Regime dependent instability as a transition mechanism in large-scale atmospheric flow

Isabella Bordi
Dipartimento di Matematica e Fisica, Università di Camerino, Italy

Abstract
Two apparently distinct approaches to studying the observed low frequency variability in the atmosphere have evolved over the past few years. One approach invokes multiple, recurrent flow regimes to explain the observed variability. An alternative approach involves the linear stability properties of the climatological mean flow. In the present study, these approaches are merged to understand the different stability properties of the basic states associated with different flow regimes. In particular, the different zonally asymmetric components of the regime basic states lead to differing stability properties that in turn may explain the transition mechanism between regimes. This is particularly true for the transition between an amplified atmospheric planetary wave flow regime and a zonal regime. The observed transition streamfunction anomaly pattern compares very well to the most unstable stationary eigenmode of a linear stability calculation for both the barotropic and two-level baroclinic case studied. However, the growth rate of the barotropic case is too slow compared to observations and it is quite sensitive to the dissipation rate and the resolution of the calculation. In the baroclinic case, the same eigenmode appears but with a faster growth rate and more structural stability. Within the constraints of a two-layer model, the effect of baroclinicity is to remove the dependence on dissipation rate of the growth rate of the most unstable barotropic mode, allowing fast growth without sensitivity to chosen parameters. The most unstable stationary baroclinic eigenmode strongly resembles the anomaly pattern of the observed transition. The energetics of the growing mode involves extraction of energy from the zonally asymmetric flow in agreement with observations. Experiments with greatly increased dissipation reveal very little sensitivity of the growth rate of this stationary eigenmode to the rate of dissipation. Alternatively, the eigenmodes for the opposite transition considered, from the zonal to amplified wave regime, are different from this former case in terms of structure, growth rate and energetics. Therefore, we conclude that the linear stability properties of the atmospheric flow are a function of the amplitude of the zonal asymmetries in the antecedent regime basic state, and that the dynamics of the transitions between regimes might be understood within the context of the linear instability properties of the regime basic states.

Key words *regimes – instability – baroclinic*

Mailing address: Dr. Isabella Bordi, Dipartimento di Fisica, Edificio Fermi, Università di Roma «La Sapienza», Piazzale A. Moro 2, 00185 Roma, Italy; e-mail: bordi@roma2.infn.it

1. Introduction

The concept of weather regimes dates back to the 1930's when regimes were classified on the basis of the patterns observed in surface synoptic charts (*e.g.*, Baur, 1951 and references therein). In his penetrating appraisal of the dominant problems of modern meteorology, Rossby (1959) emphasized the weather regimes question and gave a description of why the atmos-

pheric planetary scale circulation alternately cycles between prevalently zonal and predominantly meridional behavior. Weather regimes problem was reopened by Charney and Devore (1979) who proposed a new point of view that is based on the intrinsic nonlinear nature of the atmospheric motion. In fact, in a nonlinear system the space of its physical state may be such that a trajectory in this space may wander around a few preferred regions of phase space. For the atmospheric case, Charney and Devore went so far as to identify such regions with the attraction domain of the stable steady states of a highly truncated barotropic, non-divergent model. Since then, work has been devoted to both identifying and modeling the atmospheric circulation as a discrete set of attractive regions in the full atmospheric phase space (*e.g.*, Legras and Ghil, 1985; Hansen and Sutera, 1986; Benzi *et al.*, 1986a,b; Molteni *et al.*, 1988) including recent work by Hansen and Sutera (1995a,b,c), Malguzzi *et al.* (1996), and Molteni (1996a,b). Many of these studies concentrate on the search for multiple steady states of atmospheric models of increasing complexity. A somewhat less studied question has been that of determining the physical processes that are responsible for the transition among the quasi-steady states. There are few notable exceptions. Molteni and Palmer (1993) and Buizza and Molteni (1996) suggest that the transitions are the result of nonmodal instabilities. In most of these papers, including Charney and Devore's pioneering work, an implicit common assumption is that there exists a process that is the source for the transitions between different weather regimes. From the point of view of atmospheric low frequency variability (at least for time scales related to atmospheric intraseasonal variability), it is only a knowledge of such mechanisms that might provide for an explanation of a relevant portion of this variability. Then, the observed circulation may be interpreted, as Rossby foresaw, as the end result of the switching among these steady or quasi-steady states.

An alternative line of research (developed in parallel with the work discussed above) has studied the observed low frequency variability by means of the linear instability analysis of the observed, zonally varying climatological mean flow (Simmons *et al.*, 1983; SWB hereafter and originally discussed by Lorenz, 1972). Thereafter, this subject has flourished and has been considered by many authors (see Fredricksen and Webster, 1988 and references therein). It has also been used in improving short term predictability of rapidly evolving synoptic disturbances by exploiting the interesting properties associated with the non normality (Kato, 1966) of the evolution of the linearized meteorological operators (Borges and Hartmann, 1992). A key element of these studies is that the instability arises through the zonally asymmetric component of the climatological mean flow. Numerous studies have attempted to identify, with varying degrees of success (Branstator, 1985, 1990, and references therein), the unstable normal modes computed from linear stability analysis with the observed patterns of low frequency variability. In many cases, an appeal is made to non-modal or finite-time instability to explain the observed growth of anomalies (Molteni and Palmer, 1993; Chang and Mak, 1995). In addition, Palmer (1988) has investigated the barotropic stability properties of opposite signs of teleconnection patterns.

In this paper we will attempt to exploit the former theories (instability of a zonally asymmetric basic field) to suggest an explanation of the transition mechanisms between regimes.

However, there is one aspect of the traditional eigenanalysis of zonally asymmetric basic flows that is so far debated. It deals with how the basic state is specified and maintained. Regarding this concern, Andrews (1984) has argued that the nature of the forcing is crucial to the resulting instability problem. In fact, the perturbation forcing that arises can act to stabilize perturbations to a mean flow that would otherwise be considered unstable without the forcing.

As commonly done in instability calculations, we take the basic state to be an observed, time averaged flow corresponding to a particular weather regime, assuming the background flow to be maintained by external, *ad hoc* forcing that does not interact with the growing disturbances.

We will show evidence that physically distinct instability processes might be the source of

the transition between regimes. This distinction arises from the nature of the zonally asymmetric component of the flow in different regimes (in fact, the symmetric components are the same in the two regimes, see Hansen and Sutera, 1987). In particular, the amplitude of the flow asymmetries will be shown to be a determining factor of the nature of the instability. This approach is not new, in fact, it was foreseen by Charney (see page 81 of Lindzen *et al.*,1990). Here we will show that the approach is rewarding and allows a favorable comparison with observations.

To pursue this goal, we need to consider a suitable definition of weather regimes. Among the various approaches to identifying weather regimes, well documented and statistically robust observational evidence for the existence of multiple weather regimes is that proposed by Sutera (1986) and Hansen and Sutera (1986). Some concern has been raised by Nitsche *et al.* (1994) on the technique used by Hansen and Sutera (1986) to establish the statistical confidence in the existence of multimodality of probability density distributions, but much of their concerns have been dissipated by additional results presented in Hansen and Sutera (1995b) (which include, among other things, comparisons to different regime definitions; *e.g.*, Molteni *et al.*, 1990; Cheng and Wallace, 1993; Kimoto and Ghil, 1993).

Therefore, we will use in our stability calculations the two fields which may be derived by the definition of regimes corresponding to the two modes of the planetary wave amplitude indicator defined in Hansen and Sutera (1986, 1995b). One of the regimes corresponds to a zonal pattern on the hemispheric scale (hereafter referred to as the Zonal Regime or ZR), while the other corresponds to a wavy, planetary wave pattern featuring a major ridge over the Eastern Pacific and downstream trough over North America (hereafter referred to as the Amplified Wave Regime or AWR). As shown in Hansen and Sutera (1993), the latter regime includes nearly all Pacific blocking cases as a subset.

In Section 2 we will linearize barotropic, non-divergent atmospheric equations around these two basic states and calculate the relevant linear instabilities for these two circulation regimes. We will show evidence that the resulting instability patterns and the associated e-folding times, though strongly dependent on dissipation, are consistent with observations that describe such transitions. In Section 3 we show how the instability found may be explained in terms of a single triadic interaction. We will also show that the nonlinear closure of the triad model leads to asymptotic states that partially cancel the non symmetric component of the unstable basic field. In Section 4 we introduce a two level model to represent a baroclinic atmosphere. It is shown that, in this environment, some of the concerns (which apply also to our barotropic calculations) about instability calculations in the presence of dissipation raised by Borges and Sardeshmukh (1995) can be greatly mitigated. Conclusions and plans for future work are in Section 6.

2. Observational background

In Hansen (1988; H88 hereafter) the structure of the mean transition composite between the regimes detected in the probability density function of large scale amplitude indicator proposed by Sutera (1986) is considered. This indicator is defined as follows. Let $Z(\lambda, \varphi, t)$ be the 500 mb geopotential height at a time t, where λ and φ are longitude and latitude respectively. Then consider two latitudes φ_1 and φ_2. Let

$$\overline{Z}(\lambda, t) = \int_{\varphi_1}^{\varphi_2} Z(\lambda, \varphi, t) d\varphi \qquad (2.1)$$

and

$$\overline{Z}(\lambda, t) = \sum_k A_k(t) e^{ik\lambda}. \qquad (2.2)$$

Thus the indicator is defined as

$$I = \sum_{k=2}^{4} A_k^2(t). \qquad (2.3)$$

The data used in H88 consisted of four winters (1980-1984) of Northern Hemisphere fields, which were the same used in Sutera (1986) (statistical analysis of the indicator was later

extended to 42 winters in Hansen and Sutera, 1995b). The days that in H88 are considered as transition days, based on the composite time evolution of the wave amplitude, are those 2 days prior to the day on which the amplitude indicator mentioned above crossed the minimum of the distribution and the following 2 days. If the onset of a given case occurs within the decay of another case, or viceversa, observations occurring in the overlapping period of the two transitions are not counted.

Based on this definition, the main findings in H88 may be summarized as follows. The transition between the ZR (also referred to as «Mode 1» in previous papers) and the AWR (also referred to as «Mode 2» in previous papers) and back occurs rather quickly; in fact, in the composite of the 12 cases here considered, it takes roughly 4 days for the transition to occur. Energetics calculations indicate that the ZR to AWR transition is consistent with a strong baroclinic instability of waves shorter than planetary scale with subsequent nonlinear interactions allowing energy to cascade towards longer scales and accompanying the growth of the anomalies represented by the AWR state. The most striking difference between this and the reverse transition is that in the latter there is a predominance of barotropic conversion to even longer scale, namely zonal wave number 1. On the ground of the short time scale for transitions compared with the persistence in the regimes, we may suppose that the transitions are due to instability processes. However, the data are consistent with the instability mechanism, operating during AWR to ZR transition, which is of a different nature from that operating during the reverse one. This observational suggestion constitutes a primary motivation for the present work.

In order to be as close as possible to the H88 analysis, in this study we will consider data from ECMWF archives at T21 resolution for a period spanning from 1982-1988. Future work will extend these results to a more exhaustive climatology.

In the present paper, the stream function, at 300 mb and at 850 mb, is used instead of the geopotential height used in the previous reports. From the winter data (DJF) of 1982-1988, we construct the mean Northern Hemisphere 300 mb and 850 mb streamfunction fields for the climatology (not shown), for the zonal regime (fig. 1a,b), and for the amplified wave regime (fig. 1c,d). The main features of these composites correspond, as expected, to the appropriate flow features in the geopotential fields reported elsewhere.

In addition, we have calculated the composite average streamfunction fields for both the AWR to ZR transitions and *vice versa*. According to Sutera (1986), these fields were obtained as follows. First consider the deviation of individual daily streamfunction field from AWR (ZR). Define the transition period as the number of days during which the wave amplitude indicator, starting from values typical of AWR (ZR), is decreasing (increasing) irreversibly towards values typical of those achieved in ZR (AWR). Then the composite transition field from AWR (ZR) to ZR (AWR) is obtained as by time averaging the deviation field in that period. The number of averaging days in each transition is about 4 days though this may change in the range of 3 to 5 days for some individual cases. For our present data 12 transitions occur of each type. The composite transition streamfunction anomalies at 300 mb so obtained are shown in fig. 2a,b while the respective non-symmetric components are shown in fig. 2c,d.

In general, as mentioned in H88 and Hansen and Sutera (1993) and Hansen and Sutera (1995b), these fields were fairly consistent in a broad sense from case to case. If the transitions are due to some sort of instability, then the anomaly patterns which we will have to describe are the ones represented in figs. 2a-d.

Of course, it must be stressed that by comparing our results to the above transition maps we are not seeking the most appropriate perturbation field an individual transition, but for the one that mostly likely is present in all transitions.

3. The barotropic stability of mean states

The question we will address concerns whether the observed transition anomaly patterns can be understood in terms of linear stability theory.

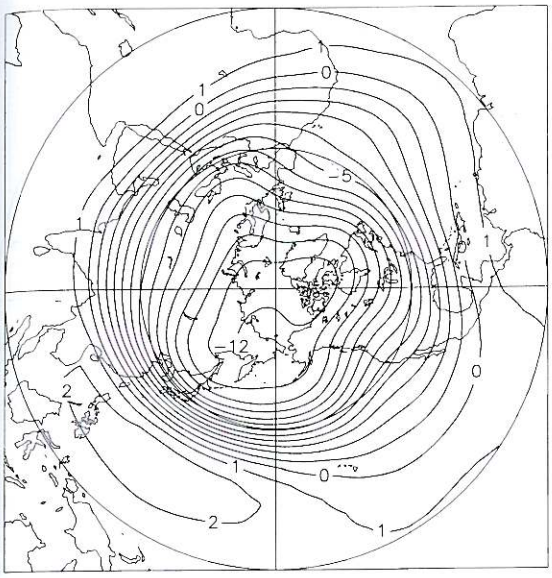

Fig. 1a. Mean streamfunction for the Zonal Regime (ZR) at 300 mb. Contour interval is 1×10^7 m^2 s^{-1}.

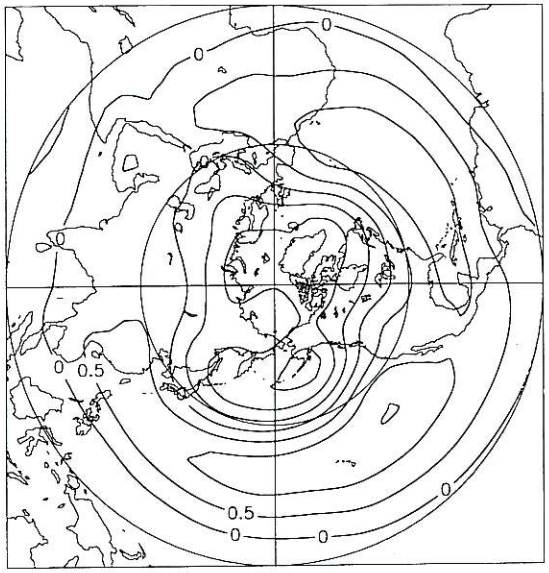

Fig. 1b. Mean streamfunction for ZR at 850 mb; contour interval is 0.5×10^7 m^2 s^{-1}.

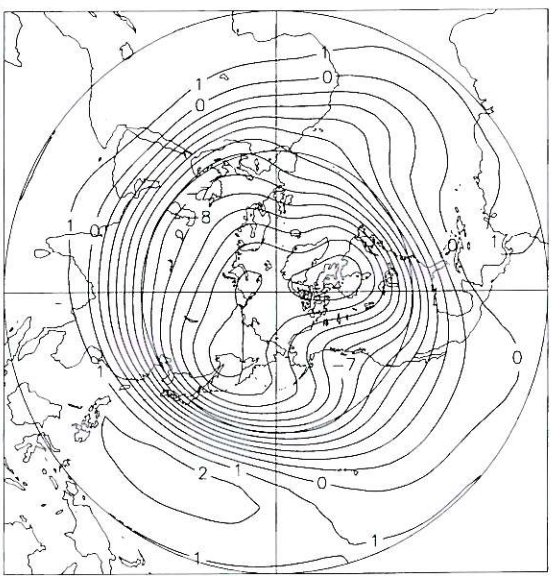

Fig. 1c. Mean streamfunction for the Amplified Wave Regime (AWR) at 300 mb. Contour interval is 1×10^7 m^2 s^{-1}.

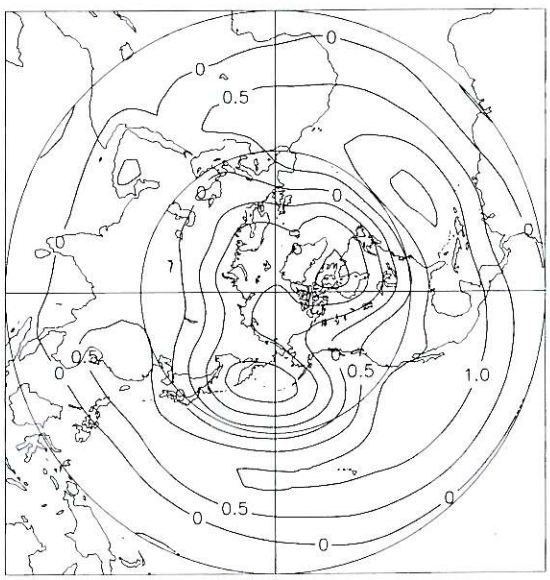

Fig. 1d. Mean streamfunction for AWR at 850 mb; contour interval is 0.5×10^7 m^2 s^{-1}.

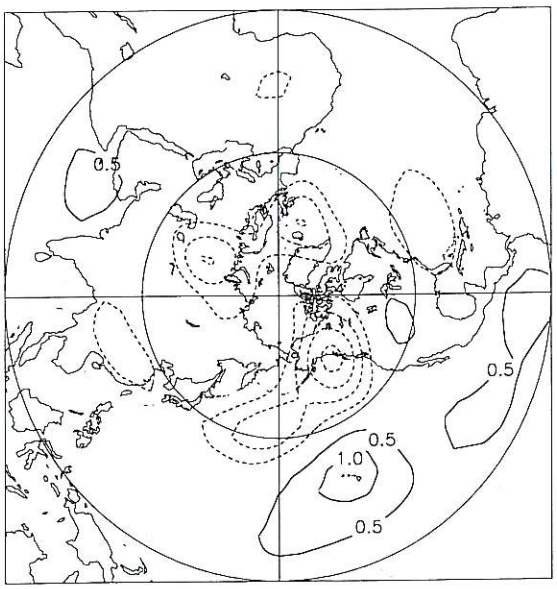

Fig. 2a. Composite streamfunction anomaly at 300 mb for the AWR to ZR transition. Contour interval is 0.5×10^7 $m^2 s^{-1}$ not including zero.

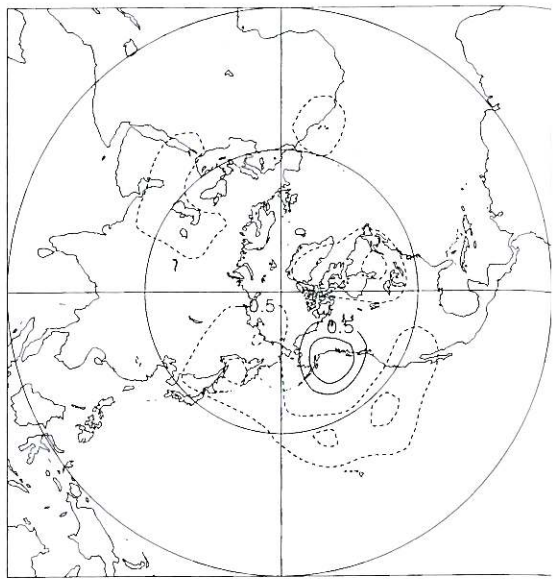

Fig. 2b. Composite streamfunction anomaly at 300 mb for the ZR to AWR transition. Contour interval is 0.5×10^7 $m^2 s^{-1}$.

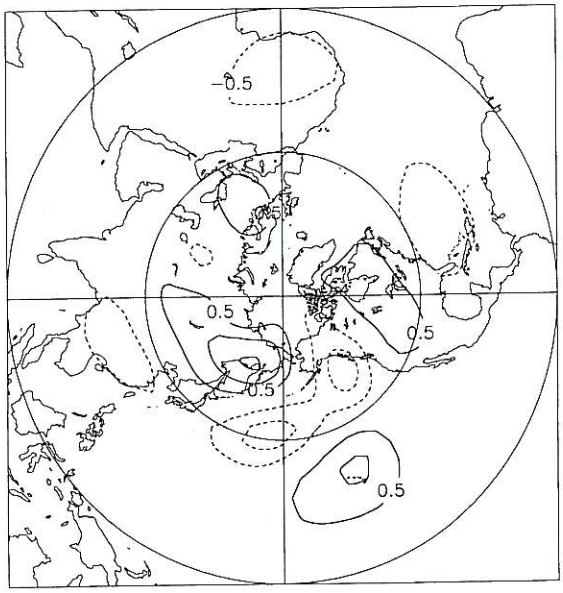

Fig. 2c. As in fig. 2a with the zonal component removed. Contour interval is 0.5×10^7 $m^2 s^{-1}$.

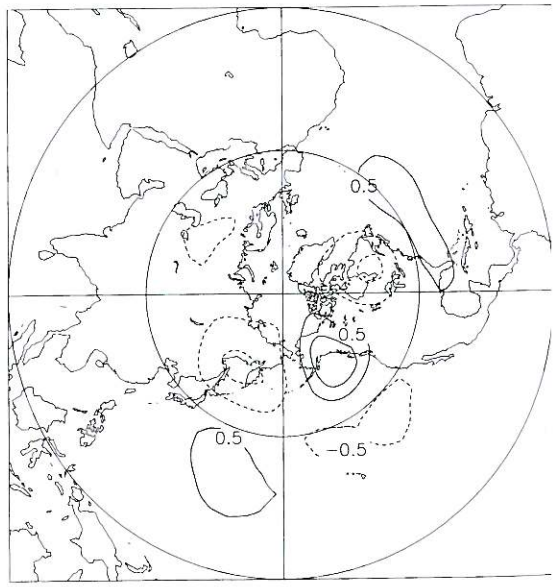

Fig. 2d. As in fig. 2b with the zonal component removed. Contour interval is 0.5×10^7 $m^2 s^{-1}$.

For this and for comparing our result to others (for example, Branstator 1985), we consider the following non-divergent barotropic model in spherical coordinates

$$\partial_t \nabla^2 \Psi + J(\Psi, \nabla^2 \Psi + 2\Omega \sin \varphi) = R \qquad (3.1)$$

where the usual quasi-geostrophic scaling is used

$$\begin{aligned} u &= U\bar{u} \\ x, y &= L\bar{x}, L\bar{y} \\ t &= \frac{L}{U}\bar{t} \end{aligned} \qquad (3.2)$$

Ψ is the 300 mb stream function ∇^2, is the Laplacian operator in spherical coordinates and Ω is the earth angular speed. R is some form of dissipation. As usually done

$$\Psi = \overline{\Psi} + \Phi \qquad (3.3)$$

where $\overline{\Psi}$ is a flow that is supposed to be a steady solution of the equation of motion and Φ is a perturbation on it. Next we expand $\overline{\Psi}$ and Φ in spherical harmonics

$$\overline{\Psi}(\text{or } \Phi) = \sum_m \sum_l \overline{\Psi}_l^m (\text{or } \Phi_l^m) P_l^m e^{im\lambda} \qquad (3.4)$$

where P_l^m are the associated Legendre polynomials and Φ_l^m are functions of time alone.

Then, we linearize (3.1) around the zonally symmetric part of each of the three 300 mb mean fields considered (*i.e.*, climatology, ZR and AWR), truncating at T21 and setting

$$R = \kappa \nabla^6 + \nu \nabla^2 \qquad (3.5)$$

where the super-dissipation $\kappa = 0.2 \times 10^{17}$ m^{4}s^{-1} and ν is the Ekman dissipation rate.

By setting

$$\Phi_l^m = \Phi_l^m (0) e^{\omega t} + (*) \qquad (3.6)$$

we can calculate the fields instability by analyzing the eigenvalue problem

$$\omega \underline{X} = A\underline{X} \qquad (3.7)$$

where $\underline{X}$ is the vector whose components are the projection coefficients of the perturbation field and A is the matrix of the projection of advection operator.

A positive (negative) sign of the real value, *i.e.* the growth rate (decay rate), of the complex number ω will tell us if the basic field is an unstable (stable) solution of the equation of motion. The imaginary part of ω will enclose information on the periodicity of the disturbance field.

For the barotropic model above, calculations may be made by simply setting $\nu = 0$, since the operator

$$\Re = \partial_t - \nu \qquad (3.8)$$

has eigenvalues $\omega - \nu$.

If the Ekman dissipation is set to zero, the three zonally symmetric basic states are weakly unstable with the largest growth rate corresponding to e-folding time of about 145 days for AWR, while ZR and the climatology are stable states. From these results a barotropic instability of the zonally symmetric flow, say of the Kuo type, is an unlikely candidate to explain any regime transition.

Next we consider the barotropic instability of the three zonally asymmetric fields mentioned above. First, consider the climatological mean flow. Using the same super-dissipation as SWB, we obtain a most unstable mode with about 17 days e-folding time and period of about 31 days.

We notice, for what will be following, that there is a mode with 86 days e-folding and infinite period. When we use their basic state, we obtain identical results to theirs.

When a 10 days time scale dissipation and a super-dissipation are added, there are no barotropically unstable modes of the climatological mean flow (see also Borges and Sardeshmukh, 1995). Similarly, the ZR's asymmetric flow, for this Ekman number, is also barotropically stable. Conversely, the AWR basic state exhibits one real eigenvalue corresponding to a non propagating disturbance. As already noted, we might

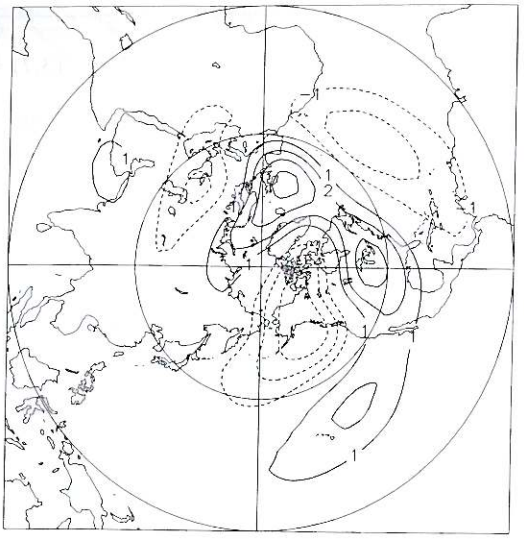

Fig. 3a. The most unstable mode of the AWR basic state in the barotropic model at T21 resolution. This eigenmode is nonpropagating. Units are arbitrary with contours every 1 unit.

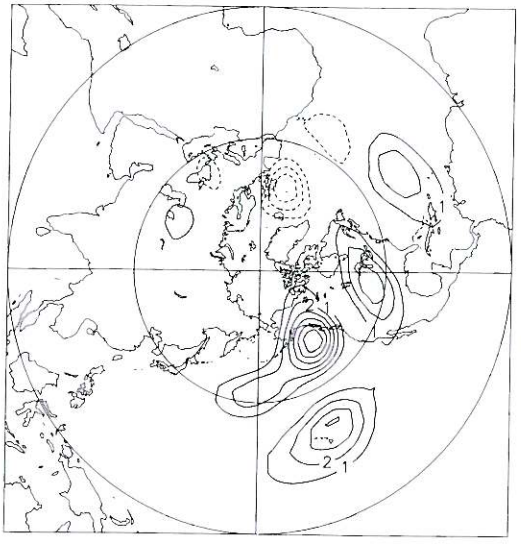

Fig. 3b. Correlation map between the barotropic eigenvector and the observed transition field. Units are arbitrary with contours every 1×10^{-3} unit.

expect that a nonpropagating (or slowly propagating) mode be relevant in the observed transition, since the maps are obtained as a composite of four days average.

In fig. 3a we show the structure of this unstable mode. Despite the e-folding time, 7.66 days, is not of the desired length, the correspondence between the main features of this eigenmode and the observations is quite striking, especially from Eastern Asia across the Pacific and North America, while the European component of the eigenvector has a stronger structure than the observed transition map (fig. 2a,c). Moreover, the structure of this mode features a mainly zonal wave number one field, which is the main eddy component that grows during the AWR to ZR transition (H88).

To make this comparison more quantitative, we also calculated the correlation map, shown in fig. 3b, between the eigenvector and the observed transition field. The strong correlations in the Pacific sector are noticeable, while the weak anticorrelations are confined mainly to the Atlantic. However, the transition map shows weak, small scale features. Differences are also noticeable in the subtropical latitudes. This may be expected since, by using the quasigeostrophic models, we can hardly accommodate for the Hadley circulation (or any other thermally direct circulation) which, instead, is included in the transition map.

It is also important that a comparison between observations and the unstable pattern be similar for the spatial distribution of their energy tendency. In fact, this field describes quantitatively the spatial structure of the relevant physical process occurring during the transition.

For this purpose, let us consider, assuming the atmosphere be a single nondivergent layer, the energy conversion field for the observations at 300 mb.

In this case the relevant equations are in Branstator (1985). For our case the resulting spatial distributions of the energetics are shown in fig. 4a,b for the observations and the eigenvector respectively. In fig. 4c the correlation map between the two fields is illustrated. Similarities (and dissimilarities), already noted between observations and the calculated pattern, are confirmed and they lead to argue that the

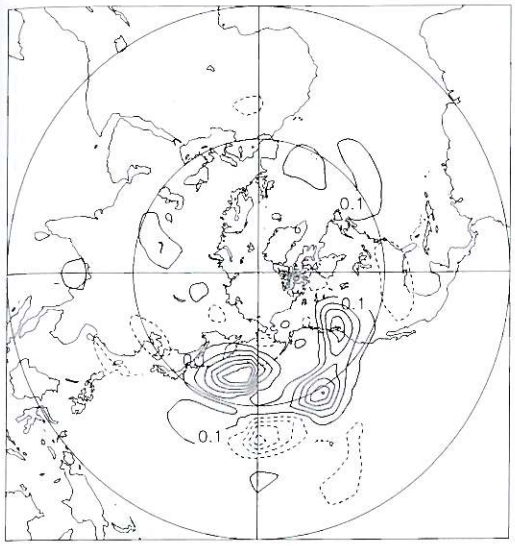

Fig. 4a. Barotropic energy conversion computed from the observations at 300 mb. Units are arbitrary with contours every 0.3 unit.

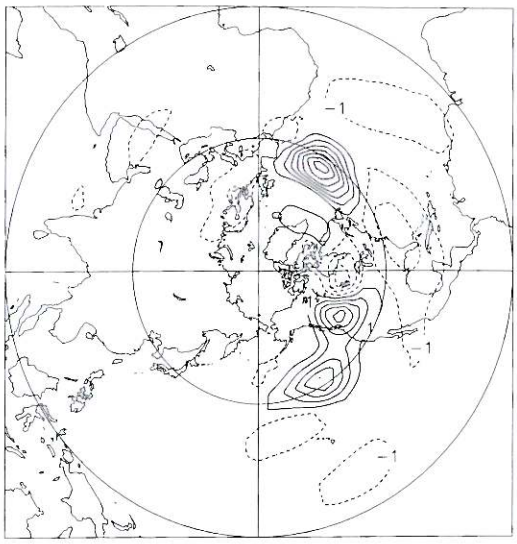

Fig. 4b. Barotropic energy conversion for the most unstable mode of the AWR in the barotropic model. Units are arbitrary with contours every 1.5 units.

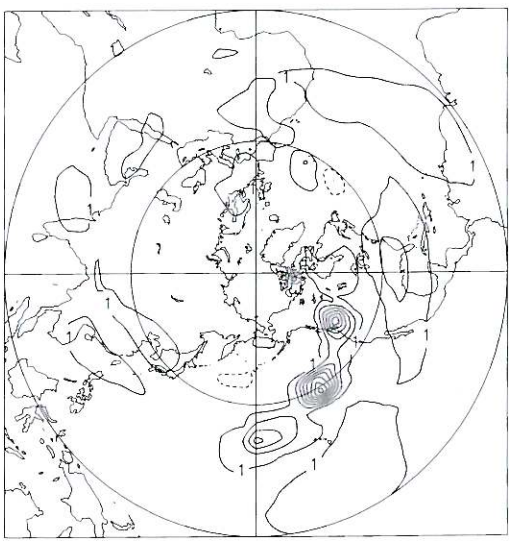

Fig. 4c. Correlation field between observed and computed barotropic energetics. Units are arbitrary with contours every 2×10^{-3} unit.

main feature of the energy conversions has been captured. In fact, the dominant contributions to the barotropic energy conversions come from the Pacific and Atlantic regions. In particular, we can notice the maximum local contribution in the northeast of the strong Pacific jet, where observations and eigenvector energetics are also strongly correlated.

Thus, the overall structure seems close enough to stimulate further steps.

To proceed further, however, we must consider two main problems that affect the instability calculations so far presented.

First, instability calculations determine eigenmodes up to a sign. Thus, if we have an interpretation for the pattern associated to a non travelling eigenvector, it can hardly be adapted also to the corresponding pattern obtained reversing the sign.

The second problem concerns the dependence of the eigenvalue on the actual value of the Ekman number. In the next two sections we will show that both questions may be answered

by considering, in order, a simple nonlinear barotropic model and instability calculations in a baroclinic environment.

4. A simple nonlinear model

Linear instability calculations, *per se*, cannot remove the limitation associated with the intrinsically arbitrary sign of the eigenvector.

However, this limitation can be removed if, by proceeding a step further, we try to understand the finite amplitude, asymptotic behavior of the disturbance.

In the present paper, we will not discuss the dynamics of the T21 model, which we intend to present on another occasion. Instead, we will show, with the following simple calculations, the way that may lead to a solution of the problem caused by the indeterminacy of the eigenvector sign.

Following Baines (1976), we know that, for a basic state which is the sum of a super rotation plus a Rossby stationary wave, the minimal triad, which allows a non-propagating instability (like the one discussed in the previous section), is the one obtained by selecting the perturbation field as follows.

Let

$$\overline{\Psi} = A_{l_s}^0 P_{l_s}^0 + A_l^m P_l^m e^{im\lambda} + (*) \qquad (4.1)$$

with $l_s = 1$ is a steady solution of the vorticity equation, then if

$$\Phi = \phi_{l_1}^0 P_{l_1}^0 + \phi_{l_2}^m P_{l_2}^m e^{im\lambda} + (*) \qquad (4.2)$$

we may study the instability Ψ under Φ for a given l, l_1, l_2 and m. In other words, the perturbation field contains a zonal component and a wave of the same zonal wave number of the basic field. For the barotropic model above, the eigenvalue problem generated by the projection onto the modes previously defined reduce to the diagonalization of a 3 by 3 matrix. In absence of dissipation the eigenvalues are

$$\omega = 0,\ \omega = \pm\sqrt{-\gamma^2 + 2\beta\alpha\left|A_l^m\right|^2} \qquad (4.3)$$

where a is the earth radius and

$$\gamma = \frac{1}{a^2}\frac{(LL_2 - LL_s)}{LL_2} mK_s A_1^0 - 2\Omega,$$

$$\beta = -\frac{1}{a^2}\frac{(LL_1 - LL)}{LL_2} mK_{m_1},$$

$$\alpha = -\frac{1}{a^2}\frac{(LL_2 - LL)}{LL_1} mK_{m_1}, \qquad (4.4)$$

$$LL_{(.)} = -\frac{l_{(.)}(l_{(.)}+1)}{a^2},$$

$$K_s = \int_{-1}^{1} (P_{l_2}^m)^2\, \partial_\mu P_1^0\, d\mu,$$

$$K_{m_1} = \int_{-1}^{1} P_l^m P_{l_2}^m\, \partial_\mu P_{l_1}^0\, d\mu$$

where $\mu = \sin(\varphi)$ and φ is latitude.

Thus instability of the Rossby wave occurs whenever $\left|A_l^m\right|^2 \geq \frac{-\gamma^2}{2\alpha\beta}$. It is easy to show (Baines, loc. cit.) that the same basic field (4.1) is unstable for a travelling disturbance composed of two waves, rather than a zonal component and a wave. Moreover, the latter instability occurs for smaller amplitudes of the asymmetric component of the basic field.

This is analogous to the situation encountered in the previous section. Namely, only AWR, *i.e.* a large amplitude state, seems favoured to be an unstable state with respect to a perturbation with a standing growth.

Regarding the energetics of the growing modes, it is plausible to argue that perturbations grow extracting kinetic energy from the non symmetric component of the basic state. In fact, the super rotation, in the absence of form drag, is stable with respect to barotropic perturbations and any form of instability of Kuo type is therefore excluded.

Moreover, from the eigenvalue structure of the linearized problem, it is possible to anticipate that, when we consider the full nonlinear problem (of course, retaining the same modes in the perturbation as in the linear case), we expect that above the threshold value of A_l^m two steady states will appear. In fact, the equations of motion are

$$\begin{aligned}
\partial_t \phi_{l_1}^0 &= -i\upsilon\phi_{l_1}^0 + \alpha(A_l^{-m}\phi_{l_2}^m - A_l^m\phi_{l_2}^{-m}) \\
\partial_t \phi_{l_2}^m &= (-i\upsilon+\gamma)\phi_{l_2}^m + \beta\phi_{l_1}^0 + \delta\phi_{l_1}^0\phi_{l_2}^m \qquad (4.5) \\
\partial_t \phi_{l_2}^{-m} &= (i\upsilon-\gamma)\phi_{l_2}^{-m} - \beta\phi_{l_1}^0 - \delta\phi_{l_1}^0\phi_{l_2}^{-m}
\end{aligned}$$

where

$$\delta = -\frac{(LL_1 - LL_2)}{LL_2} m \int_{-1}^{1} P_{l_2}^m P_{l_2}^m \partial_\mu P_{l_1}^0 d\mu. \quad (4.6)$$

It can be shown that the two new steady states are stable and, more importantly, in the limit of small dissipation they have close amplitude and phase of the nonsymmetric component of the perturbation field. It implies that, for any initial condition, the solution will converge to a new state where the nonsymmetric component is nearly the same.

As an example we show in table I the solutions for the particular setting of the parameters

$$\begin{aligned}
&A_1^0 = -19.8,\ A_l^m = 10 \times (0.18 - 0.63), \\
&l_s = 1,\ m = 2,\ l = 5,\ l_1 = 3,\ l_2 = 7
\end{aligned} \qquad (4.7)$$

and, to mimic a limiting case, $\nu = 1000$ days^{-1}. The value of A_1^0 implies a reasonable superotation of about 30 m s^{-1} at the equator (notice that the particular choice of l, m for the basic state are suggested by the large difference shown by AWR and ZR for this particular spherical harmonics).

Table I. Solutions for the simple non-linear model. Parameters setting is in the text.

	Solution 1	Solution 2
$\phi_{l_1}^0$	−9.6	7.6
Re $(\phi_{l_2}^m)$	−0.6	−0.5
Im $(\phi_{l_2}^m)$	2.5	2.0

Thus, both solutions locally modify the amplitude of the asymmetric component of the basic state field by a factor near to 1.5, while the overwhelming value of the super rotation would leave the symmetric structure of the solution unmodified except for a shift of the jet axis (as can be easily understood by considering the meridional structure of the zonal component of basic field and that of the perturbation).

In other words, for any initial condition, during the time evolution of the total field (see eq. (3.3)), we would observe changes in the pattern of the spatial distribution of the symmetric component of the flow and a reduction in some latitudinal band of the amplitude of non symmetric component of the total field. This property has to be required for a mechanism that may explain the transition AWR to ZR.

In a forthcoming paper we will expand the analysis above, adding also nonlinear calculations for a triadic interaction arising from the baroclinic model, which will be presented in the next section.

5. Instability for a baroclinic basic state

Although it is not the only possible choice, for simplicity's sake we have considered the extension to a spherical domain of the Phillips two layer model (see Pedlosky, 1979). Perhaps a two layer model is the crudest possible attempt to capture baroclinic dynamics and yet maintaining the needed simplicity in a model that attempts to isolate a particular physical process. For example, placing Ekman layer dissipation in the lower layer, as a two layer quasi-geostrophic model requires, is surely a poor approximation to the atmosphere. Nevertheless, it explores the baroclinic/barotropic dynamics and addresses whether the approach employed may lead to potentially useful results in terms of a mechanism to explain flow transitions.

Let Ψ_1 and Ψ_2 be the stream function and q_1 and q_2 be the potential vorticity at two pressure levels, say 300 mb and 850 mb respectively.

Then the equations of motion are (see Pedlosky, 1979)

$$\frac{D}{Dt} q_1 = 0$$
$$\frac{D}{Dt} q_2 = -\nu \nabla^2 \Psi_2 \tag{5.1}$$

where

$$q_1 = \nabla^2 \Psi_1 + f + F_1(\Psi_2 - \Psi_1)$$
$$q_2 = \nabla^2 \Psi_2 + f + F_2(\Psi_1 - \Psi_2) \tag{5.2}$$

and

$$F_i = \frac{f_0^2}{g \frac{\Delta\rho}{\rho_0} D_i}, \quad \frac{\Delta\rho}{\rho_0} = 0.05,$$
$$D_i = \text{layer} - \text{depth}, \quad i = 1, 2. \tag{5.3}$$

Here, we have neglected any internal dissipation on the layer's interface, which implies that the only dissipative mechanism present in the model is an Ekman pumping confined to the lower layer. F_1 and F_2 are the Froude numbers for the upper and lower layers, and f is the usual Coriolis term. As representative pressure levels of the lower and upper layer, we choose the 850 mb and 300 mb respectively. The free parameters of the problem are F_1, F_2 and, as in the barotropic case, the dissipation rate. The latter was the most sensitive parameter in the calculations so far discussed.

We wish to show that the dependence on dissipation of the growth rate of the unstable eigenmodes is weaker in the two layer model. We will consider instability calculations only for the AWR and ZR regimes. For this purpose we study the instability of the basic fields of the ZR and AWR for the averaged value of $F = F_1 = F_2 = 2$, a typical value (supposing a vertical scale height of about 10 km) for atmospheric conditions.

5.1. *Stability of the AWR*

Consider first the instability of the AWR at T21 resolution. For $F = F_1 = F_2 = 2$, and for the Ekman dissipation a value corresponding to 10 days. The first few most unstable eigenvalues are reported in table II. The eigenvalues represent propagating features with periods typical of baroclinic disturbances. There is, however, an eigenmode with infinite period and a 5.5 days e-folding time. In fig. 5a,b we show the infinite period eigenvector and the correlation map between the eigenvector and the observed transition at the upper level. Compared with the barotropic case (fig. 3b), the correlation pattern is either unchanged or ameliorated in many areas, including the Pacific sector.

To compute the energetics of the two layer model, we will extend the barotropic energetics of the unstable modes (*e.g.*, Branstator, 1985) to include also the baroclinic terms representing the conversion of available potential energy from both the symmetric and asymmetric components of the basic state to the individual growing perturbations

The time rate of change of total energy is so modified

$$\partial_t (KE + APE) = CK_x + CK_y + A + As \tag{5.4}$$

with

$$CK_x = \frac{F_0}{F_1} CK_{x_1} + \frac{F_0}{F_2} CK_{x_2}$$
$$CK_y = \frac{F_0}{F_1} CK_{y_1} + \frac{F_0}{F_2} CK_{y_2} \tag{5.5}$$

where

$$CK_{x_n} =$$
$$= -\frac{1}{a}(u_n'^2 - v_n'^2)\left(\frac{1}{\cos\varphi}\partial_\lambda u_{nb} - v_{nb} \tan\varphi\right)$$
$$CK_{y_n} = \tag{5.6}$$
$$= -\frac{1}{a} u_n' v_n' \left(\cos\varphi \partial_\varphi \left(\frac{u_{nb}}{\cos\varphi} + \frac{1}{\cos\varphi}\partial_\lambda v_{nb}\right)\right)$$

with $n = 1, 2$ referring to the two layers respec-

Table II. Few most unstable eigenvalues for the instability calculation of AWR regime at T21 resolution.

e-folding (days)	period (days)
2.9	5.9
3.8	3.9
4.1	4.3
4.4	2.9
5.1	13.1
5.4	5.3
5.5	Infinity
5.6	7.1
6.1	10.2

tively, and

$$As = \frac{F_0}{a \cos\varphi} \Phi_1 \partial_\lambda \Phi_2 (u_{1b_s} - u_{2b_s}) \tag{5.7}$$

$$A = \frac{F_0}{a \cos\varphi} \Phi_1 (\partial_\lambda \Phi_2 (u_{1b} - u_{2b}) +$$

$$+ \cos\varphi \partial_\varphi \Phi_2 (v_{1b} - v_{2b}))$$

Here F_0 is the mean Froude number of the two levels and u_{1bs} and u_{2bs} are the zonal velocity components of the symmetric basic state in the two respective layers.

Note that in this formulation, the baroclinic source terms are computed only for the conversion of potential energy from the basic state to the perturbation. The barotropic terms are the kinetic energy sources for the perturbation, while the dissipation is the computed kinetic energy dissipation by Ekman pumping in the lower layer. When energetics is calculated using observations, Ekman pumping in the lower layer of the atmosphere is assumed to have the same time scale as prescribed in the model. In fig. 6a we show the observed energetics that shows as the most prominent feature some localized patterns in the Pacific sector. In fig. 6b the energetics for the most unstable stationary baroclinic eigenmode is shown. Numerous features are

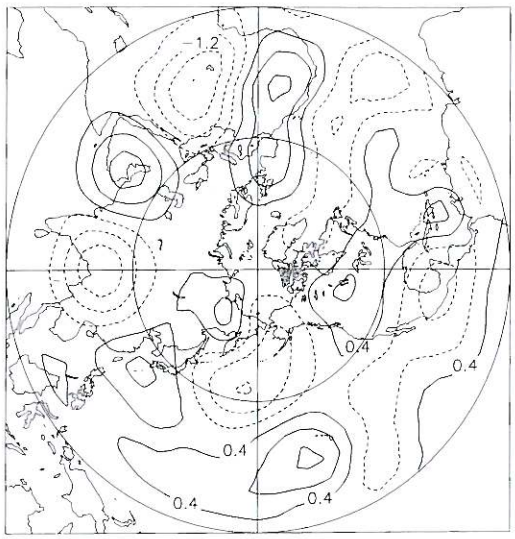

Fig. 5a. Most unstable stationary eigenmode of the AWR basic state in the baroclinic model for the upper level streamfunction pattern with Froude numbers $F_1 = F_2 = 2$. Units are arbitrary with contours every 0.8 units.

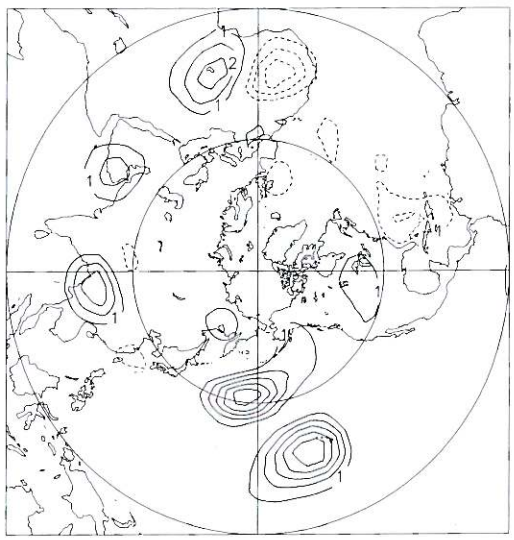

Fig. 5b. Correlation map between the baroclinic eigenvector and the observed transition at the upper level. Units are arbitrary with contours every 1×10^{-3} unit.

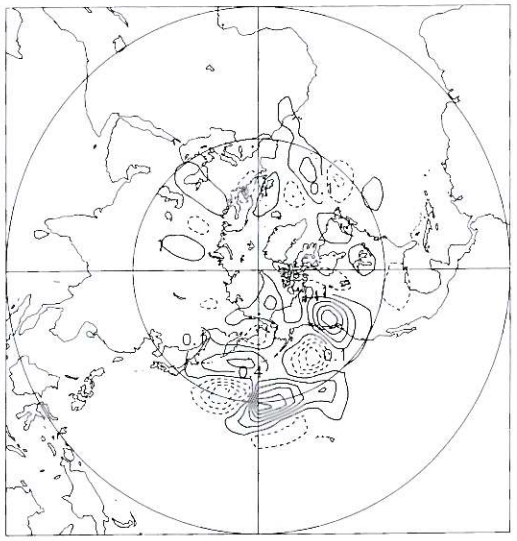

Fig. 6a. Observed baroclinic energetics. Units are arbitrary with contours every 0.3 units.

evident that are due to baroclinic activity, but the prominent patterns in the Pacific sector, noted in the barotropic calculations, are also present.

Figure 6c shows the correlation field between the observed and computed energetics for the standing baroclinic mode shown above. From the latter, we can note that the spatial scale of the physical mechanism presiding over the transition has been well captured. On the contrary, none of the other more unstable modes shows the same degree of similarity. Moreover, it appears that the unwished dependence of the e-folding time on the dissipation has been removed without altering significantly, or actually improving the similarity between the eigenmode and the observed transition map.

To investigate the latter point further, we have explored the dependence of the e-folding time of the mode that as an infinite period on the dissipation time τ. The results are summarized in fig. 7. The weaker dependence of the mode e-folding time when a baroclinic environment is

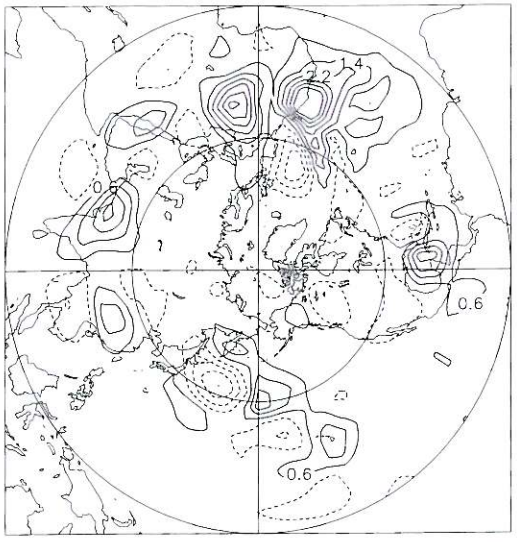

Fig. 6b. Baroclinic energy conversion for the most unstable stationary mode of AWR in the two level model. Units are arbitrary with contours every 0.8 units.

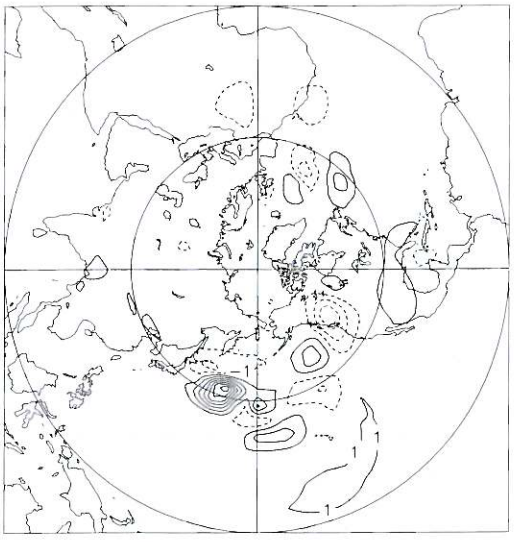

Fig. 6c. Correlation field between the observed and computed energetics for the standing baroclinic mode shown in fig. 5a. Units are arbitrary with contours every 1.5×10^{-3} unit.

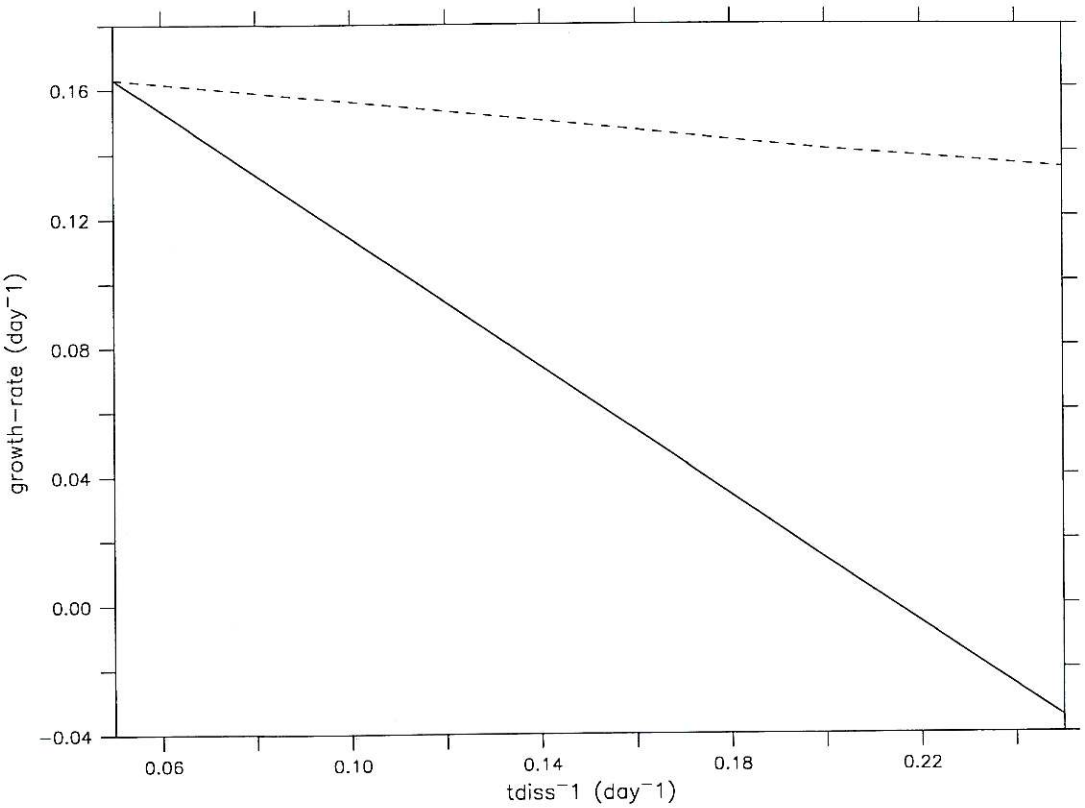

Fig. 7. Dependence of the growth-rate of the most unstable stationary mode in the baroclinic model (dashed line) and in the barotropic model (solid line) on the dissipation. Units are day^{-1}.

considered is shown. Thus the dependence of the growth rate on dissipation in barotropic calculations (Borges and Sardeshmukh, 1995) does not carry over to the baroclinic case, at least for stationary eigenmodes. Further investigations on this subject will be reported elsewhere.

5.2. *Stability of the ZR*

Consider next the stability properties of the ZR in the two layer model. Using values of $F_1 = F_2 = 2$ at T21 resolution, we find predominantly baroclinic time scale traveling modes (see table III). The most unstable nonpropagating eigenvector has an e-folding time of 15

Table III. As in table II but for ZR regime.

e-folding (days)	period (days)
15.3	3.9
14.8	Infinity
.....	.....
5.2	1.9
4.6	9.7
4.3	6.5
4.2	8.3
4.1	2.7
3.9	4.2
3.3	5.4

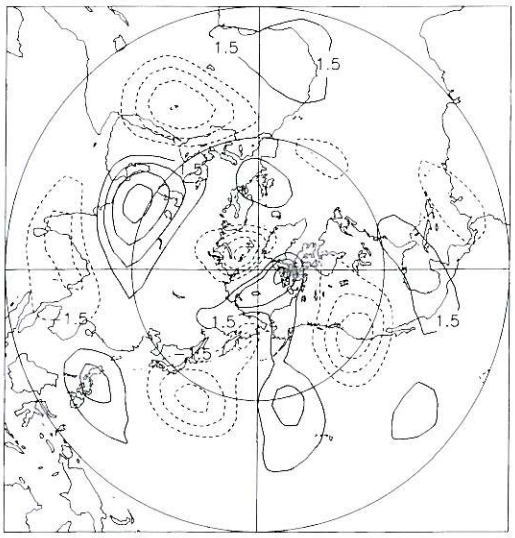

Fig. 8. Upper level streamfunction pattern for the most unstable stationary mode of ZR basic state with Froude numbers $F_1 = F_2 = 2$. Units are arbitrary with contours every 1.5 units.

days. The upper level streamfunction for this eigenvector is given in fig. 8. The pattern of this eigenmode is quite different from that of the AWR, confirming that the amplitude of the asymmetries in the basic state have a profound impact on the eigenmode structure, which is particularly evident for the stationary eigenmodes. However, it is interesting to note that the pattern of the ZR stationary eigenvector is quite different (as shown by a hemispheric correlation of about 10^{-2}) from the observed transition field (fig. 2c). The slow growth rate of this eigenmode mitigates by considering simple linear instability as a candidate to explain this transition, but its dissimilarity from the observations is interesting. In fact, observations suggest that the ZR to AWR transition is stimulated by energetic baroclinic waves featuring strong nonlinear coupling to the large scale flow (H88). Further progress on this issue awaits additional work in a nonlinear environment in which the interaction of the eigenmodes with the basic state and the equilibration of the instabilities are treated explicitly.

6. Conclusions

In this paper we considered the problem of the transition from the two weather regimes identified earlier by Hansen and Sutera (1986, 1995b) in the framework of the instability theory of non-symmetric basic states. This theory is based on the hypothesis that each weather regime is a steady state (or is maintained as such) for the flow. We found that in a barotropic flow linearized around the predominantly Zonal Regime (ZR) is a stable configuration under global perturbations. On the contrary, the Amplified planetary Wave Regime (AWR) is an unstable state with perturbations growing through significant energy extraction from the non-symmetric component of the basic field. For this case the most unstable mode is nonpropagating and the corresponding eigenvector resembles the observed composite of the transition anomaly map for the AWR to ZR transition. However, the e-folding time of this mode is too long for realistic values of the dissipation.

We have shown that in a simple nonlinear triadic model, the finite amplitude states, in the limit of small dissipation, are such that the arbitrariety of the eigenvector sign can be removed (*i.e.* any initial condition would tend to a steady stable state with a zonally asymmetric component smaller than the basic field).

Moreover, by linearizing the equations of motion of a two level Phillips model around the AWR, we found the same eigenmode with a realistic e-folding time of about 5 days. In this case, both the eigenvalue and the eigenvector are only weakly dependent on dissipation. The Ekman dissipation mechanism does not prevent the e-folding from being too long and remains comparable with other modes whose growth rates and propagation characteristics resemble baroclinic disturbances. The energetics of this mode suggests that a significant mechanism of energy extraction is associated with the available potential energy associated with the non-symmetric component of the flow. This energy source prevents the dissipation from inhibiting rapid growth. Barotropic extraction of kinetic energy from the asymmetric basic state is also a significant energy source. Similar energetics can also be found for the observed transition.

On the whole, linear instability of the AWR state is consistent with the observational studies of H88 and appears to be a physical mechanism plausible for explaining its rapid decay.

Regarding the zonal regime, our instability calculations show a plethora of modes extracting energy from the symmetric component of the available potential energy of the flow, and suggest that, perhaps, ordinary baroclinic disturbances and nonlinear interactions may lead to a transition from ZR to AWR.

In concluding we wish to stress that the work presented should be considered only as suggestive of a viable description of low frequency variability of the atmosphere. In fact, many questions have been left open and require further study. Among such questions, the most important are the understanding of the mechanism by which a particular unstable eigenvector is selected by the flow and the role that non-linearities play in damping some modes rather than others. In particular, it should be considered why unstable modes with higher e-folding times and with shorter periods are not preferred by the flow for transition from the AWR.

The work presented has exploited the merits of two different lines of research for understanding atmospheric low frequency variability which too often have been thought of as alternative approaches. These are linear instability and multiple weather regimes theories. Clearly, the approach taken in this paper should be extended to a model with better vertical resolution in order to treat the dissipation in a manner more harmonious with the atmosphere. We anticipate that the structure and dynamics of the relevant eigenmodes in such a model will be more robust and exhibit energetics more consistent with observations, but at the trade off of much more complexity (more parameters to specify) and of computational efficiency.

Acknowledgements

The author is grateful for financial support of the European Community under grant ENV4-CT96-0344 and the Italian Space Agency (ASI).

REFERENCES

ANDREWS, D.G. (1984): On the stability of forced non-zonal flows, *Q. J. R. Meteorol. Soc.*, **110**, 657-662.

BAINES, P.G. (1976): The stability of planetary waves on a sphere, *J. Fluid Mech.*, **73**, 193-213.

BAUR, F. (1951): Extended range forecasting, *Compendium of Meteorology*, Americal Meteorological Society, Boston, 814-833.

BENZI, R., P. MALGUZZI, A. SPERANZA and A. SUTERA (1986a): The statistical properties of the general atmospheric circulation: observational evidence and a minimal theory of bimodality, *Q. J. R. Meteorol. Soc.*, **112**, 661-676.

BENZI, R., A. SPERANZA and A. SUTERA (1986b): A minimal baroclinic model for the statistical properties of low frequency variability, *J. Atmos. Sci.*, **43**, 2962-2967.

BORGES, M.D. and D.L. HARTMANN (1992): Barotropic instability and optimal perturbations of observed nonzonal flows, *J. Atmos. Sci.*, **49**, 335-354.

BORGES, M.D. and P.D. SARDESHMUKH (1995): Barotropic Rossby wave dynamics of zonally varying upper-level flows during Northern Hemisphere winter, *J. Atmos. Sci.*, **52**, 3779-3796.

BRANSTATOR, G.W. (1985): Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. Part II: Eigenanalysis, *J. Atmos. Sci.*, **42**, 2242-2254.

BRANSTATOR, G.W. (1990): Low-frequency patterns induced by stationary waves, *J. Atmos. Sci.*, **47**, 629-648.

BUIZZA, R. and F. MOLTENI (1996): The role of finite time barotropic instability during transition to Blocking, *J. Atmos. Sci.*, **53**, 1675-1697.

CHANG, J.C.J. and M. MAK (1995): Nonmodal barotropic dynamics of the intraseasonal disturbances, *J. Atmos. Sci.*, **52**, 896-914.

CHARNEY, J.G. and J.G. DEVORE (1979): Multiple flow equilibria in the atmosphere and blocking, *J. Atmos. Sci.*, **36**, 1205-1216.

CHENG, X. and J.M. WALLACE (1993): Cluster analysis of the Northern Hemisphere wintertime 500-hPa height field: spatial patterns, *J. Atmos. Sci.*, **50**, 2674-2696.

FREDRICKSEN, J.S. and P.J. WEBSTER (1988): Alternative theories of atmospheric teleconnections and low-frequency fluctuations, *Rev. Geophys.*, **26**, 459-494.

HANSEN, A.R. (1988): Further observational characteristics of atmospheric planetary waves with bimodal amplitude distributions, *Mon. Weather Rev.*, **116**, 386-400.

HANSEN, A.R. and A. SUTERA (1986): On the probability density distribution of the planetary-scale atmospheric wave amplitude, *J. Atmos. Sci.*, **43**, 3250-3265.

HANSEN, A.R. and A. SUTERA (1987): The probability density distributions of the speed, horizontal shear and vertical shear of the zonal-mean flow, *J. Atmos. Sci.*, **44**, 1523-1533.

HANSEN, A.R. and A. SUTERA (1993): A comparison between planetary-wave flow regimes and blocking, *Tellus*, **45A**, 281-288.

HANSEN, A.R. and A. SUTERA (1995a): Large-amplitude flow anomalies in Northern Hemisphere mid latitudes, *J. Atmos. Sci.*, **52**, 2133-2151.

HANSEN, A.R. and A. SUTERA (1995b): The probability density distribution of planetary-scale atmospheric wave amplitude revisited, *J. Atmos. Sci.*, **52**, 2463-2472.

HANSEN, A.R. and A. SUTERA (1995c): On the role of topography in the low-frequency variability of the large-scale mid latitude circulation, *J. Atmos. Sci.*, **52**, 2497-2508.

KATO, T. (1966): *Perturbation Theory for Linear Operators* (Springer-Verlag Berlin), pp. 590.

KIMOTO, M. and M. GHIL (1993): Multiple flow regimes in the Northern Hemisphere winter. Part I: Methodology and hemispheric regimes, *J. Atmos. Sci.*, **50**, 2625-2643.

LEGRAS, B. and M. GHIL (1985): Persistent anomalies, blocking, and variations in atmospheric predictability, *J. Atmos. Sci.*, **42**, 433-471.

LINDZEN, R.S., E.N. LORENZ and G.W. PLATZMAN (Editors) (1990): The Atmosphere-A Challange. The science of Jule Gregory Charney, *Am. Meteorol. Soc. Historical Monograph Series*, pp. 321.

LORENZ, E.N. (1972): Barotropic instability of Rossby wave motion, *J. Atmos. Sci.*, **29**, 258-264.

MALGUZZI, P., A. SPERANZA, A. SUTERA and R. CABALLERO (1996): Nonlinear amplification of stationary Rossby waves near resonance. Part I, *J. Atmos. Sci.*, **53**, 298-311.

MOLTENI, F. (1996a): On the dynamics of planetary flow regimes. Part I: The role of high frequency transients, *J. Atmos. Sci.*, **53**, 1950-1971.

MOLTENI, F. (1996b): On the dynamics of planetary flow regimes. Part II: Results from a hierarchy of orographically forced models, *J. Atmos. Sci.*, **53**, 1972-1992.

MOLTENI, F. and T.N. PALMER (1993): Predictability and finite-time instability of the northern winter circulation. *Q. J. R. Meteorol. Soc.*, **119**, 269-298.

MOLTENI, F., A. SUTERA and N. TRONCI (1988): The EOFs of the geopotential eddies at 500 mb in winter and their probability density distributions, *J. Atmos. Sci.*, **45**, 3063-3080.

MOLTENI, F., S. TIBALDI and T.N. PALMER (1990): Regimes in the wintertime circulation over northern extratropics. I: Observational evidence, *Q. J. R. Meteorol. Soc.*, **116**, 31-67.

NITSCHE, G., J.M. WALLACE and C. KOOPERBERG (1994): Is there evidence of multiple equilibria in planetary wave amplitude statistics?, *J. Atmos. Sci.*, **51**, 314-322.

PALMER, T.N. (1988): Medium and extended range predictability and the stability of the Pacific North American mode, *Q. J. R. Meteorol. Soc.*, **114**, 691-714.

PEDLOSKY, J. (1979): *Geophysical Fluid Dynamics* (Springer-Verlag), pp. 624.

ROSSBY, C.G. (1959): Current problems in meteorology, *The Atmosphere and Sea in Motion*, edited by B. BOLIN (Rockefeller Inst. Press, New York), 9-50.

SIMMONS, A.J., J.M. WALLACE and G.W. BRANSTATOR (1983): Barotropic wave propagation and instability, and atmospheric teleconnection patterns, *J. Atmos. Sci.*, **40**, 1363-1392.

SUTERA, A. (1986): Probability density distribution of large-scale atmospheric motion, *Adv. Geophys.*, **29**, 227-249.

(received May 4, 1999;
accepted July 20, 1999)