Chapter 7

Baroclinic Development

Chapter 6 showed that the quasi-geostrophic equations can qualitatively account for the observed relationships among the vorticity, temperature, and vertical velocity fields in midlatitude synoptic-scale systems. The diagnostic approach used in that chapter provided useful insights into the structure of synoptic-scale systems; it also demonstrated the key role of potential vorticity in dynamical analysis. It did not, however, provide quantitative information on the origins, growth rates, and propagation speeds of such disturbances. This chapter shows how the linear perturbation analysis of Chapter 5 can be used to obtain such information from the quasi-geostrophic equations.

The development of synoptic-scale weather disturbances is often referred to as *cyclogenesis*, a term that emphasizes the role of relative vorticity in developing synoptic-scale systems. This chapter analyzes the processes that lead to cyclogenesis. Specifically, we discuss the role of dynamical instability of the mean flow in accounting for the growth of synoptic-scale disturbances. Moreover, even in the absence of such instability, we show that baroclinic disturbances may still amplify over finite periods of time if properly configured. Thus the quasi-geostrophic equations can, indeed, provide a reasonable theoretical basis for understanding the development of synoptic-scale storms, although as discussed in Section 9.2, ageostrophic effects must be included to model the development of fronts and subsynoptic-scale storms.

7.1 HYDRODYNAMIC INSTABILITY

A zonal-mean flow field is said to be hydrodynamically unstable if a small disturbance introduced into the flow grows spontaneously, drawing energy from the mean flow. It is useful to divide fluid instabilities into two types: parcel instability and wave instability. The simplest example of a parcel instability is the convective overturning that occurs when a fluid parcel is displaced vertically in a statically unstable fluid (see Section 2.7.3). Another example is inertial instability, which occurs when a parcel is displaced radially in an axisymmetric vortex with negative absolute vorticity in the Northern Hemisphere or positive absolute vorticity in the Southern Hemisphere. This instability in discussed in Section 5.5.1. A more general type of parcel instability, called *symmetric*

An Introduction to Dynamic Meteorology. DOI: 10.1016/B978-0-12-384866-6.00007-6

instability, may also be significant in weather disturbances; this is discussed in Section 9.3.

Most of the instabilities of importance in meteorology, however, are associated with wave propagation The wave instabilities important for synoptic-scale meteorology generally occur in the form of zonally asymmetric perturbations to a zonally symmetric basic flow field. In general the basic flow is a jet stream that has both horizontal and vertical mean-flow shears. *Barotropic instability* is a wave instability associated with the horizontal shear in a jet-like current. Barotropic instabilities grow by extracting kinetic energy from the mean-flow field. *Baroclinic instability*, however, is associated with vertical shear of the mean flow. Baroclinic instabilities grow by converting potential energy associated with the mean horizontal temperature gradient that must exist to provide thermal wind balance for the vertical shear in the basic state flow. In neither of these instability types does the parcel method provide a satisfactory stability criterion. A more rigorous approach is required in which a linearized version of the governing equations is analyzed to determine the structure and amplification rate for the various wave modes supported by the system.

As indicated in Problem 5.2 of Chapter 5, the traditional approach to instability analysis is to assume that a small perturbation consisting of a single Fourier wave mode of the form $\exp[ik(x - ct)]$ is introduced into the flow, and to determine the conditions for which the phase velocity c has an imaginary part. This technique, which is called the *normal modes* method, is applied in the next section to analyze the stability of a baroclinic current.

An alternative method of instability analysis is the *initial value* approach. This method is motivated by the recognition that in general the perturbations from which storms develop cannot be described as single normal mode disturbances, but may have a complex structure. The initial growth of such disturbances may strongly depend on the potential vorticity distribution in the initial disturbance. On the time scale of a day or two, such growth can be quite different from that of a normal mode of similar scale, although in the absence of nonlinear interactions the fastest-growing normal mode disturbance must eventually dominate.

A strong dependence of cyclogenesis on initial conditions occurs when a large-amplitude upper-level potential vorticity anomaly is advected into a region where there is a preexisting meridional temperature gradient at the surface. In that case, as shown schematically in Figure 7.1, the circulation induced by the upper-level anomaly (which extends downward, as discussed in Section 6.6) leads to temperature advection at the surface; this induces a potential vorticity anomaly near the surface, which in turn reinforces the upper-level anomaly. Under some conditions the surface and upper-level potential vorticity anomalies can become locked in phase so that the induced circulations produce a very rapid amplification of the anomaly pattern. Detailed discussion of the initial value approach to cyclogenesis is beyond the scope of this text. Here we concentrate primarily on the simplest normal mode instability models.

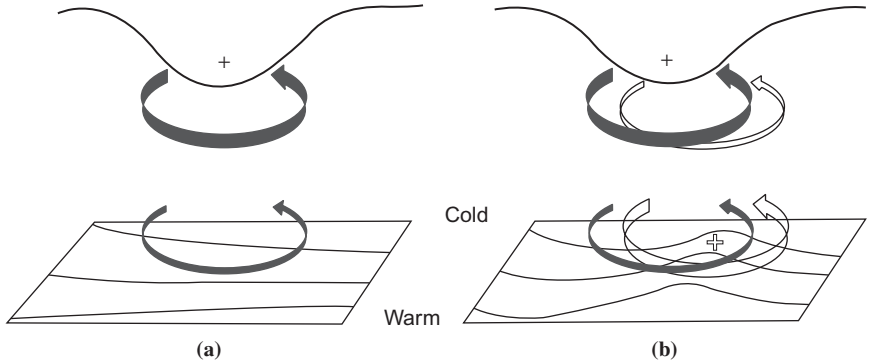

FIGURE 7.1 A schematic picture of cyclogenesis associated with the arrival of an upper-level potential vorticity perturbation over a lower-level baroclinic region. (a) Lower-level cyclonic vorticity induced by the upper-level potential vorticity anomaly. The circulation induced by the potential vorticity anomaly is shown by the *solid arrows*, and potential temperature contours are shown at the lower boundary. The advection of potential temperature by the induced lower-level circulation leads to a warm anomaly slightly east of the upper-level vorticity anomaly. This in turn will induce a cyclonic circulation as shown by the *open arrows* in (b). The induced upper-level circulation will reinforce the original upper-level anomaly and can lead to amplification of the disturbance. (*After Hoskins et al., 1985.*)

7.2 NORMAL MODE BAROCLINIC INSTABILITY: A TWO-LAYER MODEL

Even for a highly idealized mean-flow profile, the mathematical treatment of baroclinic instability in a continuously stratified atmosphere is rather complicated. Before considering such a model, we first focus on the simplest model that can incorporate baroclinic processes. The atmosphere is represented by two discrete layers bounded by surfaces numbered 0, 2, and 4 (generally taken to be the 0-, 500-, and 1000-hPa surfaces, respectively) as shown in Figure 7.2. The quasi-geostrophic vorticity equation for the midlatitude β plane is applied at the 250- and 750-hPa levels, designated by 1 and 3 in Figure 7.2, whereas the thermodynamic energy equation is applied at the 500-hPa level, designated by 2 in Figure 7.2.

Before writing out the specific equations of the two-layer model, it is convenient to define a *geostrophic streamfunction*, $\psi \equiv \Phi/f_0$, for the isobaric form of the quasi-geostrophic equations. Then the geostrophic wind and the geostrophic vorticity can be expressed respectively as

$$\mathbf{V}_\psi = \mathbf{k} \times \nabla_h \psi, \quad \zeta_\mathbf{g} = \nabla_h^2 \psi \tag{7.1}$$

The quasi-geostrophic vorticity equation and the hydrostatic thermodynamic energy equation can then be written in terms of ψ and ω as

$$\frac{\partial}{\partial t}\nabla_h^2 \psi + \mathbf{V}_\psi \cdot \nabla_h \left(\nabla_h^2 \psi\right) + \beta \frac{\partial \psi}{\partial x} = f_0 \frac{\partial \omega}{\partial p} \tag{7.2}$$

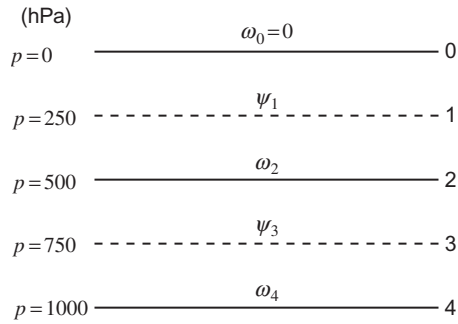

FIGURE 7.2 Arrangement of variables in the vertical for the two-level baroclinic model.

$$\frac{\partial}{\partial t}\left(\frac{\partial \psi}{\partial p}\right) = -\mathbf{V}_\psi \cdot \nabla_h\left(\frac{\partial \psi}{\partial p}\right) - \frac{\sigma}{f_0}\omega \tag{7.3}$$

We now apply the vorticity equation (7.2) at the two levels designated as 1 and 3, which are at the middle of the two layers. To do this we must estimate the divergence term $\partial\omega/\partial p$ at these levels using finite difference approximations to the vertical derivatives:

$$\left(\frac{\partial \omega}{\partial p}\right)_1 \approx \frac{\omega_2 - \omega_0}{\delta p}, \quad \left(\frac{\partial \omega}{\partial p}\right)_3 \approx \frac{\omega_4 - \omega_2}{\delta p} \tag{7.4}$$

where $\delta p = 500$ hPa is the pressure interval between levels 0 to 2 and 2 to 4, and subscript notation is used to designate the vertical level for each dependent variable. The resulting vorticity equations are

$$\frac{\partial}{\partial t}\nabla_h^2\psi_1 + \mathbf{V}_1 \cdot \nabla\left(\nabla_h^2\psi_1\right) + \beta\frac{\partial \psi_1}{\partial x} = \frac{f_0}{\delta p}\omega_2 \tag{7.5}$$

$$\frac{\partial}{\partial t}\nabla_h^2\psi_3 + \mathbf{V}_3 \cdot \nabla_h\left(\nabla_h^2\psi_3\right) + \beta\frac{\partial \psi_3}{\partial x} = -\frac{f_0}{\delta p}\omega_2 \tag{7.6}$$

where we have used the fact that $\omega_0 = 0$, and assumed that $\omega_4 = 0$, which is approximately true for a level-lower boundary surface.

We next write the thermodynamic energy equation (7.3) at level 2. Here we must evaluate $\partial\psi/\partial p$ using the difference formula

$$(\partial\psi/\partial p)_2 \approx (\psi_3 - \psi_1)/\delta p$$

The result is

$$\frac{\partial}{\partial t}(\psi_1 - \psi_3) = -\mathbf{V}_2 \cdot \nabla_h(\psi_1 - \psi_3) + \frac{\sigma\delta p}{f_0}\omega_2 \tag{7.7}$$

The first term on the right-hand side in (7.7) is the advection of the 250- to 750-hPa thickness by the wind at 500 hPa. However, ψ_2, the 500-hPa streamfunction, is not a predicted field in this model. Therefore, ψ_2 must be obtained by linearly interpolating between the 250- and 750-hPa levels:

$$\psi_2 = (\psi_1 + \psi_3)/2$$

If this interpolation formula is used, (7.5), (7.6), and (7.7) become a closed set of prediction equations in the variables ψ_1, ψ_3, and ω_2.

7.2.1 Linear Perturbation Analysis

To keep the analysis as simple as possible, we assume that the streamfunctions ψ_1 and ψ_3 consist of basic state parts that depend linearly on y alone, plus perturbations that depend only on x and t. Thus, we let

$$\begin{aligned} \psi_1 &= -U_1 y + \psi_1'(x,t) \\ \psi_3 &= -U_3 y + \psi_3'(x,t) \\ \omega_2 &= \omega_2'(x,t) \end{aligned} \tag{7.8}$$

The zonal velocities at levels 1 and 3 are then constants with the values U_1 and U_3, respectively. Hence, the perturbation field has meridional and vertical velocity components only.

Substituting from (7.8) into (7.5), (7.6), and (7.7) and linearizing yields the perturbation equations

$$\left(\frac{\partial}{\partial t} + U_1\frac{\partial}{\partial x}\right)\frac{\partial^2\psi_1'}{\partial x^2} + \beta\frac{\partial\psi_1'}{\partial x} = \frac{f_0}{\delta p}\omega_2' \tag{7.9}$$

$$\left(\frac{\partial}{\partial t} + U_3\frac{\partial}{\partial x}\right)\frac{\partial^2\psi_3'}{\partial x^2} + \beta\frac{\partial\psi_3'}{\partial x} = -\frac{f_0}{\delta p}\omega_2' \tag{7.10}$$

$$\left(\frac{\partial}{\partial t} + U_m\frac{\partial}{\partial x}\right)(\psi_1' - \psi_3') - U_T\frac{\partial}{\partial x}(\psi_1' + \psi_3') = \frac{\sigma\delta p}{f_0}\omega_2' \tag{7.11}$$

where we have linearly interpolated to express $\mathbf{V}_2$ in terms of ψ_1 and ψ_3, and have defined

$$U_m \equiv (U_1 + U_3)/2, \quad U_T \equiv (U_1 - U_3)/2$$

Thus, U_m and U_T are, respectively, the vertically averaged mean zonal wind and the mean thermal wind.

The dynamical properties of this system are more clearly expressed if (7.9), (7.10), and (7.11) are combined to eliminate ω_2'. We first note that (7.9) and

(7.10) can be rewritten as

$$\left[\frac{\partial}{\partial t}+\left(U_m+U_T\right)\frac{\partial}{\partial x}\right]\frac{\partial^2\psi_1'}{\partial x^2}+\beta\frac{\partial\psi_1'}{\partial x}=\frac{f_0}{\delta p}\omega_2' \tag{7.12}$$

$$\left[\frac{\partial}{\partial t}+\left(U_m-U_T\right)\frac{\partial}{\partial x}\right]\frac{\partial^2\psi_3'}{\partial x^2}+\beta\frac{\partial\psi_3'}{\partial x}=-\frac{f_0}{\delta p}\omega_2' \tag{7.13}$$

We now define the barotropic and baroclinic perturbations as

$$\psi_m\equiv\left(\psi_1'+\psi_3'\right)/2;\quad \psi_T\equiv\left(\psi_1'-\psi_3'\right)/2 \tag{7.14}$$

Adding (7.12) and (7.13) and using the definitions in (7.14) yield

$$\left[\frac{\partial}{\partial t}+U_m\frac{\partial}{\partial x}\right]\frac{\partial^2\psi_m}{\partial x^2}+\beta\frac{\partial\psi_m}{\partial x}+U_T\frac{\partial}{\partial x}\left(\frac{\partial^2\psi_T}{\partial x^2}\right)=0 \tag{7.15}$$

while subtracting (7.13) from (7.12) and combining with (7.11) to eliminate ω_2' yield

$$\begin{aligned}\left[\frac{\partial}{\partial t}+U_m\frac{\partial}{\partial x}\right]\left(\frac{\partial^2\psi_T}{\partial x^2}-2\lambda^2\psi_T\right)+\beta\frac{\partial\psi_T}{\partial x}\\ +U_T\frac{\partial}{\partial x}\left(\frac{\partial^2\psi_m}{\partial x^2}+2\lambda^2\psi_m\right)=0\end{aligned} \tag{7.16}$$

where $\lambda^2\equiv f_0^2/[\sigma(\delta p)^2]$. Equations (7.15) and (7.16) govern the evolution of the barotropic (vertically averaged) and baroclinic (thermal) perturbation vorticities, respectively.

As in Chapter 5, we assume that wave-like solutions exist of the form

$$\psi_m=Ae^{ik(x-ct)},\quad \psi_{\mathrm{T}}=Be^{ik(x-ct)} \tag{7.17}$$

Substituting these assumed solutions into (7.15) and (7.16) and dividing through by the common exponential factor, we obtain a pair of simultaneous linear algebraic equations for the coefficients A, B:

$$ik\left[\left(c-U_m\right)k^2+\beta\right]A-ik^3U_TB=0 \tag{7.18}$$

$$ik\left[\left(c-U_m\right)\left(k^2+2\lambda^2\right)+\beta\right]B-ikU_T\left(k^2-2\lambda^2\right)A=0 \tag{7.19}$$

Because this set is homogeneous, nontrivial solutions will exist only if the determinant of the coefficients of A and B is zero. Thus, the phase speed c must satisfy the condition

$$\begin{vmatrix}\left(c-U_m\right)k^2+\beta & -k^2U_T\\ -U_T\left(k^2-2\lambda^2\right) & \left(c-U_m\right)\left(k^2+2\lambda^2\right)+\beta\end{vmatrix}=0$$

which gives a quadratic dispersion equation in c:

$$\begin{aligned}\left(c-U_m\right)^2 k^2\left(k^2+2\lambda^2\right)+2\left(c-U_m\right)\beta\left(k^2+\lambda^2\right)\\ +\left[\beta^2+U_T^2k^2\left(2\lambda^2-k^2\right)\right]=0\end{aligned} \tag{7.20}$$

which is analogous to the linear wave dispersion equations developed in Chapter 5. The dispersion relationship in (7.20) yields for the phase speed

$$c=U_m-\frac{\beta\left(k^2+\lambda^2\right)}{k^2\left(k^2+2\lambda^2\right)}\pm\delta^{1/2} \tag{7.21}$$

where

$$\delta\equiv\frac{\beta^2\lambda^4}{k^4\left(k^2+2\lambda^2\right)^2}-\frac{U_T^2\left(2\lambda^2-k^2\right)}{\left(k^2+2\lambda^2\right)}$$

We have now shown that (7.17) is a solution for the system (7.15) and (7.16) only if the phase speed satisfies (7.21). Although (7.21) appears to be rather complicated, it is immediately apparent that if $\delta<0$, the phase speed will have an imaginary part and the perturbations will amplify exponentially. Before discussing the general physical conditions required for exponential growth, it is useful to consider two special cases.

As the first special case we let $U_T=0$ so that the basic state thermal wind vanishes and the mean flow is barotropic. The phase speeds in this case are

$$c_1=U_m-\beta k^{-2} \tag{7.22}$$

and

$$c_2=U_m-\beta\left(k^2+2\lambda^2\right)^{-1} \tag{7.23}$$

These are real quantities that correspond to the free (normal mode) oscillations for the two-level model with a barotropic basic state current. The phase speed c_1 is simply the dispersion relationship for a barotropic Rossby wave with no y dependence (see Section 5.7). Substituting from (7.22) for c in (7.18) and (7.19), we see that in this case $B=0$ so that the perturbation is barotropic in structure. The expression of (7.23), however, may be interpreted as the phase speed for an internal baroclinic Rossby wave. Note that c_2 is a dispersion relationship analogous to the Rossby wave speed for a homogeneous ocean with a free surface, which was given in Problem 7.16. However, in the two-level model, the factor $2\lambda^2$ appears in the denominator in place of the f_0^2/gH for the oceanic case.

In each of these cases there is vertical motion associated with the Rossby wave so that static stability modifies the wave speed. It is left as a problem for the reader to show that if c_2 is substituted into (7.18) and (7.19), the resulting fields of ψ_1 and ψ_3 are 180° out of phase so that the perturbation is baroclinic,

although the basic state is barotropic. Furthermore, the ω_2' field is 1/4 cycle out of phase with the 250-hPa geopotential field, with the maximum upward motion occurring west of the 250-hPa trough.

This vertical motion pattern may be understood if we note that $c_2 - U_m < 0$, so that the disturbance pattern moves westward *relative to the mean wind*. Now, viewed in a coordinate system moving with the mean wind, the vorticity changes are due only to the planetary vorticity advection and the convergence terms, while the thickness changes must be caused solely by the adiabatic heating or cooling due to vertical motion. Thus, there must be rising motion west of the 250-hPa trough to produce the thickness changes required by the westward motion of the system.

Comparing (7.22) and (7.23), we see that the phase speed of the baroclinic mode is generally much less than that of the barotropic mode since for average midlatitude tropospheric conditions $\lambda^2 \approx 2 \times 10^{-12}$ m^{-2}, which is approximately equal to k^2 for a zonal wavelength of 4300 km.[1]

As the second special case, we assume that $\beta = 0$. This case corresponds, for example, to a laboratory situation in which the fluid is bounded above and below by rotating horizontal planes so that the gravity and rotation vectors are everywhere parallel. In such a situation

$$c = U_m \pm U_T \left(\frac{k^2 - 2\lambda^2}{k^2 + 2\lambda^2} \right)^{1/2} \tag{7.24}$$

For waves with zonal wave numbers satisfying $k^2 < 2\lambda^2$, (7.24) has an imaginary part. Thus, all waves longer than the critical wavelength $L_c = \sqrt{2}\pi/\lambda$ will amplify. From the definition of λ we can write

$$L_c = \delta p \pi (2\sigma)^{1/2} / f_0$$

For typical tropospheric conditions $(2\sigma)^{1/2} \approx 2 \times 10^{-3}$ N^{-1} m^3 s^{-1}. Thus, with $\delta p = 500$ hPa and $f_0 = 10^{-4}$ s^{-1}, we find that $L_c \approx 3000$ km. It is also clear from this formula that the critical wavelength for baroclinic instability increases with static stability. The role of static stability in stabilizing the shorter waves can be understood qualitatively as follows: For a sinusoidal perturbation, the relative vorticity, and thus the differential vorticity advection, increases with the square of the wave number. However, as shown in Chapter 6, a secondary vertical circulation is required to maintain hydrostatic temperature changes and geostrophic vorticity changes in the presence of differential vorticity advection. Thus, for a geopotential perturbation of fixed amplitude, the

[1]The presence of the free internal Rossby wave should actually be regarded as a weakness of the two-level model. Lindzen et al. (1968) have shown that this mode does not correspond to any free oscillation of the real atmosphere. Rather, it is a spurious mode resulting from the use of the upper boundary condition $\omega = 0$ at $p = 0$, which formally turns out to be equivalent to putting a lid at the top of the atmosphere.

relative strength of the accompanying vertical circulation must increase as the wavelength of the disturbance decreases. Because static stability tends to resist vertical displacements, the shortest wavelengths will thus be stabilized.

It is also of interest that with $\beta = 0$ the critical wavelength for instability does not depend on the magnitude of the basic state thermal wind U_T. The growth rate, however, does depend on U_T. According to (7.17) the time dependence of the disturbance solution has the form exp(–*ikct*). Thus, the exponential growth rate is $\alpha = kc_i$, where c_i designates the imaginary part of the phase speed. In the present case

$$\alpha = kU_T\left(\frac{2\lambda^2 - k^2}{2\lambda^2 + k^2}\right)^{1/2} \tag{7.25}$$

so that the growth rate increases linearly with the mean thermal wind.

Returning to the general case where all terms are retained in (7.21), the stability criterion is understood most easily by computing the *neutral curve*, which connects all values of U_T and k for which $\delta = 0$ so that the flow is *marginally stable*. From (7.21), the condition $\delta = 0$ implies that

$$\frac{\beta^2\lambda^4}{k^4\left(2\lambda^2 + k^2\right)} = U_T^2\left(2\lambda^2 - k^2\right) \tag{7.26}$$

This complicated relationship between U_T and k can best be displayed by solving (7.26) for $k^4/2\lambda^4$, yielding

$$k^4\Big/\left(2\lambda^4\right) = 1 \pm \left[1 - \beta^2\Big/\left(4\lambda^4 U_T^2\right)\right]^{1/2}$$

In Figure 7.3 the nondimensional quantity $k^2/2\lambda^2$, which is proportional to the square of the zonal wave number, is plotted against the nondimensional parameter $2\lambda^2 U_T/\beta$, which is proportional to the thermal wind. As indicated in Figure 7.3, the neutral curve separates the unstable region of the U_T, k plane from the stable region. It is clear that the inclusion of the β effect serves to stabilize the flow, since now unstable roots exist only for $|U_T| > \beta/(2\lambda^2)$. In addition, the minimum value of U_T required for unstable growth depends strongly on k. Thus, the β effect strongly stabilizes the long wave end of the wave spectrum ($k \to 0$). Again, the flow is always stable for waves shorter than the critical wavelength $L_c = \sqrt{2}\pi/\lambda$.

This long-wave stabilization associated with the β effect is caused by the rapid westward propagation of long waves (i.e., Rossby wave propagation), which occurs only when the β effect is included in the model. It can be shown that baroclinically unstable waves always propagate at a speed that lies between maximum and minimum mean zonal wind speeds. Thus, for the two-level model in the usual midlatitude case where $U_1 > U_3 > 0$, the real part of the phase speed satisfies the inequality $U_3 < c_r < U_1$ for unstable waves. In a continuous atmosphere, this would imply that there must be a level where $U = c_r$. Such a

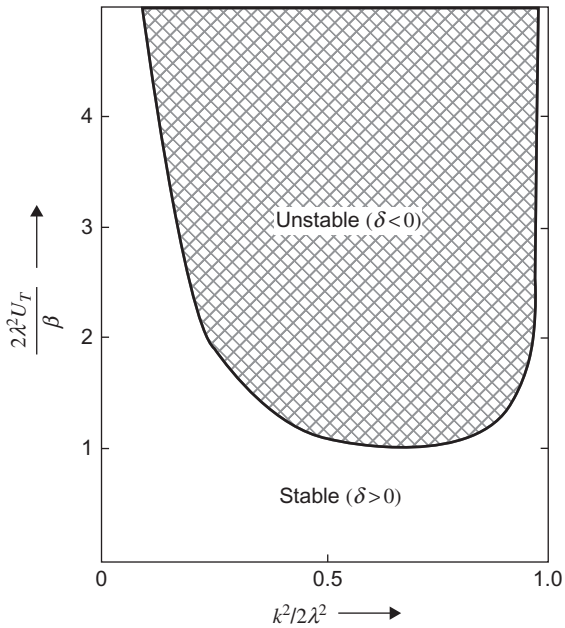

FIGURE 7.3 Neutral stability curve for the two-level baroclinic model.

level is called a *critical level* by theoreticians and a *steering level* by synopticians. For long waves and weak basic state wind shear, the solution given by (7.21) will have $c_r < U_3$, there is no steering level, and unstable growth cannot occur.

Differentiating (7.26) with respect to k and setting $dU_T/dk = 0$, we find that the minimum value of U_T for which unstable waves may exist occurs when $k^2 = \sqrt{2}\lambda^2$. This wave number corresponds to the wave of maximum instability. Wave numbers of observed disturbances should be close to the wave number of maximum instability, because if U_T were gradually raised from zero, the flow would first become unstable for perturbations of wave number $k = 2^{1/4}\lambda$. Those perturbations would then amplify and in the process remove energy from the mean thermal wind, thereby decreasing U_T and stabilizing the flow.

Under normal conditions of static stability the wavelength of maximum instability is about 4000 km, which is close to the average wavelength for midlatitude synoptic systems. Furthermore, the thermal wind required for marginal stability at this wavelength is only about $U_T \approx 4\ \mathrm{m\,s^{-1}}$, which implies a shear of $8\ \mathrm{m\,s^{-1}}$ between 250 and 750 hPa. Shears greater than this are certainly common in middle latitudes for the zonally averaged flow. Therefore, the observed behavior of midlatitude synoptic systems is consistent with the hypothesis that such systems can originate from infinitesimal perturbations of a baroclinically unstable basic current. Of course, in the real atmosphere many other factors may influence the development of synoptic systems—for example, due to lateral shear in the jet stream, nonlinear interactions of finite-amplitude perturbations,

and the release of latent heat in precipitating systems. However, observational studies, laboratory simulations, and numerical models all suggest that baroclinic instability is a primary mechanism for synoptic-scale wave development in middle latitudes.

7.2.2 Vertical Motion in Baroclinic Waves

Since the two-level model is a special case of the quasi-geostrophic system, the physical mechanisms responsible for forcing vertical motion should be those discussed in Section 6.5. Thus, the forcing of vertical motion can be expressed in terms of the sum of the forcing by thermal advection (evaluated at level 2) plus the differential vorticity advection (evaluated as the difference between the vorticity advection at level 1 and that at level 3). Alternatively, the forcing of vertical motion can be expressed in terms of the divergence of the **Q** vector.

The **Q**-vector form of the omega equation for the two-level model can be derived simply from (6.68). We first estimate the second term on the left-hand side by finite differencing in p. Using (7.4) and again letting $\omega_0 = \omega_4 = 0$, we obtain

$$\frac{\partial^2 \omega}{\partial p^2} \approx \frac{(\partial \omega / \partial p)_3 - (\partial \omega / \partial p)_1}{\delta p} \approx -\frac{2\omega_2}{(\delta p)^2}$$

and observe that temperature in the two-level model is represented as

$$\frac{RT}{p} = -\frac{\partial \Phi}{\partial p} \approx \frac{f_0}{\delta p}(\psi_1 - \psi_3)$$

Thus, (6.68) becomes

$$\sigma\left(\nabla_h^2 - 2\lambda^2\right)\omega_2 = -2\nabla_h \cdot \mathbf{Q} \tag{7.27}$$

where

$$\mathbf{Q} = \frac{f_0}{\delta p}\left[-\frac{\partial \mathbf{V_2}}{\partial x} \cdot \nabla_h (\psi_1 - \psi_3)\,, \quad -\frac{\partial \mathbf{V_2}}{\partial y} \cdot \nabla_h (\psi_1 - \psi_3)\right]$$

In order to examine the forcing of vertical motion in baroclinically unstable waves, we linearize (7.27) by specifying the same basic state and perturbation variables as in (7.8). For this situation, in which the mean zonal wind and the perturbation streamfunctions are independent of y, the **Q** vector has only an x component:

$$Q_1 = \frac{f_0}{\delta p}\left[\frac{\partial^2 \psi_2'}{\partial x^2}(U_1 - U_3)\right] = \frac{2f_0}{\delta p} U_T \zeta_2'$$

The pattern of the **Q** vector in this case is similar to that of Figure 6.14, with eastward-pointing **Q** centered at the trough and westward-pointing **Q** centered

at the ridge. This is consistent with the fact that **Q** represents the change of temperature gradient forced by geostrophic motion alone. In this simple model the temperature gradient is entirely due to the vertical shear of the mean zonal wind [$U_T \propto -\partial T/\partial y$], and the shear of the perturbation meridional velocity tends to advect warm air poleward east of the 500-hPa trough and cold air equatorward west of the 500-hPa trough so that there is a tendency to produce a component of temperature gradient directed eastward at the trough.

The forcing of vertical motion by the **Q** vector in the linearized model is, from (7.27),

$$\left(\frac{\partial^2}{\partial x^2} - 2\lambda^2\right)\omega_2' = -\frac{4f_0}{\sigma\delta p}U_T\frac{\partial\zeta_2'}{\partial x} \tag{7.28}$$

Observing that

$$\left(\frac{\partial^2}{\partial x^2} - 2\lambda^2\right)\omega_2' \propto -\omega_2'$$

we may interpret (7.28) physically by noting that

$$w_2' \propto -\omega_2' \propto -U_T\frac{\partial\zeta_2'}{\partial x} \propto -v_2'\frac{\partial\overline{T}}{\partial y}$$

Thus, sinking motion is forced by negative advection of disturbance vorticity by the basic state thermal wind (or, alternatively, cold advection of the basic state thermal field by the perturbation meridional wind), while rising motion is forced by advection of the opposite sign.

We now have the information required to diagram the structure of a baroclinically unstable disturbance in the two-level model. The lower part of Figure 7.4 shows schematically the phase relationship between the geopotential field and the divergent secondary motion field for the usual midlatitude situation where $U_T > 0$. Linear interpolation has been used between levels so that the trough and ridge axes are straight lines tilted back toward the west with height. In this example the ψ_1 field lags the ψ_3 field by about 65° in phase so that the trough at 250 hPa lies 65° in phase west of the 750-hPa trough. At 500 hPa the perturbation thickness field lags the geopotential field by one-quarter wavelength as shown in the top part of Figure 7.4, and the thickness and vertical motion fields are in phase. Note that the temperature advection by the perturbation meridional wind is in phase with the 500-hPa thickness field so that the advection of the basic state temperature by the perturbation wind acts to intensify the perturbation thickness field. This tendency is also illustrated by the zonally oriented **Q** vectors shown at the 500-hPa level in Figure 7.4.

The pattern of vertical motion forced by the divergence of the **Q** vector, as shown in Figure 7.4, is associated with a divergence–convergence pattern that contributes a positive vorticity tendency nearthe 250-hPa trough and a negative

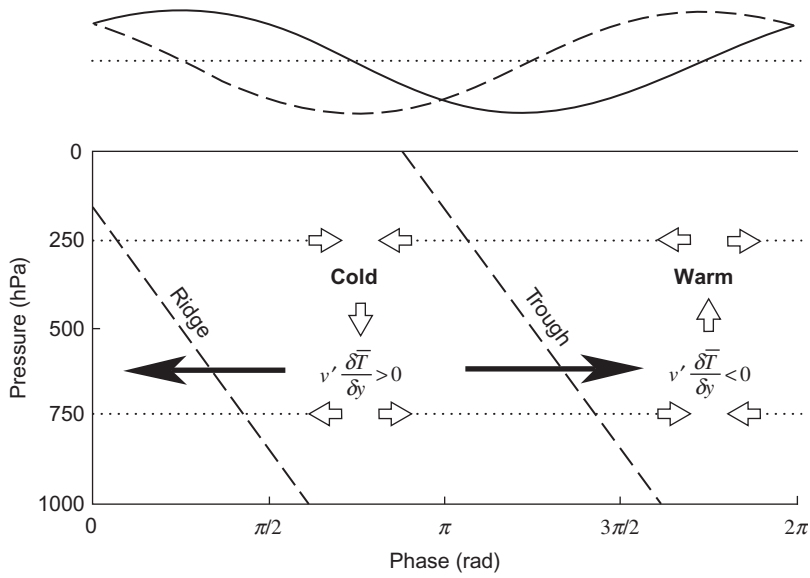

FIGURE 7.4 Structure of an unstable baroclinic wave in the two-level model. (*Top*) Relative phases of the 500-hPa perturbation geopotential (*solid line*) and temperature (*dashed line*). (*Bottom*) Vertical cross-section showing phases of geopotential, meridional temperature advection, ageostrophic circulation (*open arrows*), **Q** vectors (*solid arrows*), and temperature fields for an unstable baroclinic wave in the two-level model.

vorticity tendency near the 750-hPa ridge, with opposite tendencies at the 250-hPa ridge and 750-hPa trough. Since in all cases these vorticity tendencies tend to increase the extreme values of vorticity at the troughs and ridges, this secondary circulation system will act to increase the strength of the disturbance.

The total vorticity change at each level is, of course, determined by the sum of vorticity advection and vortex stretching due to the divergent circulation. The relative contributions of these processes are indicated schematically in Figures 7.5 and 7.6, respectively. As can be seen in Figure 7.5, vorticity advection leads the vorticity field by one-quarter wavelength. Since in this case the basic state wind increases with height, the vorticity advection at 250 hPa is larger than that at 750 hPa. If no other processes influenced the vorticity field, the effect of this differential vorticity advection would be to move the upper-level trough and ridge pattern eastward more rapidly than the lower-level pattern. Thus, the westward tilt of the trough–ridge pattern would quickly be destroyed. The maintenance of this tilt in the presence of differential vorticity advection is due to the concentration of vorticity by vortex stretching associated with the divergent secondary circulation.

Referring to Figure 7.6, we see that concentration of vorticity by the divergence effect lags the vorticity field by about 65° at 250 hPa and leads the vorticity field by about 65° at 750 hPa. As a result, the net vorticity tendencies

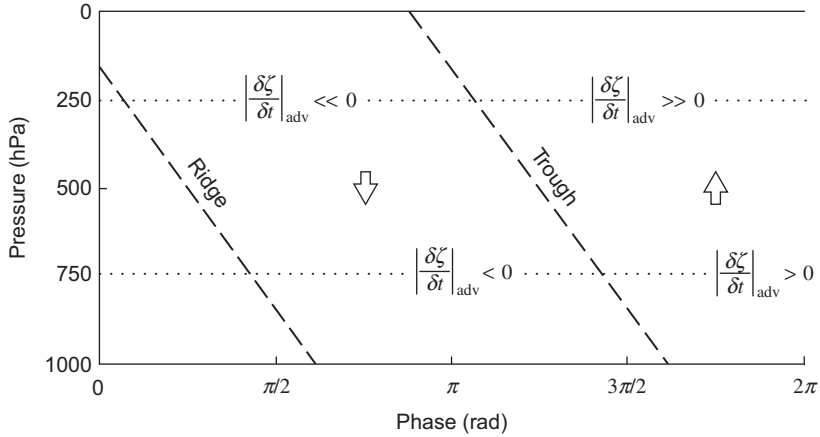

FIGURE 7.5 Vertical cross-section showing the phase of vorticity change due to vorticity advection for an unstable baroclinic wave in the two-level model.

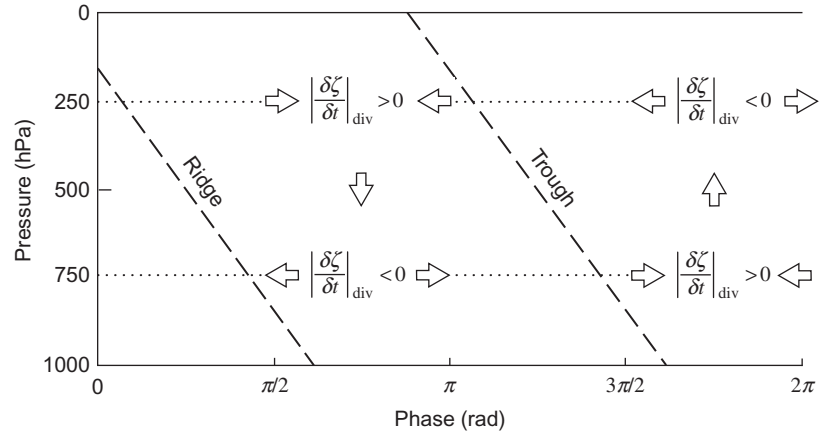

FIGURE 7.6 Vertical cross-section showing the phase of vorticity change due to divergence–convergence for an unstable baroclinic wave in the two-level model.

ahead of the vorticity maxima and minima are less than the advective tendencies at the upper level and greater than the advective tendencies at the lower level:

$$\left|\frac{\partial\zeta_1'}{\partial t}\right| < \left|\frac{\partial\zeta_1'}{\partial t}\right|_{adv}, \quad \left|\frac{\partial\zeta_3'}{\partial t}\right| > \left|\frac{\partial\zeta_3'}{\partial t}\right|_{adv}$$

Furthermore, vorticity concentration by the divergence effect will tend to amplify the vorticity perturbations in the troughs and ridges at both the 250- and 750-hPa levels as required for a growing disturbance.

7.3 THE ENERGETICS OF BAROCLINIC WAVES

The previous section showed that under suitable conditions a vertically sheared geostrophically balanced basic state flow is unstable to small wave-like perturbations with horizontal wavelengths in the range of observed synoptic-scale systems. Such baroclinically unstable perturbations will amplify exponentially by drawing energy from the mean flow. This section considers the energetics of linearized baroclinic disturbances and shows that the potential energy of the mean flow is the energy source for baroclinically unstable perturbations.

7.3.1 Available Potential Energy

Before discussing the energetics of baroclinic waves, it is necessary to consider the energy of the atmosphere from a more general point of view. For all practical purposes, the total energy of the atmosphere is the sum of internal energy, gravitational potential energy, and kinetic energy. However, it is not necessary to consider separately the variations of internal and gravitational potential energy because in a hydrostatic atmosphere these two forms of energy are proportional and may be combined into a single term called the *total potential energy*. The proportionality of internal and gravitational potential energy can be demonstrated by considering these forms of energy for a column of air of unit horizontal cross-section extending from the surface to the top of the atmosphere.

If we let dE_I be the internal energy in a vertical section of the column of height dz, then from the definition of internal energy—see Section (2.6):

$$dE_I = \rho c_v T dz$$

so that the internal energy for the entire column is

$$E_I = c_v \int_0^\infty \rho T dz \tag{7.29}$$

However, the gravitational potential energy for a slab of thickness dz at a height z is just

$$dE_P = \rho g z dz$$

so that the gravitational potential energy in the entire column is

$$E_P = \int_0^\infty \rho g z dz = -\int_{p_0}^0 z dp \tag{7.30}$$

where we have substituted from the hydrostatic equation to obtain the last integral in (7.30). Integrating (7.30) by parts and using the ideal gas law,

we obtain

$$E_P = \int_0^\infty p dz = R \int_0^\infty \rho T dz \tag{7.31}$$

Comparing (7.29) and (7.31), we see that $c_v E_P = R E_I$. Thus, the total potential energy may be expressed as

$$E_P + E_I = \left(c_p/c_v\right) E_I = \left(c_p/R\right) E_P \tag{7.32}$$

Therefore, in a hydrostatic atmosphere the total potential energy can be obtained by computing either E_I or E_P alone.

The total potential energy is not a very useful measure of energy in the atmosphere because only a very small fraction of the total potential energy is available for conversion to kinetic energy in storms. To demonstrate qualitatively why most of the total potential energy is unavailable, we consider a simple model atmosphere that initially consists of two equal masses of dry air separated by a vertical partition as shown in Figure 7.7. The two air masses are at uniform potential temperatures θ_1 and θ_2, respectively, with $\theta_1 < \theta_2$. The ground-level pressure on each side of the partition is taken to be 1000 hPa. We wish to compute the maximum kinetic energy that can be realized by an adiabatic rearrangement of mass within the same volume when the partition is removed.

Now for an adiabatic process, total energy is conserved:

$$E_K + E_P + E_I = \text{constant}$$

where E_K denotes the kinetic energy. If the air masses are initially at rest, $E_K = 0$. Thus, if we let primed quantities denote the final state

$$E'_K + E'_P + E'_I = E_P + E_I$$

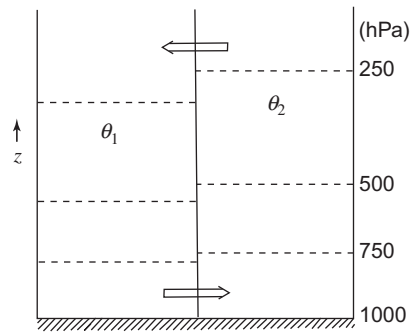

FIGURE 7.7 Two air masses of differing potential temperature separated by a vertical partition. *Dashed lines* indicate isobaric surfaces. *Arrows* show direction of motion when the partition is removed.

so that with the aid of (7.32) the kinetic energy realized by removal of the partition may be expressed as

$$E'_K = \left(c_p/c_v\right)\left(E_I - E'_I\right)$$

Because θ is conserved for an adiabatic process, the two air masses cannot mix. It is clear that E'_I will be a minimum (designated by E''_I) when the masses are rearranged so that the air at potential temperature θ_1 lies entirely beneath the air at potential temperature θ_2, with the 500-hPa surface as the horizontal boundary between the two masses. In that case the total potential energy $(c_p/c_v)E''_I$ is not available for conversion to kinetic energy because no adiabatic process can further reduce E''_I.

The *available potential energy* (APE) can now be defined as the difference between the total potential energy of a closed system and the minimum total potential energy that could result from an adiabatic redistribution of mass. Thus, for the idealized model given earlier, the APE, which is designated by the symbol P, is

$$P = \left(c_p/c_v\right)\left(E_I - E''_I\right) \tag{7.33}$$

which is equivalent to the maximum kinetic energy that can be realized by an adiabatic process.

Lorenz (1960) showed that available potential energy is given approximately by the volume integral over the entire atmosphere of the variance of potential temperature on isobaric surfaces. Thus, letting $\overline{\theta}$ designate the average potential temperature for a given pressure surface and θ' the local deviation from the average, the average available potential energy per unit volume satisfies the proportionality

$$\overline{P} \propto V^{-1} \int \left(\overline{\theta'^2}/\bar{\theta}^2\right) dV$$

where V designates the total volume. For the quasi-geostrophic model, this proportionality is an exact measure of the available potential energy, as shown in the following subsection.

Observations indicate that for the atmosphere as a whole,

$$\overline{P}/\left[\left(c_p/c_v\right)\overline{E}_I\right] \sim 5 \times 10^{-3}, \quad \overline{K}/\overline{P} \sim 10^{-1}$$

Thus, only about 0.5% of the total potential energy of the atmosphere is available, and of the available portion only about 10% is actually converted to kinetic energy. From this point of view the atmosphere is a rather inefficient heat engine.

7.3.2 Energy Equations for the Two-Layer Model

In the two-layer model of Section 7.2, the perturbation temperature field is proportional to $\psi'_1 - \psi'_3$, the 250- to 750-hPa thickness. Thus, in view of the discussion in the previous section, we anticipate that the available potential energy

in this case will be proportional to $(\psi_1' - \psi_3')^2$. To show that this in fact must be the case, we derive the energy equations for the system in the following manner: We first multiply (7.9) by $-\psi_1$, (7.10) by $-\psi_3$, and (7.11) by $(\psi_1' - \psi_3')$. We then integrate the resulting equations over one wavelength of the perturbation in the zonal direction.

The resulting zonally averaged[2] terms are denoted by overbars as done previously in Chapter 5:

$$\overline{(\,)} = L^{-1} \int_0^L (\,)\, dx$$

where L is the wavelength of the perturbation. Thus, for the first term in (7.9) we have, after multiplying by $-\psi_1$, averaging and differentiating by parts:

$$-\overline{\psi_1' \frac{\partial}{\partial t}\left(\frac{\partial^2 \psi_1'}{\partial x^2}\right)} = -\frac{\partial}{\partial x}\overline{\left[\psi_1' \frac{\partial}{\partial x}\left(\frac{\partial \psi_1'}{\partial t}\right)\right]} + \overline{\frac{\partial \psi_1'}{\partial x}\frac{\partial}{\partial t}\left(\frac{\partial \psi_1'}{\partial x}\right)}$$

The first term on the right side vanishes because it is the integral of a perfect differential in x over a complete cycle. The second term on the right can be rewritten in the form

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\frac{\partial \psi_1'}{\partial x}\right)^2}$$

which is just the rate of change of the perturbation kinetic energy per unit mass averaged over a wavelength. Similarly, $-\psi_1'$ times the advection term on the left in (7.9) can be written after integration in x as

$$\begin{aligned} -U_1 \overline{\psi_1' \frac{\partial^2}{\partial x^2}\left(\frac{\partial \psi_1'}{\partial x}\right)} &= -U_1 \frac{\partial}{\partial x}\overline{\left[\psi_1' \frac{\partial}{\partial x}\left(\frac{\partial \psi_1'}{\partial x}\right)\right]} + U_1 \overline{\frac{\partial \psi_1'}{\partial x}\frac{\partial^2 \psi_1'}{\partial x^2}} \\ &= \frac{U_1}{2}\overline{\frac{\partial}{\partial x}\left(\frac{\partial \psi_1'}{\partial x}\right)^2} = 0 \end{aligned}$$

Thus, the advection of kinetic energy vanishes when integrated over a wavelength. Evaluating the various terms in (7.10) and (7.11) in the same manner after multiplying through by $-\psi_3'$ and $(\psi_1' - \psi_3')$, respectively, we obtain the following set of perturbation energy equations:

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\frac{\partial \psi_1'}{\partial x}\right)^2} = -\frac{f_0}{\delta p}\overline{\omega_2' \psi_1'} \tag{7.34}$$

[2] A zonal average generally designates the average around an entire circle of latitude. However, for a disturbance consisting of a single sinusoidal wave of wave number $k = m/(a\cos\phi)$, where m is an integer, the average over a wavelength is identical to a zonal average.

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\frac{\partial \psi_3'}{\partial x}\right)^2} = +\frac{f_0}{\delta p}\overline{\omega_2'\psi_3'} \tag{7.35}$$

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\left(\psi_1' - \psi_3'\right)^2} = \overline{U_T\left(\psi_1' - \psi_3'\right)\frac{\partial}{\partial x}\left(\psi_1' + \psi_3'\right)} + \frac{\sigma\,\delta p}{f_0}\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} \tag{7.36}$$

where as before $U_T \equiv (U_1 - U_3)/2$.

Defining the perturbation kinetic energy to be the sum of the kinetic energies of the 250- and 750-hPa levels,

$$K' \equiv (1/2)\left[\overline{\left(\partial\psi_1'/\partial x\right)^2} + \overline{\left(\partial\psi_3'/\partial x\right)^2}\right]$$

we find by adding (7.34) and (7.35) that

$$dK'/dt = -\left(f_0/\delta p\right)\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} = -\left(2f_0/\delta p\right)\overline{\omega_2'\psi_T} \tag{7.37}$$

Thus, the rate of change of perturbation kinetic energy is proportional to the correlation between perturbation thickness and vertical motion.

If we now define the perturbation available potential energy as

$$P' \equiv \lambda^2\overline{\left(\psi_1' - \psi_3'\right)^2}/2$$

we obtain from (7.36)

$$\begin{aligned} dP'/dt &= \lambda^2 U_T\overline{\left(\psi_1' - \psi_3'\right)\partial\left(\psi_1' + \psi_3'\right)/\partial x} + \left(f_0/\delta p\right)\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} \\ &= 4\lambda^2 U_T\overline{\psi_T\partial\psi_m/\partial x} + \left(2f_0/\delta p\right)\overline{\omega_2'\psi_T} \end{aligned} \tag{7.38}$$

The last term in (7.38) is just equal and opposite to the kinetic energy source term in (7.37). This term clearly must represent a conversion between potential and kinetic energy. If on average the vertical motion is positive ($\omega_2' < 0$) where the thickness is greater than average ($\psi_1' - \psi_3' > 0$) and vertical motion is negative where thickness is less than average, then

$$\overline{\omega_2'\left(\psi_1' - \psi_3'\right)} = 2\overline{\omega_2'\psi_T} < 0$$

and the perturbation potential energy is being converted to kinetic energy. Physically, this correlation represents an overturning in which cold air aloft is replaced by warm air from below, a situation that clearly tends to lower the center of mass and thus the potential energy of the perturbation. However, the available potential energy and kinetic energy of a disturbance can still grow simultaneously, provided that the potential energy generation due to the first term in (7.38) exceeds the rate of potential energy conversion to kinetic energy.

The potential energy generation term in (7.38) depends on the correlation between the perturbation thickness ψ_T and the meridional velocity at 500 hPa,

$\partial\psi_m/\partial x$. In order to understand the role of this term, it is helpful to consider a particular sinusoidal wave disturbance. Suppose that the barotropic and baroclinic parts of the disturbance can be written, respectively, as

$$\psi_m = A_m \cos k\,(x - ct) \quad \text{and} \quad \psi_T = A_T \cos k\,(x + x_0 - ct) \tag{7.39}$$

where x_0 designates the phase difference. Because ψ_m is proportional to the 500-hPa geopotential and ψ_T is proportional to the 500-hPa temperature (or 250- to 750-hPa thickness), the phase angle kx_0 gives the phase difference between geopotential and temperature fields at 500 hPa. Furthermore, A_m and A_T are measures of the amplitudes of the 500-hPa disturbance geopotential and thickness fields, respectively. Using the expressions in (7.39), we obtain

$$\begin{aligned}\overline{\psi_T \frac{\partial \psi_m}{\partial x}} &= -\frac{k}{L}\int_0^L A_T A_m \cos k\,(x + x_0 - ct)\sin k\,(x - ct)\,dx \\ &= \frac{kA_T A_m \sin kx_0}{L}\int_0^L [\sin k\,(x - ct)]^2\,dx \\ &= (A_T A_m k \sin kx_0)/2\end{aligned} \tag{7.40}$$

Substituting from (7.40) into (7.38), we see that for the usual midlatitude case of a westerly thermal wind ($U_T > 0$) the correlation in (7.40) must be positive if the perturbation potential energy is to increase. Thus, kx_0 must satisfy the inequality $0 < kx_0 < \pi$. Furthermore, the correlation will be a positive maximum for $kx_0 = \pi/2$—that is, when the temperature wave lags the geopotential wave by 90° in phase at 500 hPa.

This case was shown schematically earlier in Figure 7.4. Clearly, when the temperature wave lags the geopotential by one-quarter cycle, the northward advection of warm air by the geostrophic wind east of the 500-hPa trough and the southward advection of cold air west of the 500-hPa trough are both maximized. As a result, cold advection is strong below the 250-hPa trough, and warm advection is strong below the 250-hPa ridge. In that case, as discussed previously in Section 6.4.2, the upper-level disturbance will intensify. It should also be noted here that if the temperature wave lags the geopotential wave, the trough and ridge axes will tilt westward with height, which, as mentioned in Section 6.1, is observed to be the case for amplifying midlatitude synoptic systems.

Referring again to Figure 7.4 and recalling the vertical motion pattern implied by the omega equation (7.28), we see that the signs of the two terms on the right in (7.38) cannot be the same. In the westward-tilting perturbation of Figure 7.4, the vertical motion must be downward in the cold air behind the trough at 500 hPa. Hence, the correlation between temperature and vertical velocity must be positive in this situation; that is,

$$\overline{\omega_2' \psi_T} < 0$$

Thus, for quasi-geostrophic perturbations, a westward tilt of the perturbation with height implies both that the horizontal temperature advection will increase the available potential energy of the perturbation and that the vertical circulation will convert perturbation available potential energy to perturbation kinetic energy. Conversely, an eastward tilt of the system with height would change the signs of both terms on the right in (7.38).

Although the signs of the potential energy generation term and the potential energy conversion term in (7.38) are always opposite for a developing baroclinic wave, it is only the potential energy generation rate that determines the growth of the total energy $P' + K'$ of the disturbance. This may be proved by adding (7.37) and (7.38) to obtain

$$d\left(P' + K'\right)/dt = 4\lambda^2 U_T \overline{\psi_T \partial \psi_m / \partial x}$$

Provided the correlation between the meridional velocity and temperature is positive and $U_T > 0$, the total energy of the perturbation will increase. Note that the vertical circulation merely converts disturbance energy between the available potential and kinetic forms without affecting the total energy of the perturbation.

The rate of increase of the total energy of the perturbation depends on the magnitude of the basic state thermal wind U_T. This is, of course, proportional to the zonally averaged meridional temperature gradient. Because the generation of perturbation energy requires systematic poleward transport of warm air and equatorward transport of cold air, it is clear that baroclinically unstable disturbances tend to reduce the meridional temperature gradient and thus the available potential energy of the mean flow. This latter process cannot be described mathematically in terms of the linearized equations. However, from Figure 7.8 we

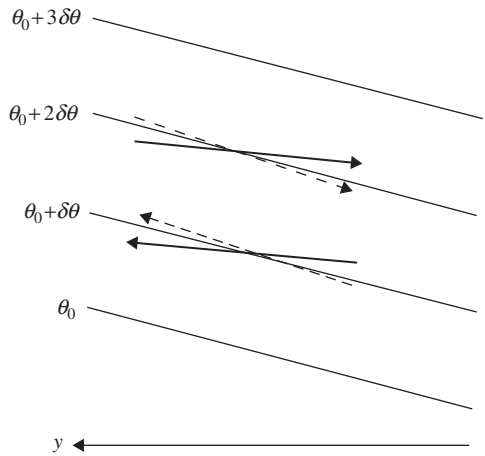

FIGURE 7.8 Slopes of parcel trajectories relative to the zonal-mean potential temperature surfaces for a baroclinically unstable disturbance (*solid arrows*) and for a baroclinically stable disturbance (*dashed arrows*).

can see qualitatively that parcels that move poleward and upward with slopes less than the slope of the zonal-mean potential temperature surface will become warmer than their surroundings, and vice versa for parcels moving downward and equatorward.

For such parcels the correlations between disturbance meridional velocity and temperature and between disturbance vertical velocity and temperature will both be positive as required for baroclinically unstable disturbances. For parcels that have trajectory slopes greater than the mean potential temperature slope, however, both of these correlations will be negative. Such parcels must then convert disturbance kinetic energy to disturbance available potential energy, which is in turn converted to zonal-mean available potential energy. Therefore, in order that perturbations are able to extract potential energy from the mean flow, the perturbation parcel trajectories in the meridional plane must have slopes less than the slopes of the potential temperature surfaces, and a permanent rearrangement of air must take place for there to be a net heat transfer.

Since we have previously seen that poleward-moving air must rise and equatorward-moving air must sink, it is clear that the rate of energy generation can be greater for an atmosphere in which the meridional slope of the potential temperature surfaces is large. We can also see more clearly why baroclinic instability has a short-wave cutoff. As mentioned previously, the intensity of the vertical circulation must increase as the wavelength decreases. Thus, the slopes of the parcel trajectories must increase with decreasing wavelength, and for some critical wavelength the trajectory slopes will become greater than the slopes of the potential temperature surfaces. Unlike convective instability, where the most rapid amplification occurs for the smallest possible scales, baroclinic instability is most effective at an intermediate range of scales.

The energy flow for quasi-geostrophic perturbations is summarized in Figure 7.9 by means of a block diagram. In this type of energy diagram each block represents a reservoir of a particular type of energy, and arrows designate the direction of energy flow. The complete energy cycle cannot be derived in terms of linear perturbation theory but will be discussed qualitatively in Chapter 10.

7.4 BAROCLINIC INSTABILITY OF A CONTINUOUSLY STRATIFIED ATMOSPHERE

In the two previous sections some basic aspects of baroclinic instability were elucidated in the context of a simple two-layer model. The dependence of growth rate on vertical shear and the existence of a short-wave cutoff were clearly demonstrated. The two-layer model, however, does have one severe constraint: It assumes that the altitude dependence of large-scale systems can be adequately represented with only two degrees of freedom in the vertical (i.e., the streamfunctions at the 250- and 750-hPa levels). Although most synoptic-scale systems in midlatitudes are observed to have vertical scales comparable

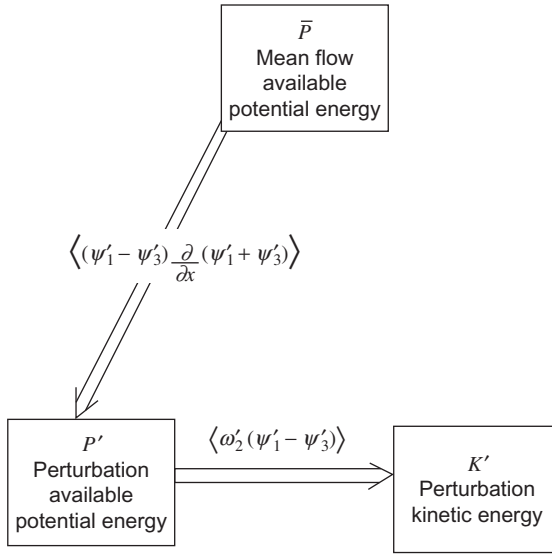

FIGURE 7.9 Energy flow in an amplifying baroclinic wave.

to the depth of the troposphere, observed vertical structures do differ. Disturbances that are concentrated near the ground or near the tropopause can hardly be represented accurately in the two-layer model.

An analysis of the structure of baroclinic modes for realistic mean zonal wind profiles is quite complex, and indeed can only be done by numerical methods. However, without obtaining specific normal mode solutions, it is possible to obtain necessary conditions for baroclinic or barotropic instability from an integral theorem first developed by Rayleigh. This theorem, which is discussed in Section 7.4.2, also shows how baroclinic instability is intimately related to the mean meridional gradient of potential vorticity and the mean meridional temperature gradient at the surface.

If a number of simplifying assumptions are made, it is possible to pose the stability problem for a continuously stratified atmosphere in a fashion that leads to a second-order differential equation for the vertical structure that can be solved by standard methods. This problem was originally studied by the British meteorologist Eady (1949) and, although mathematically similar to the two-layer model, provides additional insights. It is developed in Section 7.4.3.

7.4.1 Log-Pressure Coordinates

Derivation of the Rayleigh theorem and the Eady stability model is facilitated if we transform from the standard isobaric coordinates to a vertical coordinate based on the logarithm of pressure. In the *log-pressure coordinates*, the vertical

coordinate is defined as

$$z^* \equiv -H \ln (p/p_s) \tag{7.41}$$

where p_s is a standard reference pressure (usually taken to be 1000 hPa) and H is a standard scale height, $H \equiv RT_s/g$, with T_s a global average temperature. For the special case of an isothermal atmosphere at temperature T_s, z^* is exactly equal to geometric height, and the density profile is given by the reference density

$$\rho_0 \left(z^*\right) = \rho_s \exp \left(-z^*/H\right)$$

where ρ_s is the density at $z^* = 0$.

For an atmosphere with a realistic temperature profile, z^* is only approximately equivalent to the height, but in the troposphere the difference is usually quite small. The vertical velocity in this coordinate system is

$$w^* \equiv Dz^*/dt$$

The horizontal momentum equation in the log-pressure system is the same as that in the isobaric system:

$$D\mathbf{V}/Dt + f\mathbf{k} \times \mathbf{V} = -\nabla_h \Phi \tag{7.42}$$

However, the operator D/Dt is now defined as

$$D/Dt = \partial/\partial t + \mathbf{V} \cdot \nabla_h + w^* \partial/\partial z^*$$

The hydrostatic equation $\partial\Phi/\partial p = -\alpha$ can be transformed to the log-pressure system by multiplying through by p and using the ideal gas law to get

$$\partial\Phi/\partial \ln p = -RT$$

which after dividing through by –H and using (7.41) gives

$$\partial\Phi/\partial z^* = RT/\mathrm{H} \tag{7.43}$$

The log-pressure form of the continuity equation can be obtained by transforming from the isobaric coordinate form (3.5). We first note that

$$w^* \equiv -(H/p)\, Dp/Dt = -H\omega/p$$

so that

$$\frac{\partial \omega}{\partial p} = -\frac{\partial}{\partial p}\left(\frac{pw^*}{H}\right) = \frac{\partial w^*}{\partial z^*} - \frac{w^*}{H} = \frac{1}{\rho_0}\frac{\partial(\rho_0 w^*)}{\partial z^*}$$

Thus, in log-pressure coordinates the continuity equation becomes simply

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{1}{\rho_0}\frac{\partial(\rho_0 w^*)}{\partial z^*} = 0 \tag{7.44}$$

It is left as a problem for the reader to show that the first law of thermodynamics (3.6) can be expressed in log-pressure form as

$$\left(\frac{\partial}{\partial t} + \mathbf{V}\cdot\nabla_h\right)\frac{\partial\Phi}{\partial z^*} + w^* N^2 = \frac{\kappa J}{H} \tag{7.45}$$

where

$$N^2 \equiv (R/H)\left(\partial T/\partial z^* + \kappa T/H\right)$$

is the buoyancy frequency squared (see Section 2.7.3) and $\kappa \equiv R/c_p$. Unlike the static stability parameter, S_p, in the isobaric form of the thermodynamic equation (3.6), the parameter N^2 varies only weakly with height in the troposphere; it can be assumed to be constant without serious error. This is a major advantage of the log-pressure formulation.

The quasi-geostrophic potential vorticity equation has the same form as in the isobaric system, but with q defined as

$$q \equiv \nabla_h^2\psi + f + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\psi}{\partial z^*}\right) \tag{7.46}$$

where $\varepsilon \equiv f_0^2/N^2$.

7.4.2 Baroclinic Instability: The Rayleigh Theorem

We now examine the stability problem for a continuously stratified atmosphere on the midlatitude β plane. The linearized form of the quasi-geostrophic potential vorticity equation can be expressed in log-pressure coordinates as

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)q' + \frac{\partial\overline{q}}{\partial y}\frac{\partial\psi'}{\partial x} = 0 \tag{7.47}$$

where

$$q' \equiv \nabla_h^2\psi' + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\psi'}{\partial z^*}\right) \tag{7.48}$$

and

$$\frac{\partial\overline{q}}{\partial y} = \beta - \frac{\partial^2\overline{u}}{\partial y^2} - \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\overline{u}}{\partial z^*}\right) \tag{7.49}$$

As in the two-layer model, boundary conditions are required at lower- and upper-boundary pressure surfaces. Assuming that the flow is adiabatic and that the vertical motion w^* vanishes at these boundaries, the linearized form of the thermodynamic energy equation (7.45) valid at horizontal boundary surfaces is simply

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\frac{\partial\psi'}{\partial z^*} - \frac{\partial\psi'}{\partial x}\frac{\partial\overline{u}}{\partial z^*} = 0 \tag{7.50}$$

The sidewall boundary conditions are

$$\partial\psi'/\partial x = 0, \quad \text{hence, } \psi' = 0 \quad \text{at} \quad y = \pm L \tag{7.51}$$

We now assume that the perturbation consists of a single zonal Fourier component propagating in the x direction:

$$\psi'(x, y, z, t) = \text{Re}\left\{\Psi(y, z)\exp\left[ik(x - ct)\right]\right\} \tag{7.52}$$

where $\Psi(y, z) = \Psi_r + i\Psi_i$ is a complex amplitude, k is the zonal wave number, and $c = c_r + ic_i$ is a complex phase speed. Note that (7.52) can alternatively be expressed as

$$\psi'(x, y, z, t) = e^{kc_i t}\left[\Psi_r \cos k(x - c_r t) - \Psi_i \sin k(x - c_r t)\right]$$

Thus, the relative magnitudes of Ψ_r and Ψ_i determine the phase of the wave for any y, z^*.

Substituting from (7.52) into (7.47) and (7.50) yields

$$(\overline{u} - c)\left[\frac{\partial^2\Psi}{\partial y^2} - k^2\Psi + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\Psi}{\partial z^*}\right)\right] + \frac{\partial\overline{q}}{\partial y}\Psi = 0 \tag{7.53}$$

and

$$(\overline{u} - c)\frac{\partial\Psi}{\partial z^*} - \frac{\partial\overline{u}}{\partial z^*}\Psi = 0 \quad \text{at } z^* = 0 \tag{7.54}$$

If the upper boundary is taken to be a rigid lid at a finite height, as is sometimes done in theoretical studies, the condition (7.54) is appropriate at that boundary as well. Alternatively the upper boundary condition can be specified by requiring Ψ to remain finite as $z^* \to \infty$.

Equation (7.53), together with its boundary conditions, constitutes a linear boundary value problem for $\Psi(y, z^*)$. It is generally not simple to obtain solutions to (7.53) for realistic mean zonal wind profiles. Nevertheless, we can obtain some useful information on stability properties simply by analyzing the energetics of the system.

Dividing (7.53) by $(\overline{u} - c)$ and separating the resulting equation into real and imaginary parts, we obtain

$$\frac{\partial^2\Psi_r}{\partial y^2} + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\Psi_r}{\partial z^*}\right) - \left[k^2 - \delta_r\left(\partial\overline{q}/\partial y\right)\right]\Psi_r - \delta_i\partial\overline{q}/\partial y\Psi_i = 0 \tag{7.55}$$

$$\frac{\partial^2\Psi_i}{\partial y^2} + \frac{1}{\rho_0}\frac{\partial}{\partial z^*}\left(\varepsilon\rho_0\frac{\partial\Psi_i}{\partial z^*}\right) - \left[k^2 - \delta_r\left(\partial\overline{q}/\partial y\right)\right]\Psi_i + \delta_i\partial\overline{q}/\partial y\Psi_r = 0 \tag{7.56}$$

where

$$\delta_r = \frac{\overline{u} - c_r}{(\overline{u} - c_r)^2 + c_i^2}, \quad \text{and} \quad \delta_i = \frac{c_i}{(\overline{u} - c_r)^2 + c_i^2}$$

Similarly, dividing (7.54) through by $(\bar{u} - c)$ and separating into real and imaginary parts gives, for the boundary condition at $z^* = 0$,

$$\frac{\partial \Psi_r}{\partial z^*} + \frac{\partial \overline{u}}{\partial z^*}(\delta_i \Psi_i - \delta_r \Psi_r) = 0; \quad \frac{\partial \Psi_i}{\partial z^*} - \frac{\partial \overline{u}}{\partial z^*}(\delta_r \Psi_i + \delta_i \Psi_r) = 0 \tag{7.57}$$

Multiplying (7.55) by Ψ_i, (7.56) by Ψ_r, and subtracting the latter from the former yields

$$\begin{aligned} \rho_0 \left[\Psi_i \frac{\partial^2 \Psi_r}{\partial y^2} - \Psi_r \frac{\partial^2 \Psi_i}{\partial y^2} \right] + \left[\Psi_i \frac{\partial}{\partial z^*} \left(\varepsilon \rho_0 \frac{\partial \Psi_r}{\partial z^*} \right) \right. \\ \left. - \Psi_r \frac{\partial}{\partial z^*} \left(\varepsilon \rho_0 \frac{\partial \Psi_i}{\partial z^*} \right) \right] - \rho_0 \delta_i (\partial \overline{q}/\partial y) \left(\Psi_i^2 + \Psi_r^2 \right) = 0 \end{aligned} \tag{7.58}$$

Using the chain rule of differentiation, (7.58) can be expressed in the form

$$\begin{aligned} \rho_0 \frac{\partial}{\partial y} \left[\Psi_i \frac{\partial \Psi_r}{\partial y} - \Psi_r \frac{\partial \Psi_i}{\partial y} \right] + \frac{\partial}{\partial z^*} \left[\varepsilon \rho_0 \left(\Psi_i \frac{\partial \Psi_r}{\partial z^*} - \Psi_r \frac{\partial \Psi_i}{\partial z^*} \right) \right] \\ - \rho_0 \delta_i (\partial \overline{q}/\partial y) \left(\Psi_i^2 + \Psi_r^2 \right) = 0 \end{aligned} \tag{7.59}$$

The first term in brackets in (7.59) is a perfect differential in y; the second term is a perfect differential in z^*. Thus, if (7.59) is integrated over the y, z^* domain, the result can be expressed as

$$\begin{aligned} \int_0^\infty \left[\Psi_i \frac{\partial \Psi_r}{\partial y} - \Psi_r \frac{\partial \Psi_i}{\partial y} \right]_{-L}^{+L} \rho_0 dz^* + \int_{-L}^{+L} \left[\varepsilon \rho_0 \left(\Psi_i \frac{\partial \Psi_r}{\partial z^*} - \Psi_r \frac{\partial \Psi_i}{\partial z^*} \right) \right]_0^\infty dy \\ = \int_{-L}^{+L} \int_0^\infty \rho_0 \delta_i \frac{\partial \overline{q}}{\partial y} \left(\Psi_i^2 + \Psi_r^2 \right) dy dz^* \end{aligned} \tag{7.60}$$

However, from (7.51) $\Psi_i = \Psi_r = 0$ at $y = \pm L$ so that the first integral in (7.60) vanishes. Furthermore, if Ψ remains finite as $z^* \to \infty$, the contribution to the second integral of (7.60) at the upper boundary vanishes. If we then use (7.57) to eliminate the vertical derivatives in this term at the lower boundary, (7.60) can

be expressed as

$$c_i \left[\int_{-L}^{+L} \int_{0}^{\infty} \frac{\partial \overline{q}}{\partial y} \frac{\rho_0 |\Psi|^2}{|\overline{u} - c|^2} dy dz^* - \int_{-L}^{+L} \varepsilon \frac{\partial \overline{u}}{\partial z^*} \frac{\rho_0 |\Psi|^2}{|\overline{u} - c|^2} \bigg|_{z^*=0} dy \right] = 0 \qquad (7.61)$$

where $|\Psi|^2 = \Psi_r^2 + \Psi_i^2$ is the disturbance amplitude squared.

Equation (7.61) has important implications for the stability of quasi-geostrophic perturbations. For unstable modes, c_i must be nonzero, and thus the quantity in square brackets in (7.61) must vanish. Because $|\Psi|^2/|\overline{u} - c|^2$ is nonnegative, instability is possible only when $\partial \overline{u}/\partial z^*$ at the lower boundary and $(\partial \overline{q}/\partial y)$ in the whole domain satisfy certain constraints:

1. If $\partial \overline{u}/\partial z^*$ at $z^* = 0$ (which by thermal wind balance implies that the meridional temperature gradient vanishes at the boundary), the second integral in (7.61) vanishes. Thus, the first integral must also vanish for instability to occur. This can occur only if $\partial \overline{q}/\partial y$ changes sign within the domain (i.e., $\partial \overline{q}/\partial y = 0$ somewhere). This is referred to as the *Rayleigh necessary condition* and is another demonstration of the fundamental role played by potential vorticity. Because $\partial \overline{q}/\partial y$ is normally positive, it is clear that in the absence of temperature gradients at the lower boundary, a region of negative meridional potential vorticity gradients must exist in the interior for instability to be possible.
2. If $\partial \overline{q}/\partial y \geq 0$ everywhere, then it is necessary that $\partial \overline{u}/\partial z^* > 0$ somewhere at the lower boundary for $c_i > 0$.
3. If $\partial \overline{u}/\partial z^* < 0$ everywhere at $z^* = 0$, then it is necessary that $\partial \overline{q}/\partial y < 0$ somewhere for instability to occur. Thus, there is an asymmetry between westerly and easterly shear at the lower boundary, with the former more favorable for baroclinic instability.

The basic state potential vorticity gradient in (7.49) can be written in the form

$$\frac{\partial \overline{q}}{\partial y} = \beta - \frac{\partial^2 \overline{u}}{\partial y^2} + \frac{\varepsilon}{H} \frac{\partial \overline{u}}{\partial z^*} - \varepsilon \frac{\partial^2 \overline{u}}{\partial z^{*2}} - \frac{\partial \varepsilon}{\partial z^*} \frac{\partial \overline{u}}{\partial z^*}$$

Because β is positive everywhere, if ε is constant a negative basic state potential vorticity gradient can occur only for strong positive mean flow curvature (i.e., $\partial^2 \overline{u}/\partial y^2$ or $\partial^2 \overline{u}/\partial z^{*2} \gg 0$) or strong negative vertical shear ($\partial \overline{u}/\partial z^* \ll 0$). Strong positive meridional curvature can occur at the core of an easterly jet or on the flanks of a westerly jet. Instability associated with such horizontal curvature is referred to as *barotropic instability*. The normal baroclinic instability in midlatitudes is associated with mean flows in which $\partial \overline{q}/\partial y > 0$ and $\partial \overline{u}/\partial z^* > 0$ at the ground. Hence, a mean meridional temperature gradient at the ground is essential for the existence of such instability. Baroclinic instability can also be excited at the tropopause because of the rapid decrease of ε with height if there

is a sufficiently strong easterly mean wind shear to cause a local reversal in the mean potential vorticity gradient.

7.4.3 The Eady Stability Problem

This section analyzes the structures (eigenfunctions) and growth rates (eigenvalues) for unstable modes in the simplest possible model that satisfies the necessary conditions for instability in a continuous atmosphere given in the previous subsection. For simplicity, we make the following assumptions:

- Basic state density constant (Boussinesq approximation).
- f-plane geometry ($\beta = 0$).
- $\partial \overline{u}/\partial z^* = \Lambda =$ constant.
- Rigid lids at $z^* = 0$ and H.

These conditions are only a crude model of the atmosphere, but provide a first approximation for study of the dependence of vertical structure on horizontal scale and stability. Despite the zero mean potential vorticity in the domain, the Eady model satisfies the necessary conditions for instability discussed in the previous subsection because vertical shear of the basic state mean flow at the upper boundary provides an additional term in (7.61) that is equal and opposite to the lower boundary integral.

Using the aforementioned approximations, the quasi-geostrophic potential vorticity equation is

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right) q' = 0 \tag{7.62}$$

where the perturbation potential vorticity is given by

$$q' = \nabla_h \psi' + \varepsilon \frac{\partial^2 \psi}{\partial z^{*2}} \tag{7.63}$$

and $\varepsilon \equiv f_0^2/N^2$. The thermodynamic energy equation at the horizontal boundaries ($z^* = 0, H$) is

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right) T' = 0 \tag{7.64}$$

Letting

$$\begin{aligned} \psi'\left(x, y, z^*, t\right) &= \Psi\left(z^*\right) \cos\, ly \, \exp\left[ik\left(x - ct\right)\right] \\ \overline{u}\left(z^*\right) &= \Lambda z^* \end{aligned} \tag{7.65}$$

where as in the previous subsection $\Psi(z^*)$ is a complex amplitude and c a complex phase speed, and substituting from (7.65) into (7.62) we find that

$$\left(\overline{u} - c\right) q' = 0 \tag{7.66}$$

From (7.66) we see that either $q' = 0$ or $\overline{u} - c = 0$. Since $\overline{u} = \Lambda z^*$, $\overline{u} - c = 0$ allows for nonzero q' only at the level where the wave speed matches the flow speed. Moreover, since $\overline{u}$ is strictly real, the wave speed in this case must be as well, so this particular solution branch describes singular neutral modes (spikes of PV at the steering level of the wave). They are useful for describing the evolution of potential vorticity disturbances in the Eady model, but they do not describe instability. Therefore, unstable modes must have $q' = 0$, which from (7.63) implies that the vertical structure is given by the solution of the second-order differential equation

$$\frac{d^2\Psi}{dz^{*2}} - \alpha^2\Psi = 0 \tag{7.67}$$

where $\alpha^2 = \left(k^2 + l^2\right)/\varepsilon$. A similar substitution into (7.64) yields the boundary conditions

$$\left(\Lambda z^* - c\right) d\Psi/dz^* - \Psi\Lambda = 0, \quad \text{at } z^* = 0, H \tag{7.68}$$

valid for rigid horizontal boundaries ($w^* = 0$) at the surface ($z^* = 0$) and the tropopause ($z^* = H$).

The general solution of (7.67) can be written in the form

$$\Psi\left(z^*\right) = A\sinh\alpha z^* + B\cosh\alpha z^* \tag{7.69}$$

Substituting from (7.69) into the boundary conditions (7.68) for $z^* = 0$ and H yields a set of two linear homogeneous equations in the amplitude coefficients A and B:

$$-c\alpha A - B\Lambda = 0$$
$$\alpha(\Lambda H - c)\left(A\cosh\alpha H + B\sinh\alpha H\right) - \Lambda\left(A\sinh\alpha H + B\cosh\alpha H\right) = 0$$

As in the two-layer model, a nontrivial solution exists only if the determinant of the coefficients of A and B vanishes. Again, this leads to a quadratic equation in the phase speed c. The solution (see Problem 7.12) has the form

$$c = \frac{\Lambda H}{2} \pm \frac{\Lambda H}{2}\left[1 - \frac{4\cosh\alpha H}{\alpha H\sinh\alpha H} + \frac{4}{\alpha^2 H^2}\right]^{1/2} \tag{7.70}$$

Thus

$$c_i \neq 0 \quad \text{if} \quad 1 - \frac{4\cosh\alpha H}{\alpha H\sinh\alpha H} + \frac{4}{\alpha^2 H^2} < 0$$

and the flow is then baroclinically unstable. When the quantity in square brackets in (7.70) is equal to zero, the flow is said to be *neutrally stable*. This condition occurs for $\alpha = \alpha_c$ where

$$\alpha_c^2 H^2/4 - \alpha_c H\left(\tanh\alpha_c H\right)^{-1} + 1 = 0 \tag{7.71}$$

Using the identity

$$\tanh \alpha_c H = 2 \tanh\left(\frac{\alpha_c H}{2}\right) \Big/ \left[1 + \tanh^2\left(\frac{\alpha_c H}{2}\right)\right]$$

we can factor (7.71) to yield

$$\left[\frac{\alpha_c H}{2} - \tanh\left(\frac{\alpha_c H}{2}\right)\right]\left[\frac{\alpha_c H}{2} - \coth\left(\frac{\alpha_c H}{2}\right)\right] = 0 \qquad (7.72)$$

Thus, the critical value of α is given by $\alpha_c H/2 = /\coth(\alpha_c H/2)$, which implies $\alpha_c H \cong 2.4$. Hence, instability requires $\alpha < \alpha_c$ or

$$\left(k^2 + l^2\right) < \left(\alpha_c^2 f_0^2/N^2\right) \approx 5.76/L_R^2$$

where $L_R \equiv NH/f_0 \approx 1000$ km is the Rossby *radius of deformation* for a continuously stratified fluid—compare to λ^{-1} defined just below equation (7.16). For waves with equal zonal and meridional wave numbers ($k = l$), the wavelength of maximum growth rate turns out to be AQ:1

$$L_m = 2\sqrt{2}\pi L_R/(H\alpha_m) \cong 5500 \text{ km}$$

where α_m is the value of α for which kc_i is a maximum.

Substituting this value of α into the solution for the vertical structure of the streamfunction (7.69) and using the lower boundary condition to express the coefficient B in terms of A, we can determine the vertical structure of the most unstable mode. As shown in Figure 7.10, trough and ridge axes slope westward with height, in agreement with the requirements for extraction of available potential energy from the mean flow. The axes of the warmest and coldest air, however, tilt eastward with height, a result that could not be determined from the two-layer model where temperature was given at a single level. Furthermore, Figures 7.10a and 7.10b show that east of the upper-level trough axis, where the perturbation meridional velocity is positive, the vertical velocity is also positive. Thus, parcel motion is poleward and upward in the region where $\theta' > 0$. Conversely, west of the upper-level trough axis parcel motion is equatorward and downward where $\theta' < 0$. Both cases are thus consistent with the energy-converting parcel trajectory slopes shown earlier in Figure 7.8.

Although the Eady model yields unstable solutions that resemble the baroclinic waves that dominate the statistics of midlatitude eddy fields, individual cyclone events are highly localized and often develop at a faster rate than described by the growth rate of the most unstable mode. Both issues are resolved by considering initial value problems for finite-amplitude, *localized*, upper-level initial disturbances, as are often observed to precede the development of surface cyclones. Figure 7.11 shows the development of a surface cyclone from a

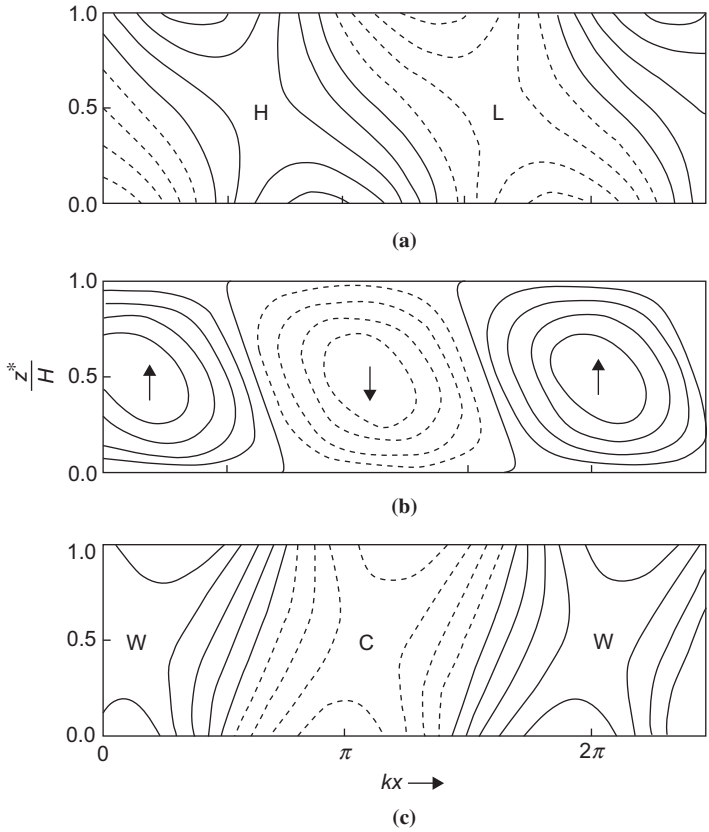

FIGURE 7.10 Properties of the most unstable Eady wave. (a) Contours of perturbation geopotential height; "H" and "L" designate ridge and trough axes, respectively. (b) Contours of vertical velocity; *up* and *down arrows* designate axes of maximum upward and downward motion, respectively. (c) Contours of perturbation temperature; "W" and "C" designate axes of warmest and coldest temperatures, respectively. In all panels 1 and 1/4 wavelengths are shown for clarity.

localized initial disturbance. Note that, after 48 hours, the full nonlinear solution resembles a mature extratropical cyclone, with warm air moving poleward and cold air moving equatorward around an area of low pressure. The linear solution for only the unstable growing modes approximates the full nonlinear solution and shows that the development may be explained by the projection of the initial disturbance onto only the unstable modes. Nonlinear effects relate to the poleward drift of the surface cyclone because of the interaction of the surface and upper-level PV anomalies and details of the development of the surface cold and warm fronts. Although the unstable modes are sufficient to explain the development of surface cyclones, it turns out that *transient* development is also possible, even when all the modes are neutral.

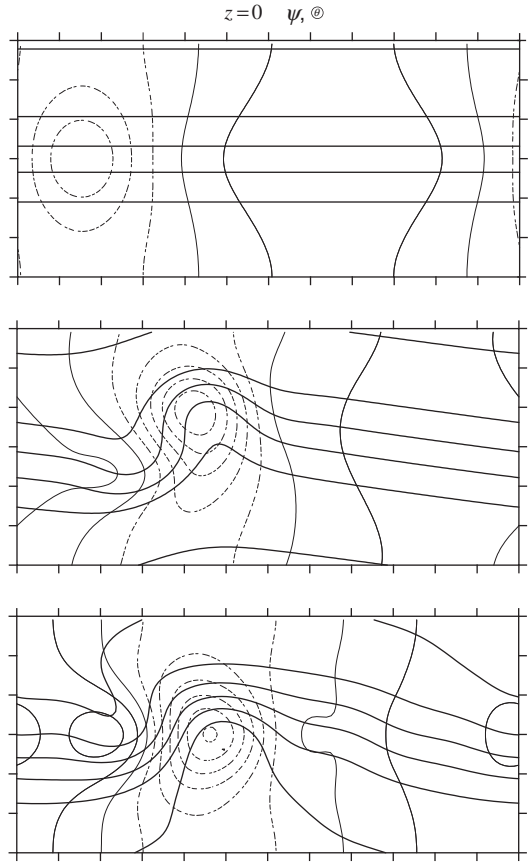

FIGURE 7.11 Development of a spatially localized disturbance in an idealized westerly jet stream. *Dashed lines* show pressure every 4 hPa and *solid lines* show potential temperature every 5 K. An initial surface disturbance (*top panel*) associated with an upper-level PV anomaly evolves into a well-developed extratropical cyclone 48 h later (*middle panel*). The linear solution for only the projection of the initial disturbance onto the unstable normal modes (*bottom panel*) captures most of the development of the full nonlinear solution. (*Adapted from Hakim, 2000. Copyright © American Meteorological Society. Reprinted with permission.*)

7.5 GROWTH AND PROPAGATION OF NEUTRAL MODES

As suggested earlier, baroclinic wave disturbances with certain favorable initial configurations may amplify rapidly even in the absence of baroclinic instability. Although the theory for the optimal initial perturbations that give rise to rapid transient growth is beyond the scope of this book, the main ideas can be exposed mathematically and illustrated with an example from the two-layer model.

A central concept for growth involves a metric, which is needed to measure the amplitude of disturbances; we call this measure a *norm*. A familiar example is the magnitude of a scalar z, $|z|$, which gives the absolute value for

real-valued z. For complex-valued $z = a+ib$, the magnitude gives the length of a vector on the complex plane, as defined by $|z| = (z^*z)^{1/2} = (a^2+b^2)^{1/2}$, where the asterisk denotes a complex conjugate. For *functions*, we adopt a natural extension of this definition:

$$|f| = [\langle f, f\rangle]^{1/2} \tag{7.73}$$

where, for functions of a single independent variable, x,

$$\langle f, g\rangle = \int f^* g\, dx \tag{7.74}$$

is an *inner product* that gives a scalar result. In particular, for $g = f$, (7.74) reveals that $|f|^2$ gives the integral over f^*f, which is a continuous analog of the sum of squared scalar components for complex numbers discussed previously. We define *amplification* over a time interval as the ratio of the norm at the time of interest, t, to the norm at the initial time, t_0:

$$A = \frac{|f|_t}{|f|_{t_0}} \tag{7.75}$$

Consider the normal modes of the Eady model, which have the form (dropping the prime)

$$\psi = \Psi(z^*)\, e^{i\phi}\, e^{kc_i t} \tag{7.76}$$

where $\phi = k(x - c_r t)$ and c_r and c_i are the real and imaginary parts of the phase speed, respectively. The amplification of a single normal mode, as defined by (7.75), for $t_0 = 0$ is given by

$$\left[\frac{\iint \Psi^* e^{-i\phi} e^{kc_i t}\, \Psi e^{i\phi} e^{kc_i t}\, dxdz}{\iint \Psi^* e^{-ikx}\, \Psi e^{ikx}\, dxdz}\right]^{1/2} = \left[e^{2kc_i t}\frac{\iint \Psi^*\Psi\, dxdz}{\iint \Psi^*\Psi\, dxdz}\right]^{1/2} = e^{kc_i t} \tag{7.77}$$

For neutral modes $c_i = 0$ and $A = 1$; otherwise, amplitude grows or decays exponentially depending on the sign of c_i. Note that, because of the space and time separation in (7.76), any function of the spatial structure of the normal modes (e.g., potential vorticity) amplifies at the same rate given by (7.77). This means that, for a single mode, the exponential growth found in (7.77) is independent of the norm.

Consider now the sum of two *neutral* modes, $\psi_1 + \psi_2$. In this case, amplification is given by

$$\left[\frac{\langle \psi_1+\psi_2, \psi_1+\psi_2\rangle_t}{\langle \psi_1+\psi_2, \psi_1+\psi_2\rangle_{t_0}}\right]^{1/2} = \left[\frac{\langle \psi_1, \psi_1\rangle_t + \langle \psi_2, \psi_2\rangle_t + 2\langle \psi_1, \psi_2\rangle_t}{\langle \psi_1, \psi_1\rangle_{t_0} + \langle \psi_2, \psi_2\rangle_{t_0} + 2\langle \psi_1, \psi_2\rangle_{t_0}}\right]^{1/2} \tag{7.78}$$

where we have used the fact that $\langle f, g\rangle = \langle g, f\rangle^*$. Since the modes are neutral, $\langle \psi_1, \psi_1\rangle$ and $\langle \psi_2, \psi_2\rangle$ are constants, and amplification is determined by the term

$$\langle \psi_1, \psi_2\rangle_t = e^{ik(c_1-c_2)t} \iint \Psi_1^* \Psi_2 \, dxdz \tag{7.79}$$

This shows that if the modes are *orthogonal in the chosen norm*—that is,

$$\iint \Psi_1^* \Psi_2 \, dxdz = 0 \tag{7.80}$$

then there is no amplification. On the other hand, if the modes are not orthogonal, then the amplitude fluctuates periodically proportional to the difference in the phase speeds of the modes. A physical interpretation is that, because the modes move relative to one another, there are times when the sum gives an upshear-tilted disturbance[3] and the amplitude of the sum increases. Eventually, the relationship changes so that the modes tilt downshear and the amplitude of the sum decreases. Therefore, the amplification effect is transient and depends on the chosen norm. In contrast, the unstable normal modes grow exponentially for all time, and measures of amplification do not depend on the norm.

7.5.1 Transient Growth of Neutral Waves

Consider now a specific example for the two-layer model of Section 7.2. If we neglect the β effect, let $U_m = 0$, and assume that $k^2 > 2\lambda^2$, the two-layer model has two oppositely propagating neutral solutions given by (7.24) with zonal phase speeds

$$c_1 = +U_T\mu; \quad c_2 = -U_T\mu \tag{7.81}$$

where

$$\mu = \left[\frac{k^2 - 2\lambda^2}{k^2 + 2\lambda^2}\right]^{1/2}$$

Then from (7.17) a disturbance consisting of these two modes can be expressed as

$$\psi_m = A_1 \exp\left[ik\left(x - c_1 t\right)\right] + A_2 \exp\left[ik\left(x - c_2 t\right)\right] \tag{7.82a}$$

$$\psi_T = B_1 \exp\left[ik\left(x - c_1 t\right)\right] + B_2 \exp\left[ik\left(x - c_2 t\right)\right] \tag{7.82b}$$

but from (7.18)

$$c_1 A_1 - U_T B_1 = 0; \quad c_2 A_2 - U_T B_2 = 0$$

[3]This interpretation is based on energy, which, as shown in Section 7.3.2, is proportional to ψ^2.

So that with the aid of (7.81)

$$B_1 = \mu A_1; \quad B_2 = -\mu A_2 \tag{7.83}$$

For an initial disturbance confined entirely to the upper level, it is easily verified that $\psi_m = \psi_T$ ($\psi_1 = 2\psi_m$ and $\psi_3 = 0$). Thus, initially $A_1 + A_2 = B_1 + B_2$, and substituting from (7.83) gives

$$A_2 = -A_1 \left[(1-\mu)/(1+\mu)\right]$$

Hence, if A_1 is real, the streamfunctions of (7.82a,b) can be expressed as

$$\begin{aligned}\psi_m(x,t) &= A_1\left[\cos\left[k(x-\mu U_T t)\right] - \frac{(1-\mu)}{(1+\mu)}\cos\left[k(x+\mu U_T t)\right]\right] \\ &= \frac{2\mu A_1}{(1+\mu)}\left[\cos kx\cos(k\mu U_T t) + \frac{1}{\mu}\sin kx\sin(k\mu U_T t)\right]\end{aligned} \tag{7.84a}$$

$$\begin{aligned}\psi_T(x,t) &= \mu A_1\left[\cos\left[k(x-\mu U_T t)\right] + \frac{(1-\mu)}{(1+\mu)}\cos\left[k(x+\mu U_T t)\right]\right] \\ &= \frac{2\mu A_1}{(1+\mu)}\left[\cos kx\cos(k\mu U_T t) + \mu\sin kx\sin(k\mu U_T t)\right]\end{aligned} \tag{7.84b}$$

The first forms on the right side in (7.84a,b) show that for small μ the barotropic mode initially consists of two waves of nearly equal amplitude that are 180° out of phase so that they nearly cancel, whereas the baroclinic mode initially consists of two very weak waves that are in phase.

As time advances, the two oppositely propagating barotropic modes begin to reinforce each other, leading to a maximum-amplitude trough 90° to the east of the initial trough at time $t_m = \pi/(2k\mu U_T)$. From (7.84) the baroclinic and barotropic modes at time t_m are then

$$\begin{aligned}(\psi_m)_{\max} &= 2A_1(1+\mu)^{-1}\sin kx \\ (\psi_T)_{\max} &= 2A_1\mu^2(1+\mu)^{-1}\sin kx\end{aligned} \tag{7.85a,b}$$

from which it is shown easily that

$$(\psi_1)_{\max} = 2A_1\left(1+\mu^2\right)(1+\mu)^{-1}\sin kx; \quad (\psi_3)_{\max} = 2A_1(1-\mu)\sin kx$$

so that for small μ the resulting disturbance is nearly barotropic. Thus, the initial disturbance not only amplifies but spreads in the vertical. The time for growth to maximum amplitude is inversely proportional to the basic state thermal wind; the maximum amplitude, however, depends only on the initial amplitude and the parameter μ.

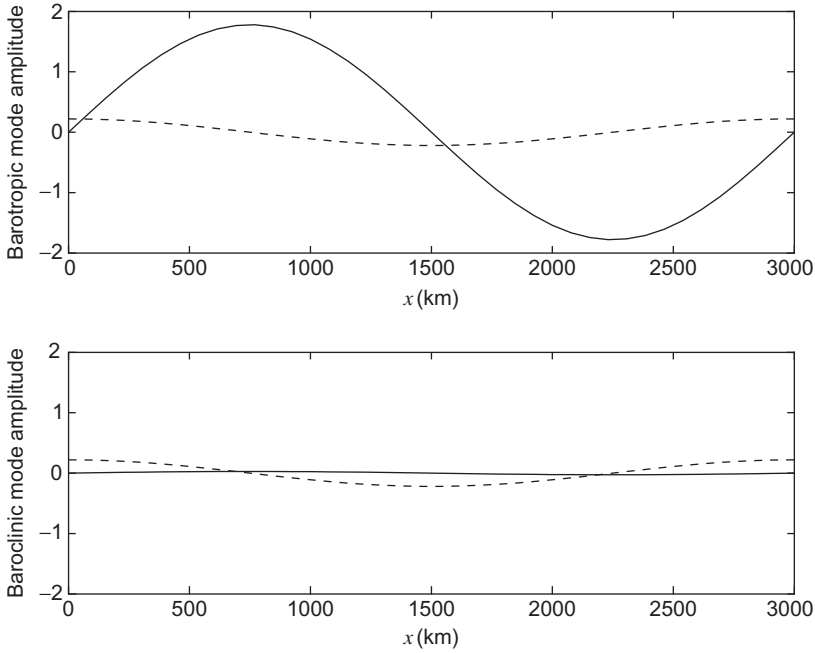

FIGURE 7.12 Zonal distribution of the amplitude of a transient neutral disturbance initially confined entirely to the upper layer in the two-layer model. *Dashed lines* show the initial distribution of the baroclinic and barotropic modes; *solid lines* show the distribution at the time of maximum amplification.

Figure 7.12 shows the initial and maximum amplitudes of the barotropic and baroclinic streamfunctions for $f_0 = 10^{-4}\,\mathrm{s}^{-1}$, $\sigma = 2\times10^{-6}\,\mathrm{m}^2\,\mathrm{Pa}^{-2}\,\mathrm{s}^{-2}$, $U_T = 35\,\mathrm{m\,s}^{-1}$, and a zonal wavelength of 3000 km. In this case, the amplitude of the barotropic disturbance increases by a factor of 8 in about 48 h. Although the most unstable normal mode for these conditions amplifies by a similar amount in less than 1 day, if the initial upper-level neutral disturbance has a velocity amplitude of a few meters per second, its growth on the time scale of a few days may dominate over a normal mode instability that grows exponentially from a much smaller initial perturbation.

As with normal mode baroclinic instability, the energy source for the transient amplification of the neutral modes is conversion of the available potential energy of the mean flow into disturbance potential energy by meridional temperature advection, followed immediately by conversion to disturbance kinetic energy by the secondary vertical circulation. This secondary circulation has a vertical velocity field that lags the upper-level streamfunction field by 90° phase. Maximum upward motion and lower-level convergence thus occur to the west of the initial upper-level ridge, and maximum downward motion and lower-level divergence occur west of the initial upper-level trough. Vorticity

tendencies associated with this convergence and divergence pattern partly balance the eastward vorticity advection at the upper level and also act to produce a lower-level trough to the east of the initial upper-level trough and lower-level ridge to the east of the initial upper-level ridge. The result is that a pattern of nearly barotropic troughs and ridges develops 90° to the east of the initial trough and ridge pattern. Unlike normal mode instability, the growth in this case does not continue indefinitely. Rather, for a given value of μ the maximum growth occurs at the time for which $t_m = \pi/(2kU_T\mu)$. As the zonal wave number approaches the short-wave instability cutoff, the total amplification increases, but the amplification time also increases.

7.5.2 Downstream Development

It was noted in Section 7.2.2 that under some circumstances the zonal group velocity may exceed the zonal phase speed so that the energy of a wave group propagates faster than the individual wave disturbances. This dispersion effect is demonstrated for barotropic Rossby waves in Problem 5.15. The β effect is not necessary, however, for the development of new disturbances to occur downstream of the original wave disturbances on the synoptic scale. The two-layer channel model with constant Coriolis parameter can be used to demonstrate the process of dispersion and downstream development of neutral waves in a simple context.

Expressed in terms of wave frequency, the dispersion relation of (7.24) becomes

$$\nu = kU_m \pm k\mu U_T \tag{7.86}$$

where μ is defined below (7.81). The corresponding group velocity is then

$$c_{gx} = \frac{\partial \nu}{\partial k} = U_m \pm U_T\mu\left(1 + \frac{4k^2\lambda^2}{k^4 - 4\lambda^4}\right) \tag{7.87}$$

Comparison of (7.87) with (7.81) shows that relative to the mean zonal flow the group velocity exceeds the phase velocity by the factor in parentheses on the right side of (7.87) for both eastward- and westward-directed neutral modes. For example, if $k^2 = 2\lambda^2\left(1 + \sqrt{2}\right)$, the group velocity equals twice the phase speed, which is the situation shown schematically earlier in Figure 7.4.

Observations of baroclinic waves in the midlatitude storm tracks show clearly the downstream development of wave energy. For example, over the North Pacific ocean, the baroclinic wave phase speed is about 10 m s^{-1}, while the group speed is about 30^{-1} (Figure 7.13). As a result, the impact of extratropical cyclones that develop near the Asian coastline may be felt along the West Coast of North America in three days, even though the initial disturbance may take five days to reach the middle Pacific during its decay phase.

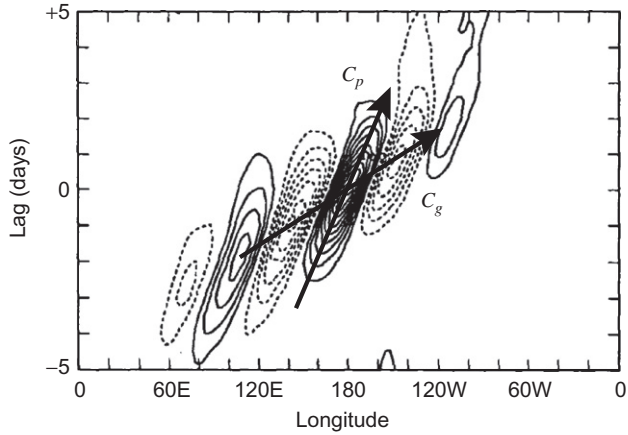

FIGURE 7.13 Time-series correlation of meridional wind as a function of longitude and time lag. Correlation is shown every 0.1 unit (i.e., a unitless number between −1 and 1), with the reference time series at 180. Results are averaged over 30 to 60°N latitude during winter. (*From Chang, 1993. Copyright © American Meteorological Society. Reprinted with permission.*)

SUGGESTED REFERENCES

Charney (1947) is the classic paper on baroclinic instability. The mathematical level is advanced, but Charney's paper contains an excellent qualitative summary of the main results that is very readable.

Hoskins, McIntyre, and Robertson (1985) discuss cyclogenesis and baroclinic instability from a potential vorticity perspective.

Pedlosky, *Geophysical Fluid Dynamics* (2nd ed.), contains a thorough treatment of normal mode baroclinic instability at a mathematically advanced level.

Pierrehumbert and Swanson (1995) review numerous aspects of baroclinic instability, including spatiotemporal development.

Vallis, *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*, covers topics in baroclinic instability similar to those considered here.

PROBLEMS

7.1. Show using (7.25) that the maximum growth rate for baroclinic instability when $\beta = 0$ occurs for

$$k^2 = 2\lambda^2\left(\sqrt{2} - 1\right)$$

How long does it take the most rapidly growing wave to amplify by a factor of e^1 if $\lambda^2 = 2 \times 10^{-12}\,\mathrm{m}^{-1}$ and $U_T = 20\,\mathrm{m\ s}^{-1}$?

7.2. Solve for ψ_3' and ω_2' in terms of ψ_1' for a baroclinic Rossby wave whose phase speed satisfies (7.23). Explain the phase relationship among ψ_1', ψ_3', ω_2' and in terms of the quasi-geostrophic theory. (Note that $U_T = 0$ in this case.)

7.3. For the case $U_1 = -U_3$ and $k^2 = \lambda^2$, solve for ψ_3' and ω_2' in terms of ψ_1' for marginally stable waves—that is, $\delta = 0$ in (7.21).

7.4. For the case $\beta = 0$, $k^2 = \lambda^2$, and $U_m = U_T$, solve for ψ_3' and ω_2' in terms of ψ_1'. Explain the phase relationships among ω_2', ψ_1', and ψ_3' in terms of the energetics of quasi-geostrophic waves for the amplifying wave.

7.5. Suppose that a baroclinic fluid is confined between two rigid horizontal lids in a rotating tank in which $\beta = 0$, but friction is present in the form of linear drag proportional to the velocity (i.e., $\mathbf{Fr} = -\mu\mathbf{V}$). Show that the two-level model perturbation vorticity equations in Cartesian coordinates can be written as

$$\left(\frac{\partial}{\partial t} + U_1\frac{\partial}{\partial x} + \mu\right)\frac{\partial^2\psi_1'}{\partial x^2} - \frac{f}{\delta p}\omega_2' = 0$$

$$\left(\frac{\partial}{\partial t} + U_3\frac{\partial}{\partial x} + \mu\right)\frac{\partial^2\psi_3'}{\partial x^2} + \frac{f}{\delta p}\omega_2' = 0$$

where perturbations are assumed in the form given in (7.8). Assuming solutions of the form (7.17), show that the phase speed satisfies a relationship similar to (7.21) with β replaced everywhere by $i\mu k$, and that as a result the condition for baroclinic instability becomes

$$U_T > \mu\left(2\lambda^2 - k^2\right)^{-1/2}$$

7.6. For the case $\beta = 0$ determine the phase difference between the 250- and 750-hPa geopotential fields for the most unstable baroclinic wave (see Problem 7.1). Show that the 500-hPa geopotential and thickness fields (ψ_m, ψ_T) are 90° out of phase.

7.7. For the conditions of Problem 7.6, given that the amplitude of ψ_m is $A = 10^7\,\mathrm{m^2\,s^{-1}}$, solve the system (7.18) and (7.19) to obtain B. Let $\lambda^2 = 2.0 \times 10^{-12}\,\mathrm{m^{-2}}$ and $U_T = 15\,\mathrm{m\,s^{-1}}$. Use your results to obtain expressions for ψ_1' and ψ_3'.

7.8. For the situation of Problem 7.7, compute ω_2' using (7.28).

7.9. Compute the total potential energy per unit cross-sectional area for an atmosphere with an adiabatic lapse rate given that the temperature and pressure at the ground are $p = 10^5$ Pa and $T = 300\,\mathrm{K}$, respectively.

7.10. Consider two air masses at the uniform potential temperatures $\theta_1 = 320\,\mathrm{K}$ and $\theta_2 = 340\,\mathrm{K}$ that are separated by a vertical partition as shown in Figure 7.7. Each air mass occupies a horizontal area of 10^4 m^2 and extends from the surface ($p_0 = 10^5$ Pa) to the top of the atmosphere. What is the available potential energy for this system? What fraction of the total potential energy is available in this case?

7.11. For the unstable baroclinic wave that satisfies the conditions given in Problems 7.7 and 7.8, compute the energy conversion terms in (7.37) and (7.38) and thus obtain the instantaneous rates of change of the perturbation kinetic and available potential energies.

7.12. Starting with (7.62) and (7.64), derive the phase speed c for the Eady wave given in (7.70).

7.13. Unstable baroclinic waves play an important role in the global heat budget by transferring heat poleward. Show that for the Eady wave solution the poleward heat flux averaged over a wavelength,

$$\overline{v'T'} = \frac{1}{L}\int_0^L v'T'dx$$

is independent of height and is positive for a growing wave. How does the magnitude of the heat flux at a given instant change if the mean wind shear is doubled?

7.14. Assuming that the coefficient A in (7.69) is real, obtain an expression for the geostrophic streamfunction $\psi'(x, y, z^*, t)$ for the most unstable mode in the Eady stability problem for the case $k = l$. Use this result to derive an expression for the corresponding vertical velocity w^* in terms of A.

7.15. For the neutral baroclinic wave disturbance in the two-layer model given by (7.85a,b), derive the corresponding ω_2' field. Describe how the convergence and divergence fields associated with this secondary circulation influence the evolution of the disturbance.

7.16. For the situation of Problem 7.15, derive expressions for the conversion of zonal available potential energy to eddy available potential energy and the conversion of eddy available potential energy to eddy kinetic energy.

MATLAB Exercises

M7.1. By varying the input zonal wavelength in the MATLAB script **twolayer_model_1A.m**, find the shortest zonal wavelength for which the two-layer model is baroclinically unstable and the wavelength of maximum instability (i.e., most rapid growth rate) in the case where the basic state thermal wind is 15 m s^{-1}. The case given corresponds to the situation given in Section 7.2.1. The code of **twolayer_model_1B.m** differs only in that a finite meridional width is assumed with meridional wave number $m = \pi/(3000\,\text{km})$. Compare the zonal wavelengths for short-wave cutoff and for maximum growth rate for these two cases.

M7.2. Use the MATLAB script **twolayer_model_2A.m** to examine the transient growth associated with a disturbance that is initially entirely confined to the 250-hPa level. Let the basic state thermal wind be 25 m s^{-1}. By examining zonal wavelengths shorter than the instability cutoff determined in Exercise M7.1, find the zonal wavelength that gives the largest amplification of the disturbance. Repeat this exercise using the script **twolayer_model_2B.m**, which includes a meridional dependence with wave number $m = \pi/(3000\ \text{km})$. Note that the program will terminate if you choose a wavelength that is long enough to be baroclinically unstable. Discuss the vertical structure of the transiently growing stable modes. What sort of situation in the real atmosphere does this solution crudely model?

M7.3. The MATLAB script **twolayer_model_3b.m** shows the propagation of a wave packet composed of the sum of nine eastward propagating neutral waves of the two-layer model, ranging in wave number from 0.6 to 1.6 k, where $k = 2\pi/L$ and $L = 1850\,\text{km}$. Run the script and estimate from the animation

the characteristic phase velocity and the characteristic group velocity for the cases of $U_T = 15$ and $30\,\mathrm{m\,s^{-1}}$.

M7.4. Suppose that a meridional dependence is included in the eddy fields in the two-layer model so that (7.17) becomes

$$\psi_m = A\cos(my)\,e^{ik(x-ct)}; \quad \psi_T = B\cos(my)\,e^{ik(x-ct)}$$

where $m = 2\pi/L_y$ and $L_y = 3000\,\mathrm{km}$. Determine the zonal wavelength of the most unstable wave. Solve for the fields ψ_m, ψ_T, and ω_2' and plot (x, y) cross-sections of these fields using the MATLAB script **contour_sample.m** as a template. *Hint:* The solutions in this case are similar to those of Problems 7.6 through 7.8 but with k^2 replaced by $k^2 + m^2$ everywhere.

M7.5. The MATLAB script **eady_model_1.m** shows vertical and meridional cross-sections of the solution for the Eady model (Section 7.4.3) for the most unstable wave mode at time $t = 0$. Modify this script to plot the solution for the Eady wave corresponding to the neutral stability condition given in (7.71). Explain the vertical structure in this case in terms of quasi-geostrophic theory.