $E = \int (Hv^2 + g\eta^2)/2 \, d\boldsymbol{x}$. The initial KE is infinite, the initial PE is zero, and the final state has $PE = KE = gL_d \eta_0^2/4$; that is, the energy is equipartitioned between kinetic and potential energies.

4.7 In the shallow water equations show that, if the flow is approximately geostrophically balanced, the energy at large scales is predominantly potential energy and the energy at small scales is predominantly kinetic energy. Define precisely what 'large scale' and 'small scale' mean in this context.

4.8 In the shallow water geostrophic adjustment problem, show that at large scales the velocity essentially adjusts to the height field, and that at small scales the height field essentially adjusts to the velocity field. Your derivation may be detailed and mathematical, but explain the result at the end in words and in physical terms.

Chapter

5

Geostrophic Theory

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

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

5.1 Scaling the Shallow Water Equations

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

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

and similarly for the other variables in the equations.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Nondimensional momentum equation

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

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

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

Nondimensional mass continuity (height) equation

The (dimensional) mass continuity equation can be written as

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

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

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

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

5.2 Geostrophic Shallow Water Equations

5.2.1 Planetary-Geostrophic Equations

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

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

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

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

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

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

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

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

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

Planetary-geostrophic potential vorticity

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

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

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

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

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

5.2.2 Quasi-Geostrophic Equations

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A consequence of (5.20) is that 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. On the left-hand side the velocities are well-approximated by using the geostrophic flow, so that we have

$$\frac{\partial\zeta_g}{\partial t} + (\boldsymbol{u}_g\cdot\nabla)(\zeta_g + f) = -f_0\nabla\cdot\boldsymbol{u}_a, \tag{5.22}$$

where 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 (5.10) the mass continuity equation is

$$\frac{\mathrm{D}h_D}{\mathrm{D}t} + (H + h_D)\nabla\cdot\boldsymbol{u} = 0, \tag{5.23}$$

and since $H \gg h_D$ (using (5.11), H is bigger by a factor $(L_d/L)^2 Ro^{-1}$), the equation becomes

$$\frac{\mathrm{D}}{\mathrm{D}t}(\eta - \eta_B) + H\nabla\cdot\boldsymbol{u}_a = 0. \tag{5.24}$$

Combining (5.22) and (5.23) gives

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\zeta_g + f - \frac{f_0(\eta-\eta_B)}{H}\right) = 0. \tag{5.25}$$

It appears that we have two dynamical 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 fields are related to the streamfunction by

Both the planetary-geostrophic and the quasi-geostrophic equations can be written in the form of an evolution equation for potential vorticity, along with an 'inversion' to determine the velocity fields and the height field. The difference in the two equation sets lies in the approximations made in the inversion. In the planetary-geostrophic system relative vorticity is ignored and $Q = f/h$. In quasi-geostrophy the variations in the height field are small and $q = \beta y + \zeta - f_0\eta_T/H$. Both equation sets assume a low Rossby number.

$$\zeta_g = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla^2\psi, \qquad \text{and} \qquad \eta = \frac{f_0\psi}{g}, \tag{5.26a,b}$$

where the second relation comes from geostrophic balance. Equation (5.25) may then be written as

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \qquad q = \nabla^2\psi + \beta y - \frac{\psi}{L_d^2} + \frac{f_0\eta_B}{H}, \tag{5.27}$$

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

5.2.3 Quasi-Geostrophic Potential Vorticity

Connection to shallow water potential vorticity

The quantity q given by (5.27) 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 let us begin with the expression for the shallow water potential vorticity,

$$Q = \frac{f + \zeta}{h}. \tag{5.28}$$

For simplicity we set bottom topography to zero and then $h = H(1+\eta_T/H)$ and assume that η_T/H is small to obtain

$$Q = \frac{f + \zeta}{H(1 + \eta_T/H)} \approx \frac{1}{H}(f + \zeta)\left(1 - \frac{\eta_T}{H}\right) \approx \frac{1}{H}\left(f_0 + \beta y + \zeta - f_0\frac{\eta_T}{H}\right). \tag{5.29}$$

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

$$q = \beta y + \zeta - f_0\frac{\eta_T}{H} \tag{5.30}$$

is materially conserved. Using geostrophic balance we have $\zeta = \nabla^2\psi$ and $\eta_T = f_0\psi/g$ so that (5.30) is identical (except for η_B) to (5.27).

The approximations needed to go from (5.28) to (5.30) 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 $H \gg \eta_T$ to enable the expansion of the height field and this approximation 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 (5.27).

5.3 Scaling in the Continuously-Stratified System

We now apply the same scaling ideas, *mutatis mutandis,* to the stratified primitive equations. We use the hydrostatic Boussinesq equations, which we write as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla_z\phi, \tag{5.31a}$$

$$\frac{\partial\phi}{\partial z} = b, \tag{5.31b}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} = 0, \tag{5.31c}$$

$$\nabla\cdot\boldsymbol{v} = 0. \tag{5.31d}$$

We will consider only the case of a flat bottom, but topography is a relatively straightforward extension. 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 (5.31c) by writing

$$b = \tilde{b}(z) + b'(x, y, z, t). \tag{5.32}$$

The thermodynamic equation then becomes

$$\frac{Db'}{Dt} + N^2 w = 0, \tag{5.33}$$

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

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

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

5.3.1 Scaling the Equations

We scale the basic variables by supposing that

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

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

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

and so the pressure scales as

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

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

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

and from this we obtain

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

where

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

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

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

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

with the scaling

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

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

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

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

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

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

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

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

and

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

The nondimensional mass conservation equation is simply

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

and the nondimensional thermodynamic equation is

$$\frac{f_0 UL}{H}\frac{U}{L}\frac{\mathrm{D}\hat{b}}{\mathrm{D}\hat{t}} + \widehat{N}^2 N_0^2\frac{HU}{L}\widehat{w} = 0, \tag{5.48}$$

or, re-arranging,

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

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

Nondimensional Boussinesq Primitive Equations

The nondimensional, hydrostatic, Boussinesq equations in a rotating frame of reference are:

Horizontal momentum: $$Ro\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} + \widehat{\boldsymbol{f}} \times \hat{\boldsymbol{u}} = -\nabla\widehat{\phi}, \tag{PE.1}$$

Hydrostatic: $$\frac{\partial\widehat{\phi}}{\partial\widehat{z}} = \hat{b}, \tag{PE.2}$$

Mass continuity: $$\frac{\partial\widehat{u}}{\partial\widehat{x}} + \frac{\partial\widehat{v}}{\partial\widehat{y}} + \frac{\partial\widehat{w}}{\partial\widehat{z}} = 0, \tag{PE.3}$$

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

5.4 PLANETARY-GEOSTROPHIC EQUATIONS FOR STRATIFIED FLUIDS

The *planetary-geostrophic equations* are appropriate for geostrophic flow at large horizontal scales. The specific assumptions we apply to the primitive equations are that:

(i) $Ro \ll 1$,

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

We also allow f to have its full variation with latitude, although we still use Cartesian coordinates. If we look at the equations in the shaded box above we see that, if we are to retain only the dominant terms, the momentum equation should simply be replaced by geostrophic balance, whereas hydrostatic and mass continuity equations are unaltered.

In the thermodynamic equation both terms are of order Rossby number, and therefore we retain both. This circumstance arises because, by assumption, we are dealing with large scales and on these scales the perturbation buoyancy varies as much as the mean buoyancy itself. The buoyancy equation reverts to the evolution of total buoyancy, namely

$$\frac{\mathrm{D}b}{\mathrm{D}t} = 0, \tag{5.50}$$

where the material derivative is fully three-dimensional. This is the only evolution equation in the system. Thus, to summarize, the *planetary-geostrophic* equations of motion are, in dimensional form,

$$\frac{\mathrm{D}b}{\mathrm{D}t} = 0, \qquad \boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi, \qquad \frac{\partial\phi}{\partial z} = b', \qquad \nabla\cdot\boldsymbol{v} = 0. \tag{5.51}$$

Potential vorticity

Manipulation of (5.51) reveals that we can equivalently write the equations as an evolution equation for potential vorticity. Thus, the evolution equations may be written as

$$\frac{\mathrm{D}Q}{\mathrm{D}t} = 0, \qquad Q = f\frac{\partial b}{\partial z}. \tag{5.52}$$

The inversion — i.e., the diagnosis of velocity, pressure and buoyancy — is carried out using the hydrostatic, geostrophic and mass conservation equations, and in a real situation the right-hand sides of the buoyancy and potential vorticity equations would have terms representing heating.

5.4.1 Applicability to the Ocean and Atmosphere

In the atmosphere a typical deformation radius NH/f is about 1000 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 scales of motion can satisfy the planetary-geostrophic scaling in the atmosphere and we should then also properly 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.

5.5 Quasi-Geostrophic Equations for Stratified Fluids

Let us now consider the appropriate equations for geostrophic flow at scales of order the deformation radius.

5.5.1 Scaling and Assumptions

The nondimensionalization and scaling are as before and so the nondimensional equations are those in the shaded box on the facing page. The Coriolis parameter is given by

$$\boldsymbol{f} = (f_0 + \beta y)\,\hat{\mathbf{k}}. \tag{5.53}$$

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

(i) The Rossby number is small, $Ro \ll 1$.

(ii) Length scales are of the same order as the deformation radius, $L \sim L_d$ or $L/L_d = \mathcal{O}(1)$.

(iii) 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:

$$\boldsymbol{u} = \boldsymbol{u}_g + \boldsymbol{u}_a, \quad \text{where} \quad f_0 \hat{\mathbf{k}} \times \boldsymbol{u}_g = -\nabla\phi \quad \text{and} \quad |\boldsymbol{u}_g| \gg |\boldsymbol{u}_a|. \tag{5.54}$$

Since the Coriolis parameter is constant in our definition of the geostrophic velocity the geostrophic divergence is zero; that is

$$\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} = 0. \tag{5.55}$$

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

$$\frac{\partial w}{\partial z} = -\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y}. \tag{5.56}$$

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,

$$W \ll \frac{UH}{L}. \tag{5.57}$$

The quasi-geostrophic equations are probably the most used set of equations in theoretical meteorology, certainly for midlatitude dynamics.

5.5.2 Derivation of Stratified QG Equations

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

$$\frac{\mathrm{D}}{\mathrm{D}t}(\zeta + f) = -(\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \left(\frac{\partial u}{\partial z}\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\frac{\partial w}{\partial x}\right). \tag{5.58}$$

We now apply the quasi-geostrophic assumption, namely:

(i) The geostrophic velocity and vorticity are much larger than their ageostrophic counterparts, and therefore we use geostrophic values for the terms on the left-hand side.

(ii) On the right-hand side we keep the horizontal divergence (which is small) only where it is multiplied by the big term f. Furthermore, because f is nearly constant we replace it with f_0.

(iii) The second term in large parentheses 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. We thus neglect it.

Given the above, (5.58) becomes

$$\frac{\mathrm{D}_g}{\mathrm{D}t}(\zeta_g + f) = -f_0\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = f_0\frac{\partial w}{\partial z}, \tag{5.59}$$

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, because w is so small.

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, (5.33) becomes,

$$\frac{D_g b'}{Dt} + wN^2 = 0. \tag{5.60}$$

We now eliminate w between (5.59) and (5.60), which, after a little algebra, yields

$$\frac{D_g q}{Dt} = 0, \qquad q = \zeta_g + f + \frac{\partial}{\partial z}\left(\frac{f_0 b'}{N^2}\right). \tag{5.61}$$

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

$$\boldsymbol{u}_g = \hat{\mathbf{k}} \times \nabla\psi, \qquad \zeta_g = \nabla^2\psi, \qquad b' = f_0 \partial\psi/\partial z. \tag{5.62}$$

Thus, we have, now omitting the subscript g,

$$\frac{Dq}{Dt} = 0, \qquad q = \nabla^2\psi + f + f_0^2 \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial\psi}{\partial z}\right). \tag{5.63a,b}$$

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

$$\frac{Dq}{Dt} = \frac{\partial q}{\partial t} + \boldsymbol{u}\cdot\nabla q = \frac{\partial q}{\partial t} + J(\psi, q). \tag{5.64}$$

where $J(\psi, q) = \partial\psi/\partial x\, \partial q/\partial y - \partial\psi/\partial y\, \partial q/\partial x$.

The quantity q is known as the *quasi-geostrophic potential vorticity,* 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 (5.63b).

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

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

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

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

By inspection of (5.63b) we see that a length scale $L_d = NH/f_0$, emerges naturally from the quasi-geostrophic 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 20 km. In the atmosphere, with

$N \approx 10^{-2}\,\mathrm{s}^{-1}$, $H \approx 10^4\,\mathrm{m}$, then $L_d \approx 1000\,\mathrm{km}$. 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

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \qquad q = \nabla^2\psi + f. \tag{5.65}$$

This is the two-dimensional vorticity equation; the high stratification of this limit has suppressed all vertical motion, and variations in the flow become confined to the horizontal plane.

Finally, if we allow density to vary in the vertical and carry through the quasi-geostrophic derivation we find that the quasi-geostrophic potential vorticity is given by

$$q = \nabla^2\psi + f + \frac{f_0^2}{\tilde{\rho}}\frac{\partial}{\partial z}\left(\frac{\tilde{\rho}}{N^2}\frac{\partial\psi}{\partial z}\right), \tag{5.66}$$

where $\tilde{\rho}$ is a reference profile of density and a function of z only, typically decreasing approximately exponentially with height.

5.5.3 Upper and Lower Boundary Conditions and Buoyancy Advection

The solution of the elliptic equation in (5.63b) requires vertical boundary conditions on ψ at the ground and at some upper boundary (e.g., the tropopause), 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

$$\frac{\mathrm{D}b'}{\mathrm{D}t} = 0, \qquad b' = f_0\frac{\partial\psi}{\partial z}, \qquad z = 0, H. \tag{5.67}$$

This equation provides the values of $\partial\psi/\partial z$ at the boundary that are then used when solving (5.63b). A heating term could be added as needed, and if there is no upper boundary (as in the real atmosphere) we would apply a condition that the motion decays to zero with height, but we will not deal with that condition further. If the bottom boundary is not flat then the vertical velocity is non-zero and the thermodynamic equation becomes, at the bottom boundary,

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{\partial\psi}{\partial z}\right) + wN^2 = 0, \qquad w = \boldsymbol{u}\cdot\nabla\eta_B, \tag{5.68}$$

where η_B is the height of the topography. In practice, the bottom boundary condition is often incorporated into the definition of potential vorticity, as we shall see in the two-level model.

5.6 The Two-Level Quasi-Geostrophic Equations

The continuously-stratified quasi-geostrophic equations have, as their name suggests, a continuous variation in the vertical. It turns out to be

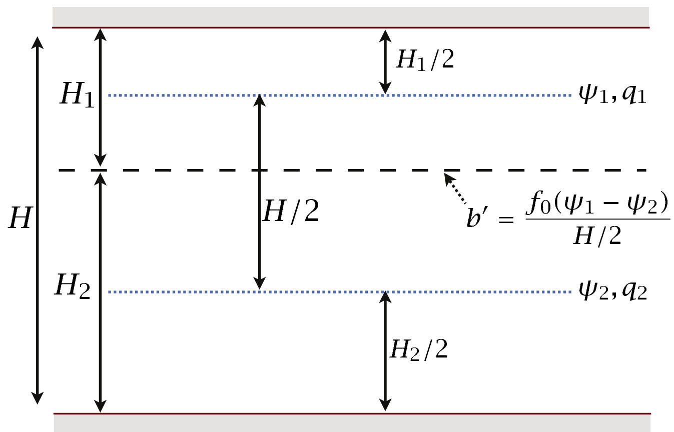

Fig. 5.1: A two-level quasi-geostrophic fluid system with a flat bottom and a rigid lid, with $w = 0$ at both. Here the levels are allowed to be of unequal thicknesses, but in the text we take $H_1 = H_2$.

extraordinarily useful to constrain this variation to that of two levels, in which the velocity is defined at just two vertical levels, and the temperature at just one level in the middle, but keeping the full horizontal variation. The resulting equations are not only algebraically much simpler than the continuous equations but they also capture many of their important features. There are two ways to derive the equations. One is by considering, *ab initio,* the motion of two immiscible shallow layers of fluid, one on top of the other, giving the 'two-layer' quasi-geostrophic equations. The second is by finite-differencing the continuous equations, and this is the procedure we follow here. Both methods give identical answers and we use 'two-layer' and 'two-level' interchangeably throught the book.

5.6.1 Deriving the Equations

The quasi-geostrophic potential vorticity may be written as

$$q = \nabla^2\psi + \beta y + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right). \tag{5.69}$$

We now apply this equation at the two levels, 1 and 2, in Fig. 5.1, taking N to be constant and $H_1 = H_2 = H/2$ for simplicity (these restrictions may be easily relaxed). Using simple expressions for vertical finite differences gives

$$q_1 = \nabla^2\psi_1 + \beta y + \frac{f_0^2}{N^2}\frac{1}{H/2}\left(\left(\frac{\partial \psi}{\partial z}\right)_{\text{top}} - \frac{\psi_1 - \psi_2}{H/2}\right). \tag{5.70}$$

In this equation $(\partial\psi/\partial z)_{\text{top}}$ is the value of $\partial\psi/\partial z$ at the top of the domain and $2(\psi_1 - \psi_2)/H$ is the value of $\partial\psi/\partial z$ in the middle of the domain. The value at the top is given by the thermodynamic equation, (5.67), and if there is no heating we set $\mathrm{D}/\mathrm{D}t\,(\partial\psi/\partial z)_{\text{top}} = 0$. The potential vorticity equation for the upper level is then given by

$$\frac{\mathrm{D}q_1}{\mathrm{D}t} = 0, \qquad q_1 = \nabla^2\psi_1 + \beta y + \frac{k_d^2}{2}(\psi_2 - \psi_1), \tag{5.71a,b}$$

where $k_d^2 = 8 f_0^2/N^2H^2$. In proceeding this way we have built the boundary conditions into the definition of potential vorticity. A thermody-

Quasi-Geostrophic Equations

Continuously-stratified

The adiabatic quasi-geostrophic potential vorticity equation for a Boussinesq fluid on the β-plane, is

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

where ψ is the streamfunction. The horizontal velocities are given by $(u, v) = (-\partial\psi/\partial y, \partial\psi/\partial x)$. The boundary conditions at the top and bottom are given by the buoyancy equation,

$$\frac{\mathrm{D}b}{\mathrm{D}t} = 0, \qquad b = f_0\frac{\partial\psi}{\partial z}. \tag{QG.2}$$

Two-level

The two-level or two-layer quasi-geostrophic equations are

$$\frac{\mathrm{D}q_i}{\mathrm{D}t} = 0, \qquad q_i = \nabla^2\psi_i + \beta y + \frac{k_d^2}{2}\left(\psi_j - \psi_i\right), \tag{QG.3}$$

where $i = 1, 2$, denoting the top and bottom levels respectively, and $j = 3 - i$.

namic source and/or frictional terms may readily be added to the right-hand of (5.71a) as needed.

Defining k_d so that $k_d^2/2$ rather than just k_d^2 is used in the two-level quasi-geostrophic potential vorticity is for later algebraic convenience rather than being of fundamental importance.

We apply an exactly analogous procedure to the lower level, and so obtain an expressions for the evolution of the lower layer potential vorticity, namely,

$$\frac{\mathrm{D}q_2}{\mathrm{D}t} = 0, \qquad q_2 = \nabla^2\psi_2 + \beta y + \frac{k_d^2}{2}\left(\psi_1 - \psi_2\right), \tag{5.72}$$

Once again the boundary conditions on buoyancy — the finite difference analogue of the buoyancy equations at the top and bottom — are *built in* to the definition of potential vorticity. If topography is present then it has an effect through the buoyancy equation at the bottom and it is incorporated into the definition of the lower level potential vorticity, which becomes

$$q_2 = \nabla^2\psi_2 + \beta y + \frac{k_d^2}{2}\left(\psi_1 - \psi_2\right) + \frac{f_0\eta_b}{H_2}. \tag{5.73}$$

This has a similar form to the shallow water potential vorticity with topography, as in (5.27). The two-level equations have proven to be enormously useful in the development of dynamical oceanography and meteorology and we will come back to them in later chapters. The quasi-geostrophic equations are summarized in the box above.

5.7 Frictional Geostrophic Balance and Ekman Layers

5.7.1 Preliminaries

Within a few hundred meters of the ground in the atmosphere, and in the upper and lower tens of metres in the ocean, frictional effects are important and geostrophic balance by itself is not the leading order balance in the momentum equations. Rather, a frictional term is important and the momentum equation becomes, for a Boussinesq fluid,

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi + \boldsymbol{F}, \tag{5.74}$$

where $\boldsymbol{F}$ is a frictional term and $\nabla\phi$ is the horizontal gradient of pressure, at constant z. The friction ultimately comes from molecular viscosity and is of the form $\boldsymbol{F} = \nu\nabla^2\boldsymbol{u}$. However, in geophysical settings the vertical derivative dominates and it is common to write (5.74) in the form

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla_z\phi + \frac{1}{\rho_0}\frac{\partial \boldsymbol{\tau}}{\partial z}, \tag{5.75}$$

where $\boldsymbol{\tau}$ is the stress, and since density is assumed constant we shall absorb it into the definition of $\boldsymbol{\tau}$. (The quantity $\boldsymbol{\tau}/\rho_0$ is the 'kinematic stress', but we shall just refer to it as the stress and also denote it $\boldsymbol{\tau}$. Also, in some other fluid dynamical contexts the stress is a tensor but here $\boldsymbol{\tau}$ is a vector.) Commonly, the regions where the stress term is important are *boundary layers* (see Fig. 5.2) and they exist because the fluid takes on values at the boundary that differ from the values in the fluid interior. Boundary layers are ubiquitous in fluids in many circumstances, and if the Coriolis term is important then the region is called an *Ekman layer.* As noted, the stress ultimately arises from molecular forces and $\boldsymbol{\tau} = \nu\partial\boldsymbol{u}/\partial z$ where ν is the coefficient of viscosity, again just including vertical derivatives. However, on the scale of the Ekman layer molecular effects are greatly amplified by the effects of small scale turbulence (as discussed more in Chapter 10) and we commonly replace ν by a much larger 'eddy viscosity', A.

In the atmosphere we imagine that the flow is nearly geostrophic above an Ekman layer, with frictional effects coming from the need to bring the speed of flow down from its high geostrophic value to zero at the ground. In the ocean, on the other hand, the stress largely arises from the wind in the atmosphere blowing over the ocean surface. The stress then decays to zero with depth, and in the deeper ocean the flow again is geostrophic. The stress itself is continuous across the ocean–atmosphere interface, usually with a maximum value at the interface, decaying with height in the atmosphere and with depth in the ocean.

The equations of motion of the Ekman layer are completed by the mass continuity equation, $\nabla\cdot\boldsymbol{v} = 0$, and the hydrostatic equation, $\partial\phi/\partial z = b$. In order to treat the simplest case we take buoyancy to be constant, and in that case, and without any additional loss of generality, we take $b = 0$. The horizontal gradient of pressure is independent of height, which implies that, even though velocity may vary rapidly in the boundary layer the pressure gradient force does not.

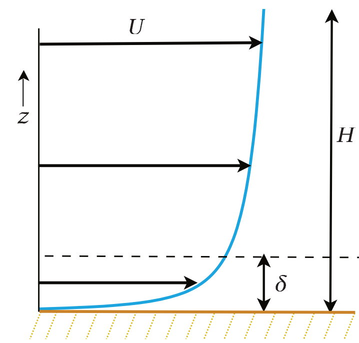

Fig. 5.2: An idealized boundary layer in which fields vary rapidly in order to satisfy the boundary conditions, here that $U = 0$. The scale δ is a measure of the thickness of the boundary layer and H is a typical scale of the variations away from the boundary.

Fridtjof Nansen (1861–1930), the polar explorer and statesman, apparently wanted to understand the motion of pack ice and of his ship, the Fram, *embedded in the ice and which seemed to move with neither the wind nor the surface current. He posed the problem to V. W. Ekman (1874–1954) who then calculated the direction of the flow in the upper ocean and published a paper on the matter in 1905.*

5.7.2 Properties of Ekman Layers

We shall now determine three important properties of Ekman layers:

(i) The transport in an Ekman layer.

(ii) The depth of the Ekman layer.

(iii) The vertical velocity at the edge of the Ekman layer.

The importance of these will become clear as we go on.

Transport

We may write the Ekman layer equation (5.74) as

$$\boldsymbol{f} \times (\boldsymbol{u} - \boldsymbol{u}_g) = \frac{\partial \boldsymbol{\tau}}{\partial z} \qquad \text{or equivalently} \qquad f\boldsymbol{u}_a = -\hat{\mathbf{k}} \times \frac{\partial \boldsymbol{\tau}}{\partial z}, \qquad (5.76\text{a,b})$$

where $\boldsymbol{u}_g$ is the geostrophic wind such that $\boldsymbol{f} \times \boldsymbol{u}_g = -\nabla\phi$, $\boldsymbol{u}_a = \boldsymbol{u} - \boldsymbol{u}_g$ is the ageostrophic wind and $\hat{\mathbf{k}}$ is the unit vector in the vertical (so that $\boldsymbol{f} = f\hat{\mathbf{k}}$). Consider an ocean with a stress that diminishes with depth from its surface values, $\boldsymbol{\tau}_0$. If we integrate (5.76) over the depth of the Ekman layer (i.e., down from the surface until the stress vanishes) we obtain

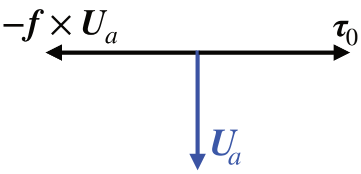

Fig. 5.3: Vertically integrated transport in a Northern Hemisphere ocean Ekman layer. A stress $\boldsymbol{\tau}_0$ at the surface induces an ageostrophic flow, $\boldsymbol{U}_a$. In order that the Coriolis force on the ageostrophic flow can balance the surface stress the ageostrophic flow must be at right angles to the stress, as illustrated.

$$\int f\boldsymbol{u}_a \,\mathrm{d}z = -\hat{\mathbf{k}} \times \boldsymbol{\tau}_0, \quad \text{or} \quad \int f u_a \,\mathrm{d}z = \tau_0^y, \quad \int f v_a \,\mathrm{d}z = -\tau_0^x, \qquad (5.77)$$

where the superscripts denote the x and y components of the stress. That is to say, *the integrated ageostrophic transport in the Ekman layer,* $\boldsymbol{U}_a = \int \boldsymbol{u}_a \,\mathrm{d}z$, *is at right angles to the stress at the surface,* as in Fig. 5.3. This result is independent of the detailed form of the stress, and it gives a partial answer to Nansen's question — if the surface stress is in the direction of the wind, the water beneath it flows at an angle to it, as a consequence of the Coriolis force. More generally, the Ekman transport is the most immediate response of the ocean to the atmosphere and is thus of key importance in the ocean circulation. In the atmosphere we *live* in the Ekman layer, and Ekman layer effects are partly responsible for determining the surface wind.

Depth

Although the transport in the Ekman layer is independent of the detailed form of the stress, the depth — namely the extent of the frictional influence of the surface — is not and to calculate it we use the Ekman-layer equation in the form

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi + A\frac{\partial^2 \boldsymbol{u}}{\partial z^2}, \qquad (5.78)$$

where A is a coefficient of viscosity. Suppose for the moment that the vertical scale of the flow is H; the ratio of size of the frictional term to the Coriolis term is then (A/fH^2), and this ratio defines the Ekman number, Ek:

$$Ek \equiv \left(\frac{A}{fH^2}\right). \qquad (5.79)$$

In the fluid interior the Ekman number is usually small. In the Ekman layer itself, however, the viscous terms must be large (otherwise we are not in the Ekman layer), and the vertical scale must therefore be smaller than H. If the viscous and Coriolis terms in (5.78) are required to be the same magnitude we obtain an Ekman layer depth, δ, of

$$\delta \sim \left(\frac{A}{f}\right)^{1/2}. \tag{5.80}$$

The Ekman layer depth decreases with viscosity, becoming a thin boundary layer as $A \to 0$. We also see that the Ekman number can be recast as the ratio of the Ekman layer depth to the vertical scale in the interior; that is, from (5.79) and (5.80), $Ek = (\delta/H)^2$.

Vertical velocity

The vertical velocity at the edge of the Ekman layer may be calculated using the mass continuity equation,

$$\frac{\partial w}{\partial z} = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right). \tag{5.81}$$

The right-hand side may be calculated from (5.76) and, noting that divergence of the geostrophic velocity is given by $\nabla \cdot \boldsymbol{v}_g = -\beta v_g / f$, we find that

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = -\frac{\beta v_g}{f} + \frac{\partial}{\partial z}\mathrm{curl}_z\left(\frac{\boldsymbol{\tau}}{f}\right), \tag{5.82}$$

where $\mathrm{curl}_z(\boldsymbol{\tau}/f) = \partial_x(\tau^y/f) - \partial_y(\tau^x/f)$. Consider an ocean Ekman layer in which the vertical velocity is zero at the surface and the stress is zero at the bottom of the layer. If we integrate (5.81) over the depth of that layer and use (5.82) we obtain

$$w_E = \mathrm{curl}_z \frac{\boldsymbol{\tau}_0}{f} - \frac{\beta V_g}{f}. \tag{5.83}$$

where $V_g = \int v_g \, dz$ is the integral of the meridional geostrophic velocity over the Ekman layer. Evidently, 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.

This result is particularly useful for the top Ekman layer in the ocean, where the stress can be regarded as a function of the overlying wind and is largely independent of the flow in the ocean; an equivalent result applies in the atmosphere, but here the surface stress is a function of the geostrophic wind in the free atmosphere. If the curl of the wind stress at the top of the ocean is anticyclonic (i.e., negative if $f > 0$, positive if $f < 0$) then a downward velocity is induced at the base of the Ekman layer, a phenomenon known as *Ekman pumping,* and this is particularly important in setting the structure of the subtropical gyres in the ocean, discussed in Chapter 14.

Notes and References

The first systematic derivation of the quasi-geostrophic equations was given by Charney (1948), although aspects of the concept, and the words'quasi-geostrophy', appeared in Durst & Sutcliffe (1938) and Sutcliffe (1947). The various forms of geostrophic equations were brought together in a review article by Phillips (1963) and since then both the quasi-geostrophic and planetary-geostrophic equations have been staples of dynamical meteorology and dynamical oceanography.

Problems

5.1 Do either or both:

(a) Carry through the derivation of the quasi-geostrophic system starting with the anelastic equations and obtain (5.66).

(b) Carry through the derivation of the quasi-geostrophic system in pressure coordinates.

In each case, state the differences between your results and the Boussinesq result.

5.2 (a) The shallow water *planetary geostrophic* equations may be derived by simply omitting ζ in the equation

$$\frac{\mathrm{D}}{\mathrm{D}t}\frac{\zeta + f}{h} = 0 \tag{P5.1}$$

by invoking a small Rossby number, so that ζ/f is small. We then relate the velocity field to the height field by hydrostatic balance and obtain:

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{f}{h}\right) = 0, \qquad fu = -g\frac{\partial h}{\partial y}, \quad fv = g\frac{\partial h}{\partial x}. \tag{P5.2}$$

The assumptions of hydrostatic balance and small Rossby number are the same as those used in deriving the quasi-geostrophic equations. Explain nevertheless how some of the assumptions used for quasi-geostrophy are in fact different from those used for planetary-geostrophy, and how the derivations and resulting systems differ from each other. Use any or all of the momentum and mass continuity equations, scaling, nondimensionalization and verbal explanations as needed.

(b) Explain if and how your arguments in part (a) also apply to the stratified equations (using, for example, the Boussinesq equations or pressure coordinates).

5.3 In the derivation of the quasi-geostrophic equations, geostrophic balance leads to the lowest-order horizontal velocity being divergence-free — that is, $\nabla \cdot \boldsymbol{u}_0 = 0$. It seems that this can also be obtained from the mass conservation equation at lowest order. Is this a coincidence? Suppose that the Coriolis parameter varied, and that the momentum equation yielded $\nabla \cdot \boldsymbol{u}_0 \neq 0$. Would there be an inconsistency?

5.4 Consider a wind stress imposed by a mesoscale cyclonic storm (in the atmosphere) given by

$$\boldsymbol{\tau} = -A\mathrm{e}^{-(r/\lambda)^2}(y\,\hat{\mathbf{i}} - x\,\hat{\mathbf{j}}), \tag{P5.3}$$

where $r^2 = x^2 + y^2$, and A and λ are constants. Also assume constant Coriolis gradient $\beta = \partial f/\partial y$ and constant ocean depth H. In the ocean, find

(a) the Ekman transport, (b) the vertical velocity $w_E(x, y, z)$ below the Ekman layer, (c) the northward velocity $v(x, y, z)$ below the Ekman layer and (d) indicate how you would find the westward velocity $u(x, y, z)$ below the Ekman layer.

5.5 In an atmospheric Ekman layer on the f-plane let us write the momentum equation as

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi + \frac{1}{\rho_a}\frac{\partial \boldsymbol{\tau}}{\partial z}, \tag{P5.4}$$

where $\boldsymbol{\tau} = A\rho_a \partial\boldsymbol{u}/\partial z$ and A is a constant eddy viscosity coefficient. An *independent* formula for the stress at the ground is $\boldsymbol{\tau} = C\rho_a\boldsymbol{u}$, where C is a constant. Let us take $\rho_a = 1$, and assume that in the free atmosphere the wind is geostrophic and zonal, with $\boldsymbol{u}_g = U\hat{\mathbf{i}}$.

(a) Find an expression for the wind vector at the ground. Discuss the limits $C = 0$ and $C = \infty$. Show that when $C = 0$ the frictionally-induced vertical velocity at the top of the Ekman layer is zero.

(b) Find the vertically integrated horizontal mass flux caused by the boundary layer.

(c) When the stress on the atmosphere is $\boldsymbol{\tau}$, the stress on the ocean beneath is also $\boldsymbol{\tau}$. Why? Show how this is consistent with Newton's third law.

(d) Determine the direction and strength of the surface current, and the mass flux in the oceanic Ekman layer, in terms of the geostrophic wind in the atmosphere, the oceanic Ekman depth and the ratio ρ_a/ρ_o, where ρ_o is the density of the seawater. Include a figure showing the directions of the various winds and currents. How does the boundary-layer mass flux in the ocean compare to that in the atmosphere? (Assume, as needed, that the stress in the ocean may be parameterized with an eddy viscosity.)

Partial solution for (a): A useful trick in Ekman layer problems is to write the velocity as a complex number, $\hat{u} = u + \mathrm{i}v$ and $\hat{u}_g = u_g + \mathrm{i}v_g$. The fundamental Ekman layer equation may then be written as

$$A\frac{\partial^2 \widehat{U}}{\partial z^2} = \mathrm{i} f \widehat{U}, \tag{P5.5}$$

where $\widehat{U} = \hat{u} - \hat{u}_g$. The solution to this is

$$\hat{u} - \hat{u}_g = [\hat{u}(0) - \hat{u}_g] \exp\left[-\frac{(1+\mathrm{i})z}{d}\right], \tag{P5.6}$$

where $d = \sqrt{2A/f}$ and the boundary condition of finiteness at infinity eliminates the exponentially growing solution. The boundary condition at $z = 0$ is $\partial\hat{u}/\partial z = (C/A)\hat{u}$; applying this gives $[\hat{u}(0) - \hat{u}_g]\exp(\mathrm{i}\pi/4) = -Cd\hat{u}(0)/(\sqrt{2}A)$, from which we obtain $\hat{u}(0)$, and the rest of the solution follows.

Chapter

6

Rossby Waves

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

6.1 Fundamentals and Formalities

6.1.1 Definitions and Kinematics

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

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

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

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

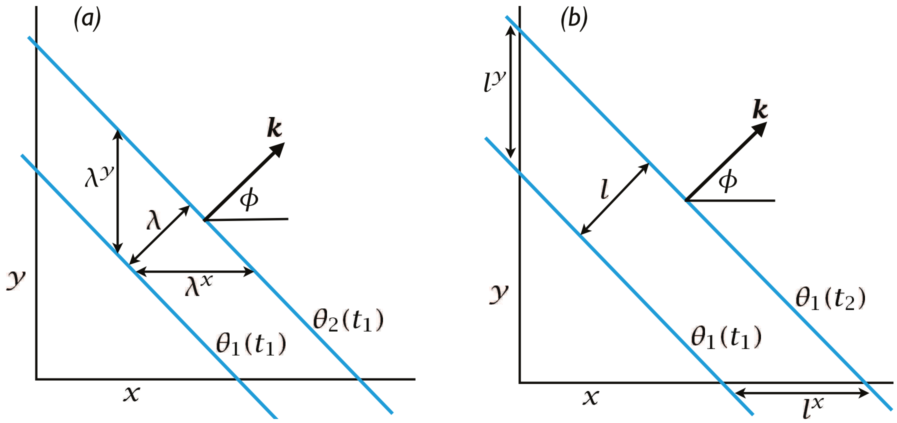

Fig. 6.1: The propagation of a two-dimensional wave. (a) Two lines of constant phase (e.g., two wavecrests) at a time t_1. The wave is propagating in the direction $\boldsymbol{k}$ with wavelength λ. (b) A line of constant phase at two successive times. The phase speed is the speed of advancement of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

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

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

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

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

6.1.2 Wave Propagation and Phase Speed

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

Phase speed

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

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

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

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

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

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

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

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

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

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

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

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

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

6.1.3 The Dispersion Relation

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

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

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

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

Wave Fundamentals

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

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

  where the second example gives rise to Rossby waves.

- Solutions to the governing equation are often sought in the form of plane waves that have the form

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

  where A is the wave amplitude, $\boldsymbol{k} = (k, l, m)$ is the wavevector, and ω is the frequency.

- The dispersion relation connects the frequency and wavevector through an equation of the form $\omega = \Omega(\boldsymbol{k})$ where Ω is some function. The relation is normally derived by substituting a trial solution like (WF.2) into the governing equation. For the examples of (WF.1) we obtain $\omega = c^2 K^2$ and $\omega = -\beta k / K^2$ where $K^2 = k^2 + l^2 + m^2$ or, in two dimensions, $K^2 = k^2 + l^2$.

- The phase speed is the speed at which the wave crests move. In the direction of propagation and in the x, y and z directions the phase speeds are given by, respectively,

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

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

- The group velocity is the velocity at which a wave packet or wave group moves. It is a vector and is given by

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

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

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

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

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

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

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

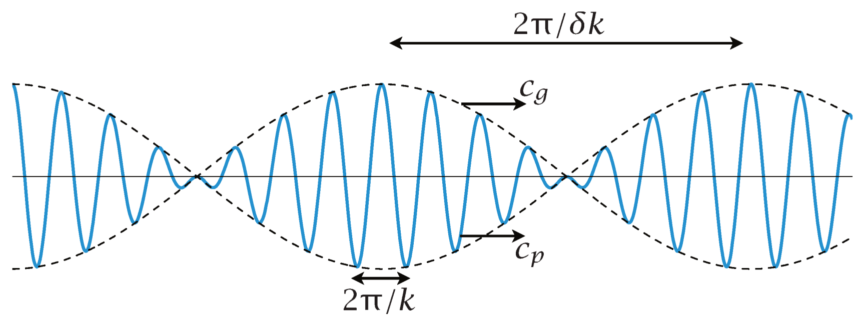

Fig. 6.2: Superposition of two sinusoidal waves with wavenumbers k and $k + \delta k$, producing a wave (solid line) that is modulated by a slowly varying wave envelope or packet (dashed). The envelope moves at the group velocity, $c_g = \partial\omega/\partial k$, and the phase moves at the group speed,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

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

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

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

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

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

6.2 Group Velocity

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

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

6.2.1 Superposition of Two Waves

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

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

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, (6.15) becomes

$$\begin{aligned}\psi &= \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx-\omega t)}[\mathrm{e}^{\mathrm{i}(\Delta k x-\Delta\omega t)} + \mathrm{e}^{-\mathrm{i}(\Delta k x-\Delta\omega t)}] \\ &= 2\,\mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx-\omega t)}\cos(\Delta k\, x - \Delta\omega\, t).\end{aligned} \tag{6.16}$$

The resulting disturbance, illustrated in Fig. 6.2 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 moves 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

$$\boldsymbol{c}_g = \frac{\partial\omega}{\partial\boldsymbol{k}} \equiv \left(\frac{\partial\omega}{\partial k}, \frac{\partial\omega}{\partial l}, \frac{\partial\omega}{\partial m}\right). \tag{6.17}$$

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.

The above derivation can be generalized to apply to the superposition of many waves. It also applies when the medium through which the waves propagate is not homogeneous, provided that changes occur on a longer space scale than the wavelength of the waves. Energy and most other physically meaningful properties of waves travel at the group velocity, as explicitly shown for Rossby waves in Section 9.2.2.

The group velocity, $\boldsymbol{c}_g$ is given by $\boldsymbol{c}_g = \partial\omega/\partial\boldsymbol{k}$ or, in Cartesian components, $\boldsymbol{c}_g = (\partial\omega/\partial k, \partial\omega/\partial l, \partial\omega/\partial m)$. It is a vector, and it gives the velocity at which wave packets, and hence energy, are propagated. The phase speed, $c_p = \omega/k$, is the speed at which the phase of an individual wave moves. For some waves the group velocity and phase speed may be in opposite directions, so that the wave crests appear to move backwards through the wave packets.

6.3 Rossby Wave Essentials

Rossby waves are the most prominent wave in the atmosphere and ocean on large scales, although gravity waves sometimes rival them. They can best be described using the quasi-geostrophic equations, as follows.

6.3.1 The Linear Equation of Motion

The relevant equation of motion is the inviscid, adiabatic potential vorticity equation in the quasi-geostrophic system, as discussed in Chapter 5, namely

$$\frac{\partial q}{\partial t} + \boldsymbol{u}\cdot\nabla q = 0, \tag{6.18}$$

where $q(x, y, z, t)$ is the potential vorticity and $\boldsymbol{u}(x, y, z, t)$ is the horizontal velocity. The velocity is in turn related to a streamfunction by $u = -\partial\psi/\partial y$, $v = \partial\psi/\partial x$, and the potential vorticity is some function

The Rossby wave, the most well-known wave in meteorology and oceanography, is named for Carl-Gustav Rossby who described the essential dynamics in Rossby (1939). The waves are in fact present in the solutions of Hough (1898) although in a rather oblique form and Rossby is normally given credit for having discovered them.

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, are

$$q = f + \zeta + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right) \qquad \text{and} \qquad q = \zeta + f - k_d^2\psi, \tag{6.19a,b}$$

Definitions of k_d and L_d differ from source to source (and book to book), often by factors of 2 or π or similar. Caveat lector.

where $\zeta = \nabla^2\psi$ is the relative vorticity and $k_d = 1/L_d$ is the inverse radius of deformation for a shallow water system.

We now *linearize* (6.18); 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 can write

$$q = \overline{q}(y, z) + q'(x, y, t), \qquad \psi = \overline{\psi}(y, z) + \psi'(x, y, z, t), \tag{6.20}$$

with a similar notation for the other variables, and $\overline{u} = -\partial\overline{\psi}/\partial y$ and $\overline{v} = 0$. Substituting into (6.18) gives, without approximation,

$$\frac{\partial q'}{\partial t} + \overline{\boldsymbol{u}} \cdot \nabla\overline{q} + \overline{\boldsymbol{u}} \cdot \nabla q' + \boldsymbol{u}' \cdot \nabla\overline{q} + \boldsymbol{u}' \cdot \nabla q' = 0. \tag{6.21}$$

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 $\overline{v} = 0$ (since $\int \partial\psi/\partial x \, dx = 0$) and $\partial\overline{q}/\partial x = 0$ we obtain

$$\frac{\partial q'}{\partial t} + U\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = 0, \tag{6.22}$$

where $U \equiv \overline{u}$. This equation or one very similar appears very commonly in studies of Rossby waves. Let us first consider the simple example of waves in a single layer.

6.3.2 Waves in a Single Layer

Consider a system obeying (6.18) and (6.19b). The dynamics are more easily illustrated on a 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 vorticity equation may be written as

$$\frac{\partial \zeta}{\partial t} + \boldsymbol{u} \cdot \nabla\zeta + \beta v = 0. \tag{6.23}$$

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

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

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

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

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

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

Substituting (6.25) into (6.24) yields

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

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

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

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

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

and

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

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

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

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

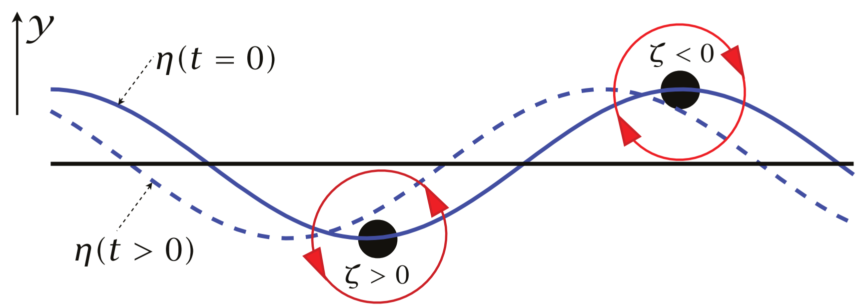

Fig. 6.3: A two-dimensional (x–y) Rossby wave. An initial disturbance displaces a material line at constant latitude (the straight horizontal line) to the solid line marked $\eta(t = 0)$. Conservation of potential vorticity, $\beta y + \zeta$, leads to the production of relative vorticity, ζ, as shown. The associated velocity field (arrows on the circles) then advects the fluid parcels, and the material line evolves into the dashed line with the phase propagating westward.

an eastward group velocity, and this has implications for the 'downstream development' of wave packets, but we will not pursue that topic here.

We see from (6.29) the group velocity is negative (westward) if the x-wavenumber is sufficiently small compared to the y-wavenumber. Essentially, *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.

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 $\nabla q = (\beta + Uk_d^2)\hat{\mathbf{j}}$. The linearized equation of motion is thus

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)(\nabla^2\psi' - \psi' k_d^2) + (\beta + Uk_d^2)\frac{\partial \psi'}{\partial x} = 0. \tag{6.31}$$

Substituting $\psi' = \widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}$ we obtain the dispersion relation,

$$\omega = \frac{k(UK^2 - \beta)}{K^2 + k_d^2} = Uk - k\frac{\beta + Uk_d^2}{K^2 + k_d^2}. \tag{6.32}$$

It is clear from the second form of the above equation that the uniform velocity field no longer provides just a simple Doppler shift of the frequency, nor does it provide a uniform addition to the phase speed. This 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, as explored further in Problem 6.1.

6.3.3 The Mechanism of Rossby Waves

The fundamental mechanism underlying Rossby waves may be understood as follows. 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. 6.3. In the displacement, the potential vorticity of the fluid parcels is conserved, and in the simplest

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 in this case the associated Rossby waves are called *planetary waves*. In the presence of non-zero β the ambient potential vorticity increases northward and the phase of the Rossby waves propagates westward. Topography is another source of potential vorticity gradients. In general, Rossby waves propagate to the left of the direction of increasing potential vorticity.

- A common equation of motion for Rossby waves is

$$\frac{\partial q'}{\partial t} + U\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = 0, \qquad \text{(RW.1)}$$

  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

$$\frac{\partial}{\partial t}(\nabla^2 + k_d^2)\psi' + \beta\frac{\partial \psi'}{\partial x} = 0, \qquad \text{(RW.2)}$$

  with dispersion relation

$$\omega = \frac{-\beta k}{k^2 + l^2 + k_d^2}\,. \qquad \text{(RW.3)}$$

- In the absence of a mean flow (i.e., $U = 0$), 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.3) the components of the group velocity are given by

$$c_g^x = \frac{\beta(k^2 - l^2 - k_d^2)}{(k^2 + l^2 + k_d^2)^2}\,, \qquad c_g^y = \frac{2\beta k l}{(k^2 + l^2 + k_d^2)^2}\,. \qquad \text{(RW.4)}$$

  The group velocity is westward if the zonal wavenumber is sufficiently small, and eastward if the zonal wavenumber is sufficiently large.

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 the material line in the manner shown, and the wave propagates westward.

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.

6.4 Rossby Waves in Stratified Quasi-Geostrophic Flow

We 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, (5.63), and linearizing this about a constant mean zonal flow, U, gives

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\left[\nabla^2\psi' + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial\psi'}{\partial z}\right)\right] + \beta\frac{\partial\psi'}{\partial x} = 0, \tag{6.33}$$

where, for simplicity, we suppose that f_0^2/N^2 does not vary with z.

6.4.1 Dispersion Relation and Group Velocity

Deferring the issues of boundary conditions for a moment, let us seek solutions of the form

$$\psi' = \operatorname{Re}\widetilde{\psi}e^{i(kx+ly+mz-\omega t)}, \tag{6.34}$$

where $\widetilde{\psi}$ is a constant. Substituting (6.34) into (6.33) gives, after a couple of lines of elementary algebra, the dispersion relation

$$\omega = Uk - \frac{\beta k}{k^2 + l^2 + m^2 f_0^2/N^2}\,. \tag{6.35}$$

As in most wave problems the frequency is, by convention, a positive quantity.

In reality we usually have to satisfy a boundary condition at the top and bottom of the fluid and these are determined by the thermodynamic equation, (5.67), which after linearization becomes

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\left(\frac{\partial\psi'}{\partial z}\right) + N^2 w = 0. \tag{6.36}$$

If the boundaries are flat, rigid, surfaces then $w = 0$ at those boundaries suggesting that instead of (6.34) we choose a streamfunction of the form

$$\psi' = \operatorname{Re}\widetilde{\psi}e^{i(kx+ly-\omega t)}\cos\left(\frac{m_j\pi z}{H}\right), \qquad m_j = 1, 2\ldots, \tag{6.37}$$

which satisfies $\partial\psi'/\partial z = 0$ at $z = 0, H$. The dispersion relation is the same as (6.35) with m equal to $m_j\pi/H$. If the domain is finite in the horizontal direction also — as it nearly always is — then the horizontal wavenumbers are also quantized. For example, if the extent of the domain in the

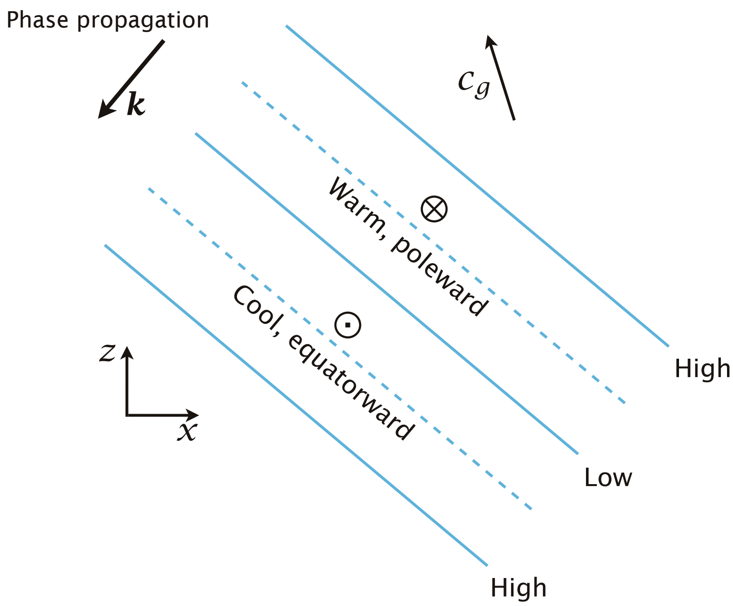

Fig. 6.4: A schematic east-west section of an upwardly propagating Rossby wave. The slanting lines are lines of constant phase and 'high' and 'low' refer to the pressure or streamfunction values. Both k and m are negative so the phase lines are oriented up and to the west. The phase propagates westward and downward, but the group velocity is upward.

x-direction is L_x and the flow is periodic in that direction, then k is restricted to the values $k = 2\pi k_j/L_x$ where k_j is an integer, and similarly for l and l_j. (It is common albeit sometimes confusing to refer to both k and k_j as the wavenumber and also to drop the subscript j on k_j with the distinction then made clear by context, and similarly for l and l_j.)

Using (6.35), the three components of the group velocity for these waves are:

$$c_g^x = U + \frac{\beta[k^2 - (l^2 + m^2 f_0^2/N^2)]}{\left(k^2 + l^2 + m^2 f_0^2/N^2\right)^2}, \tag{6.38a}$$

$$c_g^y = \frac{2\beta kl}{\left(k^2 + l^2 + m^2 f_0^2/N^2\right)^2}, \qquad c_g^z = \frac{2\beta km f_0^2/N^2}{\left(k^2 + l^2 + m^2 f_0^2/N^2\right)^2}. \tag{6.38b,c}$$

The propagation in the horizontal is similar to the propagation in a single-layer model. We see 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 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! This case is illustrated in Fig. 6.4 and to understand it better let us look at the vertical propagation in more detail.

Vertical propagation of Rossby waves was not discussed by Rossby, who was mainly interested in their properties in a horizontal plane. Rather, it was Charney & Drazin (1961) who first considered the vertical propagation of the waves in a stratified fluid in any detail.

6.4.2 Vertical Propagation of Rossby Waves

Conditions for wave propagation

With a little re-arrangement the dispersion relation, (6.35), may be written as

$$m^2 = \frac{N^2}{f_0^2}\left(\frac{\beta}{U - c} - (k^2 + l^2)\right), \tag{6.39}$$

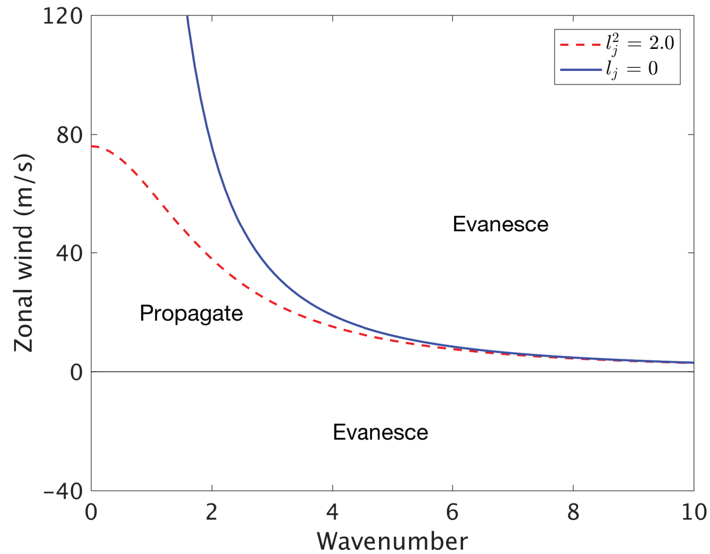

Fig. 6.5: The boundary between propagating waves and evanescent waves as a function of zonal wind and wavenumber, calculated using (6.41). The abscissa is the x-wavenumber, k_j, and the results show the boundary with two values of meridional wavenumber, l_j. Small wavenumbers (large scales) can more easily propagate into the stratosphere, and propagation is inhibited if the zonal wind is too strong or is negative.

where $c = \omega/k$. For waves to propagate upwards we require that $m^2 > 0$: if $m^2 < 0$ the wavenumber is imaginary and the wave amplitude either grows with height (which is unphysical) or is damped. This condition implies

$$0 < U - c < \frac{\beta}{k^2 + l^2}. \tag{6.40}$$

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

Stationary, vertically oscillatory modes can exist only for zonal flows that are eastward and that are less than some critical value $U_c = \beta/(k^2 + l^2)$.

$$0 < U < \frac{\beta}{k^2 + l^2}, \tag{6.41}$$

as illustrated in Fig. 6.5. That is to say, the vertical propagation of stationary Rossby waves occurs only in eastward 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. The lower limit, where $U = c$ (and so $U = 0$ if stationary) is known as a critical level and the upper limit is a 'turning level'. Since (6.39) is just a form of the dispersion relation, for any given frequency there is a criterion for propagation given by (6.40), and for any given U there may be a frequency that allows propagation. However, the zero frequency case is often regarded as the most important.

The upper critical velocity, $U_c = \beta/(k^2 + l^2)$, is a function of wavenumber and 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'. An important consequence 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 polar regions in the stratosphere are warmer than equatorial regions) 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.

For westward flow, or for sufficiently strong eastward flow, m is imaginary and the waves decay exponentially in the vertical as $\exp(-\alpha z)$ where

$$\alpha = \frac{N}{f_0}\left(k^2 + l^2 - \frac{\beta}{U}\right)^{1/2}, \tag{6.42}$$

and the decay scale is smaller for shorter waves.

An interpretation

One physical way to interpret the propagation criterion is to write (6.41) as

$$0 < |Uk| < \frac{|\beta k|}{k^2 + l^2}, \tag{6.43}$$

allowing for the fact that k may be negative. Now, in a resting medium ($U = 0$) the Rossby wave frequency has a maximum value when $m = 0$ given by

$$\omega = \frac{|\beta k|}{k^2 + l^2}, \tag{6.44}$$

and the minimum frequency is zero. Let us suppose that the stationary Rossby waves are excited by flow of speed U moving over bottom topography. If we move into the frame of reference of the flow then the waves are forced by a moving topography, and the frequency of the forcing is just $|Uk|$. Thus, (6.43) is equivalent to saying that for oscillatory waves to exist *the forcing frequency, $|Uk|$, must lie within the frequency range of vertically propagating Rossby waves.*

6.4.3 Heat Transport and Vertical Propagation

If the group velocity in the z-direction, given by (6.38) is to be positive, then we require the product $km > 0$. This has an important consequence 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

$$H = \overline{vb} = f_0 \overline{\frac{\partial\psi}{\partial z}\frac{\partial\psi}{\partial x}}, \tag{6.45}$$

where an overbar denotes a zonal average. To evaluate this expression it is simplest to suppose that $\psi = \psi_0 \cos(kx + ly + mz)$, where ψ_0 is a real constant, in which case we obtain

$$H = f_0 k m \psi_0^2 \overline{\sin^2(kx + ly + mz)} = \frac{1}{2} f_0 \psi_0^2 k m, \tag{6.46}$$

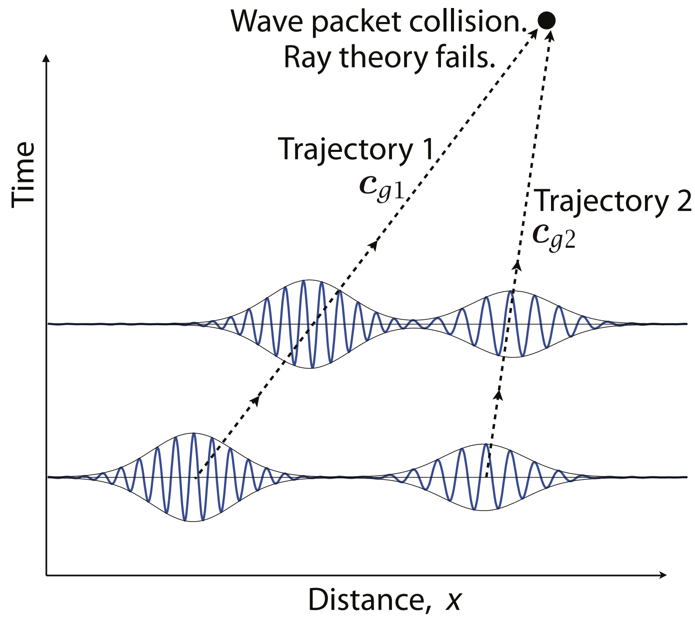

Fig. 6.6: Idealised trajectory of two wavepackets, each with a different wavelength and moving with a different group velocity, as might be calculated using ray theory. If the wave packets collide ray theory must fail. Ray theory gives only the trajectory of the wave packet, not the detailed structure of the waves within a packet.

using the standard result that the average of $\sin^2 x$ over a wavelength is equal to $1/2$. Thus, the heat flux, like the vertical component of the group velocity, is equal to km multiplied by a positive quantity, and we may conclude that an upward propagation of Rossby waves is associated with a polewards heat flux. A moment's thought will reveal that the same conclusion holds in the Southern Hemisphere, even though f_0 is negative there.

6.5 Ray Theory and Rossby Rays

In the real world most waves propagate in a medium that is inhomogeneous; for example, in the Earth's atmosphere and ocean the Coriolis parameter varies with latitude and the stratification with height. In these cases it can be hard to obtain the solution of a wave problem by Fourier methods, even approximately. Nonetheless, the idea of signals propagating at the group velocity is a robust one, and we can often obtain some of the information we want — and in particular the trajectory of a wave — using a recipe known as *ray theory*. If the background properties of the medium vary only slowly compared to the wavelength, we assume that the dispersion relation is satisfied *locally.* We then calculate the group velocity, and assume that the trajectory of the wave is along that group velocity.

To implement this recipe, we first obtain a dispersion relation from the governing equation. In the homogeneous case we obtained a dispersion relation $\omega = f(\boldsymbol{k})$; that is the frequency is some function only of the wavenumber. In the inhomogeneous case we proceed the same way, and obtain a dispersion relation *as if* the parameters were fixed, but we then allow the parameters to vary in the dispersion relation; that is, we obtain a dispersion relation of the form $\omega = f(\boldsymbol{k}; \boldsymbol{x}, t)$. For example, for barotropic

Rossby waves, the dispersion relation might now be

$$\omega = U(y) - \frac{\beta(y)}{k^2 + l^2}. \tag{6.47}$$

This is the same as the usual dispersion relation except that U and β are now taken to be slowly varying functions of position. This procedure is valid provided that U and β vary more slowly than the wavelengths of the Rossby waves, which is essentially the same condition used in the WKB approximation, discussed in the appendix below.

The utility of ray theory is that it allows us to calculate the trajectory of a wave packet without fully solving the problem. Thus, suppose that a disturbance somewhere generates Rossby waves — perhaps a sea-surface temperature anomaly in the tropics. The waves propagate away from the disturbance following the group velocity, which we calculate using the dispersion relation. Thus, if the dispersion relation is given by (6.47) then the group velocity is given by

$$c_g^x = U + \frac{\beta(k^2 - l^2)}{(k^2 + l^2)^2}, \qquad c_g^y = \frac{2kl\beta}{(k^2 + l^2)^2}, \tag{6.48a,b}$$

where β and U are functions of space. Thus, the position of the wave packet, $\boldsymbol{x}$, is given by solving (often numerically) the ray equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = c_g^x, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = c_g^y, \tag{6.49}$$

where c_g^x and c_g^y are evaluated using (6.48). Figure 6.6 shows schematically how a wave packet might then propagate. In the Rossby-wave case here, the frequency is constant (determined by the source of the waves), as is the x-wavenumber, k, because U and β are only functions of y. As we move along the ray, we recalculate the y-wavenumber, l, at each point using the dispersion relation and then calculate the group velocity at each point, and thence the trajectory. When this is done on Earth, one sometimes can see Rossby waves propagating from the tropics to midlatitude that bring about long-range correlations in weather patterns known as 'teleconnections'.

✦ Appendix A: The WKB Approximation

The WKB method is a way of finding approximate solutions to certain linear differential equations in which the term with the highest derivative is multiplied by a small parameter. In particular, WKB theory can be used to find approximate solutions to wave equations in which the coefficients vary slowly in space or time. Consider an equation of the form

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}z^2} + m^2(z)\xi = 0. \tag{6.50}$$

Such an equation commonly arises in wave problems. If m^2 is positive the equation has wavelike solutions, and if m is constant the solution has the

In 1926, G. Wentzel, H. A. Kramers and L. Brillouin each presented a method to find approximate solutions of the Schrödinger equation in quantum mechanics, and the technique became known as the WKB method. It turns out that the technique had already been discovered by Harold Jeffreys, a mathematical geophysicist, a few years prior to that, and in fact the origins of the method go back to the early nineteenth century. Evidently, methods are named after the last people to discover them.

harmonic form

$$\xi = \mathrm{Re}\, A_0 \mathrm{e}^{imz}, \tag{6.51}$$

where A_0 is a complex constant. If m varies only slowly with z (meaning that the variations in m only occur on a scale much longer than $1/m$) one might reasonably expect that the harmonic solution above would provide a decent first approximation; that is, we expect the solution to locally look like a plane wave with local wavenumber $m(z)$. However, we might also expect that the solution would not be *exactly* of the form $\exp(\mathrm{i}m(z)z)$, because the phase of ξ is $\theta(z) = mz$, so that $\mathrm{d}\theta/\mathrm{d}z = m + z\mathrm{d}m/\mathrm{d}z \neq m$. Thus, in (6.51) m is not the wavenumber unless m is constant.

The condition that variations in m, or in the wavelength $\lambda \sim m^{-1}$, occur only slowly may be variously expressed as

$$\lambda \left| \frac{\partial \lambda}{\partial z} \right| \ll \lambda \quad \text{or} \quad \left| \frac{\partial m^{-1}}{\partial z} \right| \ll 1 \quad \text{or} \quad \left| \frac{\partial m}{\partial z} \right| \ll m^2. \tag{6.52a,b,c}$$

This condition will generally be satisfied if variations in the background state, or in the medium, occur on a scale much longer than the wavelength.

The Solution

Let us seek solutions of (6.50) in the form

$$\xi = A(z)\mathrm{e}^{\mathrm{i}\theta(z)}, \tag{6.53}$$

where $A(z)$ and $\theta(z)$ are both real. Using (6.53) in (6.50) yields

$$\mathrm{i}\left[2\frac{\mathrm{d}A}{\mathrm{d}z}\frac{\mathrm{d}\theta}{\mathrm{d}z} + A\frac{\mathrm{d}^2\theta}{\mathrm{d}z^2}\right] + \left[A\left(\frac{\mathrm{d}\theta}{\mathrm{d}z}\right)^2 - \frac{\mathrm{d}^2 A}{\mathrm{d}z^2} - m^2 A\right] = 0. \tag{6.54}$$

The terms in square brackets must each be zero. The WKB approximation is to assume that the amplitude varies sufficiently slowly that $|A^{-1}\mathrm{d}^2A/\mathrm{d}z^2| \ll m^2$, and hence that the term involving $\mathrm{d}^2A/\mathrm{d}z^2$ may be neglected. The real and imaginary parts of (6.54) become

$$\left(\frac{\mathrm{d}\theta}{\mathrm{d}z}\right)^2 = m^2, \qquad 2\frac{\mathrm{d}A}{\mathrm{d}z}\frac{\mathrm{d}\theta}{\mathrm{d}z} + A\frac{\mathrm{d}^2\theta}{\mathrm{d}z^2} = 0. \tag{6.55a,b}$$

The solution of the first equation above is

$$\theta = \pm \int m\, \mathrm{d}z, \tag{6.56}$$

and substituting this into (6.55b) gives

$$2\frac{\mathrm{d}A}{\mathrm{d}z}m + A\frac{\mathrm{d}m}{\mathrm{d}z} = 0, \quad \text{with solution} \quad A = A_0 m^{-1/2}, \tag{6.57a,b}$$

where A_0 is a constant. Using (6.56) and (6.57b) in (6.53) gives us

$$\xi(z) = A_0 m^{-1/2} \exp\left(\pm \mathrm{i} \int m\, \mathrm{d}z\right), \tag{6.58}$$

and this is the WKB solution to (6.50). In terms of real quantities the solution may be written

$$\xi(z) = B_0 m^{-1/2} \cos\left(\int m \,\mathrm{d}z\right) + C_0 m^{-1/2} \sin\left(\int m \,\mathrm{d}z\right), \tag{6.59}$$

where B_0 and C_0 are real constants.

Using (6.55a) and the real part of (6.54) we see that the condition for the validity of the approximation is that

$$\left| A^{-1} \frac{\mathrm{d}^2 A}{\mathrm{d}z^2} \right| \ll m^2, \tag{6.60a,b}$$

which using (6.57b) is

$$\left| \frac{1}{m^{-1/2}} \frac{\mathrm{d}^2 m^{-1/2}}{\mathrm{d}z^2} \right| \ll m^2. \tag{6.61}$$

Equation (6.52) expresses a similar condition to (6.61).

Notes and References

Many books discuss waves in fluids with the ones by LeBlond & Mysak (1980), Pedlosky (2003) and Gill (1982) having a geophysical emphasis. Accessible introductions to WKB theory can be found in Simmonds & Mann (1998) and Holmes (2013). An example of the application of WKB theory to the atmosphere is given in Hoskins & Karoly (1981).

Problems

6.1 *Rossby waves and Galilean invariance in shallow water.* Consider the flat-bottomed shallow-water quasi-geostrophic equations in standard notation,

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\zeta + \beta y - \frac{f_0 \eta}{H}\right) = 0. \tag{P6.1}$$

(a) How is ζ related to η? Express u, v, η and ζ in terms of a streamfunction.

(b) Linearize (P6.1) about a state of rest, and show that the resulting system supports two-dimensional Rossby waves. Discuss the limits in which the wavelength is much shorter or much longer than the deformation radius.

(c) Now linearize (P6.1) about a *geostrophically balanced state* that is translating uniformly eastwards. This means that: $u = U + u'$ and $\eta = \overline{\eta}(y) + \eta'$, where $\eta(y)$ is in geostrophic balance with U. Obtain an expression for the form of $\overline{\eta}(y)$. Obtain the dispersion relation for Rossby waves in this system. Show that their speed is different from that obtained by adding a constant U to the speed of Rossby waves in part *(b)*. The problem is therefore not *Galilean invariant* — why?

6.2 *Horizontal propagation of Rossby waves.* Consider barotropic Rossby waves obeying the dispersion relation $\omega = Uk - \beta k/((k^2 + l^2)$, where U and/or β vary slowly with latitude.

(i) By re-arranging this expression to obtain an expression for the meridional wavenumber, or otherwise, show that the Rossby waves can only propagate in the horizontal if

$$0 < U - c < \frac{\beta}{k^2}, \qquad \text{(P6.2)}$$

where $c = \omega/k$.

(ii) If the waves approach a latitude where $U = c$ (a 'critical latitude') show that the meridional wavenumber l becomes large but the group velocity in the y-direction becomes small. Show that Rossby waves generated in the atmosphere in midlatitudes are unlikely to propagate into the tropics (where the mean flow is westward).

(iii) Suppose that the Rossby waves approach a latitude where $U - c = \beta/k^2$. Calculate the x- and y-components of the group velocity in this limit and infer that a wave will turn away from such a latitude.

6.3 Consider barotropic Rossby waves with an infinite deformation radius. Obtain an expression for the group velocity in the presence of a uniform eastward mean flow. Show that for stationary waves the group velocity is eastward relative to Earth's surface, and hence deduce that energy propagation is downstream of topographic sources. Is this still true if the deformation radius is finite? Is it true if the mean flow is westward?

6.4 *Rossby waves in the two-level model.*

(a) Consider a two-layer (or two-level) quasi-geostrophic system on a β-plane, and you may consider the layers to be of equal depth. Linearize the equations about a state of rest, and show that they may be transformed into two uncoupled equations. Give a physical interpretation of what these equations represent.

(b) Obtain the dispersion relations for this system (there are two). Show that one of the dispersion relations corresponds to the synchronous, depth-independent motion in the two layers. What does the other one correspond to? Obtain the phases and group velocities of these waves and determine the conditions under which they are eastward or westward.

6.5 In realistic calculations of the vertical propagation of Rossby waves one must take into account the vertical variation of density. Carry through the calculation leading to the Charney–Drazin condition, either using pressure (or log-pressure) coordinates or using the anelastic version of quasi-geostrophy, using (5.66) with $\tilde{\rho} = \rho_0 \exp(-z/H)$. Show that the condition for propagation analogous to (6.41) becomes

$$0 < U < \frac{\beta}{k^2 + l^2 + \gamma^2}, \qquad \text{(P6.3)}$$

and obtain an expression for γ.