18
Oceanic Analogues of Large-scale Atmospheric Motions

Jule G. Charney and
Glenn R. Flierl

18.1 Introduction

Newton (1687, book 2, propositions 48–50) and Laplace (1799) were aware that the principles governing the ocean tides would also govern the atmospheric tides. Helmholtz (1889) showed that ocean waves and billow clouds were manifestations of the same hydrodynamic instability, and he speculated that storms were caused by a similar instability. Had he known the structure of the Gulf Streams meanders, he might have speculated on their dynamic similarities to storms as well. Such intercomparisons were natural to the great hydrodynamicists of the past who took the entire universe of fluid phenomena as their domain. Although a degree of provincialism was introduced in the late nineteenth and early twentieth centuries by the exigencies of weather forecasting, it was just the practical requirements of weather observation that stimulated the development of modern dynamic meteorology and led to the deepening of its connections with physical oceanography. The recent explosive growth of the three-dimensional data base, the exploration of other planetary atmospheres, and the resulting increase in theoretical activity have greatly extended the list of ocean-atmosphere analogues. Indeed, it is now no exaggeration to say that there is scarcely a fluid dynamical phenomenon in planetary atmospheres that does not have its counterpart in the oceans and vice versa. This had led to the discipline, geosphysical fluid dynamics, whose guiding principles are intended to apply equally to oceans and atmospheres.

Within this discipline the dominance of the earth's rotation defines a subclass of large-scale phenomena whose dynamics may for the most part be derived from quasi-geostrophy. For several years the authors have conducted a graduate course at MIT on the dynamics of large-scale ocean and atmospheric circulations in the belief that a parallel consideration of large-scale oceanic and atmospheric motions would broaden the range of our students' experience and deepen their understanding of the principles of fluid geophysics. C.-G. Rossby (1951) put the matter well:

> It is fairly certain that the final formulation of a comprehensive theory for the general circulation of the atmosphere will require intimate cooperation between meteorologists and oceanographers. The fundamental problems associated with the heat, mass and momentum transfer at the sea surface concern both these sciences and demand a joint effort for their solution. However, an even stronger reason for that pooling of intellectual resources . . . may be found in the fact that the various theoretical analyses of the large-scale oceanic and atmospheric circulation patterns which have been called into being by our sudden wealth of observational data, appear to have so much in common that they may be looked upon as different facets of one broad general study, the ultimate aim of which might

be described as an attempt to formulate a comprehensive theory for fluid motion in planetary envelopes.

He went on to say that "comparisons between the circulation patterns in the atmosphere and in the oceans provide us with a highly useful substitute for experiments with controlled variations of the fundamental parameters."

Needless to say, we subscribe to Rossby's views and therefore have willingly undertaken the task of reviewing the fluid dynamics of a number of phenomena that have found explanation in one medium and are deemed to have important analogues in the other. Where the analogues have already been explained, we have given a brief review of some of their salient features, but where little is known, we have not refrained from interpolating simple models or speculations of our own. In doing so, we were aware that the subject has grown so large that it includes most of geophysical fluid dynamics, and that, even if it were limited to large-scale, quasi-geostrophic motions, it could not be encompassed in a review article of modest size. At least two excellent review articles on this topic, by N. Phillips (1963) and by H.-L. Kuo (1973), have already appeared, and there is now a text by Pedlosky (1979a) on geophysical fluid dynamics. For these reasons we have decided to limit ourselves to a small number of topics having to do primarily with disturbances of the principal atmospheric and oceanic currents, their propagation characteristics, their interactions with the embedding currents, and, to a lesser degree, with their self-interactions.

It is perhaps no accident that both Phillips and Kuo are meteorologists. Dynamic meteorology has benefited from a wealth of observational detail that has been denied to physical oceanography. Consequently, meteorologists have been the first to observe and explain many typical large-scale phenomena. This has by no means always been the case, but it has been so sufficiently often to justify the title of our review.

18.2 The General Circulations of Oceans and Atmospheres Compared

If it had been possible in a meaningful way, it would have been useful and instructive to make detailed dynamical comparisons of the general circulations of the atmosphere and oceans. This has not been so, and we must content ourselves with a few general remarks. Because of its relative transparency to solar radiation, the earth's atmosphere is heated from below; the oceans, like the relatively opaque atmosphere of Venus, are heated from above. The differential heating of the atmosphere produces a mean circulation that carries heat upward and poleward. Its baroclinic instabilities do likewise. These heat transports, combined with internal radiative heat transfer, tend to stabilize the atmosphere statically, and their effects are augmented by moist convection which drives the temperature lapse-rate toward the moist-adiabatic. The net result is that the atmosphere is rather uniformly stable for dry processes up to the tropopause. Above the tropopause, it is made increasingly more stable by absorption of solar ultraviolet radiation in the ozone layer. The oceans are rendered statically stable by heating from above, but heating cannot take place everywhere because the oceans, unlike the atmosphere, cannot dispose of internal heat by radiation to space; they must carry it back to the surface layers, where it can be lost by surface cooling. The most stable parts of the oceans are in the subtropical gyres, where the oceans are heated and Ekman pumping transfers the heat downward. The least stable parts are in the polar regions, where cold water is formed and carried downward by convection. The atmospheric troposphere has sometimes been compared with the waters above the thermocline, the tropopause with the thermocline, and the stratosphere with the deep waters below the thermocline (cf. Defant, 1961b). This comparison may be justified by the fact that it is in the atmospheric and oceanic tropospheres that the horizontal temperature gradients and the kinetic and potential energy densities are greatest. But from the standpoint of static stability, the absorption of radiation at the surface of the oceans makes its upper layers more analogous to the stratosphere. The static stability for the bulk of the atmosphere and oceans is determined by deep convection occurring in small regions. The analogy between deep convection in the atmosphere and deep convection in the oceans is between the narrow intertropical convergence zones over the oceans and the limited areas of cumulus convection over the tropical continents on the one hand, and the limited regions of deep-water formation in the polar seas on the other. From this standpoint, the ocean waters below the thermocline are more analogous to the atmospheric troposphere. Both volumes comprise more than 80% of the total by mass and both are controlled by deep convection.

It is remarkable that the regions of pronounced rising motion in the atmosphere and sinking motion in the oceans are so confined horizontally. Stommel (1962b) was the first to offer an explanation for the smallness of the regions of deep-water formation. His work motivated several attempts to explain the asymmetries in the circulation of a fluid heated differentially from above or below. We may cite as examples the experimental work of H. T. Rossby (1965) and the theoretical work of Killworth and Manins (1980) on laboratory fluid systems, and the papers of Goody and Robinson (1966) and Stone (1968) on the upper circulation of Venus. Because of its cloudiness, Venus has been assumed to be heated primarily from above, although it

now is known that sunlight does penetrate into the lower Venus atmosphere and that the high temperatures near the surface are due to a pronounced greenhouse effect (Keldysh, 1977; Young and Pollack, 1977; Tomasko, Doose, and Smith, 1979). When a fluid is heated from below, the rising branches are found to be narrow and the sinking branches broad; when it is heated from above the reverse is true. One may offer the qualitative explanation that it is the branch of the circulation that leaves the boundary and carries with it the properties of the boundary that has the greatest influence on the temperature of the fluid as a whole: convection is more powerful than diffusion. In the case of differential heating from below, the rising warm branch causes most of the fluid to be warm relative to the boundary and therefore gravitationally stable except in a narrow zone at the extreme of heating where the intense rising motion must occur. In the case of differential heating from above, the sinking cold branch causes the bulk of the fluid to be cold and gravitationally stable except in a narrow region at the extreme of cooling where the intense sinking motion must occur. Theoretical models of axisymmetric, thermally driven (Hadley) circulations in the atmosphere (Charney, 1973) show the same effect: a narrow rising branch and a broad sinking branch. This effect is strengthened by cumulus convection (Charney, 1969, 1971b; Bates, 1970; Schneider, 1977). The narrow rising branch of the Hadley circulation directly controls the dynamic and thermodynamic properties of the tropics and subtropics and indirectly influences the higher-latitude circulations. Similarly, the small sinking branches of the ocean circulation determine the near-homogeneous deep-water properties as well as some of the intermediate-water properties. But there is a difference: we know how the heat released in the ascending branch of the Hadley circulation is disposed of; we do not know how the cold water in the abyssal circulation is heated. Whatever the process, the existence of a preponderant mass of near-homogeneous water at depth forces great static stability in the shallow upper regions of the ocean. It demands a thermocline.

Again we have an analogy between the upper circulation of the oceans and the upper circulation of Venus. Rivas (1973, 1975) has shown that the intense circulation of Venus is confined to a thin layer within and just below the region of intense heating and cooling by radiation. The more or less independent circulation produced by the separate heat sources of the atmospheric stratosphere is similarly analogous to the upper ocean circulation. And here there is an atmosphere–ocean analogy pertaining to our knowledge of transfer processes. While the mechanisms of heat transfer in the stratosphere are fairly well known, the mechanism of transfer of thermally inactive gases or suspended particles is not, for the latter involves a knowledge of particle trajectories, which are not easily determined. The large-scale eddies in the atmosphere, insofar as they are nondissipative, cannot transfer a conserved quantity across isentropic surfaces. Thermal dissipation is required for parcels to move from the low-entropy troposphere to the high-entropy stratosphere. Atmospheric chemists sometimes postulate an ad hoc turbulent diffusion to explain the necessary vertical transfers of such substances as the oxides of nitrogen and the chlorofluoromethanes into the ozone layers. But it is doubtful whether this type of diffusion is needed to account for the actual transfer, because there already exists the nonconservative mechanism of radiative heat transfer, and this, together with the large-scale eddying motion, can by itself account for the transfers (Andrews and McIntyre, 1978a; Matsuno and Nakamura, 1979). The analogous oceanic problem has already been mentioned: how is heat or salt transferred from the deep ocean layers into the upper wind-stirred layers? The cold deep water produced in the polar seas and the intermediate salty water produced in the Mediterranean Sea must eventually find their way by dissipative processes to the surface, where they can be heated or diluted. While a number of internal dissipation mechanisms have been proposed in a speculative way (double diffusion, low-Richardson-number instability zones, internal-wave breaking) one may invoke Occam's razor, as has been done so successfully to explain the Gulf Stream as an inertial rather than a frictional boundary layer (Charney, 1955b; Morgan, 1956) and postulate no internal dissipative mechanism at all. Then, outside of the convective zones, properties will be advected by the mean flow or by eddies along isentropic surfaces, and move from one such surface to another only at the boundaries of the ocean basins where we may assume turbulent dissipation does occur. It would be interesting to see how far one could go with boundary dissipation alone. Welander (1959) and Robinson and Welander (1963) have taken a first step by investigating the motions of a conservative system communicating only with an upper mixed layer.

18.3 The Transient Motions

Meteorologists, pressed with the necessity of forecasting the daily weather, have always been concerned with the transient motions of the atmosphere. Attempts to understand the processes leading to growth, equilibration, translation and decay of these "synoptic-scale" (order 1000 km) eddies[1] have produced theories of baroclinic instability (Charney, 1947; Eady, 1949), Rossby wave motion (Rossby et al., 1939; Haurwitz, 1940b) and Ekman pumping (Charney and Eliassen, 1949). It was first recognized by Jeffreys (1926) and

demonstrated conclusively by Starr (1954, 1957) and Bjerknes (1955, 1957) that the dynamics of the mean circulation are strongly influenced by the transports of heat and zonal momentum by the eddies [for a review of these developments, see Lorenz (1967)]. This has prompted research on eddy dynamics and also on the parameterization of eddy fluxes (cf. Green, 1970; Rhines, 1977; Stone, 1978; Welander, 1973).

The study of eddy motions in the ocean is a new development. Although the existence of fluctuations in the Gulf Stream was reported by early observers such as Laval (1728) and Rennell (1832), actual *prediction* of ocean eddies has never been a particularly profitable exercise (perhaps the recent interest of yachtsmen in Gulf Stream rings may presage a change). The serious study of deep ocean fluctuations really began with M. Swallow (1961), Crease (1962), and J. Swallow (1971). At this point, oceanographers began to realize that the mid-ocean variable velocities were not, as might perhaps have been reasonably inferred from atmospheric experience, comparable to the mean flows but rather were an order of magnitude larger. This has spurred intensive experimental and theoretical investigations of the dynamics of the oceanic eddies and their roles in the general circulation (see chapter 11).

Comparisons between the oceanic and atmospheric eddies may be made in respect to their generation, propagation, interaction (both eddy-mean flow and eddy-eddy) and decay. In the sections below we shall describe some of the theoretical approaches to these problems.

In the atmosphere, energy conversion estimates (cf. Oort and Peixoto, 1974) clearly show a transformation of zonal-available potential energy into eddy-available potential energy, then into eddy kinetic energy and finally into heat by dissipation, with some transfer from eddy to zonal kinetic energy. The similarity of the growth phase of this cycle to that exhibited in the theory of small traveling perturbations of a baroclinically unstable (but barotropically stable) flow, leads naturally to the identification of the source of the waves as the baroclinic instability of the zonal flow. This idea has been supported further by the fact that the energy spectrum has peaks near zonal wavenumber six, which simple models predict to be the most rapidly growing wavenumber. However, attempts to apply these models directly to the atmosphere lead to problems: one expects that nonuniform mean flows, nonhomogeneous surface conditions, variable horizontal and vertical shears, etc., will alter the dynamics; and one may also wonder about the applicability of the small perturbation-normal mode approach.

The topographically and thermally forced standing eddies also draw upon zonal available potential energy (Holopainen, 1970). The processes by which they do this are not yet clearly understood; they may be related to the form-drag instability to be described in section 18.7.3. The standing eddies, like the transient eddies, also transport heat and momentum.

Overall energetic analyses have not been applied to oceanic data. However, budgets for basin-averaged kinetic and potential energy have been calculated for the "eddy-resolving general circulation models." These are reviewed by Harrison (1979b). In all but one of the 21 cases he considered, the eddy kinetic energy came from *both* mean kinetic and mean potential energy. Collectively, the eddies seem to be acting as dissipative mechanisms, but Harrison cautions that because the model statistics are inhomogeneous, the overall results may not be representative of the actual dynamics in any limited region. Indeed, the eddies may be acting as a negative viscosity in parts of the domain. Thus, Holland (1978) suggests that the eddies generated in the upper layer of his two-layer model drive the mean flows in the lower layer.

Oceanographers have examined local energy balances. The best known of these studies, by Webster (1961a), has often been interpreted as an indication that the Gulf Stream is accelerated by the eddies. However, Schmitz and Niiler (1969) have pointed out that the cross stream-averaged value of $\overline{u'v'}\overline{v}_x$ is not distinguishable from zero, so that Webster's results may be an indication merely of transfer of energy from the offshore to the onshore side of the jet and not a mean-flow generation. Even this result is not unambiguous, since the divergence term $(\overline{u'v'}\overline{v})_x$ is not small compared to the terms $-(\overline{u'v'})_x\overline{v}$ and $\overline{u'v'}\overline{v}_x$, representing, respectively, the eddy-mean flow and the mean flow-eddy conversions. The same problem occurs in attempts to compute regional energy budgets in numerical models (cf. Harrison and Robinson, 1978) unless considerable care is taken.

The conversion of mean-flow potential energy is sometimes inferred from the tilting of the phase lines of the temperature wave with height. This tilt is taken as evidence that the wave is growing by baroclinic instability of the mean flow. Here one must be careful: it is the lagging of the temperature wave behind the pressure wave and the consequent tilting of the phase line of the pressure trough toward the cold air that is important; the temperature phase line may tilt in any direction or not at all. Thus in the Eady model of baroclinic instability the temperature wave has the opposite slope from the pressure wave, whereas in the Charney model it has the opposite slope at low levels and the same slope at high levels (Charney, 1973, chapter IX). It is appropriate, then, to caution that the oceanically most readily available quantity, the phase change of the buoyancy (or entropy) with height, may not lead to a straightforward determination of the sign of the buoyancy flux.

This discussion should make it clear that very little is settled concerning the source of the eddies in the ocean and their effects on the mean flow. "Local" generation mechanisms, such as baroclinic or barotropic instability, flow over topography, and wind forcing are still being considered, and atmospheric analogues are much in mind. But the fact that the transient atmospheric perturbation velocities are comparable to those of the mean flow, whereas the particle speeds for mid-ocean mesoscale eddies are an order of magnitude greater than the mean speeds, suggests that energy may be generated only in limited regions (e.g., the western boundary currents) and propagate from there to other regions.

Oceanic eddies propagate in much the same manner as atmospheric eddies, although there are differences because of the upper-surface boundary conditions: atmospheric waves may propagate upward without reflection, whereas oceanic waves are reflected at the upper boundary. In the atmosphere, the potential vorticity gradients associated with the mean flow play an important role in determining the vertical structure and horizontal propagation for atmospheric waves, whereas this role is played primarily by the gradient of the earth's vorticity for mid-oceanic waves.

The interaction mechanisms may be classified as wave-mean flow interactions and wave-wave interactions. As mentioned above, the concrete evidence for significant wave-mean flow interaction is much greater for the atmosphere than for the ocean. There is not, of course, much oceanic data—Reynolds stresses have been calculated only along a few north-south sections (Schmitz, 1977). Moreover, these records are not very long and the spatial resolution is not sufficient to compute accurate gradients of the Reynolds stresses (given the great inhomogeneity).

The wave-wave interactions, however, seem similar in the two media. The crucial parameter is the Rossby wave steepness parameter $M = U/\beta L^2$, which distinguishes wavelike regimes ($M < 1$) from more turbulent ($M > 1$) regimes as illustrated in the experiments of Rhines (1975).[2] The oceans are similar to the atmosphere in that this parameter is of order unity for both, although it appears to vary considerably from one oceanic region to another.

The physical mechanisms for dissipation of atmospheric and oceanic eddies are thought to be similar with respect to bottom friction and transfer of energy to gravity wave motions or turbulence (though radiation is a further factor in damping atmosphere waves), but their relative importance may be quite different. The crucial differences for the large-scale circulations between the atmosphere and the ocean may not be in the details of the dissipation mechanisms but rather in their overall time scales. In the atmosphere, damping times are of the order of a few days, comparable to the eddy velocity advection time L/U, whereas the damping time in the ocean may be as long as several years (cf. Cheney and Richardson, 1976) while the advection time is of the order of a week.

18.4 The Geostrophic Formalism

18.4.1 The Development of the Geostrophic Formalism

The discovery that the atmospheric winds are approximately geostrophic is usually attributed to Buys Ballot (1857). Ferrel (1856) suggested that ocean currents might also have this property. But it took nearly a century before this knowledge was used dynamically. Because the geostrophic and hydrostatic equations express only a condition of balance, it is necessary to consider the slight imbalances produced by forcing, dissipation, and transience in order to predict the evolution and to understand the processes that maintain the balance. One of the first to exploit geostrophy was Bjerknes (1937) in a seminal work on the upper tropospheric long waves and their role in cyclogenesis. Basing his analysis on semiempirical considerations of the gradient wind and the variation with latitude of the Coriolis parameter, he gave the first explanation of the eastward propagation of the upper wave at a speed slower than the mean wind. It was this work that led Rossby et al. (1939) to their vorticity analysis of the upper wave as an independent entity in planar flow. Charney (1947) and Eady (1949) derived quasigeostrophic equations in their analyses of baroclinic instability for long atmospheric waves. General derivations of these equations for arbitrary motions were presented by Charney (1948), Eliassen (1949), Obukhov (1949), and Burger (1958). A particularly simple form which will be used in this review was given by Charney (1962) and Charney and Stern (1962).

In addition to these commonly used approximations, there have been a number of simplifications of the equations of motion which apply the concept of near-geostrophic balance in a less restrictive form. When flows become nongeostrophic in one horizontal dimension while remaining geostrophic in the other, as in frontogenesis, flow over two-dimensional mountain barriers, and in the western boundary currents of the oceans, a set of "semigeostrophic" equations derived from Eliassen's original formulation has often been found useful (Robinson and Niiler, 1967; Hoskins, 1975). Both the quasi- and semigeostrophic equations are special cases of the "balance equations" proposed by Bolin (1955), Charney (1955c, 1962), P. Thompson (1956), and Lorenz (1960). They may be derived from the consideration that in a large class of atmospheric flows the constraints of the earth's rotation and/or gravitational stability so inhibit vertical motion that the horizontal flow, even when it is not quasi-geo-

strophic, remains quasi-nondivergent. The equations derived by Eliassen (1952) for slow thermally and frictionally driven circulations in a circular vortex are a special case of the balance equations; they represent the laws of conservation of angular momentum and entropy and the requirement of equilibrium among the meridional components of the pressure, gravity, and centrifugal forces. For the equilibrium condition to be valid, the flow must be gravitationally and inertially stable. This implies that the potential vorticity must be positive in the northern hemisphere and negative in the southern hemisphere, and it may be shown that this condition on the potential vorticity is also required for the general asymmetric case.

One must also explain why external sources of energy excite quasi-geostrophic flows rather than gravity wave motions to begin with and why so little energy is transferred by nonlinear interactions into the gravity modes afterward. The tendency toward geostrophy is sometimes explained as an adjustment of an initially unbalanced flow by radiation of gravity waves in the manner discussed by Rossby (1938) (see also Blumen, 1968). However, since much of the forcing is applied slowly, rather than impulsively, the calculations of Veronis and Stommel (1956), who consider the nature of the exciting forces, are perhaps more relevant. They showed that the flows will be geostrophically balanced when the forcing period is very large compared to the inertial period. Thus we expect most of the energy will go into geostrophic motions.

The question of how much transfer occurs from geostrophic to nongeostrophic motions through nonlinear interactions remains a matter of concern. Errico's (1979) work suggests that equipartition of energy between gravity waves and geostrophic motions will occur in a conservative, rotating system in statistical equilibrium at sufficiently high energy. But in dissipative systems resembling the atmosphere and oceans, the energy will remain in the geostrophic modes because the gravity waves are dissipated on time scales that are small in comparison with those of their generation. This problem has elements in common with the so-called initialization problem in numerical weather prediction: to find initial values of a flow field that are at once compatible with the incomplete data and at the same time minimize the initial gravity-wave energy and its production rate (cf. Machenhauer, 1977; Daley, 1978).

The problems of transfer from geostrophic into gravity-wave energy are related to those of the production of hydrodynamic noise by a turbulent flow, first studied by Lighthill (1952). An excellent review is presented by Ffowcs Williams (1969). Here the problem is to calculate the generation of acoustic energy in a turbulent flow in which most of the energy resides in nondivergent motions. Since the turbulence is confined within a limited domain and radiates sound waves into the surrounding medium, there is no possibility of equipartition. An atmospheric or oceanographic analogy to the Lighthill problem would be the generation of internal gravity waves by turbulence in a planetary boundary layer (Townsend, 1965), except that here the generation takes place, not within the layer, but at its interface with the neighboring stable stratum.

Physicists, too, have struggled with problems in which many scales interact simultaneously (cf. Wilson, 1979), but it is not known whether their renormalization group methods can be usefully applied to atmospheric or oceanic problems.

18.4.2 Natural Oscillations of the Atmosphere and Oceans

The quasi-geostrophic equations have been derived for ranges of the various nondimensional parameters that are of interest in dealing with particular classes of atmospheric and oceanic motions. It is not to be expected that they will remain uniformly valid throughout the entire range of rotationally dominated flows, even when the primary balance is geostrophic. Burger (1958) was the first to point out explicitly that when the β-plane approximation $L/a \ll 1$, where L is the characteristic horizontal scale and a is the radius of the earth, is no longer valid, the dynamics of the motion change radically. On a planetary scale the motion becomes even more strongly geostrophic, but the vorticity balance changes. Sverdrup (1947) made implicit use of this dynamics in his treatment of the steady, wind-driven circulation of the oceans, and it has been used for the treatment of steady thermohaline circulations of the oceans by Robinson and Stommel (1959) and Welander (1959). [See the review articles by Veronis (1969, 1973b), and chapter 5.]

In this section we present a classification of natural oscillations in atmospheres and oceans in which rotation plays a dominant role, paying special attention to the domain of validity of the β-plane, quasi-geostrophic equations, the nature of the oscillations for which these equations are not valid, and the effects of nonlinearity, which, especially in the oceans, may give rise to solitary wave behavior.

Although the wave forcing is very different in the oceans and the atmosphere, there are many features of the responses that strongly resemble one another. This results because the response of a forced system depends strongly on the characteristics of the natural oscillations, and these have many similarities in the atmosphere and oceans. We shall discuss both the linear and nonlinear natural (unforced and nondissipative) oscillations of a simple model consisting of a single-layer, homogeneous, incompressible fluid with a free surface on a β-plane. This is a much oversimplified model, and we must regard the conclusions to be drawn merely as

suggestions of the way in which the fully stratified, spherical system would behave. The most interesting implication—that the dynamics of scales intermediate between the Rossby radius and the radius of the earth may be dominated by solitary waves, in which nonlinear density advection balances linear dispersive effects—may not be very sensitive to the particular model chosen.

At the beginning of each subsection to follow we shall describe briefly the methods used and the results obtained in order to make it possible for the reader to omit the more detailed derivations. We begin the discussion of the normal modes of oscillation by stating the shallow-water equations in a reference frame moving with the wave. The Bernoulli and potential vorticity integrals then give two equations relating the wave streamfunction ϕ to the surface elevation η above the mean level H. These equations contain an unknown functional $\mathcal{B}$, the Bernoulli function, which we choose by requiring the equations to hold for vanishing ϕ and η.

We then obtain two coupled nonlinear partial differential equations defining an eigenvalue problem for the phase speed c. These equations have three nondimensional parameters: ϵ, a Rossby number measuring the ratio of the inertial to the Coriolis forces; $\hat{S}$, a static stability parameter measuring the ratio of the deformation scale L_R to the wave scale L; and $\hat{\beta}$, the fractional change in the Coriolis parameter over the wave scale. In the standard quasi-geostrophic range ($\epsilon \sim \hat{\beta} \ll 1$, $\hat{S} \sim 1$), when motions have small Rossby numbers, length scales comparable to the deformation radius, the Bernoulli equation to lowest order is simply a statement of geostrophic balance, and the potential vorticity equation becomes the linear quasi-geostrophic wave equation.

The single-layer equations will be written in dimensional form

$$\begin{aligned} &\frac{Du}{Dt} - (f_0 + \beta y)v = -g\eta_x, \\ &\frac{Dv}{Dt} + (f_0 + \beta y)u = -g\eta_y, \\ &\frac{D\eta}{Dt} + (H + \eta)(u_x + v_y) = 0, \\ &\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}, \end{aligned} \tag{18.1}$$

where η is the displacement of the surface from mean sea level, H is the mean depth of the fluid, g is the gravitational acceleration (we shall use a reduced gravity value here), $f_0 = 2\Omega \sin\Theta$, $\beta = 2\Omega \cos\Theta/a$, $y = a\,\Delta\Theta$, where Θ is the central latitude and $\Delta\Theta$ is the angular distance from Θ. In nondimensional form these equations become

$$\begin{aligned} &\hat{\beta}\hat{S}\frac{Du}{Dt} - (1 + \hat{\beta}y)v = -\eta_x, \\ &\hat{\beta}\hat{S}\frac{Dv}{Dt} + (1 + \hat{\beta}y)u = -\eta_y, \\ &\hat{\beta}\frac{D\eta}{Dt} + \left(1 + \frac{\epsilon}{\hat{S}}\eta\right)(u_x + v_y) = 0, \\ &\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{\epsilon}{\hat{\beta}\hat{S}}\left(u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}\right), \end{aligned} \tag{18.2}$$

where both x and y are scaled by L, η is scaled geostrophically by LUf_0/g, and the nondimensional parameters are $\hat{\beta} = \beta L/f_0 = L\cot\Theta/a \equiv L/L_\beta$, $\epsilon = U/f_0L$, and $\hat{S} = gH/f_0^2L^2 \equiv L_R^2/L^2$. We have also introduced the definitions of two scales which turn out to be important in determining the boundaries between various types of behavior: the β-scale $L_\beta = f_0/\beta = a\tan\Theta$, at which variations in the vertical component of the earth's angular velocity are order of the angular velocity itself, and the Rossby radius of deformation $L_R = \sqrt{gH}/f_0$ (Rossby, 1938). We shall use 3500 km for L_β (corresponding to $\Theta \sim 30°$) and 50 km (oceanic) or 1000 km (atmospheric) for L_R. We have also made the choice of the long-wave period for the time scale so that $T = L/\beta L_R^2$.

The quasi-geostrophic potential vorticity equation may be derived by expanding (18.2) in powers of ϵ for $\hat{\beta} \sim \epsilon \ll 1$ and $\hat{S} \sim 1$, giving

$$\left[\frac{\partial}{\partial t} + \frac{\epsilon}{\hat{\beta}\hat{S}}\left(\eta_x\frac{\partial}{\partial y} - \eta_y\frac{\partial}{\partial x}\right)\right]\left(\nabla^2\eta - \frac{1}{\hat{S}}\eta + \hat{\beta}y/\epsilon\right) = 0. \tag{18.3}$$

The failure of this equation at small space or time scales and near the equator is well known. Meterorologists since Burger (1958) have also recognized that some larger than synoptic-scale motions also do not evolve according to this equation. Rather, the appropriate equations are derived by assuming $\hat{S}$ and ϵ to be small (because L is very large) and $\hat{\beta}$ to be of order 1. The resulting velocities remain geostrophic:

$$u = -\frac{1}{1 + \hat{\beta}y}\eta_y, \qquad v = \frac{1}{1 + \hat{\beta}y}\eta_x, \tag{18.4}$$

and the height field evolves according to

$$\frac{\partial}{\partial t}\eta - \frac{1}{(1 + \hat{\beta}y)^2}\left(1 + \frac{\epsilon}{\hat{S}}\eta\right)\eta_x = 0. \tag{18.5}$$

Equations (18.3) and (18.5) have very different properties: the quasi-geostrophic equation has uniformly propagating linear-wave solutions which are essentially dispersive even at large amplitudes (cf. McWilliams and Flierl, 1979), while the Burger equation does not have uniformly propagating linear-wave solutions and initial disturbances steepen because of non-

linearity. We shall demonstrate that there is an intermediate band of length scales in which nonlinearity and dispersion can balance to give cnoidal or solitary waves. In the ocean, as we shall see, the change from quasi-geostrophic to intermediate dynamics to Burger dynamics occurs at a relatively small scale because the deformation radius is so small compared to the radius of the earth.

We may elucidate these differences by considering the shallow-water equations under the assumption that the motions are translating steadily at speed c:

$$\begin{aligned}(\mathbf{v} - c\hat{\mathbf{x}})\cdot\nabla v + \hat{\mathbf{z}}(f_0 + \beta y) \times \mathbf{v} &= -g\nabla\eta,\\ (\mathbf{v} - c\hat{\mathbf{x}})\cdot\nabla\eta + (H + \eta)\nabla\cdot\mathbf{v} &= 0,\end{aligned} \tag{18.6}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are unit vectors in the positive x and z directions. We may define a transport streamfunction in the coordinate system moving with the wave

$$(\mathbf{v} - c\hat{\mathbf{x}})(H + \eta) = \hat{\mathbf{z}} \times \nabla\psi$$

and write the Bernoulli and potential vorticity integrals of motion:

$$\begin{aligned}&\frac{1}{2}|\nabla(\phi + cHy)|^2 + g(H + \eta)^3 + c(H + \eta)^2\left(f_0 y + \frac{\beta y}{2}\right)^2\\ &= (H + \eta)^2\mathcal{B}(\phi + cHy),\end{aligned} \tag{18.7}$$

$$\nabla\cdot\frac{\nabla(\phi + cHy)}{H + \eta} + (f_0 + \beta y) = (H + \eta)\mathcal{B}'(\phi + cHy),$$

where we have isolated the wave part of the streamfunction $\phi = \psi - cHy$. We require that (18.7) hold as $\phi, \eta \to 0$; this determines the Bernoulli functional

$$\mathcal{B}(Z) = \frac{c^2}{2} + gH + \frac{f_0}{H}Z + \frac{\beta}{2cH^2}Z^2. \tag{18.8}$$

The choice of a single-valued, well-behaved Bernoulli functional implies that only motions which reduce smoothly to linear waves will be considered; thus the solutions of Stern (1975b) or Flierl, Larichev, McWilliams, and Reznik, (1980) which involve closed streamlines and a multiple-valued $\mathcal{B}$ will not be examined here.

In nondimensional form, equations (18.7) become

$$\begin{aligned}&\frac{\epsilon}{2}|\nabla\phi|^2 + \hat{\beta}\hat{S}c\phi_y + \underline{\eta\left(1 + \frac{\epsilon}{\hat{S}}\eta\right)^2}\\ &= \hat{\beta}^2\hat{S}c^2\eta\left(1 + \frac{\epsilon}{2\hat{S}}\eta\right) + \underline{\phi(1 + \hat{\beta}y)\left(1 + \frac{\epsilon}{\hat{S}}\eta\right)^2}\\ &\quad + \frac{\epsilon}{\hat{S}}\frac{\phi^2}{2c}\left(1 + \frac{\epsilon}{\hat{S}}\eta\right)^2,\end{aligned} \tag{18.9}$$

$$\begin{aligned}&\underline{\hat{S}\nabla^2\phi\left(1 + \frac{\epsilon}{\hat{S}}\eta\right)} - \epsilon\nabla\phi\cdot\nabla\eta - \hat{\beta}\hat{S}c\eta_y\\ &= \underline{\eta(1 + \hat{\beta}y)} \times \left(1 + \frac{\epsilon}{\hat{S}}\eta\right)^2 + \underline{\frac{\phi}{c}\left(1 + \frac{\epsilon}{\hat{S}}\eta\right)^3}.\end{aligned} \tag{18.10}$$

We show in figure 18.1 the dependence upon L and U of our three basic parameters $\hat{\beta} = L/L_\beta = \beta L/f_0 = L\cot\Theta/a$ (the ratio of the wave scale to the radius of the earth scale), $\epsilon = U/f_0 L$ (the Rossby number), and $\hat{S} = L_R^2/L^2$ (the inverse of the rotational Froude number). Note immediately the differences in scale separation for oceanic versus atmospheric conditions. In the atmosphere L_β is very close to L_R, so that there is only a short range between the usual baroclinic Rossby wave scales ($\hat{S} \sim 1$, $\beta \ll 1$) and the Burger range ($\hat{S} \ll 1$, $\hat{\beta} \sim 1$); in the ocean there is a large scale gap. Thus one might expect the different dynamics to be seen more clearly in the ocean.

Linear Waves ($\epsilon = 0$) The first step toward understanding the various types of large-scale free motion is to consider the linearized solutions. When the Rossby number is very small, the two equations can be combined into a single streamfunction equation which governs both gravity and Rossby waves. The Rossby wave-phase speed increases as the length scales of the wave increase, leveling off for $L > L_R$. For still larger scales, however, the speed again increases as the wave amplitude begins to be more pronounced equatorially. We demonstrate that the natural dividing scale here is what we call the "intermediate" scale $L_I \equiv (L_\beta L_R^2)^{1/3}$, where $\hat{S} = \hat{\beta}$ (see figure 18.1). This is the scale at which

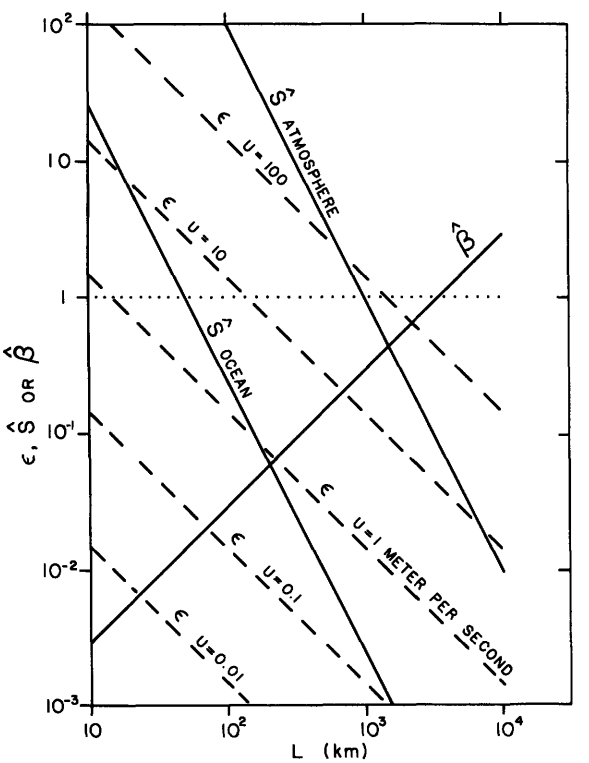


Figure 18.1 Values of the Rossby number ϵ, inverse Froude number $\hat{S}$, and beta parameter $\hat{\beta}$ as functions of the length scale L and velocity scale U for oceanic and atmospheric values of the deformation radius.

the relative vorticity changes become as small as the variations in vortex stretching due to the β term. Alternatively, one could say that the rule, "f equals a constant except when differentiated," breaks down near the intermediate scale. The phase speed continues to increase and, for large enough north-south scales, the wave domain crosses the equator. Then the wave becomes equatorially trapped and the phase speed again becomes independent of L.

For the parameters we have chosen—L_β = 3500 km, L_R = 50 km (ocean), 1000 km (atmosphere)—the intermediate scale L_I = 210 km (ocean), 1500 km (atmosphere) is not very large. It represents the upper bound to the scales for which the standard quasigeostrophic equations are valid. It may again be seen that there is a significantly greater separation among the various scales in the ocean compared to the atmosphere. This suggests that the ocean mesoscale motions may be a cleaner example of quasi-geostrophic flow than the synoptic-scale motions of the atmosphere; the approximations used for the latter are less exact.

For linear motions, the Bernoulli equation (18.9) defines η in terms of ϕ; η may then be eliminated from the potential vorticity equation (18.10) to yield a single equation for the streamfunction

$$\hat{S}\nabla^2\phi - \frac{1}{c}\phi - (1+\hat{\beta}y)^2\phi = \hat{\beta}^2\hat{S}^2c^2\phi_{xx}. \tag{18.11}$$

Since the coefficients do not involve x, we may set $\phi = e^{ix}G(y;L)$. The resulting equation together with boundary conditions presents an eigenvalue problem for $c(L)$ and the wave structure $G(y;L)$.

It has three eigenvalues, corresponding to two gravity-wave modes and one Rossby-wave mode. We can identify the gravity modes with the retention of the right-hand term in (18.11). For mid-latitude modes this term is significant only when $\hat{\beta}^2\hat{S}^2c^2 \sim \hat{S}$ or c^2 (dimensional) $\sim gH$; it is small for the Rossby mode solutions. For equatorially trapped Rossby modes, the y scale contracts so that $\hat{S}\phi_{yy}$ dominates both $\hat{S}\phi_{xx}$ and $\hat{\beta}^2\hat{S}^2c^2\phi_{xx}$. Eliminating the right-hand side corresponds to retaining only the underlined terms in (18.9)-(18.10). The filtered linear equation becomes

$$\hat{S}\nabla^2\phi - \frac{1}{c}\phi = (1+\hat{\beta}y)^2\phi, \tag{18.12}$$

which has been discussed extensively by Lindzen (1967) and others. Here we comment on the various types of solution primarily as a guide to our later discussion of the effects of nonlinearity.

Figure 18.2 shows the nondimensional phase speed as a function of L for atmospheric or oceanic parameters under the simplifying boundary conditions $\phi = 0$ at $y = \pm\pi/2$, which make the x and y scales of the domain similar. We can identify four different types of behavior.

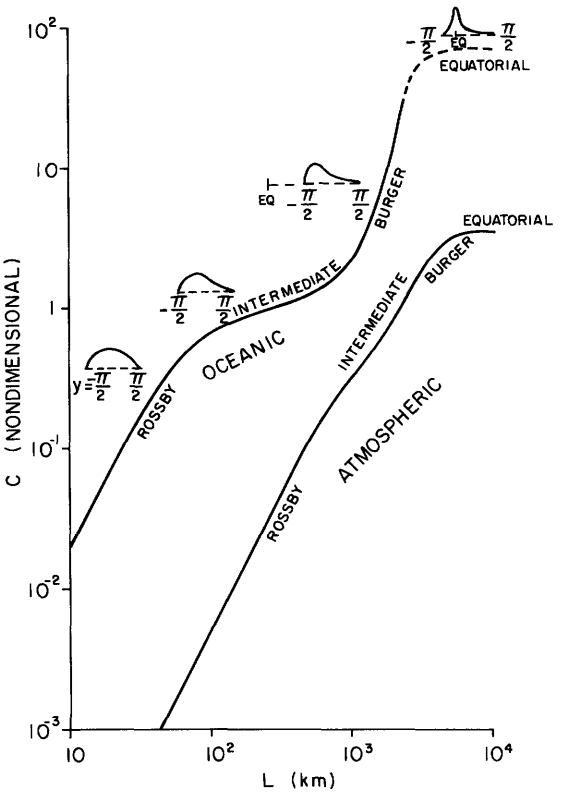


Figure 18.2 Phase speed nondimensionalized by βL_R^2 as a function of the x scale L (wavelength/2π) in a channel of width πL. Also shown are typical shapes of the y structure function $G(y;L)$ for the various classes of motion.

Midlatitude Rossby waves ($\hat{\beta} \ll 1$, $\hat{S} \sim 1$): For these motions, first described by Rossby et al. (1939), the streamfunction satisfies

$$\hat{S}\nabla^2\phi - \frac{1}{c}\phi = \phi, \tag{18.13}$$

which has solutions in the box

$$\phi = e^{ix}\cos y$$

with

$$c = -1/(1+2\hat{S}), \tag{18.14}$$

or, more generally, for waves oriented in any direction, we have

$$\phi = e^{i\mathbf{k}\cdot\mathbf{x}}$$

with

$$c = -1/(1+\hat{S}\,\mathbf{k}\cdot\mathbf{k}) \tag{18.15}$$

(see the discussion in chapter 10).

Intermediate scale waves ($\hat{\beta} \sim \hat{S} \ll 1$): The Rossby-wave dispersion relation (18.14) remains valid for $\hat{\beta} \ll \hat{S} \ll 1$ and becomes

$$c = -1 + 2\hat{S},$$

so that for a sufficiently small $\hat{S}$ the waves are nondispersive c (dimensional) $= -\beta L_R^2$. However, when L increases to the point where $\hat{\beta} \sim \hat{S} \ll 1$, the small correction in the formula above becomes invalid. This occurs when the $\phi\hat{\beta}y$ term becomes comparable to the $\hat{S}\nabla^2\phi$ term, that is, when $L \simeq (L_\beta L_R^2)^{1/3}$, which is 210 km for the oceans or 1500 km for the atmosphere. We denote this scale as the "intermediate scale" L_I. The wave structure is determined by expanding (18.12) in $\hat{S}$ (or $\hat{\beta}$). Setting $\phi = \phi^{(0)} + \hat{S}\phi^{(1)} + \cdots$ and $c = -1 + \hat{S}c^{(1)} + \cdots$, we obtain

$$(\nabla^2 + c^{(1)} - 2\hat{\beta}y/\hat{S})\phi^{(0)} = 0. \tag{18.16}$$

When $L \gtrsim L_I$ the y dependence of f can no longer be neglected, the y scale becomes order of the intermediate scale, and the solutions begin to be concentrated toward the equator (see figure 18.2). As L continues to increase, $\hat{S}$ decreases but $\hat{\beta}$ increases and the phase speed is no longer insensitive to L but begins to increase; c behaves like $-1 + O(\hat{\beta})$ rather than $-1 + O(\hat{S})$. The phase speed becomes less and less sensitive to the x wavenumber, so that the waves may still be considered approximately nondispersive. We have required that $\hat{\beta}$ and $\hat{S}$ be small, but figure 18.1 shows that these quantities are small only for a rather narrow range of L even in the oceanic case, and figure 18.2 shows that c varies perceptibly with L everywhere. For the atmospheric parameters a totally nondispersive regime ($\hat{\beta} \ll \hat{S} \ll 1$) does not exist at all.

Burger motions ($\hat{\beta} \sim 1$, $\hat{S} \ll 1$): When L increases to the point where $\hat{\beta} \sim 1$, the motions become strongly concentrated near the equator. The y scale contracts (relative to L) so that the lowest order balance includes all the terms in (18.12) and the y wave domain crosses the equator. The phase speed rapidly increases from that of the midlatitude Rossby waves to that of the equatorial waves.

Now we can see why the Burger equation (18.5), which assumes equal x and y scales, has no linear free-wave solutions: free waves with a very large x scale do not have the same y scale. Instead the unforced motions acquire a meridional scale between L_I and the (somewhat larger) equatorial scale. Forced motions, of course, may have comparable x and y scales and may therefore have evolution equations in which the terms of (18.5) contribute along with the forcing terms.

Equatorial waves: Here we can drop the 1 in the $1 + \hat{\beta}y$ term of equation (18.12) to change to the equatorial β-plane (the f_0 factors will all cancel out upon dimensionalization). The solutions are well known (cf. Lindzen, 1967) and again become nondispersive for small $\hat{S}$. Rescaling the equation for small $\hat{S}$ shows that the y wave domain is confined to a region around the equator of meridional extent $\hat{\beta}^{-1/2}\hat{S}^{1/4}$, which corresponds to the dimensional scale $L_e = (gH/\beta^2)^{1/4} = (L_\beta L_R)^{1/2}$, the well-known equatorial deformation scale. For our assumed parameters, this scale is 420 km for the oceans and 1900 km for the atmosphere; however, this estimate is not very accurate since the equivalent depth for baroclinic motions varies considerably. Moore and Philander (1977) give 325 km as an estimate of this scale for the first baroclinic mode. The phase speeds are order $\hat{\beta}^{-1}\hat{S}^{-1/2} = L_\beta/L_R$, corresponding to a dimensional speed $\beta L_e^2 = \sqrt{gH}$. (The other solutions have $c \sim \pm\hat{\beta}^{-1/2}\hat{S}^{-3/4}$.) A further discussion appears in chapter 6.

Nonlinear Waves ($\epsilon > 0$) When the motion becomes of sufficiently large amplitude, the propagation characteristics of a single wave change. We shall investigate the size of the Rossby number necessary for this to occur. This size may be quite different from the Rossby number required for significant nonlinear interactions in a full spectrum of waves. However, the nonlinear behavior of a single wave can be of interest when it allows the possibility for solitary waves. On the scale of the mid-latitude Rossby wave, this does not appear to occur and the nonlinearity gives only a correction to the phase speed and shape; the lowest-order balance remains strongly dispersive. However, as the scale becomes equal to or greater than the intermediate scale, the phase speed becomes less dependent on the x wavenumber. When the Rossby number becomes of the order L_R^4/L^4, the nonlinear advection term becomes comparable to the east-west dispersion term and the solutions propagate as solitary waves. The structure of these isolated high-pressure disturbances is found to be the same as that of the $\mathrm{sech}^2 x$ solution to the Korteweg-deVries equations. The implication of this section, then, is that the dynamics of motions of the intermediate or large scales may be quite different from that of the ordinary Rossby wave.

Let us now consider the conditions under which the nonlinear terms can alter the propagation characteristics of the free waves in our model. This can occur whenever one of the ϵ terms is comparable to one of the linear terms that have been retained in the Bernoulli or potential vorticity equations (18.9)–(18.10) (the underlined terms). This happens when $\epsilon \sim 1$, $\epsilon/\hat{\beta} \sim 1$, $\epsilon/\hat{S} \sim 1$, $\epsilon/\hat{\beta}\hat{S} \sim 1$, or $\epsilon/\hat{S}^2 \sim 1$. The velocities required for each of these conditions are shown in figure 18.3, which emphasizes again the relative complexity of the atmosphere: for 40-ms^{-1} winds at 1000-km scales, all of the nonlinear terms enter simultaneously. For the ocean, only strong meandering motions could cause each of ϵ, $\epsilon/\hat{S}$ and $\epsilon/\hat{S}^2$ to be of order unity, and in these circumstances $\hat{\beta}$ remains quite small. We shall not attempt to deal with these more complicated motions, but instead shall examine the nonlinear ef-

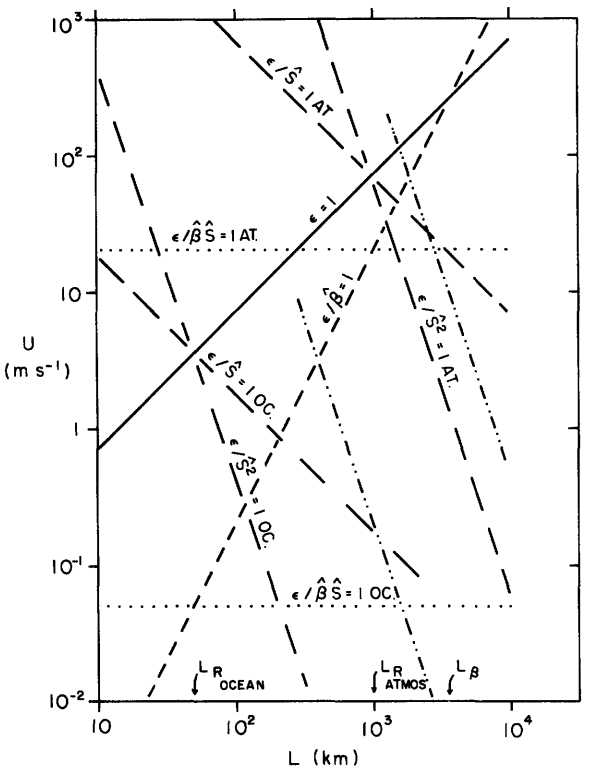


Figure 18.3 Conditions under which nonlinear terms become important. Labeled curves show relationship between U and L such that a particular parameter ratio becomes equal to one. This corresponds to one of the nonlinear terms in (18.9)-(18.10) becoming equal in magnitude to one of the underlined linear terms.

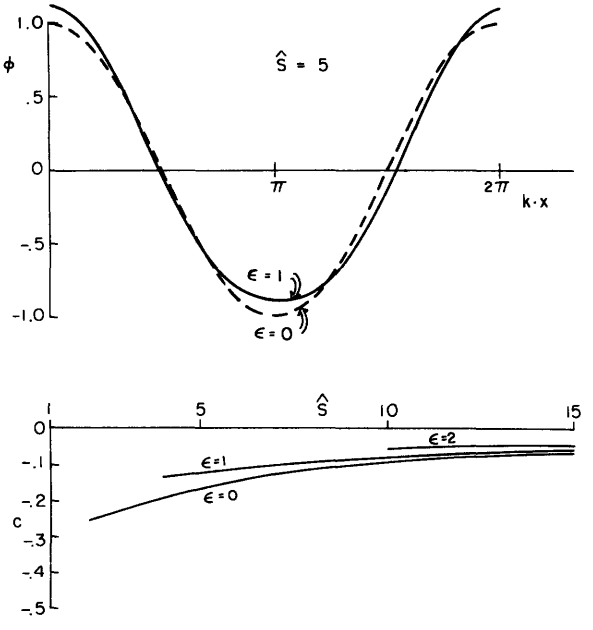


Figure 18.4 Effects of nonlinearity on a short Rossby wave. The upper figure shows the changes in the shape of the wave. The lower figure shows the changes in the dispersion relation.

fects on each of the waves that has been considered above.

Midlatitude Rossby waves: The first nonlinear condition that occurs when $\hat{S} \gtrsim 1$ is $\epsilon = \hat{\beta}$. However, since $\hat{\beta}$ remains small and does not enter the governing equation (18.13), we expect that this will not significantly alter the behavior of a single steadily propagating sinusoidal wave. When ϵ or $\epsilon/\hat{S}$ becomes order 1, nonlinearity begins to affect the structure significantly. For example, consider the parameter range $\epsilon \sim 1$, $\hat{\beta} \ll \hat{S}^{-1} \ll 1$. To lowest order in an expansion of both ϕ and c in $\hat{S}^{-1}$ (c being of order $\hat{S}^{-1}$), the potential vorticity equation gives

$$\nabla^2\phi^{(0)} = \frac{1}{c^{(0)}\hat{S}}\,\phi^{(0)}.$$

At first order we find the corrections to the phase speed and shape of the wave. The result is

$$c \simeq \hat{S}^{-1}\left[-\frac{1}{\mathbf{k}\cdot\mathbf{k}} + \hat{S}^{-1}\left(\epsilon^2 + \frac{1}{(\mathbf{k}\cdot\mathbf{k})^2}\right) + \cdots\right], \qquad (18.17)$$

as sketched in figure 18.4. The order ϵ nonlinear terms cause a sharpening of the streamfunction crests and a decrease in the propagation rate.

Intermediate scale waves: When $\hat{S} \ll 1$, nonlinear terms first enter when $\epsilon \sim \hat{\beta}\hat{S}$ or $\epsilon \sim \hat{S}^2$ (see figure 18.3). We can find the forms of the solutions by letting $\epsilon = E\hat{S}^2$ and $\hat{\beta} = B\hat{S}$ and expanding for small $\hat{S}$ assuming E, B to be of order unity or less. We get

$$\begin{aligned} &c = -1 + \hat{S}c^{(1)}, \\ &\nabla^2\phi^{(0)} + c^{(1)}\phi^{(0)} + \tfrac{3}{2}E[\phi^{(0)}]^2 - 2By\phi^{(0)} = 0 \end{aligned} \qquad (18.18)$$

for the equations governing the shape and the speed of the wave.

The simple limit here is $B = \hat{\beta}/\hat{S} \ll 1$, corresponding to the range $L_R \ll L \ll L_I$, and E of order unity, corresponding to particle speeds given by the $\epsilon/\hat{S}^2 = 1$ lines in figure 18.3. The wave equation

$$\nabla^2\phi^{(0)} + c^{(1)}\phi^{(0)} + \tfrac{3}{2}E[\phi^{(0)}]^2 = 0$$

has both one- and two-dimensional solutions on the plane. These include the cnoidal and solitary wave solutions to the Korteweg-deVries equation (Whitham, 1974) for uniformly propagating waves:

$$\begin{aligned} &\phi^{(0)} = cn^2\left(\frac{K(m)}{\pi}\frac{\mathbf{k}\cdot\mathbf{x}}{|\mathbf{k}|}; m\right) - \frac{\sqrt{1-m+m^2}-1+2m}{3m}, \\ &c = -1 + \hat{S}4K^2(m)\sqrt{1-m+m^2}/\pi^2, \\ &\epsilon\hat{S}^{-2} = 4mK^2(m)/\pi^2, \end{aligned} \qquad (18.19a)$$

and

$$\phi^{(0)} = \operatorname{sech}^2 \mathbf{k}\cdot\mathbf{x},$$

$$c = -1 - 4\hat{S}\mathbf{k}\cdot\mathbf{k}, \tag{18.19b}$$

$$\epsilon\hat{S}^{-2} = 4\mathbf{k}\cdot\mathbf{k}.$$

Plots of the shapes of the cnoidal and solitary waves and the dispersion relations are shown in figures 18.5A and 18.5B. The cnoidal waves show a phase speed decreasing with amplitude (as in the example above) while the solitary wave speed increases as the wave gets stronger.[3]

A second type of solution (cf. Flierl, 1979b) is a radially symmetric solitary wave

$$\phi^{(0)} = G(k\sqrt{\mathbf{x}\cdot\mathbf{x}}),$$

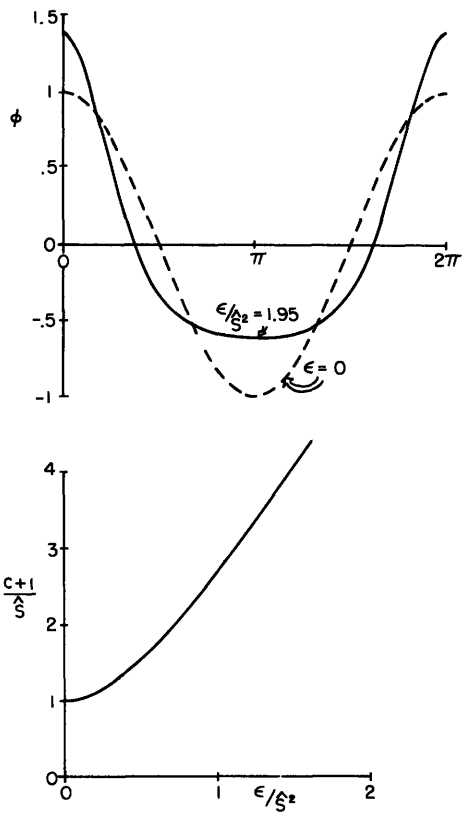
$$c = -1 - \hat{S}k^2, \tag{18.19c}$$

$$\epsilon\hat{S}^{-2} = 1.59k^2,$$

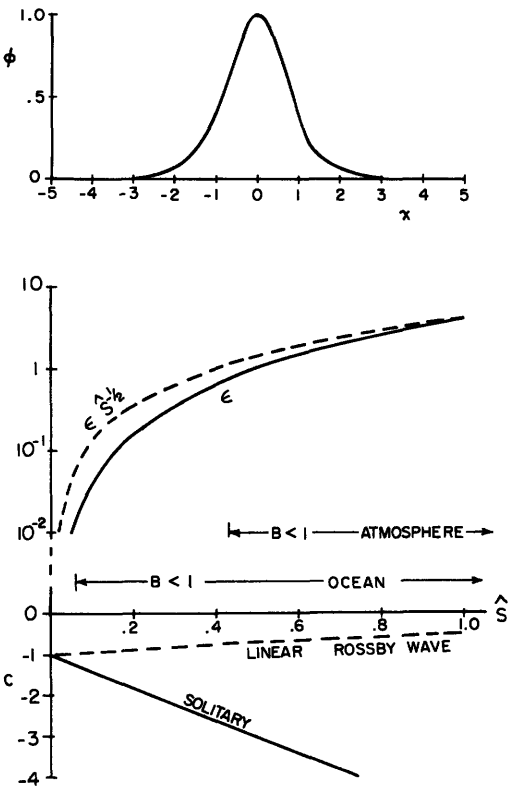
whose shape and dispersion relations are shown in figure 18.6.

It may be seen from equation (18.15) that the dynamics of large-scale motions for which $\epsilon \sim \hat{S}^2$ and $\hat{\beta} \ll \epsilon/\hat{S}$ are distinctly different from those of the quasigeostrophic eddies. We might expect, if the motions are governed by the Korteweg–deVries equation as suggested by (18.18), that solitons will be formed and dominate the subsequent evolution of the field. In the atmosphere, solitary-wave behavior would be difficult to find because of the rapid frictional decay time, the east–west periodicity for scales not so much larger than those under consideration, and the rather limited parameter range for the Korteweg–deVries regime. In the ocean, the situation is quite different; the parameter range for solitary-wave behavior is more distinct, the waves are of small scale compared to the size of the basin, and the decay rates are slow so that there is sufficient space and time for the necessary balance between nonlinearity and dispersion to develop.

For scales larger than the intermediate scale, B becomes large in (18.18). If y is rescaled by $B^{-1/3}$ (dimensionally by L_1), this equation can be solved by expansion in powers of $B^{-2/3}$. To lowest order, one obtains a linear equation for the y structure; to next order, the x dispersion and nonlinear steepening (if E is order unity) are included and the x structure is then given by an equation of the Korteweg–deVries type.



18.5(A)



18.5(B)

Figure 18.5 Effects of nonlinearity on long waves. (A) Cnoidal waves: the upper figure shows the change in shape occurring when the nonlinearity is increased while the lower figure shows the changes in the dispersion relation. (B) Solitary waves: the upper figure shows the shape of the wave while the lower figure shows the relationship between the length and amplitude ($\hat{S}$ and ϵ) and also the propagation speed. For a fixed deformation radius, $\epsilon\hat{S}^{-1/2}$ is directly proportional to the velocity scale U. The relationships are only valid for $B < 1$.

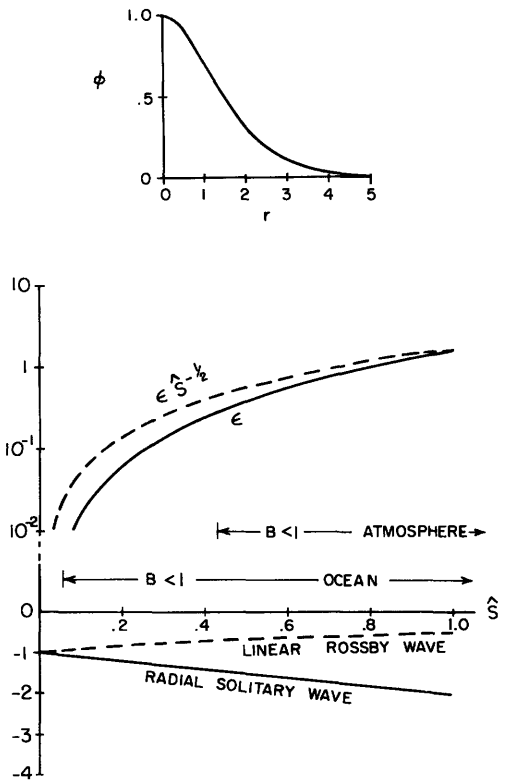


Figure 18.6 Radially symmetric solitary solutions. The upper figure shows the dependence of the pressure upon radius. The lower figure gives the relationships between amplitude, size, and propagation speed.

Burger range: Here also one can show that there are motions whose y structure is determined by a linear equation and whose x structure is determined by a nonlinear equation of the Korteweg–deVries type. We still require $\epsilon \sim \hat{S}^2$. Clarke (1971) has discussed this type of solution (and also those described above for large B) in more detail.

Equatorial motions: Boyd (1977) has shown that the long waves in this case also satisfy an equation of the Korteweg–deVries type. If we rescale the equatorial versions of (18.9) and (18.10), letting $y = \hat{\beta}^{-1/2}\hat{S}^{1/4}Y$ (so that Y has the scale L_e), $c = \hat{\beta}^{-1}\hat{S}^{-1/2}C$, and $\eta = \hat{\beta}^{1/2}\hat{S}^{1/4}N$, we can show that there are only two parameters (in the absence of north–south boundaries) of interest: $\delta = \hat{\beta}^{-1/2}\hat{S}^{1/4} = L_e/L$ and $\hat{\epsilon} = \epsilon\,\hat{\beta}^{1/2}\hat{S}^{-3/4} = UL/\beta L_e^3$. The cnoidal or solitary wave (in x) solutions are obtained when $\hat{\epsilon} \sim \delta^2 \ll 1$. This gives an equatorial velocity scale $U = \beta L_e^5/L^3$, as shown in figure 18.3.

In summary, then, we have seen three different types of natural large-scale, long-period motions in the atmosphere and ocean. For scales on the order of the deformation radius or less ($L \lesssim 50$ for the oceans and $\lesssim$ 1000 km for the atmosphere), dispersive Rossby waves dominate with nonlinear effects entering only for large Rossby number ϵ. Intermediate scales ($\hat{S} \ll 1$, $\epsilon\hat{S}^{-2} \sim 1$, $\hat{\beta} \ll \epsilon\hat{S}^{-1}$ implying $50 \ll L \ll 210$ km for the oceans and $1000 \ll L \ll 1500$ km for the atmosphere) have solitary or cnoidal wave structures as well as circular solitary highs. As the scales become larger, weak solitary or cnoidal wave structures may persist with normal-mode y shapes concentrated near the equatorward side of the domain. Stronger motions will not remain permanent but will steepen in amplitude, as do the solutions of Burger's equation (18.5). When the wave domain comes to include the equator, nonlinear equatorial wave motions satisfying a Korteweg–deVries type of equation can exist.

Korteweg–deVries Dynamics Finally we shall demonstrate that Korteweg–deVries dynamics does seem to be appropriate for general motions (not necessarily uniformly propagating waves) on the intermediate scale ($\hat{\beta} \gtrsim \hat{S}$, $\epsilon \sim \hat{S}^2$, and $\hat{S} \ll 1$). The previous derivations have shown only that the permanent form is governed by an equation that may be derived from the Korteweg–deVries equation, but it is still necessary to show that the time-dependent evolution equation is also of this type. We return to our governing equations (18.2) and set $\epsilon = E\hat{S}^2$ and $\hat{\beta} = B\hat{S}$, where B and E are assumed to be of order unity. This corresponds to $L \sim L_I$ and $U \sim f_0 L_R^2/L_\beta$ (210 km, 5 cm s^{-1} for the ocean; 1500 km, 20 m s^{-1} for the atmosphere). We note that there will be two time scales in the evolution: a fast time t corresponding to the nondispersive propagation and a slow time $T = \hat{S}t$ during which features evolve.

The lowest two orders of the expansion in $\hat{S}$ show that the flow is geostrophic and that the advection of planetary vorticity is balanced by vortex stretching, leading to the usual nondispersive propagation of very long Rossby waves. At the next order slow changes in surface height force a divergence which creates relative vorticity. The vorticity balance also is influenced by north–south variations in vortex stretching due to variations of f, while the nonlinear terms enter in the mass balance. The resulting equation is a mix between the Korteweg–deVries equation and the Rossby-wave equation. However, when L is large compared to the intermediate scale, the more detailed expansion to follow shows that the x structure indeed evolves according to a Korteweg–deVries equation.

At lowest order the flows are geostrophic

$$u^{(0)} = -\eta_y^{(0)},$$

$$v^{(0)} = \eta_x^{(0)},$$

$$u_x^{(0)} + v_y^{(0)} = 0.$$

The first-order equations,

$$u^{(1)} + Byu^{(0)} = -\eta_y^{(1)},$$

$$v^{(1)} + Byv^{(0)} = \eta_x^{(1)},$$

$$B\eta_t^{(0)} + E\mathbf{v}^{(0)}\cdot\nabla\eta^{(0)} + E\eta^{(0)}\nabla\cdot\mathbf{v}^{(0)} + \nabla\cdot\mathbf{v}^{(1)} = 0,$$

lead to Sverdrup (1947) or Burger (1958) type of balance between advection of planetary vorticity and vortex stretching,

$$u_x^{(1)} + v_y^{(1)} = -B\eta_x^{(0)},$$

and to the nondispersive wave equation

$$B\eta_t^{(0)} - B\eta_x^{(0)} = 0,$$

which implies

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial x} \qquad or \qquad \eta = \eta(x + t, y, T).$$

At second order we obtain the vorticity equation

$$B(v_x^{(0)} - u_y^{(0)}) + E\mathbf{v}^{(0)}\cdot\nabla(v_x^{(0)} - u_y^{(0)})$$
$$+ Bv^{(1)} + \nabla\cdot\mathbf{v}^{(2)} - By\,\nabla\cdot\mathbf{v}^{(1)} = 0$$

and the mass-conservation equation

$$B\eta_T^{(0)} + B\eta_t^{(1)} + E\mathbf{v}^{(1)}\cdot\nabla\eta^{(0)} + E\mathbf{v}^{(0)}\cdot\nabla\eta^{(1)}$$
$$+ \nabla\cdot\mathbf{v}^{(2)} + E\eta^{(1)}\nabla\cdot\mathbf{v}^{(0)} + E\eta^{(0)}\nabla\cdot\mathbf{v}^{(1)} = 0,$$

which jointly lead to the evolution equation [after using $\partial/\partial t = \partial/\partial x$ for the fast time, and dropping the superscript (0)]

$$B\eta_T = EB\eta\eta_x + B(\nabla^2\eta)_x$$
$$- 2B^2y\eta_x + EJ(\eta, \nabla^2\eta) \qquad (18.20)$$

[where $J(A,B)$ is the Jacobian operator] or

$$B\eta_T = EJ(\eta - \frac{B}{E}y, \nabla^2\eta + \frac{3}{2}E\eta^2 - 2By\eta).$$

One can readily show that the requirement of steady propagation leads to (18.18). Furthermore, when L is large compared to the intermediate scale L_I but E remains order one, the x structure of the solutions do satisfy a Korteweg–deVries equation. In this case B is large and E is order 1. Because the y scale becomes limited to L_I, the x dependence and the nonlinearity do not enter in the primary balance, which serves to determine the y structure and a correction to the phase speed. At the next order, the nonlinearity (from both quadratic and Jacobian terms) enters along with the third x derivative and the slow-time derivative terms to give a Korteweg–deVries equation:

$$\eta = F(x - ct, T)Ai(\mathcal{Y}),$$

$$c = -1 - \pi B - (2B)^{2/3}\mathcal{Y}_0,$$

$$\mathcal{Y} = \mathcal{Y}_0 + (2B)^{1/3}\left(y + \frac{\pi}{2}\right), \qquad (18.21)$$

$$\mathcal{Y}_0 = -2.3381 \qquad \text{(zero of Airy function)},$$

$$F_T = F_{xxx} + \frac{3}{2}E\left(\int_{\mathcal{Y}_0}^{\infty} Ai^3 \Big/ \int_{\mathcal{Y}_0}^{\infty} Ai^2\right)\frac{\partial}{\partial x}F^2.$$

This section has demonstrated that some caution must be exercised in applying the quasi-geostrophic equations (which will be discussed throughout the rest of the paper) to large-scale motions since they are valid for the oceans only for scales up to the order of 200 km. The derivations suggest that the role of nonlinearity may be very different for the intermediate and large-scale motions—leading to coherent and phase-locked structures rather than to turbulence. Clearly these inferences must be backed up by more thorough investigations which are beyond the scope of this article.

18.4.3 The Quasi-Geostrophic Equations

Because of the difficulties inherent in attacking the full equations of motion either analytically or numerically, various approximative equations have been developed. For the study of the large-scale motions, the relevant "filtering approximations" eliminate the acoustic and inertiogravity motions.[4] We have mentioned the quasi-geostrophic, semigeostrophic and balance equations and have touched on their limitations. In this section we shall discuss briefly the derivation of the quasi-geostrophic equations for a stratified fluid under oceanic conditions; details can be found in the appendix. These equations are, of course, familiar, but, since we shall use them in the rest of this chapter, we must establish our notation. We wish also to remark on differences between the standard derivation for the atmosphere (cf. Charney, 1973) and that for oceanic conditions. Finally, we include the β-effect by explicitly taking into account the two-scale nature of the problem: the planetary scale, that is, the earth's radius, and the scale of the fluid motions themselves.

For inviscid, adiabatic flow, the equations of motion and continuity expressed in modified spherical coordinates are

$$\frac{Du}{Dt} + \frac{uw}{a+z} - \frac{uv\tan\Theta}{a+z} - 2\Omega v\sin\Theta + 2\Omega w\cos\Theta$$
$$= -\frac{\alpha}{(a+z)\cos\Theta}\frac{\partial p}{\partial\Phi},$$

$$\frac{Dv}{Dt} + \frac{wv}{a+z} + \frac{u^2\tan\Theta}{a+z} + 2\Omega u\sin\Theta = -\frac{\alpha}{a+z}\frac{\partial p}{\partial\Theta},$$

$$\frac{Dw}{Dt} - \frac{u^2+v^2}{a+z} + 2\Omega u\cos\Theta = -\alpha\frac{\partial p}{\partial z} - g, \qquad (18.22)$$

$$-\frac{1}{\alpha}\frac{D\alpha}{Dt} + \frac{1}{(a+z)\cos\Theta}\frac{\partial u}{\partial\Phi} + \frac{1}{(a+z)\cos\Theta}\frac{\partial}{\partial\Theta}(v\cos\Theta)$$
$$+ \frac{1}{(a+z)^2}\frac{\partial}{\partial z}(a+z)^2w = 0,$$

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{u}{(a+z)\cos\Theta}\frac{\partial}{\partial\Phi} + \frac{v}{a+z}\frac{\partial}{\partial\Theta} + w\frac{\partial}{\partial z},$$

where Φ is the longitude, Θ the latitude, Ω the angular speed of the earth's rotation, g the acceleration of gravity, p the pressure, α the specific volume, and u, v, w the eastward, northward, upward velocity components, respectively. The radial coordinate is denoted by $a + z$, where a is the mean radius of the earth and z the height above mean sea level. This neglects the ellipticity of the geoid [see Veronis (1973b) for a discussion of this approximation].

We assume that the specific volume is determined by an equation of state as a function of absolute temperature T, salinity $\mathcal{S}$, and pressure:

$$\alpha = \alpha(T,\mathcal{S},p) \tag{18.23}$$

with salinity conserved,

$$\frac{D}{Dt}\mathcal{S} = 0, \tag{18.24}$$

and temperature changes determined from the adiabatic thermodynamics

$$\frac{DT}{Dt} - \frac{T}{c_p}\left(\frac{\partial\alpha}{\partial T}\right)_{p,\mathcal{S}}\frac{Dp}{Dt} = 0. \tag{18.25}$$

For dynamical modeling it is convenient to regard temperature as a function of specific volume, salinity, and pressure and to determine the evolution of the specific volume from

$$\frac{D\alpha}{Dt} + \frac{\alpha^2}{c_s^2}\frac{Dp}{Dt} = 0, \tag{18.26}$$

which can be derived by taking the substantial derivative of (18.23), using (18.24)–(18.25) and the definition of the sound speed:

$$c_s^2 \equiv -\alpha^2\left[\frac{T}{c_p}\left(\frac{\partial\alpha}{\partial T}\right)_{p,\mathcal{S}} + \left(\frac{\partial\alpha}{\partial p}\right)_{T,\mathcal{S}}\right]^{-1} = c_s^2(\alpha,p,\mathcal{S}). \tag{18.27}$$

Equations (18.26)–(18.27) replace (18.23) and (18.25); since the speed of sound is large compared to the mesoscale wave speeds and also is rather insensitive to its arguments (especially salinity), it plays a rather minor role in the large-scale dynamics.

In the appendix, we write the nondimensional forms of these equations based on a time scale T, a horizontal velocity scale U, a vertical velocity scale W, and a depth scale H. For the horizontal coordinates we introduce two scales of motion: the global, Θ and $\Phi \sim 1$ (the β-effect is global); and the local, $\Delta\Theta$ and $\Delta\Phi \sim L/a$, where L is a typical horizontal scale (cf. Phillips's 1973 WKB approach to Rossby waves). Thus we represent all dependent variables Q in the form $Q(\theta, \phi, z, t, \Theta, \Phi)$ with $d\phi = (a/L)d\Phi$ and $d\theta = (a/L)d\Theta$. We also explicitly introduce a basic hydrostatically balanced stratification of the ocean $\overline{T}(z)$, $\overline{\mathcal{S}}(z)$, $\overline{\alpha}(z)$, $\overline{p}(z)$ satisfying

$$\overline{\alpha}(z)\overline{p}(z) = -g.$$

[In practice, given $\overline{T}(\overline{p})$, $\overline{\mathcal{S}}(\overline{p})$ we find $\overline{\alpha}(\overline{p})$ and integrate to get $z(\overline{p})$.] We then subtract out this hydrostatic state and define the (nondimensional) geostrophic streamfunction ψ by

$$p = \overline{p} + 2\Omega\sin\Theta UL\psi/\overline{\alpha}. \tag{18.28}$$

We also define a "local" potential specific volume α_p of a fluid particle with specific volume α at pressure p and depth z as the specific volume it would acquire if the particle moved adiabatically to the horizontally averaged pressure $\overline{p}(z)$. Equation (18.26) gives

$$\alpha_p = \alpha - \frac{\overline{\alpha}^2}{\overline{c}_s^2}(\overline{p} - p) \tag{18.29}$$

(as long as α and p are not too different from their averaged values). The buoyant force per unit mass after this change becomes

$$b_p \text{ (dimensional)} = g\frac{\alpha_p - \overline{\alpha}}{\overline{\alpha}}$$

$$= g\frac{\alpha - \overline{\alpha}}{\overline{\alpha}} + g\frac{\overline{\alpha}}{\overline{c}_s^2}(p - \overline{p}).$$

This leads to a redefinition of the specific volume in terms of the nondimensional potential buoyancy:

$$\alpha = \overline{\alpha}\left[1 + \frac{2\Omega\sin\Theta UL}{gH}\left(b_p - \frac{gH}{\overline{c}_s^2}\psi\right)\right].$$

With the above scalings, we have eight nondimensional parameters (many of which vary spatially):

$$\epsilon = \frac{1}{2\Omega\sin\Theta T} \quad \text{(a time Rossby number)},$$

$$\epsilon = \frac{U}{2\Omega\sin\Theta L} \quad \text{(a velocity Rossby number)},$$

$$\hat{\beta} = (L/a)\cot\Theta,$$

$$\lambda = H/L,$$

$$\Delta = (2\Omega\sin\Theta L)^2/(gH),$$

$$\Delta_s = gH/\overline{c}_s^2,$$

$$\omega = LW/(HU),$$

$$\hat{S} = H^2\overline{N}^2(z)/(2\Omega\sin\Theta L)^2,$$

where $\overline{N}^2$ is the square of the buoyancy frequency:

$$\overline{N}^2 = (g\overline{\alpha}_z/\overline{\alpha}) - (g^2/\overline{c}_s^2).$$

Two of these parameters, ϵ and $\hat{\beta}$, are identical to those used previously with the definitions $f_0 = 2\Omega\sin\Theta$ and $\beta = 2\Omega\cos\Theta/a$. We have also explicitly separated the time scale from the Rossby wave period, whereas in the previous section ϵ was set equal to $\hat{\beta}\hat{S}$ with $\hat{S} = gH/f_0^2$. The quantity analogous to $\hat{S}$ for a continuously stratified ocean is

$\hat{S} = H^2\overline{N}^2(z)/(2\Omega \sin\Theta L)^2$.

This nondimensional variable is of order unity for motions due to baroclinic instability (Eady, 1949). It is useful to think of it as the squared ratio of two length scales, L_R^2/L^2 or H^2/H_R^2, where $L_R \sim \overline{N}H/f_0$ is the analog for a stratified ocean of the single-layer horizontal deformation radius $\sqrt{gH}/f_0$ introduced by Rossby (1938), and by analogy $H_R \sim f_0 L/\overline{N}$ may be called a vertical deformation radius. If the vertical scale is set, the natural horizontal scale will be L_R; if the horizontal scale is set, the natural vertical scale will be H_R.

We now simplify the equations of motion by making assumptions about the magnitudes of the various parameters. The first seven of our nondimensional parameters are small (for the atmosphere, Δ_s may be of order 1). However, the stability parameter $\hat{S}$ is quite variable. Taking $H \sim 1000$ m as a measure of the depth of the main thermocline, we find that $\hat{S}$ is large in the seasonal thermocline and near unity in the main thermocline. Although this variability is occasionally worrisome in making scale arguments, we shall follow the conventional choice of regarding $\hat{S} \sim O(1)$.

We begin by restricting the length scale L so that $\lambda \ll 1$ and $\Delta \ll 1$, implying that L is large compared to the ocean depth but small compared to the external deformation radius $\sqrt{gH}/f_0 \sim 3000$ km. In practice, we expect the upper limit for L to be determined by the condition that $\hat{S} \gg O(\hat{\beta})$, so that L must be less than the intermediate scale L_I defined in section 18.4.2. Using $\lambda \ll 1$ and $\Delta \ll 1$ and dropping small terms, we obtain the Boussinesq hydrostatic forms of the primitive equations (see the appendix).

Next we specify the time and velocity scale. For the standard quasi-geostrophic motions, the time scale is set by instabilities of the flow so that $T = L/U$ ($\varepsilon \sim \epsilon$) and the vertical velocity is determined by balance between local and advective changes in the vertical component of relative vorticity and stretching of the vortex tubes of the earth's rotation ($\omega = \varepsilon$). Finally, the advective changes of the relative and planetary vorticity are assumed to be comparable, so that $\hat{\beta} \sim \varepsilon$ also. Expanding in ε, we find, as expected, that the lowest-order flows are geostrophic and hydrostatic:

$$u = -\frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \psi}{\partial x}, \quad b_p = f_0 \frac{\partial \psi}{\partial z}, \tag{18.30}$$

where we have redefined the rapidly varying coordinates to look Cartesian by setting $dx = L\cos\Theta\, d\phi$ and $dy = L\, d\theta$ and have returned to dimensional variables. The full pressure is related to the streamfunction by

$$p = \bar{p}(z) + f_0\psi/\bar{\alpha}(z). \tag{18.31}$$

The vorticity equation, which is derived by cross differentiating the order-Rossby number momentum equations (with special care taken with the Θ and Φ dependence), and use of the order-Rossby number continuity equation, becomes

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)(\nabla^2\psi + \beta y) = f_0 w_z, \tag{18.32}$$

and the buoyancy equation becomes

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\psi_z + f_0 S w = 0. \tag{18.33}$$

Here $S = \overline{N}^2(z)/f_0^2$, $\mathbf{v} = (-\psi_y, \psi_x)$, and $\nabla = (\partial/\partial x, \partial/\partial y)$. These two may be combined to give the quasi-geostrophic equation

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\left(\nabla^2\psi + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\psi + \beta y\right) = 0, \tag{18.34}$$

which asserts that the quantity

$$q = \nabla^2\psi + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\psi + \beta y$$

is conserved at the projection of a particle in a horizontal plane, not, like potential vorticity, at the particle. For this reason it is called *pseudopotential vorticity* to distinguish it from potential vorticity. Because the distinction vanishes for a fluid consisting of several homogeneous incompressible or barotropic layers, there has been some confusion of terminology in the literature.

The temperature and salinity fields can be derived from the streamfunction ψ and the basic stratification $\overline{T}(z)$, $\overline{\mathcal{S}}(z)$, using the salinity and temperature equations together with the expression (18.33) for the vertical velocity:

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)(\mathcal{S} - \overline{\mathcal{S}}) + w\overline{\mathcal{S}}_z = 0,$$

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)(T - \overline{T}) + w\left(\overline{T}_z - \frac{g\overline{T}\overline{\alpha}_T}{\overline{\alpha}\overline{c}_p}\right) = 0.$$

To complete the system of equations we need the boundary conditions. At the bottom boundary vertical velocities are forced by flow over topography:

$$w = \mathbf{v}\cdot\nabla b \qquad \text{at} \quad z = -H, \tag{18.35a}$$

where $H(\Theta,\Phi)$ is the (local) mean depth, the true bottom being at $z = -H + b$. For consistency, $|b|/H$ is required to be order ε. At the upper free surface $z = \eta$, the assumption that L is small compared to the external radius of deformation implies that the boundary conditions

$$\left.\begin{aligned} D\eta/Dt &= w \\ p &= 0 \end{aligned}\right\} \qquad \text{at} \quad z = \eta$$

can be approximated simply by

$$w = 0 \qquad \text{at} \quad z = 0, \tag{18.35b}$$

with the surface displacement computed from

$$\eta = \frac{f_0}{g}\psi(x,y,0).$$

Finally, on the side-wall boundaries it is necessary to set both the order 1 and order ε normal velocities to zero, giving

$$\begin{aligned} \nabla\psi\cdot\hat{\mathbf{s}} &= 0, \\ \oint \nabla\psi_t\cdot\hat{\mathbf{n}} &= 0, \end{aligned} \tag{18.36}$$

where $\hat{\mathbf{s}}$ is the unit tangent vector and $\hat{\mathbf{n}}$ the unit normal vector to the boundary.

All of these conditions will be modified in the presence of friction: the top and bottom layers because of Ekman pumping into or out of the frictional layer (see section 18.6) and the side conditions by the necessity for upwelling layers which can feed offshore Ekman transports and can accept mass flux from the interior of the ocean.

18.5 Linear Quasi-Geostrophic Dynamics of a Stratified Ocean

The quasi-geostrophic equations (18.33)–(18.36) have been applied to large-scale, long-period, free- and forced-wave motions in the atmosphere, to the study of barotropic and baroclinic instability, to wave–mean flow and wave–wave interaction, and to geostrophic turbulence. We have mentioned already the review articles by N. Phillips (1963) and Kuo (1973) and the book by Pedlosky (1979a) in which their applications are treated. In addition, Dickinson (1978) has reviewed their application to long-period oscillations of oceans and atmospheres and Holton (1975) their application to upper-atmosphere dynamics.

In application to the mesoscale eddy range of oceanic motions (10 km $< L <$ 210 km), equations (18.32)–(18.36) exhibit a rich variety of behavior depending on the sizes of the various parameters and the initial and boundary conditions. We cannot discuss all of them here; rather we shall confine ourselves to a few topics which are also familiar in a meteorological context. We shall, whenever possible, use a typical oceanic $N^2(z)$ profile (Millard and Bryden, 1973; also see figure 18.7) rather than the constant- or delta-function profiles that are most commonly considered. This allows us to describe the vertical dependence of theoretical predictions in a way which is more directly comparable with oceanic data.

We begin with phenomena that are essentially linear—involving no transfers of energy between scales—deferring discussion of nonlinear motions to the next section.

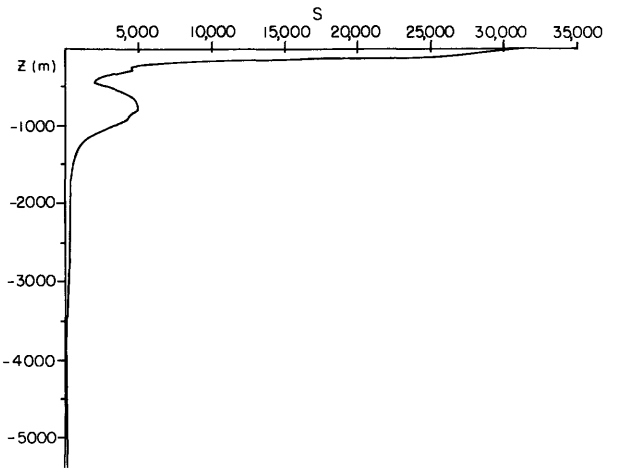


Figure 18.7 Typical oceanic structure for $S(z) = N^2/f_0^2$. The data are from Millard and Bryden (1973) and represent an average over ten stations centered on 28°N and 70°W.

18.5.1 Rossby Waves and Topographical Rossby[5] Waves

Rhines (1970) has discussed the nature of free quasi-geostrophic waves in a uniformly stratified fluid with bottom topography in some detail. We shall describe the behavior of these waves with the aid of a formalism that permits us to extend Rhines's results to real $N^2(z)$ profiles.

When the bottom slope is uniform (b_x and b_y constant), the equations are separable, so that we can write the streamfunction in the form

$$\psi = AF(z)\sin[k(x - ct) + ly], \tag{18.37}$$

where

$$c = -\beta/(k^2 + l^2 + \lambda^2) \tag{18.38}$$

and λ, the separation constant, governs the z dependence:

$$\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}F = -\lambda^2 F. \tag{18.39}$$

To close the system, we make use of the boundary conditions

$$\frac{\partial}{\partial t}\psi_z = 0, \quad z = 0,$$

$$\frac{\partial}{\partial t}\psi_z = -f_0 S(-H)J(\psi,b), \quad z = -H,$$

which become

$$F_z = 0, \quad z = 0, \tag{18.40a}$$

$$F_z = \frac{-f_0 S(-H)}{\beta}\left(b_y - \frac{l}{k}b_x\right) \times (k^2 + l^2 + \lambda^2)F, \quad z = -H. \tag{18.40b}$$

To solve (8.39) and (18.40) we proceed as follows: given $S(z)$ we integrate (18.39) with the boundary condition (18.40a) and the normalization condition

$$(1/H)\int_{-H}^{0} dz \quad F^2(z;\lambda^2) = 1$$

(using a simple staggered-grid difference scheme with 50-m vertical resolution). We then define the nondimensional function

$$R(\lambda^2) \equiv \frac{-HF_z(-H;\lambda^2)}{S(-H)F(-H;\lambda^2)} \tag{18.41}$$

in terms of which the bottom boundary condition (18.40b) becomes

$$R(\lambda^2) = \frac{f_0 H}{\beta}\left(b_y - \frac{l}{k}\, b_x\right)/(k^2 + l^2 + \lambda^2). \tag{18.42}$$

This can be used to determine λ^2 and the vertical structure $F(z)$, given the wave scale $(k^2 + l^2)^{-1/2}$ and the propagation angle $\tan\theta = l/k$.

Thus, we can summarize all of the information in one graph. Figure 18.8 shows $R(\lambda^2)$, from which λ^2 can be determined given the wave numbers and the topographic slopes by a graphical solution of (18.42). From the resulting set of λ^2 values, the values of the phase speeds of the various waves can then be determined from (18.38). The vertical structures of the waves and the dependence of their phase speeds on the slopes and wave numbers are qualitatively similar to those in Rhines's (1970) constant-N model. We shall describe these results and give useful approximate formulas for c.

When there is no slope effect, the λ_n values are simply the inverses of the deformation radii associated with the various modes $F_n(z)$ which are eigensolutions of (18.39) under the condition $F_z = 0$ for $z = 0, -H$, normalized so that $(1/H)\int_{-H}^{0} dz\, F_n^2(z) = 1$. The barotropic ($n = 0$) and first baroclinic modes ($n = 1$, $L_R = 46$ km) correspond to the structures observed in oceanic data (cf. Richman, 1976). The vertical dependence of these structures are shown also in figure 18.8. The dispersion relation (18.38) and the propagation characteristics of the various modes are described in many places (see chapter 10).

For the weak topographic slope effect, the phase speeds are altered to

$$c \simeq -\left[\beta + \frac{f_0 F_n^2(-H)}{H}\left(b_y - \frac{l}{k}\, b_x\right)\right]\Big/ (k^2 + l^2 + \lambda_n^2),$$

derived by solving (18.39) for $\lambda^2 - \lambda_n^2$ small. We find the familiar result that the bottom slope, by causing vortex stretching or shrinking, acts as an effective β. The bottom-slope effect in the baroclinic modes is weaker by a factor $F_n^2(-H)$ than the corresponding effect

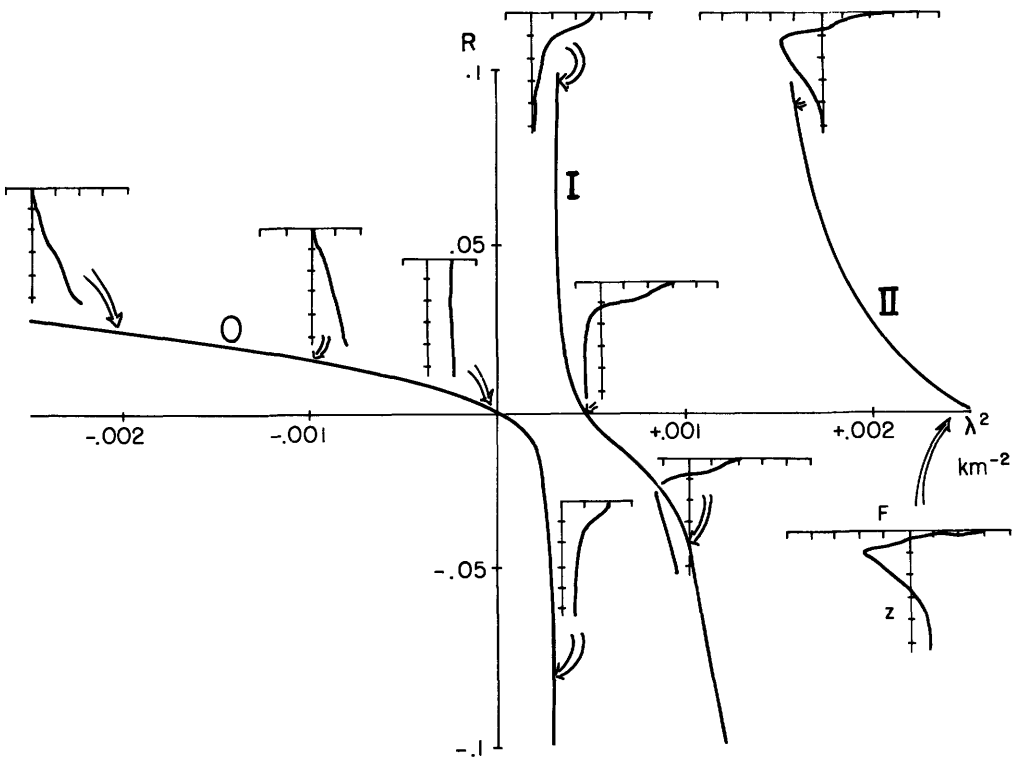


Figure 18.8 Normalized ratio of bottom shear to bottom amplitude as a function of the separation constant λ^2. For given topographic slopes and wavenumber vector, the values of λ^2 are the intersections of this curve with the line $R = (f_0 H/\beta)[b_y - (l/k)b_x](\lambda^2 + k^2 + l^2)$. Also shown is the vertical structure $F(z)$ of the streamfunction (normalized to rms value unity) at various (λ^2, R) values.

on the barotropic mode. This factor is smaller than 1 for the stratification used (about 0.4 for the first baroclinic mode).

When the slope effect $f_0 H[b_y - (l/k)b_x]/\beta$ is negative, there is also a bottom-trapped eastward-moving wave. The vertical trapping scale is $H_s = -(f_0/\beta)[b_y - (l/k)b_x]$, and the speed

$$c = \beta N^2(-H)H_s^2/f_0^2$$

can be regarded as that of a long wave in a fluid with a deformation radius based on the vertical scale H_s and the local value of N.

For large slopes the modes all have large vertical shear. Most of them are surface trapped, having essentially zero bottom amplitude and a westward component of phase speed. For a slope effect opposing the β-effect, there is also a rapidly eastward-moving wave whose vertical trapping scale is $H_R = f_0/N(-H)\sqrt{k^2 + l^2}$ and whose propagation speed is

$$c = -f_0(b_y - \frac{l}{k} b_x)/H_R(k^2 + l^2)$$

in the limit of $L \ll L_R$.

The constant N or two-layer models make significant qualitative errors in describing one or another of these types of behavior. The flat-bottom baroclinic modes in a constant-N model have bottom velocities comparable to the surface velocities; when a weak slope is added, this produces a large change in the vortex stretching and in the phase speed. Thus, the effect of the slope on the baroclinic modes is twice as great as on the barotropic mode, not, as in the realistic ocean, half as great. For large slopes the trapping scale for constant N is much smaller than the correct scale because $N(-H)$ is small compared to the average value of N which would be used in a constant N model. This results in an overestimate of the phase speed. The two-layer model generally misrepresents the eastward-traveling mode: for weak slopes, it is altogether absent, while for strong slopes the two-layer model predicts $c \sim (k^2 + l^2)^{-1}$ rather than the correct $(k^2 + l^2)^{-1/2}$ dependence.

Although we have derived the solutions (18.37)–(18.42) from the linearized equations, the individual waves are also finite amplitude solutions to the equations of motion (18.32)–(18.36) because both

$$\left(\nabla^2 + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\right)\psi \quad \text{and} \quad \psi_z$$

are proportional to ψ at each horizontal level. In fact, the nonlinearities in the equation of motion (18.34) will vanish for any set of waves having each of their wave vectors parallel or having the same total scale $(k^2 + l^2 + \lambda^2)^{-1/2}$. In the latter case, all the waves have the same phase speed so that the composite streamfunction pattern propagates uniformly. The nonlinearity in the bottom boundary condition likewise vanishes if all waves have the same value of R, that is, if the wavenumber vectors are parallel or $b_x = 0$. Thus sets of waves with horizontal scales $(k^2 + l^2)^{-1/2}$ such that both $k^2 + l^2 + \lambda^2 =$ constant and $R(\lambda) =$ constant will be full nonlinear solutions for north–south bottom slopes. Instabilities may, of course, prevent such patterns from persisting.

18.5.2 Generation of Rossby Waves by Flow over Topography

Flow over topography has been of practical interest to meteorologists attempting to forecast conditions near large mountain ranges for many years. There is an extensive bibliography of such studies (Nicholls, 1973; Hide and White, 1980). Many of these concentrate on smaller-scale lee-wave properties, although there have also been attempts to model the large-scale standing atmospheric eddies as topographically forced Rossby waves (cf. Charney and Eliassen, 1949; Bolin, 1950).

In the ocean, there are also many standing features, some of which clearly can be identified with topography (cf. Hogg, 1972, 1973; Vastano and Warren, 1976). In this section we shall discuss a simple oceanic analogy to the common atmosphere model for standing waves produced by flow over topography.

Steady solutions to (18.33)–(18.35) may be written

$$\nabla^2\psi + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\psi + \beta y = P(\psi,z),$$

$$\psi_z = T_s(\psi), \quad z = 0, \tag{18.43}$$

$$\psi_z + f_0 S(-H)b = T_b(\psi), \quad z = -H,$$

where P, T_s, and T_b are arbitrary functionals. If we allow ψ to represent a mean zonal flow plus a topographically induced deviation,

$$\psi = -\bar{u}(z)y + \phi(x,y,z),$$

we find

$$P(\psi,z) = -\left(\beta - \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial \bar{u}}{\partial z}\right)\psi/\bar{u},$$

$$T_s = T_b = 0$$

are suitable functionals to match the terms linear in y in (18.43). We have restricted ourselves to flows such that $\bar{u} \neq 0$ everywhere and $\bar{u}_z = 0$ at $z = 0, -H$. The conditions that the upper and lower surfaces be isothermal are not fundamental and could be relaxed easily; we make them simply to restrict the discussion to a reasonable number of parameters. If, however, the $\bar{u} \neq 0$ condition is violated, the analysis becomes much more difficult since the critical-layer (where $\bar{u} = 0$) problem must also be solved.

Given these restrictions, however, the fluctuation field satisfies the simple equations

$$\nabla^2\phi + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\phi + \frac{\beta - (\partial/\partial z)[1/S](\partial\bar{u}/\partial z)}{\bar{u}}\phi = 0,$$

$$\phi_z = 0, \quad z = 0, \tag{18.44}$$

$$\phi_z = -f_0 S(-H) b(x,y), \quad z = -H,$$

where no small-amplitude assumption has been made beyond the assumption that the nondimensional topographic amplitude is of the order of the Rossby number. The linearity of these equations is due to the particularly simple form of the upstream flow—a streamfunction field which is linear in y. Any horizontal shear in the upstream flow would give rise to nonlinear terms in P, T_s, and T_b, and thereby to nonlinearities in the equations (18.44). For simple sinusoidal topography $b = b_0 \sin(kx + ly)$, the topographically forced wave looks like

$$\phi = f_0 b_0 H F(z) \sin(kx + ly),$$

where

$$\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}F = \left[k^2 + l^2 - \frac{\beta - (\partial/\partial z)[1/S](\partial\bar{u}/\partial z)}{\bar{u}}\right]F,$$

$$F_z = 0, \quad z = 0, \tag{18.45}$$

$$F_z = -S(-H)/H, \quad z = -H.$$

When the zonal flow is barotropic ($\bar{u}$ = constant) the system (18.45) becomes essentially identical to (18.39)–(18.40) except that the amplitude is determined by the bottom boundary condition. We summarize the shapes and amplitudes of the forced waves in figure 18.9. We have used again the simplest form $\lambda^2 = \beta/\bar{u} - (k^2 + l^2)$ for the dependent variable.

The most striking feature in the response is, of course, the resonant behavior when

$$\bar{u} = \beta/(k^2 + l^2 + \lambda_n^2)$$

for λ_n^2 one of the inverse-square deformation radii. Near such a resonance, the amplitude becomes very large:

$$F(z) \simeq \frac{F_n(z)F_n(-H)}{[k^2 + l^2 - (\beta/\bar{u}) + \lambda_n^2]H^2}.$$

When the mean flow is eastward at a few centimeters per second, it may be near a resonance for one of the baroclinic modes and may therefore generate substantial currents even above the thermocline. Vertical standing-wave modes associated with vertical propagation and reflection at the upper boundary will be found for $\lambda^2 > 0$ or

$$0 < \bar{u} < \beta/(k^2 + l^2).$$

For $\lambda^2 < 0$ the motions are trapped and decay away from the bottom. The trapping scale becomes very small when $\bar{u}$ is nearly zero but westward or when the topographic wavelength is short. In the latter case, the vertical scale of the fluctuations is $H_R \sim f_0 L/N(-H)$, where $L = (k^2 + l^2)^{-1/2}$. When $\bar{u}$ is weak and westward ($0 < |\bar{u}| < \beta N^2(-H)H^2/f_0^2 \sim 4$ cm s), the vertical scale will again be small compared to the fluid depth.

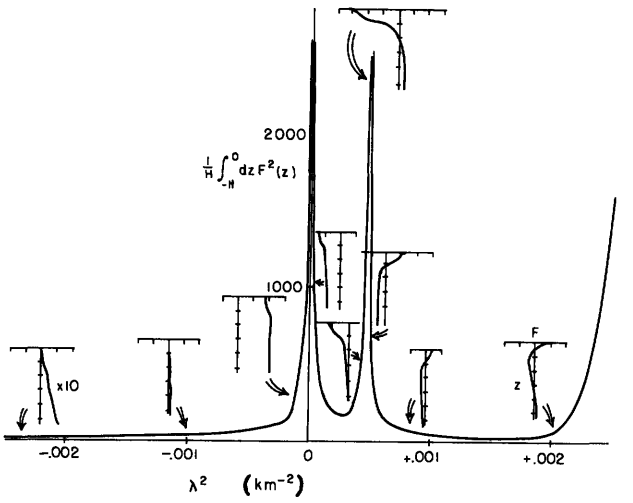


Figure 18.9 Energy and vertical structure of a topographically forced wave as a function of $\lambda^2 = (\beta/\bar{u}) - \mathbf{k} \cdot \mathbf{k}$, where $\mathbf{k}$ is the topographic wavenumber.

Although this type of problem is suggested by the atmosphere analogue, a warning about its applicability may be in order. Periodic problems are natural for the atmosphere. In the ocean, however, it is less plausible that a fluid particle periodically will revisit a topographic feature in a time less than the damping time for the excited wave. (The Antarctic Circumpolar Current could perhaps be an exception.) We can illustrate the differences between periodic and local topography by considering uniform eastward flow over a finite series of hills and valleys:

$$b = \begin{cases} b_0 \sin kx, & 0 \le x \le n\pi/k \\ 0, & x < 0 \quad \text{or} \quad x > n\pi/k. \end{cases}$$

We consider the barotropic component of flow which satisfies the depth averaged form of (18.44):

$$\nabla^2\hat{\phi} + \frac{\beta}{\bar{u}}\hat{\phi} = -\frac{f_0}{H}b(x,y).$$

Its solution is

$$\hat{\phi} = 0 \quad \text{for } x < 0,$$

$$\hat{\phi} = \frac{f_0 b_0}{H(k^2 - \beta/\bar{u})} \times \left(\sin kx - k\sqrt{\frac{\bar{u}}{\beta}}\sin\sqrt{\frac{\beta}{\bar{u}}}x\right)$$

$$\text{for } 0 < x < \frac{n\pi}{k},$$

$$\hat{\phi} = \frac{f_0 b_0}{H(k^2 - \beta/\bar{u})} \times k\sqrt{\frac{\bar{u}}{\beta}}\left[\cos\sqrt{\frac{\beta}{\bar{u}}}x - (-1)^n \cos\sqrt{\frac{\beta}{\bar{u}}}\left(x - \frac{n\pi}{k}\right)\right]$$

$$\text{for } x > \frac{n\pi}{k}.$$

Figure 18.10 shows the average energy in the far field (normalized by $\frac{1}{2}(f_0 b_0/Hk)^2$ as a function of $\bar{u}k^2/\beta$. We note that the resonance peak becomes significant only when the topography has a number of hills and valleys; this suggests that the idea of resonance, in the ocean, should be applied with caution.

When there is vertical shear in the mean flow, the situation becomes somewhat different, although equation (18.45) can still readily be integrated. However, one can gain a qualitative picture of the response for arbitrary shear and small perturbations by using the methods of Charney and Drazin (1961) as described in the next section.

18.5.3 Propagation and Trapping of Neutral Rossby Waves

In many circumstances, the ocean or atmosphere is directly forced by external conditions—heating, winds—which may have temporal and spatial variations. The forcing may generate wave disturbances in one region that propagate into a neighboring region (e.g., the propagation of tropospheric disturbances into the stratosphere). In these circumstances the motion is determined by the nature of the forcing and the refractive properties of the intervening medium. Charney (1949) first treated the vertical propagation of Rossby waves in a stratified atmosphere and Charney and Drazin (1961) first suggested the analogy between vertical propagation of Rossby waves and electromagnetic wave propagation in a medium with a variable index of refraction (possibly complex, corresponding to wave absorption). Holton (1975) has reviewed these concepts for meteorologists; oceanographers have tended to make less use of them [see, however, Wunsch (1977), who applied them to vertically propagating equatorial waves excited by monsoon winds].

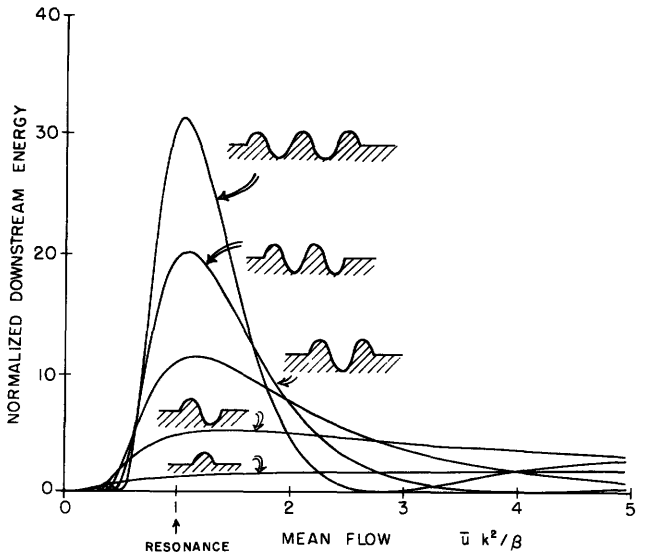


Figure 18.10 Downstream energy averaged over a wavelength of waves forced by flow over isolated bumps as a function of the normalized mean flow speed. The wavenumber of the topography in the region where it is varying is k. For an infinite topography, resonance would occur at $\bar{u}k^2/\beta = 1$. Results are shown for varying numbers of elevations and depressions in the topography.

The simplest derivation of an index of refraction is for waves in a zonal flow with vertical and horizontal shear. Consider waves of infinitesimal amplitude having east-west wavenumber k and frequency ω. The north-south and vertical dependencies of the amplitude $\psi = \Psi(y,z)e^{i(kx-\omega t)}$ are determined by the standard stability equation (cf. Charney and Stern, 1962):

$$\left(\bar{u} - \frac{\omega}{k}\right)\left(\frac{\partial^2}{\partial y^2} + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z} - k^2\right)\Psi + \left(\beta - \bar{u}_{yy} - \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\bar{u}\right)\Psi = 0.$$

If we follow the procedure that has been used for vertically propagating waves, we transform this into a Helmholtz equation with a variable coefficient of the undifferentiated term. This is quite straightforward if we are considering propagation only in the y direction with $\bar{u}_z = 0$ and $\Psi = \Phi(y)F_n(z)$, where F_n is one of the flat bottom eigenfunctions. Then the y structure is governed by

$$\frac{\partial^2}{\partial y^2}\Phi + \nu^2(y)\Phi = 0$$

where

$$\nu^2(y) = [(\beta - \bar{u}_{yy})/(\bar{u} - \omega/k)] - k^2 - \lambda_n^2. \qquad (18.46)$$

A simple illustrative example is the radiation from a meandering Gulf Stream into the neighboring Sargasso Sea (cf. Flierl, Kamenkovich, and Robinson, 1975; Pedlosky, 1977). The forcing specifies Φ at some latitude. When we have no mean flow ($\bar{u} = 0$) and the motions are barotropic, the index of refraction becomes

$$\nu^2(y) = -k^2 - (\beta k/\omega)$$

The north-south scale of the response $|\nu|^{-1}$ is shown as a function of ω (>0) and k ($\gtrless 0$) in figure 18.11. Most observations indicate eastward-traveling motions, $\omega/k > 0$, implying that the mid-ocean response will be trapped close to the Gulf Stream.

We may obtain a similar representation of the index of refraction for the full two-dimensional (y and z) problem. This result, for reasons discussed below, is probably of more interest to meteorologists than to

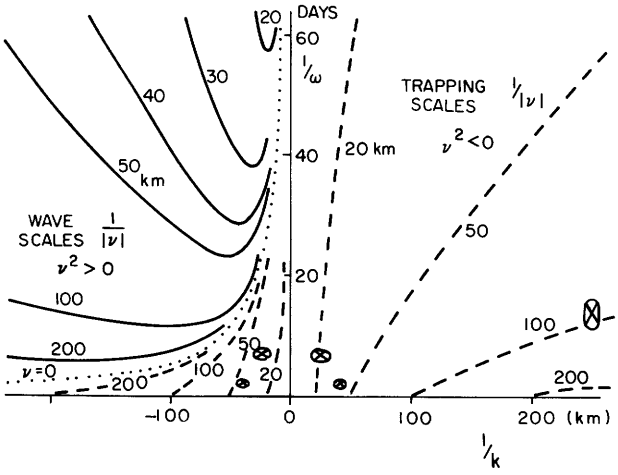


Figure 18.11 Solid lines show north-south length scales (wavelength/2π) and dashed lines shown trapping scales (e-folding distance) for barotropic waves generated by a meandering current with inverse frequency ω^{-1} and inverse wavenumber k^{-1}. Eastward going meanders ($k > 0$) produce trapped waves; westward going meanders ($k < 0$) may produce propagating disturbances. The symbols ⊗ correspond to typical observational estimates of ω^{-1} and k^{-1}.

oceanographers; however, we include it to illustrate some of the effects of the z structure. If we substitute $\Psi(y,z) = S^{1/4}\Phi(y,\zeta)$, where $\zeta = \int_{-H}^{z} S^{1/2}(z')dz'$ is a modified vertical coordinate, we find

$$\left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial \zeta^2}\right)\Phi + \nu^2(y,\zeta)\Phi = 0,$$

where the index of refraction $\nu^2(y,\zeta)$ is given by

$$\nu^2(y,\zeta) = S^{-1/4}[S^{-1}(S^{1/4})_z]_z + \frac{\beta - \bar{u}_{yy} - (\bar{u}_z/S)_z}{\bar{u} - \omega/k} - k^2. \tag{18.47}$$

When $\nu^2 > 0$ there are sinusoidal solutions and energy propagates freely, whereas when $\nu^2 < 0$ there are only exponential solutions (along the ray) and the waves die out. There are also, of course, diffraction effects and tunneling effects if the regions of negative ν^2 (or, at least, significantly altered ν^2) are relatively small. This form is useful when N is a simple function (e.g., $N_0 e^{z/d}$) so that the first term in (18.47) is also simple $[-3/(4d^2S)]$. The stratification then contributes a relatively large and negative term which increases toward the bottom, inhibiting penetration into the deep water. For our $S(z)$ profile (figure 18.7), however, numerical differentiation proved to be excessively noisy. Moreover, in the oceans, most of the motions of interest have vertical scales that are significantly influenced by the boundaries and are larger than the scales of variation of ν^2, so that a local (WKB) interpretation of ν^2 variations is not possible.

We can, however, associate modifications in ν^2 occurring on large scales with modifications in the structure of Ψ. Thus in the topographic problem, if the shear in the vertical is such that

$$\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial \bar{u}}{\partial z} > 0 \quad \text{and} \quad \frac{\partial \bar{u}}{\partial z} > 0,$$

there will be a decrease in the value of ν^2, implying that the wave will become either more barotropic ($\nu^2 > 0$) or more bottom trapped ($\nu^2 < 0$). In the example of Rossby wave radiation from a meandering Gulf Stream, (18.46) implies that the baroclinic modes ($\lambda_n^2 > 0$) become trapped even more closely than the barotropic modes.

As a final example, we note that the motions forced in the ocean by atmospheric disturbances tend to have large positive ω/k and large scales. In the absence of mean currents, the vertical structure equation, with $\Psi = e^{ily}F(z)$, becomes

$$\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}F = \left[\frac{\beta k}{\omega} - k^2 - l^2\right]F = -\nu^2 F, \tag{18.48}$$

implying that the forced currents are nearly barotropic. However, the recent work of Frankignoul and Müller (1979) suggests a possible mechanism by which significant baroclinic currents may be produced. Because the ocean is weakly damped and has resonant modes ($\nu^2 = \lambda_n^2$), even very small forcing near these resonances can cause the energy to build up in these modes. This is another example of the strong influence of the boundaries on the oceanic system.

18.6 Friction in Quasi-Geostrophic Systems

18.6.1 Ekman Layers

Ekman (1902, 1905), acting on a suggestion of Nansen, was the first to explore the influence of the Coriolis force on the dynamics of frictional behavior in the upper wind-stirred layers of the oceans. He considered both steady and impulsively applied, but horizontally uniform, winds. In an effort to understand how surface frictional stresses τ influence the upper motion of the atmosphere and, in particular, how a cyclone "spins down," Charney and Eliassen (1949) were led to consider horizontally varying winds. They showed that Ekman dynamics generates a horizontal convergence of mass in the atmospheric boundary layer proportional to the vertical component of the vorticity of the geostrophic wind in this layer. Thus a cyclone produces a vertical flow out of the boundary layer which compresses the earth's vertical vortex tubes and generates *anticyclonic* vorticity. The time constant for frictional decay in a barotropic fluid was found to be $(f_0 E^{1/2})^{-1}$, where E is the Ekman number $\nu_e/f_0 H^2$, with ν_e the eddy coefficient of viscosity and H the depth of the fluid. Greenspan and Howard (1963) investigated the time-

dependent motion of a convergent Ekman layer: if the wind is turned on impulsively, the Ekman layer is set up in a time of order f_0^{-1}; the internal flow decays in a time of order $(f_0E^{1/2})^{-1}$; and the vertical oscillations that are produced by the impulsive startup decay in a time of order $(f_0E)^{-1}$. Since f_0^{-1} is but a few hours, one may consider that for the large-scale wind and current systems of the atmosphere and oceans the Ekman pumping is produced instantly and that there is a balance in the Ekman layer among the frictional pressure and Coriolis forces. We divide the flow into a quasigeostrophic interior component (u_g, v_g, w_g) with associated pressure gradients $fv_g = \alpha p_{g_z}$ etc. and a deviation component associated with the friction (u_e, v_e, w_e) which vanishes below some small depth h. For a homogeneous fluid $p_e = 0$ because the hydrostatic assumption ensures that there can be no nontrivial pressure field which vanishes below $z = -h$. For a stratified flow a scaling argument can be made to show that buoyancy fluctuations in the upper layer will not be important enough to cause significant p_e's (unless $N^2 > \tau_0 L/\rho h^3$) so that $\rho f v_e = -(\partial/\partial z)\boldsymbol{\tau}\cdot\hat{\mathbf{x}}$, etc. If we divide by f, and compute w_{e_z} from the divergence of the Ekman horizontal velocities, we find

$$w_e = -\hat{\mathbf{z}}\cdot\mathrm{curl}(\boldsymbol{\tau}/\rho f)$$

using $\boldsymbol{\tau}(-h) = 0$, $w_e(-h) = 0$. From the surface condition $w_e(0) + w_g(0) = 0$, the Ekman pumping is therefore

$$w_g(0) \equiv w_E = \hat{\mathbf{z}}\cdot\mathrm{curl}(\boldsymbol{\tau}(0)/\rho f), \tag{18.49}$$

where $\boldsymbol{\tau}(0)$ is the wind stress at the sea surface. The same procedure can be used in the lower boundary layer:

$$w_g = \frac{\partial}{\partial x}\left(\frac{\nu_e v_{e_z}}{f}\right) - \frac{\partial}{\partial y}\left(\frac{\nu_e u_{e_z}}{f}\right), \qquad z = -H.$$

But now it is necessary to specify $\boldsymbol{\tau}(-H)$ in terms of the geostrophic velocities u_g, v_g; for this a knowledge of ν_e is required. If we assume ν_e to be constant, the pumping out of the bottom boundary layer is given by

$$w_g(-H) \equiv w_E = \frac{D_E}{2}\left[\zeta_g + u_{g_x} + v_{g_y} + \frac{1}{2}(\beta/f)(u_g - v_g)\right]\Bigg|_{z=-H},$$

where $D_E = (2\nu_e/f)^{1/2}$ and $\zeta_g = v_{g_x} - u_{g_y}$ is the vorticity of the geostrophic wind. When $L \ll a$, the divergence terms (which are equal to $-\beta v_g/f$) and the last term are negligible, so that

$$w_g(-H) = \frac{D_E}{2}\zeta_g(-H). \tag{18.50}$$

In the lower boundary layer of the deep ocean, the water is nearly homogeneous. In this case one may estimate the bulk viscosity ν_e by supposing that for this value the established boundary layer is marginally stable (cf. Charney, 1969). From the measurements of Tatro and Mollo-Christensen (1967), the condition for marginal stability is found to be that the Reynolds number based on the depth D_E, of the Ekman layer $UD_E/\nu_e = \sqrt{2}U/\sqrt{f\nu_e}$, shall be of order 100. Thus, for example, $\nu_e \sim U^2/5000f = 200\ \mathrm{cm^2\,s^{-1}}$, and $D_E \sim U/50f \sim 20$ m for a current of 10 $\mathrm{cm\,s^{-1}}$ in middle latitudes.

In a stratified atmosphere or ocean, the depth of influence of the Ekman pumping is not necessarily the depth of the fluid. If a circulation is forced from above by Ekman pumping with horizontal scale L, one expects the depth of influence to be the vertical deformation radius $H_R \sim f_0 L/N$. This depth will be comparable to the ocean depth for $L \sim L_R = 50$ km. Most surface forcing will thus excite a barotropic response. The spin-down of baroclinic mesoscale ocean eddies will be considered in Section 18.6.3.

18.6.2 Spin-Up of the Ocean

The problem of the spin-up of the entire ocean requires definition. The wind and thermally driven circulations are so coupled nonlinearly that it is not possible to treat the establishment of the wind-driven circulation independently. The important question, however, is not how the ocean circulation would be established from rest if the forcing were impulsively applied, but rather how the circulation would change if the forcing changed. The latter question has clear implications for understanding the role of the oceans in climatic change. Thus, one is led to consider first the small-amplitude adjustment of a given steady-state circulation to a change in the wind stress, with the expectation that nonadiabatic changes will require considerably long times. Even for this linearized problem, results for the spin-up of the ocean in mid-latitudes have been obtained (Anderson and Gill, 1975; Anderson and Killworth, 1977; Cane and Sarachik, 1976, 1977) only for the simplest cases of a one- or two-layer model with no preexisting circulation. The solutions for a suddenly applied wind stress are complicated, but their qualitative import can be simply stated. When a steady, east–west wind stress is suddenly applied to a two-layer ocean, initially at rest, the motion at any longitude increases uniformly with time until a nondispersive Rossby wave starting at the eastern boundary and moving with the maximum westward baroclinic group velocity $-\beta L_R^2$ reaches that longitude. When this occurs, a steady Sverdrup flow induced by the wind-stress curl will have been established in the upper layer everywhere to the east of that longitude. By the time the Rossby wave reaches the western boundary, a steady state will have been established over the entire ocean—except in the vicinity of the boundary itself, where slow-moving reflected Rossby waves influence the flow and are presumed to be dis-

sipated by friction. Thus the spin-up time is essentially the time required for a signal traveling at the speed $-\beta L_R^2$ to cross the ocean from east to west. For width of 6000 km, we obtain 1.5×10^8 or about 5 years.

We note that βL_R^2 increases toward the equator. However, as one approaches the equator the dynamics of wave propagation change. Near the equator, Rossby-gravity and Kelvin waves are generated. These have maximum group velocities of order $\sqrt{g'H}$ (g' is the reduced gravity and H the depth of the thermocline) $\sim 1\ \mathrm{m\,s^{-1}}$, giving spin-up times of the order of months rather than years. Cane (1979a) and Philander and Pacanowski (1980a) have shown that an impulsively generated uniform westward wind produces both equatorially trapped Kelvin and Rossby-gravity waves. The equatorial undercurrent is established at a given longitude when a Kelvin wave traveling eastward from the western boundary reaches that longitude. The dynamics of equatorially trapped planetary wave modes have been investigated by Rosenthal (1965) and Matsuno (1966) for the atmosphere and by Blandford (1966), Lighthill (1969), and Cane and Sarachik (1979) for the oceans. The dynamics of the equatorial undercurrent has been reviewed by Philander (1973, 1980).

A similar linear analysis for a continuously stratified ocean initially at rest leads to quite different results. In this case, a wind stress can produce a steady Sverdrup transport only in the upper frictional boundary layer. This is the result of the conservation of density, which requires $wS = 0$ or $w = 0$, and it follows from the interior geostrophic dynamics that $\beta v = f w_z = 0$. The initial application of the wind stress will produce an infinity of transient internal baroclinic modes whose sum will approach zero in time everywhere except at $z = 0$. If we consider only the barotropic and first baroclinic modes, the temporal evolution will be similar to that of the two-layer ocean, but the effect of the other modes will be such as to cause all interior velocities to vanish asymptotically in time.

However, if a perturbation in wind stress is applied to *preexisting* flow, Ekman pumping can penetrate into the interior along isopycnals and w_z need not be zero. Although this calculation has not been made in detail, it seems plausible that the final perturbation structure would be similar to the mean flow structure and, therefore, that it would be spun up in the time associated with the cross-ocean propagation of the lowest baroclinic modes.

It is also important to note that the definition of the spin-up time depends to some degree on the property one is considering. For example, the Sverdrup balance (see Leetmaa, Niiler and Stommel, 1977, for an empirical discussion) is established on relatively short time scales. If the ocean is forced by the Ekman pumping,

$$w_E = w_0 \exp[ikx - i\omega t],$$

it may be seen from the vorticity equation (18.32) that Sverdrup balance will be attained when $|\omega k| \ll \beta$. Thus fluctuations in forcing on the size of the basin with periods even as short as a few days—the time for the *barotropic* wave to cross the basin—will preserve the Sverdrup balance.

Clearly there are many unanswered questions concerning even the adiabatic response of the ocean to changes in the forcing. We know still less about the response time of the entire wind-driven thermohaline circulation, although we expect the time scales to be much longer. The heat and salt transfer processes may take as long as 50 years for transfer down to the main thermocline and 1000 years for formation of the abyssal water beneath the main thermocline.

For the atmosphere, too, the nonadiabatic spin-up or spin-down processes are slower—radiative heat-transfer processes have time constants of the order of months—than the spin-down time of a few days associated with Ekman pumping. Moreover, the calculations in section 18.6.3 indicate that Ekman spin-down will tend only to reduce the barotropic component of the kinetic energy, that is, they reduce the winds by their surface values. Other processes must be involved in the decay of the winds aloft.

18.6.3 Spin-Down of Mesoscale Eddies

As a final example of frictional quasigeostrophic dynamics, we shall consider the effects of bottom friction on mesoscale ocean eddies. In the atmosphere, friction at the ground is an important part of the dynamics of synoptic scale motions. In the ocean, however, friction is considered to be less important because the bottom currents are relatively weak. Nevertheless, it is of interest to know how much of the water column is affected by bottom friction. We do know that the surface manifestations of mesoscale motions (in particular Gulf Stream rings) can persist for longer than 2 years (Cheney and Richardson, 1976). We shall show that this time scale is consistent with predictions of the simple baroclinic spin-down time.

Holton (1965) obtained a solution to the spin-down problem for a uniformly stratified fluid in a cylindrical container, showing that the effects of Ekman pumping are confined to a height $H_R \sim f_0 L/N$. Walin (1969) completed and extended Holton's analysis by analyzing in detail the effects of the side-wall boundaries and gave a simpler illustration of the spin-down process not involving side boundaries. We shall solve the analogue of Walin's problem for the variable stratification and radial symmetry characteristic of Gulf Stream rings.

We wish to solve (18.32)–(18.34) for the streamfunction $\psi(r,z,t)$, first for $\beta = 0$, assuming $w_E(0) = 0$, $w_E(-H) = (D_E/2)\nabla^2\psi(r,-H,t)$, and the initial condition $\psi(r,z,0) = \psi_0(r,z)$. The nonlinearities vanish because of

the radial symmetry. Taking a Fourier-Bessel transform of the streamfunction

$$\hat{\psi}(k,z,t) = \int_0^\infty r\,dr\,J_0(rk)\psi(r,z,t),$$

we find

$$\left(\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z} - k^2\right)\psi(k,z,t) = \left(\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z} - k^2\right)\hat{\psi}_0(k,z),$$

$$\hat{\psi}_{zt}(k,0,t) = 0,$$

$$\hat{\psi}_{zt}(k,-H,t) = \tfrac{1}{2}k^2 D_E f_0 S(-H)\hat{\psi}(k,-H,t).$$

Solving for $\hat{\psi}$ we obtain

$$\hat{\psi}(k,z,t) = \hat{\psi}_0(k,z) - \hat{\psi}_0(k,-H)F(z;k)(1 - e^{-\sigma(k)t}),$$

where $F(z;k)$ satisfies

$$\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}F = k^2F,$$

$$F_z(0;k) = 0,$$

$$F_z(-H;k) = 1,$$

and the inverse spin-down time is given by

$$\sigma(k) = -k^2 D_E f_0 S(-H)/2F_z(-H;k)$$
$$= [-k^2 S(-H)H/F_z(-H;k)]\sigma_{BT}$$

where $\sigma_{BT} = f_0 D_E/2H$ is the inverse barotropic spin-down time. Thus the inverse baroclinic spin-down time is simply related to σ_{BT} by the factor H divided by the penetration depth.

For large scales, the motion spins down uniformly throughout the whole column with $\sigma \simeq \sigma_{BT}$. For small scales, the spin-down occurs only over a depth H_R and is much more rapid. We expect, therefore, that the smaller scales will disappear from the deep ocean, perhaps leaving a thermocline signal behind, while the larger scales will decay more slowly but also more completely. In figure 18.12, we show the structures $F(z;k)$ and inverse spin-down times $\sigma(k)$ for various scales $1/k$. Absolute decay rates depend upon D_E—for D_E = 20 m, the time scale σ_{BT}^{-1} = 89 days, so that everything happens in a few months.

For application to rings we assume $\psi_0(r,-H) = -le^{(-1/2)(r^2/l^2-1)} \times 10\ \mathrm{cm\,s^{-1}}$, which gives maximum currents of 10 cm s^{-1} at a radius l. We solve for the net change in azimuthal velocity $\psi_r(r,z,t\to\infty) - \psi_{0r}(r,z)$ and contour this change in figure 18.13. It is seen that the changes in the thermocline and shallow water are negligible so that the persistence of oceanic thermocline eddies is quite consistent with theoretical expectations.

When the beta effect is included, important differences occur in the spin-down of linear eddies. The simplest case to analyze is for weak friction. Then

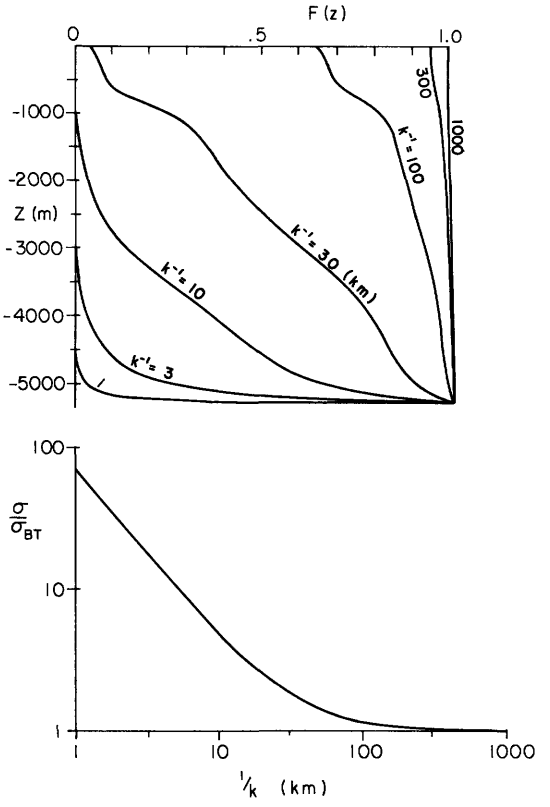


Figure 18.12 Decay in currents F as a function of depth for different radial scales k^{-1}. Actual change is given by $-F(z)$ × bottom currents. Lower figure shows ratio of decay rate to spin-down rate for a homogeneous fluid as a function of the radial scale.

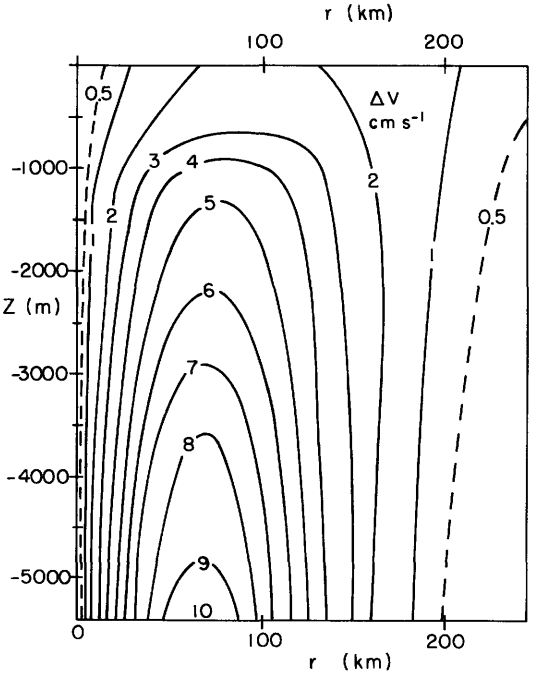


Figure 18.13 Decrease in azimuthal velocity due to bottom friction when initial bottom currents are 10 cm s$^{-1}(r/l)$ × $\exp[-\frac{1}{2}(r^2/l^2) + \frac{1}{2}]$.

there are two time scales—the period and the spin-down time. The Fourier-Bessel component with wavenumber k and (initially) vertical normal mode $F_n(z)$ behaves like

$$\hat{\psi}_n(k,t) = \hat{\psi}_n(k,0)F_n(z)J_0\left(k\sqrt{\left(x + \frac{\beta t}{k^2 + \lambda_n^2}\right)^2 + y^2}\right)$$
$$\times \exp[-\sigma_{BT}F_n^2(-H)k^2t/(k^2 + \lambda_n^2)].$$

This follows from the fact that when the solution of (18.33) is expanded in powers of period/spin-down time, the lowest order component is just a steadily propagating Bessel-function eddy. The next-order component has an inhomogeneous boundary condition due to friction and an inhomogeneous forcing of the equations of motion due to the slow time dependence. Multiplying this first-order equation of motion by $F_n(z)$ and depth averaging shows that the slow time dependence satisfies a simple exponential decay law (see also Flierl, 1978).

One important feature of this solution is the fact that baroclinic modes decay more slowly than barotropic modes both because of the increase in λ_n^2 and because of the appearance of the factor $F_n^2(-H)$. Thus a first-mode deformation-scale eddy with $F_1(-H) = -0.6$ and $k = \lambda$ has a decay rate of $0.2\sigma_{BT}$. But the important feature is that the β-plane eddy, unlike the f-plane eddy, decays completely. The β-effect permits transmission of energy downward, where it can be dissipated by friction. It appears that nonlinearity can impede this process because it slows down the dispersion of a ring (McWilliams and Flierl, 1979).

18.7 Nonlinear Motions

In this section, we shall consider mesoscale flows for which the advection of relative vorticity or density anomaly is important. This can occur either in the form of wave–wave interactions or wave–mean flow interactions. In both cases we are considering motions in which there are significant nonlinear interactions among various scales. This situation is to be contrasted with that in section 18.5.3, in which the mean flow provided a variable environment for the waves but was passive in the sense that there was no exchange of energy between the waves and the mean flow.

18.7.1 Baroclinic and Barotropic Instabilities

The problem of the instability of large-scale atmospheric motions has a long history, going back as always to Helmholtz (1888). The discoveries of the polar front and the polar-front wave by J. Bjerknes (1919) and J. Bjerknes and H. Solberg (1921, 1922) initiated several investigations of the instability of a polar-front model, notably by H. Solberg (1928) and by N. Kotschin (1932). These studies were incomplete: Solberg's avoided considering the effects of the frontal intersection with ground; Kotschin's considered various possible perturbation modes but not that of the all important baroclinic instability. E. Eliasen (1960) conducted a numerical study of a problem similar to Kotschin's, but with a vertical wall. However, the detailed exploration of Kotschin's model, a front between two fluids of different uniform densities and zonal velocities intersecting upper and lower horizontal boundaries, was left to Orlanski (1968), who considered all the four different instability modes—Helmholtz instability of vertical shear coupled with gravitational stability, Rayleigh instability of horizontal shear, baroclinic instability, and mixed baroclinic–barotropic–Helmholtz instability. Attempts to explain the long atmospheric waves observed in the troposphere were initiated by the work of J. Bjerknes (1937) to which we have already referred. Mathematical theories for the instability of a baroclinic zonal current with uniform horizontal temperature gradients were presented by Charney (1947), Eady (1949), Fjørtoft (1950), Kuo (1951), Green (1960), Burger (1962), Stone (1966, 1970) and many others—the problem is still being investigated. The stability of a horizontally shearing zonal current in two-dimensional spherical flow was studied by Kuo (1949). The stability of flows with both vertical and horizontal shear was investigated by Stone (1969), McIntyre (1970), Simmons (1974), Gent (1974, 1975) and Killworth (1980). The last named is the most comprehensive. Integral conditions for instability in more or less arbitrary zonal flows were developed by analogy with Rayleigh's condition for two-dimensional parallel flows by Kuo (1951), Charney and Stern (1962), Pedlosky (1964a, 1964b), Bretherton (1966a, 1966b), and others.

On the oceanic side, the onset of meandering of the western boundary currents has been dealt with by Orlanski (1969) and Orlanski and Cox (1973). They conclude that the meandering of the Gulf Stream between Miami and Cape Hatteras can be attributed to baroclinic instability, a result which seems to be in agreement with observations of Webster (1961a). The baroclinic instability of the free Gulf Stream extension implies a northward heat transport by the meanders and cutoff vortices. Evidence for such transports is not conclusive. The discovery of the mid-ocean mesoscale eddies initiated attempts by Gill, Green, and Simmons (1974) and Robinson and McWilliams (1974) to ascertain whether these eddies could be ascribed to baroclinic instabilities of the mid-ocean mean flows. The results have not been encouraging. Studies of the behavior of numerical ocean models also do not support this idea (Harrison and Robinson, 1978). If one merely converts available potential energy to kinetic energy while preserving the total energy density per unit area, the perturbation kinetic energies cannot exceed those

of the mean flow, and are therefore too small by an order of magnitude. Only ad hoc energy-convergence mechanisms give the right magnitudes.

Most of the studies referred to above have dealt with the instability of a zonal current with horizontal and/or vertical shear. Realistically, we must also be concerned with the instability of nonzonal and time-dependent flows, including oceanic gyres, forced and free Rossby waves and waves over topography. Thus we need to consider more general basic states.

We begin with the quasi-geostrophic potential vorticity equations (18.33)-(18.35). We attempt to find a basic solution $\bar{\psi}'(x,y,z,t)$ and investigate the growth of small perturbations $\psi'(x,y,z,t)$ around this basic state. The most straightforward basic state is a steadily translating (possibly at zero speed) unforced, nondissipative flow field

$$\bar{\psi} = \bar{\psi}(x',y,z), \qquad x' = x - \bar{c}t,$$

which satisfies the equations

$$\nabla^2\bar{\psi} + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\bar{\psi} + \beta y = P(\bar{\psi} + \bar{c}y,\ z) \tag{18.51}$$

and the boundary conditions

$$\bar{\psi}_z(x',y,0) = T_S(\bar{\psi} + \bar{c}y),$$

$$\bar{\psi}_z(x',y,-H) + f_0 S(-H)b(x' + \bar{c}t,y)$$

$$= T_b(\bar{\psi} + \bar{c}y). \tag{18.52}$$

Clearly such a solution is possible only if $\bar{c}b_x = 0$, that is, if the basic flow is independent of time or if the zonal variation in topography vanishes—waves cannot translate over varying topography without changing amplitude or shape. The basic flow is stationary in the x', y, z system and in this system the pseudopotential vorticity is constant along streamlines.

The derivation of (18.51) and (18.52) may indicate that the restrictions upon the mean flow are quite severe—no forcing or dissipation. However, our subsequent derivations will require only (18.51) and (18.52) and these can hold in much more general conditions. For example, the standard meteorological problem considers the instability of zonal flows forced by heating and perhaps Reynolds stresses and dissipated by radiation and surface Ekman pumping. Since both the mean flow and the potential vorticity are functions only of y, we can still define potential vorticity and surface functionals from (18.51)-(18.52). As long as the forcing and dissipative processes are not significant in the perturbation dynamics, the formalism below will apply. (We warn, however, that when there is topography or lateral boundaries, the stability problem for forced and dissipated flow may be quite different.)

The perturbation streamfunction $\psi' = \psi'(x',y,z,t)$ satisfies

$$\frac{\partial}{\partial t}\left(\nabla^2 + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\right)\psi'$$

$$+ (\bar{\mathbf{v}} - \bar{c}\hat{\mathbf{x}})\cdot\nabla\left(\nabla^2 + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z} - P'\right)\psi' = 0, \tag{18.53a}$$

$$P'(\Psi,z) = \frac{\partial}{\partial\Psi}P(\Psi,z),$$

$$\frac{\partial}{\partial t}\psi'_z + (\bar{\mathbf{v}} - \bar{c}\hat{\mathbf{x}})\cdot\nabla\left(\frac{\partial}{\partial z} - T'\right)\psi' = 0, \tag{18.53b}$$

$$z = 0, -H.$$

If we examine the normal modes $\psi'(x,y,z,t) = \psi'(x',y,z)e^{\sigma t}$, we have the eigenvalue equation for the growth rate

$$\sigma\left(\nabla^2 + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\right)\psi'$$

$$= -(\bar{\mathbf{v}} - \bar{c}\hat{\mathbf{x}})\cdot\nabla\left(\nabla^2 + \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z} - P'\right)\psi', \tag{18.54a}$$

with boundary conditions

$$\sigma\psi'_z = -(\bar{\mathbf{v}} - \bar{c}\hat{\mathbf{x}})\cdot\nabla\left(\frac{\partial}{\partial z} - T'_s\right)\psi', \qquad z = 0,$$

$$\sigma\psi'_z = -(\bar{\mathbf{v}} - \bar{c}\hat{\mathbf{x}})\cdot\nabla\left(\frac{\partial}{\partial z} - T'_b\right)\psi', \qquad z = -H. \tag{18.54b}$$

These equations for the perturbation streamfunction ψ' and the growth rate σ will form the basis for discussion of zonal flow instability and wave instabilities below.

Integral Theorems The classic example of an integral theorem is, of course, the Rayleigh theorem (1880). However, there is a slightly more general theorem, due originally to Arnol'd (1965) and applied to quasi-geostrophic flow by Blumen (1968), which we shall extend here to the problem of traveling disturbances and/or stationary motion over topography. This theorem states that the flow is *stable* if the potential vorticity and buoyancy along the bottom surface increase, and the buoyancy along the top surface decreases, with increasing streamfunction, that is, $P' \geq 0$, $T'_s \leq 0$, $T'_b \geq 0$ everywhere. To prove this, let us assume that $P' > 0$, $T'_s < 0$, and $T'_b > 0$ everywhere. (The cases for $P' = 0$ or $T'_s = 0$ or $T'_b = 0$ everywhere are readily proved.) First, we form an energy equation by multiplying (18.54a) by $-\psi'^*$, volume integrating, adding the conjugate equation, integrating by parts, and applying the boundary conditions. We obtain

$$(\sigma + \sigma^*)\iiint |\nabla\psi'|^2 + \frac{1}{S}|\psi'_z|^2$$

$$= \iiint q' J(\bar{\psi} + \bar{c}y,\ \psi'^*) + \text{c.c.} + \iint \psi'_z J(\bar{\psi} + \bar{c}y,\ \psi'^*)$$

$$+ \text{c.c.}\,|^0_{-H}. \tag{18.55}$$

Next, we form a normalized enstrophy equation by multiplying (18.54a) by q'^*/P' (recalling that $P' \neq 0$) and volume integrating to get

$$(\sigma + \sigma^*) \iiint \frac{|q'|^2}{P'}$$

$$= -[\iiint q'^* J(\bar{\psi} + \bar{c}y, \psi') + \text{c.c.}]. \tag{18.56}$$

Applying a similar procedure to the upper and lower boundary conditions, adding the result to (18.55) and (18.56), gives

$$(\sigma + \sigma^*) \iiint \left[|\nabla\psi'|^2 + \frac{1}{S} |\psi_z'|^2 + \frac{1}{P'} |q'|^2 \right.$$

$$\left. - \frac{1}{H} \frac{1}{T_s' S(0)} |\psi_z'(0)|^2 + \frac{1}{HT_b' S(-H)} |\psi_z'(-H)|^2 \right] = 0. \tag{18.57}$$

For the choice $P' > 0$, $T_s' < 0$, and $T_b' > 0$ the integrand is positive definite, implying that $\text{Re}(\sigma) = 0$, that is, that the flow is stable. When P', T_s', or T_b' are everywhere zero, the enstrophy or surface-temperature variance equations simply show that $|q'|$ or $|\psi_z'|$ at 0 or $-H = 0$, so that the term contributing to (18.55) can be ignored and therefore will also not enter in (18.57).

This completes the proof of the theorem. From the relation between the potential vorticity and the streamfunction (in the moving coordinate system) and the relation between the surface buoyancies and the streamfunctions at the top and bottom surfaces, we can tell whether the flow is stable or potentially unstable. In some problems (cf. Howard, 1964b; Rosenbluth and Simon, 1964) the necessary criterion for stability has been shown to be sufficient. We should also mention that the normal-mode assumption is not essential, so that the theorem applies to an arbitrary initial disturbance (Blumen, 1968).

In illustration, we note that the theorem implies that the Fofonoff (1954) inertial gyre solution,

$$P(\bar{\psi}) = \alpha\bar{\psi}, \qquad T_s(\bar{\psi}) = T_b(\bar{\psi}) = 0, \qquad \bar{c} = 0,$$

where α is a positive constant, is stable, as first pointed out by McWilliams (1977). We could find many other stable gyres by numerical means, including topographical effects, by solving (18.51), (18.52) with arbitrary functionals P and $T_{s,b}$ constrained only to satisfy the proper derivative conditions. The simplest would be to take

$$P(\bar{\psi}, z) = a(z)\psi + b(z),$$

with $a(z) > 0$ and similar linear functionals for the boundary conditions.

A second example is the flow forced by Gulf Stream meandering described in section 18.5.3. In this case, the potential vorticity equation for the forced wave (the basic state) is

$$\nabla^2\bar{\psi} + \beta y = P(\bar{\psi} + \bar{c}y).$$

The substitutions

$$\bar{\psi} = \Psi_0 e^{-\nu\mu} e^{i(kx - \omega t)}$$

and

$$\bar{c} = \omega/k$$

show that $P(Z) = \beta Z/\bar{c}$. Therefore, when the forcing propagates eastward, the trapped wave is *stable*. Unlike ordinary propagating Rossby waves, for which $\bar{c} < 0$, and which Lorenz (1972) has shown to be unstable, forced waves may be stable. We shall consider topographically forced waves in detail in section 18.7.3.

Zonal Flows We now specialize to zonal flows $\bar{\psi} = -\int^y \bar{u}(y', z)\,dy'$, $b_x = 0$. For these, we can readily find P' and $T_{s,b}'$ by taking y derivatives of (18.51) and (18.52):

$$P' = (\beta - \bar{u}_{yy} - \frac{\partial}{\partial z} \frac{1}{S} \frac{\partial}{\partial z} \bar{u})/(\bar{c} - \bar{u}),$$

$$T_s' = -\bar{u}_z/(\bar{c} - \bar{u})|_{z=0},$$

$$T_b' = (f_0 S b_y - \bar{u}_z)/(\bar{c} - \bar{u})|_{z=-H},$$

where $\bar{c}$ now is completely arbitrary (i.e., the perturbation wave speed will be simply doppler shifted by $\bar{c}$). In particular, we can choose $\bar{c}$ so that $\bar{c} - \bar{u}$ has a definite sign. Therefore we see that the flow will be stable if all the three quantities

$$\beta - \bar{u}_{yy} - \frac{\partial}{\partial z} \frac{1}{S} \frac{\partial \bar{u}}{\partial z},$$

$$\bar{u}_z(0),$$

$$f_0 S(-H) b_y - \bar{u}_z(-H)$$

have the same sign. Thus we recover the generalized Rayleigh theorem for quasigeostrophic flows: the flow is stable if

$$\bar{Q}_y = \beta - \bar{u}_{yy} - \frac{\partial}{\partial z} \frac{1}{S} \frac{\partial \bar{u}}{\partial z} + \frac{\bar{u}_z}{S} \delta(z)$$

$$+ \left(f_0 b_y - \frac{\bar{u}_z}{S} \right) \delta(z + H) \tag{18.58}$$

is uniform in sign (δ is the Dirac delta function). More conventional proofs of this theorem also can be found in Charney and Stern (1962), Pedlosky (1964a), and Bretherton (1966b).

A second standard theorem in shear flow instability theory due to Fjørtoft (1950) can also be generalized to the quasi-geostrophic flow problem. If we suppose that $\bar{Q}_y$ vanishes along some curve in the (y, z) plane and furthermore that $\bar{u} = \bar{u}_c$ = constant on this curve, the flow will be stable if $\bar{Q}_y(\bar{u} - \bar{u}_c)$ is negative everywhere. This can be demonstrated by choosing $\bar{c} = \bar{u}_c$. Clearly the requirement that $\bar{u} = \bar{u}_c$ at all points where

$$\beta - \bar{u}_{yy} - \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\bar{u} = 0$$

is highly restrictive (though it does occur for $\bar{u}_y = 0$ or $\bar{u}_z = 0$ or $\bar{u}_{yy} + (\partial/\partial z)(1/S)(\partial/\partial z)\bar{u} = K\bar{u}$).

As a practical application, we remark that the Rayleigh theorem (18.58) implies that the Eady (1949) problem (S = constant, $\bar{u}_z$ = constant, $\beta = 0$, $\bar{u}_y = 0$) can be stabilized by a sloping topography such that $b_y > \bar{u}_z/f_0 S|_{-H}$. This slope is steeper than the isopycnal slope, so that the density gradient at the bottom becomes opposite in sign to the gradient at the surface.

A second application is to demonstrate the stabilizing effort of β, especially for eastward flows. We consider zonal currents with a barotropic plus a sheared flow with the structure of the flat-bottom first-baroclinic mode

$$\bar{u}(y,z) = \bar{u}_{BT} + \bar{u}_{BC}F_1(z)$$

with $\bar{u}_{BT}$ and $\bar{u}_{BC}$ constants. (Many currents in the ocean do seem to have dominantly first-mode shears.) The Rayleigh criterion becomes

$$\bar{Q}_y = \beta + \frac{\bar{u}_{BC}}{L_R^2}F_1(z) > 0$$

for all z. This can occur only if

$$-\frac{\beta L_R^2}{|F_1(0)|} < \bar{u}_{BC} \equiv \frac{\Delta u}{|F_1(0)| + |F_1(-H)|} < \frac{\beta L_R^2}{|F_1(-H)|},$$

where Δu is the change in velocity from bottom to top. Using our N^2 profile this implies

$$-4 \text{ cm s}^{-1} < \Delta u < 22 \text{ cm s}^{-1}.$$

We see that eastward currents are considerably more stable than westward flows. Gill, Green and Simmons (1974) report on calculations which show weak growth rates for $\Delta\bar{u} \sim -5$ cm s^{-1}. Observations of actual $\Delta\bar{u}$'s are not readily available because the midocean density-field measurements are generally contaminated with eddies. However, it is not unlikely that mid-ocean mean currents away from the "recirculation region" of Worthington (1976) (see also chapters 1 and 3) are smaller than this magnitude, so that mid-ocean flows may very possibly be stable (see also McWilliams, 1975).

This result must be viewed with caution, because it is possible for forced meridional currents to be locally unstable for any value of the shear. We can see this by considering the stability of a mean flow

$$\bar{\psi} = \bar{v}(z)x - \bar{u}(z)y,$$

where we ignore the dynamics of the mechanism that supports the $\bar{v}$ component of flow on the grounds that its space and time scales are much larger than those of the perturbations we wish to consider. The perturbations satisfy

$$\frac{\partial}{\partial t}q' + J(\bar{\psi},q') + J(\psi',\bar{q}) = 0,$$

where

$$\bar{q} = \left(\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\bar{v}\right)x + \left(\beta - \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\bar{u}\right)y$$

is now *not* expressible as $P(\bar{\psi},z)$. However, we may consider perturbations of the form

$$\psi' = F(z)\exp[ik(x - ct) + ily]$$

to find

$$\left[(\bar{u} - c) + \frac{l}{k}\bar{v}\right]\left(\frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z} - k^2 - l^2\right)F$$
$$+ \left[\beta - \frac{\partial}{\partial z}\frac{1}{S}\frac{\partial}{\partial z}\left(\bar{u} + \frac{l}{k}\bar{v}\right)\right]F = 0.$$

Applying the usual Rayleigh theorem shows that the flow will be stable unless $\beta - (\partial/\partial z)(1/S)(\partial/\partial z)[\bar{u} + (l/k)\bar{v}]$ changes sign. If $\bar{v} \neq 0$, however, a proper choice of l and k (the direction of the perturbation wave) may always be made to ensure satisfying the necessary criterion for instability. Thus arguments about the zonal flow stability may not directly apply to the Sverdrup circulation.

The discussion of baroclinic instability has been extended to finite amplitudes by Lorenz (1962, 1963a) using truncated spectral expansions and by Pedlosky (1970, 1971, 1972, 1976, 1979b), Drazin (1970, 1972), and others using expansion techniques in the vicinity of critical values of the stability parameters. Thus far, the systems dealt with have been more applicable to laboratory models than the actual atmosphere or ocean. A general review has been given by Hart (1979a), who himself has contributed by experiment and analysis to the subject.

18.7.2 Wave–Mean Flow Interactions

The subject of wave–mean flow interaction in the atmosphere has been treated extensively in connection with the manner in which large-scale waves generated in the troposphere propagate vertically into the stratosphere and there interact with the mean flow. One example is the so-called sudden-warming phenomenon, the rapid breakdown of the stratospheric winter circumpolar cyclone accompanied by large-scale warming. Another example is the so-called quasi-biennial oscillation, which has been explained as a wave–mean flow interaction between vertically propagating Rossby-gravity and Kelvin waves and the zonal flow in the equatorial stratosphere (Lindzen and Holton, 1968; Holton and Lindzen, 1972). A vivid experimental and theoretical demonstration of this type of interaction has been given by Plumb and McEwan (1978).

Charney and Drazin (1961) have shown that small-amplitude steady waves in quasi-geostrophic, adiabatic, inviscid flow cannot interact to second order with the zonal flow. If there are no critical surfaces at which the zonal flow vanishes and there is no dissipation, forcing, or transience, no interaction will take place. All are present in the quasi-biennial oscillation and in Plumb and McEwan's model. The result of Charney and Drazin was originally derived by straightforward calculation. It may also be inferred from an independent study of energy transfer in stationary waves by Eliassen and Palm (1960), who derive linear relations between the horizontal Reynolds stress, the horizontal eddy heat flux, and the components of the wave energy flux. These works have been greatly extended by Andrews and McIntyre (1976), Boyd (1976), and Andrews and McIntyre (1978a,b). McIntyre (1980) reviews the subject.

There have been several suggestions of oceanic analogies: Pedlosky (1965b) and N. Phillips (1966b) have argued that westward-propagating Rossby waves can cause acceleration of the western boundary currents. Lighthill (1969) attempted to explain the onset of the Somali Current as due to the interaction of Rossby-gravity waves generated by the monsoon winds in the mid-Indian Ocean with the flow in the vicinity of the East African continent. More recently, experiments of Whitehead (1975) have shown quite clearly that mean flows may be generated by radiated Rossby waves. His work led Rhines (1977) to a theoretical reconsideration of the wave-mean flow generation problem not only when the geostrophic contours (the f/H lines which represent the streamlines for free inertial motions) are closed or periodic but also when the contours are open. Rhines's work is important for understanding large-scale forced motions in oceanic basins.

As an illustration of wave-mean flow interaction in an oceanographic context we shall ask again whether the waves produced by Gulf Stream meandering may be responsible for generating and maintaining the so-called recirculation flow found by Worthington (1976) and others. This flow occurs in a region extending some 1000 km south of the stream and contains (according to Worthington) a sizable westward transport (10^8 $m^3 s^{-1}$). This problem has been addressed by Rhines (1977), who, however, did not consider generation due to eastward-moving waves.

We consider the barotropic flow south of the Gulf Stream forced by the streamfunction $\psi(x,0,t) = A\cos(kx - \omega t)$, as in section 18.5.3, but we now include the effects of bottom Ekman friction and the second-order interaction with the mean zonal flow. The streamfunction satisfies

$$\left(\frac{\partial}{\partial t} + \sigma_{BT}\right)\nabla^2\psi + \beta\psi_x = -J(\psi,\nabla^2\psi),$$

$$\psi(x,0,t) = A\cos(kx - \omega t),$$

$$\psi \to 0, \qquad y \to -\infty.$$

The linear solution (assuming Ak^3/β small) will be

$$\psi = \mathrm{Re}\{A\exp[i(kx - \omega t) + \nu y]\},$$

$$\nu = \sqrt{k^2 + \beta\omega/(\omega^2 + \sigma_{BT}^2) - i\beta k\sigma_{BT}/(\omega^2 + \sigma_{BT}^2)},$$

if the root with positive real part is chosen to satisfy the radiation condition. The nonlinearly forced streamfunction field satisfies

$$\left(\frac{\partial}{\partial t} + \sigma_{BT}\right)\nabla^2\psi^{(1)} + \beta\psi_x^{(1)} = -2A^2\nu_r^2\nu_i k e^{2\nu_r y}$$

where ν_r and ν_i are the real and imaginary parts of ν, respectively. Its solution is

$$\psi^{(1)} = -\frac{A^2\nu_i k}{2\sigma_{BT}}\, e^{2\nu_r y}$$

or

$$\bar{u} = \frac{A^2\nu_r\nu_i k}{\sigma_{BT}}\, e^{2\nu_r y}.$$

This is, of course, just the solution to

$$\sigma_{BT}\bar{u} = -(\overline{u'v'})_y$$

with u' and v' taken from the lowest-order solution.

The mean flow is determined by a balance between friction and Reynolds-stress forcing. The importance of dissipation becomes clear: without friction, ν is either purely real or purely imaginary and $\bar{u} = 0$. With friction, we find that the waves transfer momentum into the mean flow. Moreover, we can show that the magnitude of the flow is not sensitive to the spin-down time $1/\sigma_{BT}$, as this time becomes very large.

As σ_{BT} becomes small we find

$$\nu \simeq \begin{cases} \sqrt{k^2 + \dfrac{\beta k}{\omega}} - i\beta\sigma_{BT}k/2\omega^2\sqrt{k^2 + \dfrac{\beta k}{\omega}}, & \dfrac{\omega}{k} > 0, \\ \omega/k < -\beta/k^2 \\ -i\sqrt{-k^2 - \dfrac{\beta k}{\omega}} + \beta\sigma_{BT}k/2\omega^2\sqrt{k^2 - \dfrac{\beta k}{\omega}}, \\ -\beta/k^2 < \omega/k < 0. \end{cases}$$

The forced mean flow is therefore

$$\bar{u} = -\frac{\beta A^2 k^2}{2\omega^2}\begin{cases} \exp\left(2\sqrt{k^2 + \dfrac{\beta k}{\omega}}\, y\right), & \dfrac{\omega}{k} > 0, \\ \omega/k < \beta/k^2 \\ \exp\left(\beta\sigma_{BT}ky/\omega^2\sqrt{-k^2 - \dfrac{\beta k}{\omega}}\right), \\ -\beta/k^2 < \omega/k < 0, \end{cases}$$

with amplitude independent of σ_{BT}. We can estimate the westward current speeds by relating the amplitude A to the excursions of the stream in the y direction:

$$d = -A\frac{k}{\omega}\cos(kx - \omega t) \equiv d_0\cos(kx - \omega t).$$

The maximum westward currents are $-\frac{1}{2}\beta d_0^2$. Rhines (1977) has derived from more general considerations the result that mean-flow generation is proportional to β times the square of the displacement. For typical excursions of 100–200 km, mean flows of 10–40 cm s^{-1} can be generated. [We should note that, for this problem, the eastward Stokes drift is given by

$$A^2\frac{k}{\omega}\left(k^2 + \frac{\beta k}{\omega}\right)\exp\left(2\sqrt{k^2 + \frac{\beta k}{\omega}}\,y\right),$$

which is larger than the westward Eulerian flow so that the particle drift is *eastward.*]

Observations of Gulf Stream meanders usually indicate eastward-moving disturbances; therefore much of the mean flow will be trapped in a distance one-half that shown in figure 18.12. The disturbances that generate propagating waves ($-\beta/k^2 < \omega/k < 0$) can produce mean flows over large north–south distances, but there does not seem to be enough amplitude in such disturbances. (See, however, the remarks in section 18.8)

This very simple calculation indicates that eddy radiation from the meandering Gulf Stream can generate a return flow with speeds comparable to those suggested by observations (cf. Worthington, 1976; Wunsch, 1978a; Schmitz, 1977; see also chapter 4). The predicted north–south scale of the region is quite small, however, unless there is considerably more energy in westward-going meanders than has been suggested by Hansen (1970) or by Robinson, Luyten, and Fuglister (1974).

There is another form of wave–mean flow interaction involving overreflection of waves traveling through a variable mean-flow field. Lindzen and Tung (1978) recently have demonstrated that barotropic and baroclinic instabilities may be explained as overreflection phenomena in which Rossby waves impinging upon a critical surface are reflected with a coefficient of reflection greater than unity. The combination of an overreflecting region in the mean flow with a reflecting boundary can lead to a growing disturbance in which the wave picks up energy at each passage into the overreflecting region.

This concept may be directly applicable to the problem of reflection of Rossby waves from the western boundary currents.[6] Numerous examples of Gulf Stream rings interacting with the Gulf Stream without being absorbed can be found in the data presented by Lai and Richardson (1977), and at times they appear to increase in energy as a result of the interaction (Richardson, Cheney, and Mantini, 1977). We suggest the possibility that overreflection may be involved in the dynamics of mesoscale eddies near the western boundary current. Whether or not this is so remains to be seen.

18.7.3 Wave Instability and Form-Drag Instability

The fact that Rossby waves may be unstable was first shown by Lorenz (1972) for a barotropic atmosphere. In a more detailed exploration of the problem Gill (1974) observed that there are two distinct mechanisms for the instability: a resonant triad interaction or a shear instability of the Rayleigh type. Duffy's (1978) and Kim's (1978) baroclinic studies showed that baroclinicity may also cause instability in large-scale waves. As in the instability of zonal flow, the growing baroclinic modes have the scale of the radius of deformation. Jones (1978) and Fu and Flierl (1980) have explored these ideas further as they apply to the ocean.

Wave and Form-Drag Instabilities Just as a freely propagating wave provides variations of potential vorticity which may lead to instability, topography may produce a forced flow whose variations of potential vorticity may also cause instability. Topography may be a destabilizing influence, either because the forced flow is unstable, just as a free wave, to Rayleigh or resonant instabilities, or else because the topography itself may help—via the form drag produced by the perturbation—to extract energy from the mean flow. The latter type of instability was first encountered by Charney and DeVore (1979) in their study of blocking (the persistence of anomalously high pressure in certain regions of the atmosphere) in a barotropic atmosphere. In their model, blocking occurs as an alternative flow equilibrium corresponding to a given forcing of the zonal flow in the presence of sinusoidal topography. It was found that the transition from the normal flow state to the anomalous blocking state takes place via a form-drag instability of an intermediate equilibrium state in which the zonal flow is superresonant. In this superresonant state, a small decrease of the zonal flow amplifies the forced orographic wave and increases the form drag (mountain torque), which in turn decelerates the zonal flow still further. Charney and Straus (1980) extended this study to a two-layer baroclinic atmosphere. Here again there is a form-drag instability. But when there is no lower layer flow, the instability is catalyzed by the form drag: the perturbation derives its energy from the available potential energy of the Hadley circulation generated by thermal forcing; the form drag merely establishes the necessary phase relationships.

The connections between this form-drag instability and the more familiar resonant or Rayleigh instabilities have not been previously explored. For this purpose, consider the simplest wavelike flow of a homogeneous

ocean and derive the instability conditions for both a free zonally propagating wave and a topographically forced Rossby wave in order to elucidate the similarities and differences among the respective instability mechanisms. We begin by stating the forms of the potential vorticity functionals $P(\bar{\psi} + \bar{c}y)$ and the resulting perturbation equations. For the Rossby wave, $\bar{c} = -\beta/k^2$, the potential vorticity functional in equation (18.51) is

$$P(Z,z) = (\beta/\bar{c})Z = -k^2 Z$$

with streamfunction

$$\bar{\psi} = A \sin kx.$$

The perturbation equation (18.53a) is particularly simple because P' is now a constant:

$$\sigma \nabla^2 \psi' + J(-\frac{\beta}{k^2} y + A \sin kx, (\nabla^2 + k^2)\psi') = 0, \quad (18.59)$$

where A is completely arbitrary.

For barotropic flow over topography of the form $b = b_0 \sin kx$ on the β-plane, the basic state potential vorticity equation

$$\nabla^2 \bar{\psi} + \beta y + \frac{f_0 b_0}{H} \sin kx = P(\bar{\psi})$$

has the solution

$$\bar{\psi} = -\bar{u}y + A \sin kx, \quad A = \frac{\bar{u} f_0 b_0 / H}{\bar{u}k^2 - \beta}$$

with the linear potential vorticity functional

$$P(Z) = -\frac{\beta}{\bar{u}} Z \equiv -k_u^2 Z$$

Here k_u is the wavenumber of the stationary wave (k_u^2 may be less than zero). The perturbation equation

$$\sigma \nabla^2 \psi' + J\left[-\frac{\beta}{k_u^2} y + A \sin kx, (\nabla^2 + k_u^2)\psi'\right] = 0 \quad (18.60)$$

is very similar in form to (18.59) except that k and k_u are now independent; however, the amplitude A for the topographical problem is determined by

$$A = \frac{f_0 b_0 / H}{k^2 - k_u^2}.$$

Obviously we need only to solve (18.60) and find $\sigma(\beta, A, k, k_u)$; we can then identify both the free ($k_u = k$) and forced ($k_u \neq k$) regimes. In actuality the task is even simpler since dimensional considerations show that there are only two parameters, $M = Ak^3/\beta$ and $\mu = k_u/k$, that must be varied while computing the nondimensional growth rate σ/Ak^2.

First, however, we demonstrate that only the $k_u^2 > 0$ case need be considered, since the flow is stable for $k_u^2 < 0$. This follows readily from (18.57):

$$(\sigma + \sigma^*)\left(\iint |\nabla \psi'|^2 - \frac{1}{k_u^2} |\nabla^2 \psi|^2\right) = 0,$$

which implies that Rossby waves $k_u^2 = k^2 > 0$ and waves forced by eastward flow may be unstable, while waves forced by westward flow $k_u^2 < 0$ are definitely stable. Again, as in section 18.7.1 we see that stable waves are generated when the relative motion of the forced wave with respect to the ambient flow is eastward. In addition, for eastward flow over topography or westward-propagating free Rossby waves, a necessary condition for instability is that the perturbation have components with scales larger than k_u^{-1} and components with scales smaller than k_u^{-1}. This follows from Fourier analyzing ψ' and substituting in (18.57) to get

$$\iint d\mathbf{s}\, |\mathbf{s}|^2 |\hat{\psi}'(\mathbf{s})|^2 \left[1 - \frac{|\mathbf{s}|^2}{k_u^2}\right] = 0,$$

which implies that $|\hat{\psi}'(\mathbf{s})|$ must be nonzero both for some values of $|\mathbf{s}|^2$ larger than k_u^2 and some values smaller than k_u^2.[7] This is the perturbation form of Fjørtoft's (1953) result on energy cascades.

We could readily solve the stability problem (18.60) using the mathematical techniques of Lorenz (1972), Gill (1974), Coaker (1977), or Mied (1978) to investigate growth rates. Alternatively we could discuss the limiting behavior—the Rayleigh limit for $M = Ak^3/\beta \gg 1$, the resonant interaction limit for $M \ll 1$, and the form-drag instability—by separate approximations. We shall do this for the last-named problem because it represents a relatively unfamiliar phenomenon. However, to ascertain most clearly the connections between the various types of behavior, it is most simple to employ the Fourier expansion

$$\psi' = e^{il_0 y}[A_0 e^{ik_0 x} + A_{-1} e^{i(k_0 - k)x} + A_1 e^{i(k_0 + k)x} + \cdots]$$

and truncate to the three indicated terms (cf. Gill, 1974). The resulting dispersion relation becomes just

$$(\omega K_{-1}^2 + B_{-1})(\omega K_0^2 + B_0)(\omega K_1^2 + B_1) = \frac{l_0^2 A^2 k^2}{4}(k_u^2 - K_0^2)[(k_u^2 - K_{-1}^2)(\omega K_1^2 + B_1) + (k_u^2 - K_1^2)(\omega K_{-1}^2 + B_{-1})] \quad (18.61)$$

using the notation $\omega = i\sigma$, $K_n^2 = (k_0 + nk)^2 + l_0^2$, $B_n = \beta(k_0 + nk)(k_u^2 - K_n^2)/k_u^2$. For Rossby waves we simply replace k_u by k.

We have sketched the dependence of $\hat{\sigma} = \sigma/Ak^2$ upon k_0 and l_0 for various M and μ values in figure 18.14. When $M \gg 1$, there is a broad area of (k_0, l_0) space in which the growth rates are real. The maximum occurs for $k_0 = 0$ (this is really a form of Squire's theorem) and

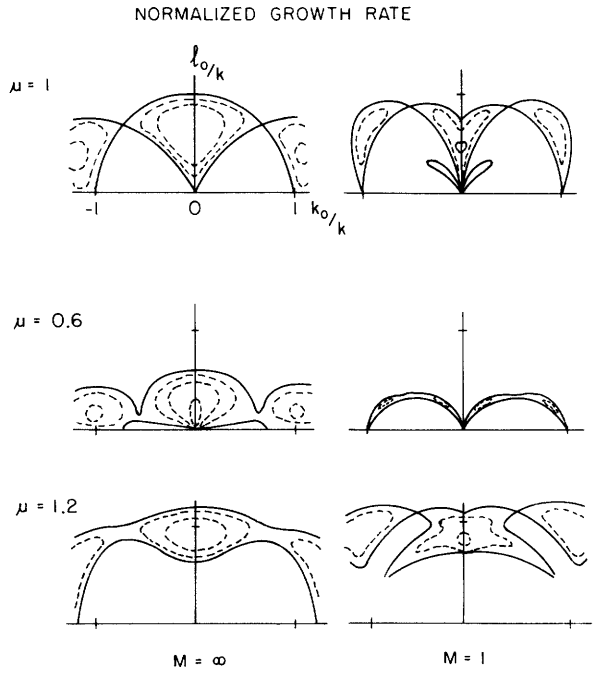


Figure 18.14 Growth rate divided by Ak^2 as a function of k_0, l_0 for various ratios of the stationary wavenumber to the topographic wavenumber μ and wave steepnesses M. The solid lines are the zero contour. The dashed contours, separated by the interval 0.1, correspond to positive growth rates.

the instability equation is similar to the barotropic instability equation. As M decreases the instability ($\hat{\sigma}$ real) becomes restricted more and more to a band around the frequency resonance line (see below). When μ decreases from 1 to a quantity smaller than 1 ($\bar{u} > \beta/k^2$), the point of maximum growth rate for $M = \infty$ moves to smaller wavenumbers l_0. For finite M, the same thing occurs—the largest growth rate moves towards $k_0 = l_0 = 0$ as μ decreases. On the other hand, when μ increases from 1 the forced wave becomes stable for small k_0 and l_0, even for small M, and the wavenumber for maximum growth migrates to smaller scales.

We are thus led to consider the three limits for (18.61):

Strong waves (A or M large): When A is large, the frequency ω is order A and we can neglect all of the B's to find

$$\sigma = \frac{A|l_0 k|}{2}\sqrt{\frac{-(k_u^2 - K_0^2)[K_1^2(k_u^2 - K_{-1}^2) + K_{-1}^2(k_u^2 - K_1^2)]}{K_0^2 K_1^2 K_{-1}^2}}.$$

Clearly the flow will be unstable for $K_0^2 < k_u^2 < K_1^2$, K_{-1}^2; this will always be possible by proper choice of l_0 and k_0. This is just the Rayleigh instability of the shear flow corresponding to the wave field. The maximum growth rate occurs at $k_0 = 0$:

$$l_0 = \begin{cases} (-1 + \sqrt{\mu^2 + \mu^4})^{1/2}, & \mu > \left(\dfrac{\sqrt{5}-1}{2}\right)^{1/2} \\ 0, & \mu < \left(\dfrac{\sqrt{5}-1}{2}\right)^{1/2}. \end{cases} \tag{18.62}$$

However, in the topographic case, A can be large not only for very strong topography ($f_0 b_0 k/H \gg \beta$), but also because the zonal flow is nearly critical. In the latter case $\mu \simeq 1$, so that the free and forced wave instabilities are indistinguishable. If the flow is not critical, but rather is forced by strong topography, the formula shows the maximum growth rate occurs at $k_0 \to 0$, $l_0 \to 0$.

Weak waves (A very small or M small): Here the critical condition is that two of the roots of the left-hand side of (18.61)—for example, $-B_0/K_0^2$ and $-B_1/K_1^2$—coalesce. This resonance condition

$$B_0/K_0^2 = B_1/K_1^2$$

or

$$\frac{k_0(k_u^2 - K_0^2)}{k_u^2 K_0^2} = \frac{(k_0 + k)(k_u^2 - K_1^2)}{k_u^2 K_1^2}$$

permits the order A part of the frequency to be complex since

$$K_0^2 K_1^2(\omega + B_0/K_0^2)^2 \simeq \frac{l_0^2 A^2 k^2}{4}(k_u^2 - K_0^2)(k_u^2 - K_1^2)$$

has a complex root for $K_0^2 < k_u^2 < K_1^2$. The growth rate is

$$\sigma = \frac{A|l_0 k|}{2}\sqrt{\frac{(k_u^2 - K_0^2)(K_1^2 - k_u^2)}{K_0^2 K_1^2}}.$$

However, the resonance condition must also hold. We can relate this to the more familiar wave-resonance conditions by defining the intrinsic frequencies of the two components of the perturbation and also that of the mean flow by

$$\hat{\omega}_0 = -\beta k_0/K_0^2,$$

$$\hat{\omega}_1 = -\beta(k_0 + k)/K_1^2,$$

$$\hat{\omega} = -\beta k/k_u^2 = -\bar{u}k.$$

The resonance conditions

$$(k_0 + k, l_0) = (k_0, l_0) + (k, 0),$$

$$\hat{\omega}_1 = \hat{\omega}_0 + \hat{\omega}$$

both hold. We show the resonance curves and growth rates for various values of k_u/k in figure 18.15.

For large μ (small mean flow speed), the resonant interaction (except for large l_0) is really a triplet interaction between the two perturbation waves and the topography itself (rather than the forced wave). This is, of course, a stable situation. As μ becomes less than

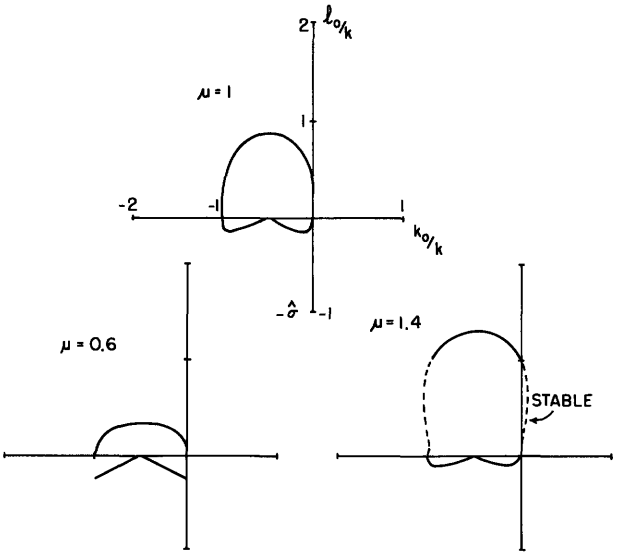


Figure 18.15 Curves above x axis show relation between k_0 and l_0 required by resonance condition. Curves below axis are plots of $-\hat{\sigma}$, showing the dependence of the growth rate upon k_0.

about 0.9, however, the maximum growth rate again occurs as $k_0 \to 0$ and $l_0 \to 0$ [$l_0 \sim (\mu^2 k_0/1 - \mu^2)^{1/2}$].

Form-drag instability: Thus in either case we are led to consider what will be shown to be a form-drag instability—the nonzero growth rates occurring at small k_0 and l_0—when $\mu < 0.79$. There is some difficulty here since the origin is a singularity for M finite; this problem would be eliminated in a bounded geometry. For convenience, we will take the limits $k_0 \to 0$ first and then $l_0 \to 0$ since this case has a simple physical interpretation. Applying these limits to (18.61) gives the frequency:

$$\omega^2 = \left[\frac{\beta}{k}\frac{k_u^2 - k^2}{k_u^2}\right]^2 + \frac{A^2 k_u^2}{2}(k_u^2 - k^2). \tag{18.63}$$

The flow is unstable when the right-hand side is negative, which cannot occur for Rossby waves $k_u = k$, but may occur for topographic waves when $k_u < k$ or $\bar{u}$ is greater than the critical speed β/k^2. In fact, the range is

$$\beta/k^2 < \bar{u} < \beta/k^2\left(1 + \left[\frac{1}{2}\left(\frac{f_0 b_0}{H}\right)^2 \frac{k^2}{\beta^2}\right]^{1/3}\right).$$

(The resonant triad instability for $\mu < 1$ does not appear here; rather, the limit $k_0 \to 0^-$ and $l_0 \sim (-k_0)^{1/2} \to 0^+$ must be used.) So far we have looked at the mathematics; let us now discuss the physics of this instability and also show that the truncation to three terms is valid.

The form-drag instability involves one component A_0 which has very large x and y scales and two components with the same scale as the topography. Examination of the individual amplitudes shows that $A_0 \sim 1/l_0$, so that $A_0 e^{il_0 y}$ contributes a term in the perturbation streamfunction which is proportional to y—a modification of the zonal mean flow. This suggests an alternative approach, which is to consider the zonal x-averaged momentum equation and the equation for the deviations. We begin with the quasi-geostrophic equations for a homogeneous fluid,

$$f_0 v = p_x,$$

$$f_0 u = -p_y,$$

$$u_t + (uu)_x + (uv)_y - \beta y v - f_0 v^{(1)} = -p_x^{(1)},$$

$$v_t + (uv)_x + (vv)_y + \beta y u + f_0 u^{(1)} = -p_y^{(1)},$$

$$u_x^{(1)} + v_y^{(1)} - \frac{1}{H}[(ub)_x + (vb)_y] = 0,$$

and consider the zonally averaged equations

$$\langle v \rangle = 0,$$

$$\langle u \rangle_t + \langle uv \rangle_y = f_0 \langle v^{(1)} \rangle,$$

$$\langle v^{(1)} \rangle_y = \frac{1}{H}\langle vb \rangle_y.$$

If the topography vanishes at some y far from the region of interest [following the arguments suggested by Hart (1979b), who showed that the Charney–DeVore truncated spectral problem was identical to that of forced flow over topography varying slowly with y], we can integrate the last equation to find

$$\langle u \rangle_t + \langle uv \rangle_y = \frac{f_0}{H}\langle vb \rangle.$$

The vorticity equation can be used to find the x-dependent part of the flow. In particular, if we assume the y scale is very large, we can drop all y derivatives to get two coupled equations:

$$\langle u \rangle_t = \frac{f_0}{H}\langle vb \rangle,$$

$$v_{xt} + \langle u \rangle v_{xx} + \beta v = -\frac{f_0}{H}\langle u \rangle b_x.$$

For the topography $b = b_0 \sin kx$, we have a steady solution

$$\langle u \rangle = \bar{u},$$

$$v = Ak \cos kx.$$

The deviations from this state satisfy

$$\langle u \rangle_t' = \frac{1}{2}\frac{f_0 b_0}{H}\langle v' \sin kx \rangle,$$

$$v_{xt}' + \bar{u} v_{xx}' + \beta v' = k\left(Ak^2 - \frac{f_0 b_0}{H}\right)\langle u \rangle' \cos kx$$

$$= k\frac{f_0 b_0}{H}\frac{\beta}{\bar{u}k^2 - \beta}\langle u \rangle' \cos kx,$$

which may be solved explicitly to give the dispersion relation (18.63). Here too one sees that it is the coupling between the change in the zonal flow induced by the wave drag and the change in the waves due to changes in zonal flow which leads to the instability. If we decrease the mean flow for a supercritical case (i.e., if we take $\langle u \rangle'$ to be negative), we produce low vorticity on the upwind slopes of the topography and high vorticity on the lee slopes. Associated with this vorticity change is high pressure on the upslope side of the mountains and low pressure on the downslope. This pressure pushes eastward on the topography so that the topography pushes westward on the fluid and decelerates the mean flow still further.

Flow in the Presence of Topography The previous section has described the influence of wavy topography upon the stability of the flows that go over it. However, there also exists topography that does not alter the mean-flow structure, either because the mean current is parallel to the topographic contours or because the currents occur only at levels above the peaks of the topography. In this section, we shall show that the stability of a parallel mean flow in the presence of topography can be quite different from that of the identical mean flow in a flat-bottomed ocean. We have been guided by the result of Charney and Straus (1980), who show that the form-drag instability can catalyze the release of available potential energy in a baroclinic shearing flow that would be stable in the absence of topography. In their study of multiple equilibria and stability in forced baroclinic flow over topography, they found that form-drag instability may occur for weaker thermal driving than conventional baroclinic instability, and that this type of instability leads to transition from one finite-amplitude, quasi-stationary equilibrium state to another. Baroclinic and barotropic instabilities of the stationary topographically perturbed flows give rise to westward-propagating, vacillating wave motions with periods of the order of 5 to 15 days. They suggest that the form-drag instability leads to transition from one stationary regime to another and that the observed westward- and eastward-propagating long planetary waves (zonal wavenumbers 1–4) are the propagating instabilities associated with these stationary regimes.

The simplest and most obvious example of the destabilizing effect of topography is the case of zonal barotropic flow with meridionally varying topography. The topography alters the effective value of β and thereby the growth rates and stability criteria: even though the energy source remains the horizontal shear, the topography can alter the possibility of extracting this energy. In particular, the Charney–Stern necessary criterion for instability [that $\beta - \bar{u}_{yy} + (f_0/H)\, b_y$ must change sign in the domain] suggests that instability may occur for lower values of shear when $b_y \neq 0$. The necessary condition may, of course, not be sufficient; in particular, when $\bar{u}_y = 0$, the flow will be stable even if $\beta + (f_0/H)b_y$ changes sign. However, in the case of sinusoidal $\bar{u}(y)$ and $b(y)$, the necessary condition seems also to be sufficient (using a simple truncated expansion in y), and the topography does destabilize the flow.

DeSzoeke (1975) discussed baroclinic flow over meridionally varying topography and found that the topography destabilizes the flow at some wavenumbers by a resonant instability involving two baroclinic waves which happen to travel at the same speed. Similar effects can be identified in the work of Durney (1977). We would like to focus our discussion, however, on the specific problem of destabilization by form-drag instability of a baroclinic flow which is neutrally stable in the absence of topography.

We shall consider the conventional two-layer model whose governing equations (cf. Pedlosky, 1979b) are

$$\left(\frac{\partial}{\partial t} + \mathbf{v}_1 \cdot \nabla\right) \times \left[\nabla^2 \psi_1 - \frac{\lambda_1^2}{1+\delta}(\psi_1 - \psi_2) + \beta y\right] = 0,$$

$$\left(\frac{\partial}{\partial t} + \mathbf{v}_2 \cdot \nabla\right) \times \left[\nabla^2 \psi_2 - \frac{\delta\lambda_1^2}{1+\delta}(\psi_2 - \psi_1) + \beta y + \frac{f_0}{H}(1+\delta)b\right] = 0,$$

where $\mathbf{v}_1$ and ψ_1 are the velocity and streamfunction in the upper layer and $\mathbf{v}_2$ and ψ_2 the corresponding quantities in the lower layer, δ is the ratio of the upper to the lower layer mean depths, and λ_1^{-1} is the layered version of the first baroclinic mode deformation radius $\lambda_1^2 = f_0(1+\delta)^2/g(\Delta\rho/\rho)H\delta$. We write the x-averaged equations

$$\frac{\partial}{\partial t}\left[-\bar{u}_{1yy} + \frac{\lambda^2}{1+\delta}(\bar{u}_1 - \bar{u}_2)\right] + \frac{\partial^2}{\partial y^2}\overline{\psi_{1x}'\left(\frac{\partial^2}{\partial y^2}\psi_1' + \frac{\lambda^2}{1+\delta}\psi_2'\right)} = 0,$$

$$\frac{\partial}{\partial t}\left[-\bar{u}_{2yy} + \frac{\delta\lambda^2}{1+\delta}(\bar{u}_2 - \bar{u}_1)\right] + \frac{\partial^2}{\partial y^2}\overline{\psi_{2x}'\left(\frac{\partial^2}{\partial y^2}\psi_2' + \frac{\delta\lambda^2}{1+\delta}\psi_1'\right)} = -\frac{f_0}{H}(1+\delta)\frac{\partial^2}{\partial y^2}\overline{\psi_{2x}' b},$$

and the equations for the deviations

$$\left(\frac{\partial}{\partial t} + \bar{u}_1\frac{\partial}{\partial x}\right)\left[\nabla^2\psi_1' - \frac{\lambda^2}{1+\delta}(\psi_1' - \psi_2')\right] + \left[\beta + \frac{\lambda^2}{1+\delta}(\bar{u}_1 - \bar{u}_2)\right]\psi_{1x}'$$

$$+J\left[\psi_1', \nabla^2\psi_1' - \frac{\lambda^2}{1+\delta}(\psi_2' - \psi_1')\right]$$

$$-\frac{\partial}{\partial y}\overline{\psi_{1x}'(\psi_{1yy}' + \frac{\lambda^2}{1+\delta}\psi_2')} = 0,$$

$$\left(\frac{\partial}{\partial t} + \bar{u}_2\frac{\partial}{\partial x}\right)\left[\nabla^2\psi_2' - \frac{\delta\lambda^2}{1+\delta}(\psi_2' - \psi_1')\right.$$

$$\left.+\frac{f_0}{H}(1+\delta)b\right] + \left[\beta - \frac{\delta\lambda^2}{1+\delta}(\bar{u}_1 - \bar{u}_2)\right]\psi_{2x}'$$

$$+J\left[\psi_2', \nabla^2\psi_2' - \frac{\delta\lambda^2}{1+\delta}(\psi_2' - \psi_1') + \frac{f_0}{H}(1+\delta)b\right]$$

$$-\frac{\partial}{\partial y}\overline{\psi_{2x}'\left[\psi_{2yy}' + \frac{\delta\lambda^2}{1+\delta}\psi_1' + \frac{f_0}{H}(1+\delta)b\right]} = 0.$$

If we now consider y scales that are order $1/\Delta$ of the x scales or the deformation scale and expand $\bar{u}_i = \bar{u}_i^{(0)} + \Delta^2\bar{u}_i^{(1)} + \cdots$, $\psi_i'(x,y) = \psi_i'^{(0)}(x) + \Delta^2\psi_i'^{(1)}(x,y) + \cdots$ (where the topography is assumed to vary only in x) we find

$$\bar{u}_{1t}^{(0)} = \bar{u}_{2t}^{(0)},$$

so that the induced changes in mean flow are barotropic. This occurs because the form-drag forcing of the mean is at much larger scales than the deformation scale. Eliminating the $\bar{u}_i^{(1)}$ terms at second order from the two mean flow equations, noting that the Reynolds stresses drop out because $\psi_i'^{(0)}$ is independent of y, we find that the barotropic component of the zonal flow is accelerated or decelerated by the form drag,

$$\frac{\partial\bar{u}_2}{\partial t} = \frac{f_0}{H}\overline{\psi_{2x}'b'}, \tag{18.64}$$

while the deviation fields are given by

$$\left[\frac{\partial}{\partial t} + (\bar{u}_2(t) + \Delta u)\frac{\partial}{\partial x}\right]\left[\psi_{1xx}' - \frac{\lambda^2}{1+\delta}(\psi_1' - \psi_2')\right]$$

$$+\left(\beta + \frac{\lambda^2\Delta u}{1+\delta}\right)\psi_{1x}' = 0, \tag{18.65}$$

$$\left[\frac{\partial}{\partial t} + \bar{u}_2\frac{\partial}{\partial x}\right]\left[\psi_{2xx}' - \frac{\delta\lambda^2}{1+\delta}(\psi_2' - \psi_1')\right]$$

$$+\left(\beta - \frac{\delta\lambda^2\Delta u}{1+\delta}\right)\psi_{2x}' = -\frac{f_0}{H}(1+\delta)\bar{u}_2 b_x. \tag{18.66}$$

Here we have dropped the superscript (0) and introduced the notation Δu for the *time-independent* shear across the interface.

From the equations (18.64)–(18.66) one could derive a single nonlinear governing equation for $\bar{u}_2$ and determine the linear instability and finite-amplitude evolution of the flow. For our purposes, however, it will be sufficient to demonstrate that the initial state $\bar{u}_2 = 0$, $\psi_1' = 0$, $\psi_2' = 0$ can be unstable in the presence of weak topography to infinitesimal perturbations even when the flow would be baroclinically stable in the absence of topography. We can do this by considering the stability in the special case $\Delta u = \beta(1+\delta)/\delta\lambda^2$. This is the maximum shear for which the Rayleigh necessary criterion for stability in the absence of topography,

$$\left(\beta - \frac{\delta\lambda^2}{1+\delta}\Delta u\right)\left(\beta + \frac{\lambda^2}{1+\delta}\Delta u\right) \geq 0,$$

is satisfied. Equations (18.14)–(18.16) simplify to the set

$$\frac{\partial\bar{u}_2}{\partial t} = \frac{f_0}{H}\overline{\psi_{2x}'b},$$

$$\left[\frac{\partial}{\partial t} + \frac{\beta(1+\delta)}{\delta\lambda_1^2}\frac{\partial}{\partial x}\right]\left[\psi_{1xx}' - \frac{\lambda_1^2}{1+\delta}(\psi_1' - \psi_2')\right]$$

$$+\beta\left(\frac{1+\delta}{\delta}\right)\psi_{1x}' = 0,$$

$$\frac{\partial}{\partial t}\left[\psi_{2xx}' - \frac{\delta\lambda_1^2}{1+\delta}(\psi_2' - \psi_1')\right] = -\frac{f_0}{H}(1+\delta)\bar{u}_2 b_x.$$

We split the streamfunction into sine and cosine parts as in section 18.7.3 and solve this system of equations to find the growth rate equation

$$\sigma^4\left(\frac{\delta}{1+\delta}\right)^2 k^2(k^2+\lambda^2)^2$$

$$+\sigma^2\left[\frac{\beta^2}{\lambda^4}\left(k^4 - \frac{\delta}{1+\delta}\lambda^4\right)^2\right.$$

$$\left.+\frac{1}{2}\left(\frac{f_0 b_0}{H}\right)^2\frac{\delta^2}{1+\delta}k^2\left(k^2 + \frac{\lambda^2}{1+\delta}\right)(k^2+\lambda^2)\right]$$

$$+\frac{1+\delta}{2}\frac{\beta^2}{\lambda^2}\left(\frac{f_0 b_0}{H}\right)^2 k^2\left(k^2 - \frac{\delta}{1+\delta}\lambda^2\right)\left(k^4 - \frac{\delta}{1+\delta}\lambda^4\right)$$

$$= 0. \tag{18.67}$$

The real solutions to (18.67) for several values of $f_0 b_0\lambda/\beta H$ are shown in figure 18.16. Notice the instability occurring for topographic scales on the order of 70 to 110 km, with growth rates proportional to the topographic height (for small heights, at least). We have thus demonstrated that the available potential energy in the flow can be tapped by the orographic instability even in situations where normal baroclinic instability is unable to extract mean-flow energy. Thus it is possible that mesoscale topography plays a role in catalyzing the conversion of mean flow potential to eddy energy in the oceans.

18.7.4 Multiple Equilibria

We have already mentioned the work of Charney and DeVore (1979) and Charney and Straus (1980), who have begun to explore the possibility that the atmosphere may possess a multiplicity of steady equilibrium

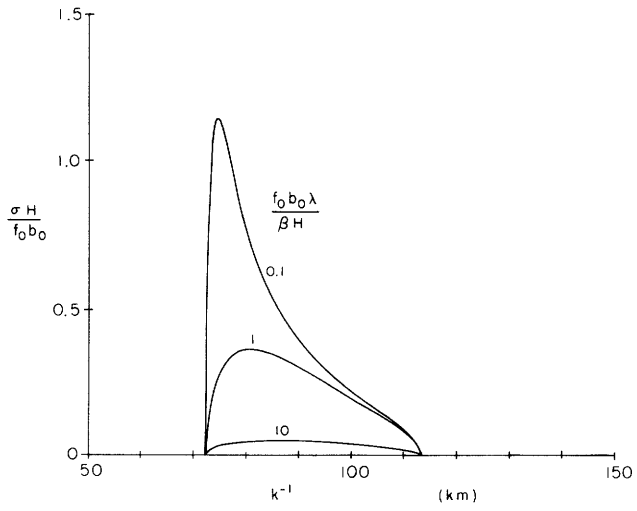


Figure 18.16 Normalized growth rates for a topographically destabilized vertical shear flow. The curves are labeled by $f_0 b_0 \lambda/\beta H$ values.

states for given external forcing in the presence of topographic inhomogeneities. In the case of sinusoidal topography in a periodic channel, they have found states resembling both the "normal" configuration in which there is a strong zonal flow and a relatively weak wave perturbation, and the "blocking" configuration in which there is a weak zonal flow and a relatively strong wave perturbation. They suggest that the blocking phenomenon is an equilibrium state which occurs by a transition via a form-drag instability from the normal to the anomalous blocking configuration. Hart (1979b) has applied similar ideas to laboratory flows and has succeeded in producing stationary multiple equilibria experimentally.

Oceanically, one phenomenon that stands out as a possible example of multiple quasi-stable equilibrium states is the large meander of the Kuroshio which sometimes occurs. Figure 18.17 shows the two quasi-stable configurations that are observed. The transitions between these configurations occur relatively rapidly. White and McCreary (1976) have considered a model for the meandering process involving flow around bumps in the Japanese coastline. Because their discussion was in terms of linear dynamics, Solomon (1978) has rightly pointed out that the model must have a smooth transition between the two states as the independent variable (the maximum inlet flow speed) varies. If, however, the phenomenon is nonlinear, catastrophic changes in the state of the Kuroshio may occur: an infinitely small change in parameters may produce a finite change in response, and several stable responses may be possible for the same set of parameters.

We propose a simple model of this process consisting of the steady, nonlinear flow of barotropic current on a β-plane along a variable coastline (see figure 18.18). Let the latitude of the coastline be $h(x)$ and let η be the north–south distance from the coastline. The potential vorticity equation becomes

$$\left[\left(\frac{\partial}{\partial x} - h_x \frac{\partial}{\partial \eta}\right)^2 + \frac{\partial^2}{\partial \eta^2}\right] \psi + \beta(\eta + h) = F(\psi).$$

If we split the streamfunction into an upstream part ($x \to -\infty$, $h \to 0$) $\bar{\psi}(\eta)$ and a topographically induced part $\phi(x, \eta)$ we find

$$\left[\left(\frac{\partial}{\partial x} - h_x \frac{\partial}{\partial \eta}\right)^2 + \frac{\partial^2}{\partial \eta^2}\right] \phi = F(\bar{\psi} + \phi) - F(\bar{\psi}) - \beta h - \bar{\psi}_{\eta\eta} h_x^2, \qquad (18.68)$$

where

$$F(\bar{\psi}(\eta)) = \beta\eta + \frac{\partial^2 \bar{\psi}}{\partial \eta^2},$$
$$\phi \to 0 \quad \text{for } \eta = 0, \quad \eta \to -\infty. \qquad (18.69)$$

When $\bar{u} = -\partial\bar{\psi}/\partial\eta$ is not constant, equation (18.69) implies that F is a nonlinear functional, so that (18.68) becomes essentially a forced nonlinear oscillator equation; it is well known that such equations may have multiple stable solutions. We note also a similarity between the equations here and the equations for flow of a barotropic fluid over topography. In the derivation below we assume that the coastline variations are small and occur on scales large compared to the cross-stream scale. We shall show that the nonlinearity plays an important role in determining the amplitude of the nonzonal flow component when the upstream flow is near the critical speed U_c. This speed is defined by the condition that long waves (x wavelength large compared to the width of the current) are stationary. Near critical speeds, the amplitude becomes large. The lowest-order dynamic equation only determines the cross-stream wave structure. The first-order equation shows a balance between advection by the mean flow, effects of the coastline variations, dispersion, and nonlinearity.

We shall work with the nondimensional forms of (18.68)–(18.69). Our scaling is guided by the versions

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial \eta^2}\right) \phi = -\beta h + F'(\bar{\psi})\,\phi = -\beta h - \frac{\beta - \bar{u}_{\eta\eta}}{\bar{u}} \phi \qquad (18.70)$$

obtained by linearizing in ϕ and h. We obtain $F'(\bar{\psi})$ by differentiating (18.69). If $h = h_0 \cos kx$, resonance occurs when

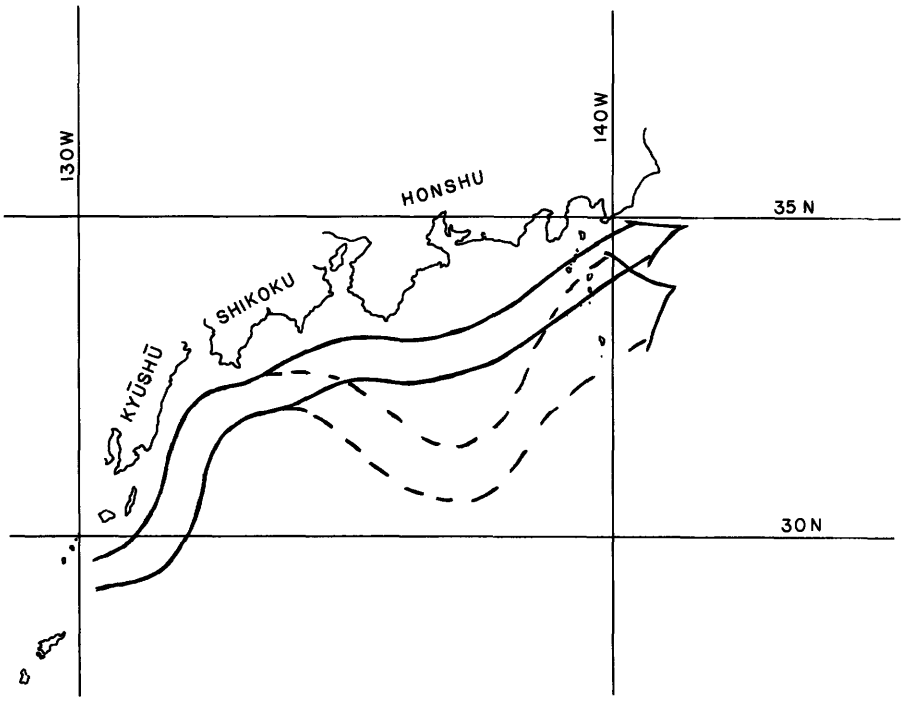


Figure 18.17 Sketch of two equilibrium positions of the Kuroshio. See Taft (1972) and White and McCreary (1976) for detailed tracks.

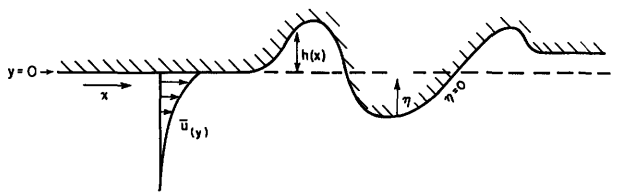


Figure 18.18 Model for coastline induced meandering. The deviation of the coastline from a latitude circle is denoted $h(x)$; the coordinate η is equal to $y - h(x)$. The upstream flow is $\bar{u}(y)$.

$$\left(\frac{\partial^2}{\partial\eta^2} + \frac{\beta - \bar{u}_{\eta\eta}}{\bar{u}}\right)\phi = k^2\phi,$$
$$\phi = 0, \quad \eta = 0, -\infty. \tag{18.71}$$

For forcing on a scale long compared to the width of the current ($|\partial/\partial x| \ll |\partial/\partial\eta|$) we expect that one of the underlined terms in (18.70) will balance the forcing from the side-wall variations $-\beta h$ giving $\phi \sim Uh_0$ or $\phi \sim \beta h_0 l^2$, where U is the scale of $\bar{u}$ and l is the cross-stream scale. When the flow profile is nearly critical for long waves—meaning that the left-hand side of (18.71) vanishes for some nonzero function ϕ which also satisfies the boundary conditions, the long-wave solutions of (18.70) have the two underlined terms nearly canceling, so that the forcing must be balanced by the ϕ_{xx} term. This gives a scale of $\phi \sim \beta h_0 L^2$, where L is the downstream scale of variation of the topography.

Therefore, we scale x by L, h by h_0, ϕ by $\beta h_0 L^2$, η by l, and $\bar{u}$ by U in (18.68)-(18.69) to find

$$\left[\frac{\partial^2}{\partial\eta^2} + \delta\left(\frac{\partial}{\partial x} - \gamma\delta^2 h_x\frac{\partial}{\partial\eta}\right)^2\right]\phi$$
$$= -\delta h + \frac{1}{\gamma\delta}\left[F\left(\bar{\psi} + \frac{\gamma\delta}{M}\phi\right) - F(\bar{\psi})\right] - \gamma M\delta^4 h_x^2\bar{\psi}_{\eta\eta},$$
$$F(\bar{\psi}(\eta)) = \eta + M\bar{\psi}_{\eta\eta},$$

with $M = U/\beta l^2$, $\gamma = h_0 L^4/l^5$, and $\delta = l^2/L^2$. If we assume that the width of the current is small compared to the downstream scale ($\delta \ll 1$) and that the variations in coastline are weak enough so that $\gamma \lesssim 1$, we can simplify to

$$\left(\frac{\partial^2}{\partial\eta^2} + \delta\frac{\partial^2}{\partial x^2}\right)\phi = -\delta h + F'(\bar{\psi})\frac{\phi}{M} + \frac{1}{2}F''(\bar{\psi})\frac{\gamma\delta\phi^2}{M^2} + O(\delta^2) \tag{18.72}$$

$$\phi \to 0, \qquad \eta = 0, -\infty,$$

with

$$F'(\bar{\psi}) = -\frac{1}{\bar{u}}(1 - M\bar{u}_{\eta\eta}),$$
$$F''(\bar{\psi}) = \frac{1}{\bar{u}}\frac{\partial}{\partial\eta}\left[\frac{1}{\bar{u}}(1 - M\bar{u}_{\eta\eta})\right], \tag{18.73}$$

which are known functions of η given the specification of the upstream ($x \to -\infty$) flow $\bar{u}(\eta)$.

We assume that the flow is nearly critical so that $U = U_c(1 + \Delta)$ where U_c is the critical speed (defined exactly below) and therefore $M = M_c(1 + \Delta)$. We expand (18.72)-(18.73) assuming $\Delta \sim \delta$ and $M_c \sim 1$, $\gamma \leq 1$, and find to lowest order

$$\frac{\partial^2}{\partial\eta^2}\phi = \frac{\bar{u}_{\eta\eta} - M_c^{-1}}{\bar{u}}\phi,$$
$$\phi = 0, \qquad \eta = 0, -\infty, \tag{18.74}$$

which defines the eigenvalue M_c and thus the critical speed U_c given the shape of the upstream flow. The η structure of ϕ must be an eigenfunction G of (18.74), $\phi = f(x)G(\eta)$. At next order in Δ and δ, the solvability condition for (18.72) gives

$$\left[\int_{-\infty}^{0} d\eta\, G^2(\eta)\right] f''(x) = -h(x)\left[\int_{-\infty}^{0} d\eta\, G(\eta)\right] + \frac{\Delta}{\delta M_c}\left[\int_{-\infty}^{0} G^2(\eta)/\bar{u}(\eta)\right] f(x) + \frac{1}{2}\left[\int_{-\infty}^{0} d\eta\, G^3(\eta)\frac{1}{\bar{u}}\frac{\partial}{\partial\eta}\frac{1}{\bar{u}}(1 - M_c\bar{u}_{\eta\eta})\right]\frac{\gamma}{M_c^2} f^2(x).$$

This ordinary differential equation for the x structure of the wave $f(x)$ is to be solved for a particular form as Δ/δ and γ vary. For convenience we shall normalize G and redefine parameters slightly to write

$$f'' - \hat{\Delta}f + \hat{\gamma}f^2 = -h(x). \tag{18.75}$$

The simplest problem to illustrate the characteristics of (18.75) is the linear case with $h(x) = \cos x$. (This topography extends to $x = -\infty$, which is not really consistent with our original model: however, it does point out some of the properties of these nonlinear flows.) The solution to (18.75) with $\hat{\gamma} = 0$ is

$$f = \frac{\cos x}{1 + \hat{\Delta}},$$

showing a resonance at $\hat{\Delta} = -1$ (see figure 18.19).

For weak nonlinearity ($\hat{\gamma}$ small), we can express f as a Fourier series

$$f = A\cos x + \hat{\gamma}(A_0 + A_2\cos 2x) + \hat{\gamma}^2\sum_{n=3}^{\infty} A_n\cos nx,$$

which implies a cubic equation for A:

$$(1 + \hat{\Delta})A - \hat{\Delta}^2A^3\left[\frac{1}{\hat{\Delta}} + \frac{1}{(4 + \hat{\Delta})}\right] = 1. \tag{18.76}$$

(One can show that the higher-order terms will not contribute, even near resonance.) Figure 18.19 also shows the solution of (18.76) for $\hat{\gamma} = 0.2$. Here we clearly see that there are three equilibrium states for $\hat{\Delta} < -1.3$. The state with intermediate amplitude is unstable; thus we see that we can have either a large positive amplitude wave (in phase with topography) or a small negative amplitude wave (out of phase).

This simple model suggests that the Kuroshio meander may be a case of multiple states depending on the flow rate at the inlet. Slight decreases in speed may cause a sudden transition to a meander state, with hysteresis effects likely, so that large increases are necessary before the Kuroshio would return to its path closer to the coast.

The above results merely suggest the possibility of multiple equilibria because, to begin with, we have required the coastline to have an infinite number of ridges and troughs in order to create the possibility of linear resonance. As shown in figure 18.10, an infinite number of periods may not be necessary, but there must be at least two ridges and a trough or vice versa. A single coastal ridge (as in the half-wave case) would not be enough to give a maximum response. One must

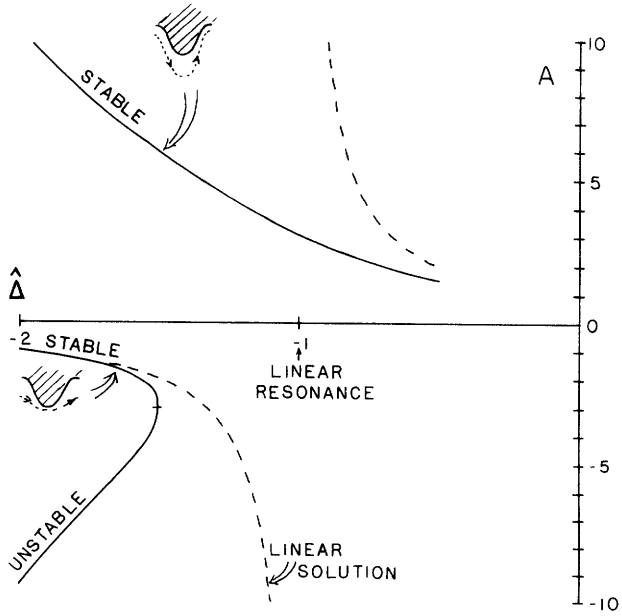


Figure 18.19 The wave amplitude as a function of $\hat{\Delta}$ (proportional to the magnitude of the upstream current). Sketches show the relationship between the streamlines and the coastline. Multiple equilibria occur for $\hat{\Delta} < -1.3$.

ask: Can a single ridge or, in the case of blocking in the atmosphere, a single mountain range, in an extended domain give rise to multiple equilibria? And is resonance needed?

In problems of nonrotating shallow-water flow over an obstacle, multiple states can exist when the Froude number $\bar{u}/\sqrt{gH}$ is greater than unity. For certain values of the Froude number and the ratio of the obstacle height to H, two states are found, one corresponding to smooth flow with no upstream disturbance and one corresponding to a permanent elevation of the free surface upstream of the obstacle—created during the approach to equilibrium by a bore traveling upstream (Baines and Davies, 1980). Similar examples of multiple equilibria are also found in transonic compressible flow past obstacles. Thus we suspect that the upstream flow in our problem cannot be specified arbitrarily but may very well be affected by the upstream propagation of energy. It may be that, as in the periodic models of Charney-DeVore and Charney-Straus, the flow must be dealt with as a global or basin-wide unit.

18.7.5 Quasi-Geostrophic Turbulence

We have considered only wave-wave or wave-mean flow interactions involving a small number of components. In particular, we have not considered energy-cascade processes involving large numbers of components and leading ultimately to turbulent dissipation. It was pointed out by Onsager (1949), Lee (1951), Batchelor (1953a), and especially by Fjørtoft (1953) that vorticity conservation in two-dimensional flow imposes a strong constraint on scale interactions. Later Charney (1966, 1971a) showed that the conservation of pseudopotential vorticity in three-dimensional quasi-geostrophic flow imposes similar constraints. Such constraints suggested to Kraichnan (1967) that there may be an inertial subrange in two-dimensional, homogeneous, isotropic turbulence in which the energy spectrum is controlled by uniform transfer of enstrophy (mean-squared vorticity) from large to small scales at scales less than the excitation scale, and by uniform transfer of energy from small to large scales at scales greater than the excitation scale. He predicted a k^{-3} spectral energy density for scalar wavenumber k in the former range, and a Kolmogorov $k^{-5/3}$ law in the latter range. In extending these ideas to three-dimensional, quasigeostrophic turbulence, Charney (1971a) also obtained a k^{-3} law at the tail of the spectrum and conjectured that in this region there would also be equipartition between the two components of the kinetic energy and the available potential energy. This conjecture has been confirmed by Herring (1980) in a homogeneous quasi-geostrophic turbulence closure model.

The topic of quasi-geostrophic turbulence has been investigated by a number of oceanographers, notably Rhines (1975) by numerical simulation, and Holloway and Hendershott (1977) and Salmon (1978) by means of closure [see also Herring (1980) and Leith (1971)]. It may be that their work is more applicable to the atmosphere than to ocean basins, where meridional boundaries play important roles and where statistical inhomogeneity of excitation cannot be ignored.

The existence of quadratic invariants, energy and enstrophy—mean-squared vorticity in two-dimensional or mean-squared pseudopotential vorticity in three-dimensional quasi-geostrophic flows—permits application of the principles of statistical mechanics. These have been applied by Onsager (1949) and Kraichnan (1975) to two-dimensional flow, and by Salmon, Holloway, and Hendershott (1976) to a two-layer quasi-geostrophic flow. In the two-dimensional case, the energy in each horizontal mode is $L^2/(b + aL^2)$, where a and b are constants depending on the total energy and enstrophy. With typical choices of these constants, the largest scale waves have the most energy. In the two-layer case, the equilibrium spectrum is dominated by the largest scales, and these motions are barotropic. The available potential-energy spectrum, corresponding to the thermocline displacement spectrum, is peaked near the deformation radius. These spectra represent the effects of the nonlinear terms alone; one expects [as Errico (1979) found in studying the partition of energy between gravity-wave and geostrophic motions] that the spectra which are actually realized in a forced and dissipative system will be determined largely by the wavenumber dependence of the forcing and dissipation.

The cornerstone for the theory of quasi-geostrophic turbulence is the conservation (in the inviscid limit) of the energy $\frac{1}{2}\iiint |\nabla\psi|^2 + (1/S)|\partial\psi/\partial z|^2$ and the enstrophy $\frac{1}{2}\iiint[\nabla^2\psi + (\partial/\partial z)(1/S)(\partial\psi/\partial z)]^2$. We emphasize the fragility of the last principle: enstrophy can increase or decrease if there are (1) temperature gradients along horizontal boundaries, (2) side walls on the domain, or (3) topography.

If we first consider the case when none of these restrictions obtain—flow in a periodic, flat-bottomed domain—we can readily argue that the energy will be transferred to large horizontal scales and to more barotropic motions by the nonlinear terms. If we expand the streamfunction in the flat-bottomed normal modes and perform a Fourier transform horizontally,

$$\psi = \sum_n F_n(z) \iint d\mathbf{k}\hat{\psi}_n(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{x}},$$

we can use the conservation principles

$$\frac{\partial}{\partial t}\sum_n \iint d\mathbf{k}\,(\mathbf{k}^2 + \lambda_n^2)\,|\hat{\psi}_n|^2 = 0,$$

$$\frac{\partial}{\partial t}\sum_n \iint d\mathbf{k}(\mathbf{k} + \lambda_n^2)^2\,|\hat{\psi}_n|^2 = 0$$

(λ_n is the reciprocal radius of deformation for the nth baroclinic mode; see section 18.5.1) in exactly the same manner as Fjørtoft (1953) or Charney (1971a) to show that the amount of energy with $k^2 + \lambda_n^2 > K_0^2$ is a small fraction of the initial energy, if K_0 is large compared to the initial mean wavenumber

$$\overline{K} = \frac{\Sigma_n \iint d\mathbf{k}\,(\mathbf{k}^2 + \lambda_n^2)^{1/2}[(\mathbf{k}^2 + \lambda_n^2)|\hat{\psi}|^2]}{\Sigma_n \iint d\mathbf{k}[(\mathbf{k}^2 + \lambda_n^2)|\hat{\psi}|^2]}.$$

In essence the nonlinearity does not transfer energy to small scales. Another way to show the reverse cascade, is to use Rhines's (1975) argument that the turbulence spreads energy out in wavenumber space

$$\frac{\partial}{\partial t} \sum_n \iint d\mathbf{k}\,[(\mathbf{k}^2 + \lambda_n^2)^{1/2} - \overline{K}]^2[(\mathbf{k}^2 + \lambda_n^2)|\hat{\psi}_n|^2] > 0.$$

Combining this with the definition of $\overline{K}$ and using the conservation laws shows that the mean total wavenumber must decrease

$$\frac{\partial}{\partial t}\overline{K} < 0.$$

Thus energy cascades to larger horizontal *and* vertical scales, implying an increase in energy for small λ_n^2, that is, a tendency for the motion to become more barotropic.

This tendency for the flow to become barotropic in the absence of topography or side walls has been well demonstrated through numerical simulation by Rhines (1977). He has also shown that rough topography can halt this cascade for flows that are not too energetic. The β-effect slows the cascade when the scale has increased so much that the wave steepness M becomes order one. This could occur while the motions are still baroclinic if the initial energy were small compared to $\beta^2 L_R^4$.

Rhines (1975) has also argued that side walls can stop the reverse cascade. This has been demonstrated in laboratory experiments by Colin de Verdière (1977). Essentially the western boundary serves as a source of enstrophy and the eastern boundary a sink. This can be understood—as Pedlosky (1967), in a slightly different context, has revealed—by considering reflection of Rossby waves from the western boundary. For linear waves with $c = -\beta/(k^2 + l^2 + \lambda_n^2)$, the x component of the group velocity is negative for $k^2 < (l^2 + \lambda_n^2)$. Therefore the reflected wave's zonal wavenumber $k_r = (l^2 + \lambda_n^2)/k$ is larger than k. One can readily show that the energy fluxes of the incident and reflected waves (c_g times the energy density) are equal and opposite, but that the enstrophy flux of the outgoing wave is larger by a factor k_r/k than the flux of the incident wave.

Rhines (1977) has discussed many of the topics above, and in particular has demonstrated clearly that the strong nonlinear interactions involved in geostrophic turbulence cannot occur unless nonlinearity is much stronger than wave dispersion, that is, unless the wave steepness $Uk^2/\beta \gg 1$. At these scales there is some, but not conclusive, evidence in support of a k^{-3} spectrum for the atmosphere (Julian and Clive, 1974). The enstrophy cascade mechanism has not yet been checked adequately either by direct measurement or by numerical simulation. It remains possible that the observed atmospheric spectra can be explained in terms of ordered (periodic) frontal structures, rather than random cascades of enstrophy, as suggested by Andrews and Hoskins (1978). They obtain a $k^{-8/3}$ dependency which is as much in accord with observations as the k^{-3} spectrum—or perhaps better. However, the predicted spectra are highly anisotropic, and this does not seem to be in as good agreement with observations or results from numerical modeling as the predictions of the theory of geostrophic turbulence.

So far the data do not exist for a corresponding oceanic check.

18.8 Summary Remarks

Our primary focus has been on oceanic analogues of transient atmospheric motions of large scale. As promised in the introduction, it has been possible to find formal oceanic analogues for most categories of large-scale atmospheric motions: indeed, it has been almost trivial to do so. Far more difficult, however, has been to demonstrate the physical reality and importance of these analogues. It is only recently, through studies of the meanders of the western boundary currents and through such concentrated, large-scale observational programs as MODE, Polygon, POLYMODE, and the oceanographic component of the GARP Atlantic Tropical Experiment (GATE), that some understanding of the nature of the transient motions has begun to emerge. Although many of the oceanic analogues we have dealt with are hypothetical to a greater or lesser degree, we have chosen them on physical grounds as at least of potential importance. We feel that contrasting them with their atmospheric counterparts, with respect to both their individual properties and their roles in the generation and maintenance of the large-scale circulation patterns, has been a useful exercise.

In section 18.4, it is shown that the quasi-geostrophic, β-plane formalism derived for the atmosphere applies to oceanic motions whose horizontal scales are on the order of the deformation radius of the first baroclinic mode L_R, that is, the scales corresponding to baroclinic instability. It is characteristic of these motions that they are dispersive at both small and large amplitudes. At larger scales the dynamics change: linear free modes may no longer have the same east-west and north-south scales and vertical density advection

becomes an important cause of nonlinearity. For motions with length scales near the "intermediate scale" $(L_R^2 a)^{1/3}$ with velocities of order $f_0 L_R^2/a$ (where a is the radius of the earth), the vertical component of vorticity changes not only because of β-effects and horizontal advection but also because of vertical density advection and variation of the undifferentiated Coriolis parameter. Thus both Rossby and Burger terms appear. For east–west scales on the order of the intermediate scale (210 km for the ocean and 1500 km for the atmosphere) or larger, dispersion and nonlinearity can balance to give solitary or cnoidal waves. For larger scales, we show that the evolution is determined by a Korteweg–deVries equation, so that solitons will be the natural end product of the evolution of an initially isolated disturbance. Because of the slower dissipation and the larger scale separation between the intermediate and basin scales, such waves are more likely to be found in the ocean than in the atmosphere. These results suggest that the larger-scale dynamics of the ocean transients may be dominated by more orderly, phase-coherent structures than are predicted by the theory of geostrophic turbulence. If this is so, then at scales larger than the excitation scale the low-wavenumber components would be more highly correlated than if they were due to a random reverse cascade.

The theory of free and forced small-amplitude Rossby waves in the oceans may be transposed almost entirely from the corresponding atmospheric theory. The MODE and Polygon experiments have provided evidence of the importance of Rossby-wave propagation, particularly on time scales greater than a month (McWilliams, 1976). In section 18.5, we present some of this theory in an oceanographic context, paying particular attention to the influence of bottom topography in altering the propagation of free waves and in generating waves from a mean flow. Rhines's results for a uniformly stratified fluid are extended to arbitrarily stratified flows. The major result, the prediction of bottom-trapped modes, has been verified from observations at site D (Thompson, 1977) although no observations of the eastward-traveling modes which should exist when the topography opposes the β-effect have been reported.

The idea of Rossby-wave propagation in a medium with a variable (real or imaginary) index of refraction was first advanced to account for the vertical trapping of shorter waves by upper easterlies and strong westerlies. We apply these ideas in an oceanographic context not only to vertical propagation from the surface or bottom but also to horizontal propagation of waves generated by the meandering of the Gulf Stream. Eastward-propagating meanders produce trapped disturbances close to the Gulf Stream, while westward-propagating meanders may give a real index of refraction and southward propagation.

In section 18.6, we describe briefly the influence of friction both on the generation of ocean currents by wind and on the decay of individual oceanic eddies. Existing theory is not adequate to account for the spin-up of the real ocean or for the decay of the real atmospheric circulation. Interest in the baroclinic spin-down problem was originally motivated by a desire to understand the long persistence of Gulf Stream rings; however, the axisymmetric models that had previously been employed do not account for the complete decay of a baroclinic eddy. We show that the β-effect permits vertical propagation of energy and therefore allows for complete spin-down.

In the last section, we consider oceanic disturbances in which advective effects are important—either through wave–mean flow interaction as in the breakdown of an unstable mean flow, through the interaction of waves generated elsewhere with the local mean flow, or through wave-wave interactions. The last type of interaction can occur when the unstable flow is itself a wave or when waves generated in any manner interact with one another as in turbulence.

The concept of baroclinic instability was developed to explain the principal traveling waves and vortices embedded in the atmospheric westerlies; it has been applied to the oceans in an effort to account for the meandering of the western boundary currents and the existence of mid-ocean mesoscale eddies. The meandering does seem to be an effect of baroclinic instability, modified by barotropic effects, but it is highly questionable on theoretical and numerical-modeling grounds whether the mid-ocean eddies are due to local baroclinic instabilities. It seems more likely that these eddies are vortices cast off from, or forced in some more general fashion by, the meandering western boundary currents and their extensions.

In our exposition of the baroclinic and barotropic instability problem, we use the methods of Arnol'd and Blumen to extend the integral theorems of Kuo, Charney-Stern, and others to a class of basic flows that need not be zonal, that may translate with constant speed, and that may be influenced by topography.

The work on wave-mean flow interactions originated by Eliassen and Palm and Charney and Drazin in an atmospheric context is applied to the problem of the rectification of Rossby waves radiated from the western boundary currents and their extensions. Of particular interest is the so-called recirculation flow found by Worthington and others south of the Gulf Stream extension. Rhines has attempted to account for this recirculation as driven by the westward-propagating Rossby waves produced by the meandering. We present a slight generalization of his work by considering also the effects of eastward-traveling meanders. The results do suggest that a relatively strong westward flow, confined fairly close to the Gulf Stream, can be produced

by eastward-propagating wave–mean flow interactions in the presence of dissipation.

In recent times, the stability analyses for the atmosphere have been extended to wavy motions in an attempt to account for nonlinear cascades of energy in large-scale motions. The stability of forced wavy motions has also been studied to account for the transition from one stationary state to another of a forced flow over topography. It has been found from a study of simple truncated spectral models that the stationary flow equilibria produced by the forcing of a zonal flow over topography in a rotating system may be indeterminate in the sense that for a given forcing, there exists a multiplicity of equilibrium states. This result has been utilized in an attempt to explain the so-called blocking phenomenon in the atmosphere—the persistence of large-amplitude anticyclonic flow anomalies in the planetary circulation. The existence of multiple equilibria for a given forcing appears to be common; it also occurs for supercritical-Froude-number flow in hydraulics and transonic flow in gas dynamics. A natural oceanic analogue of multiple, quasi-stationary equilibrium in the atmosphere is the known existence of two states of flow for the Kuroshio in the vicinity of the Japanese coast. We investigate a simple model of such a flow and find indeed that two different steady states may be produced by a given upstream flow as it passes a wavy boundary. Our model leaves much to be desired, but it does point a direction for future research.

We find that the transition between one state and another in topographically forced flows occurs via a form-drag instability in which the perturbed form drag (mountain torque) modifies the mean flow in such a way that the perturbation increases in amplitude. The instability of a forced topographic wave is thus different from that of a free wave. The latter instability was shown by Gill to be basically a Rayleigh-like shear instability or a resonant-triad wave interaction. We have presented the wave-stability analysis for both free and forced waves in a unified fashion to bring out the similarities and differences between the shear, resonant, and form-drag instabilities.

The form-drag instabilities producing transition grow in place; however, the wavy equilibrium states themselves may also exhibit traveling Rayleigh-like instabilities. It has been suggested that instabilities of this kind account for the observed eastward- and westward-propagating very long (zonal wavenumbers 1–4) planetary waves in the atmosphere. One may speculate that a careful analysis of the topographically induced meanders of the western boundary currents in the ocean will also reveal such secondary wave instabilities. If these are westward propagating, then they could contribute to a broader recirculation region.

The final topic is quasi-geostrophic turbulence. Fjørtoft's prediction that there will be a transfer of energy from the excitation scale to larger scales in 2-dimensional energy-and-enstrophy-conserving flow may be extended to 3-dimensional quasi-geostrophic flows if the bottom and top boundaries are flat and isentropic. The theory predicts that the scale will increase vertically as well as horizontally, that is, that the flow will become increasingly barotropic at large horizontal scales. Rhines has verified this result in a series of numerical experiments, and the oceanic observations of Schmitz (1978) show that the mesoscale eddies tend to be more barotropic the more energetic they are. This observation, while not verifying the inertial theory of geostrophic turbulence, is at least consistent with it. As Rhines has shown, the inertial theory applies only when the effects of nonlinearity dominate those of linear dispersion, that is, when the wave steepness Uk^2/β is much greater than unity. Thus one expects the prediction to be valid only in the energetic parts of the ocean. The similarity prediction of a k^{-3} spectrum at scales smaller than the excitation scale is not inconsistent with observations in the atmosphere, but there are as yet no data to test this theory in the ocean. The effects of topography, side boundaries, and surface gradients of entropy also have not been thoroughly explored.

We have not produced a systematic or comprehensive treatment of atmosphere–ocean analogues. Our excuse, as we have stated, is that virtually all large-scale atmospheric motions have oceanic counterparts and there are simply too many of these to discuss. We have preferred to deal with analogues for which there is observational evidence or at least some physical basis for believing they should exist. We have been forced to speculate, and, as the reader will surely have perceived, our own speculations have been guided primarily by our own experience and interests.

Appendix: The Quasi-Geostrophic Equations

Here we shall give details of the derivations of the quasi-geostrophic equations. We begin by nondimensionalizing the equations of motion (18.22)–(18.27), using the definitions of the geostrophic streamfunction and the potential buoyancy b_p in (18.28) and (18.29). Using the characteristic scales described in the text, we obtain

$$\varepsilon\frac{Du}{Dt} + \omega\epsilon\lambda\hat{\beta}\frac{uw\tan\Theta}{r} - \epsilon\hat{\beta}\frac{uv\tan^2\Theta}{r} - \lambda\omega\cot\Theta w - v$$

$$= -\frac{\alpha}{r}\frac{1}{\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)\psi, \qquad (18.A1)$$

$$\varepsilon\frac{Dv}{Dt} + \omega\epsilon\lambda\hat{\beta}\frac{vw\tan\Theta}{r} + \epsilon\hat{\beta}\frac{u^2\tan^2\Theta}{r} + u$$

$$= -\frac{\alpha}{r}\frac{1}{\sin\Theta}\left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right)\sin\Theta\psi, \qquad (18.A2)$$

$$\omega\lambda^2\varepsilon\frac{Dw}{Dt} - \epsilon\lambda\hat{\beta}\tan\Theta\frac{u^2+v^2}{r} - \lambda u\cot\Theta$$

$$= -\psi_z + \Delta\hat{S}\psi + b_p$$

$$- \epsilon\Delta(b_p - \Delta_s\psi)(\psi_z - \Delta\hat{S}\psi), \tag{18.A3}$$

$$\frac{1}{r\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)u + \frac{1}{r}\left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\frac{\partial}{\partial\Theta}\right)v$$

$$-\frac{\hat{\beta}v\tan^2\Theta}{r} + \frac{\omega}{r^2}\frac{\partial}{\partial z}r^2w$$

$$= \frac{\varepsilon\Delta}{\alpha}\frac{1}{\sin\Theta}\frac{D}{Dt}\sin\Theta(b_p - \Delta_s\psi) + \omega(\Delta_s + \Delta\hat{S})w \tag{18.A4}$$

$$\frac{1}{\alpha\sin\Theta}\frac{D}{Dt}\sin\Theta b_p + \frac{\omega}{\varepsilon}\hat{S}w + \frac{\omega\Delta_s}{\varepsilon\Delta}\left(1 - \frac{\alpha^2}{\hat{c}_s^2}\right)w$$

$$- w\psi\frac{\omega\epsilon}{\varepsilon}\frac{\partial}{\partial z}\Delta_s + \omega\frac{\epsilon}{\varepsilon}(\Delta\hat{S}$$

$$+ \Delta_s)\left[b_p - \Delta_s\left(1 + \frac{\alpha^2}{\hat{c}_s^2}\right)\psi\right]w$$

$$+ \left(\frac{\alpha^2}{\hat{c}_s^2} - 1\right)\Delta_s\frac{1}{\sin\Theta}\frac{D}{Dt}\sin\Theta\psi = 0, \tag{18.A5}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{\epsilon}{\varepsilon}\left[\frac{u}{r\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)\right.$$

$$\left. + \frac{v}{r}\left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right) + \omega w\frac{\partial}{\partial z}\right]$$

and

$$r = 1 + \lambda\hat{\beta}\tan\Theta z,$$

$$\alpha = 1 + \epsilon\Delta(b_p - \Delta_s\psi),$$

$$\hat{c}_s^2 = c_s^2/\bar{c}_s^2 = 1 + O(\epsilon\Delta)$$

are used as abbreviations.

The quasi-Boussinesq approximation entails choosing the scale of motion to be small compared to the external radius of deformation $\Delta \ll 1$. We therefore drop terms from (18.A1)-(18.A5) which are small in this sense. The definition of "small" requires some care because the various Rossby numbers are also small. Thus in equations (18.A1), (18.A2), (18.A4) we shall keep both order 1 and order ε, ϵ, $\hat{\beta}$, ω terms so that we may drop only terms of order $\epsilon\Delta$, $\varepsilon\Delta$, $\omega\Delta$. Equations (18.A1) and (18.A2) are changed only slightly: α is replaced by 1. The other equations become

$$\frac{1}{r\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)u + \frac{1}{r}\left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right)v$$

$$-\frac{\hat{\beta}v\tan^2\Theta}{r} + \frac{\omega}{r^2}\frac{\partial}{\partial z}r^2w = \Delta_s\omega w, \tag{18.A6}$$

$$\omega\lambda^2\varepsilon\frac{Dw}{Dt} - \epsilon\lambda\hat{\beta}\tan\Theta\frac{u^2+v^2}{r} - \lambda u\cot\Theta$$

$$= -\psi_z + b_p, \tag{18.A7}$$

$$\frac{1}{\sin\Theta}\frac{D}{Dt}\sin\Theta b_p + \frac{\omega}{\varepsilon}\hat{S}w - \omega\frac{\epsilon}{\varepsilon}w\psi\frac{\partial}{\partial z}\Delta_s$$

$$+ \frac{\omega}{\varepsilon}\Delta_s\epsilon\left(\frac{\hat{c}_s^2 - 1}{\epsilon\Delta}\right)w - \frac{\omega\epsilon}{\varepsilon}\Delta_s(b_p - \Delta_s\psi)w = 0. \tag{18.A8}$$

For the ocean, we can further refine these equations by noting that Δ_s is also very small, so that the continuity equation (18.A6) becomes that of an incompressible fluid, and the potential buoyancy equation becomes simply

$$\frac{1}{\sin\Theta}\frac{D}{Dt}\sin\Theta b_p + \frac{\omega}{\varepsilon}\hat{S}w = 0. \tag{18.A9}$$

The hydrostatic approximation applies to a thin layer of fluid: $\lambda \ll 1$. This allows us to drop the centrifugal terms involving w, the vertical accelerations, and to replace r by 1, giving us

$$\varepsilon\frac{Du}{Dt} - \epsilon\hat{\beta}uv\tan^2\Theta - v$$

$$= -\frac{1}{\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)\psi, \tag{18.A10}$$

$$\varepsilon\frac{Dv}{Dt} + \epsilon\hat{\beta}u^2\tan^2\Theta + u$$

$$= -\frac{1}{\sin\Theta}\left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right)\sin\Theta\psi, \tag{18.A11}$$

$$\frac{1}{\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)u + \left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right)v$$

$$- \hat{\beta}v\tan^2\Theta + \omega\frac{\partial w}{\partial z} = 0, \tag{18.A12}$$

$$\psi_z = b_p, \tag{18.A13}$$

$$\frac{1}{\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)u + \left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right)v$$

$$- \hat{\beta}v\tan\Theta + \omega\frac{\partial w}{\partial z} = 0, \tag{18.A14}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{\epsilon}{\varepsilon}\frac{u}{\cos\Theta}\left(\frac{\partial}{\partial\phi} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Phi}\right)$$

$$+ \frac{\epsilon}{\varepsilon}v\left(\frac{\partial}{\partial\theta} + \hat{\beta}\tan\Theta\frac{\partial}{\partial\Theta}\right) + \omega w\frac{\partial}{\partial z}.$$

Equations (18.A9)-(18.A14) are the Boussinesq hydrostatic equations.

Finally, the quasi-geostrophic β-plane approximation assumes that $\hat{\beta} \sim \varepsilon \sim \epsilon \ll 1$ and (by necessity) that ω

must also be of this order. We set $\omega = \varepsilon$ and expand in powers of ε to find the lowest-order balances

$$u^{(0)} = -\frac{\partial \psi^{(0)}}{\partial \theta},$$

$$v^{(0)} = \frac{1}{\cos\Theta}\frac{\partial \psi^{(0)}}{\partial \phi},$$

$$b_p^{(0)} = \frac{\partial \psi^{(0)}}{\partial z}$$

and the continuity equation

$$\frac{1}{\cos\Theta}\frac{\partial u^{(0)}}{\partial \phi} + \frac{\partial v^{(0)}}{\partial \theta} = 0$$

which is consistent with the geostrophic equations. At first order we have the momentum equations

$$\varepsilon\frac{\hat{D}u^{(0)}}{Dt} - \varepsilon v^{(1)} = -\frac{1}{\cos\Theta}\frac{\partial}{\partial \phi}\varepsilon\psi^{(1)} - \frac{\hat{\beta}\tan\Theta}{\cos\Theta}\frac{\partial}{\partial \Phi}\psi^{(0)},$$

$$\varepsilon\frac{\hat{D}v^{(0)}}{Dt} + \varepsilon u^{(1)} = -\frac{1}{\sin\Theta}\frac{\partial}{\partial \theta}\varepsilon\sin\Theta\psi^{(1)} - \frac{\hat{\beta}}{\cos\Theta}\frac{\partial}{\partial\Theta}\sin\Theta\psi^{(0)},$$

$$\frac{\hat{D}}{Dt} = \frac{\partial}{\partial t} + \frac{\epsilon}{\varepsilon}\frac{u^{(0)}}{\cos\Theta}\frac{\partial}{\partial\phi} + \frac{\epsilon}{\varepsilon}v^{(0)}\frac{\partial}{\partial\theta},$$

from which we form the vorticity equation

$$\varepsilon\frac{\hat{D}}{Dt}\left(\frac{v_\phi^{(0)}}{\cos\Theta} - u_\theta^{(0)}\right) + \varepsilon\left(\frac{u_\phi^{(1)}}{\cos\Theta} + v_\theta^{(1)}\right)$$

$$= -\hat{\beta}\frac{1}{\cos^2\Theta}\frac{\partial}{\partial\Theta}\sin\Theta\frac{\partial\psi^{(0)}}{\partial\phi} + \hat{\beta}\frac{\sin\Theta}{\cos^2\Theta}\frac{\partial}{\partial\Phi}\frac{\partial\psi^{(0)}}{\partial\theta}$$

$$= -\hat{\beta}v^{(0)} - \hat{\beta}\frac{\sin\Theta}{\cos\Theta}\frac{\partial v^{(0)}}{\partial\Theta} + \hat{\beta}\tan^2\Theta v^{(0)} - \hat{\beta}\frac{\sin\Theta}{\cos^2\Theta}\frac{\partial u^{(0)}}{\partial\Phi}.$$

Combining this with the first-order continuity equation

$$\frac{\varepsilon}{\cos\Theta}u_\phi^{(1)} + \varepsilon v_\theta^{(1)} + \hat{\beta}\frac{\sin\Theta}{\cos^2\Theta}u_\Phi^{(0)} + \hat{\beta}\tan\Theta\frac{\partial v^{(0)}}{\partial\Theta}$$

$$-\hat{\beta}v^{(0)}\tan^2\Theta + \varepsilon\frac{\partial w^{(0)}}{\partial z} = 0$$

(note that for the atmosphere, we would have an additional term $-\Delta_s\epsilon w^{(0)}$ in this equation) leads to the vorticity equation

$$\frac{\hat{D}}{Dt}\left(\frac{v_\phi^{(0)}}{\cos\Theta} - u_\theta^{(0)}\right) + \frac{\hat{\beta}}{\epsilon}v^{(0)} = w_z^{(0)}.$$

The lowest-order potential-buoyancy equation

$$\frac{\hat{D}}{Dt}\psi_z^{(0)} + \hat{S}w^{(0)} = 0$$

can now be combined with the vorticity equation to give the quasigeostrophic conservation equation:

$$\frac{\hat{D}}{Dt}\left[\frac{1}{\cos\Theta}\frac{\partial}{\partial\phi}\frac{1}{\cos\Theta}\frac{\partial}{\partial\phi}\psi^{(0)} + \frac{\partial^2}{\partial\theta^2}\psi^{(0)} + \frac{\partial}{\partial z}\frac{1}{\hat{S}}\frac{\partial}{\partial z}\psi^{(0)}\right]$$

$$+\frac{\hat{\beta}}{\epsilon}\frac{\psi_\Phi^{(0)}}{\cos\Theta} = 0.$$

For the atmosphere, the additional term in the continuity equation appears as an extra contribution

$$-\frac{\Delta_s}{\hat{S}}\frac{\partial}{\partial z}\psi^{(0)}$$

in the potential vorticity. Alternatively, the thickness term can be written as

$$\bar{\alpha}\frac{\partial}{\partial z}\frac{1}{\bar{\alpha}\hat{S}(z)}\frac{\partial}{\partial z}\psi$$

for atmospheric quasi-geostrophic motions.

Notes

1. Unfortunately, meteorologists use "mesoscale" very differently from oceanographers. We shall use mesoscale in the oceanic sense to refer to motions that are dynamically analogous to the "synoptic" scale motions of the atmosphere.

2. Note that the U here is characteristic of the disturbances, not of the mean flow.

3. Some care needs to be used in the cnoidal wave case since the mean depth of the fluid becomes $H + (1/2\pi L)\int_0^{2\pi L} dx\,\eta$ and the last term does not vanish. For the figures we have corrected for this effect to show c nondimensionalized by βL_R^2 where L_R is based on the actual average depth. Thus we have plotted $c_{\text{actual}} = (1 - \epsilon\langle\eta\rangle/\hat{S}_{\text{actual}})c$ as a function of $\hat{S}_{\text{actual}} = \hat{S} + \epsilon\langle\eta\rangle$, and also subtracted out the mean from the plots of $\phi^{(0)}$. The same process was used for the nonlinear Rossby wave but had no effect on the dispersion relation.

4. The "anelastic equations" (Batchelor, 1953b; Ogura and Charney, 1962; Ogura and Phillips, 1962) filter only acoustic waves.

5. It is customary to call any quasi-geostrophic wave a Rossby wave.

6. Geisler and Dickinson (1975) studied critical layer absorbtion in the western boundary current but did not explicitly include a reflected wave. It is also possible that the effects of the mechanisms maintaining the western boundary current are important in the interaction process.

7. It can be shown trivially that there is no nonlinear interaction in a spectrum of Rossby waves with all components having the same scale $(k^2 + \lambda^2)^{-1/2}$ (including baroclinic effects).