The Instability of Rossby Basin Modes and the Oceanic Eddy Field*

J. H. LA CASCE
Norwegian Meteorological Institute, Blindern, Oslo, Norway

J. PEDLOSKY
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

(Manuscript received 29 August 2003, in final form 10 March 2004)

ABSTRACT

Low-frequency, large-scale baroclinic Rossby basin modes, resistant to scale-dependent dissipation, have been recently theoretically analyzed and discussed as possible efficient coupling agents with the atmosphere for interactions on decadal time scales. Such modes are also consistent with evidence of the westward phase propagation in satellite altimetry data. In both the theory and the observations, the scale of the waves is large in comparison with the Rossby radius of deformation and the orientation of fluid motion in the waves is predominantly meridional. These two facts suggest that the waves are vulnerable to baroclinic instability on the scale of the deformation radius. The key dynamical parameter is the ratio $Z$ of the transit time of the long Rossby wave to the $e$-folding time of the instability. When this parameter is small the wave easily crosses the basin largely undisturbed by the instability; if $Z$ is large the wave succumbs to the instability and is largely destroyed before making a complete transit of the basin. For small $Z$, the instability is shown to be a triad instability; for large $Z$ the instability is fundamentally similar to the Eady instability mechanism. For all $Z$, the growth rate is on the order of the vertical shear of the basic wave divided by the deformation radius. If the parametric dependence of $Z$ on latitude is examined, the condition of unit $Z$ separates latitudes south of which the Rossby wave may successfully cross the basin while north of which the wave will break down into small-scale eddies with a barotropic component. The boundary between the two corresponds to the domain boundary found in satellite measurements. Furthermore, the resulting barotropic wave field is shown to propagate at speeds about 2 times as large as the baroclinic speed, and this is offered as a consistent explanation of the observed discrepancy between the satellite observations of Chelton and Schlax and simple linear wave theory. Here it is suggested that Rossby basin modes, if they exist, would be limited to tropical domains and that a considerable part of the observed midlatitude eddy field north of that boundary is due to the instability of wind-forced, long Rossby waves.

1. Introduction

Evidence of westward phase propagation in satellite altimetry data as described by Chelton and Schlax (1996) has revived interest in oceanic Rossby waves. The observations of Chelton and Schlax would seem to confirm earlier suggestions of the existence of long Rossby waves from the analysis of hydrographic data (e.g., White 1977; Kessler 1990). The waves clearly emanate from the eastern boundary of the North Pacific Ocean and traverse the entire basin. The satellite data, though, are particularly intriguing since they appear to show the unambiguous propagation of the long waves in a basin-wide region only south of about 25°N, while in midlatitudes the altimetry data seem rather to suggest an eddy rather than a wave field. Qiu et al. (1997) have suggested that the apparent confinement of the Rossby waves to the eastern boundary of the ocean in the North Pacific may be related to the dissipation of the relatively slow moving Rossby wave whose long-wave speed decreases with increasing latitude. Qiu et al. explored that process in a model in which the dissipation mechanism is specified a priori as a diffusion of momentum with a specified mixing coefficient. In the present study we examine the issue in a mechanistic manner by describing the dissipation process directly in terms of the baroclinic instability of the Rossby wave itself and describe the dissipation and confinement of the Rossby wave as the breakdown of the wave as it transfers energy to a small-scale eddy field.

Recent theoretical work on the problem of Rossby basin modes (LaCasce 2000; Cessi and Primeau 2001;
LaCasce and Pedlosky (2002) has isolated a new class of basin modes particularly resistant to dissipation mechanisms that preferentially damp small scales, such as the type employed by Qiu et al. (1997). While basin modes typically require the synthesis of long Rossby waves with westward-propagating group velocity and short Rossby waves with eastward group velocity, these new modes closely resemble free, long Rossby waves with zonal wavelengths that are integral multiples of the basin width. Such waves satisfy the boundary condition of constant streamfunction on eastern and western boundaries without the need for small-scale reflected Rossby waves. Relatively minor corrections to the free Rossby wave pattern are required in narrow boundary layer regions on the northern and southern boundaries of the idealized basins employed in the theory. The modes are dynamically completed by either boundary Kelvin waves in primitive equation models (e.g., Cessi and Primeau 2001) or by an equivalent integral mass conservation condition if quasigeostrophic dynamics are used (e.g., Flierl 1977; Kamenkovich and Kamenkovich 1993). As shown by LaCasce and Pedlosky (2002), the modes are easily excited by an oscillating wind stress and are fairly robust even to changes in basin geometry. The absence of small-scale Rossby waves as a component of these modes leads to their weak dissipation, and the suggestion has been advanced in the theoretical papers cited above that these weakly damped modes are capable of providing efficient coupling mechanisms to the atmosphere since the signal so imposed on the ocean by, say, the wind forcing can endure long enough to propagate such imposed anomalies throughout the oceanic basin and react back on the atmosphere.

The similarity of the basin mode structure to a free Rossby wave makes the observational evidence of long Rossby wave propagation particularly encouraging as to the possibility of the existence of such basin modes. However, as is evident in the theory, the constraint that the geostrophic streamfunction be constant on the eastern boundary of the basin imposes a structure on the westward-propagating waves such that the crest lines mimic strongly the shape of the eastern boundary of the basin. For basins whose boundaries are mainly north–south, this leads to advancing wave crests that are themselves oriented in the meridional direction. In quasigeostrophic theory that orientation is maintained as the wave propagates while in primitive equation models maintaining the full variation of the Coriolis parameter the waves do bend with latitude, reflecting the faster wave speed at low latitudes. In either theory the north–south scale is large in comparison with a deformation radius as is the zonal wavelength of the wave. This implies that the fluid motion in the propagating Rossby wave is largely in the meridional direction. Such broad flows are expected to be particularly vulnerable to baroclinic instability since the meridional shear is a manifestation of available potential energy in a zonal density gradient. That energy can be released by perturbation motions that are zonal and hence largely immune to the stabilizing effect of β, the planetary vorticity gradient.

We describe in this paper the basic instability process as a function of the amplitude of the basic Rossby mode. We show that since the free mode is always unstable the important parameter for this problem is the ratio of the time $T_R$ taken to traverse the basin by the Rossby wave to the $e$-folding growth time of the instability $\sigma^{-1}$, where $\sigma$ is the growth rate of the instability. Thus the critical parameter of our analysis is

$$Z = \frac{\sigma}{\sigma}. \quad (1.1)$$

When $Z < 1$ the wave traverses the basin before its instability can substantially degrade the wave. When $Z > 1$ the wave will be shown to break up into deformation-scale eddies. Our analysis addresses the instability of a plane, westward-propagating Rossby wave as well as the basin mode.

In section 2 we define the quasigeostrophic model we will use to describe the instability. Scaling of the problem exposes the centrality of $Z$. Section 3 is a description of the instability of the free, long Rossby wave over the whole range of $Z$. Both analytical and numerical methods are used. In section 4 we describe the instability of the basin modes and show how the existence of the basin mode depends on the value of $Z$. In section 5 we apply our results in a heuristic manner to delimit the latitude regions where the long Rossby wave can succeed in crossing the basin and use those results to explain the observations of Chelton and Schlax (1996).

2. The model

To simplify the analysis we will consider the problem in the context of the standard, quasigeostrophic two-layer model. It is certainly true that a more realistic multilayer model would be advantageous in describing the structure of the basic baroclinic Rossby wave and its instability but in this first approach to the problem the advantages of simplicity are compelling. The model consists of two layers on the beta plane with resting depths $H_1$ and $H_2$. The characteristic horizontal scale of the basin and of the zonal wavelength of the Rossby wave is $L$ and its characteristic velocity is $U$. In terms of these scales, the geostrophic streamfunction is non-dimensionalized with $UL$ while lengths are scaled with $L$ and time with the advective time $LU$. It is convenient to write the equations in terms of the barotropic and baroclinic streamfunctions. Thus, if $\phi_n$, $n = 1, 2$, are the (non-dimensional) streamfunctions of the upper and lower layers, respectively, the barotropic and baroclinic streamfunctions are

$$\phi_1 = h_1 \phi_1^1 + h_2 \phi_2^1 \quad \text{and} \quad \phi_1 = \phi_1^1 - \phi_2^1, \quad (2.1a)$$

where the $h_n$ are the rest-layer thicknesses divided by
the total thickness of the two layers. The equations of motion then become
\[ \frac{\partial}{\partial t} q_b + \beta \phi_{b_s} + J(\phi_b, q_b) + h_1 h_2 J(\phi_t, q_t) = 0 \]
and
\[ \frac{\partial}{\partial t} q_T + \beta \phi_{t_s} + J(\phi_t, q_b) + J(\phi_b, q_T) + (h_2 - h_1) J(\phi_t, q_T) = 0, \]
where
\[ q_b = \nabla^2 \phi_b \quad \text{and} \quad q_T = \nabla^2 \phi_t - F \phi_t, \]
and the nondimensional parameters \( F \) and \( \beta \) are defined in terms of the scales already described as well as the reduced gravity \( g' \) and the Coriolis parameter \( f_o \),
\[ F = \frac{f^2 L^2}{g' H_1 H_2 (H_1 + H_2)} = \frac{L^2}{L_d^2} \quad \text{and} \quad \beta = \frac{\beta_{\text{dim}} L^2}{U}, \]
where \( L_d \) is the deformation radius.

Note that for unequal layer thicknesses the self interaction of the baroclinic streamfunction will generate changes in the baroclinic motion.

It is convenient to anticipate certain aspects of the analysis. For a large-scale baroclinic Rossby wave, the transit time \( T_R \) for the wave to cross a basin of width \( L \) will be on the order of \( L/c_R \) where \( c_R = \beta_{\text{dim}} L_d^2 \). Thus,
\[ T_R = \frac{L}{\beta_{\text{dim}} L_d^2}, \]
which is also the characteristic period of the Rossby wave with wavelength \( L \). Also, since the motion in the wave is largely meridional, energy releasing instabilities can be anticipated for which the stabilizing effects of \( \beta \) will be weak and so the expected characteristic growth rate for an instability will be the baroclinic shear times the wavenumber of the instability, which in turn can be expected to be of the order of the deformation radius. Hence, an additional natural time scale is the growth time,
\[ T_g = \frac{L_d}{U}. \]

It is useful to rewrite the problem in terms of those two time scales. We expect the streamfunction to be a function of time on each time scales and we exploit that explicitly by writing the streamfunction as a function of the two times,
\[ t_r = \frac{t_{\text{lim}}}{T_R} = \frac{L}{U} t/T_R = \frac{\beta_{\text{dim}} L_d^2}{U} t = bt \quad \text{and} \quad (2.7a) \]
\[ t_g = \frac{t_{\text{lim}}}{T_g} = \frac{L}{U} t/T_g = \frac{L}{L_d} t = F^{1/2} t. \quad (2.7b) \]

Since the instabilities will be expected to have meridional scales on the order of the deformation radius, it is also helpful to introduce a new meridional variable,
\[ s = y F^{1/2}. \quad (2.8) \]
For all realistic parameter settings, the ratio of the basin scale to the deformation radius is large—that is, \( F \gg 1 \)—so that \( \varepsilon = F^{-1/2} \) is small. The parameter \( b \), which is a measure of the amplitude of the baroclinic wave, may be large or small depending on the wave amplitude. The ratio
\[ Z = \frac{F^{1/2}}{b} = \frac{T_R}{T_g} \quad (2.9) \]
is the same parameter as defined in (1.1). With the above definitions the equations of motion become
\[ \left( \frac{\partial}{\partial t} + Z \frac{\partial}{\partial s} \right) q_b + \phi_{b_s} + Z [J(\phi_b, q_b) + h_1 h_2 J(\phi_t, q_t)] = 0 \]
\[ + (h_2 - h_1) J(\phi_t, q_T) = 0, \quad (2.10a) \]
and
\[ \left( \frac{\partial}{\partial t} + Z \frac{\partial}{\partial s} \right) q_T + \phi_{t_s} + Z [J(\phi_b, q_T) + J(\phi_t, q_b)] \]
\[ + (h_2 - h_1) J(\phi_t, q_T) = 0, \quad (2.10b) \]
where now
\[ q_b = \phi_{b_s} + \varepsilon^2 \phi_{b_s} \quad \text{and} \quad (2.11a) \]
\[ q_T = \phi_{t_s} - \phi_t + \varepsilon^2 \phi_{t_s}, \quad (2.11b) \]
and where the Jacobian operators in (2.10a) and (2.10b) are in terms of the variables \( x \) and \( s \) instead of \( x \) and \( y \).
Each streamfunction is thus a function of \( x, s, t_r, \) and \( t_g \).

It is immediately evident that the nature of the instability problem, assuming our a priori presumptions are correct, is governed by two parameters \( Z \) and the small parameter \( \varepsilon \). Here \( Z \), which may be large or small, can be interpreted either as the ratio of the transit time of the baroclinic wave to the growth time of parasitic baroclinic instabilities or the ratio of its period to that growth time. The parameter \( \varepsilon \) is always small so that the problem in the limit of small \( \varepsilon \) really depends primarily on \( Z \).

In the following sections we shall examine the instability of a free, long baroclinic Rossby wave as an elementary model of the essential structure of the basin mode and then turn our attention to the instability of the basin mode itself. The analytical theory is supplemented by a direct numerical integration of (2.2a) and
(2.2b) that serves to test our a priori scaling assumptions and describes aspects of the problem inaccessible to analytical analysis.

Two numerical models are used, one for examining plane waves and the second for basin modes. The former is a two-layer spectral model, written by G. Flierl, with periodic boundary conditions in x and y. The latter is a two-layer basin model, derived from the barotropic model of LaCasce (2002). This uses finite differences to calculate spatial derivatives and sine transforms to invert the barotropic and baroclinic vorticities, as well as a third-order “QUICK” scheme (Leonard 1979) for advection. The basin model uses third-order Adams–Bashforth time stepping, whereas the spectral model employs a leapfrog time step, stabilized by an occasional Euler step.

We used either 128^2 or 256^2 Fourier modes in the spectral runs, and 256^2 grid points in the basin. In all cases the resolution was more than sufficient to resolve the evolution.

3. Rossby wave instability in the infinite domain

a. Triad instability for small Z

It is clear from (2.10a) and (2.10b) that for small Z the lowest-order problem is simply the problem for linear, free barotropic and baroclinic Rossby waves. The instability that follows is a type of resonant triad instability first discussed in the barotropic context by Gill (1974). The baroclinic problem has previously been discussed by Jones (1979) and Vanneste (1995). We recapitulate the essence of the triad analysis to emphasize results of particular pertinence to the basin mode problem of this paper, in particular the behavior at large F.

We expand the streamfunction in a series,

\[ \phi_f = \phi_f^{(0)} + Z\phi_f^{(1)} + \cdots \]  
and

\[ \phi_b = \phi_b^{(0)} + Z\phi_b^{(1)} + \cdots. \]  

(3.1a)

(3.1b)

In the infinite domain, one solution at lowest order, which we identify with the basic baroclinic wave whose stability is under investigation, is

\[ \phi_f^{(0)} = Ae^{i(kx - \omega t)}, \]  

where the asterisk denotes the complex conjugate of the preceding function. The solution of the linear equation yields the simple dispersion relation

\[ \omega = -k, \]  

(3.3)

corresponding to the Rossby wave traveling westward with the long wave speed. Since the lowest-order problem is linear, two further solutions may be added, corresponding to a second baroclinic and a barotropic wave, each with an s wavenumber \( l_s \). Thus the total solution at \( O(1) \) is

\[ \phi_f^{(0)} = Ae^{i(kx - \omega t)} + A_s e^{i(kx + l_s x - \omega t)}, \]  
and

\[ \phi_b^{(0)} = A_s e^{i(kx + l_s x - \omega t)} + *, \]  

(3.3a)

(3.3b)

Each wave satisfies the linear dispersion relation so that

\[ \omega = -\frac{k_o}{l_o^2 + 1} \]  

(3.4a)

\[ \omega_1 = -\frac{k_1}{l_1^2} \]  

(3.4b)

It is important to note that each amplitude, \( A, A_s, \) and \( A_1 \) is an unknown function of the growth time \( t_g \), which for small Z is a slow time as compared with \( t_r \). At the next order in Z the nonlinear interactions of the two baroclinic waves will resonate with the barotropic wave and the barotropic wave will interact with each baroclinic wave to resonantly force the other baroclinic wave. This will occur only under the conditions of resonance:

\[ k + k_s = k_1 \]  

(3.5a)

\[ k + \frac{k_o}{k_o^2 + 1} = \frac{(k + k_s)}{l_s^2}, \]  

(3.5b)

by which the wavenumbers and frequencies form a closed resonant triad set. Such resonance would render the expansion in (3.1a) and (3.1b) invalid unless the resonant forcing terms are balanced. That balance yields the evolution equations for the amplitudes on the growth time; that is,

\[ \frac{\partial A_1}{\partial t_g} = h_1 h_2 k l_A A_1, \]  

(3.6a)

\[ \frac{\partial A_s}{\partial t_g} = k l_A \frac{(1 - l_s)}{(1 + l_s)} A_s A_1^*, \]  

(3.6b)

\[ \frac{\partial A}{\partial t_g} = -k l_A A_1 A_s^*. \]  

(3.6c)

The instability for small Z of the basic wave can be easily deduced by linearizing the set (3.6) around the basic wave—for example,\[ A = \bar{A} + a, \]  

(3.7a)

\[ A_s = a_s, \]  

(3.7b)

\[ A_1 = a_1, \]  

(3.7c)

where the \( a \) are small with respect to the amplitude of the basic wave \( \bar{A} \). The resulting linear equations yield exponential growth for \( a_s \) and \( a_1 \) with growth rate

\[ \lambda = l_s \left[ h_1 h_2 \frac{(1 - l_s)}{(1 + l_s)} \right]^{1/2} k |\bar{A}|. \]  

(3.8)

Thus instability is assured if the s wavenumber is less than 1. Since the growth rate vanishes for zero s wavenumber, there is an intermediate wavenumber,

\[ l_s = (2^{1/2} - 1)^{1/2} \approx 0.644, \]  

(3.9a)

that maximizes the growth rate. The growth rate also
FIG. 1. (a) The growth rate curve for the triad instability of the long Rossby wave is shown by the dashed line. The solid curve gives the accompanying $y$ wavenumber for the instability (divided by $F^{1/2}$). In the case shown, the basic wave has a wavenumber $6\pi$ and $F = 9870$. (b) The evolution of the wave amplitudes of the triad on the growth rate time scale.

depends linearly on $k|\vec{A}|$, which is related to the amplitude of the meridional velocity in the basic wave. If the basic baroclinic wave has the form $\phi_x = (V_o/k) \cos[k(x - t)]$ then $k|\vec{A}| = V_o/2$. This leads to a maximum growth rate

$$\lambda_{\max} = 0.207(h_1h_2)^{1/2}V_o/2. \quad (3.9b)$$

The growth rate is a maximum when the two layers have equal depth, but the variation in growth rate with other realistic values changes only slightly (Fig. 5, below). The growth rate must be multiplied by $F^{1/2}$ when the growth on the advective time $t$ is considered.

The resonance conditions (3.5a) and (3.5b) imply a relation between the $x$ and $s$ wavenumbers. In particular, for a given $x$ wavenumber $k$ of the basic wave,

$$k_s/k = l_s^4 - 1 \quad \text{and} \quad k_1 = k_s + k, \quad (3.10)$$

so that, for the most unstable perturbation,

$$k_s = -0.8284k \quad \text{and} \quad (3.11a)$$

$$k_1 = 0.17156k. \quad (3.11b)$$

Hence the parasitic baroclinic wave will have an $x$ wavenumber of the same order as the basic wave while the barotropic portion of the disturbance will have a relatively small $x$ wavenumber. The $y$ wavenumber is $l_yF^{1/2}$ in original $y$ units and hence the $y$ scale will be very short. Figure 1a shows the growth rate as a function of $k_s$. The dashed curve is the growth rate (on the advective time $t$) for the parameters $F = 9870$ and $k$ of the original wave equal to $6\pi$. The solid curve shows
the corresponding \( s \) wavenumber. The peak of the growth rate curve corresponds to a value of \( l_o \) as predicted by (3.9).

The growth rate predicted by (3.8) is reminiscent of the Eady problem. It is not, however, the same as the Eady growth rate. If the current were a broad, steady meridional flow, a standard stability analysis would yield as the growth rate,

\[
\lambda_{\text{Eady}} = \frac{V}{4} \sqrt{\frac{4h_1h_2 - l_o^2}{1 + l_o^2}},
\]

which coincides with (3.8) only if \( h_1 = h_2 \). The discrepancy comes from the nonlinear term in (2.10b) that has \( h_2 - h_1 \) as a factor and that does not yield a resonant term for the triad interaction but that enters the classical Eady problem. The qualitative similarity is nevertheless clear and we can identify the triad instability as a classic baroclinic instability emerging on times long in comparison with the basic wave period. That is, the \( \beta \) effect is unable to stabilize the basic wave even for small \( Z \) where the \( \beta \) effect is dominant in size. Even though the basic wave is unstable it can still propagate several of its own wavelengths before the instability would be noticeable. The decay of the basic wave amplitude will be an order amplitude squared effect of the instability and hence will remain small at least on the transit time \( T_R \).

The evolution of the amplitude of each member of the triad is shown in Fig. 1b. During the initial, exponential growth phase given by linear theory the amplitude of the original baroclinic wave is nearly unchanged. It then diminishes as the parasitic instability waves grow and finally equilibrate. The amplitudes execute a continuing nonlinear oscillation, but we should realistically expect an irreversible effect to occur as the waves in finite amplitude lock on to other triads with which they can also exchange energy. Hence we show only the first part of the growth and equilibration phase in the figure. We should therefore expect the wave to successfully cross the basin displaying only slight alterations due to its instability when \( Z \) is small.

**b. Plane wave instability for large \( Z \)**

For large \( Z \) the \( \beta \) effect is negligible at lowest order. The basic baroclinic wave can be represented by the solution

\[
\psi_b = (V/k) \sin kx,
\]

since the phase of the wave is irrelevant in the infinite domain and the propagation of the wave on the time scale \( T_R \) is negligible on the time scale for growth for small \( Z \). If the perturbations to the basic wave are written as

\[
\phi_b = e^{i\omega} e^{ikx} A_b(x) \quad \text{and} \quad \phi_f = e^{i\omega} e^{ikx} A_f(x),
\]

then the linearized equations for the wave amplitudes in this limit are (for the case \( h_1 = h_2 \))

\[
\lambda \left[ \varepsilon^2 A_{bs} - (l^2 + 1)A_f \right] + ilV \cos kx \times \left[ \varepsilon^2 A_{bs} + A_f(1 - l^2 + \varepsilon^2 k^2) \right] = 0 \quad (3.15a)
\]

and

\[
\lambda \left[ \varepsilon^2 A_{bs} - (l^2)A_f \right] + ilh_1 h_2 V \cos kx \times \left[ \varepsilon^2 A_{bs} + A_f(-l^2 + \varepsilon^2 k^2) \right] = 0. \quad (3.15b)
\]

We have found it possible to find solutions in the form

\[
A_b = \sum_{m=0,M_{\text{max}}} A_{bs} \cos mkx \quad \text{and} \quad A_f = \sum_{m=1,M_{\text{max}}} A_{fs} \cos mkx. \quad (3.16a)
\]

The resulting matrix equations for the Fourier coefficients, after truncation at \( m = M_{\text{max}} \) yield a straightforward eigenvalue problem for the growth rate as a function of \( s \) wavenumber. Figure 2a shows the growth-rate curves for the same values of \( F \) as in the low \( Z \) triad case with a truncation corresponding to \( M_{\text{max}} = 25, \)
which tests show is an adequate resolution for convergence. The peak of the most unstable mode occurs at nearly the same meridional wavenumber as before while the maximum growth rate in comparison with the low $Z$ triad case is somewhat increased; it is 17.3 instead of 10.35, an increase of about 68%. Figure 2b shows the $x$ structure of the barotropic part of the disturbance streamfunction. The mode is dominated by a structure that has 2 times the $x$ wavenumber of the fundamental wave. This wavenumber is still very small when compared with the meridional wavenumber, which, again, is on the order of $L_d^{-1}$. Thus, over the whole range of $Z$ the growth rate is an order-1 constant multiplied by the amplitude of the baroclinic velocity in the wave divided by the deformation radius. An example of the evolution, from a numerical run with $Z = 5$, is shown in Fig. 3. The baroclinic wave (left panels) is seen to develop zonally oriented wiggles before dissolving entirely into eddies. The barotropic field (right panels) quickly evolves from the isotropic random initial state to one dominated by zonally elongated eddies.

We have carried out calculations over a wide range of $Z$, and the resulting growth rates as a function of $Z$ are shown in Fig. 4a. The corresponding analytical rates in the limits of large and small $Z$ are indicated by the lines. The experimental values are close to predicted in the two limits, although somewhat smaller. The difference varies with the choice of initial barotropic field;
somewhat faster growth is obtained when initializing with the most unstable barotropic wave. Note that the growth rates at intermediate $Z$ lie between those at the two extremes (i.e., nothing unusual happens at $Z = 1$). The weak variation with $Z$ is remarkable.

In Fig. 4b, we show the variation of growth rate with $F$ for two values of $Z$. In both cases, the $F^{1/2}$ dependence, expected from theory, is apparent. Again, the growth is faster with larger $Z$. As mentioned above, the growth rates are a very weak function of the relative layer depths. Figure 5 shows the dependence of the growth rates, as determined by the numerical model, over the range of layer depth ratio for both small ($Z = 0.25$) and large ($Z = 2$) values of $Z$.

Interestingly, the value of $Z$ also determines the character of the nonlinear evolution. If $Z > 1$, triadic interactions among other barotropic modes, subsequent to the initial growth, produce a nearly isotropic barotropic eddy field. These eddies merge, producing larger eddies; that is, there is an inverse energy cascade to larger scales. The cascade is halted at a larger scale by beta, the so-called arrest described by Rhines (1975) and others. The end state is one of weakly zonally elongated barotropic eddies that exhibit westward phase propagation at the barotropic phase speed corresponding to their size. The eddies are larger than deformation scale, the more so for larger values of $Z$ ($Z$ determines the arrest scale).

If, on the other hand, $Z$ is order-1 or smaller, the barotropic eddies retain the zonally elongated aspect of the barotropic perturbation and intensify as such. In this case beta is large enough so that the barotropic field is already “arrested” at the outset and no further spectral evolution occurs. The barotropic eddies at late times thus retain their deformation scale meridional width. This is essentially what has happened in Fig. 3.

4. The instability of the basin modes

We turn our attention now to the instability of the low-frequency, large-scale basin modes described by LaCasce (2000) and Cessi and Primeau (2001). As we have already noted, the form of such basin modes closely resembles the latitude-independent baroclinic wave whose instability was examined in the previous section (see, e.g., Fig. 3 in LaCasce and Pedlosky 2002). We therefore can expect strong similarities between the stability properties of the basin modes and the free, infinite-domain baroclinic Rossby wave. We can make that similarity explicit by first examining, as did LaCasce (2000)
a simple model in which the basin is replaced by a meridional channel with solid boundaries at \( x = 0 \) and \( x = 1 \) (\( L \) in dimensional units). An exact solution of (2.2a) and (2.2b) that satisfies the condition that the streamfunction be spatially constant on the meridional boundaries and that satisfies the integral condition

\[
\frac{\partial}{\partial t} \int_0^1 \psi_f \, dx \, dy = 0 \quad (4.1)
\]
is

\[
\psi_f = \frac{V_o}{k} \sin(kx - \omega t), \quad (4.2a)
\]
\[
\omega = -\frac{\beta k}{k^2 + F}, \quad \text{and} \quad (4.2b)
\]
\[
k = 2j\pi, \quad \text{for } j = 1, 2, 3 \ldots \quad (4.2c)
\]

For large \( Z \) the stability problem is described by (3.15a) and (3.15b) except that now the basic wave has an arbitrary phase \( \vartheta \) with respect to the channel boundary at \( x = 0 \):

\[
\lambda [\varepsilon^2 A_{\tau,\vartheta} - (I^2 + 1)A_f] + ilV_o \cos(kx + \vartheta)[\varepsilon^2 A_{\vartheta,\vartheta} + A_{\vartheta}(1 - I^2 + \varepsilon^2 k^2)] = 0 \quad (4.3a)
\]

and

\[
\lambda [\varepsilon^2 A_{\vartheta,\vartheta} - (I^2)A_{\vartheta}] + ilh_1h_2V_o \cos(kx + \vartheta) \times [\varepsilon^2 A_{\tau,\tau} + A_{\tau}(-I^2 + \varepsilon^2 k^2)] = 0, \quad (4.3b)
\]

where on the time scale of the instability the phase of the wave, \( \vartheta = -\omega t \), is a constant to the first approximation.

The boundary conditions are that the perturbation must vanish at \( x = 0 \) and 1. Thus solutions can be sought in the form

\[
A_f(x) = \sum_{n=1} A_{\tau,\vartheta} \sin n \pi x \quad \text{and} \quad (4.4a)
\]
\[
A_{\vartheta}(x) = \sum_{n=1} A_{\vartheta,\tau} \sin n \pi x. \quad (4.4b)
\]

This leads to two coupled matrix equations of the form

\[
\lambda [A_f] = [Q][A_{\vartheta}] \quad \text{and} \quad (4.5a)
\]
\[
\lambda [A_{\vartheta}] = [R][A_f], \quad (4.5b)
\]

where the matrices \( Q \) and \( R \) are given in the appendix. Combining the equations leads to a single matrix eigenvalue problem (A.4). The eigenvalues can be easily determined by any standard matrix eigenvalue package such as MATLAB and the results are shown in Fig. 6 for the growth rate for the unstable modes is shown as a function of (scaled) meridional wavenumber. The important result is that for large \( Z \) the growth rate is essentially the same as the growth rate for the instability of the baroclinic wave in the infinite domain (cf. with Fig. 2). Perhaps this is not too surprising since the instability is of such small meridional scale that the perturbations are largely insensitive to the detailed structure of the flow near the meridional boundaries. In the example shown, the phase \( \vartheta \) was chosen arbitrarily to be \( \pi/4 \) but the result is also insensitive to the phase.

The similarity of the growth rates for the channel model and the infinite-domain model suggest that the general estimates for the latter will hold for the basin modes for a wide range of \( Z \). If that is the case, then for small \( Z \) we would expect the basin mode to survive the instability as the interior long Rossby wave crosses the basin while for large \( Z \) the mode should suffer strongly from the small-scale instabilities.

To test this hypothesis we have initialized our numerical basin model with a baroclinic basin mode. The latter was calculated in the manner described by LaCasce and Pedlosky (2002) for a square (untilted) basin. A weak, barotropic perturbation (generated from a white-noise PV field) was superimposed to hasten unstable growth, and weak Rayleigh damping imposed on the baroclinic and barotropic relative vorticities; the damping coefficient was the same as that used in calculating the mode, and corresponded to a damping time of several hundred eddy turnover times (i.e., much longer than the typical instability time seen hereinafter).

To begin, we examined the evolution without beta, that is, infinite \( Z \). As with the plane wave, the initial stage is characterized by a nearly exponential growth in the barotropic energy. We examined how the growth rate scaled with \( F \); the case shown in Fig. 7, for the \( n = 3 \) basin mode, is typical. The growth is somewhat suppressed at smaller \( F \), evidently as the deformation radius approaches a significant fraction of the basin scale. At larger \( F \), however, the increase in growth rate asymptotically approaches an \( F^{1/2} \) dependence. Note that these growth rates are very close to those obtained for the plane wave in the spectral model at large \( Z \).
An example of the evolution with nonzero beta is shown in Fig. 8, for which \( Z = 5 \). The baroclinic basin mode (upper-left panel) resembles a plane wave except for the boundary layers at the north and south walls. In these regions, barotropic eddies form in the early stages (upper-right panel) because of the self-advection of baroclinic vorticity [the fourth term in (2.2a)]. These eddies remain near the boundary and do not significantly affect the interior evolution. However, because of them we calculated the growth rates in Fig. 7 using the barotropic kinetic energy only in a latitudinal band about the basin midsection. As time progresses, the baroclinic wave develops wiggles, as in the plane wave case, and deformation-scale barotropic eddies appear. The eddies grow initially to largest amplitude in the west, but soon the basic wave completely succumbs, leaving a nearly isotropic field of barotropic eddies throughout the interior (lower-right panel).

As with the plane wave, an inverse cascade also occurs in the basin, provided \( Z \) is large enough. The cascade is again halted by beta, at a scale determined by \( Z \). Now, however, the barotropic eddies remain isotropic. As discussed by LaCasce (2002), the arrest in a basin is accomplished by the barotropic normal modes rather than plane waves; this means that with \( Z > 1 \), we have the interesting situation of a baroclinic wave mode evolving to (finite amplitude) barotropic wave modes.

With \( Z < 1 \), the baroclinic waves propagate a significant distance before instability sets in, as suggested earlier. With \( Z = 0.25 \) (upper panels of Fig. 9), the baroclinic wave propagates nearly to the western boundary, and barotropic eddies are visible only there. With \( Z = 0.5 \), instability sets in roughly halfway across, and so the baroclinic wave is coherent only in the eastern half of the basin; west of that, the barotropic eddies dominate. So we may identify \( Z = 0.5 \) as an approximate boundary between the existence of the mode and its breakdown into eddies.

To reiterate, the basin baroclinic normal mode remains intact when \( Z \) is small; when \( Z < 0.25 \), the constituent baroclinic waves propagate unharmed across the basin. When \( Z > 0.5 \), the baroclinic mode dissolves by instability into barotropic eddies; the final flow, following an inverse cascade and a beta arrest, is dominated by finite-amplitude, barotropic basin modes.

### 5. Consequences of the instability

The parameter \( Z \) is the ratio of the travel time of a long Rossby wave to the \( \epsilon \)-folding time of the baroclinic instability of that wave and we have seen that the nature of the instability and the order of magnitude of the growth rate can be simply scaled with the amplitude of the wave and the deformation radius. More precisely, the dimensional growth rate \( \sigma \) can be written as the parameter \( \lambda \), the growth rate nondimensionalized with the advective time \( V/L \) of quasigeostrophic theory, where \( V \) is the characteristic fluid velocity in the wave; thus

\[
Z = \sigma T_a = \left( \frac{\lambda V L}{f L_d \beta_{\text{dim}}} \right)
\]

where \( \lambda \) is the growth rate on the advective time scale and \( V \) is the amplitude of the meridional shear in the wave. Note that \( L \) may be considered as both the characteristic zonal scale of the wave (hence the role of \( Z \) in determining the dynamical balances) and the order of magnitude of the basin width for large-scale modes.

At \( Z = 1 \), for which the transit time and the growth time are equal, the length \( L \) is given by

\[
L = \left( \frac{Z F^{1/2}}{\lambda} \right) L_d \beta_{\text{dim}} F^{3/2}.
\]

The ratio \( \lambda F^{1/2} \) is an order-1 number, ranging between 0.1 for small \( Z \) to about 0.17 for large \( Z \). In quasigeostrophic theory, \( L \), as defined by (3.2), is a constant. However, we may heuristically consider the variation with latitude of the critical \( L \) that makes \( Z = 1 \) by allowing the variation of \( \beta \) and of the deformation radius. Since \( L_d = ND/f \), the deformation radius is inversely proportional to the sine of latitude while \( \beta \) is proportional to the cosine of latitude. Figure 10 shows the resulting variation of \( L \) in thousands of kilometers as a function of latitude for \( Z = 0.5 \) and \( \lambda F^{1/2} = 0.17 \). We have chosen this value of \( Z \) since our results in Fig. 9 indicate this is a reasonable transition point between propagation and destruction of the wave. What is clear from Fig. 10 and relatively independent of the specific values of \( Z \) and \( \lambda \) is that the distance \( L \) is small at all latitudes outside the Tropics. Only the weakest waves can propagate large distances unaffected by the insta-
A long Rossby wave with a velocity amplitude of 5 cm s$^{-1}$ will reach a critical value of unit $Z$ at less than a thousand km, much less than the width of the Pacific basin. If the critical value of $Z$ is even smaller, as evidenced by the results of section 4, say $Z = 0.25$, the length will be one-half of that. We believe that this is why the large-scale wave features seen by Chelton and Schlax (1996) appear to be confined to low latitudes (see their Fig. 4). The dissipation of the waves suggested by Qiu et al. (1997) is, we believe, actually due to their instability, and that instability is enhanced at higher latitudes. The limitation to lower latitudes is due partly to
the lower propagation speed of the Rossby waves at higher latitudes but also due to the enhanced instability at higher latitudes; both effects combine to limit the distance the baroclinic Rossby wave can propagate. We emphasize that this argument holds whether the wave forms part of the low-frequency basin mode, resistant to ordinary dissipation (e.g., LaCasce 2000; Cessi and Primeau 2001), or whether it is a free Rossby wave generated in the midocean by, say, wind forcing.

We may draw further conclusions. In addition to altering the large-scale coherence and vertical structure of the eddy field, baroclinic instability would also change the dominant eddy length scales. Our results suggest this happens essentially in two stages. First, as the instability calculations of sections 3 and 4 show, the emerging barotropic field has a $y$ wavenumber of about $0.66L_{d}^{1/4}$, so that the daughter barotropic eddies are about 2 times the deformation radius. Second, as seen in the numerical experiments, those barotropic eddies merge if $Z > 1$, leading to an inverse energy cascade and a beta-arrest at a still larger scale. If $Z > 1$, the final eddies are barotropic and possibly much larger than the deformation radius.

Stammer (1997) calculated dominant eddy length
scales from Ocean Topography Experiment (TOPEX)/Poseidon data. His eddy scales were comparable to the deformation radius in the Tropics but were larger at higher-latitudes. However the scales at higher latitudes appeared to be proportional to the deformation radius (see his Fig. 25). In the latitude range of 20°–50°, the scales are approximately 2 times the deformation radius; at higher latitudes, they are somewhat larger (perhaps 3 times the radius). In addition, his length scales display no relation to the Rhines scale. As Stammer pointed out, it appeared that no beta arrest was occurring in the extratropical ocean.

Rossby wave instability could explain the length scales at higher latitudes, provided that the inverse cascade and beta arrest, seen in the models when $Z > 1$, is somehow defeated. The models, having flat bottoms and no mean flows, are ideal environments for an inverse cascade, but the real ocean may be less accommodating. If no cascade occurs, we would expect the eddy field to reflect that which emerges from the instability, with a scale of roughly 2 times the deformation radius. Stammer’s results suggest in addition that the tropical wave field is of deformation scale, something we could not have predicted; our results simply say that the tropical baroclinic waves should retain their structure.

A second point concerns the observed phase speeds. Because the eddies generated by instability are larger than the deformation radius and are barotropic, they will propagate faster than the baroclinic long wave speed. The emerging barotropic eddies, with a $y$ wavenumber of $0.66L_d^2$, would have a westward phase speed of about $2.25\beta L_d^2$ or about 2 times the baroclinic wave speed. A graphic example, shown in Fig. 11, reveals this. Shown is a Hovmöller diagram constructed from the upper-layer streamfunction from a plane wave computation with $Z = 0.25$, that is, within the wave regime. The propagation at early times is at the baroclinic phase speed corresponding to the initial wave (for which $k = 14\pi$). Following instability, however, the apparent phase speed increases to a value comparable to the aforementioned estimate (as indicated by the solid line in the figure). The numerical experiments in which $Z > 1$ reveal still faster phase speeds for the late-time barotropic field, because the cascade produces larger, faster-propagating eddies. If the cascade is defeated in the ocean, however, as suggested above, the late time phase speeds would be roughly 2 times the long-wave speed.

One of the most puzzling results of the Chelton and Schlax (1996) altimeter study was an apparent systematic increase in the Rossby phase speeds outside of the Tropics. This is plainly seen in their Fig. 5b. Numerous theoretical studies followed that invoked various effects to explain this increase [e.g., mean flow effects (Killworth et al. 1997); homogenized subsurface potential vorticity (Dewar 1998); topography (Tailleux and McWilliams 2001)]. While each explanation was appealing in its own right, it is difficult to see how any one would produce a systematic change over the entire Pacific basin. The present explanation—that is, a change in the wave structure due to instability—could produce such a systematic change. Waves emanating from the eastern boundary or generated in the interior would break up, and the eddy field at higher latitudes should be more barotropic and propagating faster. No mean flow is required.

In addition, however, the present explanation is the only one that also predicts a change in eddy scale, as seen by Stammer (1997). Indeed, the coincidence that the waves are propagating 2 times too fast in the extratropics but are also 2 times the deformation radius seems to have been overlooked. Furthermore, we would expect to see the phase speeds to accelerate in the basin interior where $Z = O(1)$ because long waves would be present in the eastern portion of the basin and barotropic eddies in the west. In the Pacific such an acceleration evidently occurs near 180°, between 15° and 25°N (Chelton and Schlax 1996; Leeuwenburgh and Stammer 2001). Last, of course, the wave instability conjecture predicts that the propagating eddies should be much more barotropic in the extratropics, and this, of course, is testable.

Perhaps most important though is that the parasitic instability of Rossby waves, no matter how they are generated, can contribute substantially to the vigor of the midocean eddy field in mid- and higher latitudes. We may go so far as to suggest that the eddy field at higher latitudes might in fact be a manifestation of the underlying presence of large-scale Rossby waves.
6. Summary and conclusions

We have examined the instability of long baroclinic Rossby waves, both in the infinite domain and as components of low-frequency baroclinic Rossby basin modes and have shown that such waves are unstable at all wave amplitudes. The $\beta$ effect is unable to render the wave stable. The central parameter that emerges is the product of the growth rate of baroclinic instability and the time of the Rossby wave to traverse a distance $L$. That parameter, $Z$ in our notation, distinguishes a parameter regime in which the instability only slightly affects the basic Rossby wave (small $Z$) from the regime ($Z > 1$) in which the wave is destroyed by the parasitic baroclinic instability. That criterion holds for both free waves and for the low-frequency basin modes previously suggested as being particularly immune to dissipation. We have found that contrary to such expectations these long waves will naturally give up much of their energy in latitudes outside the Tropics so that, at the very least, it would seem unlikely that such long-lived basin modes could persist in midlatitudes. Indeed, we suggest that the instability that should occur where $Z > 1$, which we identify with high latitudes, can be a significant contributor to the midlatitude eddy field.

Baroclinic wave instability has further consequences. For one, it would produce eddies greater than deformation scale. The model simulations suggest that the eddies are about 2 times the deformation scale if $Z = 1$, as predicted by theory, but grow to still larger scales because of an inverse cascade of energy. The final scale...
was determined by beta in those experiments because the cascade arrested to barotropic Rossby waves, and the larger Z is, the larger the equilibrated eddy scale will be. Satellite results (Stammer 1997), however, do not support the existence of a beta arrest in the extratropical ocean, but suggest the dominant eddy scale at higher latitudes is about 2 times the deformation radius. So it may be that instability is occurring, but not the cascade.

The other consequence of baroclinic wave instability is that the observed phase speed at higher latitudes should be at least 2 times the baroclinic long wave speed (reflecting larger, barotropic waves). It would be greater still if an inverse cascade occurs. Satellite results (Chelton and Schlax 1996) suggest the observed phase speed is about 2 times as fast outside the Tropics. That would be consistent with our results, providing the inverse cascade is somehow inhibited.

It would be of particular interest to extend the present model beyond quasigeostrophy to include the equatorial domain to examine whether Rossby/Kelvin basin modes of the sort discussed here exist and are stable. That problem is currently under study.

Acknowledgments. This research was partially supported (JP) by a grant from the Ocean Sciences section of the National Science Foundation and by Norsk Forskningsradet (JL).

APPENDIX

The Matrices $Q$ and $R$

Projecting (4.3a) and (4.3b) leads to the two matrix equations in (4.5a) and (4.5b). Writing

$$L_{11}(m) = (m\pi e)^2 + l^2,$$

(A.1a)

$$L_{21}(m) = (m\pi e)^2 + l^2 + 1,$$

(A.1b)

$$L_{12}(m) = h_1 h_2 l V_e [(m\pi e)^2 + l^2 - \varepsilon^2 k^2],$$

(A.1c)

and

$$L_{22}(m) = i l V_e [(m\pi e)^2 + l^2 - 1 + \varepsilon^2 k^2]$$

(A.1d)

and

$$
\begin{align*}
P(n, m) &= 0.5 \cos \vartheta, \quad m \pm n = \pm 2j, \tag{A.2a} \\
P(n, m) &= \frac{-2j}{\pi} \sin \vartheta [1 - (-1)^{m+n}]
\times \left[ \frac{1}{(n + m)^2 - 4j^2} + \frac{1}{(n - m)^2 - 4j^2} \right], \tag{A.2b}
\end{align*}
$$

then

$$Q(n, m) = \frac{L_{12}(m)P(n, m)}{L_{11}(n)} \quad \text{and} \quad (A.3a)$$

$$R(n, m) = \frac{L_{22}(m)P(n, m)}{L_{22}(n)}. \quad (A.3b)$$

This leads to the single matrix eigenvalue problem:

$$\lambda^2 [A_n] = [QR] [A_n], \quad (A.4)$$

whose solution leads to the growth rates shown in Fig. 6.

REFERENCES


