

The counter-propagating Rossby-wave perspective on baroclinic instability. Part III: Primitive-equation disturbances on the sphere

Article

Published Version

Methven, J. ORCID: <https://orcid.org/0000-0002-7636-6872>, Heifetz, E., Hoskins, B. J. and Bishop, C. H. (2005) The counter-propagating Rossby-wave perspective on baroclinic instability. Part III: Primitive-equation disturbances on the sphere. Quarterly Journal of the Royal Meteorological Society, 131 (608). pp. 1393-1424. ISSN 1477-870X doi: 10.1256/qj.04.22 Available at <https://centaur.reading.ac.uk/95/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://www.interscience.wiley.com>

To link to this article DOI: <http://dx.doi.org/10.1256/qj.04.22>

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

The counter-propagating Rossby-wave perspective on baroclinic instability. Part III: Primitive-equation disturbances on the sphere

By J. METHVEN^{1*}, E. HEIFETZ², B. J. HOSKINS¹ and C. H. BISHOP³

¹*University of Reading, UK*

²*Tel-Aviv University, Israel*

³*Naval Research Laboratories/UCAR, Monterey, USA*

(Received 16 February 2004; revised 28 September 2004)

SUMMARY

Baroclinic instability of perturbations described by the linearized primitive equations, growing on steady zonal jets on the sphere, can be understood in terms of the interaction of pairs of counter-propagating Rossby waves (CRWs). The CRWs can be viewed as the basic components of the dynamical system where the Hamiltonian is the pseudoenergy and each CRW has a zonal coordinate and pseudomomentum. The theory holds for adiabatic frictionless flow to the extent that truncated forms of pseudomomentum and pseudoenergy are globally conserved. These forms focus attention on Rossby wave activity.

Normal mode (NM) dispersion relations for realistic jets are explained in terms of the two CRWs associated with each unstable NM pair. Although derived from the NMs, CRWs have the conceptual advantage that their structure is zonally untilted, and can be anticipated given only the basic state. Moreover, their zonal propagation, phase-locking and mutual interaction can all be understood by ‘PV-thinking’ applied at only two ‘home-bases’—potential vorticity (PV) anomalies at one home-base induce circulation anomalies, both locally and at the other home-base, which in turn can advect the PV gradient and modify PV anomalies there. At short wavelengths the upper CRW is focused in the mid-troposphere just above the steering level of the NM, but at longer wavelengths the upper CRW has a second wave-activity maximum at the tropopause. In the absence of meridional shear, CRW behaviour is very similar to that of Charney modes, while shear results in a meridional slant with height of the air-parcel displacement-structures of CRWs in sympathy with basic-state zonal angular-velocity surfaces. A consequence of this slant is that baroclinically growing eddies (on jets broader than the Rossby radius) must tilt downshear in the horizontal, giving rise to up-gradient momentum fluxes that tend to accelerate the barotropic component of the jet.

KEYWORDS: Momentum fluxes Phase-locking Realistic jets Wave-activity orthogonality

1. INTRODUCTION

Baroclinic growth is the essential driving mechanism behind the existence of extra-tropical weather systems. The first quantitative theories of weather systems considered the baroclinic instability of normal mode (NM) perturbations to steady zonally-symmetric flows described by quasi-geostrophic (QG) dynamics (Charney 1947; Eady 1949). The necessary conditions for the baroclinic instability of a basic state in the QG framework are that the meridional potential-vorticity (PV) gradient must take opposite signs somewhere within the domain (Charney and Stern 1962) and the zonal wind and PV gradient should be positively correlated (Fjørtoft 1950).

Bretherton (1966b) showed how baroclinic instability in the Eady and two-layer (Phillips 1954) models can be viewed in terms of the zonal propagation and interaction of two counter-propagating Rossby waves (CRWs). In these models, the PV gradient is only non-zero at two levels and a CRW is associated with zonal propagation on each one. In general, growth on basic states with three or more distinct locations with non-zero PV gradient can be described by interactions between CRWs propagating at each location (e.g. Heifetz *et al.* 1999). Heifetz *et al.* (2004a) (hereafter Part I) extended this QG perspective to general zonal flows where the PV gradient can vary continuously throughout the atmosphere but, for particular initial conditions, any disturbance consisting of a superposition of a growing NM and its decaying complex conjugate (CC) can

* Corresponding author: Department of Meteorology, University of Reading, PO Box 243, Earley Gate, Reading RG6 6BB, UK. e-mail: J.Methven@reading.ac.uk,

alternatively be viewed in terms the evolution of a *pair of CRWs*. Since these CRWs are derived from the NMs, they cannot describe initial-value problems involving the continuous spectrum (Farrell 1982). The CRW structures are zonally untitled by definition and can be obtained objectively given the basic state and growing NM structure by imposing the constraints that the CRWs are orthogonal with respect to pseudomomentum and eddy energy growth. Evolution equations can be obtained for the CRWs by considering global integrals of the PV equation or from the Hamiltonian equations for a system where the CRWs are the basic components with distinct zonal positions and pseudomomenta.

Alternatively, CRW structures can be obtained by the ‘home-base method’ where the NM and its CC are superposed such that the resulting untitled CRW has zero PV at a chosen location. The home-base of CRW-1 is defined to coincide with the zero PV line of CRW-2 so that perturbation PV there is given entirely by the PV of CRW-1, and vice versa for the home-base of CRW-2. Evolution equations for the CRWs can then be obtained from the PV equation at the chosen two home-bases. These equations are homomorphic to the equations for the ‘orthogonal CRWs’ but there are different expressions for the coefficients determining the propagation rates and growth of the CRWs. Although the choice of home-bases is subjective, this method has the advantage that the interaction coefficients depend only upon the PV gradient, the perturbation PV and the meridional wind that the CRWs induce at these two home-bases. Simple ‘PV-thinking’ arguments can then be used to deduce the effect of interaction on the CRW propagation rates, enabling phase-locking and the mutual growth of the CRWs, just as for the two-layer model (Hoskins *et al.* 1985).

Heifetz *et al.* (2004b) (hereafter Part II) calculated CRW structures for the NMs of the Charney model using both the orthogonality and home-base methods. Provided that the choice of home-bases is motivated by the structures of the orthogonal CRWs, both methods give very similar results for the CRW structure, phase difference and phase speeds. While wave-activity orthogonality can be used to obtain CRWs objectively, their interaction can be discussed in terms of the meridional wind that each induces at the other CRW’s home-base. The self-induced phase speed of a CRW is composed of advection by the zonal wind and its propagation rate on the PV gradient at its own home-base, given the meridional wind that it induces there. Phase-locking is achieved when the phase speeds of both CRWs are made equal by interaction.

Here CRW theory is generalized to describe linearized primitive-equation dynamics on a rotating sphere, making clear the necessary assumptions (section 2). CRW forms of the necessary conditions for instability are obtained. In section 3, the NM dispersion relations and their associated CRW structures are investigated for a zonal jet resembling the mid-latitude jet in a storm-track region, including a realistic potential-temperature distribution and tropopause. The basic state, referred to as Z1, is the same as used for the LC1 experiment of Thorncroft *et al.* (1993) (hereafter THM). The effect of surface meridional shear on CRW structures, phase difference and phase speeds is then investigated using the basic state of THM’s LC2 experiment (Z2).

In Part II it was shown how the difference in the scale-dependence of PV inversion between interior and boundary PV anomalies moulds the dispersion relation of the Charney modes. CRWs must adopt an increasingly ‘hindering’ phase-locked configuration as the wavelength increases. In particular, the Rossby-wave westward propagation of the upper CRW relative to the zonal flow is so strong that it has to be increasingly hindered by the coupling of the two CRWs. This implies that their phase difference ϵ^+ must increase. It is shown in section 4 that the hindering configuration ($\pi/2 < \epsilon^+ < \pi$) of the Charney CRWs also applies to CRWs growing on realistic zonal jets without surface shear. Also, the change to air-parcel displacement structures of CRWs expected

in an environment with meridional shear explains why the horizontal momentum flux of baroclinically growing NMs is always up-gradient when the width of the jet is greater than the width of the CRW structures at the ground. Balasubramanian and Garner (1997) proposed that the distribution of NM momentum flux influences the direction of Rossby wave breaking at upper levels. The companion paper, Methven *et al.* 2005 (hereafter Part IV), analyses the properties of CRW theory that provide robust predictions for the nonlinear evolution of baroclinic wave life cycles.

2. CRW THEORY FOR THE PRIMITIVE EQUATIONS

(a) *Conserved properties*

The primitive equation form of the theory of CRWs makes use of the same principles as in the quasi-geostrophic (QG) case; namely the conservation of mass, potential temperature (θ) and PV following air parcels, and the global conservation of the disturbance quantities pseudomomentum and pseudoenergy. In this section, CRW evolution equations are derived and the procedure for obtaining CRWs from a growing NM is outlined, stating the necessary assumptions about the flow and choices concerning CRW structure that must be made.

Assumption 1. The flow is adiabatic and frictionless so that mass, θ and Ertel PV (P) are conserved. Together these properties imply that air moves along isentropic surfaces and can be traced using PV contours.

Assumption 2. Disturbances are small amplitude and can be linearized about a steady zonally-symmetric basic state. The PV equation on the sphere in isentropic coordinates becomes:

$$\left[\frac{\partial}{\partial t} + \overline{U} \frac{\partial}{\partial \lambda} \right] P' + v' \overline{P}_y = 0. \quad (1)$$

Here λ is longitude, ϕ is latitude, $y = a\phi$ is meridional distance, a is the earth's radius, $\overline{U} = \overline{u}/(a \cos \phi)$ is the zonal angular velocity of the basic state, and $\overline{P}_y$ is the meridional derivative of the basic-state potential vorticity. Since the meridional wind $v' = [\partial/\partial t + \overline{U}\partial/\partial \lambda]\eta$, the meridional displacement of air parcels, η , can be related to PV anomalies on isentropic surfaces or θ -anomalies on the lower boundary through advection of the basic-state gradients:

$$\begin{aligned} \eta(\phi, \theta) &= \frac{-P'}{\overline{P}_y} \quad \text{along isentropic surfaces} \\ \eta(\phi) &= \frac{-\theta'}{\overline{\theta}_y} \quad \text{along lower boundary.} \end{aligned} \quad (2)$$

Wave-like disturbance quantities take the form $q' = \text{Re}\{q_m(\phi, \theta) \exp(im\lambda)\}$ where m is zonal wave number and q_m is a complex structure function. An inner product can be defined for global integrals of quadratic disturbance quantities weighted by mass:

$$\begin{aligned} \langle p', q' \rangle &= \iint \overline{p'q'} \bar{r} a^2 \cos \phi \, d\phi \, d\theta \\ &= \iint \frac{1}{4} (p_m^* q_m + p_m q_m^*) \bar{r} a^2 \cos \phi \, d\phi \, d\theta, \end{aligned} \quad (3)$$

where the overbar denotes the zonal average, $*$ represents the complex conjugate and $r = -(1/g)\partial p/\partial \theta$ is the pseudodensity in isentropic coordinates. Andrews (1983) and

Haynes (1988) derived expressions for the pseudomomentum and pseudoenergy associated with flow described by the hydrostatic primitive equations. Pseudomomentum is defined for disturbances to zonally-symmetric basic states and pseudoenergy relates to disturbances to steady basic states. Both quantities are globally conserved, even for large-amplitude disturbances, if the basic state is both steady and zonal and the flow is adiabatic and frictionless. Magnusdottir and Haynes (1996) extended the pseudomomentum conservation law to include an important boundary contribution associated with surface θ -anomalies. If two waves are present, labelled i and j , a pseudomomentum inner product for small amplitude waves can be defined as:

$$\mathcal{P}_{ij} = \frac{1}{2} \langle \eta_i, \overline{Q}_y \eta_j \rangle + \left\langle \frac{-r_i}{\bar{r}}, \cos \phi u_j \right\rangle + \langle \eta_i, \bar{\theta}_y \cos \phi \delta_b u_j \rangle, \quad (4)$$

where $\delta_b = \delta\{\theta - \theta_b(\phi)\}$ is a delta function (with units K^{-1}) picking out the ground in isentropic coordinates at each latitude. Note that the quantities inside the brackets have units of $m s^{-1}$ and equal the pseudomomentum density divided by $\bar{r}$ (see (28) for the conservation law). The generalized PV gradient:

$$\overline{Q}_y = \bar{r} \overline{P}_y \cos \phi + \bar{r} \overline{P} \bar{\theta}_y \cos \phi \delta_b \quad (5)$$

plays the same role as the meridional gradient of quasi-geostrophic PV in the QG theory presented in Part I. The first term in (4) is large wherever meridional displacements can advect large PV gradients and is, therefore, related to Rossby wave propagation and is denoted R_{ij} . The second term of (5) only contributes to pseudomomentum at the lower boundary (denoted R_{bij}), through the first term in (4) and has a QG β -plane analogue, $\frac{1}{2} \bar{\eta}^2 \bar{r} f_0 \bar{\theta}_y = \frac{1}{2} \bar{\eta}^2 g^{-1} (f_0^2/s^2) \partial \bar{u} / \partial p$, where s^2 denotes the stratification parameter in pressure coordinates (James 1994). Thus the boundary term is directly equivalent to the contribution of Bretherton's delta function of quasi-geostrophic potential vorticity (QGPV) to the pseudomomentum (Bretherton 1966a). The other two terms in (4) are not related to PV and can contain contributions from gravity-wave activity, and will thus be denoted G_{ij} and G_{bij} .

Following the methods of Magnusdottir and Haynes (1996), a new expression for pseudoenergy can be derived for small-amplitude disturbances to the hydrostatic primitive equations, including the intersection of isentropic surfaces with the ground. The corresponding pseudoenergy inner product (divided by a to simplify later expressions) is defined as:

$$\begin{aligned} \mathcal{H}_{ij} = & \frac{1}{2a} \langle u_i, u_j \rangle + \frac{1}{2a} \langle v_i, v_j \rangle + \frac{1}{2a} \left\langle p_i, \frac{1}{\bar{r} \bar{\rho} g} \delta_b p_j \right\rangle + \frac{1}{a} \left\langle \eta_i, \frac{g(\bar{r} \bar{\theta}_y)^2}{\bar{\rho}} \delta_b \eta_j \right\rangle \\ & - \frac{1}{2} \langle \eta_i, \overline{U} \overline{Q}_y \eta_j \rangle \\ & - \left\langle \frac{-r_i}{\bar{r}}, \overline{U} \cos \phi u_j \right\rangle - \langle \eta_i, \overline{U} \bar{\theta}_y \cos \phi \delta_b u_j \rangle, \end{aligned} \quad (6)$$

where p_i is disturbance pressure, ρ is density and g is gravitational acceleration. The top line is the positive-definite wave energy and will be denoted by E_{ij} . The second line is similar to R_{ij} , but the integral is weighted by the wind $\overline{U}$ and will be called W_{ij} . The third line is similar to G_{ij} , but again weighted by $\overline{U}$ and will be denoted as F_{ij} (terms defined such that $\mathcal{H}_{ij} = E_{ij} - W_{ij} - F_{ij}$).

Choice 1. CRWs will be defined to be untilted in the zonal direction, both meridionally and vertically. In Part I, it was shown how untilted structures can be obtained by linear combinations of a growing NM and its decaying complex conjugate. Furthermore, it was shown that the evolution of any disturbance described by a superposition of one NM and its CC can alternatively be described by the evolution of the amplitude and phase of any two (different) untilted structures defined in this way. In general the disturbance PV structure can be written as:

$$P_m = a_1 e^{i\epsilon_1} P_1 + a_2 e^{i\epsilon_2} P_2 \quad (7)$$

where a_1, a_2 are the wave amplitudes, ϵ_1, ϵ_2 are their phases and P_1, P_2 are functions describing untilted structures and are, therefore, real. All other disturbance quantities will also be untilted but may be shifted in phase relative to the PV. For example, if the meridional wind is shifted by phase ϵ_v relative to the PV, then one can write:

$$v_m = (a_1 e^{i\epsilon_1} v_1 + a_2 e^{i\epsilon_2} v_2) e^{i\epsilon_v}, \quad (8)$$

where v_1 and v_2 are also real structure functions.

Assumption 3. Focus on Rossby waves. In order to make progress some assumptions must be made about the relative phases of disturbance variables. Clearly the parcel displacement η_i is in antiphase with P_i since PV is conserved and (2) applies (i.e. $\epsilon_\eta = \pi$).

The disturbance flow is assumed to be balanced and, therefore, related to the PV-anomaly distribution. In this way our attention is focused on Rossby waves. In order that isolated positive PV anomalies correspond to cyclonic centres, we set $\epsilon_v = -\pi/2$ (i.e. maximum northward flow occurs midway between cyclonic and anticyclonic centres, to the east of positive PV anomalies). Note that the structure $v_i(\phi, \theta)$ can be negative, indicating that the wind at that point is more influenced by far-field PV than the local PV anomaly.

Similarly, PV inversion can be used to argue that positive PV anomalies are associated with locally reduced mass between isentropic surfaces so that $r_i < 0$ ($\epsilon_r = \pi$). Inversion would also indicate $u_i > 0$ to the south and $u_i < 0$ to the north of each positive PV-anomaly ($\epsilon_u = 0$). Therefore, these regions would tend to cancel each other in the ‘gravity wave’ terms G_{ij} and F_{ij} , both in the interior and on the boundary. For a basic state with uniform vertical wind shear and static stability on a β -plane, these terms would be exactly zero for a Rossby wave. In contrast, Kelvin waves on an equatorial β -plane contribute zero R_{ij}, R_{bij}, G_{bij} but large positive G_{ij} to the pseudomomentum. Sakai (1989) used R_{ij} and G_{ij} to distinguish Rossby waves from gravity waves and to investigate their coupling in an ageostrophic version of the Phillips (1954) two-layer model.

It is assumed hereafter that the truncated pseudomomentum and pseudoenergy obtained by dropping the ‘gravity-wave terms’ (G_{ij} and F_{ij}) are approximately conserved. Thus, although no explicit PV inversion relation has been assumed, the neglect of these terms focuses attention on Rossby wave motions. Note that this is not the same as making the QG approximation. In particular, meridional displacements associated with ageostrophic motions can contribute to R_{ij} , whilst the effects of sphericity on geostrophic motions can contribute to G_{ij} . In the cases examined, the neglected terms are indeed found to be small (see section 3(a)). Since the remaining terms are symmetric with respect to wave labels:

$$R_{12} = R_{21}; \quad W_{12} = W_{21}; \quad E_{12} = E_{21},$$

then the total truncated pseudomomentum for a CRW pair is:

$$R_{\text{tot}} = \frac{1}{2} \langle \eta, \overline{Q_y} \eta \rangle = a_1^2 R_{11} + a_2^2 R_{22} + 2a_1 a_2 R_{12} \cos \epsilon, \quad (9)$$

where the phase difference $\epsilon = \epsilon_2 - \epsilon_1$. The truncated pseudoenergy is:

$$H_{\text{tot}} = a_1^2 (E_{11} - W_{11}) + a_2^2 (E_{22} - W_{22}) + 2a_1 a_2 (E_{12} - W_{12}) \cos \epsilon. \quad (10)$$

(b) *CRW construction from normal modes using wave-activity orthogonality*

The CRWs can be determined from the growing NM by considering their amplitudes and relative phases when phase-locked. Phase-locking is always approached at long times for any disturbance composed of a growing NM and its decaying CC. The growing NM displacement structure is tilted and is therefore complex. For example, the meridional displacement $\eta_m = \eta_r + i\eta_i$ can be related to the CRW structures by taking $\text{Re}(\eta_m)$ and $\text{Im}(\eta_m)$ where η_m is partitioned into CRWs as in (7). This relationship can be inverted to find:

$$\begin{aligned} \eta_1 &= b_1(\eta_r + e_1 \eta_i) \\ \eta_2 &= b_2(\eta_r + e_2 \eta_i), \end{aligned} \quad (11)$$

where $b_1 a_1 = \sin \epsilon_2^+ / \sin \epsilon^+$, $b_2 a_2 = -\sin \epsilon_1^+ / \sin \epsilon^+$, $e_1 = -1 / \tan \epsilon_2^+$ and $e_2 = -1 / \tan \epsilon_1^+$. Here, ϵ_1^+ and ϵ_2^+ indicate the CRW phases required to reconstruct the growing NM, and $\epsilon^+ = \epsilon_2^+ - \epsilon_1^+$ is their phase difference. In order to find the CRWs, we must determine four unknown coefficients: b_1 , b_2 , e_1 and e_2 . The wave-activity orthogonality method is an objective means to determine the coefficients from any NM structure.

Choice 2. The CRW pair is chosen to be orthogonal with respect to truncated pseudomomentum

$$R_{12} = 0, \quad (12)$$

so that the total pseudomomentum reduces to $R_{\text{tot}} = a_1^2 R_{11} + a_2^2 R_{22}$. It is clear from (4) and (5) that the pseudomomentum takes on the sign of $\overline{P_y}$ in the interior and $\overline{P} \overline{\theta}_y$ on the boundary. Thus, if R_{tot} is constant in time, the amplitude of particle displacements can only increase everywhere if $\overline{P_y}$ and/or boundary $\overline{P} \overline{\theta}_y$ take on different signs where the wave activity is located, so that the CRWs have opposite signed pseudomomentum. This is an extension of the Charney and Stern (1962) conditions for instability that is valid for the hydrostatic, primitive equations to the extent that the truncated pseudomomentum is conserved.

Note that an unstable NM on its own has zero pseudomomentum, but the superposition of a NM with its CC has non-zero pseudomomentum in general (Held 1985). From the CRW perspective, the total pseudomomentum is composed of fixed contributions from each CRW, scaled by their amplitude squared. In the phase-locked configuration $R_{\text{tot}} = 0$. Hayashi and Young (1987) interpreted instability on shear flows in the shallow-water model in terms of two ‘resonating wave components’; instability can occur when the components have opposite signed pseudomomentum, their Doppler-shifted ‘intrinsic phase speeds’ are similar and interaction is possible. Sakai (1989) showed that the necessary conditions for baroclinic instability in the ageostrophic two-layer model are essentially the same. He partitioned NMs of the full problem into ‘wave components’ that were obtained as eigenvectors of the one-layer systems where the perturbation pressure in one or the other of the layers is zero. These wave components propagate almost as they would in a single layer, but with an interaction mediated by the interface between the layers (since it is displaced by both waves). Although his method

of partition is not possible for multi-layer models, the orthogonal CRWs obtained for the two-layer model are similar to his Rossby wave components. Sakai (1989) showed that, in addition to baroclinic and Kelvin–Helmholtz instability arising from the interactions of Rossby-wave pairs and gravity-wave pairs, a mixed Rossby wave—Kelvin wave (RW–KW) instability also occurs in the two-layer model. It arises when a Kelvin wave in the upper layer ($R_{22} = 0$, $G_{22} > 0$) interacts with a Rossby wave in the lower layer ($R_{11} < 0$, $G_{11} = 0$). It is possible that RW–KW interaction occurs in the hydrostatic PEs, but it cannot be analysed using our version of CRW partition because the truncation of pseudomomentum and pseudoenergy deliberately neglects the gravity-wave terms.

It is now possible to consider the CRWs as fundamental components of a conservative system with pseudomomenta $(p_1, p_2) = (a_1^2 R_{11}, a_2^2 R_{22})$ and positions $(x_1, x_2) = (\epsilon_1/m, \epsilon_2/m)$, with a Hamiltonian equal to the truncated pseudoenergy (10).

The Hamiltonian equations

$$\frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial x_i}; \quad \frac{\partial x_i}{\partial t} = \frac{\partial H}{\partial p_i} \quad (13)$$

reduce to the CRW evolution equations (as shown in Part I):

$$\begin{aligned} \dot{a}_1 &= -\gamma_2^1 a_2 \sin \epsilon \\ \dot{a}_2 &= \gamma_1^2 a_1 \sin \epsilon \\ \dot{\epsilon}_1 &= -mc_1^1 + \gamma_2^1 \frac{a_2}{a_1} \cos \epsilon \\ \dot{\epsilon}_2 &= -mc_2^2 + \gamma_1^2 \frac{a_1}{a_2} \cos \epsilon, \end{aligned} \quad (14)$$

where the CRW interaction coefficients are $\gamma_2^1 = m(E_{12} - W_{12})/R_{11}$ and $\gamma_1^2 = m(E_{12} - W_{12})/R_{22}$. The CRW self-induced phase speeds are $c_1^1 = (W_{11} - E_{11})/R_{11}$ and $c_2^2 = (W_{22} - E_{22})/R_{22}$.

In addition, substituting (11) into the orthogonality condition gives the first equation to determine the CRW coefficients:

$$R_{rr} + e_1 e_2 R_{ii} + (e_1 + e_2) R_{ri} = 0, \quad (15)$$

where $R_{ir} = R_{ri}$ has been used.

Choice 3. CRWs are made orthogonal with respect to the wind-weighted term

$$W_{12} = 0. \quad (16)$$

As discussed in Part I, this makes the CRWs orthogonal with respect to eddy energy production to the extent that the truncated pseudoenergy is conserved. Although each CRW enables the other to grow, the total eddy-energy growth can be partitioned between the two CRWs*. Energy growth is large if $W_{11} + W_{22}$ is large and positive, implying that the zonal angular velocity and PV gradient must be positively correlated in the regions where the wave activity of each CRW is concentrated. This can be regarded as the primitive equation CRW version of the Fjørtoft (1950) condition for instability.

* CRWs are untilted and, therefore, could not contribute to eddy-energy production in isolation. However, (33) in Part I shows that eddy-energy production occurs when CRW interaction enables growth in wave amplitudes, a_1 and a_2 , provided that $W_{11} + W_{22} > 0$, from (44) of Part I.

Substituting from (11) into this second orthogonality condition gives the second equation for the CRW coefficients:

$$W_{\text{tr}} + e_1 e_2 W_{\text{ii}} + (e_1 + e_2) W_{\text{ri}} = 0. \quad (17)$$

Together (15) and (17) determine e_1 and e_2 (or, equivalently, the CRW phases ϵ_1^+ and ϵ_2^+) required to reconstruct the growing NM.

Choice 4. CRWs are normalized with respect to truncated pseudomomentum

$$R_{11} = -R_{22}, \quad (18)$$

which implies that the interaction coefficients are equal and opposite,

$$\gamma_1^2 = -\gamma_2^1 = m E_{12} / R_{22} = \sigma, \quad (19)$$

and proportional to the interaction energy (using $W_{12} = 0$). Hereafter, the CRW with negative pseudomomentum is always defined as CRW-1 so that the interaction strength σ is positive. CRW-1 will tend to have greatest displacements near the ground where the negative temperature gradient results in negative $\bar{Q}_y$, and thus will be referred to as the lower CRW.

In addition, note that the self-induced phase speeds can be expressed as

$$c_i^i = U_i - \gamma_i^i / m, \quad (20)$$

where $\gamma_i^i / m = E_{ii} / R_{ii}$ is the self-induced counter-propagation rate and

$$U_i = \frac{W_{ii}}{R_{ii}} \quad (21)$$

is referred to as the ‘home-base wind’ of CRW- i . It describes the globally integrated effect of advection by the zonal flow on this CRW.

When the CRWs are phase-locked in the NM configuration they grow at the same rate, $\dot{a}_1 = \dot{a}_2$. Equations (14) with (19) then indicate that the CRWs must have equal amplitudes, $a_1 = a_2$. Moreover, the growing NM is reconstructed with its original amplitude when $a_1 = a_2 = 1$, giving the following relations:

$$b_1 = \frac{\sin \epsilon_2^+}{\sin \epsilon^+}, \quad b_2 = -\frac{\sin \epsilon_1^+}{\sin \epsilon^+}, \quad (22)$$

completely specifying the CRWs from the growing NM structure via (11).

In a phase-locked state $\dot{\epsilon}_1 = \dot{\epsilon}_2$, so that (14) relates the phase difference to the CRW self-induced phase speeds and interaction strength

$$\cos \epsilon^+ = \frac{m(c_2^2 - c_1^1)}{2\sigma}. \quad (23)$$

There are two such states corresponding to the growing NM and its decaying complex conjugate. The growing mode has growth rate σ^+ and phase speed c_r and is obtained when the phase difference between the CRWs is in the range $0 < \epsilon^+ < \pi$. Equation (14) can be rearranged to deduce the relations:

$$\begin{aligned} \sigma &= \sigma^+ / \sin \epsilon^+ \\ c_1^1 &= c_r - (\sigma / m) \cos \epsilon^+ \\ c_2^2 &= c_r + (\sigma / m) \cos \epsilon^+. \end{aligned} \quad (24)$$

In synopsis, given a growing NM structure on a zonally-symmetric basic state we can construct two orthogonal CRWs. Their coupling describes the evolution of perturbations that can be constructed initially from that growing NM and its decaying CC. The amplitude ratio and phase difference of the CRWs changes with time, but always approaches the growing NM configuration for $\sigma t > 1$. Furthermore, it is possible to deduce the maximum possible instantaneous growth rate, σ , and the self-induced phase speeds of the CRWs directly from the NM growth rate and phase speed, once the phase difference in the locked configuration, ϵ^+ , has been determined from the orthogonality constraints.

(c) *Application of PV-thinking to orthogonal CRWs*

The chief advantage of the CRW description of the evolution is that each CRW has a simple structure, being zonally untilted, and the mechanism of both modal and non-modal baroclinic growth can be understood using PV-thinking concepts, as described by Bretherton (1966b), Hoskins *et al.* (1985) and Bishop (1993). In order to apply PV-thinking we need to be able to describe CRW evolution in terms of their zonal propagation and interaction at only two locations in the meridional plane—the CRW home-bases. This description applies exactly to CRWs obtained by the home-base method (see the introduction and Part I) if the initial conditions can be described by the superposition of a growing normal mode and its complex conjugate. The method is readily generalized to the adiabatic primitive equations by identifying the meridional displacements of air parcels with Ertel PV anomalies on isentropic surfaces and θ anomalies on the ground (2). The CRW evolution equations obtained are the same as those from the orthogonality method (14) with (20), except that the expression for the interaction coefficients is replaced by:

$$\gamma_j^i = -\frac{v_j}{\eta_i} \Big|_i^i, \quad (25)$$

where the meridional wind induced by CRW- j and the meridional displacement of CRW- i are evaluated at the home-base of CRW- i . In section 3(a) it is shown that the CRW structures and phase difference, ϵ^+ , obtained by the two methods are very similar, provided that the choice of home-bases is motivated by the PV structure of the lower orthogonal CRW, P_1 . The lower CRW's home-base is located at the ground at the latitude of maximum P_1 and the upper CRW's home-base is located where the zonal flow matches the home-wind U_2 (21), close to the $P_1 = 0$ line. The consequence is that the PV-thinking analysis of interaction applies to a good approximation to the orthogonal CRW structures.

3. CRW STRUCTURES AND THEIR COUPLING FOR REALISTIC JETS

(a) *Method for calculating CRWs in a global model*

The CRW analysis of the last section is now applied to a realistic zonally symmetric jet. The aim is to clarify the features of the flow that are important to perturbation growth, and to understand the structures of the perturbations. The jet studied in this section, Z1, was used originally by Simmons and Hoskins (1980) and more recently by THM in their LC1 experiment. Figure 1(a) shows a latitude–pressure cross-section of the basic-state potential temperature and zonal flow which are in thermal wind balance; the tropopause is also indicated. Note that the jet maximum and the maximum $|\bar{\theta}_y|$ are at 46°N and that the zonal wind is zero at the surface.

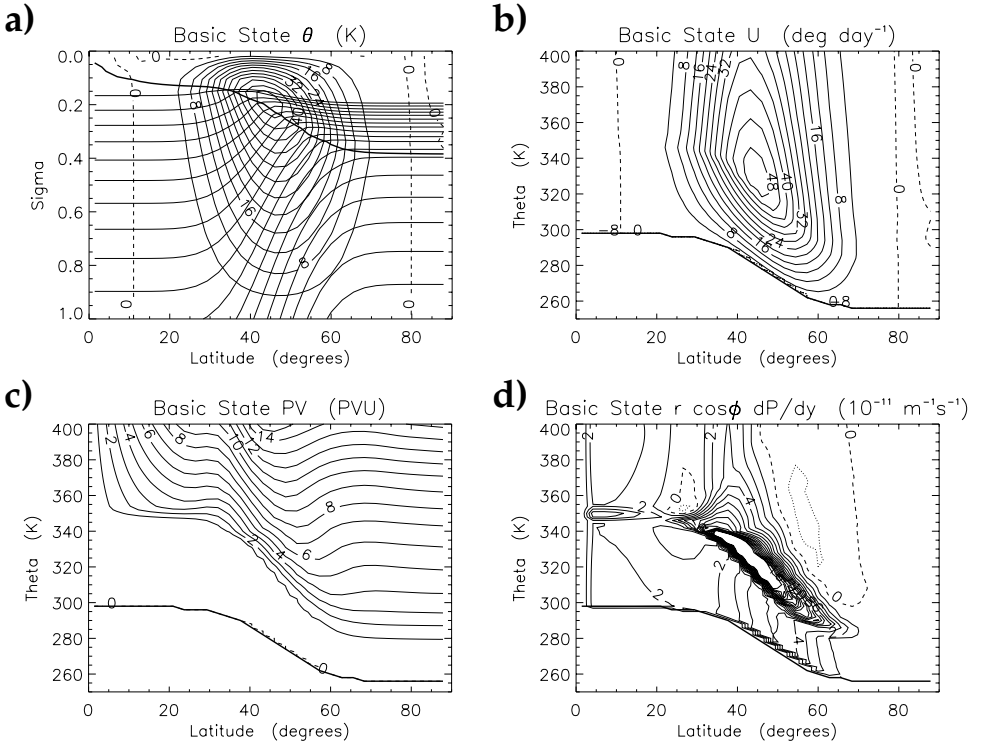


Figure 1. The basic state Z1. (a) θ (contour interval 5 K up to 350 K) and zonal angular velocity $\bar{U} = \bar{u}/(a \cos \phi)$ (contour interval 4 deg day^{-1}) in σ -coordinates. The thick line marks the 2 PVU ($= 2 \times 10^{-6} \text{ K kg}^{-1} \text{ m}^2 \text{ s}^{-1}$) tropopause. (b) $\bar{U}$ in isentropic coordinates. The heavy line marks the ground. (c) The Ertel potential vorticity (PV) (contour interval 1 PVU). (d) The isentropic PV-gradient measure $\bar{Q}_y$ (contour interval $1 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$ up to 17). In all panels, positive contours are solid, the zero contour is dashed and negative contours are dotted.

In order to perform the CRW analysis it is necessary to interpolate the basic-state fields onto isentropic surfaces using linear vertical interpolation from the σ -coordinates of the primitive-equation spectral model of Hoskins and Simmons (1975). For the CRW calculations the model has spectral truncation T42 and there are 15 levels located at $\sigma = 0.018, 0.060, 0.106, 0.152, 0.197, 0.241, 0.287, 0.338, 0.400, 0.477, 0.569, 0.674, 0.784, 0.887, 0.967$. The fields are interpolated onto isentropic surfaces with a spacing of 2 K and a maximum value $\theta_K = 400$ K. The zonal angular velocity, $\bar{u}/(a \cos \phi)$, and PV are shown in isentropic coordinates in Figs. 1(b) and (c). Note that the isentropic coordinates stretch the top of the model domain, since isentropic density, $\bar{\rho}$, falls with height from about $300 \text{ kg m}^{-2} \text{ K}^{-1}$ at the surface to about $120 \text{ kg m}^{-2} \text{ K}^{-1}$ at 2 PVU*, and then rapidly across the tropopause zone to $30 \text{ kg m}^{-2} \text{ K}^{-1}$ at about 6 PVU. The isentropic PV gradients are calculated using centred differencing of the interpolated PV field and are set to zero at the grid points closest to the lower boundary on each θ -surface, giving rise to the step-like nature of the field close to the lower boundary. Figure 1(d) shows the PV gradient term that appears in the definition of wave-activity density, $\bar{\rho} \bar{P}_y \cos \phi$. Although it has a clear maximum along the tropopause, the PV-gradient measure is also significant in the mid-latitude troposphere and stratosphere

* Potential vorticity unit: $1 \text{ PVU} = 10^{-6} \text{ K kg}^{-1} \text{ m}^2 \text{ s}^{-1}$.

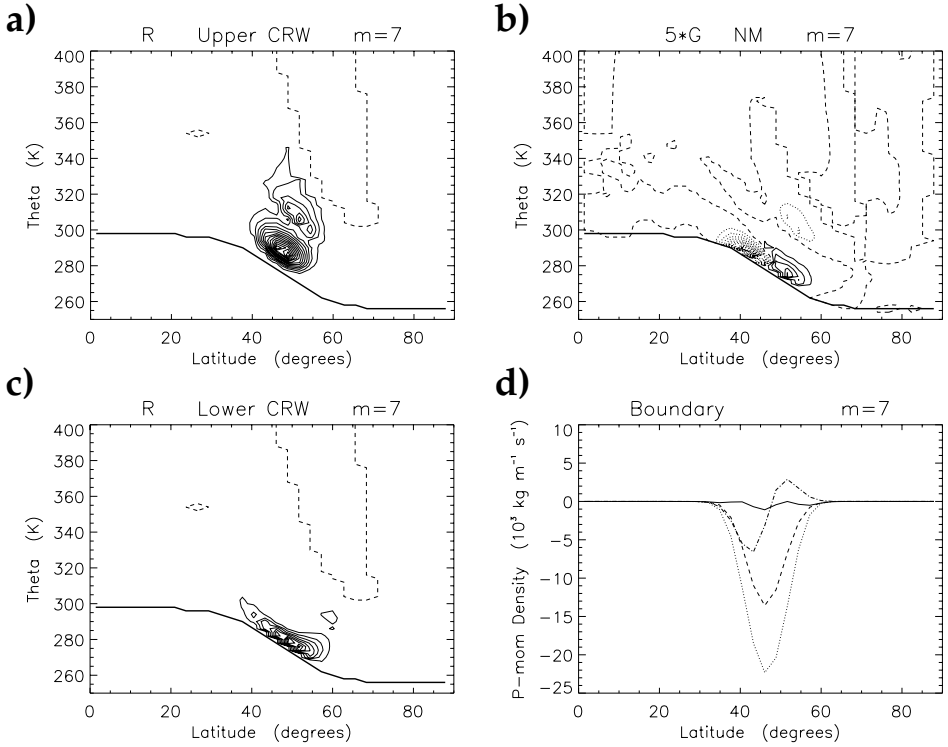


Figure 2. Wave-activity density for $m=7$ disturbances to the Z1 jet. Panels (a) and (c) show the interior pseudomomentum density, $\frac{1}{2}\eta_i^2\bar{\tau}^2\bar{P}_y\cos\phi$, for the upper and lower CRWs with an amplitude ratio appropriate for normal mode (NM) reconstruction. (b) Shows the gravity-wave component for the real part of the NM structure, $-r_r u_r \cos\phi$ (magnitude multiplied by five). (d) The boundary pseudomomentum density. The solid line shows the contribution from the upper wave, $\frac{1}{2}\eta_2^2\bar{\tau}^2\bar{P}_y\cos\phi$, the dotted line is that from the lower wave, $\frac{1}{2}\eta_1^2\bar{\tau}^2\bar{P}_y\cos\phi$, and the dashed line is for the real part of the NM, $\frac{1}{2}\eta_r^2\bar{\tau}^2\bar{P}_y\cos\phi$. The dash-dot line is the neglected G-term $\bar{\tau}\bar{\theta}_y\cos\phi\eta_r u_r$.

where its value is several times larger than the planetary-vorticity gradient ($\beta \approx 1.6 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$). There are also two regions of weak negative PV gradient in the stratosphere.

The normal modes (NMs) for this jet were obtained as the eigenvectors of the tangent-linear propagation matrix for the primitive-equation model, as described by Wyatt (1981). The complex eigenvalues obtained gave the growth rate and phase speed of each mode. For details of the method used to calculate values for the elements of the propagation matrix refer to Hoskins and Karoly (1981). In the global spectral model the complex structure functions for each NM, P_m , are obtained directly from the spectral coefficients, P_m^n , by summing over total wave number n :

$$P'(\lambda, \phi, \sigma) = P_m(\phi, \sigma) e^{im\lambda} = \left\{ \sum_{n=m}^N 2P_m^n(\sigma) L_m^n(\phi) \right\} e^{im\lambda}, \quad (26)$$

where $L_m^n(\phi)$ are associated Legendre functions. Only the structure of the fastest growing NM for each m was calculated. The displacement structure functions were found by interpolating the real and imaginary parts of the NM PV structure onto isentropic surfaces, as for the basic state, and then using (2). The inner products (3) were

calculated using Gaussian quadrature for integration over latitude and the trapezium rule for integration over θ .

As explained by Magnusdottir and Haynes (1996), wave-activity density is only defined in closed regions where particle displacements do not cross the $\bar{P}_y = 0$ lines. Within these regions there is a unique relationship between particle displacements and PV anomalies. Consequently, for their analysis of pseudomomentum fluxes in nonlinear baroclinic life cycles they modified the Z1 basic state by an adiabatic PV rearrangement so that the interior PV gradient was everywhere positive. This modification was found to have minimal effects on the life cycles, even in the small-amplitude stage. In the linear analysis here, particle displacements are arbitrarily small and we expect the pseudomomentum density either side of a $\bar{P}_y = 0$ line to cancel to a large extent in the integral (4). Therefore, it was not thought to be necessary to rearrange the basic state but, at neighbouring grid points with latitudes spanning a $\bar{P}_y = 0$ line, the displacements were set to zero in order to avoid huge values in the displacement field there. It was found that the phase-locked angle, ϵ^+ , and, therefore, the CRW structures were insensitive to this modification of the displacement field.

Although the NM amplitudes are arbitrary we have scaled them such that their maximum surface pressure perturbations are all 10 hPa. This enables comparison of different CRW structures under the same norm.

(b) *The fastest-growing structures*

Figures 2(a) and (c) show the interior pseudomomentum density for the upper and lower CRW associated with the fastest growing NM for the Z1 jet (which has $m = 7$). It is immediately apparent that both CRWs have greatest pseudomomentum density where the baroclinicity, $|\bar{\theta}_y|$, is large. Figure 2(d) shows the boundary contribution to the pseudomomentum density for each CRW. The lower CRW has a distinct minimum at 46°N where $\bar{\theta}_y$ is most negative. Its interior pseudomomentum density is also greatest in the vicinity of this latitude and is confined to the lower troposphere. The upper CRW has an almost negligible signal at the boundary and two distinct maxima in pseudomomentum density: one in the mid-troposphere and the other at the tropopause.

In order to construct orthogonal CRWs we have neglected the ‘gravity wave’ pseudomomentum density terms, G_{ij} . Its interior and boundary contributions are shown for the real part of the growing NM in Figs. 2(b) and (d). The boundary contribution has the same sign as the interior contribution near the boundary since there is not the sign reversal associated with the PV gradient terms. It is clear from the form of G that a Rossby wave that is symmetric about a reference latitude on a β -plane would be associated with an equal and opposite sign in G to either side of this latitude, and thus G would not contribute to the pseudomomentum. However, the spherical geometry lends greater weight to the equatorward lobe, resulting in net negative contributions from the interior and boundary. However, the magnitude of the integral G_{tr} is about 30 times smaller than the pseudomomentum of the upper CRW, R_{22} .

The meridional-wind structure function, v' , is positive for the upper CRW (Fig. 3(a)), consistent with the cyclonic circulation induced by its positive interior PV anomalies (Fig. 3(b)). However, the PV structure for the lower CRW is mainly negative (Fig. 3(d)) even though v' is positive (Fig. 3(c)), implying that its cyclonic circulations are centred on negative interior PV anomalies. This is because the velocity induced by the interior PV anomalies of the lower CRW is overpowered by the velocity associated with its surface potential-temperature anomalies, which are in antiphase. A poleward meridional displacement at the ground must result in both a positive boundary PV (θ)

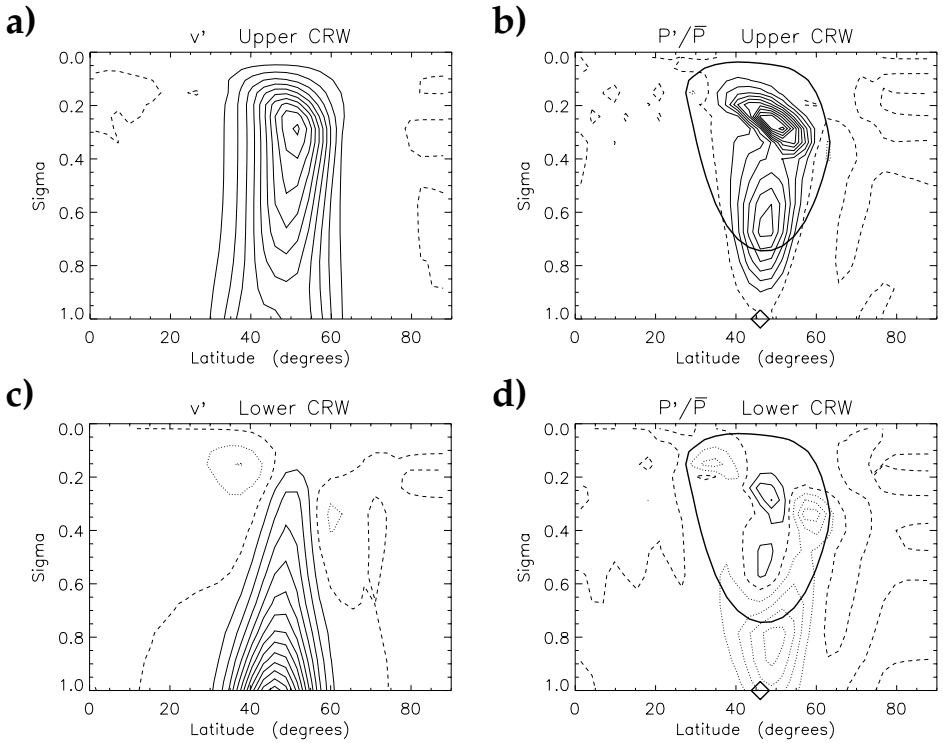


Figure 3. Counter-propagating Rossby-wave (CRW) structures in meridional wind, v' , (left-hand panels) and perturbation PV for $m = 7$ (right-hand panels), shown in σ -coordinates. Panels (a) and (b) show the upper CRW, and (c) and (d) show the lower CRW. The PV perturbations, P' , are divided by the basic-state PV, $\bar{P}$, for reasons that are discussed in the text. The contour interval corresponds to 1 m s^{-1} for v' and 0.05 for $P'/\bar{P}$ when the NM surface-pressure amplitude is 10 hPa . On these PV plots the bold curve indicates the steering surface, $U = c_r$, and the diamond shows the latitude of the maximum boundary wave-activity density of the lower CRW—see Fig. 2(d).

anomaly and an opposing negative anomaly in interior PV. The partial cancellation of the cyclonic wind induced by boundary PV by interior PV anomalies results in a rapid decay of the lower CRW's velocity with height (Fig. 3(c)) compared with the upper CRW's velocity.

Figure 4 shows vertical cross-sections along the latitude of maximum $|\bar{\theta}_y|$. The meridional wind and PV of each CRW are shown with the appropriate phase and amplitude to reconstruct the growing NM. The phase difference between the upper wave's PV anomalies and the lower wave's θ anomalies is $\epsilon^+ = 112.4^\circ$, so that the NM's velocity field tilts westwards with increasing height (see Fig. 4(e)). Since the interior PV of the lower wave is in antiphase with its surface θ anomalies, the interior PV field of the NM tilts eastwards with height (Fig. 4(f)).

Note how the PV of the upper wave (Fig. 4(b)) is almost zero on the ground, indicating that it has almost no displacement amplitude there. This in turn implies that it must have almost no boundary wave activity (as seen in Fig. 2(d)). The lower CRW (Fig. 4(d)) has one level of zero PV at $\sigma \approx 0.62$, which closely corresponds with the tropospheric maximum PV amplitude of the upper CRW. These two levels can be associated with the 'home-bases' of the CRWs. If CRWs were to be constructed from the NMs by the home-base method, the upper CRW would be defined as the untilted

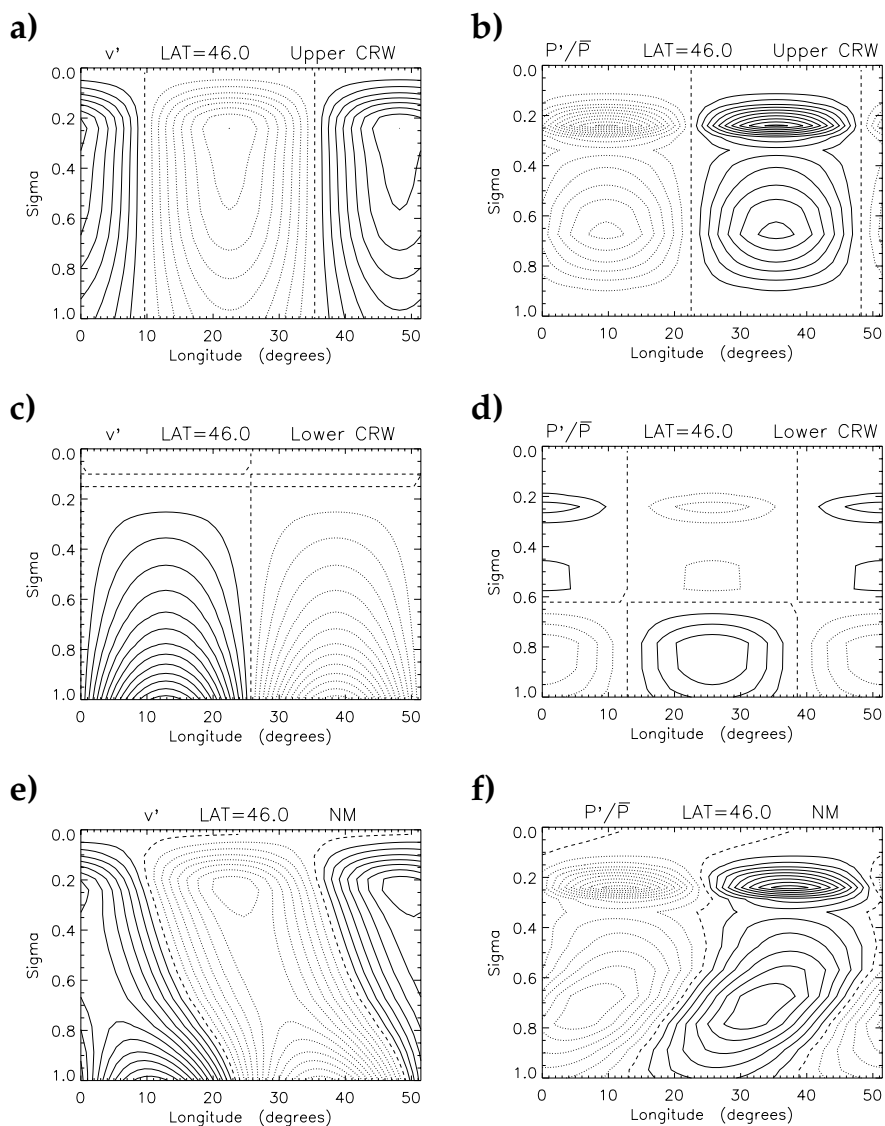


Figure 4. Zonal structure at the latitude of maximum surface $\bar{\theta}_y$ (46°N) in meridional wind (left-hand panels) and PV (right-hand panels) for $m = 7$, shown in σ -coordinates for (a) and (b) the upper CRW, (c) and (d) the lower CRW, and (e) and (f) the fastest-growing NM. Contour intervals are the same as in Fig. 3.

structure with identically zero displacement at the ground (and, therefore, positive-definite pseudomomentum). The PV of the upper CRW obtained by this method closely resembles Fig. 4(b). The lower CRW would be defined as the untilted structure with zero displacement at the level of the upper CRW's tropospheric PV maximum. As a test, two home-bases were used for the upper CRW corresponding to the model level above and below its mid-tropospheric PV maximum at 46°N ($\sigma = 0.569, 0.674$). The lower CRWs obtained were very similar to the orthogonal CRW in Fig. 4(d) but with $P_1 = 0$ lines displaced upwards and downwards, respectively. The phase-locked angle also spanned

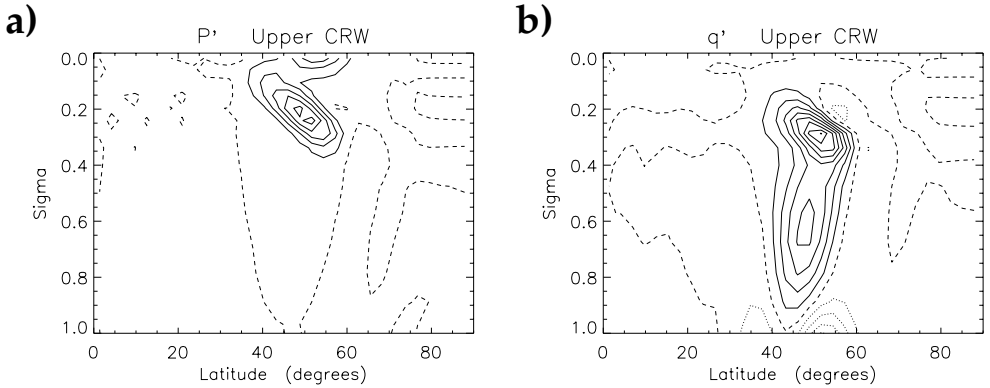


Figure 5. (a) The perturbation Ertel PV, P' , and (b) the quasi-geostrophic PV, q' , for the upper CRW (compare with Fig. 3(b) in which $P'/\bar{P}$ is shown).

the value obtained from the wave-activity (WA) orthogonality method ($\epsilon^+ = 111.4^\circ$ and $\epsilon^+ = 131.9^\circ$ from the home-base method as opposed to $\epsilon^+ = 112.4^\circ$ from WA orthogonality).

The propagation and interaction of ‘home-base CRWs’ can be described exactly by considering only the zonal flow, the PV gradient, their PV anomalies and the induced meridional wind at the two home-bases (as in Fig. 1 of Part I). Since the phase-locked angle deduced by the home-base method is very close to the result from the orthogonality method, the interaction strength and self-induced phase speeds obtained by both methods using (24) must also be similar. Therefore, the interaction between ‘orthogonal CRWs’ can be described to a good approximation by the meridional wind that their PV anomalies induce at the ground and the level where the zonal flow on the jet axis matches the home-base wind (21) of the upper CRW.

Note that $P'/\bar{P}$ has been used to display the PV anomalies in Figs. 3 and 4. By way of comparison, we show the Ertel PV anomaly P' and the QGPV anomaly q' for the upper CRW in Figs. 5(a) and (b). Although the pseudomomentum density maximum (Fig. 2(a)) is clearly located in the mid-troposphere, P' only picks out the anomaly at the tropopause. In contrast, q' , which has a known inversion relationship with geostrophic wind, picks out both pseudomomentum density maxima. Since the mid-tropospheric maximum is closer to the lower boundary and QGPV inversion is linear, it is clear that this maximum will be at least as important for the baroclinic coupling and growth as the anomaly at the tropopause. The upper CRW’s velocity structure (Fig. 3(a)) extends throughout the troposphere, which also highlights the importance of the mid-tropospheric PV anomaly. The diagnostic $P'/\bar{P}$ has been used to reduce the effects of the basic-state pseudodensity’s ($\bar{P}$) decrease with height and to highlight both maxima. Its structure (Fig. 3(b)) is similar to the QGPV structure because under QG scaling $q' = \bar{P}P' \approx f_0 P'/\bar{P}$. Since $P'/\bar{P} = (\ln P)'$ it is reasonable to suppose that $\ln P$, which is also a conserved quantity, may be a more revealing diagnostic than Ertel PV itself when examining nonlinear wave development. However, its use is restricted to regions where $P > 0$. Lait (1994) has suggested other functions of Ertel PV and θ in order to combat this problem.

(c) *Variation with zonal wave number*

(i) *Dispersion curves.* The growth rates for all the growing NMs obtained from the eigenvalue solution* for the Z1 jet fall into four distinct branches and are shown in Fig. 6(a); Fig. 6(c) shows their phase speeds. It is clear that the slower-growing branches are associated with smaller phase speeds and thus have steering levels closer to the ground. Since these modes have a smaller-scale structure they are expected to be more sensitive to dissipation, and indeed this is found to be the case. Table 1 explores the sensitivity of the growing modes for wave number 6 to a change in the strength of the diffusion coefficient. It is clear that branches 1 and 2 are quite insensitive to the dissipation but that branches 3 and 4 are very sensitive. Since the CRW theory relies on the conservation of PV, θ , pseudomomentum and pseudoenergy, and in each case the conservation is violated by dissipation, it is consistent only to look at modes that are affected weakly by the diffusion. In the following discussion of disturbances to the Z1 jet we concentrate only on the first branch.

The CRW phases in the locking configuration were obtained for all zonal wave numbers in the first branch of NMs using the orthogonality constraints (15) and (17). (24) then gives the interaction strength, σ , and self-induced phase speeds of the CRW structures. σ also quantifies the maximum instantaneous growth rate that can occur for a disturbance described by this pair of CRWs or, alternatively, the growing NM and its complex conjugate. Although super-modal growth is possible, its rate does not greatly exceed the NM growth rate (see Fig. 6(a)) since $\sin \epsilon^+$ is almost one. Although the home-base wind (21) of the upper CRW is much greater than that of the lower CRW (Fig. 6(d)), its self-induced phase speed (advection plus self-propagation) is typically less (Fig. 6(c)). Phase-locking can only be achieved if the strong westward self-propagation of the upper CRW against the eastward flow is hindered by interaction so that its phase speed, c_2^2 , increases to that of the NM, c_r . The strong eastward self-propagation of the lower CRW is reduced by the same amount so that its phase speed, c_1^1 , decreases to c_r . Hindering behaviour, together with growth, occurs whenever the phase-locked angle satisfies $\pi/2 < \epsilon^+ < \pi$, as seen for $m \leq 12$ in Fig. 6(b).

The home-wind (21) of the lower CRW, U_1 , is negative because its negative boundary pseudomomentum dominates R_{11} , while W_{11} is positive since $\bar{U} = 0$ at the ground. As the wave number decreases the Rossby scale height, $af/(Nm)$, increases and the influence of the lower CRW extends further above the lower boundary. Since $\bar{U}$ rises with height, the lower CRW's home-wind, W_{11}/R_{11} , becomes more negative (Fig. 6(d)). The upper CRW structure also shifts upwards so that its home-wind increases (both W_{22} and R_{22} are positive). At the lowest wave numbers its wave activity is concentrated near the jet maximum, which also coincides with the tropopause in the region of greatest baroclinicity.

(ii) *Long-wave structure.* As an example of long-wave behaviour, the CRW structures with zonal wave number 4 are shown in Figs. 7(a) and (c). The upper CRW (Fig. 7(a)) has two strong PV maxima just above the steering level and at the tropopause near the

* Note that only 45 modes are growing out of a total of 15×420 modes for zonal wave numbers $m \geq 2$ in the T42, L15 spectral model. This is a result of the four-hour ∇^6 hyperdiffusion that is applied to the prognostic variables (vorticity, divergence and temperature) in order to make the full model computationally stable. It is added in the linear NM calculation in order to damp unrealistic modes with structure close to the truncation limit of the model. The effect of diffusion is to subtract the same factor from the growth rate of a growing NM, σ^+ , and the growth rate of its decaying CC, $\sigma^- = -\sigma^+$. If dissipation is strong compared with the baroclinic growth then both modes will become decaying modes. Simmons and Hoskins (1977) have shown similar results for a slightly different jet.

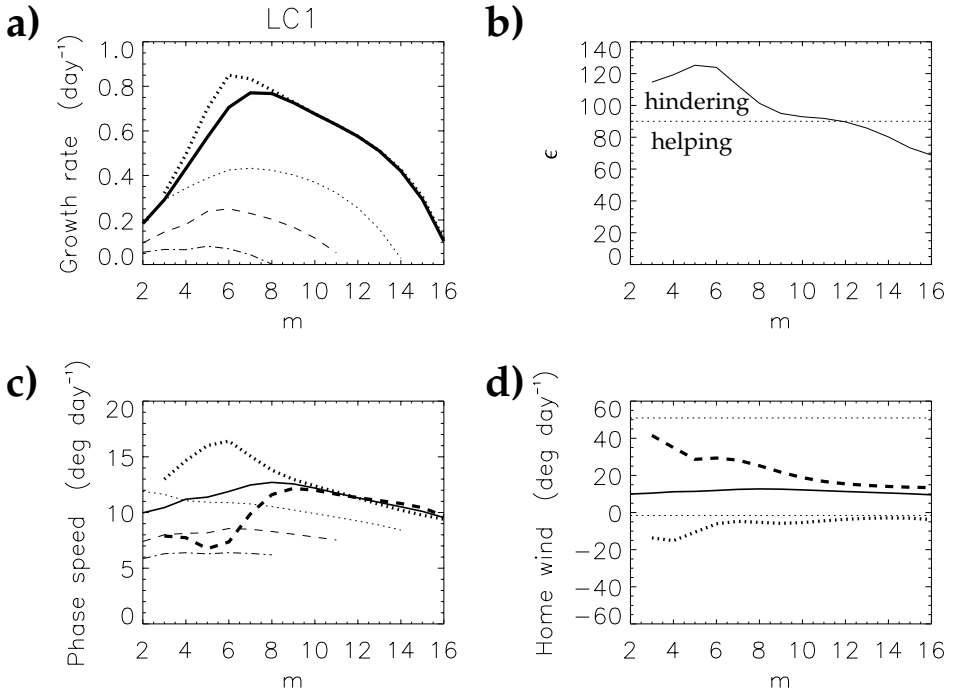


Figure 6. NM and CRW dependence on zonal wave number m for the Z1 jet. (a) The NM growth rates, σ^+ . The heavy solid line marks branch 1 (the fastest-growing modes) and the heavy dotted line shows the interaction strength, σ , for this branch. The dotted, dashed and dash-dot lines show σ^+ for branches 2, 3 and 4, respectively. (b) The CRW phase difference, ϵ^+ , when phase-locked into the growing NMs of branch 1. (c) The NM phase speeds, using the same line style for each branch as used for growth rate. The heavy dashed and dotted lines show the self-induced phase speeds of the upper and lower CRWs, respectively, for branch 1. (d) The heavy dashed and dotted lines show the home-wind of the upper CRW, U_2 , and the lower CRW, U_1 , and the solid line is the NM phase speed for branch 1. The dotted lines indicate the maximum and minimum $\bar{U}$ in the basic state.

TABLE 1. THE EFFECTS OF DIFFUSION ON $m = 6$ NORMAL-MODE PERTURBATIONS TO THE Z1 JET

Diffusion timescale (hr)	σ_1	σ_2	σ_3	σ_4	c_1	c_2	c_3	c_4
0.4	0.693	0.340	0.060	< 0	11.60	9.96	6.69	
4	0.705	0.424	0.250	0.071	11.89	10.89	8.56	6.40
40	0.706	0.442	0.302	0.174	11.94	11.13	9.22	7.63

The first four columns show the growth rate (day^{-1}) and the second four show the phase speed (deg day^{-1}) of the only growing modes for this wave number when the diffusion timescale is 4 hr. The diffusion timescale refers to diffusive decay rate of the highest retained wave number, here 42, in the spectral model.

latitude of maximum baroclinicity. The lower CRW (Fig. 7(c)) has most negative tropospheric PV at around 50°N , coincident with the maximum in PV gradient (Fig. 1(d)). Its greatest boundary wave activity (marked by the diamond) has shifted polewards of the maximum $|\bar{\theta}_y|$ (cf. $m = 7$ in Fig. 3(d)), indicating that boundary θ -anomalies become less dominant over the lower CRW's interior PV anomalies in the induction of surface wind as wavelength increases. This behaviour was crucial to the existence of westward propagating neutral planetary-scale modes in the Charney model (see Part II).

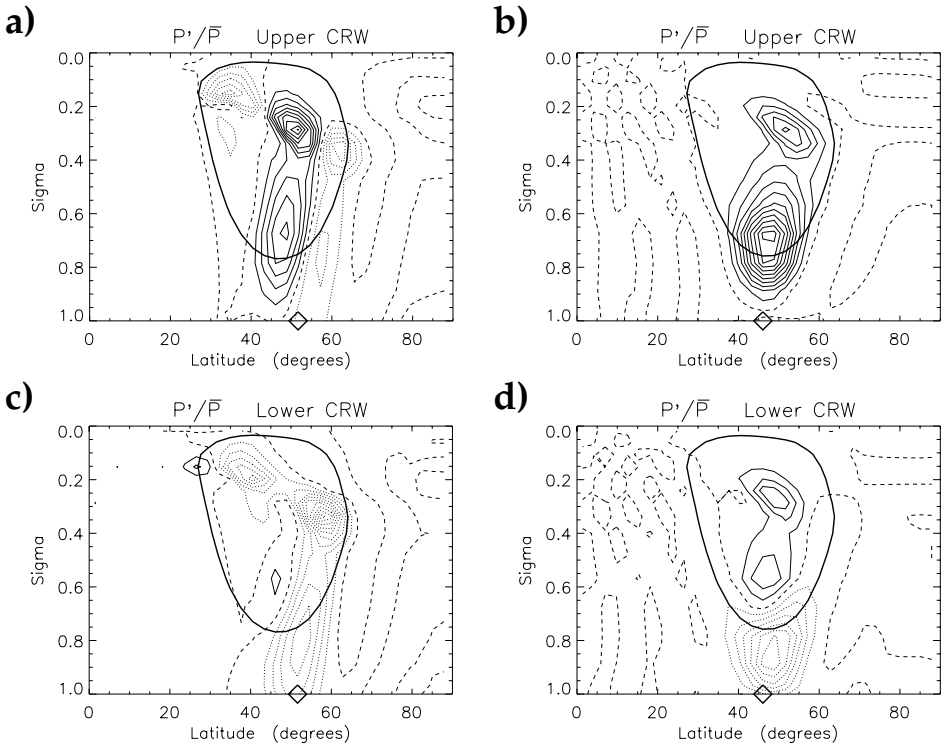


Figure 7. CRW structures for zonal wave numbers 4 and 11 from the fastest-growing branch for the Z1 basic state. Panels (a) and (c) show the PV of the upper and lower CRWs for $m = 4$, and (b) and (d) the PV of the upper and lower CRWs for $m = 11$. Contouring as for Fig. 3.

In addition, the wave number 4 upper CRW (Fig. 7(a)) has pronounced negative PV anomalies that flank its positive PV anomalies, particularly at the tropopause. These flanks are also negative in the induced meridional wind structure (not shown), implying westward Rossby wave propagation on the positive PV gradient in those regions, maintaining the untilted CRW structure. It is also clear that these negative flanks oppose the positive PV maximum, thus reducing the winds that the upper CRW induces and the strength of baroclinic interaction. Note the dramatic drop in interaction strength, σ , for $m < 6$ in Fig. 6(a), such that $m = 2$ modes are almost neutral.

(iii) *Short-wave structure.* For short-waves ($m > 7$), as the wave number increases the strength of interaction decreases (Fig. 6(a)) so that the propagation rate of each CRW is less affected by interaction. In addition, the westward self-propagation rate of upper CRW diminishes more rapidly with wave number ($\gamma_2^2/m \sim m^{-2}$) than its home-wind, or the eastward, self-propagation rate of the lower CRW ($c_1^1 \propto m^{-1}$) so that phase-locking can only occur when the PV of the upper CRW is concentrated just above the level where $\bar{U}$ matches c_1^1 (i.e. $c_2^2 \approx U_2 \approx c_1^1$). Since the NM phase speed $c_r = (c_1^1 + c_2^2)/2 \approx c_1^1$, the steering level of the NM must be located immediately below the upper CRW (Fig. 7(b)). The ‘home-wind’ of the upper CRW (Fig. 6(d)) asymptotes towards the NM phase speed as wave number increases. The lower CRW clearly becomes dominated by boundary wave activity so that $U_1 \rightarrow \bar{U}(0)$.

Figure 7(d) shows the PV structure of the lower CRW for $m = 11$. Meridional displacements are greatest near the ground in the mid-latitudes. Consequently, the

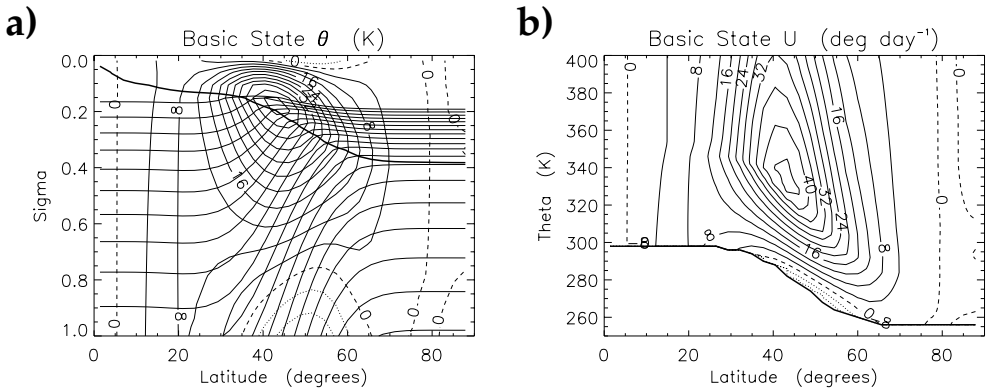


Figure 8. Basic state Z2. (a) $\bar{\theta}$ and zonal angular velocity $\bar{U} = \bar{u}/(a \cos \phi)$ in σ -coordinates, and (b) $\bar{U}$ in isentropic coordinates. Contouring as for Fig. 1.

interior PV maximum is centred at the latitude of greatest tropospheric PV gradient (49°N) but the boundary wave activity (marked by the diamond) is centred where $|\bar{\theta}_y|$ is greatest (46°N). For short waves, the mid-tropospheric PV maximum of the upper CRW (Fig. 7(b)) clearly overwhelms the PV anomalies at the tropopause, making the similarity with Charney modes transparent.

(d) The effect of surface shear on CRW structures

Here, we examine the effects of additional barotropic shear associated with surface pressure gradients on CRW structures using the Z2 jet from THM. The Z2 basic state has an identical temperature structure to that of Z1, but a meridional surface pressure gradient is imposed. Zonal wind is then obtained iteratively from temperature and surface pressure using thermal wind balance. The resulting basic state has surface westerlies centred at about 20°N and surface easterlies centred at about 50°N and is shown in Fig. 8(a). The meridional PV gradient is changed very little by the shear and its pattern closely resembles Fig. 1(d) except that the tropospheric maximum is slightly smaller.

The growing NM growth rates and phase speeds were obtained by the same method as for Z1 (using four-hour ∇^6 diffusion). In this case, five clear branches were identified containing 54 NMs in total. As for Z1, only branches 1 and 2 are insensitive to the diffusion coefficient. Branches 4 and 5 have nearly constant, low phase speeds ($c_r \approx 4.0$ deg day⁻¹, 2.0 deg day⁻¹) and the growth rate maxima are also low (0.137 day⁻¹, 0.074 day⁻¹). For clarity, only the first three branches are shown in Figs. 9(a) and (c). Note that branch 2 is the fastest growing for $m = 2, 3$ and 12 , but branch 1 is faster at all other wave numbers.

Branch 1 shares similar characteristics to those of the modes of the Z1 jet. However, phase speeds are much lower for the Z2 case because the surface zonal wind at the latitude of maximum baroclinicity is negative. The phase difference in the locking configuration, ϵ^+ , decreases with wave number (Fig. 9(b)), as expected for Charney-like modes, but even shifts into the ‘helping’ regime for $m \geq 8$, which cannot occur in the absence of surface meridional shear (explained in section 4(a)).

The lower CRWs for branch 1 of the Z2 jet shift equatorwards with wave number (see Fig. 1(a) in Part IV), resulting in an increase in U_1 (Fig. 9(d)) due to the cyclonic surface shear (Fig. 8(b)). Therefore, the decrease in the self-induced speed of the lower

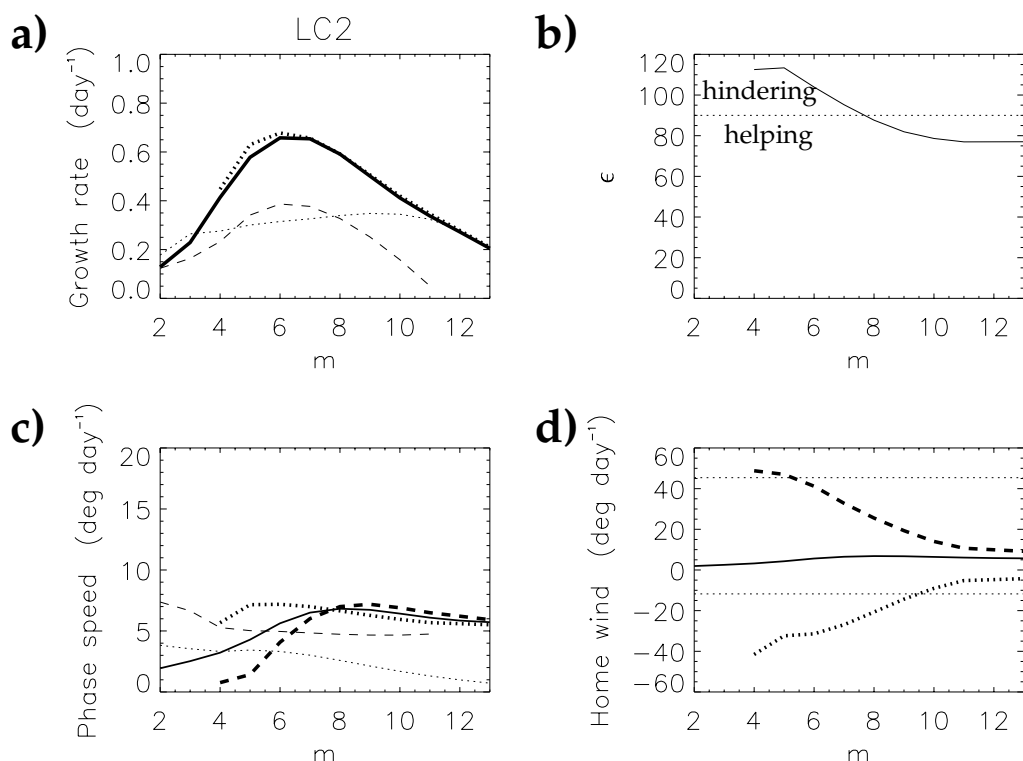


Figure 9. The NM and CRW dependence on zonal wave number m for the Z2 jet. (a) Normal-mode growth rate, (b) CRW phase difference, (c) phase speeds, and (d) home-winds of the two CRWs. Details as in Fig. 6.

CRW, c_1^1 , with wave number for $m > 6$ (the bold dotted line in Fig. 9(c)) is not as marked as for the Z1 CRWs (Fig. 6(c)). As a result, the NM phase speed barely decreases at high wave numbers. Note that the CRWs for branch 2 shift polewards with wave number, resulting in a steady decrease in phase speed with wave number (the dotted line in Fig. 9(c)).

As the wave number increases, the PV structures of the branch 1 CRWs growing on the Z2 jet slant* increasingly polewards with height. In the planetary-wave limit the additional surface shear appears to have little influence; the upper and lower CRWs for the $m = 4$ mode growing on the Z2 jet have almost identical structure to the $m = 4$ mode of the Z1 jet (Figs. 7(a) and (c)). However, Figs. 10(b) and (d) show that the slant of the PV of the upper and lower CRWs for the fast growing $m = 7$ mode is the most obvious difference from the $m = 7$ CRWs on the Z1 jet (Fig. 3(b) and (d)). Note that the positive interior PV of the upper CRW lies equatorward of the negative interior PV of the lower CRW at all heights.

Figures 11(a) and (c) show that the slant of the short waves in branch 1 is more extreme, and still the PV anomalies of the upper CRW lie equatorward of the lower CRW at all heights. In contrast, the $m = 11$ CRWs on the Z1 jet exhibit almost no slant (Figs. 7(b) and (d)).

* The term ‘slant’ is used exclusively to describe structures in the y - z plane, as opposed to ‘vertical tilt’ in the x - z plane and ‘horizontal tilt’ in the x - y plane.

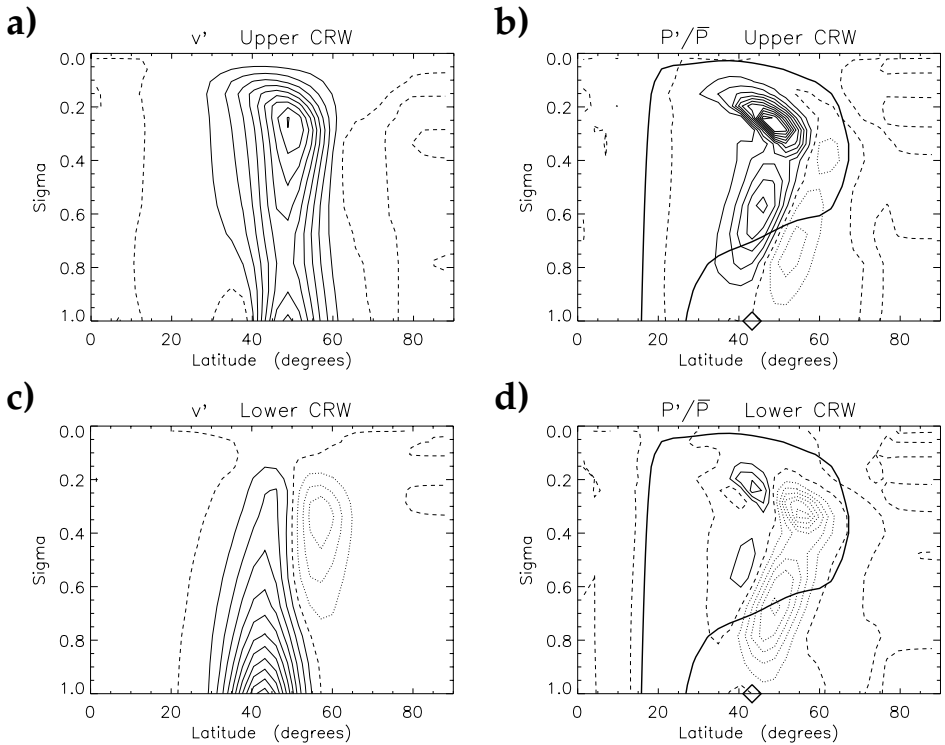


Figure 10. As Fig. 3, but for the fastest-growing mode on the Z2 jet for $m = 7$.

The slant of CRW displacement structures is influenced by three features of the basic state: the slopes of its $\bar{U}$, $\bar{P}_y$ and $\bar{\theta}$ surfaces. An explanation is as follows. The phase speed of a CRW along any zonal line is comprised of advection by the zonal angular velocity, self-induced counter-propagation (20) and an interaction term (14). As wavelength decreases the interaction term decreases most rapidly and the CRW can only support meridional displacements across a region where its self-induced phase speed, $c_i(\phi, \theta) = \bar{U} - \bar{P}_y v_i / (m P_i)$, hardly varies. In the troposphere where $\bar{P}_y$ is approximately constant with height (Fig. 1(d)), the displacement structure of CRWs will tend to be coherent along $\bar{U}$ surfaces. At the tropopause $\bar{P}_y$ is much greater, counteracting the greater zonal flow there, so the upper CRW can support structure along the tropopause too. Close to the ground, meridional displacements result in horizontal advection of θ , and air parcels must cross basic-state $\bar{\theta}$ surfaces. However, away from boundaries a displaced air parcel will tend to move parallel to basic-state isentropic surfaces, rather than across them (Hoskins 1974). Therefore, CRW meridional displacement structures tend to show coherence along $\bar{\theta}$ surfaces in the upper troposphere.

The PV structure of the upper CRW on the Z2 jet follows $\bar{U}$ contours in the lower troposphere and $\bar{\theta}$ contours in the upper troposphere (Fig. 11(a)), resulting in the strong poleward slant. For the Z1 jet the $\bar{U}$ contours are horizontal at the latitude of maximum baroclinicity, so that the CRWs only slant in the upper troposphere in sympathy with isentropic surfaces (e.g. Fig. 7(b)). The PV anomaly extremum at the tropopause naturally follows the tropopause slope equatorwards (Fig. 1(d)) above the

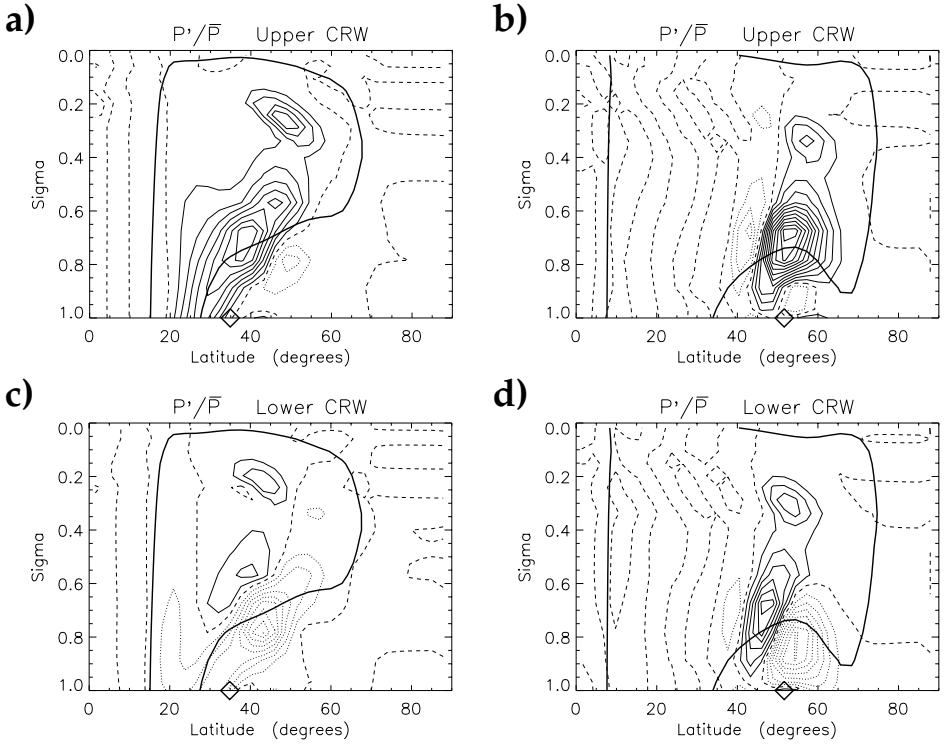


Figure 11. CRW structures of the fastest-growing modes on the Z2 jet at $m = 11$ (left panels) and $m = 12$ (right panels). Note that the $m = 11$ mode belongs to branch 1, while the $m = 12$ mode belongs to branch 2. Contouring as for Fig. 3.

maximum boundary wave activity of the lower CRW. The influence of the basic state is particularly clear for the CRW structures of the $m = 12$ mode from branch 2 growing on the Z2 jet (Figs 11(b) and (d)). Both CRWs for this branch shift polewards with wave number. The PV of the upper CRW slants with isentropic surfaces in the mid-troposphere but follows the $\bar{U}$ contours in the lower troposphere at higher latitudes where the shear is anticyclonic (Fig. 8(a)).

4. CONSEQUENCES OF CRW STRUCTURES FOR GROWING BAROCLINIC WAVES

(a) *The phase-locking configuration*

An important result of Part II, was that the CRWs of the Charney model can only phase-lock in a *hindering* configuration which corresponds to a large westward phase difference between the upper and lower CRWs ($\pi/2 < \epsilon^+ < \pi$). In the limit of zero wavelength the CRW phase difference is $\pi/2$ but, as wavelength increases, the CRWs must lock in an increasingly hindering configuration. Similar behaviour was noted for the Z1 jet, which has no surface meridional shear, but helping behaviour was seen for the short-wave branch 1 modes on the jet with additional cyclonic shear (Z2). The following argument explains how surface shear enables the short waves to lock in a helping configuration.

In the lower troposphere, the zonal flow increases roughly linearly with height, $\bar{U}(y, z) = \bar{U}(y, 0) + \Lambda z/a$, as in the Charney model. Vital to the short-wave limit is

the fact that the zonal flow at a Rossby scale height above the ground varies linearly with wavelength while the westward counter-propagation rate of the upper CRW varies as wavelength squared (or higher order), so that $c_2^2 \rightarrow U_2$ as $1/m \rightarrow 0$. In addition, the lower CRW becomes dominated by its boundary θ signature so that $c_1^1 \rightarrow \overline{U}(y_1, 0) + \alpha \Lambda/m$, where α is a dimensionless constant depending only on the surface meridional wind induced by the θ -anomalies of the lower CRW ($\alpha = -\gamma_1^1/\Lambda = [-\overline{\theta}_y/\Lambda]v_1/\theta_1 \approx [f\overline{\theta}/g]v_1/\theta_1$). In the Charney model, phase-locking occurs when $\cos \epsilon^+ = 0$ because the upper CRW is focused at the Rossby scale height $\alpha\Lambda/m$, so that $c_2^2 \approx U_2 \approx c_1^1$.

In the Z1 case the vertical wind shear varies with latitude (the jet is confined), but the latitude of the jet axis does not vary with height because there is no surface meridional shear. The CRWs of branch 1 are centred on the jet axis and the boundary wave activity of the lower CRW has a tighter structure than the jet itself (an important feature discussed further in the next section). A meridional shift of the upper CRW's home-base would result in a smaller shear between them and weaker interaction due to the increased separation. Both effects would tend to make $\cos \epsilon^+$ more negative (23) relative to the case of no meridional displacement—a more 'hindering' configuration. Thus, all CRWs are expected to be hindering when phase-locked if the jet axis does not slant.

However, wave numbers $m \geq 13$ for branch 1 on the Z1 jet are found to be in a helping configuration (Fig. 6(b)). This is thought to be spurious and due to the fact that these modes are adversely affected by the model's vertical resolution. In the Charney model the upper CRW's wave activity asymptotes towards the ground as the wavelength approaches zero but in the discretized model the level spacing prohibits this and as a result the coupling rapidly becomes weaker resulting in a rapid drop in growth rate (Fig. 6(a)). Also, the CRW counter-propagation weakens but the upper CRW cannot move closer to the lower boundary where the zonal flow is weaker (as happens in the Eady model where the vertical separation of the CRWs is fixed), resulting in $c_2^2 - c_1^1 > 0$, and therefore $\cos \epsilon^+ > 0$ (23).

In the Z2 case, a barotropic wind shear, which is cyclonic across the latitude of maximum baroclinicity (Fig. 8(a)), is superposed on the Z1 jet. The shear between the CRWs, $U_2 - U_1$, increases relative to its value in the absence of surface shear (compare Figs. 6(d) and 9(d)) because the global integrals (21) are weighted by the displacement amplitudes of each CRW and the displacement structure of the upper CRW lies equatorward of the lower CRW at every level. Therefore, in the short-wave limit $c_2^2 - c_1^1$ becomes positive, rather than zero, resulting in $\cos \epsilon^+ > 0$ (23)—a 'helping' configuration. It is natural that the displacements of the upper CRW should lie equatorward of the lower CRW's displacements at all levels because the displacement structures of both CRWs slant polewards with the $\overline{U}$ surfaces and the upper CRW is focused above the lower CRW.

(b) Horizontal momentum fluxes

An important feature of the general circulation of the atmosphere is that baroclinically growing eddies possess horizontal momentum fluxes that tend to be up-gradient, i.e. towards the latitudes where the zonal flow is greater, implying that the horizontal structure of eddies must be tilted downshear. McIntyre (1970) considered the changes to NMs that occur when the basic state of the Eady model is perturbed by non-zero interior PV gradients and barotropic meridional shear. He showed that these perturbed Eady NMs are indeed tilted downshear in the sense expected from differential advection, but then notes that, 'the result and its physical interpretation are not really obvious without

the (perturbation) analysis, since the unstable mode involves a subtle balance between advection and propagation effects.* Similar results were obtained by Stone (1969) for the two-layer model. Here, an explanation for the direction of momentum fluxes is sought in terms of CRW structures.

(i) *Vertically integrated momentum fluxes.* The zonal-mean zonal flow in isentropic coordinates evolves according to the equation (from Eq. (3.9.7) of Andrews *et al.* 1987):

$$\bar{r} \bar{u}_t + \bar{r} \bar{v} \left\{ \frac{(\bar{u} \cos \phi)_\phi}{a \cos \phi} - f \right\} + \bar{r} \bar{Q} \bar{u}_\theta = \frac{\nabla \cdot \mathbf{F}}{\cos \phi} - (\bar{r}' \bar{u}')_t + \bar{r} \bar{X}, \quad (27)$$

where Q is heating, X represents frictional effects and $\mathbf{F}$ is the generalized Eliassen–Palm flux with components $F^{(\phi)} = -\cos \phi \overline{(rv)'u'}$ and $F^{(\theta)} = (ga)^{-1} \overline{p'M'_\lambda} - \cos \phi \overline{(rQ)'u'}$. The pseudomomentum conservation law reduces for adiabatic, frictionless, small-amplitude disturbances to:

$$\frac{\partial}{\partial t} \{R + G\} + \nabla \cdot \mathbf{F} = 0, \quad (28)$$

where $R = \frac{1}{2} \bar{r} \overline{Q_y \eta'^2}$, $G = -\cos \phi \overline{r'u'} + \bar{r} \bar{\theta}_y \cos \phi \overline{\eta'u'} \delta_b = G_I + G_b$ (cf. (4))* . Integrating (28) vertically and using the result that the EP-flux on the lower boundary can be re-interpreted as the rate of change of boundary pseudomomentum (appendix C of Magnusdottir and Haynes 1996) we find:

$$\int_{\theta_b}^{\infty} \frac{(F^{(\phi)} \cos \phi)_\phi}{a \cos \phi} d\theta - F^{(\theta)}|_{\theta_b} + \int_{\theta_b}^{\infty} \frac{\partial G_I}{\partial t} d\theta = - \int_{\theta_b}^{\infty} \frac{\partial}{\partial t} \{R_I + R_b + G_b\} d\theta. \quad (29)$$

Since the left-hand side of (29), divided by $\cos \phi$, equals the vertical integral of the right-hand side of (27) with $X = 0$, it is clear that the barotropic component of the flow accelerates where the vertical integral of $\partial(R_I + R_b + G_b)/\partial t$ is negative.

Growing normal modes maintain a fixed structure while amplitude increases exponentially at rate σ^+ . Since R is quadratic in displacement, we have $\partial R/\partial t = 2\sigma^+ R$. Therefore, (29) shows that the acceleration of the barotropic flow depends on the vertical integral of R_I plus the boundary pseudomomentum density. The boundary density is dominated by R_b , which is negative given that $\bar{\theta}_y < 0$, and the vertical integral of R_I is everywhere positive, since predominantly $\overline{Q_y} > 0$ in the interior. Following Held and Andrews (1983) (their section 2) consider slowly growing modes with weak boundary wave activity. In this case, R_I dominates so that the RHS of (29) is negative, implying that there is vertically integrated EP-flux convergence (or, equivalently, momentum-flux divergence), assuming that the G_I term in (29) is small. The wave activity density, R_I , is greatest on the jet, since $\overline{Q_y}$ is largest there, so that (27) implies that the barotropic component of the jet will decelerate.

However, Held (1985) showed that the global pseudomomentum of a growing normal mode is zero. Consequently, if the boundary wave activity is focused under the jet, but the interior wave activity is spread more widely along surfaces of constant $\bar{U}$, it must be the case that the right-hand side of (29) is positive at the latitude of maximum R_b and negative to either side so that the global integral is zero. In this case, the jet would tend to accelerate as a result of vertically integrated momentum-flux convergence there (Fig. 12(a)).

* The definitions for $\mathbf{F}$ in Andrews *et al.* (1987) and pseudomomentum density from Magnusdottir and Haynes (1996) have been divided by a so that $R/\bar{r}$ equals pseudomomentum per unit mass (m s^{-1}).

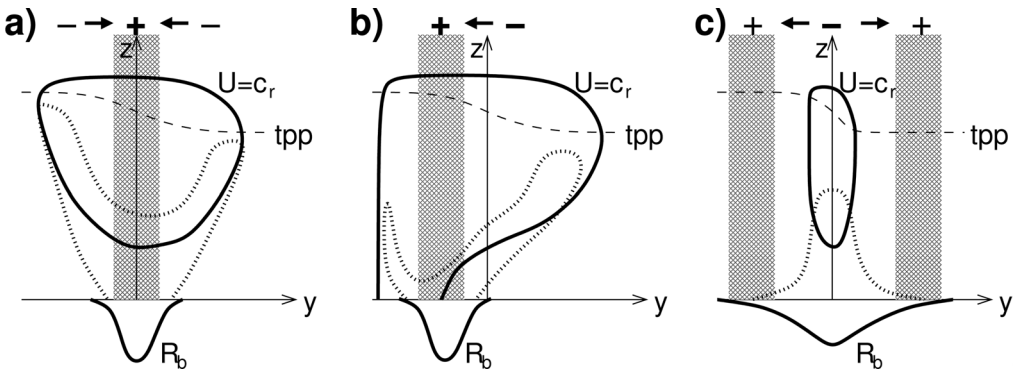


Figure 12. Schematics of jet structure and implications for CRW displacement structure and momentum fluxes in three cases: (a) broad jet ($r > 1$) with no surface meridional shear, (b) broad jet with cyclonic surface shear and (c) narrow jet ($r < 1$) with no surface shear. The steering surface $\bar{U} = c_r$ outlines the jet shape and the tropopause (tpp) is also drawn. R_b denotes boundary wave-activity density and has a narrow structure relative to jet width in (a) and (b). The dotted line indicates the displacement structure for the lower CRW of a fast-growing NM. The hatched regions highlight where boundary wave-activity density dominates the vertical integral of interior wave-activity density (for the growing NM) such that their sum is negative. This implies that the right-hand side of (29) is positive (marked by the +) and there is depth-integrated momentum-flux convergence, as indicated by the arrows.

The tight structure of boundary wave activity relative to the width of the jet is essential in the argument above. Juckes (2000) developed a theory for the meridional scale, b , of the meridional displacement structure by considering the Eady problem modified by a slow variation in $\bar{\theta}_y$ with latitude, prescribing the width of the jet. Heuristic arguments and asymptotic expansions led to the result $b \sim r^{1/2}$ where r is the jet width normalized by the Rossby radius of deformation. Therefore, when $r > 1$ we expect $b < r$, i.e. the wave structure is narrow relative to the jet width. This is indeed the regime relevant to the Z1 and Z2 jets because they are broad relative to the Rossby radius. Figure 12 illustrates the generic structure expected for the lower CRW of a fast-growing mode and its implications for the vertically integrated momentum flux in cases of broad and narrow jets and on the addition of surface meridional shear. The general rule that growing NMs have upgradient momentum fluxes on broad jets ($r > 1$) but downgradient fluxes on narrow jets was shown by Held and Andrews (1983) and Fig. 12 illustrates why this occurs (although some exceptions exist).

(ii) *Momentum fluxes along isentropic surfaces.* The tendency for the air-parcel displacement structures of CRWs to be coherent along $\bar{U}$ lines (when the jet is broad, $r > 1$) has implications for the horizontal momentum fluxes of NMs on isentropic surfaces in addition to the vertical integral.

First, consider the structure of CRWs on broad jets without a slanting axis (e.g. Z1). The boundary displacements of the lower CRW are focused close to the latitude of maximum baroclinicity ($|\bar{\theta}_y|$) with scale smaller than the jet width (e.g. Fig. 2(d)). Poleward meridional displacements near the ground are associated with positive θ -anomalies and negative interior PV anomalies (P_1). The negative anomalies extend upwards on the jet axis to the curve $P_1 = 0$, which approximately coincides with the home-base of the orthogonal upper CRW, where $\bar{U} = U_2$ (Fig. 3). The coincidence is exact by construction for CRWs obtained by the home-base method. Since $U_1 < c_r < U_2$ for growing modes, the negative PV anomalies must cross the steering level. The tendency for meridional displacements to be coherent along $\bar{U}$ contours, as argued in section 3(d),

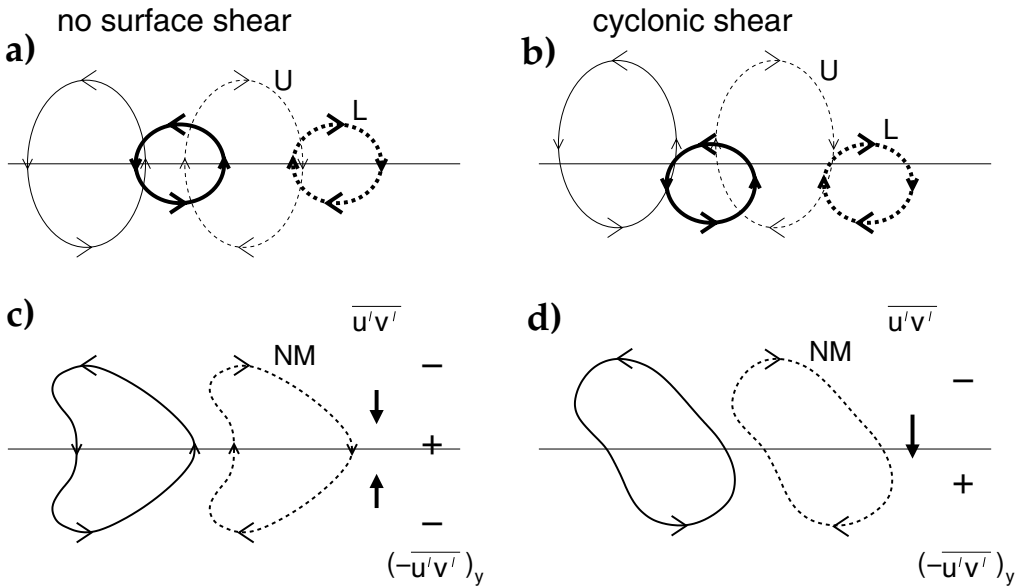


Figure 13. Illustrations of growing normal-mode (NM) eddy shape resulting from the westward phase difference between its upper (U) and lower (L) CRW. On a jet without surface meridional shear (panels (a) and (c)), both CRWs are centred on the jet axis but the circulation anomalies of the lower CRW are meridionally confined by its negative PV flanks. On adding cyclonic surface shear (panels (b) and (d)), the upper CRW's circulation anomalies shift polewards relative to those of the lower CRW. The sign of the NM's horizontal-momentum flux and its convergence are determined by the resultant eddy tilt.

results in negative PV anomalies extending along the curve $\overline{U} = c_r$ (e.g. Fig. 3(d)). Since the jet is broad, the negative PV flanks slant outwards from the surface activity to the tropopause where the PV gradient is larger and, therefore, the same displacement can advect a stronger PV anomaly. The negative PV anomalies oppose the positive boundary θ -anomaly such that the v -structure of the lower CRW (Fig. 3(c)) falls much more rapidly with height than the upper CRW (Fig. 3(a)) and is confined meridionally by its negative PV flanks.

Figure 13(a) depicts the horizontal circulation induced by each CRW. The two CRWs are centred on the jet axis but the circulation anomalies of the lower CRW are more confined. Furthermore, the upper CRW must be positioned to the west of the lower CRW for baroclinic growth. Since CRWs are untilted, in isolation each is associated with zero horizontal momentum flux. However, when combined the resulting eddy must tilt down-shear giving rise to momentum fluxes which converge on the jet axis (i.e. up-gradient fluxes). The momentum-flux structure for the fastest-growing wave number 6 NM is shown in Fig. 7(b) of Simmons and Hoskins (1977) for a jet very similar to Z1. They noted that there is a poleward bias in the NM momentum fluxes that was shown by Hollingsworth *et al.* (1976) to result from spherical geometry. This bias can be viewed as a distortion of the circulation anomalies of each CRW so that they are broader towards the equator (but still untilted).

The existence of PV flanks associated with the lower CRW can be deduced from previous results in the literature. McIntyre (1970) perturbed the Eady basic state with the weak jet, $\overline{U} = z + \mu \sin^2 \pi y$ ($\mu < 1$), and showed that the meridional θ -fluxes of the NM (his Fig. 5) are more meridionally confined at the ground than at the upper lid.

Davies *et al.* (1991) examined baroclinic growth on jets in a semi-geostrophic model with uniform interior PV and a rigid lid (similar to the Eady model). The jets were sufficiently strong that the steering surface $\overline{U} = c_r$ intersected the upper lid either side of the jet, although still broad in the sense that $r > 1$. Their Fig. 6(a) clearly shows that NM θ -fluxes on the ground are confined close to the jet axis, but on the upper lid they extend out across the 'critical lines' $\overline{U} = c_r$. Although displacements in the interior are not associated with PV anomalies because the interior PV gradient is zero in these models, on the boundaries poleward advection results in positive θ -anomalies which are equivalent to positive boundary PV anomalies at the ground and negative PV anomalies at the lid (Bretherton 1966a). The displacements at the upper lid associated with the flanks of the lower CRW result in negative PV anomalies either side of the jet and also broaden the positive NM θ -flux, $\overline{v'\theta'} = a_1 a_2 \sin \epsilon \{v_2 \theta_1 - v_1 \theta_2\}$ (from Eq. (20) in Part I), since $\theta_1, v_1, v_2 > 0$ but $\theta_2 < 0$. At the ground the upper CRW has no displacements ($\theta_2 = 0$) so that the θ -flux is confined close to the jet axis where $\theta_1 > 0$. The flanks of the lower CRW at the upper lid act as the analogue of the negative PV flanks at the tropopause in more realistic models (e.g. the lower CRW shown in Fig. 3(d)).

When cyclonic surface shear is added (the Z2 jet), the displacement structures of both CRWs are clearly coherent along $\overline{U}$ -surfaces, especially for short waves (Fig. 11) where the interaction strength is weaker and, therefore, has less impact on the propagation against the flow. The interior PV of the lower CRW is negative below the steering level and in the flanks, which lie just above the steering level either side of the jet. Therefore, the v -structure of the lower CRW is again more confined than that of the upper CRW, both vertically and meridionally (Fig. 10).

The poleward slant of the upper CRW's displacements also has ramifications. Its meridional displacements are now non-zero at the ground either side of the maximum boundary wave activity of the lower CRW. On the equatorward side this results in a negative θ -anomaly which opposes the positive interior PV above resulting in almost zero induced wind at the boundary there ($\phi < 40^\circ\text{N}$ in Fig. 10(a)). On the poleward side the upper CRW gains a weak positive θ -anomaly which induces circulation in the same sense as the interior PV anomalies slanting overhead. The net effect is that the upper CRW's v -structure is upright, but shifted polewards of the lower CRW's v -structure at every level.

The schematic in Fig. 13(b) depicts the circulation anomalies of the two CRWs under the influence of cyclonic shear. Again, the lower CRW is more meridionally confined but, in addition, the upper CRW has been shifted polewards. Although each CRW has zero momentum flux in isolation, when combined with the upper CRW to the west, the resulting eddies tilt from north-west to south-east, which gives rise to an equatorward momentum flux and convergence where the barotropic component of the zonal flow is greatest.

Similar behaviour can be deduced for the model of Davies *et al.* (1991) when uniform barotropic cyclonic shear is added. Their Fig. 6(b) shows how the NM θ -fluxes at the ground are confined close to the latitude of maximum baroclinicity but shift slightly equatorwards, reflecting the influence of the upper CRW's θ -anomaly at the ground. In contrast, the NM θ -fluxes at the upper lid are much broader (extending across the steering lines) and are strongly biased towards the poleward side, reflecting the θ -anomalies associated with the flanks of the lower CRW and the poleward shift of the upper CRW's wind anomalies.

Importantly, Davies *et al.* (1991) show that perturbations have symmetry with respect to change in the sign of surface shear in their semigeostrophic model, such that they are reflected about the jet axis. Thus we expect that anticyclonic shear

encourages an equatorward slant in the displacement structures of CRWs (especially when interaction strength is weak), inducing an equatorward shift in the upper CRW's v -structure relative to the lower CRW and, therefore, south-west to north-east tilt to the resulting eddies, and a poleward momentum flux. Again, this flux would be in the direction of increasing barotropic wind, i.e. up-gradient.

The up-gradient momentum fluxes are a characteristic of baroclinic growth. Barotropic growth results in down-gradient momentum fluxes, as can be seen from the following CRW argument. Consider a strong eastward flow positioned to the south of a weaker eastward flow, separated by a linear shear zone with positive vorticity (the Rayleigh problem). CRW-1 propagates eastwards on the northern edge of the strip where the vorticity gradient is negative ($|\overline{Q}_y|^1 < 0$) while CRW-2 propagates westwards on the southern edge where the vorticity gradient is positive ($|\overline{Q}_y|^2 > 0$). Clearly $U_2 - U_1 > 0$ satisfying the CRW-form of the Fjørtoft condition for instability. CRW theory tells us immediately that, in order to grow, CRW-2 must be shifted to the west of CRW-1 and, therefore, that barotropically growing eddies must have a south-west to north-east tilt, giving rise to a northward momentum flux towards the weaker zonal flow. A further corollary is that the momentum fluxes associated with baroclinic growth must also contribute to energy conversions that are characteristic of barotropic decay and could be interpreted as slowing baroclinic growth—this was referred to as the ‘barotropic governor effect’ by James (1987).

5. CONCLUSIONS

The CRW perspective on baroclinic instability has been generalized to describe the linearized dynamics of disturbances to steady zonally-symmetric flows governed by the adiabatic frictionless primitive equations on the sphere. The evolution of any disturbance that can be described by the superposition of a growing normal mode (NM) and its decaying complex conjugate can alternatively be viewed as the zonal propagation and interaction of a pair of untilted CRWs. The only approximation required is that a truncated form of pseudomomentum and pseudoenergy should be globally conserved, which focuses attention on Rossby waves. The global integral of the terms neglected during the truncation was about 3% of the pseudomomentum of a CRW (for the fastest-growing mode).

The CRWs are obtained objectively from the NM by defining them to be zonally untilted structures, and then imposing the constraints that they are orthogonal with respect to pseudomomentum and eddy energy production. They can be viewed as components of a Hamiltonian system with separate zonal positions and (pseudo)momenta. Instability is possible when the CRWs have opposite signs of pseudomomentum and their self-induced phase speeds are sufficiently similar that modification of zonal propagation by interaction can achieve phase-locking (condition (51) of Part I quantifies how similar). These are the same conditions as described by Hayashi and Young (1987) for the shallow-water model and Sakai (1989) for the ageostrophic two-layer model.

The evolution of these ‘orthogonal CRWs’ can be described to a good approximation by the zonal propagation and interaction characteristics of the untilted wave structures at only two locations in the meridional plane, called their ‘home-bases’, even though the NMs are complicated three-dimensional structures and the basic state varies in latitude and height. Discussion of their phase-locking and mutual growth rate reduces to consideration of their PV (and boundary θ) structures and v induced by each CRW at the two home-bases, as well as $\overline{U}$ and $\overline{Q}_y$ (see (5)) at these two locations.

Instability is only possible if $\overline{Q}_y$ has opposite signs at the home-bases and if $\overline{U}$ is greater at the home-base where $\overline{Q}_y > 0$. These are CRW forms of the Charney and Stern (1962) and Fjørtoft (1950) necessary conditions for baroclinic instability that are valid for the primitive equations, to the extent that the truncated pseudomomentum and pseudoenergy are conserved.

Since poleward advection near the ground forms both a positive θ -anomaly and negative interior PV-anomaly (unless the interior PV gradient is zero), the lower CRW must always have interior PV anomalies that act to oppose its boundary PV anomalies in the induction of meridional wind and confine its wind structure closer to the ground. The negative PV-anomaly extends up to the home-base of the upper CRW (since the lower CRW's PV goes to zero there) and thus crosses the steering level. Therefore, it is important to associate part of the interior PV, especially in the lower troposphere, with the boundary θ -wave when attempting to attribute baroclinic development to the interaction of upper and lower CRWs. Hoskins and Berrisford (1988) and Davis and Emanuel (1991) have also shown that lower-tropospheric PV anomalies associated with non-conservative effects can be important in the interaction. Positive PV is produced in the warm sector during frontal occlusion, which induces cyclonic circulation adding to that associated with the positive θ -anomaly. Note that the occlusion process always results in interior PV production, even in dry models, where the surface θ -anomalies are dissipated (with cancelling effects on the strength of circulation anomalies) but that latent-heat release from condensation associated with large-scale ascent enhances the interior PV production rate in the warm sector (e.g. Fig. 13 of Ahmadi-Givi *et al.* 2004).

Note that it is important that $P'/\overline{P}$ is used to examine the PV structure of CRWs rather than the Ertel PV perturbation, P' . $P'/\overline{P}$ is proportional to QGPV under QG scaling and its structure more closely resembles wave-activity density. P' gives misleading weight to perturbations of the tropopause which is not reflected in the induced circulation anomalies. This suggests that $\ln P$ would generally be a better diagnostic in the extratropics than P . Also, CRWs of realistic jets with synoptic wavelengths, or shorter, clearly resemble Charney modes much more closely than Eady modes, since the mid-tropospheric PV gradient is crucial to their growth and phase speeds.

For broad jets without a slanting axis (e.g. Z1), CRW behaviour for wavelengths shorter than the fastest growing NM ($m > 7$) resembles that of the Charney modes. As wavelength decreases, interaction weakens and the westward self-propagation rate of the upper CRW becomes negligible compared with advection by its home-base wind, so that phase-locking is achieved when $c_2^2 \approx U_2 \approx c_1^1$. The phase-locked angle ϵ^+ tends to $\pi/2$ (from (23)). Conversely, as wavelength increases, the westward self-propagation rate of the upper CRW becomes stronger relative to eastward advection at its home-base, so that the CRW interaction must increasingly 'hinder' the self-propagation for phase-locking to occur. This is achieved by an increasing phase difference between the CRWs. Thus the CRWs must be in a hindering configuration ($\pi/2 < \epsilon^+ < \pi$) at all wavelengths.

Consequently, when the jet axis is almost vertical, cases where an upper-level positive PV-anomaly phase-locks above a positive θ -anomaly on the boundary (almost no westward tilt) cannot be described by the evolution of a single CRW pair. Such cases could be described by the aggregation of CRWs in wave packets (Davies and Bishop 1994), or by the dynamics of the 'continuous spectrum modes' that are neutral but can result in finite-time baroclinic growth (Farrell 1982). Their interaction with the CRWs of the discrete (unstable normal mode) spectrum is under investigation.

The displacement structures of CRWs generally slant in sympathy with surfaces of constant $\overline{U}$ if the jet is broad relative to the boundary displacement structure, which occurs if the jet is wider than the Rossby radius (Jukes 2000). This has important

implications for the structure of baroclinic waves growing on broad jets. Pseudomomentum conservation was used to show that vertically integrated EP-flux divergence only occurs at latitudes where boundary pseudomomentum density dominates the vertical integral of interior pseudomomentum density. Therefore, momentum-flux convergence and acceleration of the barotropic component of the jet occurs above the maximum in boundary wave activity of the NM and compensating divergence occurs on either side (see Fig. 12). Moreover, the slant of the lower CRW displacement structure around the flanks of the jet is associated with negative PV anomalies there. The wind anomalies induced by the lower CRW are meridionally confined by its negative PV flanks. Consequently, the horizontal momentum fluxes of baroclinically growing eddies typically converge on the jet axis at all levels (see Fig. 13).

If additional cyclonic barotropic shear is imposed on a jet (e.g. Z2), it results in poleward slanting $\overline{U}$ surfaces and encourages a poleward slant in CRW displacement structures. Consequently, on horizontal surfaces the displacement anomalies of the upper CRW lie equatorward of those of the lower CRW. This enhances the difference between the 'home-base winds' so that $c_2^2 - c_1^1$ can become positive as wavelength decreases, enabling phase-locking of short waves in a 'helping' configuration ($0 < \epsilon^+ < \pi/2$).

Due to the poleward slant, the maximum of interior wave activity is displaced polewards of the boundary wave activity so that there is vertically integrated momentum-flux convergence on the equatorward side of the jet maximum and divergence on the poleward side, accelerating the barotropic component of the jet. Also, the meridional-wind anomalies of the upper CRW lie poleward of those of the lower CRW at all levels. Growing eddies described by the CRW pair must tilt from north-west to south-east, and the horizontal momentum flux is negative on all levels.

Therefore, CRW theory provides an explanation for why eddies growing baroclinically on broad jets typically tilt down-shear in the horizontal and, therefore, have up-gradient momentum fluxes (Stone 1969; McIntyre 1970; Held and Andrews 1983). The crucial ingredients are that the displacement structures of CRWs tend to be coherent along $\overline{U}$ -surfaces, the lower CRW has a narrow structure at the ground (relative to jet width) but must have negative interior PV flanks following the steering surface around the jet, and the upper CRW must be shifted to the west of the lower CRW for growth.

ACKNOWLEDGEMENTS

John Methven is grateful for an Advanced Fellowship sponsored jointly by the Natural Environment Research Council and the Environment Agency. Craig Bishop acknowledges support from National Science Foundation grants ATM-96-12502 and ATM-98-14376. Support from the Office of Naval Research grant N00014-00-1-0106 and Program Element 0601153N Project number BE-0333-0345 are also gratefully acknowledged. This work would not have been possible without the tools developed by the UK Universities Global Atmospheric Modelling Programme (UGAMP) and the expertise of Mike Blackburn. Thanks to Martin Jukes and the referees for their insightful reviews which resulted in substantial improvements to Section 4.

REFERENCES

- | | | |
|--|------|--|
| Ahmadi-Givi, F., Craig, G. C. and Plant, R. S. | 2004 | The dynamics of a mid-latitude cyclone with very strong latent-heat release. <i>Q. J. R. Meteorol. Soc.</i> , 130 , 295–323 |
| Andrews, D. G. | 1983 | A finite-amplitude Eliassen–Palm theorem in isentropic coordinates. <i>J. Atmos. Sci.</i> , 40 , 1877–1883 |
| Andrews, D. G., Holton, J. R. and Leovy, C. B. | 1987 | <i>Middle atmosphere dynamics</i> . Academic Press |

- Balasubramanian, G. and Garner, S. 1997 The role of momentum fluxes in shaping the life cycle of a baroclinic wave. *J. Atmos. Sci.*, **54**, 510–533
- Bishop, C. H. 1993 On the behaviour of baroclinic waves undergoing horizontal deformation. Part 2: Error bound amplification and Rossby wave diagnostics. *Q. J. R. Meteorol. Soc.*, **119**, 241–268
- Bretherton, F. P. 1966a Critical layer instability in baroclinic flows. *Q. J. R. Meteorol. Soc.*, **92**, 325–334
- 1966b Baroclinic instability and the short wavelength cut-off in terms of potential vorticity. *Q. J. R. Meteorol. Soc.*, **92**, 335–345
- Charney, J. G. 1947 The dynamics of long waves in a baroclinic westerly current. *J. Meteorol.*, **4**, 135–163
- Charney, J. G. and Stern, M. E. 1962 On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172
- Davies, H. C. and Bishop, C. H. 1994 Eady edge waves and rapid development. *J. Atmos. Sci.*, **51**, 1930–1946
- Davies, H., Schär, C. and Wernli, H. 1991 The palette of fronts and cyclones within a baroclinic wave development. *J. Atmos. Sci.*, **48**, 1666–1689
- Davis, C. and Emanuel, K. 1991 Potential vorticity diagnostics of cyclogenesis. *Mon. Weather Rev.*, **119**, 1929–1953
- Eady, E. T. 1949 Long waves and cyclone waves. *Tellus*, **1**, 33–52
- Farrell, B. 1982 The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663–1686
- Fjørtoft, R. 1950 Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex. *Geofys. Publ.*, **17**, 1–52
- Hayashi, Y.-Y. and Young, W. 1987 Stable and unstable shear modes of rotating parallel flows in shallow water. *J. Fluid Mech.*, **184**, 477–504
- Haynes, P. 1988 Forced, dissipative generalisations of finite-amplitude wave-activity conservation relations for zonal and non-zonal basic flows. *J. Atmos. Sci.*, **45**, 2352–2362
- Heifetz, E., Bishop, C. H. and Alpert, P. 1999 Counter-propagating Rossby waves in the barotropic Rayleigh model of shear instability. *Q. J. R. Meteorol. Soc.*, **125**, 2835–2853
- Heifetz, E., Bishop, C., Hoskins, B. and Methven, J. 2004a The counter-propagating Rossby wave perspective on baroclinic instability. Part I: Mathematical basis. *Q. J. R. Meteorol. Soc.*, **130**, 211–231
- Heifetz, E., Methven, J., Hoskins, B. and Bishop, C. 2004b The counter-propagating Rossby wave perspective on baroclinic instability. Part II: Application to the Charney model. *Q. J. R. Meteorol. Soc.*, **130**, 233–258
- Held, I. M. 1985 Pseudomomentum and the orthogonality of modes in shear flow. *J. Atmos. Sci.*, **42**, 2280–2288
- Held, I. M. and Andrews, D. G. 1983 On the direction of the eddy momentum flux in baroclinic instability. *J. Atmos. Sci.*, **40**, 2220–2231
- Hollingsworth, A., Simmons, A. and Hoskins, B. 1976 The effect of spherical geometry on momentum transports in simple baroclinic flows. *Q. J. R. Meteorol. Soc.*, **102**, 901–911
- Hoskins, B. 1974 The role of potential vorticity in symmetric stability and instability. *Q. J. R. Meteorol. Soc.*, **100**, 480–482
- Hoskins, B. J. and Berrisford, P. 1988 A PV perspective on the storm of 15–16 October 1987. *Weather*, **43**, 122–129
- Hoskins, B. and Karoly, D. 1981 The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179–1196
- Hoskins, B. and Simmons, A. 1975 A multi-layer spectral model and the semi-implicit method. *Q. J. R. Meteorol. Soc.*, **101**, 637–655
- Hoskins, B. J., McIntyre, M. E. and Robertson, A. W. 1985 On the use and significance of isentropic potential vorticity maps. *Q. J. R. Meteorol. Soc.*, **111**, 877–946
- James, I. N. 1987 Suppression of baroclinic instability in horizontally sheared flows. *J. Atmos. Sci.*, **44**, 3710–3720
- 1994 *Introduction to circulating atmospheres*. Cambridge University Press
- Juckes, M. N. 2000 Linear instability of broad baroclinic zones. *Q. J. R. Meteorol. Soc.*, **126**, 1065–1098
- Lait, L. 1994 An alternative form for PV. *J. Atmos. Sci.*, **51**, 1754–1759
- McIntyre, M. E. 1970 On the non-separable baroclinic parallel flow instability problem. *J. Fluid Mech.*, **40**, 273–306
- Magnusdottir, G. and Haynes, P. 1996 Wave activity diagnostics applied to baroclinic wave life cycles. *J. Atmos. Sci.*, **53**, 2317–2353

- | | | |
|---|------|--|
| Methven, J., Hoskins, B. J.,
Heifetz, E. and Bishop, C. H. | 2005 | The counter-propagating Rossby-wave perspective on baroclinic instability. Part IV: Nonlinear life cycles. <i>Q. J. R. Meteorol. Soc.</i> , 131 , ???–??? |
| Phillips, N. | 1954 | Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. <i>Tellus</i> , 6 , 273–286 |
| Sakai, S. | 1989 | Rossby-Kelvin instability: a new type of ageostrophic instability caused by a resonance between Rossby waves and gravity waves. <i>J. Fluid Mech.</i> , 202 , 149–176 |
| Simmons, A. and Hoskins, B. | 1977 | Baroclinic instability on the sphere: Solutions with a more realistic tropopause. <i>J. Atmos. Sci.</i> , 34 , 581–588 |
| | 1980 | Barotropic influences on the growth and decay of nonlinear baroclinic waves. <i>J. Atmos. Sci.</i> , 37 , 1679–1684 |
| Stone, P. H. | 1969 | The meridional structure of baroclinic waves. <i>J. Atmos. Sci.</i> , 26 , 376–389 |
| Thorncroft, C., Hoskins, B. and
McIntyre, M. | 1993 | Two paradigms of baroclinic wave life cycle behaviour. <i>Q. J. R. Meteorol. Soc.</i> , 119 , 17–55 |
| Wyatt, L. | 1981 | Linear and nonlinear baroclinic instability of the Northern Hemisphere winter zonal flow. <i>J. Atmos. Sci.</i> , 38 , 2121–2129 |