

Barotropic Instability and Optimal Perturbations of Observed Nonzonal Flows

MARK D. BORGES AND DENNIS L. HARTMANN

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

(Manuscript received 28 January 1991, in final form 8 July 1991)

ABSTRACT

An eigenvalue analysis of a divergent barotropic model on a sphere is extended to the formulation of a global optimization problem, whose solution selects an initial perturbation that evolves into the most energetic structure at a finite time interval, τ . The evolution of this perturbation is obtained from companion linear and nonlinear global spectral time-dependent models, and the optimization prediction of perturbation size at time τ is verified. Two zonally asymmetric flows defined by time-mean ECMWF global 300-mb analyses during winter 1985/86 are used to illustrate the application and insights provided by the optimization problem.

The dependence of the optimal perturbations on the parameter τ is examined. The optimal perturbations become increasingly localized as τ is decreased to periods on the order of three days. The initial growth rates of these perturbations greatly exceed that of the most unstable normal mode, and also exceed the growth rate of a disturbance with maximum projection onto the most unstable mode (i.e., the adjoint structure). Furthermore, the development of the optimal perturbations in the nonlinear model is in reasonable agreement with the available observations. The optimal perturbations may thus be more important than either the eigenmode or adjoint structure for determining the stability and expected behavior of anomalies to some time-mean flows.

1. Introduction

Long-duration, high-amplitude anomalies in the midlatitude flow field are a practically important and scientifically challenging meteorological problem. Because the amplitude and duration of these anomalies are at the extremes of normal weather variability, they can produce extreme weather events. The importance of these extreme events is amplified because they tend to remain geographically fixed, thus subjecting fixed areas to prolonged periods of extreme temperatures or precipitation. Rex (1950a,b) described the phenomenon of blocking, which emphasizes the development of an extreme high pressure anomaly in the westerlies. Other ways of describing low-frequency anomalies with time scales between 10 and 90 days are teleconnections in low-pass filtered data (Wallace and Gutzler 1981) and persistent anomalies (Dole and Gordon 1983). The structure of these disturbances is often very nearly that of a quasi-stationary barotropic Rossby wave group.

It is known that the wave energy released during the growth of transient baroclinic waves will quickly evolve to a nearly barotropic structure (Gall 1976; Simmons and Hoskins 1980), and that through nonlinear dynamics this energy will migrate toward larger spatial scales (Fjortoft 1953) and eventually collect on the

zonal mean flow or in quasi-stationary Rossby waves (Rhines 1975). For these reasons low-frequency variability can be produced in an atmosphere with no stationary wave forcing (e.g., Hendon and Hartmann 1985). Nonetheless, because of large-amplitude stationary zonal asymmetries forced by topography and surface heating, the low-frequency variability in the Northern Hemisphere has well-defined spatial structures and tends to occur in preferred geographical regions. One method of understanding these spatial structures is to consider a linearization about a zonally varying time-mean state and examine the structures associated with the barotropic (Simmons et al. 1983, hereafter referred to as SWB) or baroclinic (Frederiksen 1983) instabilities of this mean state. These and related studies reveal that barotropic instability associated with east-west variations in jet intensity can give rise to structures that resemble teleconnection patterns and the observed low-frequency variability. Therefore, it appears that barotropic wave propagation and instability within the context of a zonally varying time-mean flow form an important part of the explanation for low-frequency anomalies and a guide to their potential predictability.

The use of unstable modes as a tool for interpreting the development of persistent anomalies presents several problems. The growth rate and structure of the unstable modes is very sensitive to the mean structure on which they develop. In nature, the mean field evolves on a time scale that is comparable to the e-folding time of the unstable modes. One is faced with

Corresponding author address: Dr. Mark D. Borges, University of Washington, Department of Atmospheric Sciences AK-40, Seattle, WA 98195.

the reality that atmospheric flow is fundamentally nonlinear. Moreover, often several modes can exist with different structure and similar growth rates. The structures tend to be global in nature, and it is difficult to associate a particular unstable mode with a local high-amplitude event. Some of these difficulties can be avoided by applying some additional constraints to the instability problem, and by considering mean states defined on a shorter time interval than a month or a season (e.g., Frederiksen 1989). Observational studies suggest that low-frequency circulation anomalies develop quickly, with consistent precursor events identifiable in composites at most about five days before the event matures (Dole 1986; Dole and Black 1990; Nakamura and Wallace 1990). This time is short compared to the e -folding times of the quasi-stationary modes most often associated with low-frequency anomalies, although SWB showed that local apparent growth rates can be much greater than that of the global structure associated with the mode. From this perspective, then, it might be better to seek structures that achieve maximum growth rate over a finite period of time, rather than the infinite period of time allowed for unstable normal modes in a traditional instability analysis.

Farrell (1989a) has argued that initial perturbations selected to achieve the maximum growth after a finite time interval provide a guide to understanding wave development that improves upon traditional analysis of unstable modes and is more applicable to observed developments of synoptic disturbances. As Farrell (1982) and Boyd (1983) showed, properly configured initial disturbances can exhibit rapid initial growth rates even on a mean state that is stable to small perturbations, although this growth rate is not sustained. It is this initial growth rate that is of greatest importance, since it is only during this period that the linear theory is valid. Once the large anomaly is established, it may be sustained by nonlinear processes (e.g., Illari and Marshall 1983; Haines and Marshall 1987; Holopainen and Fortelius 1987).

In this study we investigate the unstable modes and optimal initial perturbations of the global 300-mb flow. The dynamics of low-frequency variability has been frequently studied with a barotropic model (e.g., Hoskins et al. 1977) and is usually justified by the equivalent barotropic structure of the features under study. In our study we will use a variant of the barotropic model as a tool to investigate the rapid development of persistent anomalies. As noted by Dole (1986), the precursors of his PAC persistent anomaly pattern are most easily defined near 300 mb and possess a westward tilt with height along the 45°N latitude circle. We certainly cannot and do not discount the possible importance of baroclinic processes during this early stage of the anomaly's life cycle. Our aim is to focus attention on the barotropic aspects of the anomaly that may previously have been overlooked. To the extent that vortex

tube stretching by the perturbation is negligible, this model is an adequate first approximation of the vorticity dynamics at upper tropospheric levels. Frederiksen and Bell (1990) studied the instability of instantaneous flow states and found stationary modes dominated by barotropic instability growth mechanisms with growth rates comparable to those of baroclinic cyclogenesis modes in some cases.

The optimal initial perturbation is formed from the complete set of linear modes associated with the mean state and model resolution. Damped and neutral modes are as important as the growing modes in the formation of this initial state, and allow the initial condition to have a localized structure, in contrast to the global structure of the individual modes. When the period allowed for growth is limited to a few days, the optimal initial perturbation tends to be highly localized, thereby identifying the critical location for initial growth. The technique is applied to time means for specific 16-day periods. We have chosen 16 days to have a separation between the time scales present in the mean state and the phenomena we wish to explain using the linearization. Though the choice is somewhat arbitrary, we feel a 16-day average is an appropriate mean state for this problem. The optimal initial states evolve into large-amplitude flow anomalies that closely resemble the observed anomalies for the periods studied. Nonlinear integrations from the optimal initial conditions also provide realistic predictions of the observed anomaly development.

The following section summarizes some previous stability studies and their application to explain observed low-frequency anomalies. The motivation for this study and model governing equations are also presented and defined there. Section 3 discusses some pertinent concepts concerning stability analyses of zonally asymmetric basic flows, and is followed in the next section by a formulation of the global optimization problem for nonzonal flows. The results of the application of the method to two particular time-mean 300-mb flows are discussed in section 5. The final section summarizes our results and conclusions.

2. Traditional stability analyses

The decomposition of geophysical flows into a zonal mean and departures therefrom has proven to be a valuable framework for the study of flow instability. By considering an idealized basic flow with vertical shear alone, Charney (1947) and Eady (1949) isolated the baroclinic instability mechanism and showed that the zonal mean component of realistic midlatitude tropospheric flows is unstable. The resulting exponentially amplifying perturbation structure (i.e., a "normal" mode) has length and time scales similar to observed atmospheric cyclogenesis, and thus provides one explanation of day-to-day weather.

One can also consider an idealized flow with hori-

zonal shear alone to isolate the barotropic instability mechanism (Kuo 1949). When the zonal mean component of observed tropospheric flows is analyzed for barotropic instability, however, the general result has been that such flows are stable to perturbations of normal-mode form (Lorenz 1972; see Zhang 1988 for a recent calculation). Consequently, the relevance of barotropic instability to atmospheric variability received little attention prior to the 1970s. In a pioneering paper on planetary wave stability, Lorenz (1972) re-examined the barotropic instability mechanism, but with zonal variation added to the basic flow in the form of a single Rossby wave. He was able to analytically show that this two-dimensional flow is strongly unstable and suggested its implications for extended-range predictability. Further analytic study of more general zonally inhomogeneous flows was hindered by the added complexity such flows introduce. Moreover, the nonzonal flow one would typically like to study (e.g., a time-mean flow) is not a solution to the steady-state barotropic vorticity equation; an external forcing field is usually required to render the nonzonal basic flow a stationary solution. This external forcing, which may be viewed as representative of the effects of transient eddies, topography, or diabatic effects on the basic flow, is assumed constant in its simplest form. The linearized perturbation equation is then autonomous; that is, the forcing that maintains the time-mean flow is not permitted to interact with or otherwise influence the developing perturbations. As computing power increased, many numerical studies of the instability of zonally asymmetric flows were carried out using this paradigm (e.g., Branstator 1982; Fredericksen 1982, 1983, 1989; Simmons 1982; SWB; Branstator 1985; Zhang 1988). A brief review of the results obtained applying this method to low-frequency variability follows.

Simmons (1982) was one of the first to investigate the behavior of perturbations placed in a nonzonal flow derived from climatology. He found a much larger forced response in a nonlinear barotropic time-dependent model by using a nonzonal basic flow instead of the zonal mean. The geographic location of the enhanced response was found to be remarkably close to regions of maximum observed low-frequency variability, namely, the North Pacific and North Atlantic. Furthermore, Simmons was able to demonstrate a strong sensitivity to the location of the specified perturbation forcing in the model, with the largest responses in the North Pacific occurring when the forcing was located on the equatorward flank of the subtropical Asian jet.

Later, Branstator (1982) and SWB found the behavior to be consistent with the most rapidly growing normal mode associated with barotropic instability of the climatological January mean flow used by Simmons. The inviscid e -folding growth rate of about a week and period of 45 days for the fastest-growing mode agreed well with companion initial-value sim-

ulations. Through an energetics analysis, they determined that zonally elongated structures straddling the jet exit region could extract energy more efficiently from the mean flow, and from this argued that those finite-amplitude anomaly patterns that resembled such normal modes were in a more energetically favorable position, and hence more likely to appear in low-pass filtered statistics. The relevance of barotropic instability to low-frequency variability was strongly suggested.

Branstator (1985) and Zhang (1988) have shown that the response of the fastest-growing barotropic mode to forcing in specific geographic locations is related to the structure of its adjoint. As discussed later (see the Appendix and section 3), the projection onto the adjoint determines the spectral characteristics of a particular structure. It turns out that the adjoint of the fastest-growing mode of the January climatology has most of its amplitude concentrated in the subtropical latitudes, south of the Asian jet (Zhang 1988; Ferranti et al. 1990). Therefore, knowledge of the adjoint mode structure helps to answer the question of how best to excite a *particular* eigenfunction; if this mode dominates the response (e.g., as in SWB) the model atmosphere will exhibit strong sensitivity to forcing in this region. This result was also obtained by Simmons (1982) through a less direct method.

One of the remaining problems acknowledged by SWB was whether these relatively slowly growing modes¹ can compete with baroclinic instability or even exist in more general flows with vertical shear. Fredericksen and Bell (1990) have recently studied aspects of blocking by conducting an eigenanalysis of a five-level quasigeostrophic model linearized about instantaneous Northern Hemisphere flow fields. As in earlier work (see the review article of Fredericksen and Webster 1988 and references therein), they bin the normal modes according to period into classes such as "monopole cyclogenesis," "onset-of-blocking," and "mature anomaly." By using instantaneous flows the structures of the latter two classes appear more localized and have larger e -folding times (on the order of two days) compared to a parallel analysis of climatological flows. Further, they assert the mature anomaly modes are initiated by the onset modes, which they find to be eastward-propagating dipoles with westward tilts with height and that this combined baroclinic-barotropic instability is more relevant than pure barotropic instability for events such as the onset of blocking. We feel some caution should be exercised in interpreting these results, however, as the use of an instantaneous basic state is formally valid only when the time tendency of the basic state is small compared to the tendencies of

¹ As compared to barotropic instability growth rates. Using instantaneous 300-mb flow patterns, Branstator (1982) found fastest mode e -folding times on the order of two days. Their structure, however was not shown.

the unstable modes. It is not obvious that this condition is satisfied for the modes in question.

There are two aspects of the traditional eigenanalysis of zonally asymmetric basic flows that should concern the investigator. The first, as alluded to before, deals with how the basic state is specified and maintained. The second aspect is more philosophical and concerns the interpretation of the eigenmodes as distinct physical entities.

Regarding the first concern, Andrews (1984) has argued that the mathematical form the forcing takes is crucial to the resulting stability problem. As an example, Andrews shows that if the external forcing takes the form usually thought to mimic topographic forcing in a β -plane barotropic model, the resulting linearized perturbation equation is not autonomous. Furthermore, the perturbation forcing that arises can act to stabilize perturbations to a mean flow that would otherwise be considered unstable without the forcing. Zhang (1988) has found a similar effect of decreasing growth rates in a global barotropic model by using an equivalent mountain forcing to maintain the basic state as compared to the constant forcing artifice.

An alternative to the topographically forced barotropic model that sidesteps the artificial forcing device is the "divergent barotropic model" discussed by Sardeshmukh and Hoskins (1988). This model arises from consideration of an upper tropospheric vorticity balance, where modeling and observational evidence suggests that vertical advection and twisting terms are of secondary importance, at least in the tropics (Sardeshmukh and Held 1984; Sardeshmukh and Hoskins 1985). This balance may be expressed,

$$\frac{\partial \zeta}{\partial t} + \nabla \cdot (\mathbf{V}(\zeta + f)) = F_d - (C_d + \nu \nabla^4) \zeta$$

or,

$$\frac{\partial \zeta}{\partial t} + \mathbf{V}_\psi \cdot \nabla(\zeta + f) = F_d - (C_d + \nu \nabla^4) \zeta - \{\nabla \cdot (\mathbf{V}_x(\zeta + f))\}. \quad (1)$$

Here, F_d represents an external forcing or a simple parameterization of the neglected vertical advection and twisting terms and the total wind has been decomposed into its irrotational, \mathbf{V}_x , and nondivergent, \mathbf{V}_ψ , parts. Linear damping and diffusion are included in the form of a constant Rayleigh damping coefficient C_d and fourth-order diffusivity, ν . Other symbols have their usual meaning. In all cases to be considered ν has been set to provide a damping time scale of 1/4 day for spatial scales of total wavenumber 42, the same value used in many other studies (e.g., Branstator 1982; Simmons 1982; Branstator 1985; SWB; Hoskins and Sardeshmukh 1987).

The steady-state balance of (1) is,

$$\bar{\mathbf{V}}_\psi \cdot \nabla(\bar{\zeta} + f) = \bar{F}_d - (C_d + \nu \nabla^4) \bar{\zeta} - \{\nabla \cdot [\bar{\mathbf{V}}_x(\bar{\zeta} + f)] + \nabla \cdot (\bar{\mathbf{V}} \bar{\zeta}')\} \quad (2)$$

where the overbar represents a time average and a prime the departure therefrom. The linearized equation governing small perturbations to the time-mean basic flow is,

$$\frac{\partial \zeta'}{\partial t} + \mathbf{V}'_\psi \cdot \nabla(\bar{\zeta} + f) + \bar{\mathbf{V}}_\psi \cdot \nabla \zeta' = F'_d - (C_d + \nu \nabla^4) \zeta' - \{\nabla \cdot (\bar{\mathbf{V}}_x \zeta') + \nabla \cdot [\mathbf{V}'_x(\bar{\zeta} + f)]\}. \quad (3)$$

If we ignore the terms in braces in (1)–(3) and define $F_d = \bar{F}_d$ by (2), we obtain the set of equations studied by SWB. The consistent linearization of this model for initial-value simulations would make $F'_d \equiv 0$ in (3). The advection of absolute vorticity by the time-mean divergent wind in (2) contributes to the maintenance of the time-mean basic state, and leads to an additional term in (3) that will affect the perturbations. In fact, one could set $\bar{F}_d \equiv 0$ in (2) and solve for $\bar{\mathbf{V}}_x$ given $\bar{\mathbf{V}}_\psi$ and the transient vorticity flux (the χ problem solved by Sardeshmukh and Hoskins 1987), thus eliminating F_d altogether. In the linear equation (3) inclusion of the mean divergent wind acts like a time-dependent forcing on the perturbation. One could also specify a perturbation divergent wind, \mathbf{V}'_x (e.g., to represent tropical intraseasonal variability as in Ferranti et al. 1990), to obtain the final perturbation forcing term in (3).

For this study we will use the ECMWF analyzed $\bar{\mathbf{V}}_x$, and define $F_d = \bar{F}_d$ as a residual of (2) after neglecting the transient vorticity flux, that is:

$$F_d = \bar{\mathbf{V}}_\psi \cdot \nabla(\bar{\zeta} + f) + (C_d + \nu \nabla^4) \bar{\zeta} + \{\nabla \cdot [\bar{\mathbf{V}}_x(\bar{\zeta} + f)]\}. \quad (4)$$

Equation (1) with $\mathbf{V}_x = \bar{\mathbf{V}}_x$ and F_d defined by (4) is the nonlinear governing equation of the model adopted for this study. For nonlinear initial-value simulations a Rayleigh damping coefficient $C_d = (10 \text{ days})^{-1}$ is included. The corresponding linearized equation is (3) with $F'_d = \mathbf{V}'_x = C_d \equiv 0$, namely,

$$\frac{\partial \zeta'}{\partial t} + \mathbf{V}'_\psi \cdot \nabla(\bar{\zeta} + f) + \bar{\mathbf{V}}_\psi \cdot \nabla \zeta' + \nu \nabla^4 \zeta' + \nabla \cdot (\bar{\mathbf{V}}_x \zeta') = 0 \quad (5)$$

so that the mean divergent wind, which helps maintain the basic state, is also felt by the developing perturbations via the last term in (5).

The model equations are expanded in spherical harmonics using the spectral-transform technique in the manner of Bourke (1972). The resulting ordinary differential equations for vorticity are time differenced using a third-order Adams–Bashforth method that

eliminates the need for a time filter (Durran 1991). For all cases to be discussed a triangular truncation retaining 31 wavenumbers (T31) was used to obtain the final results.

The second area of concern deals with the interpretation of the eigenmodes themselves. Since an eigenanalysis of an N -component system in general produces N different normal modes, one might ask whether consideration of each of these modes in isolation has any physical significance. In zonally symmetric problems with latitudinal shear we know some of these modes will be discrete, while the remainder approximate the continuum (Yanai and Nitta 1968; Kasahara 1980). Attention has generally been focused on the discrete modes, on the basis that the continuum modes must asymptotically decay as t^{-2} (Orr 1907; Case 1960). Farrell (1982) and Boyd (1983) show, however, that the continuum spectrum in simple linear shear flow problems (where no normal modes can exist) allows some perturbations to grow significantly before decaying. Here we investigate whether the same potential for transient growth exists in zonally asymmetric unstable flows by searching for favorably configured (optimal) perturbations.

Farrell (1988, 1989a) formulated two problems to determine the structure of the optimal perturbations that will 1) most efficiently excite a particular linear mode, and 2) have the largest "size" at the end of a fixed time interval. He chose norms representing streamfunction amplitude and energy as measures of perturbation size, and showed the solution dependence on the particular measure used. Farrell (1988) solved the optimization problems for a barotropic β -plane channel model linearized about a constant shear basic flow, while Farrell (1989a) solved the optimization problems for a quasigeostrophic β -plane model with a linear height-dependent basic flow that allowed for unstable modes. The solution of the energy optimization problem in this latter work indicated that a structure tilting against the vertical shear would grow most rapidly, and was in broad agreement with common precursors of cyclogenesis. Farrell (1989a) found that these optimal perturbations can exhibit growth rates that, initially, greatly exceed those of exponentially unstable modes (see also Lacarra and Talagrand 1988; Mak and Cai 1989) and argued that even in unstable flows the optimal perturbations are likely to limit short-term and medium-range forecasts.

The solution of the optimization problem for these idealized basic flows provides a direct, objective way to find the most rapidly growing structures (for a fixed time interval), without any constraint on their time behavior. By using observed zonally asymmetric flows we hope the method will provide similar insights for studies of low-frequency variability. During the course of this work, we became aware that Frederiksen and Bell (1990) have recently done a similar calculation. They found the optimal perturbation that maximizes

the initial instantaneous growth in the squared streamfunction amplitude (L_2 norm squared) for their R15 five-level quasigeostrophic model linearized about instantaneous synoptic flows. They discovered that the optimal structure had largest amplitude at the surface and was localized in the blocking region. Its growth in a linear model was not sustained and by day 10 had an order of magnitude smaller amplitude and two orders of magnitude smaller kinetic energy than the adjoint to the most unstable eigenmode initial condition. From this they concluded that the adjoint to the eigenmode, not the optimal initial state, was most relevant. Our study presents a range of optimizing intervals, τ , in a barotropic model with higher horizontal resolution, which we feel is necessary to adequately capture both the adjoint and optimal initial condition structures.

3. Nonself-adjoint systems and projectability

A full derivation of the eigenvalue problem for zonally asymmetric basic states, formulated using a spherical harmonic expansion, may be found in Branstator (1982), SWB, and more recently Zhang (1988). Branstator (1985) and Zhang (1988) have also discussed the concept of "projectability" and its importance in general linear systems. Our formulation of the eigenvalue problem associated with (5) follows Branstator (1985). Schematically, we write

$$\frac{d\psi}{dt} - \mathcal{L}\psi = 0 \quad (6)$$

where ψ is a complex vector of spectral streamfunction coefficients and \mathcal{L} denotes a spectral-space linear operator that includes advective terms and vortex tube stretching by the mean divergent wind acting on the perturbation vorticity field. We truncate the spherical harmonic expansion of (6) triangularly at total wavenumber 31 (T31) and refer to the resulting matrix of order 1023 by \mathbf{L} . As shown in the Appendix, the nonself-adjoint matrix, \mathcal{L} , is associated with an adjoint matrix, \mathbf{M} , whose eigenvectors $\{y_k\}$ are biorthogonal to the eigenvectors $\{x_j\}$, of \mathbf{L} , that is,

$$\langle y_k, x_j \rangle = 0, \quad \text{if } \sigma_j \neq \lambda_k^* \quad (7)$$

where σ_j is an eigenvalue of \mathbf{L} , λ_k^* is the complex conjugate of an eigenvalue of \mathbf{M} , and $\langle \cdot, \cdot \rangle$ denotes an inner product operator.

The biorthogonality has important consequences for the spectral representation of an arbitrary field, G , in the space spanned by the eigenvectors. For example, consider the spectral expansion of G ,

$$G = \sum_{j=1}^N g_j x_j. \quad (8)$$

To determine the projection coefficients, $\{g_j\}$, we pro-

ceed in the usual manner by forming the inner product and, using (A7), find that

$$g_k = \langle \mathbf{y}_k, G \rangle \gamma_k. \quad (9)$$

The projection coefficients in general linear systems depend on the structure's projection onto the *adjoint*, not the eigenmode. The quantity $\gamma_k = \langle \mathbf{y}_k, \mathbf{x}_k \rangle^{-1}$, termed the "projectability" by Zhang (1988), acts as an amplification factor. The important result is that even though G may have a weak projection onto the k th adjoint structure, its modal expansion coefficient, g_k , may be large if the k th mode has a large projectability. The projectability is a measure of the degree of orthogonality of the normal modes; a large γ_k means that \mathbf{x}_k is nearly parallel to the hyperplane defined by $\{x_j\}_{j \neq k}$ in phase space. This is clearly favorable for putting a large spectral amplitude in a *particular* mode under constrained conditions. For instance, suppose that G must be of unit size. One can think of putting large amplitude in the desired mode, then masking its amplitude by using the other, nearly parallel modes to "cover" it up. We note that setting G equal to the adjoint structure, \mathbf{y}_k , will give the maximum projection onto a particular mode.

These ideas have recently been applied in a study by Ferranti et al. (1990). In that investigation the adjoint mode (defined by the barotropic global kinetic energy inner product) for a particular time-mean flow (the January climatology used by SWB) was found to be very similar to the vorticity forcing thought to be associated with the tropical 30–60-day oscillation. They found that the extratropical response to this forcing in a simple barotropic model matched the observed low-frequency anomalies reasonably well and suggested that this interaction may play a role in the observed mid-latitude intraseasonal variability.

4. The optimization problem for the global spectral barotropic model

Following Farrell (1988, 1989a), the general solution process consists of performing a traditional stability analysis of the flow to determine a set of eigenfunctions. Assuming that distinct eigenvalues exist, the eigenfunctions form a complete basis set in which to expand an arbitrary disturbance. The evolution of the disturbance can then be determined by merely advancing the amplitude and phase of each eigenfunction according to its respective eigenvalue. Using this framework an optimization problem can be formulated and solved for the spectral expansion coefficients. Note that this method of superposing normal-mode solutions is essentially equivalent to a Laplace (or Fourier) transform method (Tung 1983). Since we are actually solving the initial-value problem in this manner, the presence of continuum modes and their algebraic time dependence requires the use of many exponential functions to adequately capture the time evolution.

To be specific, we want to determine the initial condition of fixed energy that will evolve into the most energetic structure after a specified time interval, τ . That is, we seek to maximize $K(\tau)$ subject to the constraint that $K(0) = 1$, where $K(t)$ is the global mean kinetic energy associated with the rotational part of the flow at time t . This problem is equivalent to maximizing the "auxiliary" function, φ (Hildebrand 1963, pp. 352–355),

$$\varphi = K(\tau) + \lambda(K(0) - 1) \quad (10)$$

with *no* constraints, λ being a Lagrange multiplier. If $\varphi = \varphi(x)$ we can find the extrema in the usual way by setting $d\varphi/dx = 0$. With this in mind we now proceed to find the functional form for $K(t)$.

The global mean kinetic energy is

$$K(t) = \langle \boldsymbol{\psi}(t), \boldsymbol{\psi}(t) \rangle_{\text{KE}} = \frac{1}{2} (\boldsymbol{\psi}^H \mathbf{D} \boldsymbol{\psi}) \quad (11)$$

where $\boldsymbol{\psi}_N$ is a (complex) vector of spherical harmonic expansion coefficients for the streamfunction, $\mathbf{D}_{N \times N}$ is a real positive-definite diagonal matrix defining the energy norm, and $()^H$ denotes the Hermitian. Assuming the eigenvalue problem (6) provides a complete set of eigenvectors, $\{x_j\}$, we can expand the streamfunction,

$$\boldsymbol{\psi}(t=0) = \mathbf{X} \boldsymbol{\alpha} \quad (12)$$

where $\mathbf{X}_{N \times N}$ is a matrix with the eigenvectors as columns and $\boldsymbol{\alpha}_N$ is a vector of projection coefficients. Substituting (12) in (11), the matrix expression for kinetic energy becomes,

$$K(t=0) = (\mathbf{X} \boldsymbol{\alpha})^H \mathbf{D} (\mathbf{X} \boldsymbol{\alpha}) = \boldsymbol{\alpha}^H (\mathbf{X}^H \mathbf{D} \mathbf{X}) \boldsymbol{\alpha}. \quad (13)$$

The eigenvectors will of course evolve in time according to $x_j e^{\sigma_j t}$ so that at any particular time $\boldsymbol{\psi}(t) = \mathbf{X} \boldsymbol{\Lambda}_t$, and

$$K(t) = \boldsymbol{\alpha}^H (\boldsymbol{\Lambda}_t^H \mathbf{X}^H \mathbf{D} \mathbf{X} \boldsymbol{\Lambda}_t) \boldsymbol{\alpha} \quad (14)$$

where $\boldsymbol{\Lambda}_{N \times N}$ is a diagonal matrix with $e^{\sigma_j t}$ as elements. For brevity, we define:

$$\mathbf{A}_0 \equiv \mathbf{X}^H \mathbf{D} \mathbf{X} \quad (15)$$

$$\mathbf{A}_t \equiv \boldsymbol{\Lambda}_t^H (\mathbf{X}^H \mathbf{D} \mathbf{X}) \boldsymbol{\Lambda}_t \quad (16)$$

and note that \mathbf{A} so defined is positive definite and Hermitian. Using (13), (14), (15), (16), the auxiliary function (10) may be written in the form,

$$\varphi = \boldsymbol{\alpha}^H \mathbf{A}_t \boldsymbol{\alpha} + \lambda (\boldsymbol{\alpha}^H \mathbf{A}_0 \boldsymbol{\alpha} - 1) = \varphi(\boldsymbol{\alpha}). \quad (17)$$

As $\boldsymbol{\alpha}^H$ is not an analytic function of $\boldsymbol{\alpha}$, perhaps the simplest way to proceed at this point is to set

$$\boldsymbol{\alpha} = \boldsymbol{\alpha} + \delta \boldsymbol{\alpha}, \quad \varphi = \varphi + \delta \varphi \quad (18)$$

in (17) and search for those $\boldsymbol{\alpha}$ that make $\delta \varphi = 0$ for any $\delta \boldsymbol{\alpha}$. Substituting (18) into (17) we obtain after some algebraic manipulation the condition for $\delta \varphi = 0$:

$$(\mathbf{A}_\tau + \lambda \mathbf{A}_0) \mathbf{a} = 0. \quad (19)$$

The projection coefficients we seek are also the eigenvectors of a generalized eigenvalue problem with the Lagrange multiplier, λ , serving as the eigenvalue. We also see that this eigenvalue, λ , represents the fractional increase in kinetic energy at time τ , that is, $\lambda = K(\tau)/K(0)$.

To avoid confusion with the eigenvalues of the original problem, the set of λ arising from the solution to the eigenvalue problem (19) will be referred to as optimal values hereafter. Because of the properties of the matrix \mathbf{A} , the set of optimal initial states are orthogonal to each other under the chosen inner product.

5. Results

The methodology outlined above has been applied to the analysis of time-mean 300-mb flows derived from daily 1200 UTC global ECMWF initialized analyses in the following manner. After bicubic spline interpolation to a T42 Gaussian grid,² the 300-mb wind components were Fourier analyzed along each latitude circle. These Fourier coefficients were then used to obtain spectral forms of vorticity and divergence via Gaussian quadrature. Following Sardeshmukh and Hoskins (1984), the spectral vorticity and divergence were then spatially smoothed to retain scales larger than total wavenumber 24, and various time averages taken. These time-mean spectral vorticity and divergence coefficients are the required input to both the eigenanalysis and time-dependent barotropic models.

As an example of the application of the method, we focus on a particular season. The Northern Hemisphere winter of 1985/86 was notable for its strong low-frequency anomalous behavior. Beginning in the last week of January 1986, an episode of persistent negative height anomalies occurred in the North Pacific (Dole and Black 1990) leading to an anomalous zonal extension of the Asian jet and enhanced ridging over the west coast of North America. The flow was also highly unusual in the eastern Atlantic sector. Temperatures in England changed abruptly from mild to extremely cold between January and February, a change that was unparalleled in the previous 249 years (Ratcliffe 1986). Furthermore, the skill of ECMWF forecasts was below average during this transitional period (Hoskins and Sardeshmukh 1987).

The period 1 January–28 February 1986 was divided into several 16-day time-mean flows to study its stability. Sixteen-day means were chosen as a compromise to reduce the synoptic-scale signal in the data and yet retain the separate character of the two flow regimes

during this period. This choice of averaging interval is by no means unique, and it is of interest to examine the dependence of the analysis on this interval. Nevertheless, the results for two basic flows defined by the periods 1–16 January 1986 and 24 January–8 February

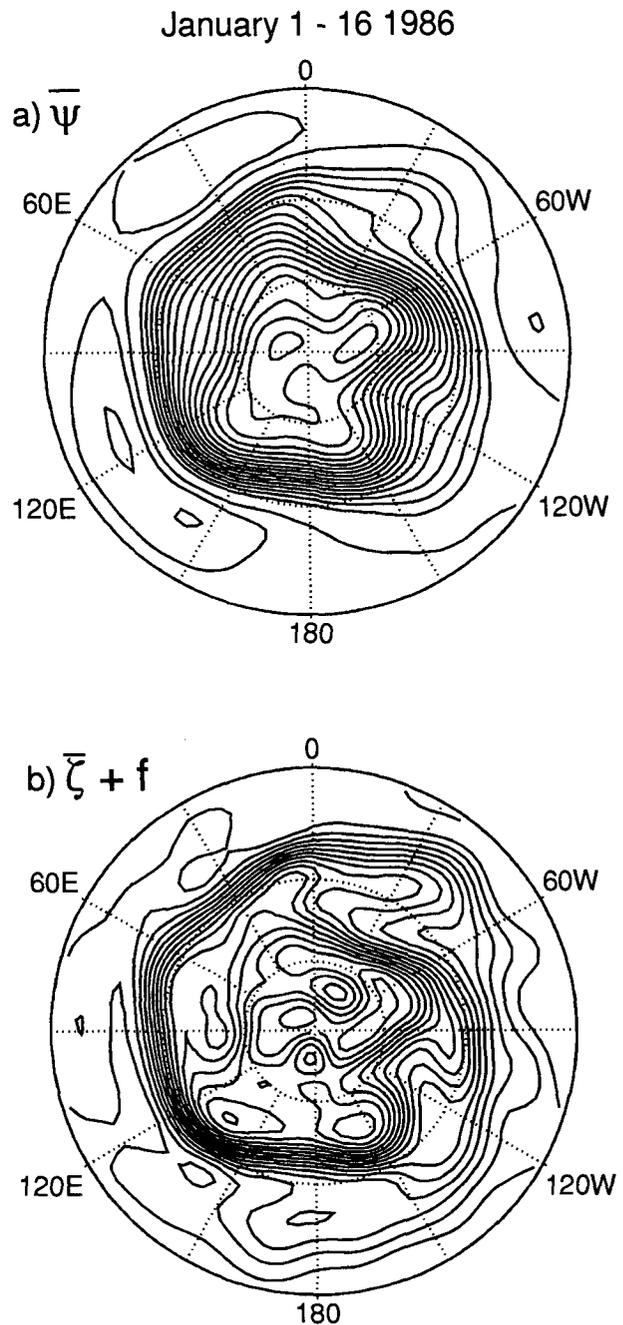


FIG. 1. Observed 300-mb mean flow for the period 1–16 January 1986: (a) streamfunction (contour interval $7.5 \times 10^6 \text{ m}^2 \text{ s}^{-1}$); (b) absolute vorticity (contour interval 10^{-5} s^{-1}). The zero contour is omitted in this and all succeeding figures. All map projections are polar stereographic, the bounding circle is the equator, and grid lines are drawn every 30 degrees.

² The T42 grid consists of 128 equally spaced longitudes and 64 Gaussian latitudes giving an effective resolution nearly equal to the original ECMWF $2.5^\circ \times 2.5^\circ$ grid.

1986 will serve as a useful illustration for the method. In the next two subsections, a brief discussion of the T31 eigenanalysis for each flow is followed by the results of the optimization problem. The linear evolution of the optimal perturbation as revealed by its kinetic energy and streamfunction is presented, followed by some results of the behavior of the optimal anomaly in a nonlinear model.

a. Mean flow for 1–16 January 1986

The first flow to be considered is the time average for the period 1–16 January 1986 (Fig. 1). There are local regions of negative absolute vorticity gradient in the Northern Hemisphere, but the Rayleigh–Kuo necessary condition for instability of the zonal mean flow is satisfied only in the Southern Hemisphere.

An eigenanalysis of the complete flow produces 2 ω unstable modes, as revealed in the plot of growth rate versus frequency shown in Fig. 2a. As noted by Branstator (1985), the shape of the eigenvalue scatter in such a plot is related to the scale-selective diffusion, with high-frequency, small-scale modes being preferentially damped. The effects of diffusion on the large-scale unstable modes, however, is negligible. Figure 2b indicates that the decaying modes have the largest projectabilities, so we should be aware that these modes may have significant spectral amplitudes in an eigenbasis expansion. The most unstable mode is stationary with an e -folding time of 4.5 days (Fig. 3a). Most of its amplitude is north of 60°N, not in the north-central

Pacific region where low-frequency anomalies subsequently develop in the latter half of the month. The adjoint (Fig. 3b) also has a global-scale structure with considerable amplitude in both the north polar region and the subtropical latitudes. This suggests that forcing (or perturbations) from subtropical latitudes may asymptotically lead to a response with the most unstable mode signature, consistent with the previous work of Simmons (1982), SWB, Branstator (1985), and more recently Ferranti et al. (1990).

To find the structure that leads to the greatest possible increase in global mean kinetic energy at a specified time, τ , we solve the optimization problem for representative values of $\tau = 1, 3, 6,$ and 20 days. The spectrum of optimal values (the fractional increase in kinetic energy at $t = \tau$) is shown in Figs. 4a,b. There are many amplifying perturbations, on the order of 150 for $\tau = 3$. This number is of course a function of τ , and must equal the number of unstable eigenmodes as $\tau \rightarrow \infty$. The number of perturbations that amplify significantly is much less, and is shown more clearly in Fig. 4b. For instance, only 22 of the possible 1023 perturbations have more than twice the energy of the eigenmode at $t = \tau = 3$ days. The optimal anomalies to the 1–16 January 1986 mean flow are able to extract energy more quickly than those for the 24 January–8 February 1986 mean flow. At $\tau = 3$ days the maximum kinetic energy growth in the latter is $\sim 20\%$ smaller than in the former.

Figure 5 shows the projection of the optimal perturbation for $\tau = 3$ days onto each of the eigenmodes.

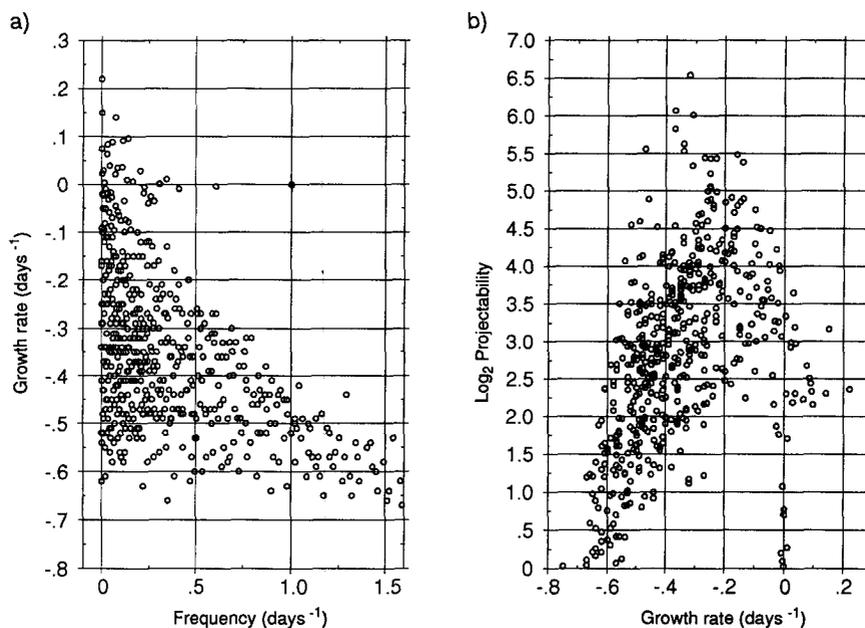


FIG. 2. The eigenvalue spectrum for the 1–16 January 1986 total mean flow: (a) growth rate vs frequency in units of days $^{-1}$; (b) \log_2 projectability vs growth rate in units of days $^{-1}$. The abscissa in (a) spans a portion of the total range of eigenfrequencies, whose highest frequency is 2.85 cycles day $^{-1}$.

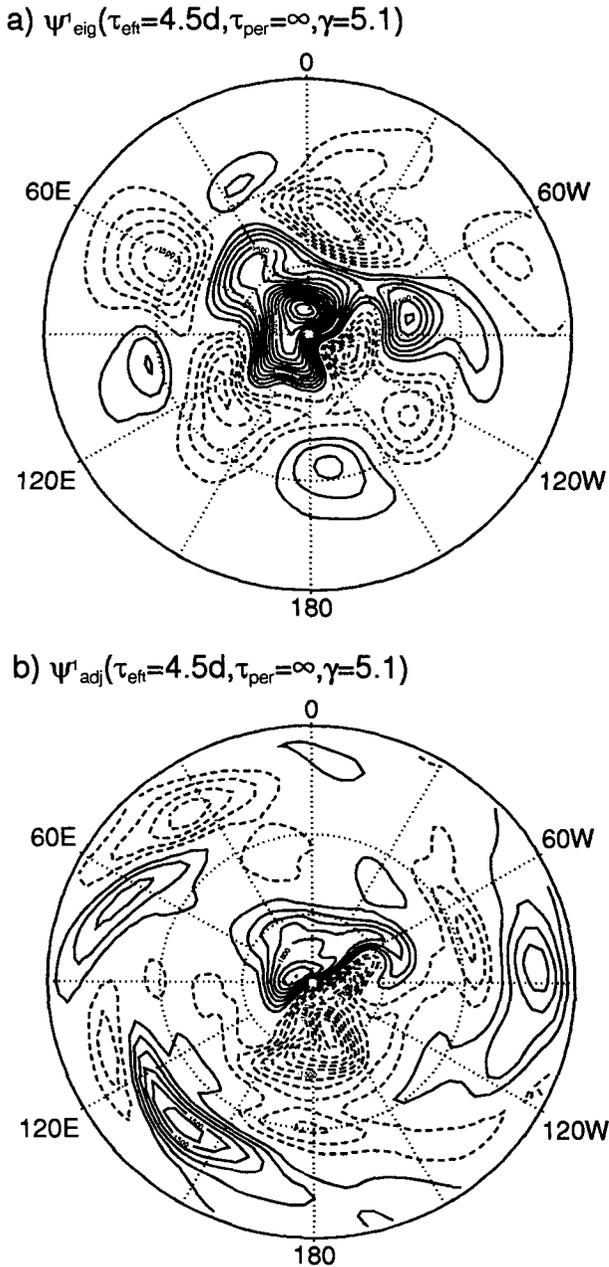


FIG. 3. The most unstable mode of the zonally asymmetric time-mean flow for 1–16 January 1986: (a) the stationary eigenmode; (b) the adjoint of the eigenmode. Contour interval $3 \times 10^5 \text{ m}^2 \text{ s}^{-1}$. Both the eigenmode and the adjoint are normalized such that the global mean kinetic energy is $0.5 \text{ m}^2 \text{ s}^{-2}$.

It is apparent that there are significant projections onto many modes. In fact, as one might anticipate from Fig. 2b, the largest projection is onto a mode that is highly damped, with a decay time scale of 2.5 days and projectability of 67. If one were to calculate the optimum based on only a subset of the eigenmodes (say, the near-neutral and unstable modes), there would be little hope of obtaining the correct structure. The nonor-

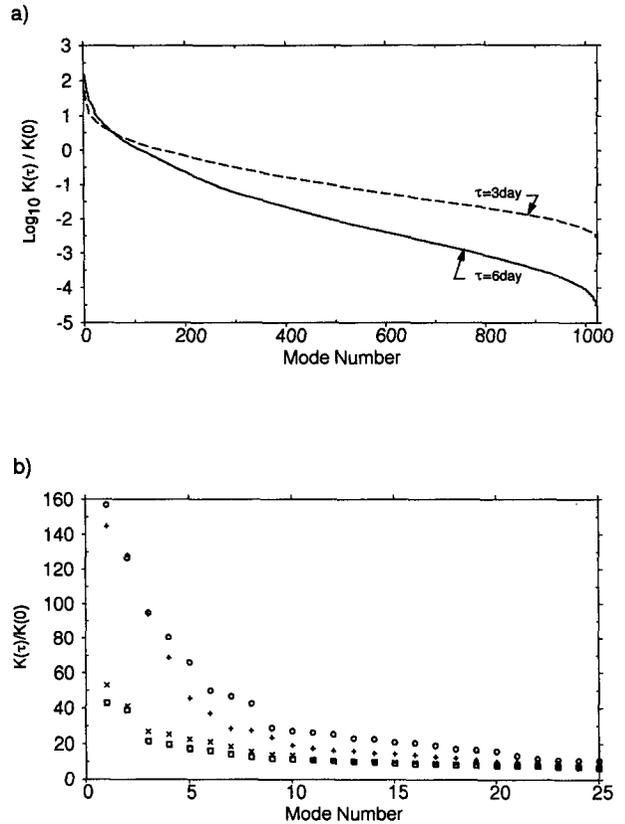


FIG. 4. (a) The “optimal value” spectrum for 1–16 January 1986 for two values of the optimizing interval, τ . The ordinate is the ratio of the kinetic energy at $t = \tau$ to the initial kinetic energy. (b) The first 25 optimal values for the 1–16 January 1986 and 24 January–8 February 1986 mean flows. The open circles and crosses denote the $\tau = 6$ and $\tau = 3$ day optima, respectively, for 1–16 January 1986. Plus symbols and squares denote the $\tau = 6$ and $\tau = 3$ day optima, respectively, for 24 January–8 February 1986.

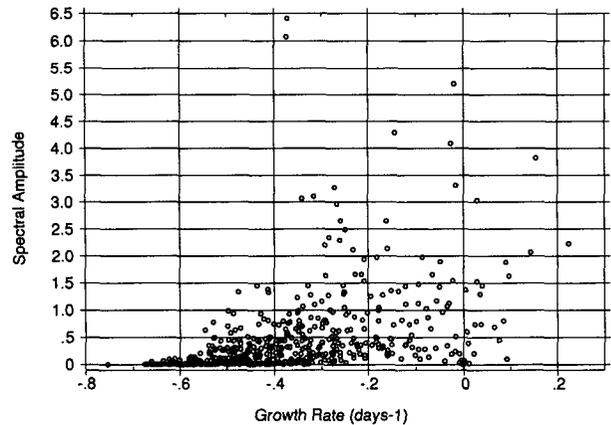


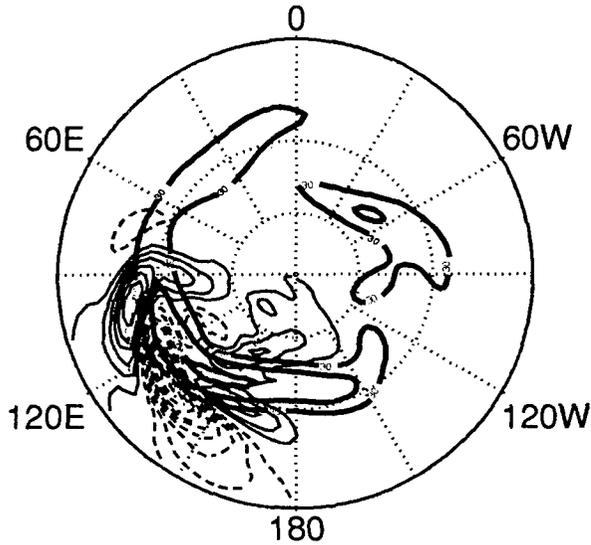
FIG. 5. The initial amplitude of the streamfunction spectral expansion coefficients for the first $\tau = 3$ day optimal perturbation to the 1–16 January 1986 mean flow as a function of the eigenmode growth rate, in units of days^{-1} . The amplitude values are normalized such that the total kinetic energy of the initial disturbance is $0.5 \text{ m}^2 \text{ s}^{-2}$.

thogonality of the system dictates that a complete set of modes be used. As the flow evolves, of course, the unstable modes increasingly dominate the solution, and

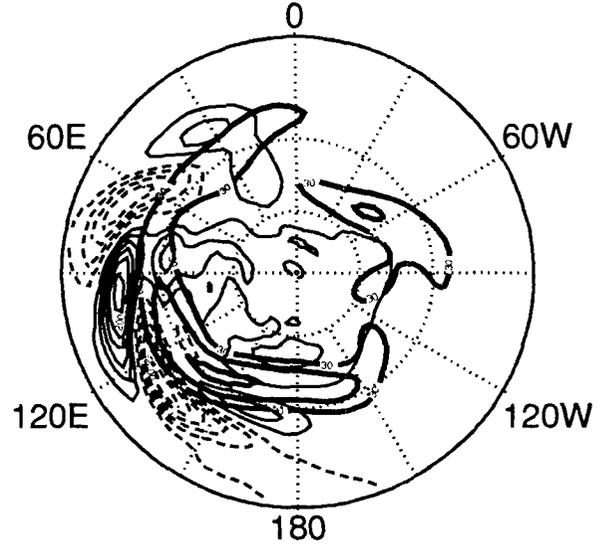
the length of this time scale determines the importance of the unstable modes in the real atmosphere.

The dependence of the optimum on the interval, τ ,

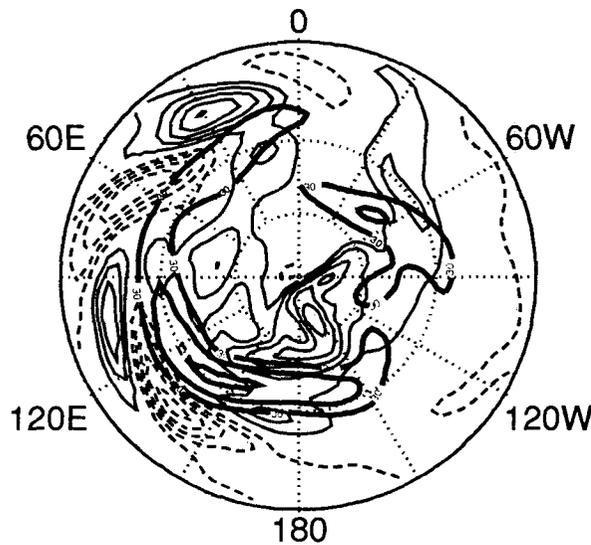
a) $\psi'(\tau=1\text{d}, K/K_0=8.3)$



b) $\psi'(\tau=3\text{d}, K/K_0=53.1)$



c) $\psi'(\tau=6\text{d}, K/K_0=157)$



d) $\psi'(\tau=20\text{d}, K/K_0=10^5)$

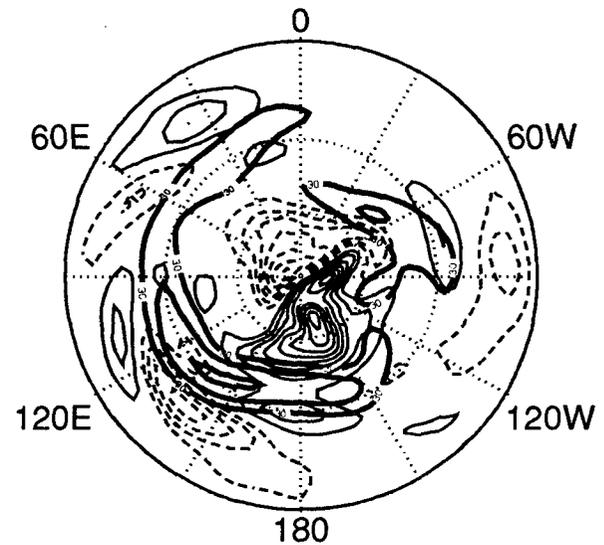


FIG. 6. Streamfunction of the first optimal perturbations as a function of the optimizing interval τ : (a) $\tau = 1$ day; (b) $\tau = 3$ days; (c) $\tau = 6$ days; (d) $\tau = 20$ days. The global mean kinetic energy associated with each pattern is $0.5 \text{ m}^2 \text{ s}^{-2}$, and the ratio of the kinetic energy at time $t = \tau$ to the initial kinetic energy is noted above each figure. Contour interval is $4 \times 10^5 \text{ m}^2 \text{ s}^{-1}$ (zero contour suppressed) throughout. The 1–16 January 1986 time-mean zonal wind, contoured every 15 m s^{-1} , beginning at 30 m s^{-1} is superposed in thick contours for reference.

is illustrated in Fig. 6. As the optimizing interval is increased it becomes more advantageous to put greater amplitude into the unstable modes, and the optimal perturbation becomes increasingly similar to the most unstable mode's adjoint. Of course, as $\tau \rightarrow \infty$ we recover this adjoint structure exactly. At $\tau = 1$ day, the optimum assumes the simple and energetically favorable form of a structure tilted against the shear of the Asian jet entrance region. As τ increases, the tilt of this main feature becomes more severe, and more structures are added upstream on the equatorward flank of the jet axis. Figure 6 indicates that the crucial feature of the perturbation for its initial growth is its tilt in the Asian jet region, a feature that is present in the adjoint structure but certainly not prominent.

The time dependence of the global mean kinetic energy for several initial states in the linear model is displayed in Fig. 7. The added complexity of the optimal perturbation when τ increases from three to six days does not become important from an energetics standpoint until after day 4. At day 3 the kinetic energy of the $\tau = 3$ day optimum initial condition is an order of magnitude larger than that of the eigenmode initial condition and over twice that of the adjoint mode initial

condition. It is not until after day 5 that the kinetic energy of the adjoint mode initial condition exceeds that of the $\tau = 3$ day optimum. The initial growth of the optimal perturbations is quasi exponential with a time scale on the order of 1.5 days. It appears that this rapid growth is associated with the shear component of the basic flow, although further analysis is warranted. Farrell (1989b) points out that properly configured disturbances in regions of diffluence and confluence may also exhibit substantial growth.

Having verified that optimal perturbations do in fact develop into the most energetic structures at a finite time, we now investigate whether these structures resemble observed anomalous features in the atmosphere. Figure 8 shows the time development of the $\tau = 3$ day optimal perturbation for a period of six days as governed by the linear equation (5). During the first three days of rapid growth the perturbation gains energy as it becomes sheared over and propagates along the jet axis. In a baroclinic model some of this early growth could be attributed to baroclinic processes. After this time the growth rate decreases and the primary feature is a dipole structure near 120°W . Integrations with the $\tau = 6$ day optimal perturbation (not shown) reveal a

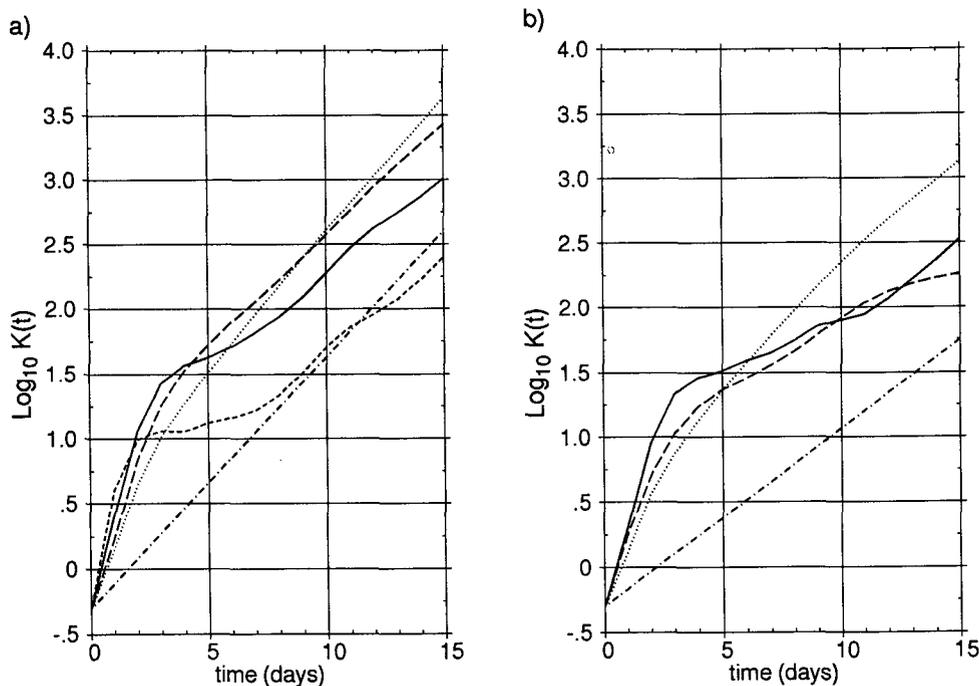


FIG. 7. The global-mean perturbation kinetic energy in the linear model as a function of time on a semilogarithmic (base 10) plot, for various initial conditions using (a) 1–16 January 1986 mean flow and (b) 24 January–8 February 1986 mean flow. The initial conditions in (a) are identified by the following line patterns: most unstable eigenmode (dot-dash); most unstable eigenmode's adjoint (dotted); first $\tau = 6$ day optimum (dash); first $\tau = 3$ day optimum (solid); first $\tau = 1$ day optimum (short dash). The initial conditions in (b) are identified by the following line patterns: most unstable eigenmode (dot-dash); most unstable eigenmode's adjoint (dotted); first $\tau = 3$ day optimum (solid); third $\tau = 3$ day optimum (dash).

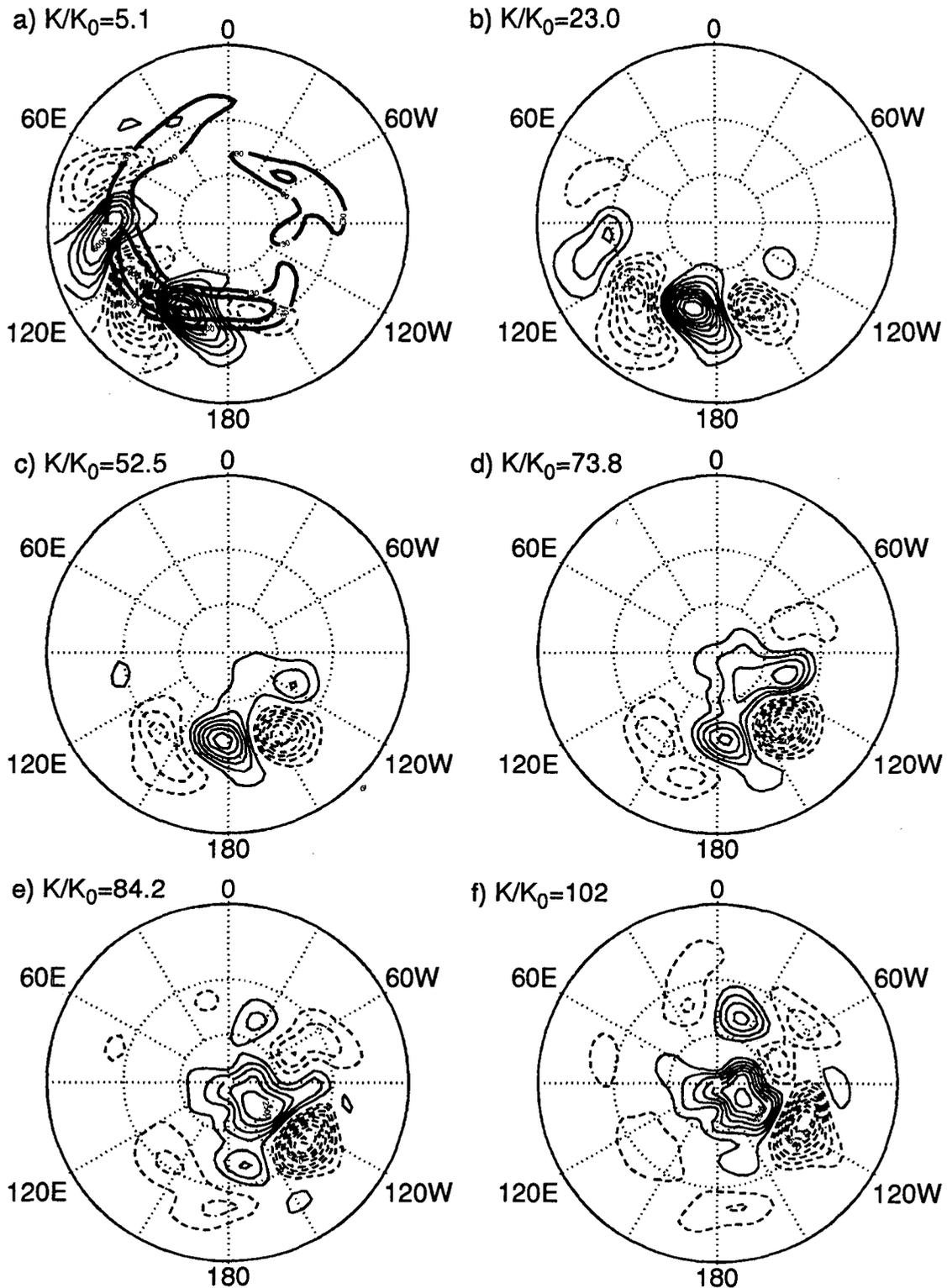


FIG. 8. (a-f) The linear evolution of the first $\tau = 3$ day optimal perturbation at one-day intervals, as represented by the perturbation streamfunction. The basic state is the 1-16 January 1986 time-mean flow. The ratio of the kinetic energy to the initial kinetic energy is noted above each figure; the contour interval is arbitrary and changes throughout. In (a) the time-mean zonal wind, contoured every 15 m s^{-1} beginning at 30 m s^{-1} , is superposed in thick contours for reference.

similar development through day 3, but afterward a more zonally elongated and persistent dipole structure near 120°W develops.

The nonlinear behavior of the optimal anomaly is summarized in Figs. 9 and 10. Figure 9a shows the departure of the global mean kinetic energy from the basic-state value (i.e., the “perturbation” in the nonlinear model) as a function of time. The initial amplitude of the perturbation is scaled to be 0.25% of the total basic flow kinetic energy ($\sim 1 \text{ m}^2 \text{ s}^{-2}$). This provides maximum zonal wind anomalies on the order of 5 m s^{-1} on the equatorward flank of the jet entrance region. Dole and Black (1990) show that the zonal wind anomalies prior to their composite onset appear to be most significant near the jet’s entrance region and its poleward flank over the Asian continent, with a composite amplitude of $5\text{--}10 \text{ m s}^{-1}$.

Some intriguing patterns developed during the first few days of the nonlinear integrations. As an example, we show the total streamfunction for the first six days using the $\tau = 3$ day optimal perturbation initial condition in the nonlinear model (Fig. 10). Although the observed zonal extension of the Asian jet is poorly captured by the model, the pronounced ridging that was

observed over the west coast of North America during the latter half of January (Dole and Black 1990) does in fact occur by day 3 and persists through day 6 with little phase propagation.

b. Mean flow for 24 January–8 February 1986

Because of the anomalous zonal extension of the Asian jet, the Northern Hemisphere mean 300-mb flow for the period 24 January–8 February 1986 (Fig. 11) satisfies the Rayleigh–Kuo necessary condition for normal-mode barotropic instability; the meridional gradient of zonal mean absolute vorticity changes sign at 50°N (not shown). On this basis, one might expect it to be more unstable than the 1–16 January 1986 mean flow.

An eigenanalysis of the complete asymmetric flow, however, shows it to be more stable to perturbations of normal-mode form than the previous flow. This suggests again that the instability associated with the zonally asymmetric portion of the flow dominates. The fastest growing of the 17 unstable modes is again stationary, but with a longer e -folding time scale of 6.3 days (Fig. 12a). Although this eigenmode bears some

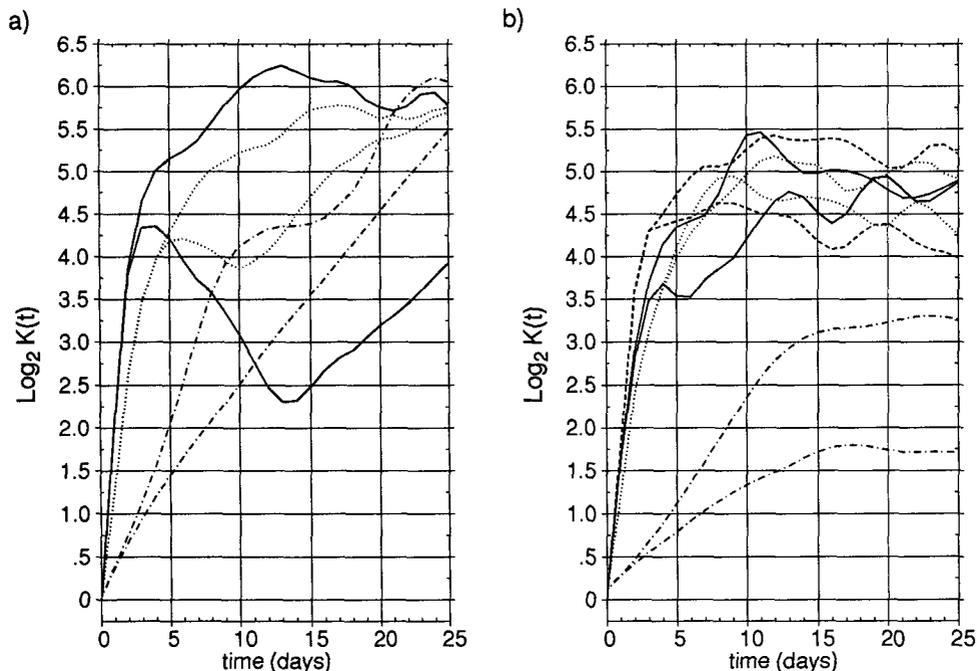


FIG. 9. The global-mean perturbation kinetic energy (defined as the departure from that of the initial basic flow) in the nonlinear model as a function of time on a semilogarithmic (base 2) plot, for various initial conditions using (a) 1–16 January 1986 initial basic flow and (b) 24 January–8 February 1986 initial basic flow. The initial conditions are constructed by adding an initial anomaly with 0.25% of the global-mean basic state kinetic energy to each basic flow. The anomalies in (a) are identified by the following line patterns: most unstable eigenmode (dot-dash); most unstable eigenmode’s adjoint (dotted); first $\tau = 3$ day optimum (solid). The anomalies in (b) are identified by the following line patterns: most unstable eigenmode (dot-dash); most unstable eigenmode’s adjoint (dotted); first $\tau = 3$ day optimum (solid); third $\tau = 3$ day optimum (long dash). The two curves plotted in each line pattern are for opposite polarity initial anomalies.

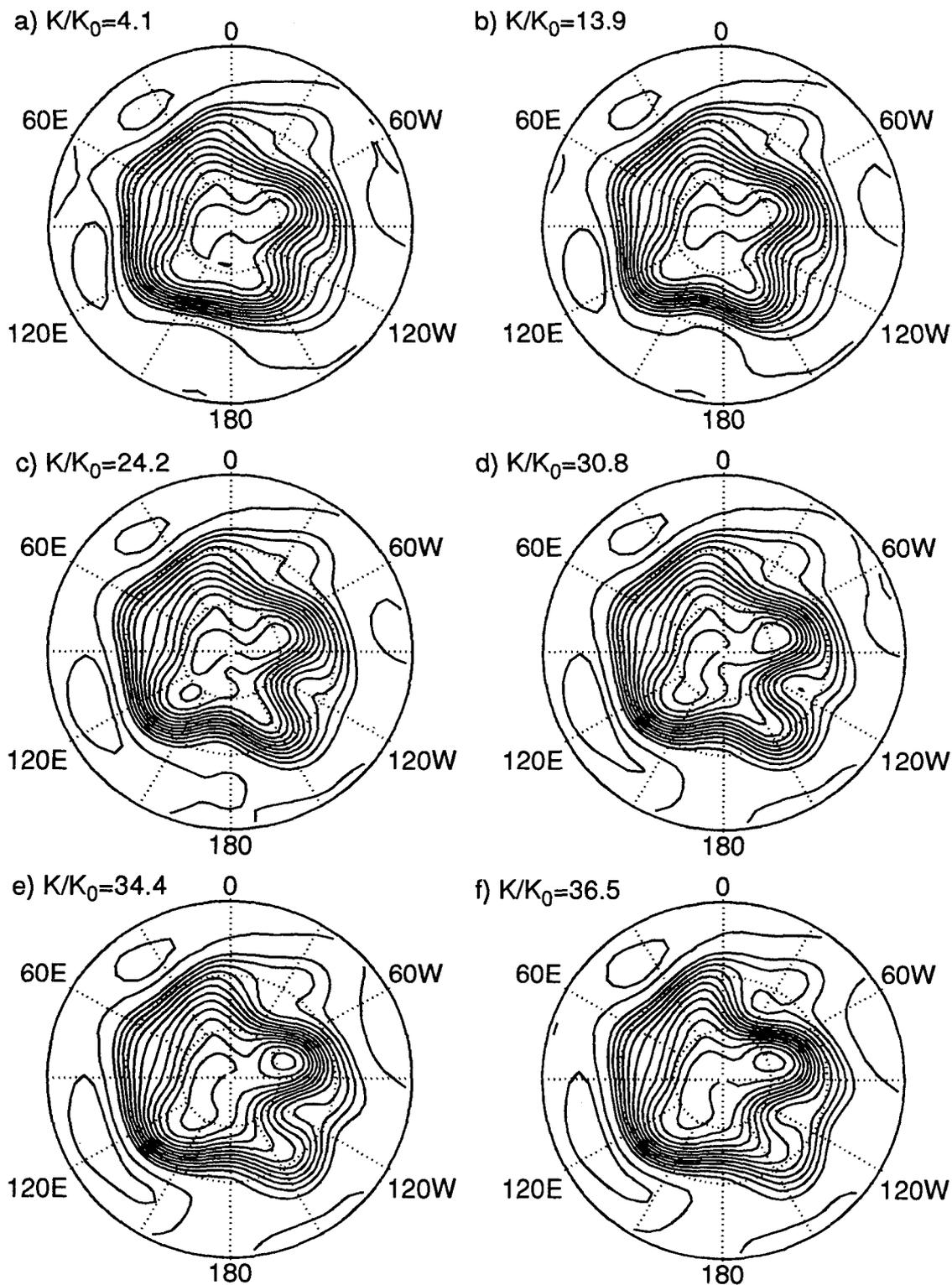


FIG. 10. (a-f) The nonlinear evolution of the first $\tau = 3$ day optimal perturbation at one-day intervals from day 1 to day 6. The initial condition is the 1-16 January 1986 time-mean flow plus the optimal perturbation with kinetic energy of 0.25% of the global-mean basic state kinetic energy. The contour interval is $7.5 \times 10^6 \text{ m}^2 \text{ s}^{-1}$.

January 24 - February 8 1986

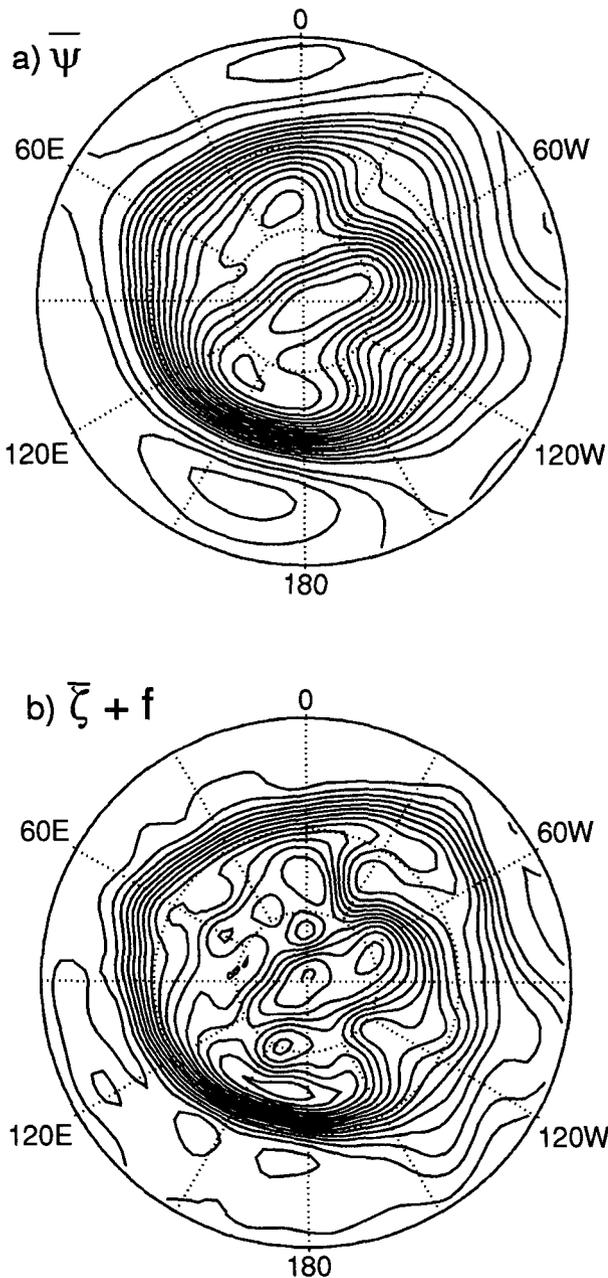


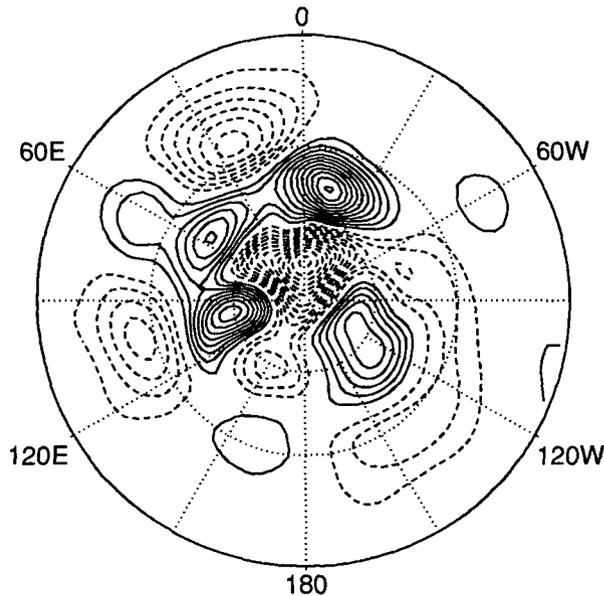
FIG. 11. As in Fig. 1 but for the period 24 January-8 February 1986.

resemblance to the streamfunction anomalies shown by Hoskins and Sardeshmukh (1987), its relevance remains unclear. The mode is not well separated from the rest of the spectrum (the second most unstable mode has an e -folding time of 6.5 days) and its projectability ($\gamma = 5.3$) is not particularly large. In addition, the adjoint (Fig. 12b) suggests that the most un-

stable normal mode is not particularly sensitive to the region bounded by $10^{\circ}\text{S}-30^{\circ}\text{N}$, $90^{\circ}\text{W}-45^{\circ}\text{W}$, an area Hoskins and Sardeshmukh (1987) suggest played a role in the development of the Atlantic block.

Compared to some of the optimal initial states, the complicated global structure of the adjoint does not result in greater perturbation kinetic energy in the linear

a) $\psi^1_{\text{eig}}(\tau_{\text{eff}}=6.3\text{d}, \tau_{\text{per}}=\infty, \gamma=5.3)$



b) $\psi^1_{\text{adj}}(\tau_{\text{eff}}=6.3\text{d}, \tau_{\text{per}}=\infty, \gamma=5.3)$

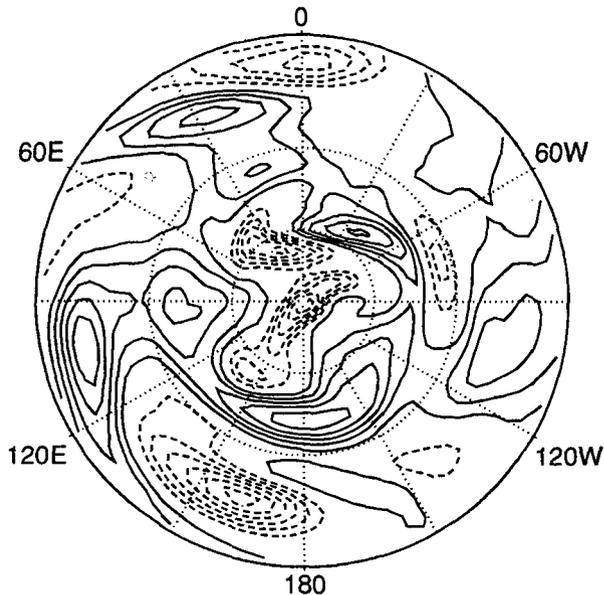


FIG. 12. As in Fig. 3 but for the period 24 January-8 February 1986.

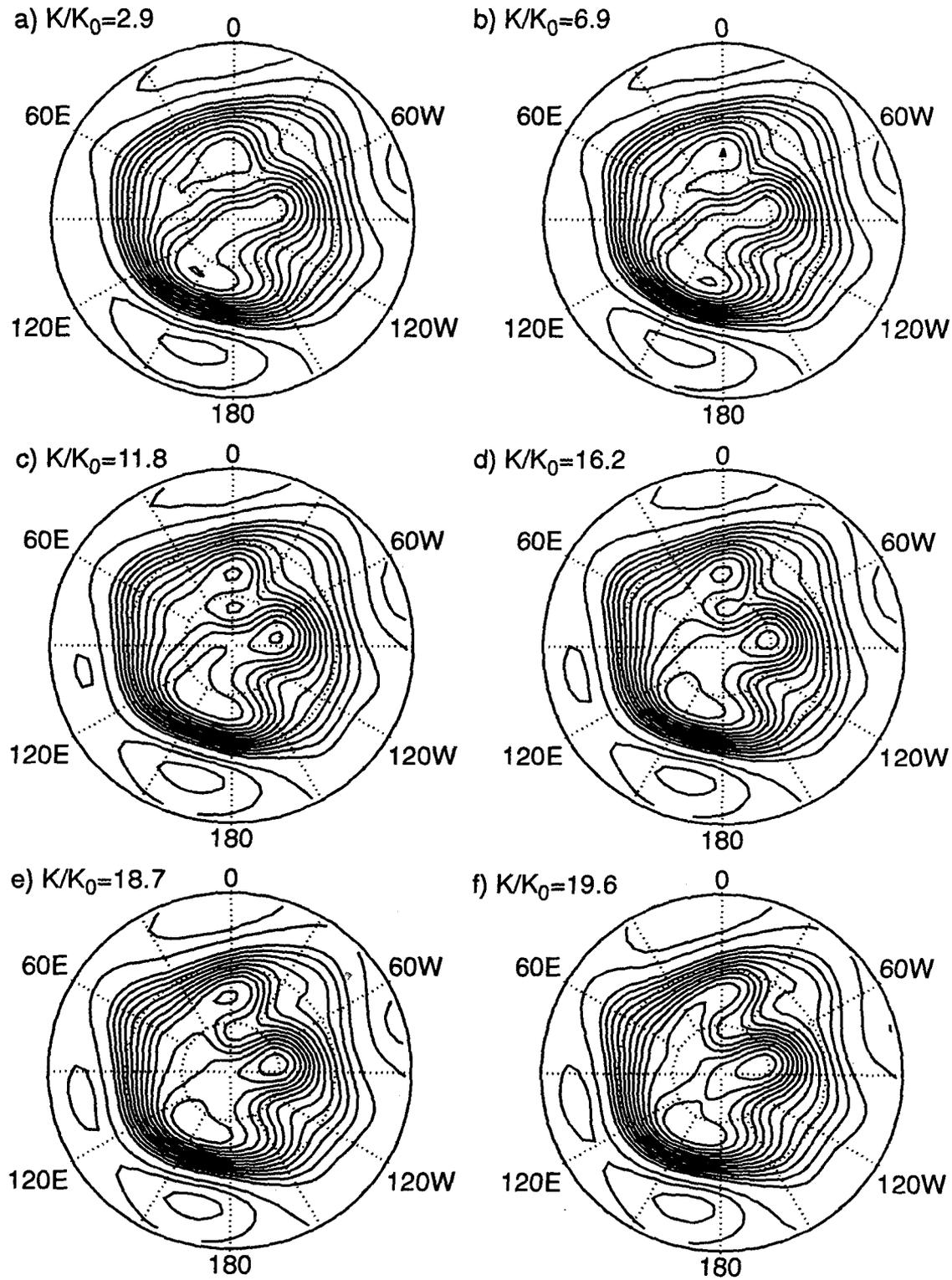


FIG. 13. As in Fig. 10 but for the third $\tau = 3$ day optimal perturbation superposed on the 24 January–8 February 1986 time-mean flow.

model until after day 5 (Fig. 7b). From Fig. 9b we see, however, that nonlinear effects cannot be completely ignored beyond day 5. Furthermore, the adjoint and first three $\tau = 3$ day optimal initial states reach comparable global-mean perturbation kinetic-energy finite amplitudes of $32 \text{ m}^2 \text{ s}^{-2}$, while the eigenmode initial

condition integrations reach one third to one half of these values. Even though the optimization problem indicates the third $\tau = 3$ day optimum can extract only half the energy the first optimum can (Fig. 4b), this third optimal initial state seems equally important in the nonlinear model.

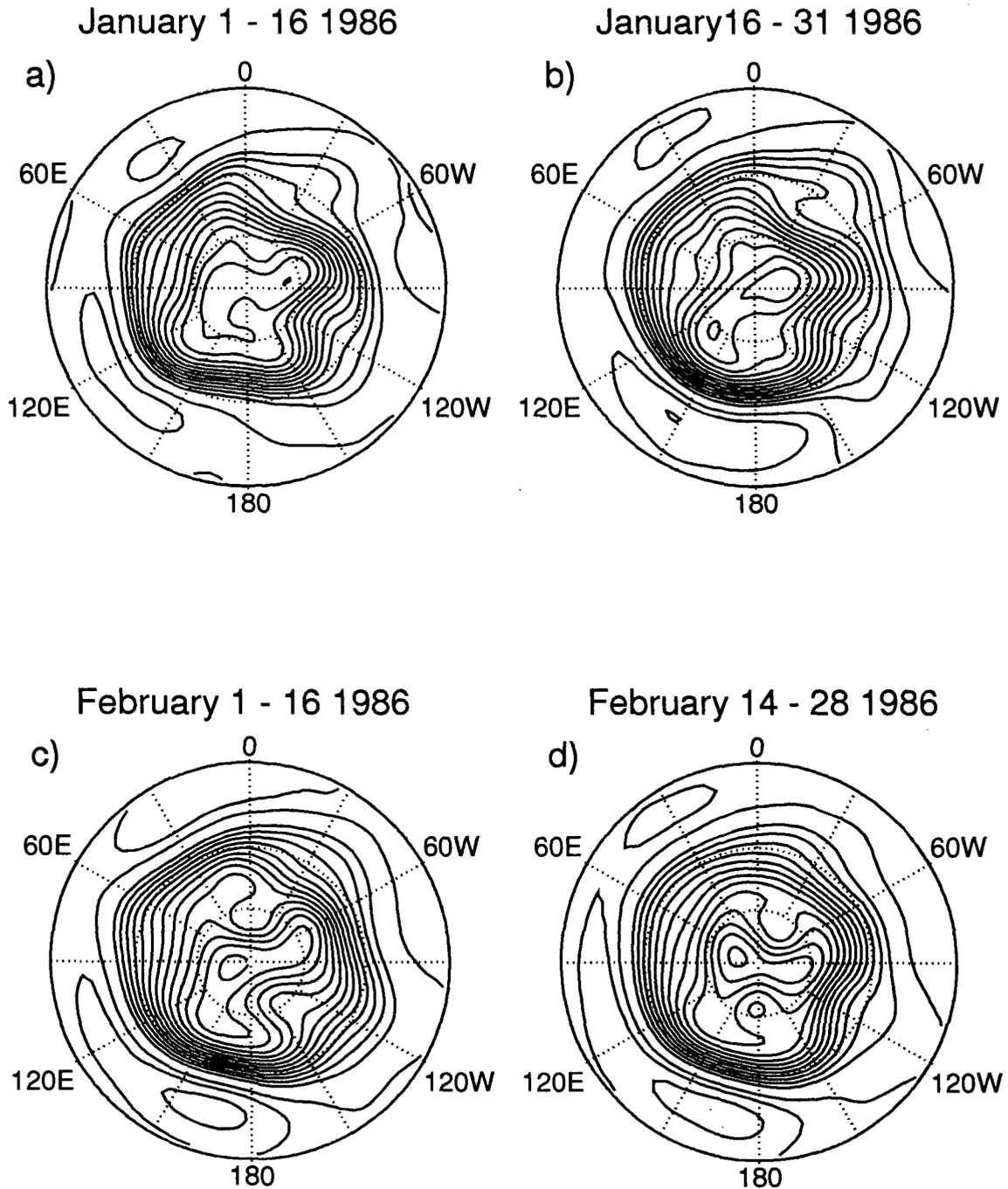


FIG. 14. ECMWF analyzed mean 300-mb flows for the period 1 January–28 February 1986: (a) 1–16 January; (b) 17–31 January; (c) 1–16 February; (3) 14–28 February. Contour interval $10^7 \text{ m}^2 \text{ s}^{-1}$.

In contrast to the adjoint structure, the optimal initial states show a more localized character. The first two optimal initial states for $\tau = 3$ days are similar to the corresponding 1–16 January 1986 first optimum; most of their amplitude is concentrated in subtropical latitudes south of the Asian jet. The subsequent linear evolution in both cases, however, is characterized more by wave propagation than the quasi-stationary dipole feature found using the 1–16 January 1986 basic flow (not shown).

The third $\tau = 3$ -day optimal initial state emphasizes the region north of 30°N , in particular the poleward flank of the North American jet exit region near 60°N , 30°W . It is similar to the mid- and high-latitude structure of the adjoint, but with virtually zero amplitude south of 30°N . The nonlinear evolution of this initial state is shown in Fig. 13. During the early stages of rapid initial growth (through day 6) the streamfunction field develops a blocking ridge in the North Atlantic ($\sim 60^\circ\text{N}$, 30°W) as well as an enhanced ridge along the 130°W meridian. Comparison to the ECMWF analyzed 300-mb streamfunction pattern averaged over the first half of February 1986 (Fig. 14c) shows close agreement of position with both of these features. Neither of these anomalous ridging patterns is as closely duplicated in the integrations using the adjoint and first two optimal perturbations as initial states.

6. Summary and conclusions

We have studied the barotropic stability of observed 300-mb nonzonal flows during the winter of 1985/86, a period associated with strong low-frequency variability. Rather than exclusively seeking exponentially amplifying and/or periodic solutions, we allow unrestricted time dependence and search for the disturbance that grows most rapidly over a finite time interval. This disturbance is found following Farrell (1988, 1989a), using the fact that the eigenfunctions from a traditional stability analysis may form a complete basis set with which to expand an arbitrary disturbance. The obvious advantage of this basis set is that the energy of the disturbance at any time is easily obtained after the spectral expansion coefficients are known, so that an optimization problem can be formulated. Conceptually, this method of superposing normal-mode solutions is equivalent to a Laplace (or Fourier) transform method (Tung 1983).

From the solution of a global optimization problem, we have determined the structure of optimal perturbations that extract the greatest amount of globally integrated energy from nonzonal flows over a specified time interval. These optimal perturbations were found to assume the energetically favorable position of being tilted against the strongest shear of the flow, and became rather localized as the optimizing interval was decreased. Only 2% of the total number of $\tau = 3$ day

optimal perturbations extract at least twice the energy of the most unstable eigenmode at $t = \tau$.

Consistent with previous studies of simpler basic flows (Farrell 1989a; Mak and Cai 1989), the initial growth of the optimum greatly exceeded that of the most unstable eigenmode, and was also significantly greater than an adjoint mode initial condition. Moreover, the nonlinear evolution of the optimal perturbations suggests that they are at least as important for stability considerations as the adjoint and eigenmode structures.

The nonlinear evolution of the 3-day optimal anomaly to the 1–16 January 1986 mean flow leads to the development of strong ridging over the west coast of North America. This development is apparently associated with the fastest way for a perturbation to gain energy, and is in reasonable agreement with the observed behavior. Preliminary sensitivity studies indicate this behavior remains when the optimal perturbation is longitudinally shifted 10° or less. The time evolution of the first two optimal anomalies to the 24 January–8 February 1986 mean flow did not exhibit a quasi-stationary development. For this flow it appears the third $\tau = 3$ day optimal initial condition leads to a development that most closely mimics the observed pattern.

Palmer (1988) has recently studied the predictive skill of the ECMWF forecast model, and found evidence that it varies with the extratropical flow regime as categorized by the PNA index defined by Wallace and Gutzler (1981). Paradoxically, the most barotropically stable regime was found to be the least predictable. Zhang (1988) has suggested that this paradox may be resolved by remembering that the projectability is an important quantity to consider in nonself-adjoint systems, since it essentially determines how easily a mode can be excited. He shows that the most unstable mode has a larger projectability in the more stable flow regime, and hence the projection of an arbitrary field onto the most unstable mode is likely to be greater in this stabler regime. He argues that the initial behavior of perturbations to a basic flow can thus depend more on the projectability than on the exponential growth rate.

Farrell (1990) has argued that exponential instability is not even required to give the flow greater unpredictability. He postulates that the error growth may be related to the spectrum of optimal perturbations, and that an assessment of this error can be made for an ensemble of perturbations once this spectrum is known. Our particular application of the optimization problem may be used to directly examine Farrell's hypothesis and its implications for predictability, as well as its use for the general stability analysis addressed here.

Acknowledgments. This research was supported by the Climate Dynamics Program, Atmospheric Sciences

Division of the National Science Foundation under Grants ATM 86-20077 and ATM 90-06123. Computational support was supplied by the San Diego Supercomputer Center. M.D.B. is grateful to the members of the Dynamics Reading Group at the University of Washington, in particular E. Sarachik and J. Barsugli for many stimulating discussions. J. L. Anderson and an anonymous reviewer provided helpful comments that improved the original manuscript.

APPENDIX

Eigenanalysis of Nonself-adjoint Systems

Consider the archetypal ordinary differential equation,

$$\frac{d\psi}{dt} - \mathcal{L}\psi = 0 \quad (\text{A1})$$

where \mathcal{L} denotes a general linear operator that can be represented as a matrix, \mathbf{L} , of order N . The corresponding eigenvalue problem,

$$\mathbf{L}\mathbf{x}_j = \sigma_j\mathbf{x}_j \quad (\text{A2a})$$

will provide solutions to the discretized version of (A1) in the form, $\psi = x_j e^{\sigma_j t}$. Now consider the additional matrix eigenvalue problem:

$$\mathbf{M}\mathbf{y}_k = \lambda_k\mathbf{y}_k \quad (\text{A2b})$$

where \mathbf{M} is yet to be specified. It is easy to show that if \mathbf{x} is a vector of spherical harmonic coefficients, then the matrix representation of the inner (scalar) product may be written

$$\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^H \mathbf{D} \mathbf{x} \quad (\text{A3})$$

where $\langle \cdot, \cdot \rangle$ denotes a suitable inner product, \mathbf{D} is a real positive-definite diagonal matrix that defines its matrix equivalent, and $(\)^H$ denotes the Hermitian (complex conjugate transpose). For instance, setting $\mathbf{D} = \mathbf{I}$, the identity matrix, yields the squared-amplitude inner product. Another choice of \mathbf{D} will return a kinetic energy quantity. If we multiply (A2a, b) by the other's eigenvector, take the inner product, and subtract the resulting equations we obtain

$$y_k^H (\mathbf{D}\mathbf{L} - \mathbf{M}^H \mathbf{D}) \mathbf{x}_j = (\sigma_j - \lambda_k^*) y_k^H \mathbf{D} \mathbf{x}_j \quad (\text{A4})$$

where $*$ denotes the complex conjugate. If we identify \mathbf{M} as the adjoint matrix to \mathbf{L} we will be led to a biorthogonality relation, since by definition the adjoint operator satisfies

$$\langle \varphi, \mathcal{L}\psi \rangle = \langle \mathcal{M}\varphi, \psi \rangle \quad (\text{A5})$$

or in matrix notation,

$$y^H \mathbf{D} \mathbf{L} \mathbf{x} = (\mathbf{M} \mathbf{y})^H \mathbf{D} \mathbf{x} = \mathbf{y}^H \mathbf{M}^H \mathbf{D} \mathbf{x} \rightarrow \mathbf{D} \mathbf{L} = \mathbf{M}^H \mathbf{D}. \quad (\text{A6})$$

Since the eigenvalues of \mathbf{M} are the same as those for

\mathbf{L} (Friedman 1956, pp. 199–200), we can obtain the biorthogonality relation from (A3), (A4), and (A5): the adjoint vectors, \mathbf{y} , are orthogonal to the eigenvectors, \mathbf{x} , that is, $\langle \mathbf{y}_k, \mathbf{x}_j \rangle = 0$ if $\sigma_j \neq \lambda_k^*$. For the squared-amplitude inner product $\mathbf{M} = \mathbf{L}^H$ so that if \mathbf{L} is Hermitian the operator is self adjoint, and the eigenvectors are mutually orthogonal. Inhomogeneities in the flow generally destroy this symmetry in \mathbf{L} and one must consider the adjoint problem to determine the spectral representation of a field in the space spanned by the eigenvectors. For a nonself-adjoint matrix, \mathbf{L} , our normalization of the eigenvectors and adjoint vectors follows Zhang (1988) so that the projectability, γ_j , belongs to the set of real numbers,

$$\gamma_j \langle \mathbf{y}_k, \mathbf{x}_j \rangle = \delta_{jk} \quad (\text{A7})$$

$$\langle \mathbf{x}_j, \mathbf{x}_j \rangle = 1 \quad (\text{A8})$$

where $\delta_{jk} = 1$ if $j = k$, 0 otherwise.

REFERENCES

- Andrews, D. G., 1984: On the stability of forced non-zonal flows. *Quart. J. Roy. Meteor. Soc.*, **110**, 657–662.
- Bourke, W., 1972: An efficient, one-level, primitive-equation spectral model. *Mon. Wea. Rev.*, **100**, 683–689.
- Boyd, J. P., 1983: The continuous spectrum of linear Couette flow with the beta effect. *J. Atmos. Sci.*, **40**, 2304–2308.
- Branstator, G., 1982: Horizontal energy propagation and barotropic instability in a quasi-linear model of the atmosphere. Ph.D. thesis, University of Washington.
- , 1985: Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. Part II: Eigenanalysis. *J. Atmos. Sci.*, **42**, 2242–2254.
- Case, K. M., 1960: Stability of inviscid plane Couette flow. *Phys. Fluids*, **3**, 143–148.
- Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, **4**, 135–162.
- Dole, R. M., 1986: The life cycles of persistent anomalies and blocking over the North Pacific. *Adv. Geophys.*, **29**, 31–69.
- , and N. D. Gordon, 1983: Persistent anomalies of the extratropical Northern Hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. *Mon. Wea. Rev.*, **111**, 1567–1586.
- , and R. X. Black, 1990: Life cycles of persistent anomalies. Part II: The development of persistent negative height anomalies over the North Pacific ocean. *Mon. Wea. Rev.*, **118**, 824–846.
- Durran, D. R., 1991: The third-order Adams–Bashforth method: An attractive alternative to leapfrog time differencing. *Mon. Wea. Rev.*, **119**, 702–720.
- Eady, E. T., 1949: Long waves and cyclone waves. *Tellus*, **1**, 33–52.
- Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663–1686.
- , 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45**, 163–172.
- , 1989a: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46**, 1193–1206.
- , 1989b: Transient development in confluent and diffluent flow. *J. Atmos. Sci.*, **46**, 3271–3288.
- , 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, **47**, 2409–2416.
- Ferranti, L., T. N. Palmer, F. Molteni, and E. Klinker, 1990: Tropical–extratropical interaction associated with the 30–60 day oscillation and its impact on medium and extended range prediction. *J. Atmos. Sci.*, **47**, 2177–2199.

- Fjortoft, R., 1953: On the changes in spectral distribution of kinetic energy for two-dimensional non-divergent flow. *Tellus*, **5**, 225–230.
- Frederiksen, J. S., 1982: A unified three-dimensional theory of the onset of blocking and cyclogenesis. *J. Atmos. Sci.*, **39**, 967–987.
- , 1983: A unified three-dimensional theory of the onset of blocking and cyclogenesis. II. Teleconnection patterns. *J. Atmos. Sci.*, **40**, 2593–2609.
- , 1989: The role of instability during the onset of blocking and cyclogenesis in Northern Hemisphere synoptic flows. *J. Atmos. Sci.*, **46**, 1076–1092.
- , and P. J. Webster, 1988: Alternative theories of atmospheric teleconnections and low-frequency fluctuations, **26**, 459–494.
- , and R. C. Bell, 1990: North Atlantic blocking during January 1979: Linear theory. *Quart. J. Roy. Meteor. Soc.*, **116**, 1289–1313.
- Gall, R., 1976: Structural changes of growing baroclinic waves. *J. Atmos. Sci.*, **33**, 374–390.
- Haines, K., and J. Marshall, 1987: Eddy-forced coherent structures as a prototype for atmospheric blocking. *Quart. J. Roy. Meteor. Soc.*, **113**, 681–704.
- Hendon, H. H., and D. L. Hartmann, 1985: Variability in a nonlinear model of the atmosphere with zonally symmetric forcing. *J. Atmos. Sci.*, **42**, 2783–2797.
- Hildebrand, F. B., 1963: *Advanced Calculus for Engineers*. Prentice-Hall, 646 pp.
- Holopainen, E. O., and C. Fortelius, 1987: High frequency transient eddies and blocking. *J. Atmos. Sci.*, **44**, 1632–1645.
- Hoskins, B. J., and P. D. Sardeshmukh, 1987: A diagnostic study of the dynamics of the northern hemisphere winter of 1985–1986. *Quart. J. Roy. Meteor. Soc.*, **113**, 759–778.
- , A. J. Simmons, and D. G. Andrews, 1977: Energy dispersion in a barotropic atmosphere, **103**, 553–567.
- Illari, L., and J. C. Marshall, 1983: On the interpretation of eddy fluxes during a blocking episode. *J. Atmos. Sci.*, **40**, 2232–2242.
- Kasahara, A., 1980: Effect of zonal flows on the free oscillations of a barotropic atmosphere. *J. Atmos. Sci.*, **37**, 917–929.
- Kuo, H. L., 1949: Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Meteor.*, **6**, 105–122.
- Lacarra, J., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81–95.
- Lorenz, E. O., 1972: Barotropic instability of Rossby wave motion. *J. Atmos. Sci.*, **29**, 258–264.
- Mak, M., and M. Cai, 1989: Local barotropic instability. *J. Atmos. Sci.*, **46**, 3289–3311.
- Nakamura, H., and J. M. Wallace, 1990: Observed changes in baroclinic wave activity during the life cycles of low-frequency circulation anomalies. *J. Atmos. Sci.*, **47**, 1100–1116.
- Orr, W. M., 1907: The stability or instability of the steady motions of a liquid. *Proc. R. Irish Acad.*, **A27**, 69–138.
- Palmer, T. N., 1988: Medium and extended range predictability and stability of the Pacific/North American mode. *Quart. J. Roy. Meteor. Soc.*, **114**, 691–713.
- Ratcliffe, R. A. S., 1986: Review of winter 1985/86 over the Northern Hemisphere. *Weather*, **41**, 170–172.
- Rex, D. F., 1950a: Blocking action in the middle troposphere and its effects on regional climate. I. An aerological study of blocking. *Tellus*, **2**, 196–211.
- , 1950b: Blocking action in the middle troposphere and its effects on regional climate. II. The climatology of blocking action. *Tellus*, **2**, 275–301.
- Rhines, P. B., 1975: Waves and turbulence on a beta-plane. *J. Fluid Mech.*, **69**, 417–443.
- Sardeshmukh, P. D., and I. M. Held, 1984: The vorticity balance in the tropical upper troposphere of a general circulation model. *J. Atmos. Sci.*, **41**, 768–778.
- , and B. J. Hoskins, 1984: Spectral smoothing on the sphere. *Mon. Wea. Rev.*, **112**, 2524–2529.
- , and —, 1985: Vorticity balances in the tropics during the 1982–83 El Niño–Southern Oscillation event. *Quart. J. Roy. Meteor. Soc.*, **111**, 261–278.
- , and —, 1987: On the derivation of the divergent flow from the rotational flow: The χ problem. *Quart. J. Roy. Meteor. Soc.*, **113**, 339–360.
- , and —, 1988: On the generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228–1251.
- Simmons, A. J., 1982: The forcing of stationary wave motion by tropical diabatic heating. *Quart. J. Roy. Meteor. Soc.*, **108**, 503–534.
- , and B. J. Hoskins, 1980: Barotropic influences on the growth and decay of nonlinear baroclinic waves. *J. Atmos. Sci.*, **37**, 1679–1684.
- , J. M. Wallace, and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.*, **40**, 1363–1392.
- Tung, K. K., 1983: Initial value problems for Rossby waves in a shear flow with critical level. *J. Fluid Mech.*, **133**, 443–469.
- Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784–812.
- Yanai, M., and T. Nitta, 1968: Finite difference approximations for the barotropic instability problem. *J. Meteor. Soc. Japan*, **46**, 389–403.
- Zhang, Z., 1988: The linear study of zonally asymmetric barotropic flows. Ph.D. thesis, University of Reading.