Baroclinic instability in a long wave environment. Part I: Review

By RICHARD GROTJAHN

Department of Land, Air and Water Resources, University of California

(Received 21 April 1983; revised 23 February 1984. Communicated by Professor B. C. Weare)

SUMMARY

A brief survey of past theoretical studies of the instability of curving flows is presented. Most of the studies consider linear instability. Finite amplitude, nonlinear studies are not considered here. Most of the early papers were concerned with finding instabilities; later papers include more discussion of the structural and other properties of the unstable modes. It is found that nearly all these studies lack a proper discussion of the energetics of the unstable modes. Proper energy conversions are derived and examined in part II.

1. A REVIEW OF LINEAR THEORY FOR CURVING BASIC FLOWS

Over the years there have been many studies of the linear stability properties of geophysical flows. Yet there are comparatively few articles describing the linear stability of synoptic-scale flows which vary in both horizontal directions. The principal reason is that this class of instability studies usually requires solving non-separable differential equations. Thus obtaining a solution awaited the development of sophisticated analytical or numerical techniques. Since the number of studies is still small, a comprehensive review is a manageable task. While many of the later articles summarize some previous work, the author is unaware of a summary as comprehensive as the one presented here. A comprehensive survey is useful to provide a context for the current research and perhaps some direction to future research on this subject. We will not discuss nonlinear ‘finite amplitude’ studies in this review. We will refer to energy conversions as being ‘barotropic’ and ‘baroclinic’. The former refers to conversion between basic flow and perturbation kinetic energy; ‘baroclinic’ refers to potential energy conversion between the basic flow and the perturbation. Also a ‘baroclinic’ basic wave has an amplitude change with height whereas a ‘barotropic’ basic wave does not.

The original paper on the subject appears to be the numerical study by Egger (1969). He used a quasi-geostrophic two-layer model with a fixed long wave basic state. That article has been largely overlooked, perhaps because it is published in German.

The de facto classic paper on the subject is Lorenz (1972). This is an essentially analytic treatment of the linear stability of a basic state composed of a zonal mean flow and a barotropic, long, Rossby wave on a mid-latitude $\beta$ plane (where the Coriolis parameter varies linearly). He assumed a Fourier series type solution and obtained a series of linear equations that are coupled through the presence of wavenumber one in the basic state. He discussed his solution using results truncated after the first few pairs of wavenumbers. He found that there were some exponentially amplifying solutions when the amplitude and wavelength of the long wave were sufficiently large. His discussion focuses mainly upon the wave interaction building a stronger zonal ‘jet’ and upon atmospheric predictability. This paper inspired several later studies. Hoskins and Hollingsworth (1973), Gill (1974), Coaker (1977), Meid (1978) and Jones (1978, 1979) were all in some way interested in investigating the nature of the instability found in Lorenz’s model. Most of these studies (including Lorenz) demonstrated the validity of Fjortoft’s (1953) theorem regarding the flow of energy between interacting waves. That theorem could be deduced from a necessary condition for unstable modes. Jones (1978) claimed to generalize the theorem for a stably stratified fluid. Hoskins and Hollingsworth re-examined the unstable modes and concluded that they were tied to a balance between
the dispersive effects of $\beta$ and wave interaction. Gill generalizes Lorenz's work to examine the instability as a function of a parameter $M$, which expresses the relative importance of the wave interaction and variable Coriolis effects. He found that there were two different types of instabilities. When the $\beta$ effect was strong and the long wave amplitude was comparatively weak ($M \ll 1$), the unstable modes looked part of a resonant triad with the basic wave and mean flow. When the long wave amplitude was strong and the $\beta$ effect weak ($M \gg 1$), the instability looked like a shear (Rayleigh type) instability. Coaker objected to the smallness of the truncated spectrum used by previous authors. Instead, he solved the problem by transforming the time variable and obtaining a Floquet system. From this he was able to calculate neutral stability curves using parameters similar to those of Gill. He also pointed out that the resonant interaction type instability found by Gill is more precisely called 'parametric' instability. (Parametric instability is discussed by Kelly (1965), for example.) True resonant interaction only occurs in the limit of vanishing long wave amplitude. This analysis was further reinforced by Meid, who emphasized that the two types of instability are distinct and both types can coexist for $M \sim 1$. Jones (1978, 1979) appeared to be mainly concerned with the wave interaction ('parametric') instability in his study. he includes an analysis for the two-layer model generalized from Phillips (1954) to include a long wave in the basic state. Finally, Pedlosky (1975), using his two-layer model and apparently restricting his attention to the 'wave interaction' modes, showed that even though his (marginally stable) long wave had infinite meridional extent, the amplifying short waves had zonal and meridional scales that were of the order of the deformation radius.

Along with these barotropic studies, there were studies considering the stability of a 'baroclinic' long wave, by using a two-layer model in which the upper wave was 180° out of phase with the lower wave. An early paper is by Kim (1978) who uses a generalization of Phillips's $\beta$ plane model. Yamagata (1976) uses a similar model, but restricts his attention to cases where the $\beta$ effect is small. He examines how the stability properties change (in his three-wave system) as a function of the rotational Froude number, $F$. For $F > \frac{1}{2}$ he finds a purely barotropic and a mixed baroclinic–barotropic unstable mode. For $F < \frac{1}{2}$ he claims to find two other types of modes, one purely barotropic (but the wave has a phase shift between the layers) and another whose barotropic amplification is reduced by baroclinic energy conversion. Kim noted that the character of the instability (barotropic versus baroclinic) depended upon whether the long wave length scale was greater or shorter than the deformation radius. His discussion is hampered by a curious choice of parameters for the axes of his diagrams. The method of solution is similar to Lorenz's in that a Fourier expansion is used. Merkine and Israeli (1978) used a similar model, but allowed enough terms in the Fourier series to assure convergence. Their main interest was in showing that a lee trough could locally destabilize an otherwise stable zonal mean flow.

Lin (1980a) attempted to synthesize the earlier work on a $\beta$ plane. Using a two-layer model similar to Kim (1978), he varied the relative amplitudes of the baroclinic wave, barotropic wave and zonal mean flow in his basic state. At various limits, the Phillips (1954), Lorenz (1972) and Kim (1978) problems were reproduced. He also varied the length scales of basic state wave. He derived approximate relations for the length scales of the most unstable wave by taking the appropriate derivative of the growth rates in his dispersion relation. He showed volume-averaged quasi-geostrophic energy conversions for several different combinations of his parameters.

In a companion paper, Lin (1980b) used atmospheric values to define a 'Green' (1960) type of long wave basic state. From this he compared the observed spectrum of eddy heat fluxes with those produced by the most rapidly growing mode. Both spectra
have a maximum at wavenumber three, but some care is needed on two points: first, it is clear from his Fig. 6 that the zonal and meridional structure is principally that of wavenumber 6, which is close to the deformation radius. Second, the time average heat fluxes over a nonlinear eddy life-cycle can be quite different from linear results such as these (e.g. Simmons and Hoskins 1978).

Duffy (1975) re-examined the Lorenz (1972) problem but used the shallow water equations, so that an additional pair of inertial–gravity modes are included. These additional eigenmodes turned out to be non-amplifying (he uses the ambiguous term ‘stable’). Later Duffy (1978) used a basic state composed solely of a long wave whose structure was similar to the most unstable Eady (1949) mode but then formulated the stability calculation as a two-layer model after Phillips (1954). As an additional wrinkle, he used the semi-geostrophic coordinate transformation (Hoskins 1975) which includes some ageostrophic effects after one transforms back to physical space. However, he presents no quantitative results attributable to ageostrophy. He does claim that perturbations are only baroclinically growing if their meridional wavelengths are smaller than the zonal wavelength of the basic wave (which has infinite meridional wavelength); if they are larger, then they grow by both baroclinic and barotropic means. Unfortunately, his quantitative evidence is ambiguous and limited to a figure that is missing its line labels. Yet Lin (1980a) confirms a similar relationship (in very general terms) between perturbation meridional wavenumber and the principal instability mechanisms. Sasamori and Youngblut (1981) sought to complement their diagnostic work regarding observed, forced, atmospheric long waves by making barotropic and two-layer baroclinic analyses with the assumption that the long waves are stationary due to forcing. This removes the $\beta$ dependence of the long wave phase speed that Hoskins and Hollingsworth (1973) suggested might be important for the location of the neutral stability curve. All these studies assume plane wave solutions which have an infinite wavelength in one direction, so remarks about meridional and zonal scales actually refer to the orientation of a plane wave. From this point of view it seems more useful to look at waves that have similar wavelengths along and perpendicular to their direction of motion. This can be done on a $\beta$ plane (as is done in part II), or more appropriately on a sphere.

The earliest study using spherical geometry appears to be Hoskins (1973). He looked at the stability of a barotropic Rossby–Haurwitz wave to a superimposed zonal mean flow and perturbation. He describes numerical, initial value integrations using wavenumber multiples of 4 for the basic wave and perturbation. A basic state wavenumber 4 stayed about the same while wavenumber 8 created a low latitude jet and lost a lot of energy. He also included a truncated linear analysis to accompany the numerical integrations. From both he concluded that long waves of zonal wavenumber 5 or less were stable on a sphere. Baines (1976) objected to Hoskins’ numerical procedure and was able to show that all barotropic waves with meridional ‘wavenumber’ (for the Legendre polynomial) greater than or equal to three could be unstable if their amplitude was great enough. Baines uses Fjortoft’s criteria to demonstrate the stability of meridional ‘wavenumbers’ one and two. The energy flow was greater towards the higher than towards the lower wavenumbers. The critical amplitude for instability was linked with the classical Rayleigh criterion of vanishing potential vorticity gradient. Along with a linear eigenfunction analysis, Baines includes some nonlinear integrations. Another barotropic study is by Simmons et al. (1983). They sought unstable solutions to the barotropic vorticity equation linearized about the northern hemisphere climatological January 300 mb pattern. They found that a few modes could grow rapidly; they used the projection of these modes upon specified forcings to note the similarities in appearance between these modes and the so-called ‘teleconnection’ patterns.
Frederiksen (1978) followed up the linear study of Baines with some analyses of a two-layer quasi-geostrophic model. This eigenvalue problem was formulated by assuming that all the spherical harmonics comprising a given eigenmode have the same phase frequency. This yields solutions that consist of an amplitude envelope which modulates the travelling wave solutions. This envelope is the only means by which the modes (1) can have spatially varying growth rates, and (2) can propagate in a direction other than zonally. For example, when the gradient of the envelope has its largest component parallel to the direction of the eddy motion then the modes have their largest growth rates (note that Frederiksen somewhat misleadingly states that the most active development is where the envelope amplitude is largest). Frederiksen found that a purely barotropic long wave reduced, and a purely baroclinic wave increased, the growth rates when superimposed upon a zonal mean flow. In this and in subsequent papers he attempts to relate the regions of maximum eigenmode growth with a measure of the latent instability of the basic flow. At first this stability criterion is merely the difference in wind speed between the two layers weighted by the cosine of latitude. He finds the largest envelope amplitudes of streamfunction, heat and momentum flux occurring just downstream from the regions of largest stability criterion. This is consistent with our interpretation above, since the positive downstream gradient of the amplitude envelope in those regions will locally increase the growth rate there. He does not show a diagram of the horizontal variation of the stability criterion, but judging by his basic state streamfunction map, the agreement seems very good. Frederiksen first shows some features of the envelope in his second (1979a) paper using the same model but emphasizing a different basic state long wave. He refers to the modulation by the envelope in his early papers as the ‘non-oscillatory’ part of the quantity and in his later papers as the “random phase ensemble averages” or ‘RPEA’. Frederiksen (1980) uses the two-layer model again, and suggests the natural generalization of the stability criterion to include horizontal shear by using the basic state potential vorticity. Frederiksen (1979b) extends his model to five levels and finds that the lower level eigenmode streamfunction amplitude and momentum fluxes are larger than in the two-layer model. He also points out that in the five-level model the upper-level amplitude is larger when a long wave is present. More recently, Frederiksen (1982a, 1982b, 1983) has used basic states derived from northern hemisphere climatology. The first of these uses the two-layer model, winter climatology and considers changes in the volume average static stability. When the static stability was larger the fastest growing mode moved more slowly and had a meridionally oriented dipole (high–low pair) structure; Frederiksen claimed that this result may be related to observed atmospheric blocking. The second uses a basic state derived from a numerical simulation of stratospheric sudden warming. The third paper uses the five-level model with January and July mean patterns; he compares the heat and momentum fluxes with observed patterns. Frederiksen does not calculate any local energy conversions, though he does calculate meridional (but not zonal) heat and momentum fluxes. Apparently Frederiksen was more interested in reproducing observed vertical distributions of meridional heat and momentum fluxes than investigating the energetics of his solutions.

Finally, there have been a few other studies that used a constant Coriolis parameter. Niehaus (1980, 1981) used an \( f \) plane, quasi-geostrophic system. She solves this system as an eigenvalue problem in the former paper and using the method of multiple scales in the latter (where widely different basic state and perturbation length scales are assumed). She also finds the largest amplitude downstream from where her basic state vertical shear is greatest. Drazin (1978) also used the Eady (1949) problem, but formulated it in approximate cylindrical coordinates and calculated neutral stability curves. Grotjahn
(1983) pointed out that the cross-flow horizontal shear neglected by Drazin is an important part of the curvature problem. He examined the stability properties of a circular jet using an initial value, high resolution, spectral model. The cylindrical symmetry of the problem afforded easy analysis of the energetics of the unstable modes. Curvature was found to reduce the growth rate through barotropic damping.

One thing that becomes clear when reviewing these articles is that few of the authors calculate proper energy conversions. Several authors calculate eddy heat fluxes and eddy momentum fluxes (sometimes only the northeard component of each!). In order to understand properly the dynamics of the eddies it is necessary to look at the actual energy conversions, for several reasons. First, when using height as a vertical coordinate, one must include a density weighting. Second, and more importantly, one must include a weighting by the basic flow or by the basic flow shear (depending upon how the chain rule is used in the definition of the energy conversion). These weightings can make the energy conversion field look quite different from the eddy heat or momentum flux field. Third, for a curving flow, the energy conversions will depend upon the eddy heat or momentum flux in the direction locally orthogonal to the basic flow direction. As mentioned above, Lin (1980a) calculated energy conversions but was interested in the different vertical variations caused by different horizontal variation of the basic state, so he took horizontal averages of these conversions.

In part II of this paper we seek to fill part of this gap by presenting and emphasizing the properly defined energy conversions for eddies growing upon a long wave basic state. These energy conversions are derived from the governing equations. In the process, three apparently new 'ageostrophic' energy conversions are identified which redistribute the eddy properties in the three coordinate directions. Two of the ageostrophic energy conversions are found to cause cross-isobaric motion of eddies, allowing the eddies to track along the axis of maximum vertical shear in the two cases examined. The third ageostrophic conversion redistributes the energy in the vertical. To the extent that comparisons can be made with previous work, the results are consistent.

ACKNOWLEDGMENT

Support for this article was provided by the National Science Foundation under grant ATM 8305091.

REFERENCES


Eady, E. T. 1949 Long waves and cyclone waves. Tellus, 1, 33-52


Fjørtoft, R. 1953 On the changes in the spectral distribution of kinetic energy for two-dimensional, nondivergent flow. Tellus, 5, 225-237
Frederiksen, J. S.


1979b Baroclinic instability of zonal flows and planetary waves in multi-level models on a sphere. *ibid.,* 36, 2320–2335

1980 Zonal and meridional variations of eddy fluxes induced by long planetary waves. *Quart. J. R. Met. Soc.,* 106, 63–84


1982b Instability of the three-dimensional distorted stratospheric polar vortex at the onset of the sudden warming. *ibid.,* 39, 2313–2329


Gill, A. E.


Green, J. S. A.


Grotjahn, R.


Hoskins, B. J.


1973 On the simplest example of the barotropic instability of Rossby wave motion. *ibid.,* 30, 150–153


Hoskins, B. J. and Hollingsworth, A.

1979 Rossby wave interactions and instabilities in a rotating, two-layer fluid on a beta-plane. Part II: Stability. *ibid.,* 12, 1–33

Jones, S.


1980b Eddy heat fluxes and stability of planetary waves, Part II. *ibid.,* 37, 2373–2380

Lorenz, E. N.

1972 Barotropic instability of Rossby wave motion. *ibid.,* 29, 258–264


1980 Instability of non-zonal baroclinic flows. *ibid.,* 37, 1447–1463


Meid, R. P.


Merkine, L. and Israeli, M.


Niehaus, M. C. W.

1983 Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *ibid.,* 40, 1363–1392


Pedlosky, J.


Phillips, N. A.


Sasamori, T. and Youngblut, C. E.


Simmons, A. J., Wallace J. M. and Branstator, G. W.


Simmons, A. J. and Hoskins, B. J.
