BAROCLINIC INSTABILITY (MS 076)

Richard Grotjahn

Department of Land, Air, and Water Resources,

One Shields Ave

University of California,

Davis, California 95616-8627, USA

grotjahn@ucdavis.edu

5307522246 (V)

5307521793 (F)

Abstract/Synopsis

Baroclinic instability refers to a process by which perturbations draw energy from the mean flow potential energy. Conditions in the middle latitude atmosphere are conducive for zonally-varying structures (eddies) to grow by this process. The baroclinic energy conversions are proportional to eddy heat fluxes, these eddies also accomplish some of the necessary poleward transport of heat, especially in middle latitudes. Baroclinically unstable solutions arising in simple linear quasi-geostrophic formulations have properties similar to observed frontal cyclones. Improving those simplifying assumptions (such as allowing nonlinearity) improve the similarity between simulated and observed properties.

1.1 INTRODUCTION

Baroclinic instability draws energy from the portion of the potential energy available to be converted (referred to as 'available potential energy' or APE) is dependent upon a horizontal gradient of temperature. The conversions of energy are proportional to perturbation heat fluxes in the horizontal and vertical. From thermal wind balance, a horizontal temperature gradient implies the presence of vertical shear. So, baroclinic instability is also an instability of the vertical shear.

Another view of baroclinic instability emphasizes interacting potential vorticity (PV) anomalies. Baroclinic instability is often studied by linearizing the dynamics equations and using eigenvalue or initial value techniques. These alternative views and analysis procedures generally provide complementary means to understand better baroclinic instability.

The atmosphere requires heat fluxes to maintain the observed pattern of net radiation (positive in the tropics, negative poleward of 38 degrees on an annual average). A zonal mean meridional circulation, such as a tropical Hadley cell, can generate these heat fluxes. The poleward moving air in the Hadley cell accelerates while conserving angular momentum. In contrast, lower tropospheric air is slower-moving. Hence, vertical shear builds towards the poleward edge of each Hadley cell. In middle latitudes, baroclinic instability provides a mechanism to explain how the eddies form and evolve whilst including and accomplishing the necessary heat fluxes. Theoretical models of baroclinic instability can simulate various observed properties of midlatitude eddies including: the dominant length scales, propagation speed, vertical structure, and energetics.

1.2 AN ILLUSTRATIVE MODEL

An illustrative model provides mathematical relations and archetype solutions for the concepts that follow.

1.2.1 Mathematical formulation

The model uses quasi-geostrophic (QG) approximations and nondimensional scaling appropriate for midlatitude frontal cyclones. PV has contributions from the interior and from temperature gradients at rigid bottom (z=0) and top ($z=Z_T$) boundaries. PV in the QG system can be written:

$$q=\underbrace{\underbrace{\nabla^2\psi}_{RV}+f_o+\beta y}_{AV}+\underbrace{\gamma\frac{\partial\psi}{\partial z}+\varepsilon\frac{\partial^2\psi}{\partial z^2}}_{TV}+\underbrace{\left\{\varepsilon\left.\frac{\partial\psi}{\partial z}\right|_{z=0}-\varepsilon\left.\frac{\partial\psi}{\partial z}\right|_{z=Z_T}\right\}}_{BPV} \qquad (1)$$

Where

$$\varepsilon=\frac{f_0{}^2 L}{g\,\kappa\,D},\quad \kappa=D\frac{\partial\ln\theta_s}{\partial z},\quad and \quad \gamma=\frac{\varepsilon}{\rho}\frac{\partial\rho}{\partial z}+\frac{\partial\varepsilon}{\partial z} \qquad (2)$$

ψ is the horizontal velocity streamfunction, □is density, g is the acceleration of gravity, □ is the static stability from the horizontal mean potential temperature. The coordinates are: x eastward, y northward, and z upward. Nondimensional length scales are L in the horizontal and D in the vertical. f_o is the constant part while □is the meridional derivative (approximated as a constant) of the Coriolis parameter.

An inherent horizontal length scale is the Rossby radius of deformation ($L_R = N H f_0^{-1}$) where N is the Brunt Väisälä frequency ($N^2 = g\square D^{-1}$) and $H = RTg^{-1}$ is the scale height (an inherent vertical length scale). Thus, $\square= (LH)^2(L_RD)^{-2}$ relates the assumed scales L and H to L_R and D.

QGPV includes three distinct parts: absolute vorticity (AV) which includes relative vorticity (RV), “thermal” vorticity (TV), and boundary PV (BPV). Positive PV is associated with an interior trough (in geopotential) and/or a warm surface (i.e. boundary) temperature anomaly.

When the vorticity and potential temperature conservation equations are combined, one obtains a time dependent equation for QGPV conservation:

$$\left[\frac{\partial}{\partial t}+U\frac{\partial}{\partial x}\right]\left(\nabla^2\Psi+\frac{1}{\rho}\frac{\partial}{\partial z}\left(\rho\varepsilon\frac{\partial\Psi}{\partial z}\right)\right)+\frac{\partial Q}{\partial y}\frac{\partial\Psi}{\partial x}=0 \qquad \text{(3a)}$$

with boundary conditions at the bottom and top

$$\left[\frac{\partial}{\partial t}+U\frac{\partial}{\partial x}\right]\left(\frac{\partial\Psi}{\partial z}\right)-\frac{\partial U}{\partial z}\frac{\partial\Psi}{\partial x}=0 \ \text{ at } z = 0, Z_T \qquad \text{(3b)}$$

"Basic state" variables are specified: U (independent of x) is zonal wind, and Q is the interior part of the QGPV; meridional and vertical velocities are zero. One can solve (3) as an initial value problem by specifying an initial streamfunction or potential vorticity.

An eigenvalue problem can be formulated from (3). A common approach assumes time and space dependence in this form:

$$\psi(x,y,z,t)=\mathrm{Re}\{\phi(y,z)\exp(ik(x-ct))\} \qquad \text{(4)}$$

for the "perturbation" streamfunction being sought. This solution has zonal wavenumber k and complex phase speed c. The growth rate is given by $k\mathrm{Im}\{c\}$. If U has no meridional variation, then one can assume a wave-like y dependence too: $\exp(ily)$. When meridional wavenumber l equals zonal wavenumber k, the solution is a "square wave". Perturbation velocities are defined as $u = -\partial\psi/\partial y$ and $v = \partial\psi/\partial x$.

Additional simplifying approximations are often made. A particularly simple form, commonly labelled the “Eady model”, was described by E.T. Eady in 1949. The Eady model assumes wavelike meridional structure, $\partial Q/\partial y = 0$, $U = z$, incompressibility (ρ = constant), and $\varepsilon = 1$. Then (3a) is reduced simply to solving $q' = 0$ in the interior where the prime denotes the “perturbation” sought. The Eady eigenvalue problem can be solved analytically, yielding a pair of “normal modes” one growing, one decaying for scaled wavenumber $\alpha < \sim 2.4$. The scaled wavenumber:

$$\alpha = \left\{ (k^2 + l^2)\, \varepsilon^{-1} \right\}^{1/2} \tag{5}$$

is proportional to absolute wavenumber and static stability.

Equations for perturbation kinetic energy, K_e and available potential energy, A_e are:

$$\begin{aligned} \frac{\partial A_e}{\partial t} &\equiv \frac{\partial}{\partial t} \int\int\int \frac{\rho_S}{2} \left\{ \varepsilon \left(\frac{\partial \psi}{\partial z} \right)^2 \right\} dx\,dy\,dz \\ &= \underbrace{\int\int\int \varepsilon\, \rho_S \frac{\partial U}{\partial z} \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial z}\, dx\,dy\,dz}_{(A_z \to A_e)} - \underbrace{\int\int\int \rho_S\, w \frac{\partial \psi}{\partial z}\, dx\,dy\,dz}_{(K_e \to A_e)} \end{aligned} \tag{6a}$$

$$\begin{aligned} \frac{\partial K_e}{\partial t} &\equiv \frac{\partial}{\partial t} \int\int\int \frac{\rho_S}{2} \left\{ \left(\frac{\partial \psi}{\partial y} \right)^2 + \left(\frac{\partial \psi}{\partial x} \right)^2 \right\} dx\,dy\,dz \\ &= \underbrace{-\int\int\int \rho_S\, U \left\{ \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial y} \right\}_y dx\,dy\,dz}_{(K_z \to K_e)} + \underbrace{\int\int\int \rho_S\, w \frac{\partial \psi}{\partial z}\, dx\,dy\,dz}_{(A_e \to K_e)} \end{aligned} \tag{6b}$$

The volume integrals are over a closed domain. In the QG system, $\partial\psi/\partial z$ is proportional to potential temperature θ, making the first term on the RHS of (6a) proportional to a meridional heat flux, while the second term is proportional to a vertical heat flux. The specified vertical shear, $\partial U/\partial z$ is proportional to the available potential energy, A_z of the basic state and is the energy upon which the baroclinic instability mechanism feeds. The first term on the RHS of (6b) is a barotropic energy conversion. The barotropic conversion is proportional to the divergence of eddy momentum flux and also draws energy from the mean flow. The second term on the RHS of (6a) and (6b) is the same but with opposite sign indicating a conversion between A_e and K_e.

1.2.2 Example solutions

This QG eigenmodel of baroclinic instability is applicable to the midlatitudes. In these regions zonal flow increases with height reaching a maximum near the tropopause. **Fig. 1a** is a representative nondimensional profile of *U* where the tropopause is at nondimensional *z = 1.0*. The growth rate and phase speed spectra along with the (growing normal mode) eigenfunction structures for different *k* are shown in **Fig. 1** as well. The growth rate has maximum value at a specific value of k. The vertical structure tends to have relative maxima at the surface and near the tropopause, but it becomes progressively more bottom-trapped for shorter waves. The phase varies such that unstable modes tilt upstream with height, i.e. against the mean flow shear. Other solutions to (3), labeled “continuum modes”, are relevant to “nonmodal” growth.

<near Fig 1>

For shorter waves the eigenmodes with lower level maximum tend to dominate (when compressibility is included) and the solution decays rapidly away from the boundary. For longer waves the tropopause level maximum tends to dominate (**Fig. 1d**). Eady model normal modes have interior $q' = 0$; from (1): the LaPlacian increases as k^2 requiring a rapid change with height for short waves to make the thermal term comparable to the LaPlacian term (this leads to boundary trapping of the solutions). For longer waves, the LaPlacian becomes small and the vertical structure is more evenly spread in the vertical, hence these modes are "deeper".

Typical geopotential patterns observed prior to frontal cyclone development have separate surface and upper troposphere troughs, each equivalent-barotropic (vertical trough axis), with the upper level trough more prominent. A crude simulation of that initial state is used to generate solutions shown in **Fig. 2.** Time series of the growth rates of several quantities are tracked over several days. The time series include potential enstrophy ($H = (q')^2$) and total energy (TE = $A_e + K_e$) integrated over the whole domain. The solutions asymptote to the most unstable normal mode growth rate as that eigenmode emerges to dominate the solution. The growth rate has transient peak values that can exceed the asymptotic (normal mode) value.

<near Fig. 2>

1.3 CLASSICAL VIEW

Baroclinic instability draws upon the APE of the environment in which an eddy sits. Since APE is related to a horizontal temperature gradient, and that in turn to the vertical shear, it can be viewed as a type of shear instability. One advantage of doing so is to make comparisons with barotropic instability which draws energy from the horizontal shear. This view provides a link to eddy fluxes that are observed and necessary for each conversion.

As demonstrated in (6) heat fluxes are necessary to have a baroclinic energy conversion. Horizontal heat fluxes imply that the temperature and mass (here ψ) fields are offset. The offset implies that the trough and ridge axes tilt upstream with elevation.

<near Fig. 3>

The QG formulation above is adiabatic, so individual parcels conserve their potential temperature (θ) over time. For unstable modes, the horizontal and vertical eddy heat fluxes must distort the θ field over time as suggested schematically in **Fig. 3.** An isentropic (θ) surface is drawn in three-dimensional perspective; it curves up and over colder areas and dips down over warmer areas. Prior to eddy development, the isentropic surface did not vary in the *x* direction and had a shape like its intersection with the wall at $x=0$. The isentropic surface is distorted by flow around the high and low pressure centers and representative cold {c} and warm {w} trajectories are also drawn. When these trajectories are projected onto the $x=0$ wall they appear to cross the initial zonal mean isentrope and have a slope that is typically half the slope of the mean isentrope. In fact, they are changing the zonal mean of the isentrope to become more horizontal, thereby reducing the horizontal temperature gradient and thus reducing A_z. In this classical view, A_z is reduced while A_e is increased by increasing the zonal undulations of the isentropic surface. Another aspect is that colder air is sinking while warmer air is rising, a process that lowers the center of mass and thus converts A_e into K_e. To lower the center of mass, the parcel paths must have the vertical component indicated but they must also be less than the slope of the mean isentropes for instability to occur.

The classical view can incorporate latent heat release as follows. The bulk of the precipitation in a developing cyclone forms in the warm air sector of the storm. The release of latent heat further depresses the isentropic surfaces where there is poleward motion implying additional conversion of A_z into A_e and K_e.

1.4 POTENTIAL VORTICITY VIEW

The potential vorticity view of instability tracks how two or more PV anomalies interact in a way that causes growth of the PV anomalies. PV is a fundamental conserved quantity for adiabatic motions. The illustrative model is designed around QGPV conservation.

A PV pattern has an associated streamfunction and horizontal wind field. In general, (1) implies that PV emphasizes smaller scale variations than the streamfunction field. Inverting (1) obtains broad patterns of □associated with isolated packets of *q*. For example, PV anomalies in the upper troposphere have corresponding streamfunction extending through the whole troposphere (but somewhat larger amplitude at the level where q has maximum magnitude. The associated winds are displaced from PV anomaly center by ¼ wavelength (~1 kkm). A similar depiction can be deduced for a PV anomaly associated with a surface temperature gradient.

<near Fig. 4>

PV anomalies are created by flow across PV contours. **Fig. 4** illustrates how two sinusoidal PV anomalies can amplify each other. The PV gradient is reversed between the two levels, increasing with y at upper levels and decreasing with y at the surface. This pattern is consistent with upper tropospheric PV and the surface temperature gradient, respectively. (Recall that q is positive for either lower geopotential heights or warmer surface temperature.) The associated winds distort the PV pattern in a way that causes the PV pattern to propogate. However, the meridional wind associated with a PV anomaly is in quadrature with that anomaly so the PV cannot amplify itself. Growth is described simply as advection at the PV extrema that further amplifies the PV pattern. Since the associated winds extend beyond the elevation of the PV anomaly, there can be interaction between the PV anomaly and a second PV anomaly at another level. When the second PV anomaly is offset from the first as in **Fig. 4a**, the associated winds amplify the first anomaly.

This mechanism also explains how developing cyclones maintain a preferred tilt (i.e. become “phase locked”). The lower anomaly is shifted horizontally to the right in **Fig. 4b** so that upper and lower anomalies are 180 degrees out of phase. The two PV anomalies no longer amplify each other’s PV anomalies (shutting off the instability mechanism).

Furthermore, the two anomalies reinforce the velocities midway between their positive and negative extremes thereby enhancing the propagation at each level; but the propagation is directed oppositely at each level thereby reducing the phase shift to reestablish the pattern in **Fig. 4a**. As with the classical view, normal modes are a special case where this phase locking is optimized.

The PV view provides theoretical weight to a classic description of how cyclones develop: an upper level trough (PV anomaly) approaches a low level baroclinic zone (another PV anomaly) then growth commences. This paradigm is commonly labeled “type B” cyclogenesis.

<near Fig. 5>

The ‘type B’ cyclogenesis is illustrated in **Fig. 5** using a QG nonlinear model. A nearly non-developing, nearly coherent upper tropospheric trough is propogating in a flow with vertical shear and is approaching a localized region of warmer surface temperature. The surface warm anomaly is also a positive anomaly of PV in the lower troposphere. The trough has maximum amplitude in the upper troposphere, so there are associated cold anomaly in the troposphere and warm anomaly in the stratosphere. The trough is a region of positive PV. Differential vorticity advection and warm air advection cause rising motion (of cold air) ahead of the trough. Analogously, sinking motion occurs behind the trough. Hence there is positive baroclinic conversion behind and negative ahead of the trough. Integrated over the whole system the net baroclinic conversion is zero. The vertical motion requires divergence and convergence above the peak rising motion. From the QG vorticity equation convergence opposes the positive vorticity advection ahead of the trough at upper levels (reinforces it at lower levels) and in so doing maintains the vertical tilt of the trough even though the advecting wind speed increases with height. When the upper trough begins to encounter the low level warm anomaly, the warm advection in the poleward flow ahead of the trough is increased. Initially at lower levels, this warmer air rises (instead of the cold air) causing the baroclinic conversion ahead of the upper trough to become less negative and even positive, and the system begins to amplify.

Observations show independent troughs at upper and lower troposphere prior to development with the upper approaching the lower. Neither trough has upstream tilt initially, such tilt emerges only after the two become favorably aligned and growth has commenced.

A necessary condition for instability is that the across-flow mean gradient of PV change sign within the domain. In the illustrative model (3) and (4), $\square > 0$ means that $Q_y > 0$ in the interior, and the surface temperature gradient ($d\square/dy < 0$) implies that $Q_y < 0$ at the ground. In the Eady model $Q_y = 0$ everywhere in the interior, so the normal mode instability comes from BPV having opposite sign at top and bottom boundaries.

A necessary condition for instability is that a steering level, where $U = \mathrm{Re}\{c\}$, lie within the domain. A supportive kinematic argument is that air parcels remain with the system (rather than blow through it or be left behind) and are more easily mixed laterally. For really long waves, strong retrogressive motion caused by the □term leads to a different class of unstable eigenmodes for □< ~1.1 (note cusp in **Fig. 1b**) than for larger □

The eddy meridional flux of potential voriticy is also linked to **∇·F** in quasi-geostrophic theory, where **F** is the Eliassen-Palm flux. In addition, due to the strong meridional eddy heat flux present in a baroclinically-growing eddy, **F** has an upward pointing component. From wave theory, specifically the 'generalized Eliassen-Palm relation', the upward pointing EP flux provides an explanation for upper level amplification of the eddy as it grows nonlinearly.

1.5 NORMAL MODES

Normal modes are physically meaningful eigenfunctions. As in the illustrative model, the equations are linearized about a specified basic state and perturbation solutions are sought. Most commonly, the time and one or more space dependencies are assumed. By assuming a form like (4), unstable solutions grow exponentially. Simple enough models may be solved analytically. More commonly, the eigenvalue problem is solved numerically.

Normal modes are consistent with many observed features:

i. Unstable modes tend to be lined up along the jet axis (if present) in the mean flow.

ii. The most unstable wavelength is similar to the observed median size. The normal mode scale can be manipulated by varying the choices made for nondimensional parameters, but is on the order of 4500 km.

iii. Solutions tend to develop similar zonal and meridional lengths, the latter responds to the width of the jet providing one natural scale in the model. Other properties (like static stability) also influence the length scales.

iv. The vertical structure of the most unstable modes tends to have relative maxima at the surface and upper troposphere.

v. In growing normal modes the temperature lags the mass field in the lower troposphere (typically by 20-50 degrees of phase for the most unstable mode). Two consequences are:

First, troughs and ridges in the mass field must be displaced (i.e. tilt) upstream with increasing elevation. There is typically ¼ to ½ wavelength (1-2 kkm) between the trough location at the surface and at tropopause level.

Second, the lag allows across-flow heat fluxes down the temperature gradient, as expected from (6a), even for geostrophic winds. In the Eady model the heat flux is uniform with height. Model improvements, most notably compressiblity, can emphasize the eddy heat flux in the lower troposphere (where observations find it most prominent).

vi. The rate of propagation is ~10-20 m/s: slower than jet stream level winds, but faster than (zonal average) surface winds. A steering level is defined as where the propagation speed of the storm equals the wind component along the storm's track. The steering level for the most unstable normal modes is typically between 700-500 mb depending on the assumptions made. For shorter waves, the steering level is closer to the surface since these modes move slower. Longer waves respond to competing effects: they have greater upper level amplitude (where U is faster) but greater sensitivity to □(which enhances retrograde motion).

vii. The rate of growth is similar to but slower than that of observed cyclones. Observed doubling times are typically 1-2 days at upper levels.

viii. Instability is inversely proportional to static stability. For example, the peak growth rate depends on □(= 2.0 in **Fig. 1).** From (2) and (5), □is proportional to static stability □Hence, smaller □places the most unstable peak at larger k making the growth rate (k Im$\{c\}$) larger. Kinematically, vertical motion needed in (6) becomes easier for smaller □

The fact that normal modes have fixed tilt is not necessarily unrealistic. Observations of the vorticity equation terms support an approximately fixed structure for a developing low because the divergence term opposes the horizontal advection at upper levels but reinforces the horizontal advection at low levels. The normal modes (**Fig. 1**) are special structures where the net advection is exactly uniform throughout the depth of the fluid.

Tracking observed frontal cyclone troughs over time shows evidence that such storms maintain a roughly fixed tilt during their growth. The vorticity equation also illustrates instability whereupon the divergence term has positive vorticity tendency at a trough where vorticity is a maximum thus amplifying the peak vorticity (and vice-versa for ridges).

In addition to the normal modes, the eigenfunctions include a class of solutions called "continuum" modes. For an adiabatic model continuum modes have equivalent barotropic structure (no tilt) making them neutral. In the Eady model, continuum modes have zero PV at all levels except at the critical level, where their amplitude has a "kink". Continuum modes play a role in nonmodal growth.

1.6 NONMODAL GROWTH

Nonmodal growth is seen when solving initial value problems like (3). The formulation can be linear as in (3) or nonlinear. This approach is more general than eigenanalysis since the time dependence is not assumed as it is in (4).

The solution at any time can be decomposed into a combination of eigenfunctions. For an arbitrary initial state, continuum and normal modes are present. These modes move at differing speeds. In a linear formulation the modes operate independently; as modes disperse, positive and negative reinforcement varies. The interference between modes decays algebraically asymptotically. However, for some initial conditions it is possible to have sizable growth over a limited time period.

For the Eady model, analytic solutions can be found which illustrate the process. Using an initial condition with upstream tilt ($\square \sim$ exp(imz) in (4) where m>0) yields solutions with normal mode and algebraic parts. The algebraic part has time dependence proportional to $\{(m - kt)^2 + \square^2\}^{-1}$ and $\exp\{\, i\,(m - kt)z\}$. The amplitude increases while the upstream tilt becomes more vertical until $t = mk^{-1}$. After that, the wave tilts downstream and decays.

Initial upstream tilt becoming more vertical with time has led to an expectation that RV increases at the expense of TV while interior PV remains conserved. However, exceptions can be found where large nonmodal growth occurs (in H) as upstream tilt *develops* from an initial state with no tilt. The explanation lies in a rough cancellation between RV and TV leaving the BPV evolution to dictate rapid growth in H.

A robust interpretation of nonmodal growth is the progressively more favorable superposition of constituent modes. Continuum modes having mainly upper level amplitude tend to move fast, while modes with mainly lower level amplitude move slowly. Decomposition into eigenmodes of an initial state with upstream tilt finds faster continuum modes located upstream of slower continuum modes. Over time, the modes become more favorably lined up; the tilt becomes more vertical and the total amplitude increases. **Fig. 6** illustrates the process.

<near Fig. 6>

Nonmodal growth can be quite strong in simple models like Eady's. However, most improvements to the model such as adding compressibility, variable Coriolis, and realistic vertical shear of *U* reduce nonmodal growth. Using more realistic initial states also tends to reduce nonmodal growth (e.g. using a wave packet instead of a wavetrain; using separate un-tilted upper and lower features instead of connecting them with a tilt).

1.7 OTHER ISSUES

Baroclinic instability has links with barotropic instability. First, each instability draws energy from mean flow shear. Second, barotropic instability has a similar stability criterion (absolute vorticity gradient changing sign in the domain). Third, there can be interference between the two instabilities. The most unstable baroclinic eigenmode has optimal structure for a flow having the vertical shear alone, but when horizontal shear is added to that flow a different structure is needed otherwise the eddy will be sheared apart. The subsequent structure is unlikely to be as optimal for baroclinic energy conversion. Hence, the baroclinic conversion will usually be reduced, though the barotropic growth mechanism may compensate. **Fig. 7c** illustrates such a calculation; in this case adding a purely barotropic flow reduced the growth rate even though the barotropic growth mechanism was activated.

<near Fig. 7>

Baroclinically unstable frontal cyclones prefer to develop in certain regions. The preference may arise from local conditions such as: lower static stability or locally greater vertical shear. The illustrative model above assumes a wavetrain solution; when more localized development is considered, a variety of issues are raised.

For example, if one uses a single low as the initial condition, the solution typically evolves into a chain of waves as the modal constituents of the initial state disperse. Alternatively, a wave packet initial condition might be used consisting of a "carrier wave" multiplied by an amplitude envelope. The packet evolution depends upon the mean flow properties and assumptions made in the model. However, for reasonable choices of parameters, one might find a packet that spreads while propagating downwind. The leading edge of the packet has mainly faster, wider, and deeper modes. The trailing edge has slower, shorter, and shallower waves. It is possible to construct a localized structure which resists this dispersion by making a judicious combination of eigenmodes having similar phase speed, but different zonal wavenumber. One such example was used when discussing 'type B' cyclogenesis (**Fig. 5**). **Fig. 8** illustrates another example using neutral continuum modes. When this model is solved as an initial value problem the packet maintains a localized shape for a long time and almost no growth occurs since the normal modes were filtered out and there is very slow phase shifting of the constituent modes. However, when nonlinear advection is allowed, modes interact and soon amplitude is injected into all the eigenmodes including the growing normal modes which grow rapidly in this example.

<near Fig. 8>

Studies of regional development spawned sub-categories of baroclinic instability. "Absolute" instability occurs when the wave packet expands faster than it propagates; the amplitude at a point keeps growing. "Convective" (in the advection sense) instability occurs when the packet moves faster than it spreads so that growth then decay occurs as the packet moves past a point. "Global" instability (like the eigensolutions shown here) has growth that is invariant to a Gallilean transform. Such is not the case for "locally" unstable modes. Normal modes for zonally-varying basic states look like carrier waves modulated by a spatially-fixed amplitude envelope; the envelope locally modifies the growth rate (sometimes called "temporal" instability); enhancing the global growth locally where the carrier wave propagates from lower to higher amplitude of the envelope. "Spatial" instability allows wavenumber to be complex while phase speed remains real.

Nonlinear calculations raise other issues related to baroclinic instability. One issue concerns equilibration. The growing wave modifies the mean flow while drawing energy from it. This places a limit upon the cyclone development. In PV theory, this may be where the distortion shown in **Fig. 4** becomes comparable to the cyclone width. Waves longer than the most unstable wave tend to reach larger amlitude than the linearly most unstable mode. One reason why is that they are deeper and so can potentially tap more APE in the mean flow. Another reason may be the larger scale in both horizontal dimensions provides a longer time for PV contour distortion. Another reason concerns the inversion of a PV anomaly: the streamfunction amplitude is larger for a broader PV anomaly.

"Life-cycle" studies model cyclones from birth to peak amplitude to decay. These studies typically find baroclinic growth followed by barotropic decay. This cycle fits the observed facts that eddies have a net heat flux and a net momentum convergence. These studies also reveal a characteristic evolution of the eddy structure: upper level amplification compared to the linear eigenmodes. An explanation is that saturation is reached sooner at the critical level and at the surface while upper levels continue to grow. Another was given above regarding the Eliassen-Palm flux **F**. When averaged over the life-cycle, the vertical distribution of the zonal mean eddy heat and momentum fluxes becomes more realistic.

Finally, the atmosphere has higher order processes than the QG system. The biggest impact of ageostrophy is to break symmetries in the solutions. **Fig. 7d** shows the leading order ageostrophic effects for a linear model. Ageostrophy: causes enhanced eddy development on the poleward side (mainly by negative baroclinic conversion on the equatorward side), builds mean flow meridional shear, and slows down the wave. Ageostrophy also causes contours to be more closely spaced around a low and more widely spaced around a high.

Further Reading

Gill, A. (1982) Atmosphere-Ocean Dynamics. New York: Academic Press.

Grotjahn, R. (1993) Global Atmospheric Circulations: Observations and Theories. New York: Oxford University Press.

Holton, J. (2004) An Introduction to Dynamic Meteorology, 4th Edition. San Diego, Academic Press.

Hoskins, B., McIntyre, M., and Robertson, A. (1985) On the use and significance of isentropic potential vorticity maps. Quarterly Journal of the Royal Meteorological Society, 111, 877-946.

Pedlosky, J. (1987) Geophysical Fluid Dynamics, 2nd Edition. New York: Springer-Verlag.

Pierrehumbert, R. and Swanson, K. (1995) Baroclinic instability. Annual Reviews of Fluid Mechanics, 27, 419-467.

Vallis, G. (2006) Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulations, New York: Cambridge University Press.

(Suggested cross-references)

See also:

Barotropic Flow and Barotropic Instability (0076). **Cyclogenesis** (0129). **Cyclones**: Extratropical cyclones (0128). **Dynamic Meteorology**: Balanced Flow (0484); Overview (0138); Potential Vorticity (0140); Waves and Instabilities (0141). **Quasi-geostrophic Theory** (0326), **Vorticity** (0449).

FIGURE CAPTIONS

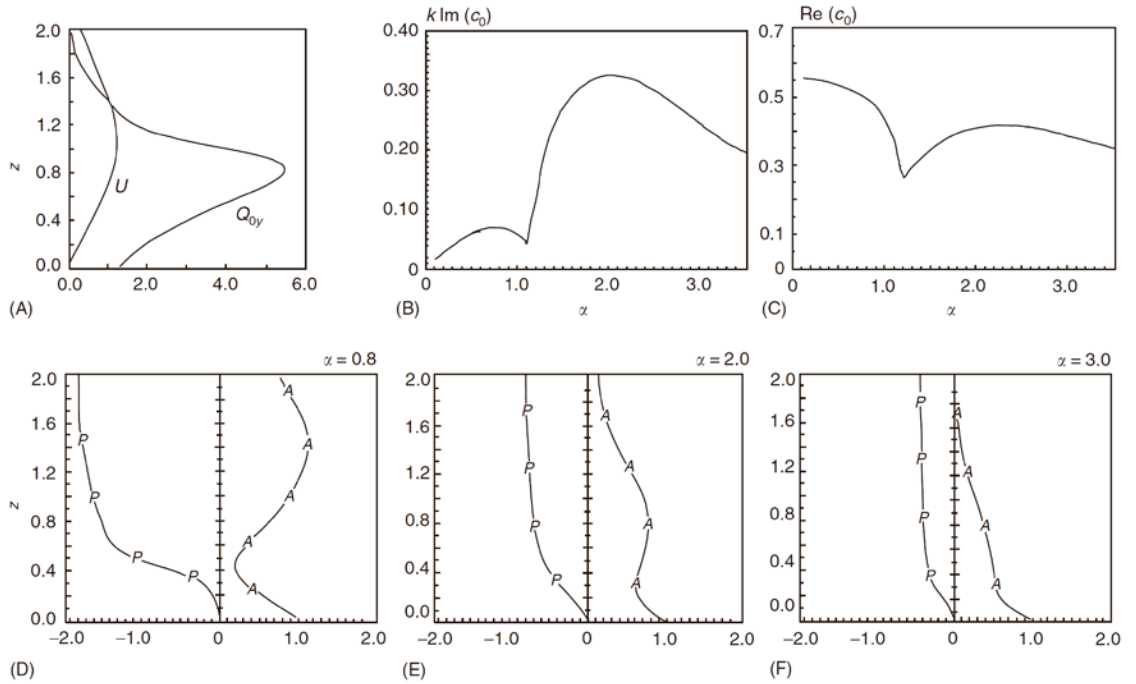

Figure 1 Quasi-geostrophic eigenanalysis. (A) Specified zonal wind *U*, and meridional gradient of interior potential vorticity Q_{0y} versus scaled height. $Z = 1$ is 10km. (B) Growth rate and (C) phase speed versus absolute wavenumber α (D)-(F) Amplitude *A*, and phase *P*, for the growing normal mode for α= 0.8, α= 2.0, and α= 3.0, respectively. All three modes tilt westward (upstream) with increasing height. Dimensional wavelengths depend upon scaling assumptions, but reasonable choices imply that α= 0.8, α= 2.0, and α= 3.0 correspond to 11., 4.4, and ~2.9 kkm wavelengths respectively. (Zonal and meridional scales are set equal.) The same scaling implies phase speed of 9 ms^{-1} and doubling time of ~1.2 days for α= 2.0. Adapted with permission from Grotjahn R (1980) *Journal of the Atmospheric Sciences*, 37: 2396-2406.

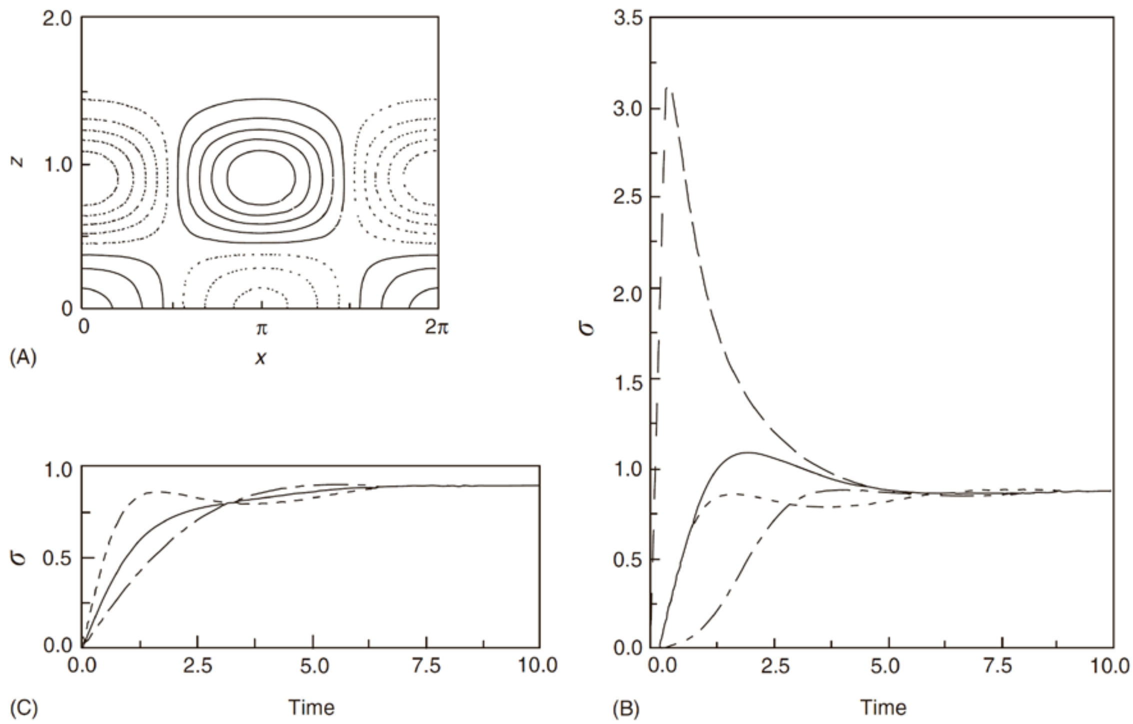

Figure 2 Initial value calculation. (A) Zonal cross section of initial streamfunction, dashed contours used for negative values. (B) Time series of growth rates for domain average potential enstrophy (solid line) and its components: RV^2 (short dashed line), TV^2 (dot-dashed line), and BPV^2 (long dashed line). Growth rates asymptote to the most unstable normal mode rate for this wavenumber (a=2.0). (C) Similar to (B) except for total energy (solid line), kinetic energy (short dashed line), and available potential energy (dot-dashed line).

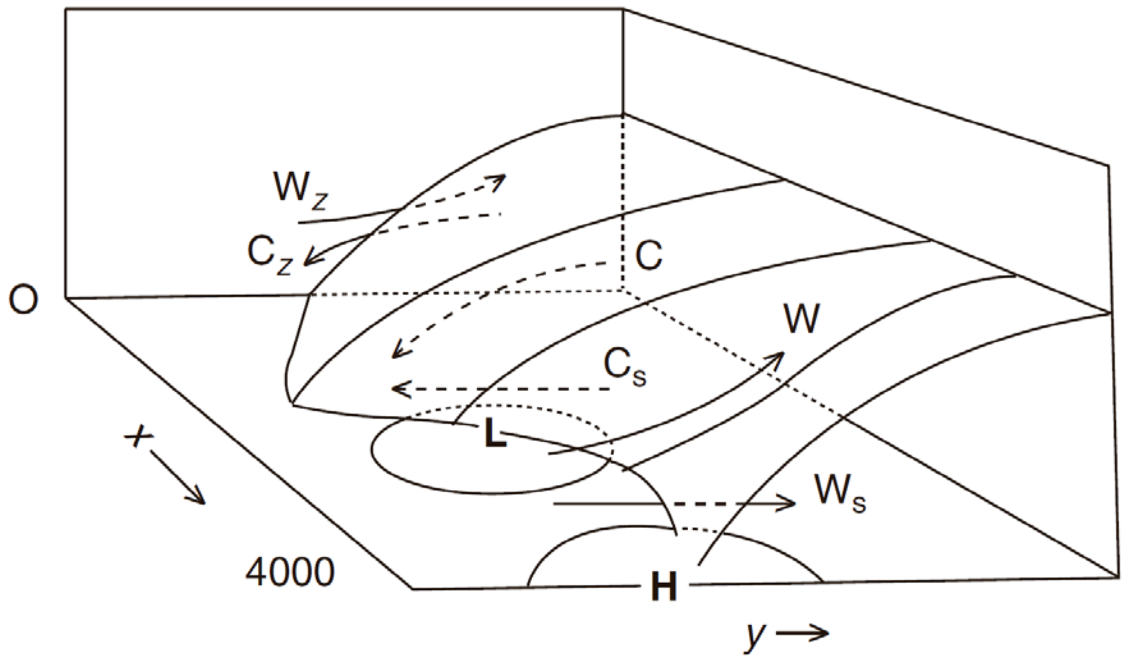

Figure 3 Schematic diagram showing distortion of an isentropic surface by a baroclinically amplifying frontal cyclone. Dotted lines used for objects underneath the three dimensional isentropic surface. Surface high H, and low L, marked along with 2 representative contours of surface pressure. Trajectories of representative parcels in the warm air W, and cold air C. Subscript s denotes projection onto the bottom surface while z denotes projection onto the meridional plane (where $x = 0$). The trajectories do not cross the isentropic surface but distort it. Initially the isentropic surface had negligible variation with x and looked like the current pattern at $x = 0$. The projections W_Z and C_Z seem to cross the initial isentropic surface but in fact are flattening it (which reduces A_z). Rising air is warm while sinking air is cold which lowers the center of mass converting A_e into K_e

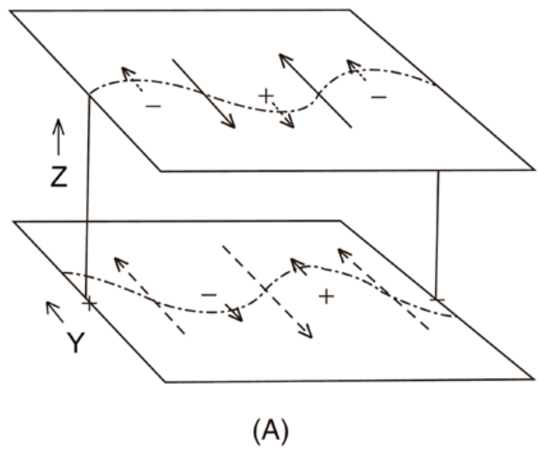

Figure 4 Baroclinic instability from interacting PV anomalies at two levels. A representative PV contour (dot-dashed line) is drawn at each level. Note that the meridional gradient (*Y* direction) is opposite at the two levels. The offset is (A) ¼ wavelength and (B) ½ wavelength. A typical wavelength might be 4 kkm. Each anomaly has associated wind component parallel to the PV gradient; dashed arrows are winds from the lower PV anomaly, while solid arrows are from the upper anomaly. The winds created by each anomaly propagate that anomaly. In (A) each PV anomaly has a wind component that amplifies the undulation in the other anomaly (by having a non-zero wind at the center of the other anomaly) thereby causing growth. In (B) each PV anomaly has a wind component that augments the propagation of the other anomaly in the manner indicated by the broad arrows; this causes the anomalies to migrate to a phase offset like diagram (A).

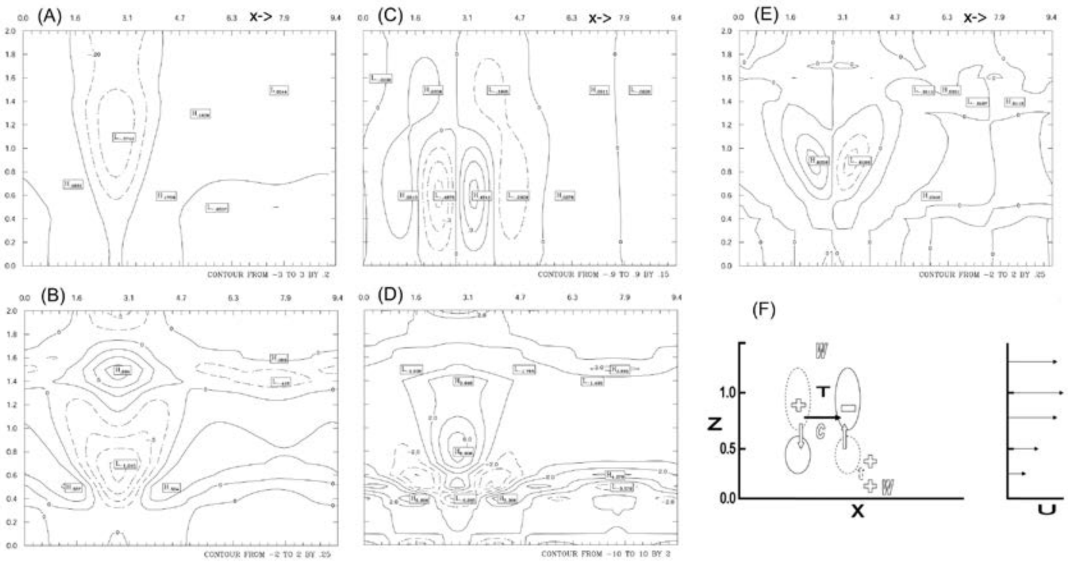

Figure 5 (A) **–** (E) Zonal cross sections (East-West direction, *x* versus elevation, *z*) showing properties across the midpoint of a nearly-coherent, non-growing, upper level trough similar to those observed. *z* is scaled by 10 km and *x* by 1 kkm. Eddy: (A) streamfunction; (B) temperature; (C) vertical motion, positive upwards; (D) quasi-geostrophic potential vorticity; and (E) baroclinic energy conversion shown. The baroclinic energy conversion averaged over the whole upper trough is zero. (F) Schematic zonal cross section of interaction between upper trough encountering a near surface warm anomaly. The upper trough, T moves in the direction of the solid arrow. The trough has eddy temperatures indicated by *C* for colder air, and *W* for warmer air. Hollow arrows show vertical motions. In the along-flow direction (*x*) there is upward motion ahead of the trough reaching a maximum near $z \sim 0.6$. Vertical motion is driven both by temperature advection (from $0.6<z<1.4$) and differential vorticity advection (from $0.3<z<0.7$). The associated divergence is indicated by solid ovals and convergence by dashed ovals. From the quasi-geostrophic vorticity equation, these divergence fields oppose vorticity advection by the mean flow at upper levels, and enhance that advection at lower levels. Hence, the trough maintains its vertical tilt in the presence of vertical shear in the zonal mean wind (U(z)). The sign of the baroclinic energy conversion, BCE is indicated by open + and – signs. The upper pattern of BCE is similar to panel (E). However, when the upward and poleward motion ahead of the upper trough encounters the warm anomaly, the vertical motion is locally enhanced as is the meridional heat flux. There is net generation of voriticty, net BCE, and the eddy begins to grow. Adapted with permission from Grotjahn, R. (2005) *Quarterly Journal of the Royal Meteorological Society*, 131: 109–124. doi: 10.1256/qj.03.163

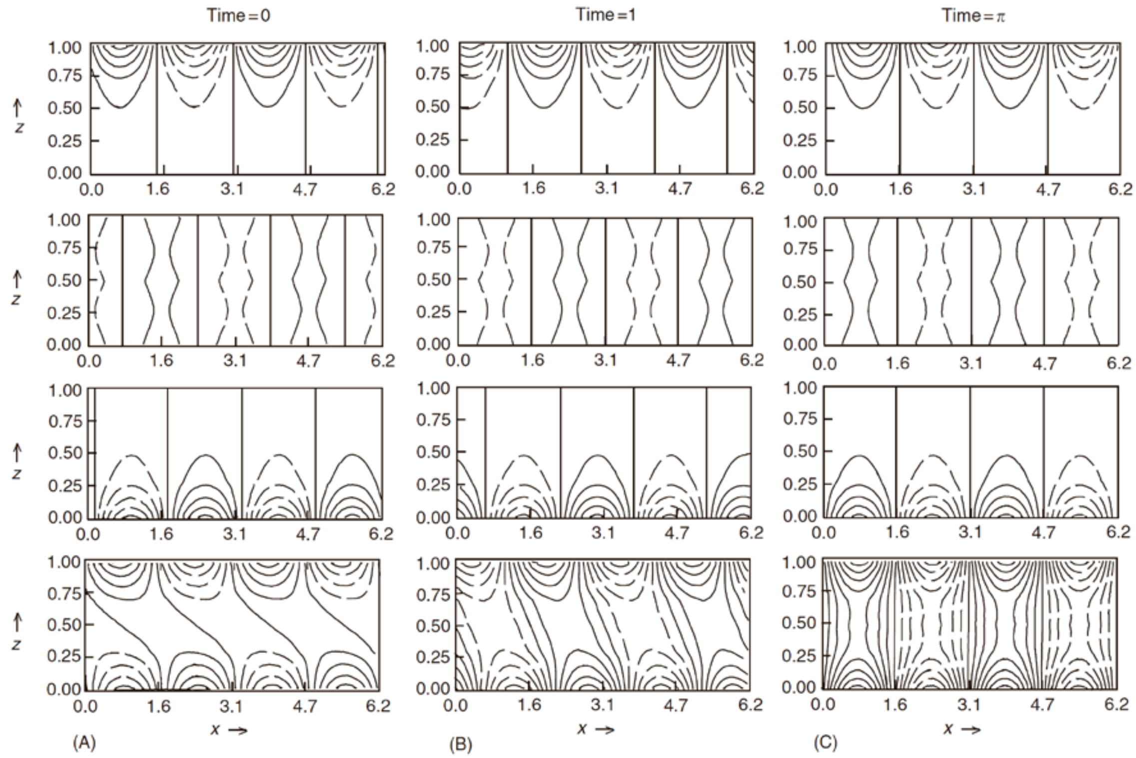

Figure 6 Nonmodal growth as a superposition process. Four initial value linear calculations are shown. The top three rows show three individual neutral continuum modes at three times. The bottom row used the sum of the three modes at the initial time. (A) initial condition, (B) time when energy growth is a maximum in the sum, *Time* = 1, (C) time when growth rate is zero in the sum, *Time* = π. Adapted with permission from Grotjahn R, Pedersen R, Tribbia J (1995) *Journal of the Atmospheric Sciences*, 36: 764-777.

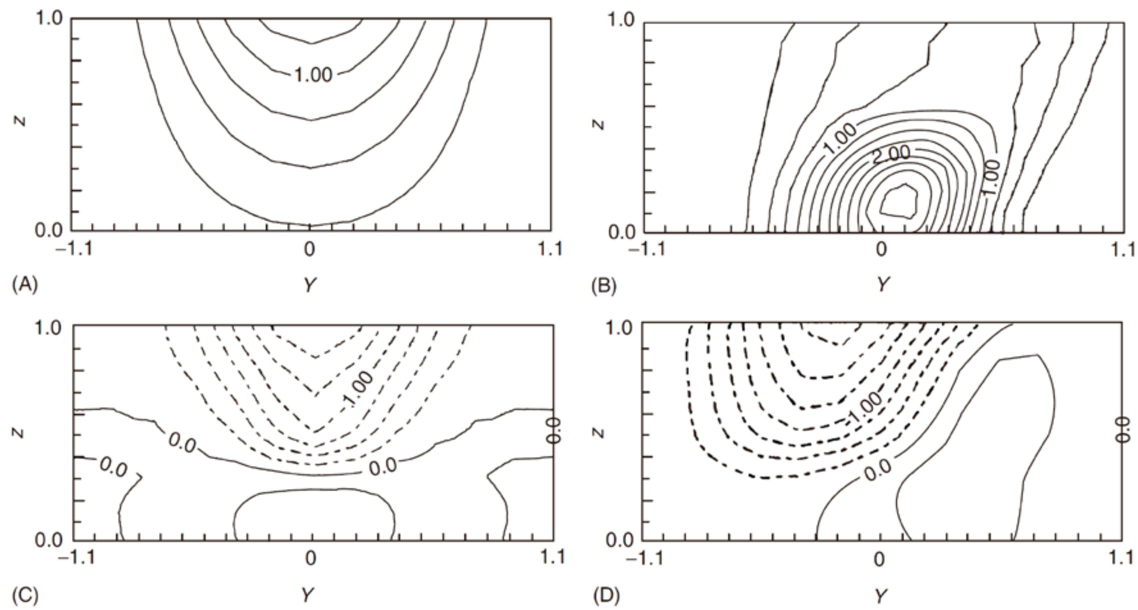

Figure 7 Baroclinic energy conversion ($A_z \rightarrow A_e$) for four models. (A) Lowest-order, square wave solution for an Eady-type model but including compressibility, increasing vertical shear in U, □= 1. (B) Solution when a surface frontal zone, centered at $Y = 0$, is added to the lowest order mean flow U_0 and leading ageostrophic advective effects are included (using geostrophic coordinates). The frontal zone adds wind field: $0.2(2z\text{-}z^2)\ U_1$ where $U_1 = b_1(1 - \tanh^2(aY) - b_2 - 3b_3Y^2$ to U_0. The geostrophic coordinate transform causes the asymmetry. (C) Correction to the conversion shown in (a) when barotropically unstable horizontal shear U_1 is added to U_0. If the total wind is $U = U_0 +$ □U_1 then the total conversion is (A) + □(C). The barotropic shear reduces the growth rate. (D) Modification due to all leading order ageostrophic corrections. If those corrections are order □, then the total conversion is (A) + □(D). Ageostrophic conversions reduce the conversion and introduce asymmetry. Adapted with permission from Grotjahn, R., 1979: *Journal of the Atmospheric Sciences*, 36: 2049-2074.

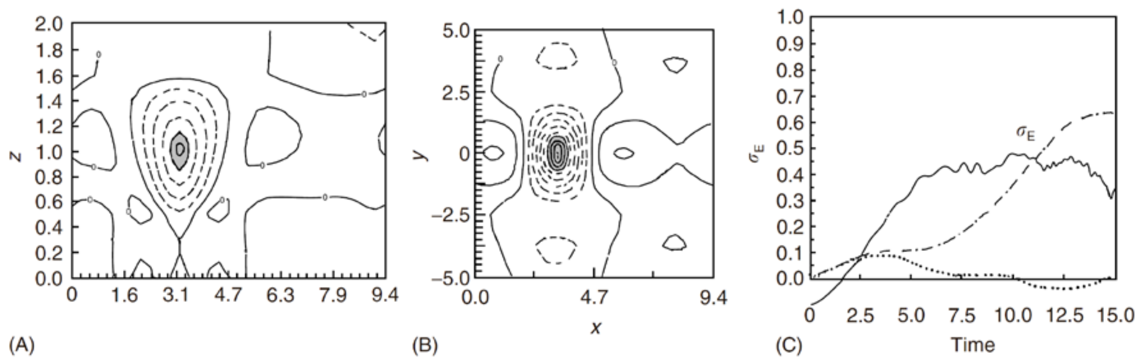

Figure 8 Initial value calculations for a linearly localized initial condition. (A) Zonal cross section showing contours of streamfunction initially. Values < -1.0 are shaded. (B) Horizontal pattern of streamfunction at tropopause level (z = 1.0) initially. Initial condition constructed from neutral modes having similar phase speed. Growing or decaying normal modes are excluded. (C) Time series of energy growth rate for three integrations. Linear model (dotted line) showing little growth since the nonmodal mechanism is weak and growing normal modes cannot develop. Also shown are nonlinear calculations for two amplitudes of the initial condition, where the solid line uses three times the initial amplitude of the dot-dashed line. Growing normal modes are activated by nonlinear interaction. Some evidence of nonlinear saturation is seen.

Keywords for MS 076:

Ageostrophy

Effects on baroclinic development

Baroclinic Instability

Conversion

Interaction with barotropic instability

Necessary conditions for,

Barotropic conversion

Continuum modes

Diabatic heating

Energy:

Available potential

Conversions

Kinetic

Potential

Quasi-geostrophic

Extratropical cyclones

Instability

Absolute

Baroclinic

Barotropic

Convective

Global

Spatial

Temporal