Chapter Thirteen

Instabilities, Fronts, and the General Circulation

13.1 Introduction

Halley's (1686) idea of the atmospheric circulation, with hot air rising in the tropics and cooler air descending at higher latitudes, was presumably based on experience with nonrotating fluids. In a nonrotating system, however, a zonally symmetric distribution of heating and cooling would give us no eastward or westward motion, and it was in this respect that Halley's scheme was deficient. The importance of rotation was later recognized by Hadley (1735), who showed that the tendency to conserve angular momentum can explain the eastward component of the trade winds (although he wrongly used conservation of angular velocity rather than angular momentum). Subsequent developments of models of the circulation are discussed by Lorenz (1967). In the nineteenth century, these were largely attempts to construct models that were qualitatively consistent with the observed surface distributions and with principles such as those outlined by Hadley. A rather different approach was the attempt by Vettin (1857) [see Fultz *et al.* (1959)] to model the circulation by using a rotating vessel that contained air as the working fluid and sources and sinks (such as ice) of heat to drive the motion. This approach was promising, but appears not to have been followed up until nearly a century later.

A major obstacle to progress was the lack of appreciation of the role of instabilities and of the transient and nonaxisymmetric motions that develop thereby. Helmholtz (1888) saw that instabilities could be important, but he placed most emphasis on the sort that gives rise to billow clouds and thereby aids vertical mixing. The instability idea was later utilized by V. Bjerknes (1937) [see Lorenz (1967)] to develop a picture

in close accord with present ideas. He concluded that the circulation would be rather different if it were forced to be zonally symmetric, and that such a zonally symmetric flow would be unstable to small longitude-dependent disturbances. Hence the observed circulation contains fully developed disturbances that take the form of cyclones and anticyclones.

Mathematical models of the instability that leads to cyclone development were developed by Charney (1947) and Eady (1949), and these are discussed in Sections 13.4 and 13.3, respectively. The process that they studied is called baroclinic instability, and the source of energy for the disturbances is the available potential energy (see Section 7.8) of the original zonally symmetric flow. The mere presence of available energy does *not*, however, imply instability, as the counterexample of Section 13.2 shows. In fact, certain conditions are necessary for instability to be possible, and these are considered in Section 13.5.

Another form of instability of geophysical interest is called barotropic instability. In this case, the source of energy is associated with horizontal variations in the velocity of the mean flow. The example that is chosen (in Section 13.6) to illustrate this process is based on Rayleigh's (1880) study of parallel-flow instability. As well as being directly applicable to the barotropic instability problem, the mathematics of this example is very similar to that of the Eady problem studied in Section 13.3.

The instability theories deal only with the initial development of small disturbances, whereas the role of eddies in the general circulation depends on their mean effect over a life cycle. The life cycle of a baroclinic disturbance is discussed in Section 13.9 for a model pertinent to the atmospheric circulation. The eddies (i.e., cyclones and anticyclones) transport heat poleward, as expected from the fact that they take available potential energy from the mean flow. However, they also transport zonal momentum poleward, apparently because planetary waves propagate upward and equatorward from the seat of the instability and tend to be absorbed in the equatorward side of the jet stream. The eddy momentum transports have direct consequences for the surface wind distribution because of the angular momentum balance requirement. This is discussed in Section 13.10 along with other aspects of the circulation problem. An important feature is that the zonal flow is close to being in hydrostatic and geostrophic balance, as realized by Ferrel (1859/1860) (see Section 7.6).

Baroclinic eddies are a prominent feature of the ocean as well as of the atmosphere, and these are discussed in Section 13.7. Although they are dynamically similar to their atmospheric counterparts, their horizontal scale is about a tenth of that for the atmosphere (100 km instead of 1000 km) and their time scale is much longer. Another phenomenon of great interest is that of fronts. In the atmosphere, they are usually associated with developing baroclinic disturbances, and an example of a front forming through nonlinear development of an Eady wave is considered in Section 13.8.

13.2 Free Waves in the Presence of a Horizontal Temperature Gradient

In Chapter 7 it was found that a rotating fluid adjusts to a geostrophic equilibrium rather than to a state of rest, and this equilibrium state is characterized by having

potential energy that is available for conversion into other forms. It is of great interest to examine the behavior of small disturbances to such equilibria to see whether the dynamic constraints allow the disturbances to draw on this supply of available potential energy. If they can, such disturbances will grow spontaneously and become an important feature of the flow. If not, the behavior of the disturbances is still of interest, and it is useful to consider why potential energy is *not* released.

To begin with, the problem will be studied in a uniformly rotating system (f plane) and complications due to the beta effect will be ignored. The fluid will be assumed to have reached an equilibrium state in which the temperature Θ has a uniform gradient in both the y (horizontal) and z_* (vertical) directions. It will be convenient to refer to the y direction as northward, although the significance of this direction is due to the temperature gradient rather than to the beta effect. Because of the horizontal temperature gradient, the system has available potential energy, as discussed in Section 7.8, and this energy could be released if the isotherms could be made horizontal.

By the thermal wind equation (7.7.10), the x component of velocity U has uniform shear in the vertical that is related to the horizontal temperature gradient by

$$f\,dU/dz_* = -\alpha_* g\,\partial\Theta/\partial y, \tag{13.2.1}$$

where f is the Coriolis parameter, g the acceleration due to gravity, and α_* the effective "expansion" coefficient defined by (6.17.21). The fluid is assumed to be incompressible (so the scale height H_s is infinite), log-pressure coordinates will be used, and only quasi-geostrophic processes will be considered.

The equations satisfied by small perturbations to this basic state are given in Section 12.9 and have an especially simple form in the present case in which $\beta = 0$, dU/dz_* is uniform, and the frequency N_* is constant. In fact (12.9.1) and (12.9.2) reduce to

$$\frac{\partial^2\Phi'}{\partial x^2} + \frac{\partial^2\Phi'}{\partial y^2} + \frac{f^2}{N_*^2}\frac{\partial^2\Phi'}{\partial z_*^2} = 0, \tag{13.2.2}$$

which becomes Laplace's equation if the stretched vertical coordinate $z_s = N_* z_*/f$ [see (8.8.25)] is used in place of z_*. Solutions exist that are wavelike in the horizontal and in time, and have a form such as

$$\Phi' = \Phi_0 \sin ly \sin(k(x - ct)) \exp(-z_*/H_R), \tag{13.2.3}$$

where (k, l) is the horizontal wavenumber and c is the phase speed of the disturbance in the x direction (i.e., in the direction of surface level isotherms). The solution decays with altitude on the scale [see (8.7.22)] of the Rossby height H_R, which is given by

$$H_R = f/N_*\kappa_H, \tag{13.2.4}$$

where $\kappa_H = (k^2 + l^2)^{1/2}$ is the horizontal wavenumber. The associated potential temperature perturbation θ is given by the hydrostatic equation (6.17.20), i.e., by

$$\alpha_* g\theta = \partial\Phi'/\partial z_* = -\Phi'/H_R. \tag{13.2.5}$$

Thus for perturbations that decay upward, *cold* is associated with *high* geopotential

(i.e., with high pressure on level surfaces) and *warm* is associated with *low* pressure. Lines of constant phase are vertical.

Laplace's equation does not allow solutions that are wavelike in all directions, so the only possible form of wave is a "surface" wave trapped against a boundary, such as a deep-water surface gravity wave. For such a wave to exist, the appropriate surface condition must be satisfied. In the present case, the condition to be applied at the horizontal boundary $z_* = 0$ is $w_* = 0$, i.e., by the temperature equation [or (12.9.6)]

$$(\partial/\partial t + U\,\partial/\partial x)\theta + v\,\partial\Theta/\partial y = 0 \qquad \text{at} \quad z_* = 0, \tag{13.3.8}$$

where v is the perturbation velocity component in the y direction. Using (13.2.3) for v and (13.2.5) for Φ' in terms of θ, this becomes

$$\partial\theta/\partial t + (U(0) + H_R\,dU/dz_*)\,\partial\theta/\partial x = 0 \qquad \text{at} \quad z_* = 0, \tag{13.3.9}$$

showing that the wave translates at a speed c given by

$$c = U(0) + H_R\,dU/dz_*, \tag{13.3.10}$$

i.e., at the wind speed one Rossby height (or one e-folding scale) above the surface. This level, where phase speed equals wind speed, is called the "*steering level*."

The structure of the boundary wave is shown in Fig. 13.1 in a frame of reference that is stationary relative to the wave. At ground level, the streamlines are sinusoidal and air is warm for its latitude when displaced furthest poleward. Consequently, by (13.2.5) the surface pressure is low, and this is consistent with the streamline pattern only if the flow is easterly as shown in Fig. 13.1d. At the steering level there is no mean flow relative to the wave, so the motion is purely "coldward" east of the low and purely "warmward" (i.e., in the direction of the mean horizontal temperature gradient) west of the low. This is just the situation found in model A of Section 12.10, and consequently rising motion is expected in the "coldward" jet and descent in the "warmward" jet. Such is indeed the case. The value of w_* is given by (12.9.6), using (13.2.3) and (13.2.8), with the result (for $c = 0$)

$$w_* = \Phi_0 N_*^{-2} k\,(dU/dz_*) \sin ly \cos kx\,(z_*/H_R) \exp(-z_*/H_R). \tag{13.3.11}$$

Streamlines of the ageostrophic motion in the east–west vertical plane are shown in Fig. 13.1b and are obtained by integrating (13.2.9) with respect to x. Figure 13.1c shows isotherms (dashed lines) and contours of meridional velocity in the east–west vertical plane. Where isotherms are depressed the most, the air at that level is warmest, but the meridional velocity is zero. Figure 13.1e shows particle trajectories in the meridional plane. These can be calculated from (13.2.9), (12.9.3), and (13.2.3), which give

$$\frac{w_*}{v} = \frac{fz_*}{N_*^2 H_R}\frac{dU}{dz_*} = -\frac{z_*}{H_R}\frac{\partial\Theta/\partial y}{\partial\Theta/\partial z_*}. \tag{13.3.12}$$

The last equality makes use of the definition (6.17.24) of N_*^2. The formula shows that trajectories are less steep than are isotherms below the steering level and steeper

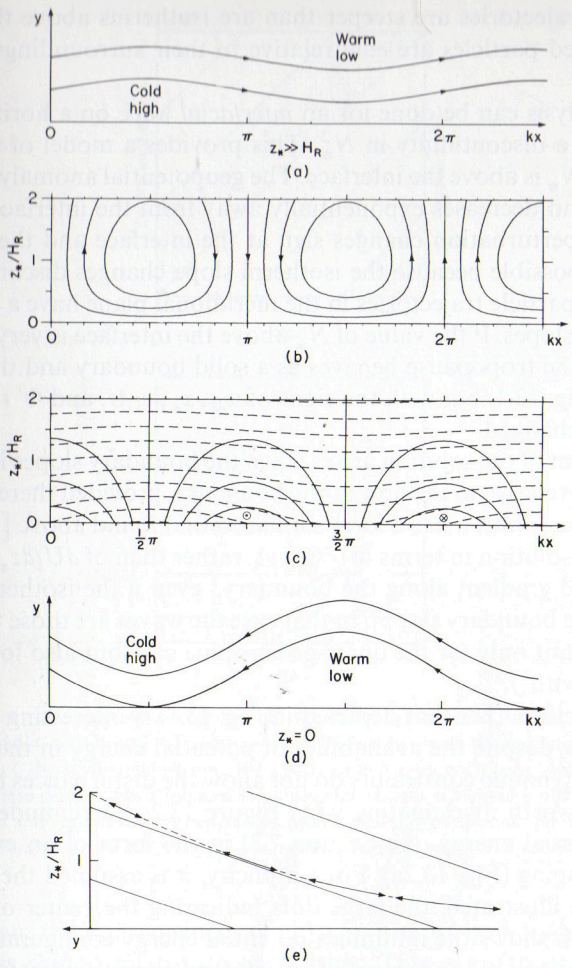


Fig. 13.1. Properties of a wave trapped against a horizontal boundary in a uniform shear flow in a uniformly rotating environment. Isotherms are uniformly sloping in the y–z plane as shown by the solid lines in (e). (a) Streamlines (which coincide with isobars and isotherms) in the horizontal for the flow relative to the wave at a high level where the disturbance is weak. As at all levels high pressure (or high geopotential) is associated with cold air. The coldness is due to the air being displaced upward. (b) Streamlines of the ageostrophic flow (i.e., of the disturbance) in the y–z_* plane. Ascent is associated with a "coldward" flow, coldward meaning that there is a horizontal component in the y-direction, i.e., toward the direction where the mean temperature at a given level is colder (see Section 12.10). (c) Contours of v, the y component of velocity (solid) and of potential temperature in the x–z_* plane. Where the air is warmest (isentropes most depressed), there is no poleward flow, and where the poleward flow is strongest, the temperature perturbation is zero. Thus there is no poleward heat transfer by the wave. (d) Surface streamlines relative to the wave. The flow is easterly, and high pressure (where streamlines are displaced furthest equatorward) is associated with cold air, the coldness being due to the equatorward displacement. (e) Particle trajectories (arrows) in the y–z_* plane relative to the isentropes (solid sloping lines). Near the ground, where the amplitude is large, the slope of these trajectories is more nearly horizontal than are those of the isentropes, so equatorward-displaced air is cold. At high levels where the amplitude is small, the slope of the trajectories is greater than that of the isentropes, so equatorward-displaced air is warm because of its relatively large downward displacement.

above. Because trajectories are steeper than are isotherms above the steering level, poleward-displaced particles are *cold* relative to their surroundings, as seen in Fig. 13.1a and c.

A similar analysis can be done for an *interfacial wave* on a horizontal boundary at which there is a discontinuity in N_*^2. This provides a model of the *tropopause*, the high value of N_* is above the interface. The geopotential anomaly Φ' is continuous at the interface and decreases exponentially away from the interface. It follows that the temperature perturbation changes sign at the interface and therefore is discontinuous! This is possible because the isotherm slope changes discontinuously at the interface and the particle trajectories in the meridional plane have a slope in between the two isotherm slopes. If the value of N_* above the interface is very large compared with that below, the tropopause behaves as a solid boundary and the solution is the same as that in Fig. 13.1, but with the signs of w_*, z_*, v, U, and Φ' reversed, those of θ and x being unchanged.

Another variant of the solution arises when the boundary slopes in the y direction. Waves propagate relative to the flow at the boundary, provided there is a temperature gradient there. The solution has a form similar to that found above [it is more appropriate to give the solution in terms of $(\partial\Theta/\partial y)_b$ rather than of dU/dz_*, where $(\partial\Theta/\partial y)_b$ is the temperature gradient along the boundary] even if the isothermal surfaces are horizontal and the boundary slopes. In that case the waves are those that were studied by Rhines (1970) not only for the quasi-geostrophic case but also for frequencies not small compared with f.

The above solution (i.e., that depicted in Fig. 13.1) is interesting because disturbances do *not* grow despite the availability of potential energy in the mean flow. For some reason, the dynamic constraints do not allow the disturbances to tap this energy source, and it is worth investigating why. Figure 13.2 is a reminder of the concept of available potential energy (see Section 7.8) in the form of an example in which isopycnals are sloping (Fig. 13.2a). For simplicity, it is assumed there are six homogeneous layers as illustrated, the large dots indicating the center of gravity of each layer. Figure 13.2b shows the minimum potential energy configuration of the layers with the new positions of the centers of gravity and arrows indicating the change in position of these centers. This illustrates the principle that release of available potential energy is associated with heavy (cold) fluid moving equatorward and light (warm) fluid moving poleward.

In the wave solution above, the dynamic constraints do not allow this to happen. The mean poleward heat flux over a wave is proportional to

$$\overline{v\theta} = \frac{1}{f}\overline{\frac{\partial\Phi'}{\partial x}\theta} = \frac{1}{\alpha_* g f}\overline{\frac{\partial\Phi'}{\partial x}\frac{\partial\Phi'}{\partial z_*}}, \tag{13.2.11}$$

where the overbar denotes an average over a wavelength in the x direction. The relation (13.2.5), however, requires this to be identically zero at all levels, because θ is proportional to Φ', and so

$$\overline{v\theta} \propto \overline{\Phi'\,\partial\Phi'/\partial x} = \overline{\partial(\tfrac{1}{2}\Phi'^2)/\partial x} = 0. \tag{13.2.12}$$

Thus no heat is carried poleward and no energy is released.

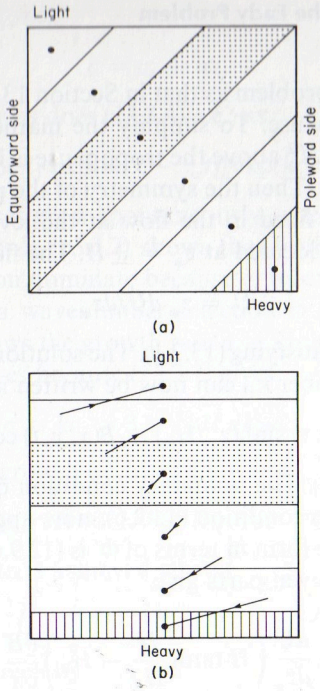


Fig. 13.2. Illustration of displacements associated with rearranging a fluid with sloping isentropes (a) to a minimum energy condition with horizontal isentropes (b). The fluid has been divided into six layers, each treated as being homogeneous. The solid circles denote the center of gravity of each layer and the arrows in (b) show the displacement of these centers that is required to achieve the state of minimum potential energy. Heavy (cold) fluid moves downward and equatorward, whereas light (warm) fluid moves upward and poleward. Consequently, there is a net poleward transfer of heat. One layer is shown hatched and another stippled for ease of identification.

To escape from the consequences of (13.2.5), it must be possible for the v and θ fields to be phase-shifted relative to each other. This can happen in various ways. One is confining the disturbance between *two* horizontal boundaries or between the surface and tropopause as shown in Eady's (1949) celebrated paper, for then there are two solutions in the vertical, one decaying away from the upper surface and one from the lower surface. If the two surfaces are many Rossby heights apart, the waves on the two boundaries affect each other only slightly, so no new effects are found. For the waves to have a large effect on each other, the Rossby height must be comparable with the height of the model tropopause, i.e., the inverse wave number must be comparable with the Rossby radius based on the tropopause height (i.e., N/f times the tropopause height, or about 1000 km). The phase shift can have different signs, depending on the relative positioning of the waves on the upper and lower boundaries. The sign that gives release of potential energy can be seen from the last expression in (13.2.11) because if *phase lines tilt westward with height* (as in Fig. 12.10, for instance), Φ' increases with x when it increases with z_*, so the poleward heat flux is positive. When the two waves are able to coexist with this phase shift, it will be shown in Section 13.3 that they grow spontaneously. Conversely, if the phase shift results in eastward tilt with height, the disturbances decay.

13.3 Baroclinic Instability: The Eady Problem

Now consider the same problem as that in Section 13.2, but with an upper boundary provided by the tropopause. To simplify the mathematics, the limiting case is taken for which the value of N_*^2 above the tropopause is large, so the upper boundary behaves as a solid boundary. Then the symmetry of the problem can be exploited by taking a frame of reference fixed in the flow at the level $z_* = 0$, midway between the two boundaries that are located at $z_* = \pm H$. The flow is thus given by

$$U = z_* \, dU/dz_*, \tag{13.3.1}$$

where dU/dz_* is a constant satisfying (13.2.1). The solution of (13.2.2) for a disturbance of fixed horizontal wavenumber κ_H can now be written in the form

$$\Phi' = A(x, y, t)\sinh(z_*/H_R) + B(x, y, t)\cosh(z_*/H_R), \tag{13.3.2}$$

where H_R is given by (13.2.4). The boundary condition of no vertical motion, which gives rise to the temperature condition (13.2.6), now applies on the two boundaries at $z_* = \pm H$. The alternative form in terms of Φ' is (12.9.6) with $w_* = 0$. Substituting (13.3.2) and taking odd and even parts give

$$\frac{\partial A}{\partial t} + \frac{dU}{dz_*}\left(H\tanh\frac{H}{H_R} - H_R\right)\frac{\partial B}{\partial x} = 0,$$
$$\frac{\partial B}{\partial t} + \frac{dU}{dz_*}\left(H\coth\frac{H}{H_R} - H_R\right)\frac{\partial A}{\partial x} = 0. \tag{13.3.3}$$

Solutions exist in which A and B are proportional to

$$\cos ly \exp(ik(x - ct)), \tag{13.3.4}$$

where the wave speed c is given by

$$c^2 = (dU/dz_*)^2(H\tanh(H/H_R) - H_R)(H\coth(H/H_R) - H_R). \tag{13.3.5}$$

This is one of the results obtained by Eady (1949).

In the limit, in which the two boundaries are many Rossby heights apart, i.e., $H \gg H_R$ [by (13.2.4), this corresponds to the *shortwave limit*], the tanh and coth functions both tend to unity and (13.3.5) gives

$$c \approx \pm(dU/dz_*)(H - H_R). \tag{13.3.6}$$

In other words, c is equal to the wind speed one Rossby height from the boundary, i.e., these are the boundary waves found in Section 13.2.

In the *long-wave limit*, where the boundaries are a small fraction of a Rossby height apart, i.e., $H \ll H_R$, (13.3.5) gives

$$c^2 \approx -\tfrac{1}{3}H^2(dU/dz_*)^2, \tag{13.3.7}$$

i.e., c has become purely imaginary. In general, if c is expressed in terms of its real part c_r and imaginary part c_i, i.e.,

$$c = c_r + ic_i, \tag{13.3.8}$$

the amplitudes A and B vary as does (13.3.4), i.e., as

$$\cos ly \exp(ik(x - c_r t))\exp(kc_i t). \tag{13.3.9}$$

If c_i is positive, disturbances grow spontaneously, whereas $c_i < 0$ corresponds to a decaying disturbance. Formula (13.3.7) shows that both types of disturbance exist, but the growing one will soon dominate because of its exponentially growing amplitude. In fact, there will be a wavenumber selection in favor of the fastest-growing disturbance. Figure 13.3 shows the growth rate $\sigma \equiv kc_i$ as a function of wavenumber (k, l). It is small for long waves because (13.3.7) gives

$$\sigma \equiv kc_i \approx 3^{-1/2}kH\, dU/dz_*, \tag{13.3.10}$$

whereas (13.3.5) shows that growth occurs only when

$$H\tanh(H/H_R) < H_R, \quad \text{i.e.,} \quad H < 1.1997H_R \quad \text{or} \quad N_*\kappa_H H < 1.1997f \tag{13.3.11}$$

by (13.2.4). Maximum growth is achieved when

$$l = 0 \quad \text{and} \quad H = 0.8031H_R, \quad \text{i.e.,} \quad N_*kH = 0.8031f, \tag{13.3.12}$$

the maximum value being given by

$$\sigma_{\max} \equiv 0.3098(f/N_*)\, dU/dz_*. \tag{13.3.13}$$

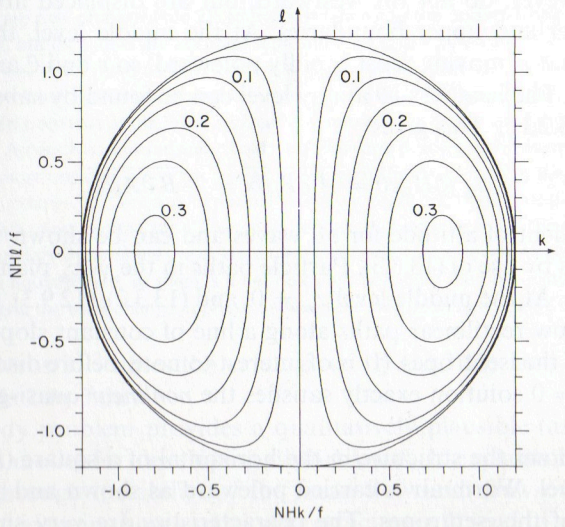


Fig. 13.3. Growth rate σ of an Eady wave as a function of wavenumber (k, l). Contours are shown in units of $(f/N_*)\,dU/dz_*$. Values are zero on $k = 0$ and when the magnitude κ_H of the wavenumber equals 1.1997. The maximum value $0.3098(f/N_*)\,dU/dz_*$ is achieved when $l = 0$ and $N_*H\kappa_H/f = 0.8031$. For fixed ratio k/l, the maximum is at the same value of κ_H. The maximum for a fixed l (corresponding to a baroclinic zone of fixed width) is at a value of k that decreases as l increases (longer unstable waves for narrower baroclinic zones).

For this value, (13.3.3) shows that

$$\frac{A}{B} = \left(\frac{H_R - H\tanh(H/H_R)}{H\coth(H/H_R) - H_R}\right)^{1/2} = 1.502. \tag{13.3.14}$$

This gives the general formula for A/B. For the fastest-growing mode it is also equal to $\coth(H/H_R)$, and thus the solution (13.3.2), taking the real part in (13.3.9), has the form

$$\Phi' = \left[\cos kx\,\frac{\sinh(z_*/H_R)}{\sinh(H/H_R)} + \sin kx\,\frac{\cosh(z_*/H_R)}{\cosh(H/H_R)}\right]\exp(\sigma t),$$

$$\alpha_* g H_R \theta = \left[\cos kx\,\frac{\cosh(z_*/H_R)}{\sinh(H/H_R)} + \sin kx\,\frac{\sinh(z_*/H_R)}{\cosh(H/H_R)}\right]\exp(\sigma t). \tag{13.3.15}$$

Figure 13.4 shows the structure of this fastest-growing mode solution in a format similar to that in Fig. 13.1. Near the lower boundary, the Eady wave is very much like the boundary wave except for the very important phase shift of 21° between the positions of isotherms and isobars, i.e., the warmest air is just ahead (eastward) of the surface trough. Conversely, the Eady wave is similar to a trapped wave on an upper boundary at that boundary. For the growing wave, the highs and lows are *displaced* 90° *to the west* on the upper boundary relative to the lower boundary. Thus the phase lines for the Φ' field tilt westward with height and the same applies to v, which is proportional to the x derivative of Φ'. The same is also true for the ageostrophic field because u_a is proportional to $U\,\partial v/\partial x$ by (12.2.24). The phase lines for temperature, however, do *not* tilt westward, but are displaced about 48° eastward between the lower and upper boundaries. At the middle level, the warmest air is actually the air that is moving most rapidly poleward, so v and θ are perfectly correlated at this level. The heat flux $v\theta$ at any level can be found by substituting the solution (13.3.2) in (13.2.11). This gives

$$\alpha_* f g H_R \overline{v\theta} = \overline{A\,\partial B/\partial x} = -\overline{B\,\partial A/\partial x}, \tag{13.3.16}$$

which is independent of altitude for all waves and can be shown to be positive for all growing waves by use of (13.3.3). Particle paths in the y–z_* plane are ellipses that expand with time. At the middle level $z_* = 0$, and (13.3.2), (12.9.3), and (12.9.6) show that particles follow rectilinear paths along a line of constant slope equal to 0.534 times the slope of the isentropes. (It is of interest to note, before discussing cases with $l = 0$, that the $l = 0$ solution exactly satisfies the *nonlinear* quasi-geostrophic equations.)

Figure 13.5 shows the structure in the horizontal of a square ($k = l$) Eady wave at the steering level. Warm air is carried poleward as shown and ascends at 0.534 times the slope of the isentropes. The characteristics are very similar to those of observed cyclonic and anticyclonic disturbances at the 700- or 600-mb level (see, e.g., Fig. 12.17). The zonal wavelength of this wave is $2^{1/2}$ times the value given by (13.3.12), namely, $11.1NH/f \approx 4000$ km or about zonal wavenumber 6 at the latitudes where disturbances are generally found. The shortest e-folding time for growth of a

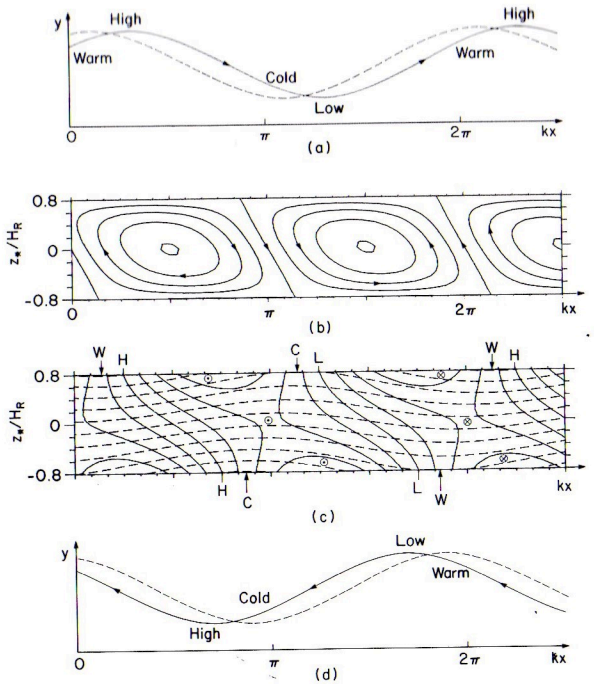


Fig. 13.4. Properties of the most unstable Eady wave, i.e., the most unstable wave in a uniform shear flow between two horizontal boundaries in a uniformly rotating environment. The solution is independent of y. (a) The pattern on the upper surface and (d) the pattern on the lower surface, the solid line being an isobar and the dashed line an isotherm. The pattern on the lower surface is very similar to that of the boundary wave shown in Fig. 13.1d, except that isotherms are now phase-shifted 21° to the east relative to the isobars. (The phase shift is exaggerated in (a) and (d) for clarity, but (b) and (c) are accurate representations.) The poleward flow is now on the warm side and the equatorward flow is on the cold side so there is net poleward heat flux. At the furthest poleward point of the isotherm, the flow is still poleward because displacements are increasing with time. The pattern at the upper surface can be obtained from that at the lower surface by symmetry. (b) The stream function for the ageostrophic flow in the x–z_* plane. Ascent is associated with "coldward" flow and descent with "warmward" flow as was found in Section 12.10. The most unstable wave has wavenumber such that there are 1.6 Rossby heights H_R between the two horizontal boundaries. (c) Contours of normal velocity v (solid) and isentropes (dashed) in the x–z_* plane. The lines marked H (high geopotential) and L (low geopotential) are zero lines for v. The points marked W (warm) and C (cold) on the boundary show where the air is warmest and coldest. At all levels, air going poleward is generally warmer than air going equatorward, so there is a net poleward heat flux. The phase lines of the v field (as for Φ') tilt westward with height, the total change in phase between the two boundaries being 90°.

square Eady wave is $2^{1/2}$ times the value given by (13.3.13), and a typical value for the atmosphere is 2 or 3 days.

Thus the Eady problem provides a qualitatively plausible (and mathematically straightforward) model for developing disturbances. It explains how they can form spontaneously through instability of the mean flow and draw on the energy available there. It also explains how a particular structure will tend to emerge by selection of the disturbance that grows most quickly. The properties of this disturbance also agree well with observation, i.e., (a) there are growing waves of wavenumber 6, (b) there is usually a warm tongue a little ahead of the surface trough, (c) at the steering level (located roughly midway between the surface and the tropopause) warm air

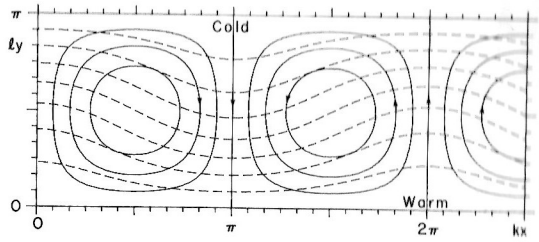


Fig. 13.5. Geopotential anomaly contours (solid) and temperature contours (dashed) for a growing ($k = l$) Eady wave at the steering level. The warm poleward jet is descending at about half the slope of the isentropes and the cold equatorward jet ascends at the same angle, so there is a net poleward heat transfer and release of potential energy. The relationship between the two fields is similar to that in the synoptic situation shown in Fig. 12.17 although the fields are considerably distorted in the real situation.

flows poleward and descends at about half the slope of the isentropes, etc. (*Note:* The observed steering level tends to be a bit lower than the Eady model predicts. Models that include beta effects can explain this lowering—see Section 13.4.)

Laboratory models have been found to be very useful for studying the properties of baroclinic systems and, in particular, the variety of behavior of disturbances that can be found in such systems. A review of this work has been given by Hide and Mason (1975). Numerical models [e.g., Williams (1971)] show that the initial structure of freshly growing disturbances is very similar to that predicted by Eady's analysis, especially if this is generalized to take account of nonuniform gradients (but still with uniformly sloping isentropes (G. P. Williams, 1974)), and of Ekman friction at the horizontal boundaries (Williams and Robinson, 1974). Some very interesting phenomena occur when the disturbances reach finite amplitude, and these are discussed in Hide and Mason's review.

The Eady problem also points to the possibility of eddylike disturbances that grow spontaneously in the ocean, although the model of the mean flow is not very realistic. However, since the preferred scale of the disturbances is of the order of the Rossby radius, it might be expected that the preferred wavelength in the ocean would be of the order of $2\pi \times 30 \approx 200$ km and this is in agreement with observed scales. Also, since the shear is so much smaller in the ocean, the growth rate given by (13.3.13) is small and therefore e-folding times of order 100 days are to be expected (Gill *et al.*, 1974).

13.4 Baroclinic Instability: The Charney Problem

The previous two sections ignored the beta effect, and it is important to find out how this affects the stability problem. Consider first the situation of Section 13.2, i.e., that with uniformly sloping isentropes over a single horizontal boundary, but with the beta effect included. This is the problem (with compressibility effects also included) considered in Charney's (1947) pioneering paper on baroclinic instability.

The equation for a small wavelike disturbance on any zonal mean flow $U(y, z_*)$ with the y variations sufficiently slow is (12.9.8) with the potential vorticity gradient

given by (12.9.4). The boundary condition $w_* = 0$ on a horizontal boundary gives, after substitution of the wavelike form (12.9.7) in (12.9.6),

$$(U - c)(\partial\psi/\partial z_* + \tfrac{1}{2}\psi/H_s) = (\partial U/\partial z_*)\psi. \tag{13.4.1}$$

For "purely baroclinic" instability problems, U is a function only of z_*, so the partial derivatives in (12.9.8) and (13.4.1) become ordinary derivatives. In the Eady problem, $\partial\bar{q}/\partial y$ was zero, so m^2 given by (12.9.9) was constant, and analytic solutions of (12.9.8) were thus easy to obtain. If $\partial\bar{q}/\partial y$ is nonzero, m^2 is not constant, but solutions are easily obtained numerically when the imaginary part c_i of c is nonzero. For "neutral" disturbances, i.e., freely propagating waves with $c_i = 0$, the calculation is complicated by the singularity that exists at the steering level, but this does not occur for growing disturbances.

In the Charney problem

$$d\bar{q}/dy = \beta \tag{13.4.2}$$

is a constant and there is no upper boundary. Solutions are discussed, e.g., by Kuo (1952, 1973), Charney (1973), and Pedlosky (1979). The solution for the fastest-growing mode (which has $l = 0$) is shown in Fig. 13.6 for the incompressible limit $H_s \to \infty$, so comparison can be made with the f-plane solution of Fig. 13.1 and the Eady solution of Fig. 13.4. The maximum growth rate is achieved for $l = 0$ and is independent of β, being given by

$$\sigma_{\max} = 0.286(f_0/N_*)\, dU/dz_*. \tag{13.4.3}$$

Surprisingly, this is almost the same result as that (13.3.13) found for the Eady

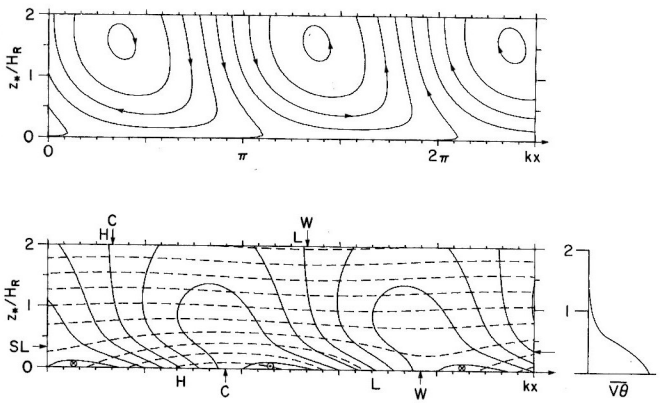


Fig. 13.6. Properties of the most unstable Charney wave, i.e., of the most unstable perturbation to a uniform shear flow on a beta plane, in the incompressible limit. The solution is independent of y. (a) The stream function for the ageostrophic flow in the x–z_* plane. Near the ground, ascent is associated with "coldward" flow and descent with "warmward" flow as found in Section 12.10. (b) Contours of normal velocity v (solid) and isentropes (dashed) in the x–z_* plane. The lines marked H (high geopotential) and L (low geopotential) are zero lines for v. The points marked W (warm) and C (cold) show where the air is warmest and coldest on the lower boundary and also at $z_* = 2H_R$, which is the edge of the picture but not of the flow. The small panel at the side shows how the poleward heat flux varies with height. The steering level (SL) is also marked. The Rossby height H_R for this wave is given by (13.4.4). (Solution courtesy of P. D. Killworth.)

problem. The functional dependence is exactly the same, with no dependence on β, and the value is only about 8% less than that in the Eady case. Thus an e-folding time of about 2 days for the mid-latitude atmosphere is also predicted by this model.

Qualitatively, the solution looks like the Eady solution (Fig. 13.4) near the ground. Perturbation geopotential contours tilt westward with height, warm air is ahead of the trough line, poleward-moving air is warmer than average and descends at an angle less than that of the isentropes. Quantitatively, of course, there are differences. For example, the surface warm tongue is now 41° ahead of the surface trough. The phase difference between temperature and pressure increases to a maximum of 65° at $z_* = 0.5H_R$ and then falls off toward zero. The steering level is at $z_* = 0.33H_R$ (this is lower than for the Eady problem, in which it was $0.40H_R$ above the surface). The poleward heat flux falls off with height and is less than 7% of its surface value at $z_* = H_R$. At higher levels, the solution approaches asymptotically the f-plane solution shown in Fig. 13.1 and the visual similarity can be seen in the figures. In particular, the phase differences between Φ' and θ become small. This is also a feature of observed transient disturbances in the atmosphere (Lau and Wallace, 1979).

The important feature of the incompressible Charney solution that is not in the Eady solution is that the horizontal and vertical scales are set by the value of β and not by the tropopause height. Scales for the fastest-growing wave (which has $l = 0$) are given by

$$k^{-1} = 1.26(f_0/\beta N_*)\, dU/dz_*, \tag{13.4.4}$$

$$H_R \equiv (f_0/kN_*) = 1.26(f_0^2/\beta N_*^2)\, dU/dz_*$$

$$= 1.26 f_0/\beta \times \text{isentrope slope}.$$

The last expression comes from using the thermal wind equation (13.2.1) and the definition (6.17.24) of N_*^2. It so happens that the value of H_R, at midlatitudes, that is given by this formula is of order 10 km, so these scales are very similar to those found by Eady.

Green (1960) studies the problem when both the beta effect *and* a lid are present, thereby combining the Eady and Charney problems. The parameter that determines the solutions is the ratio of the height scale given in (13.4.4) to the height of the lid. The maximum growth rate varies little, as expected from the fact that it is very similar in the two limiting cases. However, the beta effect does strongly influence the properties of the very short and very long waves. There are now short wave modes that are rendered unstable by the beta effect, whereas the longest waves are unstable through a new mode with a more complicated vertical structure and a weaker growth rate.

The height scale given by (13.4.4) also is comparable with the scale height H_*, so compressibility effects should also be included. These produce differences in detail but not in the basic structure of the fastest-growing mode [see Lindzen *et al.* (1980) for the compressible case].

The Charney and Eady problems provide good illustrations of the baroclinic instability process in a continuously stratified fluid. The instability can also be obtained in the single situation of two superposed homogeneous layers of different densities and differing mean velocities. This was studied by Phillips (1951) for the f-plane case [see also Pedlosky (1979)], with layers of equal depth, and was extended, e.g., by Gill *et al.* (1974), to include effects of beta, bottom slope, and unequal layer depth. The model has been applied by P. C. Smith (1976) to explain the energetic fluctuations observed in the very strong (0.6 m s^{-1}) bottom current (Fig. 10.8) that flows through Denmark Strait between Iceland and Greenland. The fluctuations have a period of 1.8 days and the e-folding time given by the model is about 4 days.

13.5 Necessary Conditions for Instability

In Sections 13.3 and 13.4 it was shown by studies of particular cases how disturbances can grow spontaneously by drawing on the supply of potential energy available in the mean flow. This is a very important demonstration because it shows a means by which atmospheric depressions and oceanic eddies can be formed. However, the example of Section 13.2 shows that the mere presence of available energy is *not* sufficient to ensure instability since dynamical constraints may not allow the energy to be released. This raises the question about what conditions are required for instability to occur. It turns out that it is possible to find conditions that are *necessary* in order for instability to occur. These are very useful because if they are *not* satisfied, it can be concluded that the dynamical constraints will *not* allow the energy to be released. If they *are* satisfied, the possibility of instability is indicated, but this cannot be verified without making detailed calculations.

The necessary conditions for instability are most easily derived for perturbations with no y dependence and for flow between two horizontal boundaries. Suppose that the flow is unstable, so that the imaginary part c_i of c is nonzero. Then multiplying (12.9.8) by the complex conjugate of ψ, integrating by parts, and using the boundary condition (13.4.1), an equation results whose imaginary (divided by c_i) and real parts yield, respectively,

$$\int \frac{N_*^2|\psi|^2}{f_0^2|U-c|^2}\frac{\partial \bar{q}}{\partial y}\,dz_* - \left[\frac{\alpha_* g|\psi|^2}{f_0|U-c|^2}\frac{\partial \Theta}{\partial y}\right] = 0, \tag{13.5.1a}$$

$$\int (U-c_r)\frac{N_*^2|\psi|^2}{f_0^2|U-c|^2}\frac{\partial \bar{q}}{\partial y}\,dz_* - \left[(U-c_r)\frac{\alpha_* g|\psi|^2}{f_0|U-c|^2}\frac{\partial \Theta}{\partial y}\right] = P, \tag{13.5.1b}$$

where P is a positive definite expression.

In each equation, the first term is an integral from the lower to the upper boundary and the square brackets enclosing the second term denote the value at the upper boundary minus the value at the lower boundary. The thermal-wind equation (13.2.1) has been used to express the boundary contributions in terms of the temperature gradient on the boundary. The result is easily generalized to the case in which y dependence is allowed, in which case (13.5.1) is integrated over the y domain, at the boundaries of which either periodicity or a condition of no normal flow is assumed. Also, the upper and lower boundaries can have a small slope, provided that $\partial\Theta/\partial y$ in (13.5.1) is interpreted as the temperature gradient along the boundary rather than

that at a fixed level. It is also possible to remove one or both of the boundaries to infinity by taking the appropriate limit.

The application of (13.5.1a) to the problems already studied can now be made by considering the *signs* of the various terms on the left-hand side. In the example of Section 13.2, the only contribution is a negative one from the lower boundary, which cannot equal zero, and therefore the flow must be stable, as is indeed the case. In the Eady problem, the negative contribution from the lower boundary is balanced by a positive contribution from the upper boundary, thereby allowing the possibility of instability. In the Charney problem, there is no upper boundary contribution, but there is instead a positive contribution from the interior, so instability is again possible. In general, the *necessary condition for instability* that is required for (13.5.1) to hold is (Green, 1960; Charney and Stern, 1962) that the set of functions

$$(\partial\bar{q}/\partial y)_{\text{interior}}, \qquad (\partial\Theta/\partial y)_{\text{lower}}, \qquad -(\partial\Theta/\partial y)_{\text{upper}} \tag{13.5.2}$$

must *not* have the same sign throughout, but must include both positive and negative values. Conversely, a sufficient condition for stability is that the set of functions (13.5.2) have the same sign everywhere. An alternative derivation and interpretation of this result is given by Bretherton (1966).

The condition can be strengthened by taking into account (13.5.1b), in which c can be replaced by an arbitrary constant U_r because an arbitrary multiple of (13.5.1a) can be added to (13.5.1b). It follows that the *flow is stable if* (Pedlosky, 1964) a number U_r can be found such that the functions

$$((U - U_r)\,\partial\bar{q}/\partial y)_{\text{interior}}, \qquad ((U - U_r)\,\partial\Theta/\partial y)_{\text{lower}}, \qquad -((U - U_r)\,\partial\Theta/\partial y)_{\text{upper}} \tag{13.5.3}$$

are nowhere positive. The result can also be obtained [see Pedlosky (1979)] without assuming a wavelike dependence on x and t by balancing the rate of change of disturbance energy with rates of change of other integrals that are negative definite when the quantities (13.5.3) are everywhere negative or zero. The technique can be generalized to nonparallel, quasi-geostrophic flows. Then the geopotential perturbation Φ'' acts as a stream function for the steady flow whose stability is being investigated, since the horizontal flow is geostrophic and thus horizontally nondivergent at the leading order of approximation. Consequently (12.8.13) shows that the potential vorticity $\bar{q}$ is constant on streamlines $\Phi'' = \text{const.}$, i.e.,

$$\bar{q} = \bar{q}(\Phi''). \tag{13.5.4}$$

Similarly, the temperature Θ of a fluid particle on the boundary is conserved, and so

$$\Theta = \Theta(\Phi'') \tag{13.5.5}$$

on the upper and lower boundaries. A sufficient condition for stability that was found by Blumen (1968), using a method due to Arnold (1965), is that

$$(\partial\bar{q}/\partial\Phi'')_{\text{interior}}, \qquad (\partial\Theta/\partial\Phi'')_{\text{lower}}, \qquad -(\partial\Theta/\partial\Phi'')_{\text{upper}} \tag{13.5.6}$$

be positive everywhere. The relationship with (13.5.3) can be seen by noting that for

parallel flow

$$\frac{\partial\bar{q}}{\partial\Phi''} = \frac{\partial\bar{q}/\partial y}{\partial\Phi''/\partial y} = -\frac{\partial\bar{q}/\partial y}{f_0 U} \tag{13.5.7}$$

The above conditions are useful, for example, in discussing baroclinic instability in the ocean. Gradients are generally weak below the surface layers, so $\partial\bar{q}/\partial y$ is close to beta except near the surface. If $\partial\bar{q}/\partial y$ does not change sign and the bottom temperature gradient is negligible, (13.5.2) implies that instability is possible only if the surface temperature increases toward the poles! This rarely happens, but in regions of westward surface flow, the temperature increases poleward at the thermocline level, as can be seen in meridional temperature sections (see also Fig. 12.5). This can produce a change in sign of $\partial\bar{q}/\partial y$ with quite modest currents. For instance, Gill *et al.* (1974) found instabilities with e-folding times of order 100 days for westward surface currents of 0.05 m s^{-1}. A quite different situation exists in the cold waters of Drake Passage, where the observed eddies appear to be due to an instability associated with the equatorward increase of temperature on the bottom (Wright, 1981).

13.6 Barotropic Instability

The stability problems examined so far have been for the case in which U is a function only of z_*, i.e., the "pure baroclinic" case. In general U is a function of both y and z_*, and the potential vorticity gradient (12.9.4), which can be responsible for instability, as it is in the Charney problem, involves y derivatives as well as z_* derivatives. The condition for the terms involving y derivatives to be small relative to those involving z_* derivatives is that the y scale L should satisfy

$$L \gg N_* H/f_0, \tag{13.6.1}$$

where H is the z_* scale of the U profile, assumed not to be larger than the scale height H_*. In other words, the pure baroclinic problem is generally applicable only when the undisturbed flow has horizontal scales large compared with the Rossby radius. The condition (see Section 12.3) is also the one for the energy of the undisturbed flow to be principally available potential energy rather than kinetic energy.

The opposite limit occurs when U is a function only of y, and is called the "pure barotropic" case. This applied when variations with height can be neglected. When both types of instability are potentially present, a quantitative assessment of their relative importance can be made by using the disturbance energy equation. For the incompressible case, this equation can be deduced by starting from the disturbance form of the quasi-geostrophic momentum equations (12.2.24) and (12.2.25), which give

$$f_0 u_a = -\beta y u_g - (\partial/\partial t + U\,\partial/\partial x)v_g, \tag{13.6.2}$$

$$f_0 v_a = -\beta y v_g + (\partial/\partial t + U\,\partial/\partial x)u_g + (\partial U/\partial y)v_g, \tag{13.6.3}$$

where $U(y, z_*)$ is the undisturbed flow,

$$u_g = -f_0^{-1}\,\partial\Phi'/\partial y, \qquad v_g = f_0^{-1}\,\partial\Phi'/\partial x \tag{13.6.4}$$

is the disturbance geostrophic velocity, and (u_a, v_a) is the disturbance ageostrophic velocity. If u_g times (13.6.3) is subtracted from v_g times (13.6.2), and (13.6.4) is used, the result [cf. (12.2.33)] can be written

$$\tfrac{1}{2}(\partial/\partial t + U\,\partial/\partial x)(u_g^2 + v_g^2) + (\partial U/\partial y)u_g v_g + u_a\,\partial\Phi'/\partial x + v_a\,\partial\Phi'/\partial y = 0. \tag{13.6.5}$$

Adding $N_*^{-2}\,\partial\Phi'/\partial z_*$ times (12.9.6) and using (12.8.2), (13.6.4), (13.2.5), and (13.2.1) gives

$$\frac{1}{2}\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\left(u_g^2 + v_g^2 + \left(\frac{\alpha_* g\theta}{N_*}\right)^2\right) + \frac{\partial U}{\partial y}u_g v_g + \left(\frac{\alpha_* g}{N_*}\right)^2 \frac{\partial\Theta}{\partial y}v_g\theta$$

$$+ \frac{\partial}{\partial x}(u_a\Phi') + \frac{\partial}{\partial y}(v_a\Phi') + \frac{\partial}{\partial z_*}(w_*\Phi') = 0. \tag{13.6.6}$$

This can be averaged with respect to x over a wavelength, the average being denoted by an overbar and periodicity in x being assumed, to give on integration with respect to y and z_*

$$\frac{1}{2}\frac{\partial}{\partial t}\int\int \overline{\left(u_g^2 + v_g^2 + \left(\frac{\alpha_* g\theta}{N_*}\right)^2\right)}\,dy\,dz_* = -\int\int \frac{\partial U}{\partial y}\overline{u_g v_g}\,dy\,dz_*$$

$$-\int\int \left(\frac{\alpha_* g}{N_*}\right)^2 \frac{\partial\Theta}{\partial y}\overline{v_g\theta}\,dy\,dz_*. \tag{13.6.7}$$

The normal velocity is assumed to vanish on the boundary of the domain of integration. The right-hand side contains two terms representing sources or sinks of disturbance energy. Only the second arises in a "pure baroclinic" problem, in which this represents conversion of mean available potential energy to disturbance energy. Only the first arises in "pure barotropic" problems, in which it represents conversion of mean kinetic energy to disturbance energy. When both processes are active, the ratio of the terms may be taken to define the relative importance of the two effects.

The "purely barotropic" problem will now be illustrated by a simple f-plane example. When beta is zero, the problem has exactly the same form as in the nonrotating case, so the classical theory of the stability of unidirectional flows can be applied (Lin, 1955; Drazin and Howard, 1966; Drazin and Reid, 1981). The example chosen is that of a uniform shear flow

$$U = y\,dU/dy \qquad \text{for} \quad |y| < L \tag{13.6.8}$$

(with dU/dy constant) sandwiched between two regions of uniform flow as shown in Fig. 13.7. This problem was first studied by Rayleigh (1880) and is very similar mathematically to the Eady problem studied in Section 13.3. This is because the potential vorticity gradient $\partial\bar{q}/\partial y$, given by (12.9.4), is zero (except at $y = \pm L$), so (12.9.1) again reduces to (13.2.2), although this time there is no z_* variation. The solution has structure similar to that of (13.3.2), and for a growing mode it takes the

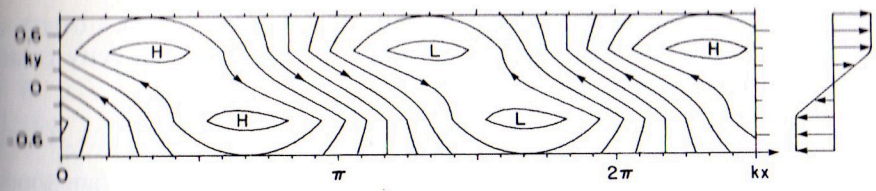


Fig. 13.7. Perturbation geopotential (or perturbation pressure) for the most unstable disturbance to the split-flow velocity profile shown at the right. The tilt of the phase lines is such that it is correlated with $-v$, i.e., if y points northward, eastward momentum is carried southward and westward momentum is carried northward.

form

$$\Phi' = (a\cos kx\sinh ky + b\sin kx\cosh ky)\exp(\sigma t) \qquad \text{for} \quad |y| < L. \tag{13.6.9}$$

For $|y| > L$, Φ' decays exponentially away from the value given by continuity of Φ' at $y = L$. The relationships between a, b, and σ follow from the requirement of continuity of v_a, as given by (13.6.3), at $y = \pm L$. They give

$$\begin{aligned} \sigma a &= b(dU/dy)(\tfrac{1}{2} - kL + \tfrac{1}{2}\exp(-2kL)), \\ \sigma b &= a(dU/dy)(kL - \tfrac{1}{2} + \tfrac{1}{2}\exp(-2kL)), \end{aligned} \tag{13.6.10}$$

and hence

$$\sigma^2 = (dU/dy)^2(\tfrac{1}{4}\exp(-4kL) - (\tfrac{1}{2} - kL)^2). \tag{13.6.11}$$

The maximum value for the growth rate σ is given by

$$\sigma_{\max} = 0.2012\,dU/dy \qquad \text{when} \quad kL = 0.3984, \tag{13.6.12}$$

and the corresponding solution is shown in Fig. 13.7. As for the Eady problem, instability occurs only for wavenumbers that are less than a cutoff value, which in this case is given by $kL = 0.6392$.

The geopotential perturbation Φ' shown in Fig. 13.7 has phase lines tilting in the direction opposite to the shear profile shown in the panel at the side, i.e., if eastward mean velocity increases to the north, phase lines tilt westward with increasing latitude. This is characteristic of a growing wave because (13.6.7) shows $\overline{u_g v_g}$ must have the sign opposite to dU/dy for growth. Equations (13.6.4) then show that where Φ' increases with x, the y gradients of Φ' and U must have the same sign, and this implies that phase lines tilt the way opposite to that of the shear flow. In the present example $\overline{u_g v_g}$ has a constant value for $|y| < L$ and is zero for $|y| > L$. The sign is that associated with bringing momentum from outside the shear layer in toward the center of the layer. This has the effect of reducing the mean flow energy by transfer to the disturbance.

The above example is useful for illustrating the characteristics of barotropic instability, but for geophysical applications the beta effect is often of prime importance. The required modification is discussed by Kuo (1949, 1973). The first condition (13.5.2), found in Section 13.5, shows that barotropic zonal flow is stable if

$$\beta - d^2U/dy^2$$

does not change sign, i.e., if the maximum value of the vorticity gradient d^2U/dy^2

is less than β. To give an indication of the implications of this, define a length scale L and velocity scale ΔU such that $\Delta U/L^2$ equals the maximum vorticity gradient. The instability is possible only when

$$L < (\Delta U/\beta)^{1/2}, \tag{13.6.13}$$

i.e., for atmospheric examples with $\Delta U = 20$ m s^{-1}, L must be less than 1000 km for instability, whereas for an oceanic example with $\Delta U = 20$ cm s^{-1}, L must be less than 100 km for barotropic instability to be possible.

Stability calculations have also been made for barotropic planetary waves (Lorenz, 1972; Gill, 1974; Coaker, 1977; Ripa, 1981). Except for the special case of zonal flow, planetary waves on an infinite beta plane are always unstable (Gill, 1974). If their inverse wavenumber is well below the value given by (13.6.13) (or, looked at from another viewpoint, if their amplitude ΔU is sufficiently large), the instability is just like that of parallel flow in the absence of beta. For large scales (or small amplitude), on the other hand, the unstable disturbance consists of two waves that form a resonant triad (see Section 8.13) with the primary wave. The results have been generalized to a two-layer system by Jones (1979). On a sphere, the geometric constraints reduce the possibilities of resonant interactions, so not all waves are unstable (Hoskins, 1973; Baines, 1976).

In practical examples, there is often a mix of baroclinic and barotropic effects. Parameters that affect the situation are the ratios of the length scale L on which the flow varies to the Rossby radius and to the beta scale (13.6.13). In oceanographic examples the ratio of the depth scale, on which the mass flow varies, to the ocean depth is also important, and Killworth (1980a) has discussed the stability properties for the variety of limiting cases that are possible. The presence of side boundaries is also important in the case of boundary currents like the Gulf Stream, and bottom topography also has an effect. Laboratory experiments on the stability of boundary currents are discussed, e.g., by Griffiths and Linden (1981).

The remaining sections show how disturbances can grow spontaneously in certain circumstances, and there are many other examples that occur in nature. In fact instabilities are very common and give rise to such natural features as turbulence in wind and water, both tropical and extratropical storms, a great variety of cloud forms, rain bands, thermals, etc. The instabilities are not studied here, but the following authors have written pertinent books: Betchov and Criminale (1967), Chandrasekhar (1961), Charney (1973), Drazin and Reid (1981), Gossard and Hooke (1975), Lilly (1979), Turner (1973), Wallace and Hobbs (1977), Woods (1982), and Yih (1980).

13.7 Eddies in the Ocean

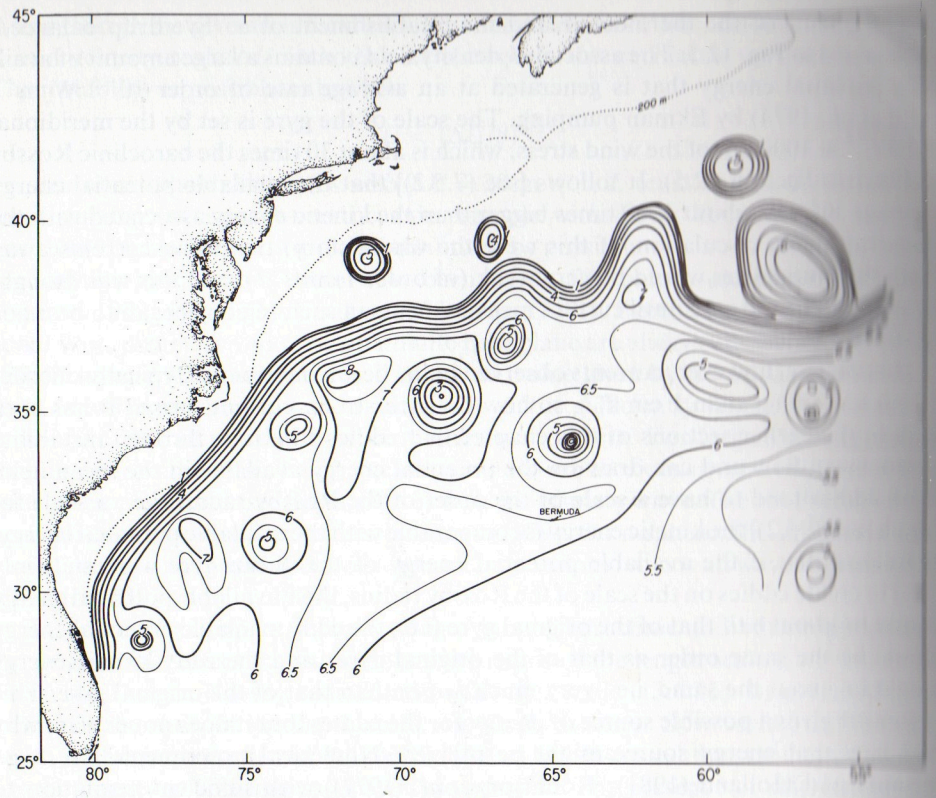
Many aspects of the dynamics of the ocean circulation were discussed earlier in Sections 10.14 (eastern boundary currents), Sections 11.12, 11.14, and 11.16 (tropical currents), and Sections 12.5 and 12.6 (extratropical currents). In particular, it was seen that in mid latitudes, Ekman pumping (introduced in Section 9.4) causes vertical

displacements of the thermocline and the establishment of a "Sverdrup balance," as depicted in Fig. 12.5. The associated density field contains a large amount of available potential energy that is generated at an average rate of order 10^{-3} W m^{-2} (Gill *et al.*, 1974) by Ekman pumping. The scale of the gyre is set by the meridional scale $l^{-1} \approx 1000$ km of the wind stress, which is about 30 times the baroclinic Rossby radius (see Section 12.5). It follows [see (7.5.2)] that the available potential energy is about 30^2, i.e., about 1000 times bigger than the kinetic energy associated with the Sverdrup mean circulation. If this were the whole story, the ocean currents away from the boundaries would be very weak (of order 1 cm s^{-1}), and this was thought by many to be the case until current measurements in such regions began to be made in the late 1950s.

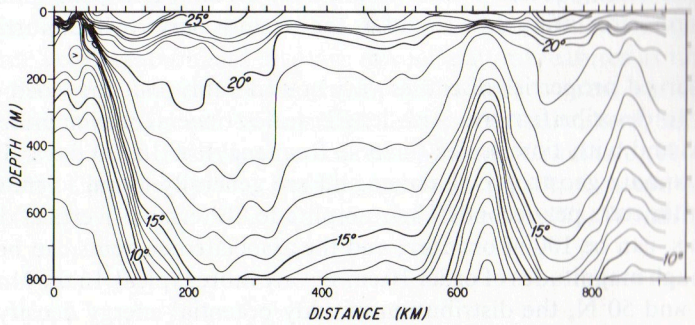
In practice, however, currents observed in the ocean interior are typically of order 10 cm s^{-1} rather than 1 cm s^{-1}, so how do such strong currents arise? It has been seen in the earlier sections of this chapter that eddies can arise through instability of the mean flow and can draw on the potential energy available in the mean field. Such eddies tend to have a scale of the order of the Rossby radius, i.e., a scale for which [see (7.5.2)] the kinetic energy is comparable with the available potential energy. It follows that if the available potential energy of the ocean gyre were suddenly used to create eddies on the scale of the Rossby radius, their available potential energy would be about half that of the original gyre (i.e., the eddy available potential energy would be the same order as that of the original gyre), and the eddy kinetic energy would be about the same, i.e., very much *larger* than that of the original gyre. This argument gives a possible source of energy for the eddies, but it does not explain why and how that energy source might be utilized. Numerical experiments [see, e.g., Schmitz and Holland (1981); Robinson *et al.* (1977)] with sufficient resolution to simulate eddies demonstrate how these can be generated and also contain many features that can be compared with observation. In particular, the geographical distribution of eddies indicates that the major production zones for eddies are in the regions of strong currents such as the Gulf Stream. Also, the models indicate that the deep recirculating flow found in the western basin of the North Atlantic is eddy-driven.

The observed properties of eddies vary considerably, but the word eddy is generally used to describe features with length scales (inverse wavenumbers) of order 10–100 km, and with time scales (inverse frequency) of 10–30 days. They tend to be in approximate geostrophic balance and are generally found to move westward at a few centimeters per second. Their amplitude, in terms of vertical displacement of isopycnals, can be 100 m or more, and the associated currents can be 1 m s^{-1} or more, although magnitudes of order 10 cm s^{-1} are more typical. In the North Atlantic between 0° and 50°N, the distribution of eddy potential energy density, calculated by Dantzler (1977), indicates that the largest values (500–2000 cm^2 s^{-2}) are confined to the general area of the Gulf Stream. Although there is little doubt that eddies in this region are due mostly to instability (Schmitz and Holland, 1981), other generation mechanisms such as wind forcing (see Section 9.11) or flow over topography (Sections 8.7–8.10) could be significant elsewhere (Müller and Frankignoul, 1981).

Instabilities may take a variety of forms, and the mechanisms are often difficult to identify observationally. However, one form clearly seen in observations is due



(a)



(b)

Fig. 13.8. (a) Chart of the depth, in hundreds of meters, of the 15° isothermal surface, showing the Gulf Stream, nine cyclonic rings, and three anticyclonic rings. Contours are based on data obtained between 16 March and 9 July 1975. (b) A temperature section through the Gulf Stream and two cyclonic (cold-core) rings south of the stream. The section is a "dog-leg" from 36°N, 75°W (left-hand end) to 35°N, 70°W (middle) and then to 37°N, 65°W (right-hand end). [From Richardson *et al.* (1978, Figs. 1a and 4a, Section 3).]

to Gulf Stream meanders forming large loops that pinch off and form Gulf Stream "rings." If they break off on the equatorward side, they have cold water in the middle (cold-core eddies), and the Labrador Sea water they contain may be trapped in the eddy for a year or so, in which time the ring may move a considerable distance to the west and south. Similarly, warm-core eddies form on the poleward side and drift westward also, sometimes being reabsorbed into the Gulf Stream. For example, Fig. 13.8 shows the Gulf Stream, nine cold-core (cyclonic) rings and three warm-core (anticyclonic) rings observed in the spring of 1975, together with a cross section through the Gulf Stream and two of the rings. A summary of the properties of rings is given by Richardson (1982).

Further information about eddies may be found in the survey article by Wunsch (1981) and in the comprehensive volume on eddies edited by Robinson (1982). Reviews of the observational base are given by Richman *et al.* (1977) and the MODE Group (1978). Modeling of eddies and their interactions with the mean flow is discussed by Rhines (1977, 1979) and by Rhines and Holland (1979). Analogies between ocean eddies and baroclinic disturbances in the atmosphere are considered by Charney and Flierl (1981).

13.8 Fronts

The solutions found in Section 13.3 nicely illustrate the *initial* development of baroclinic disturbances, but sooner or later effects that have been neglected in the model will come into play and give rise to new features. One is the development of the sharp fronts that are such a familiar feature of surface weather charts. These have a variety of forms and develop in a variety of ways. Descriptions are given e.g., by Wallace and Hobbs (1977) and by Palmén and Newton (1969); many ideas of structure go back to Bjerknes (1919). Here attention will be concentrated on one illustrative example, namely, the fronts formed by the fastest-growing Eady wave (see Fig. 13.4). This is somewhat special because the disturbance is independent of y and in fact satisfies the nonlinear quasi-geostrophic equations. However, the quasi-geostrophic approximation itself breaks down when the ageostrophic velocity component u_a becomes comparable with its geostrophic counterpart, which in this case is U. Then the derivative D/Dt, following the motion, can no longer be approximated by D_g/Dt.

A formal procedure for finding the equations that apply in the neighborhood of a front is to introduce appropriate nondimensional variables (in a frame of reference moving with the front), allowing for the fact that the cross-front scale L_x is small compared with the along-front scale L_y. Suitable scales for t, z_*, u, v, w, Φ', and θ are $(L_y/L_x)f^{-1}$, $(f/N_*)L_x$, $(L_x/L_y)fL_x$, fL_x, $(fL_x)^2/N_*L_y$, $(fL_x)^2$, and $fN_*L_x/\alpha_* g$. The result of introducing these scalings is that the only term that can be neglected in the equations is the acceleration term Du/Dt in the x component of the momentum equation, this term being of order $(L_x/L_y)^2$ relative to the Coriolis term. Thus flow along the front is in geostrophic balance with the cross-front pressure

gradient, i.e.,

$$v = v_g = f^{-1}\,\partial\Phi'/\partial x. \tag{13.8.1}$$

The cross-front flow is *not* in geostrophic balance, but it is still useful to express it as the sum of a geostrophic part and an ageostrophic part u_a, where

$$fu_a = -Dv/Dt = -Dv_g/Dt, \tag{13.8.2}$$

the equality on the far right using (13.8.1). The operator D/Dt has its full three-dimensional form (4.1.7), i.e., it includes the term $w_*\,\partial/\partial z_*$. The remaining equations to be satisfied are the hydrostatic equation (6.17.20), the continuity equation (12.8.2), and the potential temperture equation (4.10.8). These equations were used by Sawyer (1956) and Eliassen (1962) to discuss the vertical circulation at fronts.

It so happens that this set of equations can be transformed into the quasi-geostrophic equations by a change of variables (Eliassen, 1962; Hoskins and Bretherton, 1972; Hoskins and Draghici, 1977), i.e., by a replacement of the variables on the left of the following tabulation with those on the right (denoted by changing lowercase into capitals and vice versa)

$$\begin{aligned}
\Phi' \qquad & \phi' = \Phi' + \tfrac{1}{2}v^2,\\
x \qquad & X = x + v_g/f,\\
z_* \qquad & Z_* = z_*,\\
t \qquad & T = t,\\
u_a \qquad & U_a = u_a + (fJ)^{-1} w_*\,\partial v/\partial z_*,\\
w_* \qquad & W_* = w_*/J\\
N_*^2 \qquad & n_*^2 = \frac{\alpha_* g}{f}\left[\left(f + \frac{\partial v}{\partial x}\right)\frac{\partial\theta}{\partial z_*} - \frac{\partial v}{\partial z_*}\frac{\partial\theta}{\partial x}\right],
\end{aligned} \tag{13.8.3}$$

where

$$J = 1 + \frac{1}{f}\frac{\partial v}{\partial x} = \frac{\partial X}{\partial x} = \frac{\partial(X, Z_*)}{\partial(x, z_*)} \tag{13.8.4}$$

is the Jacobian of the coordinate transformation. The dependent variables θ, u_g, and v are unchanged, so there is no need to introduce new symbols for them. The independent variables z_* and t are not changed either, but capital letters are used to indicate that partial derivatives such as $\partial/\partial Z_*$ that involve capital letters, are for fixed X, whereas $\partial/\partial z_*$ is for fixed x.

The proof of the result about the transformation is a matter of manipulation of partial derivatives. For instance, by the definition (13.8.3) of X, it follows that its derivative, following the motion, is

$$DX/Dt = Dx/Dt + f^{-1}\,Dv/Dt = u + f^{-1}\,Dv/Dt = u_g, \tag{13.8.5}$$

the last equality using (13.8.2) and the definition of u_a. It follows that the expression

for the derivative following the motion is, in the new coordinates,

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial T} + \frac{DX}{Dt}\frac{\partial}{\partial X} + v\frac{\partial}{\partial y} + \frac{DZ_*}{Dt}\frac{\partial}{\partial Z_*} = \frac{D_g}{Dt} + w_*\frac{\partial}{\partial Z_*}, \tag{13.8.6}$$

where D_g/Dt is defined by

$$D_g/Dt = \partial/\partial T + u_g\,\partial/\partial X + v\,\partial/\partial y \tag{13.8.7}$$

and $\partial/\partial T$ represents the time derivative with X (not x) kept constant. The transformed versions of (13.8.2) and (4.10.8) are now

$$fU_a = -D_g v/Dt, \qquad D_g\theta/Dt + (n_*^2/\alpha_* g)W_* + v\,\partial\Theta/\partial y = 0, \tag{13.8.8}$$

and the remaining equations are

$$fv = \partial\phi'/\partial X, \qquad \alpha_* g\theta = \partial\phi'/\partial Z_*, \qquad \partial U_a/\partial X + \partial W_*/\partial Z_* = 0. \tag{13.8.9}$$

These follow by expanding derivatives with respect to X and Z_* in terms of x and z_* derivatives, using for any function F

$$\frac{\partial F}{\partial Z_*} = \frac{\partial(F, X)}{\partial(Z_*, X)} = \frac{1}{J}\frac{\partial(F, X)}{\partial(z_*, x)} = \frac{1}{J}\left[\frac{\partial F}{\partial z_*}\left(1 + \frac{1}{f}\frac{\partial v}{\partial x}\right) - \frac{\partial F}{\partial x}\frac{1}{f}\frac{\partial v}{\partial z_*}\right] \tag{13.8.10}$$

and a similar result for $\partial F/\partial X$.

Now the quantity n_*^2 is proportional to Ertel's potential vorticity Q, defined by (7.11.17), and therefore is conserved following the motion by (7.11.13). The semi-geostrophic equivalent can be derived from the above equations. For the Eady problem, n_*^2 is uniform for the undisturbed flow and thus remains uniform throughout the motion. It follows [as was first shown by Hoskins and Bretherton (1972)] that the solution in the transformed coordinate has the same form (13.3.15) as that found in section 13.3. In particular θ and v are given on a section $y = 0$ by

$$\alpha_* g\theta = n_*^2 Z_* + \frac{\exp(\sigma T)}{H_R}\left[\cos kX\,\frac{\cosh(Z_*/H_R)}{\sinh(H/H_R)} + \sin kX\,\frac{\sinh(Z_*/H_R)}{\cosh(H/H_R)}\right], \tag{13.8.11}$$

$$f^2(X - x) \equiv fv = k\exp(\sigma T)\left[-\sin kX\,\frac{\sinh(Z_*/H_R)}{\sinh(H/H_R)} + \cos kX\,\frac{\cosh(Z_*/H_R)}{\cosh(H/H_R)}\right]. \tag{13.8.12}$$

This solution is shown in Fig. 13.9, which was constructed by first calculating contours in (X, Z_*) space and then using (13.8.12) to transform to (x, z_*) space. It has the property that a singularity develops on the boundaries (at the points marked L) after a finite time, the singularities corresponding to the formation of fronts with infinite gradients. The singularity on the ground is at the point where $J = \infty$, i.e., $\partial x/\partial X = 0$, i.e., by (13.8.12) it occurs at $kX = \frac{7}{4}\pi$ when $k^2\exp(\sigma T) = 2^{-1/2}f^2$. Figure 13.9 shows the solution when the disturbance amplitude is 90% of the value at this time. For comparison, Fig. 13.10 shows an observed section through a front, and the

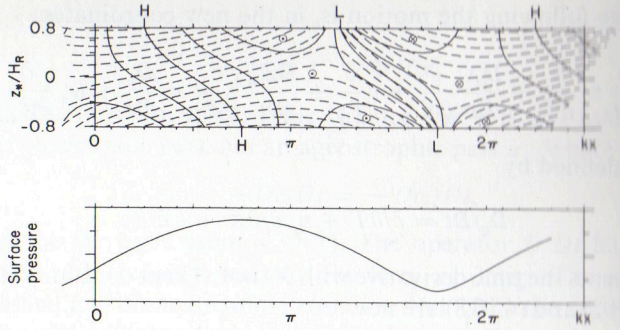


Fig. 13.9. Large-amplitude solution for the fastest-growing (two-dimensional) Eady wave. Contours are of normal velocity (solid) and potential temperature (dashed) in a vertical section. The amplitude of the perturbation is 90% of the value at which singularities develop at the low pressure (L) points on the boundary. At these points, which corresponds to the fronts, the normal velocity is zero. The contour interval in potential temperature is the same as in the small-amplitude solution shown in Fig. 13.4c. The solution was obtained by coordinate transformation from the latter, so the boundaries of the contoured region do not coincide with the frame. The vertical exaggeration is N_*/f. Note that the velocity and potential temperature contours become nearly parallel at the front. The lower panel shows the corresponding variation of surface pressure with distance x. Note the sharpness of the trough as compared with the broadness of the ridge.

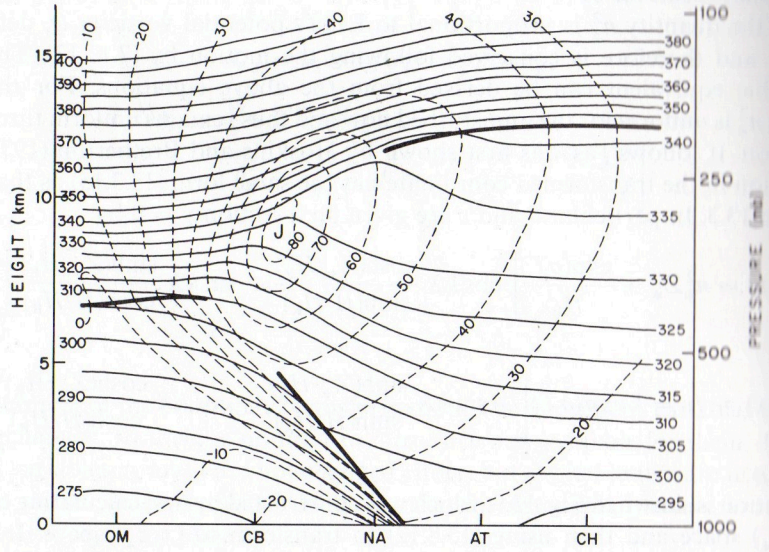


Fig. 13.10. A vertical cross section through a frontal zone from Omaha, Nebraska (OM), to Charleston, South Carolina (CH), showing contours of velocity in meters per second (dashed) and potential temperature in degrees Kelvin (solid). The horizontal distance across the section is about 2000 km. The time was 00 GCT on 20 November 1964. The figure is from Wallace and Hobbs (1977, Fig. 3.20), where further details are described. The front may be compared with the one obtained for an ideal fluid in Fig. 13.9.

similarities in the lower troposphere are apparent. (The structure near the tropopause can be modeled in a similar way, as discussed below.)

There are several features of the front that are independent of the details of the formation process. By definition, the front is a region of large gradients, so the dominant terms in the potential vorticity balance [the last equation of (13.8.3)] give

$$\frac{\partial v}{\partial x}\frac{\partial \theta}{\partial z_*} - \frac{\partial v}{\partial z_*}\frac{\partial \theta}{\partial x} \equiv \frac{\partial(v,\theta)}{\partial(x,z_*)} = 0, \tag{13.8.13}$$

i.e., *v and θ contours tend to coincide* in the front. This feature can be seen in both the theoretical and the observed fronts. Also, potential vorticity conservation requires $f + \partial v/\partial x$ to remain positive, so it can be large only when the relative vorticity $\partial v/\partial x$ is large and *cyclonic*. To give an idea of possible magnitudes, detailed observations by Sanders (1955) of a front in the United States showed $\partial v/\partial x$ with a value of $10f$ at a height of 300 m, and $\partial\theta/\partial x$ at the same height was 0.5 deg km^{-1}.

Another general statement concerns the cross-frontal flow, for (13.8.2) gives on the surface

$$\partial v/\partial t + u(f + \partial v/\partial x) = u_g = -f^{-1}\,\partial\Phi'/\partial y. \tag{13.8.14}$$

Assuming that the along-front pressure gradient and along-front acceleration do not change much, the cross-front flow u must become small when $f + \partial v/\partial x$ becomes large. This statement can be made more precise for the Eady case since v (and hence $\partial v/\partial t$) is zero at the point L of minimum pressure and maximum cyclonic vorticity at which the front forms. From Fig. 13.4b it can be seen that the ageostrophic flow opposes the geostrophic flow at this point. The latter remains constant, but u_a increases exponentially with time so (13.8.14) implies that $(f + \partial v/\partial x)$ will become infinite in a finite time. Because there is little cross-front flow, the poleward flow ahead of the front continues to bring in warm air, whereas the equatorward flow behind brings in cold air, and so builds up the temperature contrast even further. Other features of the front shown in Fig. 13.9 follow from the properties of baroclinic disturbances, e.g., the slope to the west with height, with a value for the slope of order f/N. Typical observed values are of this order [the value is about f/N in Fig. 13.10 and was about 0.3 f/N in the case studied by Sanders (1955)].

In reality, infinite gradients are not observed, so clearly other effects such as y variations (see below), friction, mixing, and latent heat release modify the structure of fronts, which can become quite complicated when viewed on a small scale [see, e.g., Bennetts and Hoskins (1979)]. The maximum velocity in the solution of Fig. 13.9 is $0.90k^{-1}f = 1.1N_*H$ or about 100 m s^{-1}. Sanders found changes in *geostrophic* velocity v_g of this magnitude across the front, but the change in the *measured* velocity v was only 20 m s^{-1} at 300 m. At 1200 m the changes in measured and geostrophic velocities were both about 35 m s^{-1}. One obvious discrepancy in the theory is the neglect of the surface boundary layer. The large shear (see Section 9.5) leads to a large Ekman flux into the surface low (Sanders measured convergence at a rate equal to $5f$ at 300 m) that enhances the temperature gradient and leads to increased upward velocity at the front. Descriptions of the vertical motion observed

at fronts have been given, e.g., by Browning *et al.* (1975). Blumen (1980) has added surface friction effects to the solution shown in Fig. 13.9 and has made comparisons with Sanders' observations. Friction effects outside the boundary layer have been considered by Gill (1981) in another context and by Garrett and Loder (1981) in connection with fronts in the ocean. Shapiro (1981) has discussed the effects of turbulent fluxes in the free atmosphere and R. T. Williams (1974) obtained a steady-state front numerically. Latent heat release may also be important and has been modeled, e.g., by Orlanski and Ross (1978).

The development of fronts in *three-dimensional flows* can be handled by the same methods as those used above. The y component of the momentum equation is approximated in the same way as is (13.8.2), i.e.,

$$fv_a = Du_g/Dt, \tag{13.8.15}$$

and will provide a good approximation in frontal regions, whatever their orientation, provided that the along-front scale is large compared with the cross-front scale. The approximation is known as the *geostrophic momentum* approximation, first introduced by Eliassen (1949), and the equations are known as the semigeostrophic equations. The equations can be transformed into the quasi-geostrophic equations (Hoskins, 1975; Hoskins and Draghici, 1977) by using the variables of (13.8.3) and also by using

$$Y = y - u_g/f. \tag{13.8.16}$$

The coordinates X, Y, called *geostrophic coordinates*, were used by Yudin (1955) and the transformation has been further discussed by Blumen (1981). The quantity n_*^2 is proportional to Ertel's potential vorticity as before, and J is the ratio of total vorticity to f. The approximation has been applied to the development of the square Eady wave (shown in Fig. 13.5) by Hoskins (1976). Hoskins and West (1979) added the effect of horizontal shear in the mean flow, and the types of fronts that develop have been discussed by Hoskins and Heckley (1981). The mathematical theory of frontogenesis has been reviewed by Hoskins (1982).

As Figs. 13.9 and 13.10 indicate, fronts can also form on the tropopause. The rigid lid model is not very good for describing their development, but representation of the tropopause as a discontinuity in N_*^2 (or rather n_*^2) gives remarkably realistic results, as found, e.g., by Hoskins (1972). A significant feature of these fronts is the descent of a tongue of stratospheric air well below normal tropopause heights, and such tongues can be traced, e.g., by measuring ozone concentration. Descriptions may be found, e.g., in Reed and Danielsen (1959) and Shapiro (1974).

In the ocean, fronts are produced by a variety of mechanisms. Sometimes they are very distinct in the temperature and salinity fields but *not* in the density field, and this distinction is important so far as the dynamics is concerned. A survey of temperature fronts as seen by satellites is given by Legeckis (1978). Figure 13.11 shows the main climatological frontal regions (i.e., regions where fronts are most commonly observed) in the North Pacific, and these are discussed by Roden (1975). One important type of front is that produced by *Ekman convergence* in the surface layer, the *subtropical fronts* found at about 30°N and 40°S being important examples. Roden and Paskausky (1978) have studied changes in such fronts due to changing Ekman

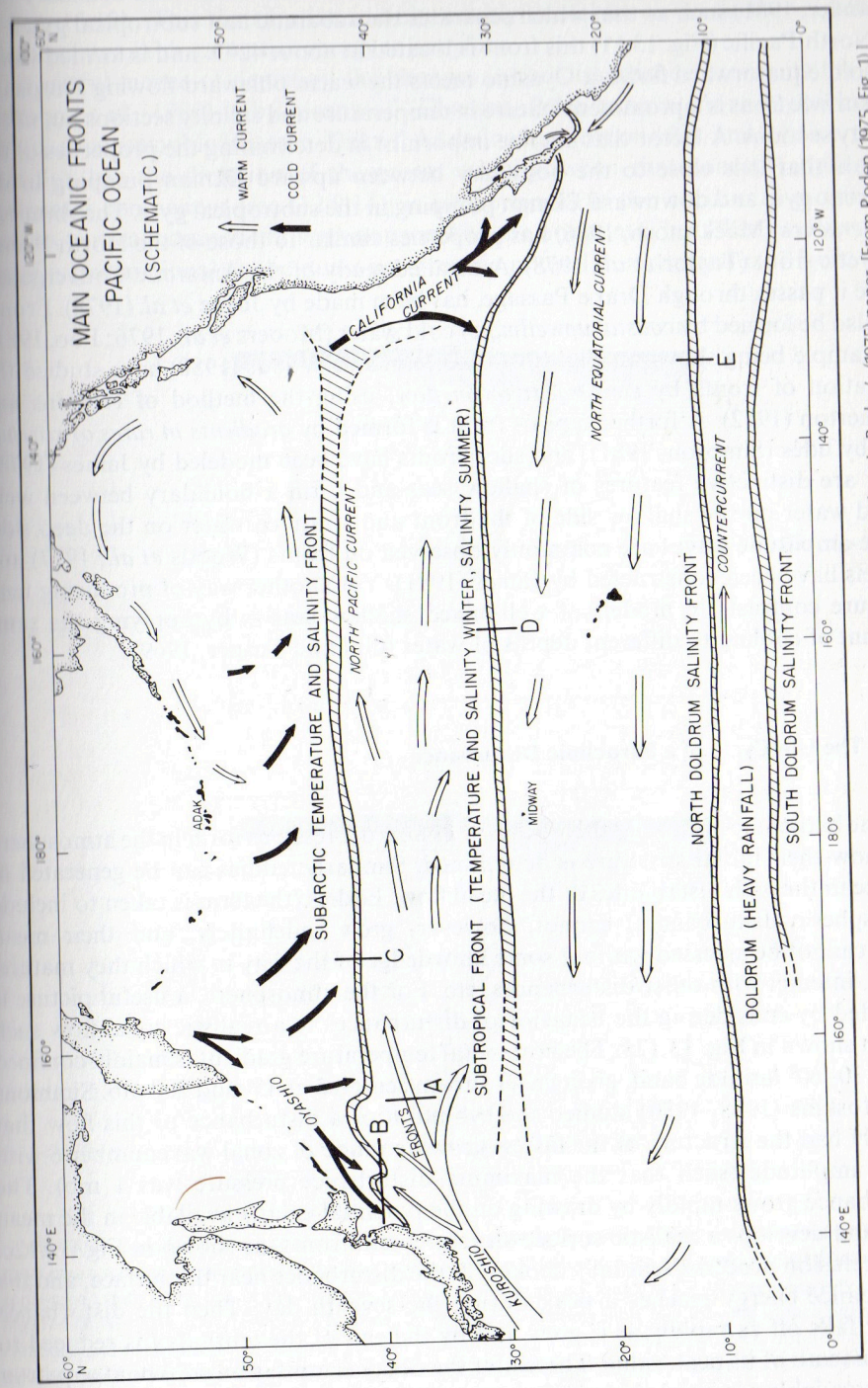


Fig. 13.11. A schematic map of the main North Pacific fronts. The cross sections marked A–E are given in Roden's (1975) paper. [From Roden (1975, Fig. 1).]

convergence. A second type of front is that formed at a *water mass boundary* (see Welander, 1981) such as that which separates the subarctic and subtropical gyres. In the North Pacific (Fig. 13.11) this front is located at about 42°N and is formed where the cold equatorward flowing Oyashio meets the warm poleward flowing Kuroshio. The surface front is a prominent feature in temperature and salinity sections but not in density sections. A factor that may be important in determining the properties of the front is that it is close to the boundary between upward Ekman pumping in the subarctic gyre and downward Ekman pumping in the subtropical gyre. The *Antarctic Convergence* (Mackintosh, 1946) has properties similar to those of the North Pacific subarctic front (Taylor *et al.*, 1978). A detailed study of the Antarctic Convergence, where it passes through Drake Passage, has been made by Joyce *et al.* (1978). Fronts can also be formed by *coastal upwelling* of cold water (Mooers *et al.*, 1976; Foo, 1981), an example being shown in Fig. 10.16. McVean and Woods (1980) have studied the formation of fronts by *convergent mean flow*, using the method of Hoskins and Bretherton (1972). A further type of front is formed by *gradients in rates of stirring*, e.g., by tides (Simpson, 1981), and such fronts have been modeled by James (1978). They are distinctive features of shallow seas and form a boundary between well-mixed water on the shallow side of the front and stratified water on the deep side. Large-amplitude waves are commonly observed on fronts (Woods *et al.*, 1977) and models have been constructed by James (1981). Yet another way of producing temperature contrasts in models of well-mixed shallow seas is by applying the same amount of cooling to different depths of water (Gill and Turner, 1969).

13.9 The Life Cycle of a Baroclinic Disturbance

The baroclinic instability theory indicates how depressions form in the atmosphere and how their initial structure is determined. Similarly, eddies can be generated in the ocean through instabilities of the mean flow. Eddies (the term is taken to include atmospheric disturbances) cannot, however, grow indefinitely, and their mean effect cannot be assessed without some knowledge of the way in which they mature, decay, interact with other disturbances, etc. For the atmosphere, a useful picture is provided by considering the behavior of disturbances to a realistic zonal flow such as that shown in Fig. 13.12a. The horizontal temperature gradient is mainly confined to the 30–60° latitude band, giving a jet stream centered at 45° and 200 mb. Simmons and Hoskins (1978, 1980) studied the evolution of a disturbance to this flow that initially had the structure of the fastest growing mode of zonal wavenumber 6 with small amplitude (such that the maximum disturbance pressure was 1 mb). The disturbance grows rapidly by drawing on the potential energy available in the mean flow, and develops a realistic surface structure with fronts, as shown in Fig. 13.12c. The occlusion process eventually chokes off the disturbance near the surface, and the disturbance energy reaches a peak during the seventh day. Then the disturbance energy falls off as rapidly as it grew and by the end of the tenth day is reduced to about a tenth of its peak value. Thereafter the decay is much slower. About a quarter of the available potential energy is released during this process, the zonal mean flow

and zonal mean temperature distribution at the end of the life cycle being as shown in Fig. 13.12b. This distribution is only weakly unstable, so no further significant release of available potential energy takes place.

A useful means of depicting the changes in the disturbance during its life cycle (Edmon *et al.*, 1980) is by means of "Eliassen–Palm cross sections," which show the quasi-geostrophic Eliassen–Palm flux, defined in Section 12.9, and its divergence. This flux has the direction of the group velocity (when that concept applies); its horizontal component [see (12.9.14)] is proportional to the horizontal momentum flux, its vertical component is proportional to the horizontal heat flux, and its divergence is proportional to the quasi-geostrophic potential vorticity flux. Figure 13.13

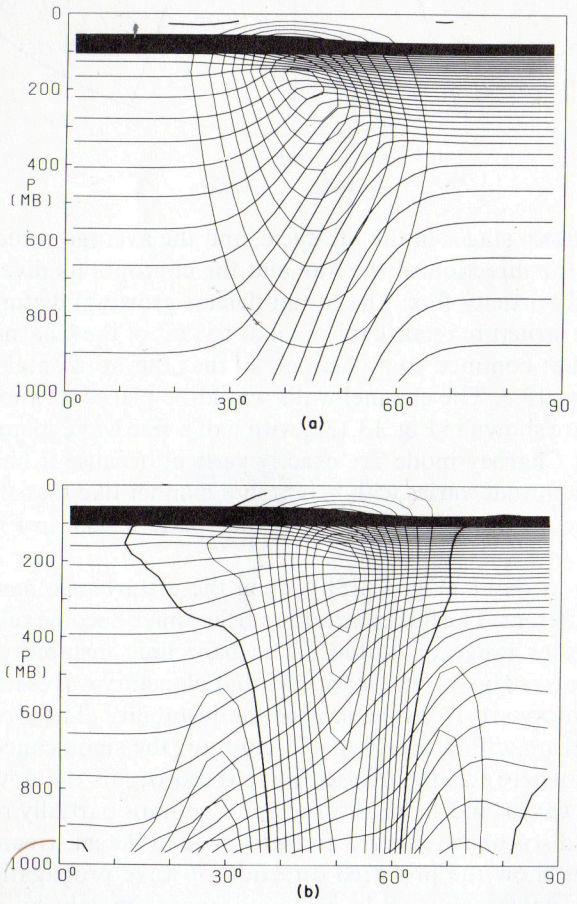
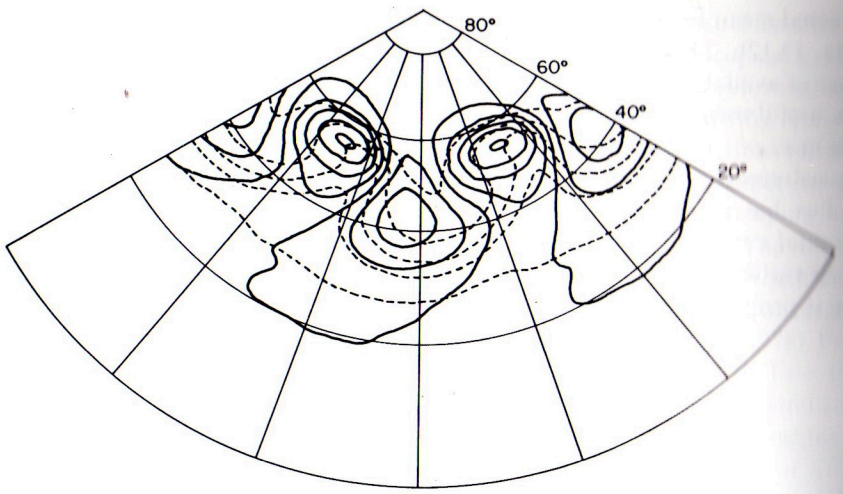


Fig. 13.12. Meridional cross sections of zonal-mean potential temperature and zonal wind for (a) The basic zonal-mean state at the beginning of the life cycle experiment. [From Simmons and Hoskins (1980, Fig. 1a).] (b) The zonal-mean state at the end of the life cycle. Contour intervals are 5 K and 5 m s^{-1}. The zero velocity contour is drawn relatively dark. [Courtesy of B. J. Hoskins.] (c) North polar stereographic plot showing surface pressure (solid contours) and near-surface temperature (dashed contours) after 5 days of integration for a disturbance of zonal wavenumber 6 to the flow shown in (a). The initial surface pressure amplitude was 1 mbar and the initial perturbation was the fastest-growing normal mode. Contour intervals are 8 mb and 8 K.



(c)

Fig. 13.12. *(continued)*

shows this flux at two stages in the life cycle, and the average values over the cycle. The arrows give the direction of the flux and the contours its divergence, i.e., they give the potential vorticity flux. The initial (fastest-growing) disturbance, shown in Fig. 13.13a, has a structure remarkably similar to that of the Charney mode, studied in Section 13.4, but confined to a channel, so that the horizontal planform is like that shown in Fig. 13.5. The channel walls would be placed at about 35 and 55° in match the structure shown in Fig. 13.13a, with half a sine wave spanning the channel. The arrows for a Charney mode are exactly vertical because it has no momentum flux, and their magnitude varies with height in a manner like that shown in the inset at the right of the lower panel in Fig. 13.6. The pattern seen in Fig. 13.13a is very close to this.

The structure at day 8 (Fig. 13.13b), when the disturbance has begun to decay rapidly, is very different. The fluxes near the surface have become relatively small and spread over a wider range of latitude. The baroclinic instability process is now ceasing to operate (see above), and it appears that planetary waves have been radiated upward and equatorward from the seat of the instability. The strongest fluxes are now in the jet stream within the 10 m s^{-1} contour, the significance of this contour being that this is where the disturbance wave speed equals the flow speed. At such places planetary waves tend to be absorbed and perhaps partially reflected (see Section 8.9), so that disturbance energy can be trapped in the jet stream by the "critical line." The arrows show the preferred direction of wave propagation for these disturbances (which presumably could be found approximately by ray-tracing techniques), namely, equatorward and upward. This determines where the waves are absorbed, namely, on the equatorial flank of the jet stream just below the tropopause. The energy lost from the disturbance as it propagates to regions where the velocity relative to the wave is weaker is largely converted to mean flow kinetic energy. A calculation of the rates of transfer of disturbance energy to zonal kinetic energy in fact shows that it begins to rise rapidly at the end of the sixth day, reaches a maximum 2 days later (the time corresponding to Fig. 13.13b), then falls to near zero during the

next 2 days. The resultant zonal flow at the end of the cycle is shown in Fig. 13.12b. The westerlies are too strong at the surface, but this can be rectified by including surface friction (Simmons and Hoskins, 1980).

The mean Eliassen–Palm cross section for the cycle is shown in Fig. 13.12c. The upward arrows represent a poleward heat flux, and this is large over the region of large initial horizontal gradient. The equatorward pointing component of the flux represents a *poleward* transfer of westerly momentum, corresponding to equatorward planetary-wave propagation [see (12.9.14)]. The flux is particularly large between 150 and 400 mb within the jet stream. This transfer is extremely important in

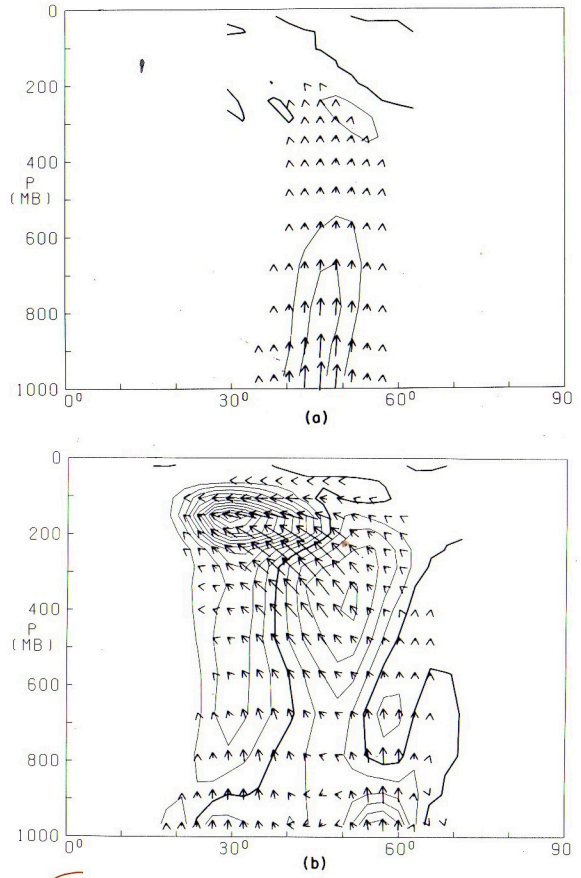
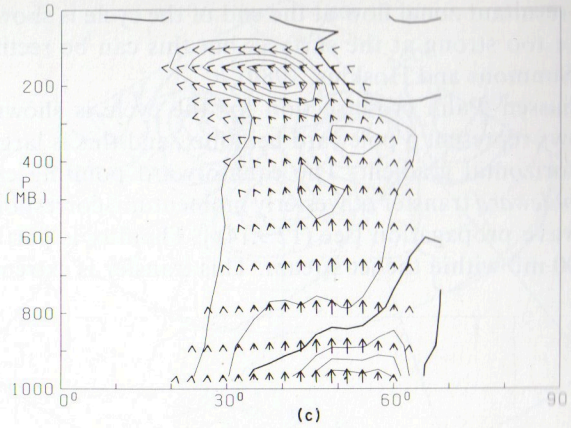


Fig. 13.13. Eliassen–Palm cross sections during the life cycle of a baroclinic disturbance. The vectors, shown by arrows, give the Eliassen–Palm flux (see Section 12.9), which for linear waves is in the direction of the group velocity. Because of the coordinates used, the horizontal and vertical components of the vector are defined, respectively, by $-2\pi r^2 g^{-1}\overline{u'v'}\cos^2\varphi$ and $2\pi r^3 g^{-1}\overline{v'T'}\cos^2\varphi/(\partial\Theta/\partial p)$, where r is the earth's radius, φ the latitude, g the acceleration due to gravity, and f the Coriolis parameter. [This is $2\pi r^2\cos^2\varphi/\rho_s g$ times the Eliassen–Palm flux as defined in (12.9.14).] With this choice, the "divergence," defined as the sum of the φ derivative of the horizontal component and the p derivative of the vertical component, is equal to $2\pi r^3 g^{-1}\overline{v'q'}\cos^2\varphi$ and therefore is proportional to the quasi-geostrophic potential vorticity flux. The "divergence," so defined, is shown by contours [contour interval 1.5 × 10^{15} m^3 for (a) and (b) and 4 × 10^{15} m^3 for (c)]. The zero contour is drawn dark. (a) At day zero for the life cycle, i.e., represents the fastest-growing mode; (b) at day 8; and (c) the average over the cycle. [from Edmon *et al.* (1980, Fig. 3).]



(c)

Fig. 13.13. (*continued*)

the angular momentum budget of the earth, and is responsible for the midlatitude surface westerlies that are such an important feature of the circulation. Further discussion may be found in Section 13.10.

13.10 General Circulation of the Atmosphere

13.10.1 Introduction

An introductory discussion about the way in which the atmosphere responds to the radiative driving from the sun was given at the beginning of the book, and it seems fitting to conclude with a further discussion in light of the concepts that have been developed in the intervening chapters. In Chapter 1 it was shown that a purely radiative equilibrium is unstable in the sense that air at the surface would be lighter than that above the surface. Thus convection takes place over the depth of the troposphere and radiative–convective equilibrium models can be constructed to simulate a local balance between radiative and convective effects. One could imagine an atmosphere in which such a balance applied on average at each latitude independently. This would give much higher temperatures at the equator and much lower ones at the poles, with the equator-to-pole temperature difference being perhaps four times the 30 K that appears in the zonal average distribution shown in Fig. 7.9. Such a temperature distribution could be in thermal-wind balance with the zonal flow, giving a very much stronger jet stream than that observed.

Such a state is not observed, however, because of the action of baroclinic disturbances as described in Section 13.9. These derive their energy from the potential energy available in the meridional temperature gradient, and act to reduce this gradient as illustrated in Fig. 13.12. Radiation, on the other hand, tends to restore the gradient, so the mean temperature field observed is largely due to a balance between the competing effects of the instability and of radiation.

To gain further insight into the zonal mean circulation, consider the equations for $[\bar{u}]$, $[\bar{v}]$, and $[\bar{w}_*]$, the temporally and zonally averaged velocity components

associated with the coordinates λ, φ, z_*, where λ is longitude, φ is latitude, and z_* is the log-pressure coordinate introduced in Section 6.17, and for the time and zonally averaged temperature $[\bar{T}]$. The square brackets denote a zonal mean (and the departure from this mean will be denoted by a superscript asterisk) and the overbar denotes a mean with respect to time (the depature from the time mean will be denoted by a prime). The two components $[\bar{v}]$ and $[\bar{w}_*]$ can be expressed [cf. (9.15.10)] in terms of a stream function ψ associated with the meridional circulation. From the continuity equation (6.17.11) and the definition (6.17.29) of ρ_*, it follows that the stream function can be defined by

$$\rho_*[\bar{v}] \cos\varphi = -\partial\psi/\partial z_*, \qquad \rho_*[\bar{w}_*] \cos\varphi = r^{-1}\,\partial\psi/\partial\varphi, \tag{13.10.1}$$

where r is the radius of the earth (the spherical polar form of the equations is discussed in Section 4.12). Similarly, the two components $[\bar{u}]$ and $[\bar{T}]$ can be expressed in terms of a single variable $[\bar{\Phi}]$, the time and zonally averaged geopotential, because the "rapid" adjustment processes studied in earlier chapters will keep them in approximate thermal-wind balance. Taking from (4.12.15), the meridional momentum equation in spherical polar coordinates, the terms that contribute to the geostrophic balance (7.6.7) (but also including a nonlinear term that can be significant in low latitudes), and appending the hydrostatic equation (6.17.20), there results

$$\begin{aligned}(2\Omega + r^{-1}[\bar{u}]\sec\varphi)[\bar{u}]\sin\varphi &= -r^{-1}\,\partial[\bar{\Phi}]/\partial\varphi,\\ H_s^{-1}R[\bar{T}] &= \partial[\bar{\Phi}]/\partial z_*.\end{aligned} \tag{13.10.2}$$

For simplicity, the distinction between virtual temperature T_v and temperature T will be dropped, i.e., moisture effects on buoyancy will be ignored.

Thus the zonal mean circulation can be described in terms of two variables ψ and $[\bar{\Phi}]$. The observed field of ψ for the two extreme seasons is shown in Fig. 1.7, whereas the fields of $[\bar{u}]$ and $[\bar{T}]$ that are associated with $[\bar{\Phi}]$ are shown in Fig. 7.9. The equations that determine the distributions of $[\bar{\Phi}]$ and ψ are the remaining two equations, namely, the averaged forms of the zonal momentum equation (4.12.14) and of the temperature equation [see (4.4.6) and (6.17.13)]. These are coupled equations and can be solved only when taken together, but for discussion purposes it is useful to consider the two equations separately.

13.10.2 The Zonal Momentum and Angular Momentum Balances

A useful form of the zonal momentum equations (4.12.14) is in terms of the angular momentum

$$M = (\Omega r\cos\varphi + [\bar{u}])r\cos\varphi. \tag{13.10.3}$$

The averaged form of this equation can then be written as

$$r^{-1}[\bar{v}]\,\partial M/\partial\varphi + [\bar{w}_*]\,\partial M/\partial z_* = r\cos\varphi(\mathscr{V}_{\text{eddy}} + \mathscr{V}_{\text{friction}}), \tag{13.10.4}$$

where

$$\mathscr{V}_{\text{eddy}} = -r^{-1}\sec^2\varphi\,\partial\{(uv)_{\text{eddy}}\cos^2\varphi\}/\partial\varphi - \rho_*^{-1}\,\partial\{\rho_*(uw)_{\text{eddy}}\}/\partial z_* \tag{13.10.5}$$