Relating Overreflection and Wave Geometry to the Counterpropagating Rossby Wave Perspective: Toward a Deeper Mechanistic Understanding of Shear Instability

N. HARNIK AND E. HEIFETZ

Tel-Aviv University, Tel Aviv, Israel

ABSTRACT

The two major approaches to understanding plane parallel shear instability—the counterpropagating Rossby wave (CRW) and overreflection (OR) approaches—illuminate very different fundamental aspects. This work examines how such seemingly different views can explain the same phenomena and why these differences arise, with the overall goal of deepening our understanding of shear instability.

This is done by rationalizing the OR theory in terms of CRW thinking, using a generalization of the CRW approach to multiple infinitesimal potential vorticity (PV) kernels. First the mechanics of cross-shear wave propagation and the exponential decay in an evanescent region are qualitatively explained. Then the cross-shear behavior in different wave geometries (full, partial, and overreflection) is examined. It is found that overreflection is basically a mutual amplification between PV kernels, with the main kernels being at the two sides of the evanescent region, which forms on one side of the critical level. The nondimensional overreflection coefficient is obtained by assuming a balance between the mutual amplification on the two sides of the evanescent region, and the rate at which this amplification is carried away. The critical level is found to be the only point at which there is a mutual amplification between two adjacent kernels.

A further examination of the sources of energy growth yields a generalized view of the restoking of the Orr mechanism, thought to be central to OR: growth depends explicitly on the shear, as in the Orr mechanism, while mutual PV kernel interactions are responsible for holding the perturbation at a configuration that allows it to contribute to growth (a restoking).

1. Introduction

Over the last 40 years, two major approaches have been taken in an attempt to understand the basic mechanism of plane parallel shear instability, namely, the counterpropagating Rossby wave (CRW) and the overreflection (OR) approaches.

In the CRW perspective, initiated by Bretherton (1966), the instability is understood in terms of a constructive, action-at-a-distance interaction between two CRWs, which are located in regions of opposite signs of mean potential vorticity (PV) gradient. Each CRW by itself is neutral and propagates counter to the local mean flow, via the Rossby (1939) mechanism. The waves interact by inducing velocity that advects the mean PV in the region of the opposed wave. Bretherton showed that these counterpropagating Rossby wave's phase speeds and growth rates are determined solely by their phase difference and interaction coefficient. He then concluded that normal-mode instability is obtained when the two CRWs phase lock each other to propagate together in a growing configuration.

The main criticism however against Bretherton's explanation, and later studies that expanded it (e.g., Hoskins et al. 1985; Davies and Bishop 1994; Heifetz et al. 1999), was that it can only be applied to models in which the CRWs are edge waves, supported by a corresponding structure of mean PV gradient delta functions. This motivated Heifetz et al. (2004a), to extend the CRW description to a general conservative plane parallel shear flow (baroclinic or barotropic with general mean shear and PV distributions), which is linearly unstable. Then the generalized CRW equations become the Hamilton equations where the Hamiltonian, the generalized momenta, and the coordinates, are the eddy pseudoenergy, the CRW pseudomomenta, and the phases, respectively. This generalization also rationalized the necessary conditions for instability of Rayleigh (1880) [and its baroclinic analog of Charney and Stern (1962) and Fjørtoft (1950)]. The former condition

Corresponding author address: Nili Harnik, Department of Geophysics, Tel-Aviv University, Tel Aviv, 69978, Israel.
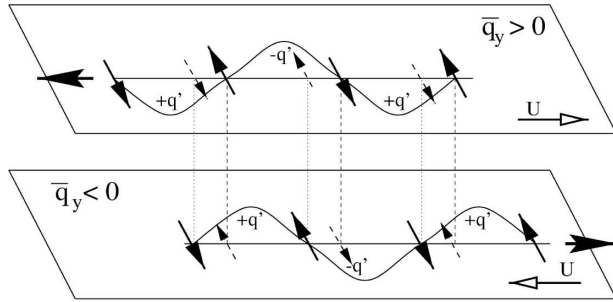E-mail: harnik@tau.ac.il

FIG. 1. Schematic illustration of the CRWs in the Eady model in the configuration of the most unstable normal mode. Blank arrows indicate the basic-state velocity U. Bold horizontal arrows represent the CRWs propagation direction. The edge positive and negative PV anomalies are indicated by $\pm q'$ and the circulation they induce is illustrated by the circled arrows and by the meridional solid arrows located $\mp \pi/2$ out of phase of $\pm q'$. The undulating solid lines illustrate the CRWs' meridional displacement. The two CRWs are phase locked in a hindering configuration with phase difference $\pi/2 < \epsilon < \pi$. The dashed arrows indicate the velocity induced by each CRW on the opposite edge. Each CRW advects the basic-state PV on the opposite edge in a way that makes the other CRW grow, while hindering its natural propagation.

FIG. 2. A schematic description of the wave geometry for an unstable normal mode. The $\overline{q}_y$ (thick line) and $U - c$ (thin line) profiles are shown on the left. Wave propagation and evanescence regions and the critical and turning levels are marked by WV, EV, CL, and TL, respectively. Arrows mark wave propagation, and U-shaped arrows mark reflection, with the widening one representing overreflection. Two lines narrowing toward each other mark wave evanescence.

requires a change of sign of the mean PV gradient within the domain and thus enables mutual instantaneous growth of the two CRWs. The latter condition requires a positive correlation between the mean PV gradient and the mean velocity and thus allows the two CRWs to counter propagate against the local mean flow and consequently maintain a phase locking. Hence, satisfying the two conditions together enables a sustained modal growth (cf. Fig. 1).

More recently, Heifetz and Methven (2005) applied the CRW approach to the nonmodal problem and optimal growth. They explained the differences in the optimal growth between the energy and the enstrophy norms in terms of the role of the Orr (1907) mechanism. In addition, they extended the CRW formulation to allow a study of the continuous spectrum, by using the Greens function approach of Robinson (1989). In this approach, the two CRWs are replaced by an infinite number (or a finite number, after discretization) of CRW kernels (hereafter KRWs). The complete spectrum solution is then described in terms of a multi-interaction between all KRWs, where the fundamental action-at-a-distance interaction between each of the KRW pairs is similar to the two CRW basic interaction.

This approach has been used to study Eady-like basic states with no interior mean PV gradients (e.g., Bishop and Heifetz 2000; Morgan 2001; Morgan and Chen 2002; Kim and Morgan 2002; Dirren and Davies 2004; de Vries and Opsteegh 2005, 2006), and recently also to
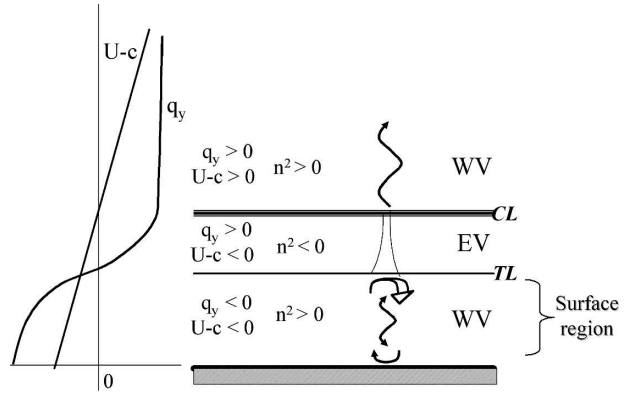
study more general basic states (de Vries and Opsteegh 2007a,b).

The OR approach, suggested by Lindzen and coauthors (summarized in the review article by Lindzen 1988), describes shear instability in terms of the evolution of waves propagating across the shear. In this framework, a central property is the wave geometry, which states the ability of waves to propagate across the shear as determined by the index of refraction. Under certain conditions, a wave approaching a reflecting level (where the index of refraction vanishes) will be overreflected, meaning the reflected amplitude will be larger than the incident one. Sustained growth can occur if the overreflected wave is fully reflected back, with the two waves interfering constructively (Fig. 2). For a given wavelength and basic state, the condition for constructive interference will be satisfied only for a certain phase speed. Thus, a quantization condition is at the heart of the dispersion relation.

In the OR wave geometry perspective the existence of a critical level (where the mean flow equals the wave phase speed) is crucial for the onset of shear instability. In the absence of a critical level, an incident wave, approaching a reflecting level will be fully reflected (Fig. 3a). If another reflecting level exists beyond it, then the incident wave will partly be reflected and partly tunneled to the next reflecting level (i.e., in the region between the two reflecting levels the index of refraction is negative and there is an exponential evanescence of the wave amplitude with zero group velocity across the shear) and will propagate away (Fig. 3b). In both configurations the overall integrated energy of the wave is

preserved and thus no instability occurs. If, on the other hand, a critical level (where the index of refraction changes sign rapidly) exists beyond a reflecting level then the wave incident upon the reflecting level will be overreflected, and in addition, some wave energy will propagate onwards beyond the critical level (Fig. 3c). In this case, more energy leaves the critical level than was incident upon it, meaning the overall energy has grown. If the overreflected wave is reflected back by another reflecting level (e.g., the ground) we can have sustained growth. Hence, Lindzen concluded that normal-mode instability will occur when the (southward/downward) overreflected wave and its subsequent (northward/upward) reflection interfere constructively. Lindzen and Barker (1985) also suggested that the source of wave energy in OR essentially comes from the Orr (1907) mechanism at the vicinity of the critical level—a perturbation that is tilted against the shear and is moving with the flow at some level will be tilted to a more vertical configuration, causing growth of the perturbation. The wave geometry configuration necessary for the instability is that which manages to hold the perturbations tilted against the shear near the critical surface, in order to continuously restoke the Orr mechanism. Lindzen and Barker (1985) and following papers did not, however, explain how mechanistically the wave geometry allows overreflection to occur, in particular, how the Orr mechanism works in the presence of a mean meridional PV gradient.

The CRW and the OR perspectives explain many aspects of shear instability, each by illuminating very different fundamental aspects of it. It is maybe not surprising, therefore, that there have been no attempts (to our knowledge) of relating the two. In this work we aim to explain the essence of the OR mechanism and wave geometry from the CRW perspective, with the overall goal of deepening our understanding of shear instability. After briefly reviewing the two approaches (section 2), we formulate a simple model that includes the necessary ingredients to exemplify OR and wave geometry, and use a numerical solution that is based on the KRW scheme, and an approximate analytical solution [which avoids exact matching conditions and application of Wentzel–Kramers–Brillouin (WKB) approximation], to capture the essential features of full, partial, and overreflection (section 3). In section 4 we examine the sources of energy growth using the CRW approach, and discuss the special role of the critical level. We then explain how the nondimensional overreflection coefficient R emerges from the CRW description, and discuss normal modes (section 5). We conclude and discuss our results in section 6.

2. CRW versus OR

One of the basic reasons for the fundamental differences between the CRW and the OR perspectives lies in the interpretation of the linearized PV equation[1]

$$\frac{\partial q}{\partial t} = -U\frac{\partial q}{\partial x} - v\overline{q}_y.$$

(1)

Equation (1), together with the inverse elliptic relation between meridional velocity and PV, forms the basis of the CRW approach. In the CRW framework, the PV anomaly is divided into localized PV structures (CRWs), which evolve both through a local advection by the zonal mean flow and because of meridional advection of the background PV field. The meridional flow associated with each of these CRW structures is nonlocal, hence it affects other CRWs.

In OR, on the other hand, the "PV thinking" is replaced by a "streamfunction thinking" so that Eq. (1) is written in terms of a streamfunction ψ. Assuming a pure zonal wave of the form $e^{ik(x-ct)}$, where the phase speed c is a priori assumed to exist, we get

$$\mathcal{L}(\psi) + \frac{\overline{q}_y}{U-c}\psi = 0,$$

(2)

where $q = \mathcal{L}(\psi)$, and $v = \partial\psi/\partial x = ik\psi$.

For simplicity we will consider a 2D barotropic shear, however the results can be generalized to the baroclinic case. Under these conditions $\mathcal{L}(\psi) = -k^2\psi + (\partial^2\psi/\partial y^2)$, and $\overline{q}_y = \beta - (\partial^2 U/\partial y^2)$. For this setup, Eq. (2) can be rewritten as

$$\frac{\partial^2\psi}{\partial y^2} + n^2\psi = 0, \quad n^2 = \frac{\overline{q}_y}{U-c} - k^2,$$

(3a,b)

where Eq. (3a) is a wave-propagation equation for the streamfunction, in the meridional direction, with an index of refraction n. The OR theory shows how, under certain conditions, waves can be overreflected.

As a first step in unifying the two approaches, we will examine how the cross-shear propagation of Rossby waves is manifest in terms of the interaction between many individual PV kernels. We note here that since OR is based on wave dynamics, it provides a more general description of shear instability, since it addresses not only instabilities arising from vortical Rossby wave interactions (the CRW approach), but

---

[1] We assume a basic zonal flow, $U(y, z)$, in the x direction with a mean PV gradient, $\overline{q}_y(y, z)$, only in the meridional direction, y, where q and v are the PV and the meridional velocity perturbations, respectively. The equation could describe the advection of a function of Ertel–Rossby PV on isentropic surfaces, or quasigeostrophic PV on horizontal surfaces.

also instabilities arising from divergent gravity wave interactions (described, e.g., by the Taylor–Goldstein equation; cf. Drazin and Reid 1981). While it may be possible to extend the CRW thinking to include gravity waves, it has not been done so far. Here, however, we restrict our discussion only to the framework of Eq. (1).

In the next two subsections we summarize briefly those features of the two perspectives that are essential for relating them. For a more detailed description, the reader is referred to the papers listed in the introduction.

a. CRW

In most previous CRW works the perturbation is divided into two CRW structures, based on the PV field. The CRWs have a distinct PV peak centered at the positive and negative PV gradient regions, respectively (although in the general case the structures extend throughout the domain). This approach stems from Eady-type models, where all of the PV gradients are concentrated at two levels (e.g., the surface and a lid).

In the following calculations we will use a different variant of the CRW approach (presented in the appendix of Heifetz and Methven 2005), where the PV perturbation is divided into multiple infinitesimal PV kernels, referred to as kernel Rossby waves (KRW), each of which induces its part of the meridional flow, as in a Green function approach. The PV kernels differ from the CRW structures in that they are multiple localized nonoverlapping PV building blocks, rather than two untilted PV structures, which extend throughout the domain. The basic dynamical evolution and the corresponding mathematical framework, however, are similar.

We therefore present the essential CRW dynamics using the simpler two CRW structures, and then present the relevant generalized KRW equations. Our analysis is done for the case of meridional propagation and barotropic instability, but all of it naturally generalizes to vertical propagation and baroclinic instability.

1) THE TWO-CRW PARADIGM

CRW structures were constructed originally for the framework of unstable basic states whose mean PV gradient changes sign and is located in two confined regions (say, in region 1, $\overline{q}_{y_1} < 0$, where in region 2, $\overline{q}_{y_2} > 0$). Each of the two CRWs has an untilted structure with PV perturbation concentrated in one of these regions; that is

$$q_j = Q_j(y, t)e^{i[kx + \epsilon_j(t)]}, \quad j = 1, 2 \qquad (4a)$$

and an associated meridional velocity field that fills the whole space

$$v_j = V_j(y, t)e^{i[kx + \epsilon_j(t) - \pi/2]}. \qquad (4b)$$

By construction, Q_j and V_j have the same sign such that a positive PV anomaly induces counterclockwise circulation and the meridional velocity is shifted $\pi/2$ to the east of the PV anomaly (Fig. 1). The details of how the two CRWs are constructed from the discrete spectrum solution is given in Heifetz et al. (2004a). Without CRW interaction, the advection of the mean PV by the velocity field, induced by the CRW of the other region, is ignored. Substituting then Eq. (4) into Eq. (1) reveals that each CRW is neutral and propagates with the Rossby phase speed

$$c_{jj} = -\frac{1}{k}\frac{\partial \epsilon_j}{\partial t} = U_j - \frac{\overline{q}_{y_j}}{k}\left(\frac{V_j}{Q_j}\right) \equiv U_j - \frac{\overline{q}_{y_j}}{k}G_{jj}, \quad (5)$$

where c_{jj} denotes the local propagation of CRW j, which is affected both by the mean flow advection [first term in the middle of Eq. (5)] and the self-induced Rossby wave propagation rate (second term), directed to the left of the mean PV gradient direction. If the Fjørtoft (1950) condition is satisfied, so that U_j and $\overline{q}_{y_j}$ are of the same sign, the Rossby propagation is counter the mean flow.

The Green function

$$G_{ij} = \frac{V_j}{Q_i} \qquad (6)$$

represents the PV inversion to the meridional velocity field. The first index at the suffix relates to the CRW that induces G whereas the second index indicates the CRW that the Green function is induced upon. Here, G is symmetric, $G_{ij} = G_{ji}$, where the evanescence of the velocity field away from the inducing CRW yields that $G_{ij} < G_{ii}$, when $i \neq j$.

When the two CRWs exist together we obtain their interaction equations by substituting Eqs. (4) and (6) for the two regions into Eq. (1). Taking the real and imaginary parts yields

$$\dot{Q}_1 = -\overline{q}_{y_1}G_{21}Q_2 \sin\epsilon, \quad \dot{Q}_2 = \overline{q}_{y_2}G_{12}Q_1 \sin\epsilon,$$

$$(7a,b)$$

$$\dot{\epsilon}_1 = -kU_1 + \overline{q}_{y_1}G_{11} + \overline{q}_{y_1}G_{21}\frac{Q_2}{Q_1}\cos\epsilon, \qquad (8a)$$

$$\dot{\epsilon}_2 = -kU_2 + \overline{q}_{y_2}G_{22} + \overline{q}_{y_2}G_{12}\frac{Q_1}{Q_2}\cos\epsilon. \qquad (8b)$$

Here, $\epsilon = \epsilon_2 - \epsilon_1$ is the CRW phase difference. Note that the quantity $\overline{q}_{y_i}G_{ji}$ is effectively an action-at-a-distance interaction coefficient.

Equations (7a) and (7b) indicate that in order to obtain growth of both waves, $\overline{q}_{y_1}$ and $\overline{q}_{y_2}$ must have oppo-

site signs (since $G_{ij} = G_{ji}$), as expected from the Rayleigh criteria and the Charney and Stern (1962) condition. Normal modes are the specific solutions of Eqs. (7a), (7b) and Eqs. (8a), (8b) for which the two CRWs are phase locked to have equal growth rates and phase speeds. The growing normal-mode configuration occurs when the CRWs are tilted against the shear ($0 < \epsilon < \pi$) either in a helping ($0 < \epsilon < \pi/2$) or hindering manner ($\pi/2 < \epsilon < \pi$).

2) KRW INTERACTION

In this section we present a generalization of the CRW dynamics to an arbitrary number of PV kernels. This allows us to relate the cross-shear wave dynamics, which is the basis of OR theory, to the Rossby wave interactions, which are the basis of CRW theory. By doing this we will gain a new mechanistic understanding of shear instability and the associated energy growth, as well as of the fundamental processes of cross-shear wave propagation, evanescence, reflection, and overreflection.

For a zonal Fourier component, the PV perturbation can be written in the integral form

$$q(y, t) = \int_{y'} [q(y', t)\delta(y' - y)] \, dy' \equiv \int_{y'} \tilde{q}(y', t) \, dy',$$

(9)

where $\tilde{q}(y', t)$ denotes the vorticity density kernel at y', which induces the meridional velocity (per unit width)

$$\tilde{v}(y, y', t) = ik\mathcal{L}^{-1}\tilde{q}(y', t) \equiv -iq(y', t)G(y', y), \quad (10a)$$

where $\quad -\dfrac{1}{k}\mathcal{L}_{(y,k)}G(y', y) = \delta(y' - y).$ (10b)

The Green function $G(y', y)$, defined by Eq. (10b) subject to the appropriate boundary conditions, is real, dimensionless and positive definite. It plays a similar role to G_{ij} of Eq. (6), namely it is the streamfunction anomaly induced by the KRW at y' (which is a function of y). Note however that G_{ij} has units of length in the two CRW case.

The $-\pi/2$ phase difference between the PV and the meridional velocity indicates that a positive PV kernel induces counterclockwise circulation. Then the inversion of q becomes

$$v(y, t) = -i\int_{y'} q(y', t)G(y', y) \, dy'.$$

(11)

Substituting Eqs. (9) and (11) into (1), and taking the real and imaginary parts, yields

$$\dot{Q}(y) = \overline{q}_y(y) \int_{y'} Q(y')G(y', y) \sin\epsilon(y, y') \, dy', \quad (12a)$$

$$\dot{\epsilon}(y) = -kU(y)$$

$$+ \overline{q}_y(y) \int_{y'} [Q(y')/Q(y)]G(y', y) \cos\epsilon(y, y') \, dy',$$

(12b)

where an overline indicates a zonal average, and $\epsilon(y, y') = \epsilon(y) - \epsilon(y')$. Equation (12) is the continuous analog to Eqs. (7a), (7b), (8a), and (8b). Heifetz and Methven (2005) showed that for practical purposes Eq. (12) can be formulated in a matrix form when the domain is discretized into a finite number of KRWs with a small width Δy.[2]

b. Overreflection

Overreflection theory assumes the existence of waves propagating in the cross-shear direction, with a well-defined phase speed, and correspondingly, a well defined index of refraction. For a given basic state, the index of refraction determines the wave propagation characteristics for the specific zonal wavenumber and phase speed. The configuration in terms of wave propagation and evanescence regions is the wave geometry. According to OR theory, the wave geometry that supports unstable normal modes consists of two basic elements: first, the part that supports overreflection as the means of drawing energy from the mean flow, and second, the conditions that allow the overreflected wave to be reflected back constructively to the critical level, thus insuring a continuous overreflection as part of a sustained normal mode.

In this section we discuss the basic conditions leading to overreflection (and shear instability), but first we review some basic relations of wave propagation that are relevant to our discussion.

1) SOME BASIC SIGN RELATIONS

In the following sections we will examine the wave propagation on a few basic wave geometry configura-

---

[2] The multiple KRW approach has similarities with that of Waugh and Dritschel (1991) for a piecewise continuous PV profile. Waugh and Dritchel used a contour dynamics approach with Green functions at each PV jump. The main variable there was the PV contour displacement (η), while our variable is PV. In the linear limit these are related by the following equation: $[(d/dt) + U(d/dx)]q' = -v'\overline{q}_y = -[(d/dt) + U(d/dx)]\eta\overline{q}_y$; therefore the contour displacement in the linear regime is $\eta = -q'/\overline{q}_y$. This, in fact, was the rational behind the basic CRW formulation, developed in Heifetz et al. (2004a).

tions, and use wave conservation relations to deduce the conditions leading to overreflection.[3] To do this, we will make use of some basic sign relations between the basic state, the wave phase structure and the energy flux. To simplify later discussion, we will first briefly review these sign relations.

The solution to the wave Eqs. (3a) and (3b) depends on the basic-state structure. For the simplest case of constant U and β, we have a wavy solution for positive n^2 and an exponential solution for negative $n^2 = -m^2$ ($n = -im$, and $m > 0$):

$$\psi = (Ae^{-iny} + Be^{iny})e^{ik(x-ct)}, \qquad (13a)$$

$$\psi = (Ae^{-my} + Be^{my})e^{ik(x-ct)}. \qquad (13b)$$

Plugging Eq. (13a) into Eqs. (3a) and (3b), we obtain the well-known Rossby wave dispersion relation:

$$c = U - \frac{\beta}{k^2 + n^2}, \qquad (14)$$

with the group velocity:

$$Cg_y \equiv \frac{\partial \omega}{\partial n} = \frac{2nk\beta}{(k^2 + n^2)^2}, \quad \omega = kc. \qquad (15a,b)$$

We note the following relations between the wave structure, group velocity, and wave fluxes. For a wave solution of the form $e^{iny} e^{ikx}$, the sign of nk determines the phase tilt in the x–y plane. Assuming a positive k, the wave tilt is westward/eastward with y for positive/negative n. Using the geostrophy relations, it is also clear that the meridional momentum flux $\overline{u_g v_g}$ is negative/positive for positive/negative values of the product nk.

Extending the Eliassen and Palm (EP) theorem (Eliassen and Palm 1961) for stationary waves to an arbitrary real phase speed c (e.g., Lindzen and Tung 1978) yields the following relation between the wave energy and momentum fluxes:

$$\overline{pv_a} = -(\overline{U} - c)\rho_o \overline{u_g v_g} \qquad (16)$$

(the overline represents zonal averaging, p is perturbation pressure, and v_a is the ageostrophic meridional velocity) so that for $U - c > 0$, a westward/eastward phase tilt with y is associated with a positive/negative wave energy flux ($\overline{v_a \psi}$), while the opposite phase tilt-energy flux direction relation is true for $U - c < 0$. From Eqs. (15a) and (15b) we also see that the direction of the group velocity Cg_y depends on the signs of nk and the PV gradient, so for $\beta > 0$, Cg_y, and $\overline{u_g v_g}$ have

TABLE 1. Basic wave propagation and wave geometry relations for a barotropic pure plane Rossby wave. The meridional group velocity and wave energy flux directions are indicated as a function of the wave phase tilt and basic-state properties. Phase lines and energy flux arrows are drawn on a latitude–longitude plane. Note that only basic states that support wave propagation [$\beta/(U - c) > 0$] are indicated.

Basic state/wave phase tilt direction	$nk > 0$ $\overline{u'v'} < 0$	$nk < 0$ $\overline{u'v'} > 0$
$\beta > 0$ $U - c > 0$	$Cg_y > 0$ $\overline{pv_a} > 0$	$Cg_y < 0$ $\overline{pv_a} < 0$
$\beta < 0$ $U - c < 0$	$Cg_y < 0$ $\overline{pv_a} < 0$	$Cg_y > 0$ $\overline{pv_a} > 0$

opposite signs, while for $\beta < 0$ they are of the same sign. For wave propagation conditions ($n^2 > 0$), for which Cg_y is defined, we expect the group velocity and wave energy flux to be in the same direction. This is indeed the case when the basic state is constant, since according to the dispersion relation [Eq. (14)] β and $U - c$ have the same sign. The above relations are summarized in Table 1. In what follows, we will explicitly state the sign of nk in the solution, and assume $k > 0$, $n > 0$. Thus $e^{iny}(e^{-iny})$ will correspond to a negative (positive) momentum flux.

Next we note that the sign relation between q and ψ depends on $\beta/(U - c)$ [Eq. (1)], or correspondingly, on $k^2 + n^2$ [by direct evaluation of $q = \mathcal{L}(\psi)$, assuming Eq. (13a)]. A wave solution [for which $\beta/(U - c) > 0$ by definition] exhibits the normal relation of cyclonic flow around a positive PV anomaly ($q \propto -\psi$, since $k^2 + n^2 > 0$). If, on the other hand, there is a region in the flow for which the basic state and zonal phase speed are such that $\beta/(U - c) < 0$, the resulting flow solution (which will necessarily be exponential) will be anomalous in the sense of anticyclonic[4] flow around a positive PV anomaly [$q \propto \psi$ since $k^2 + n^2 = k^2 - m^2 = \beta/(U - c) < 0$]. In such an anomalous case, the meridional flow is necessarily dominated by nonlocal PV anomalies. While it may seem that the above q–ψ relations are trivial, the implications for overreflection and shear instability, to the best of our knowledge, have not been explicitly noted or appreciated before. We will come back to this point later on.

---

[3] Our analysis is somewhat similar to what is done in Lindzen and Tung (1978) for gravity waves.

[4] Since in the barotropic case q is equal to the vorticity, we cannot, strictly speaking, have anticyclonic flow around a positive PV anomaly. By anticyclonic, we mean that the meridional wind contribution to vorticity is anticyclonic around positive PV centers $-(\partial v/\partial x) = -k^2\psi < 0$ for $q > 0$, yielding $q \propto \psi$. In the baroclinic case (e.g., below the critical level in the unstable Charney modes) this anomalous relation implies that positive PV anomalies are correlated with negative vorticity anomalies.

Finally, if the basic state varies with y, the solution to Eqs. (3a) and (3b) cannot in general be written analytically, however, if the basic state varies slowly enough with y, we can use the WKB approximation, which yields an approximate plane wave/exponential solution with an amplitude and wavenumber (n), which are slowly varying in y (slower than a wavelength $2\pi/n$). In this case relations [Eqs. (13)–(15a), (15b)] apply to the regions in which the solution is wavy, where a group velocity can be defined, and Eq. (16) holds for an arbitrary basic state, provided a pure real phase speed c. In particular, the group velocity and the EP relations [(15a) and (16)] are still consistent, since the wave solution requires β and $U - c$ to have the same sign.

2) WAVE GEOMETRY

The wave geometry, which can be deduced from the index of refraction [n^2, Eq. (3b)], is the configuration of the basic state in terms of wave propagation and wave evanescence regions ($n^2 > 0$, $n^2 < 0$ respectively), which can be separated in two ways: by a turning level, at which $n^2 = 0$, or by a critical level, at which $n^2 \to \pm\infty$ (when $U = c$).

Lindzen and Tung (1978) showed that an incident Rossby wave of a given zonal wavenumber and phase speed, that reaches a turning level, will be overreflected if a critical level exists beyond the turning level, with a wave-propagation region somewhere beyond it. Normal-mode growth occurs when the overreflected wave is then reflected back from another turning level, with the right zonal wavenumber to yield constructive interference between the reflected and overreflected wave.[5] This is shown schematically in Fig. 2, and explained in detail in the papers by Lindzen and coauthors (see review article Lindzen 1988).

We will mostly concentrate on the process of overreflection, and will briefly discuss the conditions necessary for normal-mode formation later on (section 5). We assume a steady wave source with a specified zonal phase speed c, located at the south of our domain (y increases northward) and look for the steady-state solution (characterized by reflection and transmission coefficients) in an unbounded domain. We examine simple basic state configurations which support full, partial, and overreflection (Fig. 3), by assuming a piecewise continuous basic state, with regions of positive and negative β and $U - c$. Physically, a step function in β can represent a triangular topography. This yields a

[5] This is the basic wave geometry that allows shear instability, although different variations are possible [e.g., two critical levels exist in the barotropic model of Kuo (1949)].
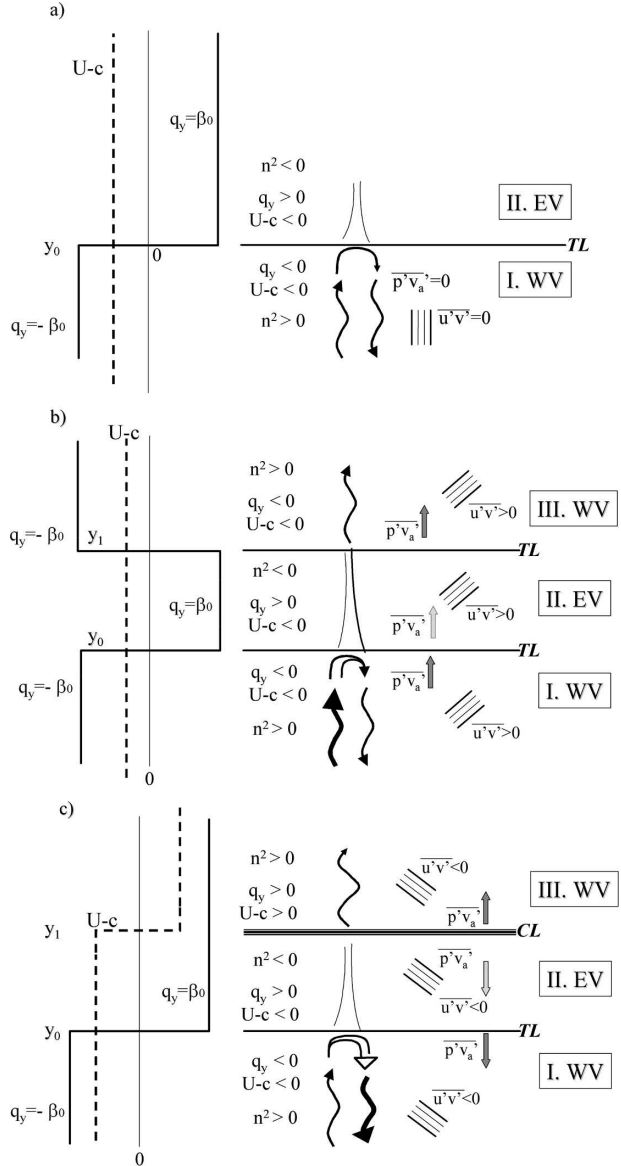

FIG. 3. The wave geometry configurations for the cases of (a) full, (b) partial, and (c) overreflection, for the simplified basics states discussed in the text. The different regions and wave processes are marked as in Fig. 2, except that the $U - c$ and $\overline{q}_y$ profiles are drawn in dashed and solid lines, respectively. Also marked are the directions of the wave energy flux $\overline{p'v'_a}$ (with arrows) and the momentum flux $\overline{u'v'}$, denoted by the direction of phase lines (see Table 1).

piecewise continuous n^2 with a turning level at the jumps in β and a critical level at the jumps in $U - c$. We note that a steady overreflecting solution is possible under radiation boundary conditions, because the excess energy which the wave draws from the basic state at the critical level is radiated away (and in linear theory the basic state is an infinite energy source).

In principle, the full solution can be obtained by

matching the solution in each region's interfaces. However, the problem is not very well posed at these interfaces since the mean PV gradient is not defined at the jump in β, and it is a dipole delta function at the jump in $U - c$. We will therefore not solve the equation fully, rather, we will make use of the EP integral conservation relations, to determine whether we have full, partial, or overreflection. This is sufficient for the qualitative understanding we want to obtain.[6] We also verify the qualitative nature of the following solutions using numerical integrations, including more smoothly varying basic states.[7]

The two EP flux relations are

$$\frac{\partial}{\partial t}\left(\frac{\overline{q^2}}{2\overline{q}_y}\right) = -\overline{vq} = \frac{\partial}{\partial y}\overline{u_g v_g}, \tag{17a}$$

$$\frac{\partial}{\partial t}\left(\frac{\overline{u_g^2 + v_g^2}}{2}\right) = (U - c)\frac{\partial}{\partial y}\overline{u_g v_g} = \frac{\partial}{\partial t}\left(\frac{(U-c)}{\overline{q}_y}\frac{\overline{q^2}}{2}\right), \tag{17b}$$

where the subscript g represents the geostrophic wind, and for the second relation we have used Eq. (16), which is strictly valid only for real c, but can be used qualitatively for sufficiently small growth rates (cf. Lindzen and Tung 1978; Lindzen et al. 1980). From these relations we see that under steady-state conditions, the momentum flux is constant everywhere:

$$\overline{u_g v_g} = -k/2 \operatorname{Im}(\psi_y \psi^*) = \text{const}, \tag{18}$$

where the asterix denotes the complex conjugate, provided $(U - c)$ and $\overline{q}_y$ do not vanish at the same level.[8] For growing normal modes, for which steady state does not apply, the domain integral of Eq. (17a) is assumed to vanish, based on the boundary conditions, yielding the Charney and Stern (1962) condition for instability [where Eq. (17b) yields the Fjortoft (1950) condition].

For the full and partial reflection setups our approach is similar to Lindzen and Tung (1978); however, for the OR configuration we avoid the complexity of the local analysis near the critical level, by taking advantage of the steady-state assumption, which does not hold for instability.

(i) Full reflection

We divide the domain into two regions, region I, for $y < y_0$, where $\beta = -\beta_o < 0$, and region II, for $y > y_0$, where $\beta = -\beta_o > 0$. We also have a constant $U - c < 0$, so that $n^2 > 0$ for $y < y_0$ and $n^2 < 0$ for $y > y_0$. At $y = y_0$ there is a turning level (n^2 necessarily goes through zero there). For this configuration, we expect a wave emanating from the south to be fully reflected from the turning level, since there is no other place for the wave energy to go.[9] Away from the turning level, the solution is

$$\psi = \begin{cases} (e^{-iny} + Re^{iny})e^{ik(x-ct)} & \text{for } y < y_0 \\ (Ae^{-my} + Be^{my})e^{ik(x-ct)} & \text{for } y > y_0 \end{cases}, \tag{19}$$

where R, A, B are complex constants, and $m^2 \equiv -\beta_0/(U - c) + k^2 > 0$ is also a constant. We have assumed that the incident wave has unit amplitude, and the reflected wave amplitude is R. The propagation direction corresponds to the direction of the group velocity, with positive being northward. Thus, since $\beta < 0$ in region I, the northward propagating wave has an eastward phase tilt with height ($\psi \propto e^{-iny}$).

From requiring that ψ be bounded at $y \to \infty$, we get $B = 0$. Thus the solution in region II, which is denoted as an evanescence region, is exponentially decaying. Using Eq. (18), the momentum fluxes in each of the regions are

$$\overline{u_g v_g} = \begin{cases} -\dfrac{kn}{2}(|R|^2 - 1) & \text{for } y < y_0 \\ -\dfrac{km}{2}\operatorname{Im}(A^*B) = 0 & \text{for } y > y_0 \end{cases}. \tag{20}$$

Equating the momentum fluxes in both regions yields $|R| = 1$, which means the reflected wave amplitude is equal to the incident wave amplitude; that is, full reflection, as expected. In this case, the net momentum fluxes, as well as wave energy fluxes [cf. Eq. (16)] are zero in both regions. This is illustrated schematically in Fig. 3a.

---

[6] A more rigorous treatment would be to assume a more physical basic state such as a smoothed out step function, for example, a tanh. This will require using a WKB approximation, with asymptotic matching to Airy and Bessel equations at the turning and critical levels, respectively, where the WKB approximation breaks down (cf. Lindzen and Rosenthal 1981). For the jump in β the contour dynamics approach of Waugh and Dritschel (1991) can also be applied.

[7] Although we do not obtain a strict steady-state solution because of the complexity of having both a constant source and no reflection from the southern boundary (see appendix B for details).

[8] Please note that although our example excludes Eady (1949) type basic states, the conclusions we draw for the OR mechanism should still be valid.

[9] This configuration is similar, for example, to the surface wave region of the Charney model (but with no critical level), and to the Charney model configuration for the external neutral modes, which have a negative phase speed c.

(ii) Partial reflection

We now add another region (III) of wave propagation beyond the wave evanescence region, by setting $\beta = -\beta_o < 0$ for $y > y_1 > y_0$, leaving the zonal mean wind the same ($U - c < 0$). Since the evanescent wave in region II does not decay to zero at $y = y_1$, some perturbation will emerge in region III and propagate onward—resulting in partial transmission (wave tunneling through the evanescent region) and correspondingly partial reflection back from the turning surface at $y = y_0$. This configuration differs from the OR configuration (described in the next section) in that the wave incident on the evanescent region from the south will tunnel to the next wave-propagation region through a turning surface, and not a critical surface.

Again, the solution in each of the three regions is

$$\psi = \begin{cases} (e^{-iny} + Re^{iny})e^{ik(x-ct)} & \text{for} \quad y < y_0 \\ (Ae^{-my} + Be^{my})e^{ik(x-ct)} & \text{for} \quad y_0 < y < y_1, \\ Te^{-iny}e^{ik(x-ct)} & \text{for} \quad y > y_1 \end{cases}$$
(21)

where T is a complex constant transmission coefficient, and we have assumed that there are no additional wave sources besides the one to the south [we set to zero the coefficient of e^{iny}, which denotes a southward-propagating wave in region III, since $\beta = -\beta_o < 0$, Eq. (15a)]. Note that now, since region II does not extent to infinity, we do not set $B = 0$.

The momentum fluxes in each of the regions are

$$\overline{u_g v_g} = \begin{cases} -\dfrac{kn}{2}(|R|^2 - 1) & \text{for} \quad y < y_0 \\ -\dfrac{km}{2}\operatorname{Im}(A^*B) & \text{for} \quad y_0 < y < y_1. \\ \dfrac{kn}{2}|T|^2 & \text{for} \quad y > y_1 \end{cases}$$
(22)

Equating the momentum fluxes in regions I and III yields $|R|^2 = 1 - |T|^2 < 1$, which means only part of the wave reflects back $|R| < 1$, and part of it is transmitted northward, with the total power leaving the system being equal to the total power entering it ($|R|^2 + |T|^2 = 1$). Although not necessary for this derivation, note that the momentum flux in the evanescent region is not zero (meaning there is a nonzero phase tilt with latitude there). For this case, the momentum fluxes, as well as the wave energy fluxes, are positive in all three regions [cf. Eq. (16)]. This is illustrated schematically in Fig. 3b.

(iii) Overreflection

We now set $\beta = \beta_o > 0$ for both regions II and III, and have a step function in $U - c$, so that $U - c > 0$ for $y > y_1$. For simplicity, we choose $-(U - c)_{y<y_1} = (U - c)_{y>y_1}$. This will result in a critical level at $y = y_1$. According to Lindzen and Tung (1978), this configuration should yield a southward overreflection of the incident wave, with some additional wave energy also being emitted northward from $y = y_1$.

Again, the solution in each of the 3 regions is

$$\psi = \begin{cases} (e^{-iny} + Re^{iny})e^{ik(x-ct)} & \text{for} \quad y < y_0 \\ (Ae^{-my} + Be^{my})e^{ik(x-ct)} & \text{for} \quad y_0 < y < y_1. \\ Te^{iny}e^{ik(x-ct)} & \text{for} \quad y > y_1 \end{cases}$$
(23)

We have again assumed that there are no additional wave sources besides the one to the south, hence for $y > y_1$ there are only northward-propagating waves, but note that now the transmitted wave is proportional to e^{iny} since $\beta > 0$, which changes the sign of the momentum flux in region III:

$$\overline{u_g v_g} = \begin{cases} -\dfrac{kn}{2}(|R|^2 - 1) & \text{for} \quad y < y_0 \\ -\dfrac{km}{2}\operatorname{Im}(A^*B) & \text{for} \quad y_0 < y < y_1. \\ -\dfrac{kn}{2}|T|^2 & \text{for} \quad y > y_1 \end{cases}$$
(24)

Equating the momentum fluxes in regions I and III yields $|R|^2 = 1 + |T|^2 > 1$, which means the reflected wave has a larger amplitude than the incident one. and the total power leaving the system is larger than what entered ($|R|^2 + |T|^2 > 1$).

The momentum flux is negative in all regions, but the wave energy flux is directed outward from $y = y_1$. Comparing to the case of partial reflection, for which the basic state in regions I and II is the same, the net wave energy flux has changed direction (as expected), and correspondingly, the momentum flux has changed direction. In region III, the momentum flux has also changed direction, but since now $U - c > 0$ the wave energy flux has not [cf. Eq. (16)]. This is illustrated schematically in Fig. 3c. Our numerical calculations [see section 3c(3)] also indicate that the results hold for a slightly less idealized basic state where the jumps in U and β are smoothed using a tanh.

This example demonstrates the OR view of the necessary conditions for instability $-\beta$ has to change sign for $|R| > 1$ (Charney and Stern 1962). The necessary existence of two wave propagation regions (I and III)

then requires $(U - c)$ to change sign as well [so that β and $(U - c)$ are positively correlated—the Fjortoft condition]. Except for the case where β changes sign at the critical level, such a configuration necessarily has a region of wave evanescence (region II) for which q and ψ are anomalously related (i.e., a positive PV anomaly is circulated anticyclonically).

3. The mechanistics of overreflection from the CRW perspective

While overreflection is shown to exist mathematically, based on conservation relations, to the best of our knowledge, an explanation of why it occurs, mechanistically, has not yet been found. The KRW formulation provides an appropriate framework for examining the basic processes of cross-shear wave propagation, evanescence, reflection, and overreflection.

a. Wave propagation

We begin with a uniform basic state of constant β and U, and an initial KRW anomaly at $y = y_0$ with a zonal wavenumber k, moving with a phase speed c such that $U - c$ has the sign of β and $n^2 = \beta/(U - c) - k^2 > 0$. The evolution of this anomaly is illustrated in Fig. 4a. For the case of positive β the initial KRW (denoted on the figure by A) will induce similar anomalies (B and B′) to the north and south, but shifted westward by $\pi/2$ (due to the advection of the mean PV by the meridional velocity induced by the initial kernel). At the same time B and B′ will induce similar westward shifted anomalies (C and C′) to the north and south, respectively, but will also act to weaken the initial anomaly A (the dominant interactions are between adjacent kernels, due to the exponential decay of the Green functions). Similarly, C and C′ will induce new westward shifted anomalies to the north and south, while reducing B and B′. This process will continue, resulting in two pulses propagating away from y_0. The northward-propagating branch will be tilted westward with y while the southward one will be tilted eastward, consistent with wave propagation theory [cf. section 2b(1) and Table 1]. Similar arguments lead to the opposite phase tilts in the case of negative β. We note that numerical runs we have done confirm this picture of two pulses propagating away from the source.

This heuristic description explains the general relation between the cross shear propagation direction and phase tilt, however to obtain a more quantitative structure one needs to take into account the multiple kernel interactions (e.g., how the initial kernel A interacts with kernels C and C′), as well as the advection by the mean flow. More specifically, once kernels B and B′ are cre-

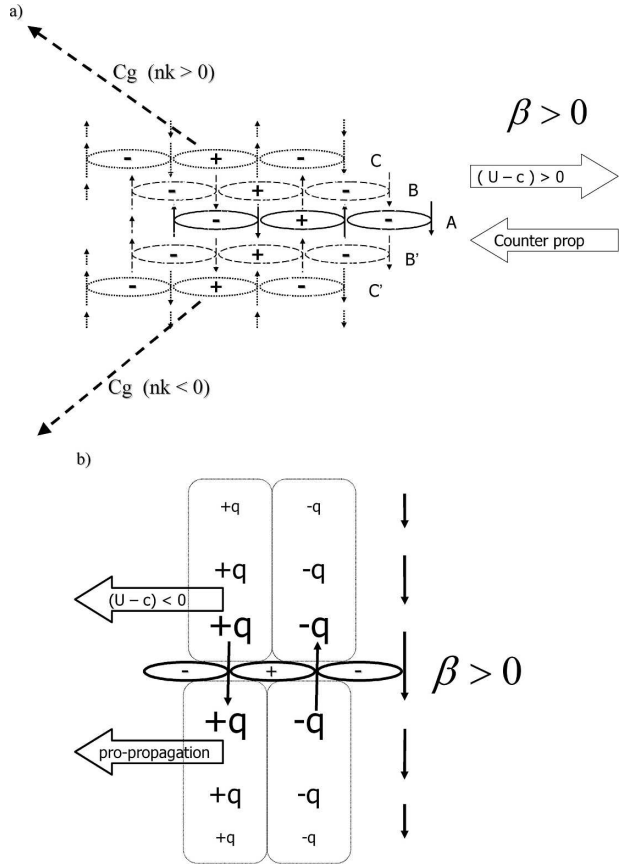

FIG. 4. A KRW schematic of the evolution of a wave source with a specified zonal phase speed c, shown on the zonal-meridional plane. (a) Wave propagation (basic states with $n^2 > 0$). The ovals are PV kernels, with the PV anomaly sign marked, and the source denoted by A. Arrows show the meridional wind induced by the PV kernel of similar line type. (b) Wave evanescence $[(\bar{q}_y/U - c) < 0]$. Bold ellipses are the source, and the dashed rectangles are the induced evanescent structures. The mean flow advection and self-induced propagation directions are marked in large arrows. See text for details.

ated $\pi/2$ to the west of A (assuming $\beta > 0$), they will be advected eastward by the mean flow (assuming $U - c > 0$), with this advection being reduced by a counter self-propagation to the west. Once B (and B′) and A are less than $\pi/2$ out of phase (recall that kernel A is forced to move with a phase speed c), kernel A will tend to shift kernel B back to $\pi/2$, at a rate that is proportional to β. In addition, the interaction with all other kernels has to be taken into account, but this will not vary the picture much since the dominant interaction stems from the adjacent kernels. The final phase tilt $(-k/n)$ will depend on the ratio $\beta/(U - c)$, consistent with the dispersion relation (14).

In appendix A we show how the KRW description of a free Rossby wave is indeed consistent with this view. Moreover, we show that once a full wavelength in y is

established ($\lambda = 2\pi/n$), the perturbation can become self-sustained,[10] in the sense that its structure is maintained, and all kernels move together with a phase speed that is very close to c (it is exactly c for an infinite wave train). This means that if we now shut off the initial source at kernel A, the wavepacket will continue to propagate in the y direction on its own with a zonal phase speed of approximately c (numerical runs do indeed support this). The ability for the anomaly to be self sustaining is crucial for wave propagation, and in fact, on basic states for which $\beta/(U - c) - k^2 < 0$, no wavepacket can be supported, hence the basic state is evanescent. This is shown in the next subsection and in appendix B.

b. Wave evanescence

We now examine how a similar initial KRW evolves on an evanescent basic state, for which β and $U - c$ are of opposite signs (let us assume a positive β and a negative $U - c$). In appendix B, we also examine the less relevant (for overreflection) evanescent basic state for which $0 < \beta/(U - c) < k^2$. As with the wave-propagation case above, the KRW anomaly will initially induce a column of KRW anomalies to the north and to the south, with an amplitude that is decaying with distance, and shifted westward by $\pi/2$ (see Fig. 4b). Unlike the wave propagation case, the induced KRWs are now propropagating (viewed from a frame of reference moving with the source, the self-induced phase speed, which depends on β, is in the direction of the mean flow), as a result of the signs of β and $U - c$ being opposite [cf. Eq. (5)]. This means that the columns are not self-sustainable, and they will be shifted westward relative to the initial KRW. Once this happens, the circulation induced by the initial kernel will induce a counterpropagation of the column. If the initial KRW is strong enough compared to the induced columns of kernels, a sustainable steady state will be reached. As we show in appendix B, this is possible if the initial KRW is a vorticity delta function, and the induced column is in anti phase and decays away from it. This evanescent configuration can be expressed mathematically as follows:

$$\psi = \Psi e^{-m|y|} e^{ikx} \quad \text{and} \tag{25a}$$

$$q = \left(\left| \frac{\beta}{U - c} \right| e^{-m|y|} - 2m\delta(y) \right) \Psi e^{ikx}, \tag{25b}$$

where the first term on the rhs of Eq. (25b) represents the decaying KRW columns and the second term rep-

---

[10] Depending on the aspect ratio k/n, may be more than a single wavelength is needed [see appendix A, Eq. (A.2c)].

resents the vorticity delta function of the initial KRW. As discussed previously, the configuration in which $\beta/(U - c) < 0$ necessarily leads to the anomalous sign relation between q and ψ. This is made possible by the dominance of the initial KRW. Note that $\beta/(U - c) < 0$ also means $k^2 < m^2$, and therefore the initial KRW influence (via the induced meridional velocity which decays as the Green function, $e^{-k|y|}$; cf. appendix A) decays slower than the induced PV evanescence $(e^{-m|y|})$.

Unlike the wave-propagation case where the induced anomaly can eventually sustain itself with a phase speed c, now the induced anomaly can only be sustained at this phase speed with the help of the initial kernel, hence, we cannot have a signal propagating away form it. This also turns out to be the case for the normal (in terms of the ψ–q relation) evanescence (cf. appendix B).

c. Full, partial, and overreflection

We now examine the evolution of anomalies when both wave propagation and wave evanescence regions are present. With the aim of understanding the process of overreflection, we examine how an initial KRW anomaly, with a set zonal wavenumber and phase speed, evolves under the three configurations discussed in section 2b(2).

1) FULL REFLECTION

We turn on a KRW wave source (KRW 0) in region I of the basic state of Fig. 3a. A wave train with an eastward phase tilt with y ($\beta < 0$ in region I) will emanate from KRW 0, as described in section 3a, until it reaches the turning surface at $y = y_0$ (we denote the kernel that has just formed to the south of the jump in β as KRW A). Once created, KRW A (Fig. 5a) will induce a cyclonic circulation around its positive PV centers, but now, since β has changed sign, it will induce a westward-shifted kernel (KRW B) to the north, and an eastward-shifted KRW B' to the south. Since the kernel created to the north will be propropagating, it will evolve, according to section 3b, into an exponentially decaying, untilted KRW column, in antiphase with KRW A. During the above adjustment process, in which kernel B shifts westward into its antiphase position, it will also alter the location of kernel A by hindering its counterpropagation, with the end result being a phase shift to the west, relative to the original position when it just reached y_0. During this process, kernel A will also be strengthened, compared to the case of pure wave propagation (in which kernel B weakened A). The combination of this strengthening and the phase shift will alter the southward effect on kernel B', relative to the pure propagation case, resulting in a south-
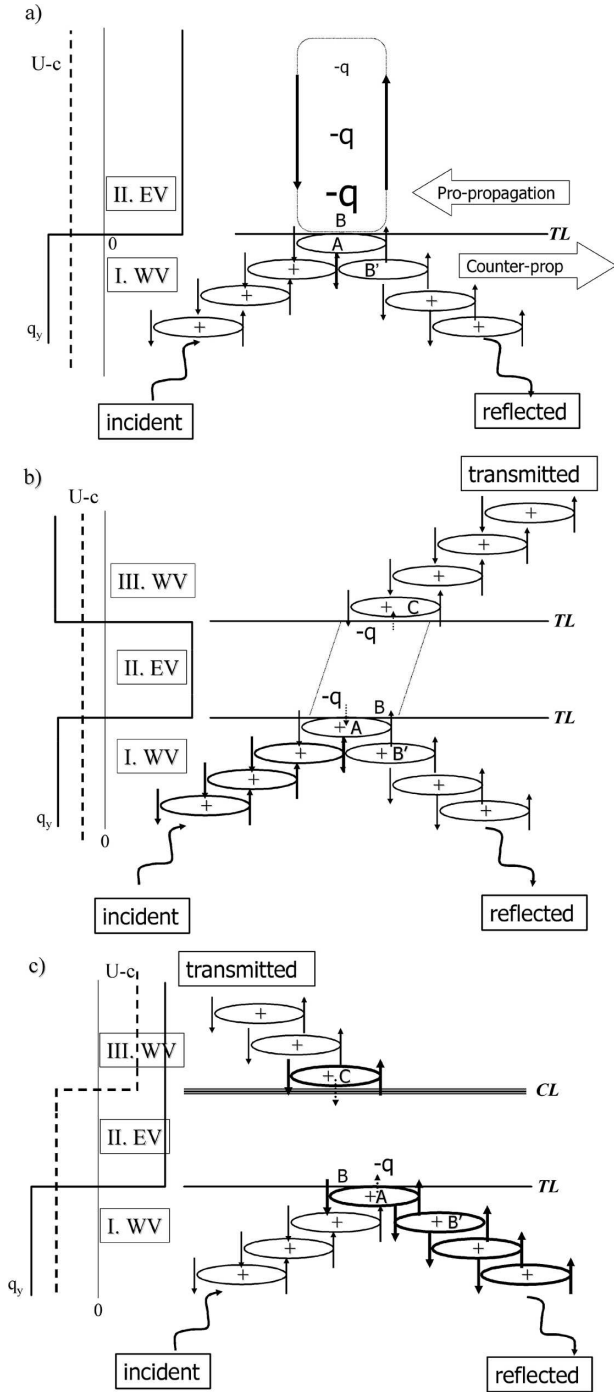
FIG. 5. The KRW schematic evolution of the three cases of Fig. 3. The different wave geometry regions, wave processes, and PV kernels are marked similar to Figs. 3 and 4. The KRWs just below and above the turning level, and just above the critical level are marked by A, B, and C, respectively. (b), (c) The width of the ellipse lines represent the PV anomaly strength.
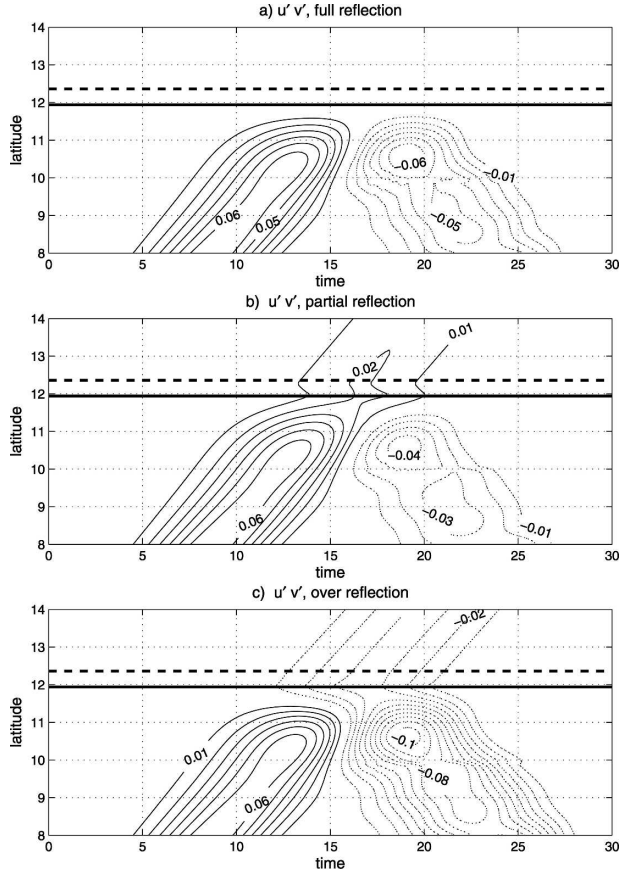

FIG. 6. Latitude–time sections of the momentum flux for the numerical runs done on the basic states of Figs. 3 and 5, as described in the text. (a) Full reflection, (b) partial reflection, (c) overreflection. The turning level and the critical level above it are marked by the horizontal lines. Negative values are dashed. Note that the wave propagation direction associated with a given sign of momentum flux depends on the sign of β. Note also that not all the domain is shown.

ward-propagating signal–reflection. The end result will be equal amounts of wave energy propagating north and south in region I, to form a standing wave, with an added phase shift at the turning level. This is illustrated schematically as well as numerically, in Figs. 5a and 6a, respectively. Also, in appendix C, we show that such a solution (a cosine/sine wave to the south of $y = y_0$ and exponential decay above) is consistent with the KRW formulation. The requirement for a sustainable structure also yields an expression for the phase shift at $y = y_0$ as a function of the problem parameters.

As we will see later on, an important point to note is that the anomaly in region II, for which $\beta/(U - c) < 0$, will have an anomalous ψ–q relation, meaning circulation in the evanescent region is dominated by the KRWs in region I below (with KRW A being the most dominant).

2) PARTIAL REFLECTION

When we now add another wave-propagation region (III), beyond $y = y_1 > y_0$ (as in Fig. 5b), the evolution will start as in the case of full reflection above, only now, the evanescent anomaly in region II will eventually reach $y = y_1$ and will induce a PV anomaly just to the north of it, in region III (KRW C; Fig. 5b). KRW C is created via the meridional wind field induced by the anomalies to its south, in this case, the evanescent region II. As shown above, however, the meridional wind field in region II is essentially controlled by KRW A, hence we can regard the whole evanescent region, plus KRW A, as one integrated region, and replace this whole region by a forward-directed evanescent kernel structure (EKS$_\uparrow$), which is a weaker version of KRW A. KRW C, which is in a region of $\beta < 0$, will therefore be a weaker eastward-shifted version of A. It in turn will create an eastward-shifted anomaly to the north, which will propagate onward—the partially transmitted wave train. KRW C will also weaken A (via a similar EKS$_\downarrow$) and the southward reflected wave train which it sets up (compared to the case of full reflection above). The end result will be partial transmission, and partial reflection, as expected. This is illustrated schematically as well as numerically, in Figs. 5b and 6b, respectively.

3) OVERREFLECTION

We now examine the case where wave region III is due to a jump in $U - c$, with β remaining the same as in region II (Fig. 5c). The anomalies will evolve in a manner similar to the case of partial reflection above, except that now, since $\beta > 0$ in region III, KRW C, which is induced by EKS$_\uparrow$, will be shifted westward, rather than eastward, with respect to KRW A. It in turn will induce a northward-propagating wave train, now with a westward phase tilt with y. In contrast to the case of partial reflection, KRW C will now strengthen KRW A (via an EKS$_\downarrow$), resulting in a stronger southward reflected wave train, compared to the case of full reflection—yielding overreflection. This is illustrated schematically as well as numerically, in Figs. 5c and 6c, respectively. As mentioned before, in order to examine the process of overreflection, we need to ignore the contribution of the jump in U to the PV gradient (which will in itself result in barotropic instability). Numerically, we do this by setting $\overline{q}_y = \beta$ [rather than $\overline{q}_y = \beta + \delta(y - y_1^-) - \delta(y - y_1^+)$]. We therefore also verified the qualitative nature of the results, by smoothing the jump in U, using a tanh, and taking into account the full contribution of U_{yy} to the PV gradient. These results are shown in Fig. 7.
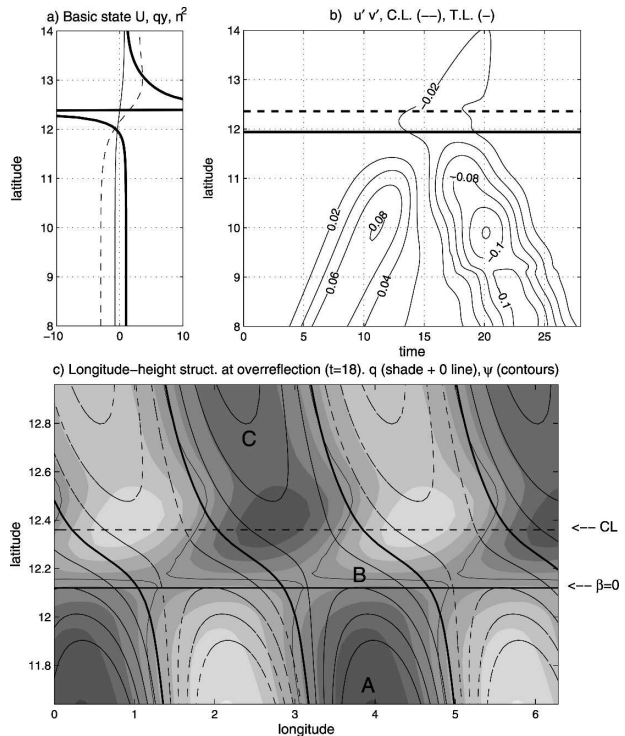
When the domain is bounded, this emitted energy

FIG. 7. (a) Numerical runs based on the KRW formulation, for the tanh basic state. Shown are U (thin solid line), $\overline{q}_y$ (dashed line), and n^2 (thick solid line). (b) Time–latitude plot of the momentum fluxes (similar to Fig. 5c). The turning level and critical level are marked by the horizontal dashed and solid lines, respectively. (c) A longitude–latitude plot of the PV (shaded grays, negative values are dark) and streamfunction (contours, negative values are dashed) anomalies, showing the X structure in the vicinity of the critical level (see text). The critical level and $\beta = 0$ levels are marked by the horizontal lines, and the approximate locations corresponding to kernels A, B, and C are marked (see text). (c) Note that the meridional domain is much smaller than (a) and (b).

can be reflected back, to form, under suitable wave geometry conditions, an exponentially growing normal mode (e.g., Lindzen 1988; see discussion later on). In an unbounded domain, however, this configuration will reach a steady state, in which KRWs A and C mutually reinforce each other, and at the same time, they pass the added wave energy northward and southward of the evanescent region.

The structure of PV in the evanescent region and in the vicinity of the critical surface is slightly more complex (not illustrated in Fig. 5c), as follows. While kernels A and C are tilted westward with y, the PV in the evanescence region and in the immediate vicinity of the critical level tilts eastward with y, because of the π phase shift at $y = y_0$. The meridional wind and streamfunction, on the other hand, do not change sign at $y = y_0$, hence they tilt westward with height. The ψ and q fields therefore form an X structure around the critical

level (illustrated numerically for tanh in Fig. 7c, and schematically in Fig. 8), so that they are essentially in phase (anomalous relation) at kernel B, and out of phase at kernel C. In between, they reach a $\pi/2$ phase difference at the critical level (for $q_y > 0$, ψ is $\pi/2$ to the west of q, so that v and q are in antiphase). This phase relation at the critical level is necessary to allow the PV perturbation to propagate with the flow at this level ($c = U$), because there is no self-induced phase propagation. Note, that according to Eq. (17a), this also necessarily leads to growth of the PV anomaly, via advection of the mean PV, and moreover, this is the optimal ψ–q configuration for growth. Furthermore, mathematically, having a $\pi/2$ phase between ψ and q requires c to have a nonzero imaginary component $\psi/q = -(U - c)/\beta$. The above steady state, which is used here for illustrative purposes only, therefore requires ψ to vanish at the critical level. Figure 7c shows these relations hold qualitatively, in a numerical integration, for a slightly less idealized basic state.

The above picture of overreflection is very similar to the two CRW approach. As an example, in the Charney model, wave propagation region I is squeezed into the surface region (see Fig. 2), and evanescent region II extends between the surface and the critical level. Indeed, in the Charney model, PV tilts eastward with height in most of the domain (with the strongest tilt near the critical surface), and the X structure between q and v is very clear (see Fig. 2 of Heifetz et al. 2004b). In the Charney model, the lower CRW (represented here by KRWs A and B via EKS↑) is essentially comprised of the surface temperature PV delta function, and the evanescent region PV above (which is in anti phase with the surface PV), while the upper CRW (here represented by KRW C) is essentially comprised of the wave region above the critical level.

4. Energy growth, the Orr mechanism, and the role of the critical level

One of the major differences between the OR and CRW approaches is the role of the critical level—it is central to the former, but is only implicit in the latter. In the previous section we showed, however, that the basic CRW interaction essentially applies to the vicinity of the critical level, suggesting that the mutual amplification of PV anomalies via their induced mean PV gradient advection is a source of growth in overreflection. On the other hand, Lindzen (1988) argued that the ultimate source of OR energy growth is the Orr mechanism, which operates in the vicinity of the critical level, but the exact way in which this works is left obscure. The classical Orr mechanism Orr (1907) is the change
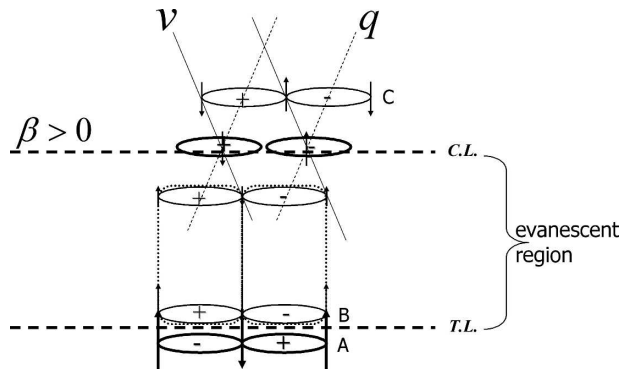

FIG. 8. A schematic of the X structure: The longitude–latitude structure of PV kernels (ovals) and total meridional wind (vertical arrows) in the vicinity of the turning level (TL; bold dashed line) and critical level (CL; bold dashed line) and the evanescent region in between. The slanted dashed and solid lines represent the PV and meridional wind phase lines, respectively. Note that at the turning level, PV changes sign but the meridional wind does not. Here, A, B, and C denote the PV kernels discussed in the text.

in kinetic energy that accompanies the tilting of PV anomalies by advection in a shear flow. In Orr's example of a pure plane wave the PV and streamfunction perturbations were aligned so that in a growing configuration they were both tilted against the shear. In the classical shear instability models, which include a mean PV gradient, however, we have an inherently different anomaly structure—an X configuration—where for a growing anomaly in the vicinity of the critical level the streamfunction is always tilted against the shear, while the PV anomaly is tilted with the shear [e.g., see Fig. 2 of Heifetz et al. (2004b) for the Charney model, and Fig. 7c here].

For the purpose of examining the sources of nonnormal optimal growth, Heifetz and Methven (2005) derived expressions for the energy and enstrophy growth in the two-CRW system for the Rayleigh problem, and showed a clear distinction between the mechanisms of PV amplitude growth via mutual CRW amplification, and growth due to a change in phase tilt that does not involve the mean PV gradient (the Orr mechanism). In the following section we expand this approach to clarify the role of the critical level, and to explicitly determine the source(s) of energy growth.

a. Energy growth in the CRW framework

The domain integrated perturbation energy[11] can be written as

---

[11] In the case of barotropic flow, the perturbation energy is the kinetic energy, but the analysis also holds for the baroclinic case, with energy being kinetic plus potential energy.

$$E = \frac{1}{2} \int_{y_s}^{y_n} \overline{(u_g^2 + v_g^2)} \, dy = -\frac{1}{2} \int_{y_s}^{y_n} \overline{(\psi q)} \, dy - \left[\frac{\overline{\psi u_g}}{2} \right]_{y_s}^{y_n},$$

$$(26)$$

where (y_s, y_n) indicate the southern and northern boundaries. When the integration is taken to be over the entire domain the second term on the RHS vanishes due to the boundary conditions (boundedness of the anomalies at infinity, the vanishing of meridional velocity at sidewalls, or pure plane waves under radiation boundary conditions). Note that evanescence regions, for which the ψ–q relation is anomalous, contribute negatively to the domain integrated energy (via the first term on the rhs).

For the two-CRW paradigm of section 2a(1), using Eqs. (4a), (4b), and (6); Eq. (26) yields (with the second term on the rhs assumed zero)

$$E = \frac{1}{4k} [Q_1^2 G_{11} + Q_2^2 G_{22} + Q_1 Q_2 (G_{12} + G_{21}) \cos\epsilon].$$

$$(27)$$

Recall that $G_{12} = G_{21}$, then

$$\dot{E} = \frac{1}{2k} [\dot{Q}_1 (Q_1 G_{11} + Q_2 G_{12} \cos\epsilon)$$

$$+ \dot{Q}_2 (Q_2 G_{22} + Q_1 G_{12} \cos\epsilon) - \dot{\epsilon} Q_1 Q_2 G_{12} \sin\epsilon].$$

$$(28)$$

Hence, the energy growth results either from enstrophy growth (the terms involving $\dot{Q}$) or from the change of the phase tilt between the two CRWs (the term involving $\dot{\epsilon}$). The latter term can be referred to as a generalized Orr mechanism, however, in the classical Orr paper (which discussed the evolution in the absence of a mean PV gradient) the change in phase tilt resulted solely from differential advection by the shear. When a mean PV gradient is present, Eqs. (8) and (5) yield

$$\dot{\epsilon} = -k\Delta U + (\overline{q}_{y_2} G_{22} - \overline{q}_{y_1} G_{11})$$

$$+ \left(\overline{q}_{y_2} G_{12} \frac{Q_1}{Q_2} - \overline{q}_{y_1} G_{21} \frac{Q_2}{Q_1} \right) \cos\epsilon, \quad (29)$$

where the first term on the rhs (with $\Delta U = U_2 - U_1$) represents the classical Orr mechanism, and the two additional terms are the contributions from the difference in the self–counter propagation (second term), and the mutual interaction (third term) of the two CRWs.

Substituting Eqs. (7a), (7b) for $\dot{Q}_1$ and $\dot{Q}_2$, and Eq. (29) for $\dot{\epsilon}$, Eq. (28) yields the initially unexpected result, that the energy growth due to enstrophy growth, exactly cancels the two CRW terms in $\dot{\epsilon}$, leaving only the shear term:

$$\dot{E} = \frac{1}{2} Q_1 Q_2 G_{12} \sin\epsilon \Delta U. \quad (30)$$

This means that the instantaneous energy growth does not depend explicitly on the existence of a mean PV gradient, but on the strength of the shear (ΔU) and the phase tilt between the two CRWs. The mean PV gradient affects the energy growth only indirectly, through Q and ϵ.

The following picture emerges. Energy growth occurs when the PV anomaly is tilted against the shear, but the shear itself will tend to destroy this growing configuration. The CRW interaction can, under the right conditions of counterpropagation, counteract the shearing, to maintain the growing configuration. What Eq. (30) implies, is that although this counterpropagation will also necessarily change the magnitude of the PV anomalies, the net effect of the CRW interactions on the energy growth will be zero, leaving only the shear to directly contribute to energy growth. This is true even for the case of normal-mode growth, where the phase locking is accompanied by a mutual CRW amplification. While this is maybe surprising, it can be thought of as a CRW mechanistic explanation of Lindzen's suggestion that "the wave geometry acts to restoke the Orr mechanism" (although strictly speaking the Orr mechanism is a nonmodal process, which requires the decrease of phase tilt with time).

The above derivation was done for an arbitrary CRW pair, but as is shown in appendix C, a time derivative of Eq. (26) indicates that the ψ–q correlation term in Eq. (26) depends only on the meridional variations of the zonal mean wind for the continuous case as well:

$$\dot{E} = \frac{1}{4} \int_{y=y_s}^{y_n} \left\{ \int_{y'=y_s}^{y_n} Q(y)Q(y')G(y,y') \sin\epsilon(y,y')[U(y) - U(y')] \, dy' \right\} dy. \quad (31)$$

Note that for the case of two localized CRWs the double integration yields the factor of 2 in Eq. (30). For clarity, we have dropped the boundary term $-[\overline{\psi u_g}/2)]_{y_s}^{y_n}$, which

simply represents the net contribution to $\dot{E}$ of all the anomalies outside of the domain of integration, since it vanishes under most standard boundary conditions

(boundedness, rigid sidewalls, pure plane wave radiation boundary conditions).

For comparison, the more common derivation of the energy growth includes a Reynolds stress term and a boundary wave energy flux term:

$$\dot{E} = -\int_{y_s}^{y_n} \overline{u_g v_g} U_y \, dy - [\overline{v_a \psi}]_{y_s}^{y_n}. \tag{32}$$

In this formulation, the interior term also depends explicitly on the shear and not on the PV gradient. The boundary term (wave energy flux divergence) also vanishes under boundedness or sidewall conditions. Under pure plane wave radiation boundary conditions, however, it does not vanish, and it is even necessary for maintaining steady state conditions. For example, we look at the special steady-state overreflection case of Fig. 3c. From standard wave considerations (cf. Table 1), the momentum flux is negative, thus the Reynolds stress contribution is positive, and comes only from the critical level (since all the shear is concentrated there). Also, $\overline{v_a \psi}$ is constant and directed northward in region III, and southward in region I (divergence), thus the energy flux term contributes negatively, allowing the steady state to be maintained.

A few points arise from comparing the two formulations [Eqs. (31) and (32)]. While in the common formulation energy growth depends on local shear, in the CRW formulation it depends on the meridional difference in the zonal mean wind between two regions, regardless of whether there is a local shear in each of them (provided of course there is a phase difference ϵ). For example, in the piecewise constant profile of Fig. 3c, the contribution to total energy growth comes from the interaction of PV anomalies between the two wave propagation regions (I and III). We note however, that the CRW interactions also account for processes like a wave energy flux, via the mechanism described in section 3a. In steady state, therefore, the amplification via the CRW interaction between regions I and III should also be carried away via the CRW interactions within each region [the integral of Eq. (31) includes both terms on the rhs of Eq. (32), and should vanish]. This is stated more explicitly in section 5.

b. The role of the critical level

Looking more closely at the role of the critical level, from Eq. (31), it seems that this region contributes negatively to the energy growth, since q is tilted with the shear there. This reasoning, however, is not quite correct, since it assumes we take both integrals, over y and y', only in the vicinity of the critical surface. Physically it means we have only considered the ψ anomalies

induced by PV anomalies in the vicinity of the critical level. In section 3b we have shown on the other hand, that in the evanescent region the streamfunction is always induced by PV anomalies outside it. We therefore need to take into account ψ induced by the entire PV field [take the y' integral in Eq. (31) over the entire domain]. When this is done, the contribution of the critical level region to $\dot{E}$ can become positive, as expected for growth.

What is then the uniqueness of the critical level? In the standard formulation, it is the only place where $\overline{\psi v_a}$ is divergent [cf. the EP theorem, Eqs. (16), (18)].[12] In the CRW framework, the critical level is the only place where two adjacent kernels can be mutually amplifying. It can be considered an energy source in the sense that in all other regions, two adjacent kernels interact such that one amplifies and one weakens the other—an energy transfer. To illustrate we again use piecewise constant basic states. In the special case where the jump in β and in U occur at the same point, it is clear that the two kernels on both sides of the critical levels are the only ones that can mutually amplify. In all other regions, the interaction is as illustrated in the case of wave propagation (section 3a; Fig. 4). When the jumps in β and U are not at the same place, this picture is slightly more complicated. The region between the two jumps is evanescent, meaning the PV inside it (e.g., kernel B in Fig. 8) is directly linked to the kernels just outside it (e.g., kernel A in Fig. 8), with an anomalous ψ–q relation externally imposed. Taking this into account (that an amplification of kernel B results in a stronger cyclonic circulation around negative PV centers, via the direct link to kernel A) we get that the kernels on two sides of the critical level are now the only two adjacent kernels which are mutually amplifying. We note that this mutual amplification, which occurs even though the kernels are in regions of similar-signed PV gradients, requires the influence of kernel A (e.g., a wave-propagation region), which is in an oppositely signed PV gradient region (the Charney–Stern condition for instability).

5. Normal-mode growth from the combined CRW–OR perspective

So far we have discussed the mechanism by which an anomaly propagating toward the critical surface ampli-

---

[12] For the overreflection example of Figs. 3c and 5c, wave considerations (Table 1) imply that $\overline{\psi v_a}$ is northward in region III and southward in region I. In the evanescent region II, where wave considerations do not apply, the streamfunction at the edges of the domain is dominated by the wave region closest to it, so that the general phase tilt has to be eastward with y.

fies upon reaching it, yielding overreflection. In addition, OR theory discusses how this process yields normal-mode instability (e.g., Lindzen 1988)—if the over-reflected anomaly is reflected back from some surface, under the right quantization conditions (on zonal wave-number and phase speed), the reflected and overre-flected waves will constructively interfere to yield a growing normal mode (Fig. 9). Lindzen and Tung (1978) further suggest an upper limit to the normal-mode growth rate based on this picture:

$$kc_i = \frac{\ln(R)}{2\bar{\tau}}. \tag{33}$$

This formula, which is essentially a laser formula, is obtained by assuming the perturbation amplifies by a factor of R each time it travels back and forth in region I between the overreflecting and the reflecting surfaces (distanced H from each other); $\bar{\tau} = H/C_{g_I}$ in Eq. (33) is the time it takes the wave to propagate this distance.

It is not yet clear how the KRW amplification [cf. Eq. (12a)] relates to the nondimensional overreflection co-efficient R. Some physical insight can be gained by more explicitly formulating the steady-state example from a KRW perspective. In section 3c(3) we saw that the overreflection case is essentially a mutual amplifi-cation of the two KRWs A and C, on the two sides of the evanescent region II (cf. Fig. 5c). A steady state can be obtained despite this mutual amplification, provided the extra perturbation energy is carried away by the overreflected and transmitted waves. This can be for-mulated schematically by considering the amplitude evolution of two KRWs A and C, via Eqs. (7a) and (7b) with additional sources and sink terms:

$$\frac{\dot{Q}_A}{Q_A} = -\bar{q}_{y_A} G_{CA} \sin\epsilon \frac{Q_C}{Q_A} + \frac{1 - |R|}{\tau_A} = 0, \quad (34a)$$

$$\frac{\dot{Q}_C}{Q_C} = \bar{q}_{y_C} G_{AC} \sin\epsilon \frac{Q_A}{Q_C} - \frac{1}{\tau_C} = 0. \quad (34b)$$

In addition to the KRW mutual amplification, A is also affected by the approach of the incident wave from the south (a source term) and the southward propagation via the reflected wave (a sink term) whose amplitude is R times the incident wave. Likewise, C is affected by the northward propagation via the transmitted wave (a sink). The KRW amplitudes are normalized by the am-plitude of the incident wave source. We have assumed that the perturbations are transmitted to and from the two KRW locations with a time scale τ, which is a pa-rameter of the problem. The only time scale that ap-pears in this problem (except for the mutual amplifica-tion one) is $\tau = 2\pi/|nC_g|$, corresponding to the time it
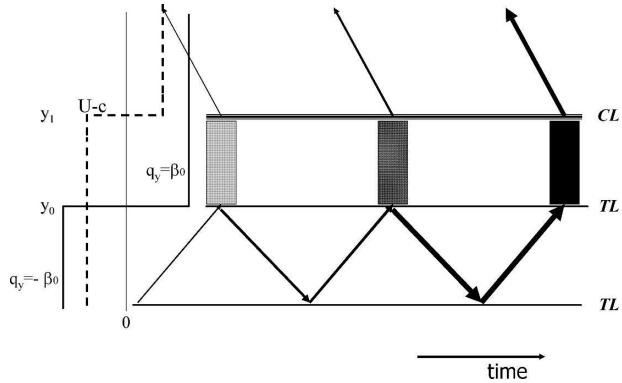

FIG. 9. A time–latitude schematic of normal-mode structure from the OR point of view (as suggested by Lindzen and Coau-thors, see text), for the basic state of Figs. 3c and 5c. Arrows represent wave propagation, with arrow width representing am-plitude. The rectangles represent the tunneling through the eva-nescent region. Time increases to the right.

takes the wave to propagate a full wavelength at group speed. This can be justified by noting that we also need to take into account the mutual interactions between their immediate vicinities when considering the inter-action between kernels A and C. In appendix A we saw that in a wave train, the interactions within a single wavelength dominate. Equations (34a) and (34b) can then be solved to yield $|R|$

$$|R| = 1 - \bar{G}^2 \sin^2\bar{\epsilon} \bar{q}_{y_A} \bar{q}_{y_C} \tau_A \tau_C, \tag{35}$$

where $\bar{G} \sin\bar{\epsilon} \equiv [G_{AC} G_{CA} \sin^2\bar{\epsilon}]^{1/2} > 0$ is essentially the area weighted contribution of the Green function and phase tilt between A and C, taking into account the evanescent region in between [EKS↑ and EKS↓; cf. sec-tion 3c(3)].

The resulting expression for $|R|$ makes physical sense in a few respects. When the PV gradients in the two regions are of different signs, we get overreflection ($|R| > 1$), consistent with the Charney and Stern (1962) criterion for instability, while same signed PV gradients yield partial reflection (as in the examples of section 3c). A barotropic structure, meaning there is no phase tilt between A and C ($\epsilon = 0$) is associated with full reflection $|R| = 1$. A smaller evanescent region (closer A and C) will result in larger overreflection, via a larger $\bar{G}$. To the extent that the anomaly in the evanescent region is independent of the basic states in the two wave-propagation regions (most likely this is not the case), larger wave-propagation time scales τ_A, τ_C, result in larger overreflection, because there is more time for the mutual amplification to work. We note that it is hard to numerically confirm Eq. (35), since there is no way (to the best of our knowledge) to numerically ob-

tain a strict steady-state solution of the type presented here. This is due to the complications of dealing with the boundaries and representing both a steady source and radiation out of the domain.[13]

How expression (35), which was obtained under steady-state assumptions, is related to the overreflection coefficient in the unstable normal modes [e.g., in Eq. (33)] is not entirely obvious, since the solution, and therefore $\bar{\epsilon}$ may depend differently on the basic-state parameters. However, the essence of the above picture still holds—that $|R|$ depends on the balance between the mutual amplification on the two sides of the evanescent region, and the rate at which this amplification is carried away.

Finally, this picture of a mutual amplification between two regions, is basically captured by the two-CRW picture of Heifetz et al. (2004a,b) where the northern CRW essentially represents the northern wave region (plus the evanescent region contribution, $EKS_\downarrow$), while the southern CRW represents the southern wave region (also including the evanescent region, $EKS_\uparrow$). We note that each of these CRWs includes the mutual interactions of kernels within its own region, so that the propagation of energy away from KRWs A and C (the overreflected, reflected, and transmitted waves), as well as the restoking from the south (the incident wave), are embedded into the CRWs.

6. Concluding remarks

Shear instability is a fundamental mechanism in fluid dynamics in general and atmospheric dynamics in particular, yet our mechanistic understanding of how it works is not as complete as our understanding for instance of buoyant convection. Moreover, the two main approaches to understanding shear instability, seem at a first glance, fundamentally different.

While the CRW approach stems from a PV thinking where the conserved quantity is enstrophy, in OR the PV equation is converted to a streamfunction wave equation, and the central source of growth, thought to be the Orr mechanism, is associated with energy growth. The nonlocality in CRW is the action-at-a-distance nature of PV anomalies that induce a flow everywhere, while in OR, nonlocality is a fundamental

aspect of wave propagation which is affected nonlocaly by the wave geometry. While CRW is associated with a phase propagation in the direction of the mean flow, OR is associated with a group velocity (energy propagation) across the shear. In the CRW approach the phase speed at a given point results from the multiple phase locking interactions between the CRWs, while in the OR approach, the phase speed is a parameter of the perturbation (similar to the zonal wavenumber), which directly affects the wave geometry. Correspondingly, the critical level in the CRW view is simply the level at which the total induced velocity along the mean positive (negative) PV gradient is exactly antiphase (in phase) with the PV perturbation, but in the OR view it is at the heart of the instability.

To gain some understanding of how such seemingly different views can explain the same phenomena and why these differences arise, we rationalize the OR theory in terms of CRW thinking. To do this we use the generalization of the CRW approach to multiple infinitesimal PV kernels (KRWs) to understand the building blocks of OR. There are two basic types of KRW interactions. The first is a symmetric mutual interaction, where each KRW enhances or weakens the other leading correspondingly to growth or decay. This interaction occurs when the two KRWs are in regions of opposite-signed PV gradients. The second, which occurs when the KRWs are in same-signed regions of PV gradient, is an asymmetric interaction, where one KRW enhances while the other weakens its counterpart. This interaction is essentially a transfer of energy, which can result in cross shear-wave propagation.

Using this thinking, we first explain qualitatively the mechanistics of cross-shear wave propagation, as well as the exponential decay in an evanescent region, for a KRW source of given zonal wavenumber and phase speed. We also examine quantitatively how a pure plane wave, as well as exponential decay are maintained. In particular, we explain how the anomalous $\psi - q$ relation, which necessarily arises in evanescent regions for which $(\bar{q}_y/U - c) < 0$, is sustained. A side result is a quantitative understating of what is needed for coherent self-sustained waves to exist, with a given phase speed.

We then examine the cross-shear behavior in different wave geometries (full, partial, and overreflection). Although our calculations were done for simplified piecewise continuous basic states where the shear is concentrated into one point, we expect our results to hold for slowly varying basic states for which the WKB approximation can be applied. This is supported by numerical calculations we have done using a tanh profile

---

[13] We have also examined the relation, in the numerical solution, between an estimated overreflection coefficient (obtained from the momentum flux; cf. Fig. 6) and the various parameters appearing in Eq. (35), but no clear relation has emerged. This is not surprising though, given that the solution is not a steady state, and it is not clear how the steady state $\bar{G}$ and ϵ relate to the time-dependent solution.

(Fig. 7). We find that overreflection is basically a two-KRW mutual amplification, with the main KRWs being at the two sides of the evanescent region, which forms on one side of the critical level.

The critical level is unique in that it is the only point at which we have a mutual amplification between two adjacent KRWs, taking into account the anomalous $\psi - q$ relation on one side of the critical surface. This anomalous relation, which is critical for this interaction to be a mutual amplification rather than an asymmetric transfer of energy, can only exist because of the existence of another wave propagation region, which in turn, can only exist if the PV gradient changes sign. Thus, although the critical surface appears to be the center of energy growth (as it is the level at which the EP flux always diverges), it is not so by itself. Rather, the entire evanescent region, between the turning surface and the critical surface appears to be the center of the instability; that is, the amplification between KRWs at its two sides contributes most significantly to the growth.

We further examine the sources of energy growth, in particular the role of the Orr mechanism, which Lindzen (1988) suggested as the ultimate source. The picture in terms of energy growth is complicated by the fact that the wave geometry yields an X structure between PV and streamfunction (the PV is tilted with the shear while the streamfunction is tilted against the shear), in the vicinity of the critical level. The southern part of the X (the evanescence region) contributes negatively to the energy. Furthermore, since the PV is tilted eastward with shear at the critical level, the near KRW interaction contributes negatively to energy growth there. Only the far field action at-a-distance contributes positively. While the source of growth is essentially a CRW interaction, the kinetic energy even for such an interaction stems from the shear, in an Orr-like mechanism. Thus we obtain a generalized view of the restoking of the Orr mechanism: Growth depends explicitly on the shear, as in the Orr mechanism, but KRW interactions are responsible for holding the perturbation at a configuration that allows it to contribute to growth (a restoking).

Overreflection theory relates the maximum normal growth rate to the overreflection coefficient via a laser formula. By assuming that the increase in PV amplitude at both sides of the critical level via the CRW mutual amplification is balanced by an outward propagation, we can relate the overreflection coefficient to the CRW viewpoint. Moreover, for the steady-state case, we obtain an expression for R that qualitatively makes sense. The two-CRW picture, for example in the Charney

problem, and its relation to overreflection, becomes clearer. Growth (the overreflection) occurs as a result of the mutual interaction between the KRWs at the sides of the evanescent region. These KRWs form the central parts of the two CRWs (cf. Heifetz et al. 2004a,b). The rest of the CRWs encompass the wave regions, and the propagation away from the critical level is embedded into them via the mutual KRW interactions within each CRW. Thus, for instance the lower Charney CRW includes the surface PV delta function KRW (which represents the lower region of the wave propagation) and a decaying structure of interior PV with opposite sign, which represents the exponential decay in the evanescent region. The upper CRW, is concentrated at the wave-propagation region just above the critical level.

To conclude, by relating between the two approaches we gain additional understanding of the mechanistic process behind shear instability and the sources of energy, as well as basic wave behavior. This study has dealt with Rossby wave interactions, however OR can also explain a more general range of shear instabilities, including those associated with gravity waves. Since CRW interaction has proven to be a suitable approach to interpreting Rossby OR, it is likely that a similar gravity wave building block approach can explain the gravity wave instabilities as well.

APPENDIX A

A KRW Description of Simple Wave Configurations

Here we demonstrate the consistency of the KRW approach in some simple examples.

a. Wave propagation

Consider a Rossby plane wave, in an unbounded horizontal domain of constant β and U, of the form $\psi = \Psi e^{iny}e^{ik(x-ct)}$, whose PV structure is $q = -k^2\psi + \psi_{yy} = Qe^{i\epsilon}e^{ikx}$, where $Q = -(k^2 + n^2)\Psi$, and $\epsilon = ny - kct$. For such an infinite domain the Green function $G(y, y')$ which satisfies: $-(1/k)(-k^2 + \partial/\partial y^2)G(y, y') = \delta(y' - y)$ [cf. Eq. (10b)], is $G(y, y') = e^{-k|y-y'|}/2$. First, in

order to verify indeed that $\psi = -(1/k)\int_{-\infty}^{\infty} q(y')G(y, y') \, dy'$ (Eq. 11), one needs to show [after substitution of q, $G(y', y)$, and ψ above], that $\int_{-\infty}^{\infty} e^{iny'}e^{-k|y-y'|} \, dy' = (2k/k^2 + n^2)e^{iny}$, which is straightforward to verify.

Next, we show that the wave preserves its constant PV amplitude, via Eq. (12a). We recall that $\sin \epsilon(y, y', t) = \sin[\epsilon(y, t) - \epsilon(y', t)] = \sin[n(y - y')]$; thus Eq. (12a) becomes

$$\frac{\dot{Q}}{Q}(y) = \frac{\beta}{2}\left\{ \int_{-\infty}^{y} e^{-k(y-y')} \sin[n(y - y')] \, dy' \right.$$
$$\left. + \int_{y}^{\infty} e^{-k(y'-y)} \sin[n(y - y')] \, dy' \right\}$$
$$= \frac{\beta(n - n)}{2(k^2 + n^2)} = 0. \tag{A1a}$$

For the case of northward propagation (positive β and n; cf. Table 1), at a given latitude y, the overall meridional velocity induced from below (the first integral) contributes to amplify the PV kernel amplitude $Q(y)$, whereas the overall contribution from above (the second integral) acts to decrease it. The two contributions exactly cancel each other, so that the amplitude remains constant. Moreover, because of the symmetry of the wave, the contributions from above and below cancel out for any finite interval Δy; that is,

$$\frac{\dot{Q}}{Q}(y) = \beta \int_{y-\Delta y}^{y+\Delta y} \frac{Q(y')}{Q(y)} G(y, y') \sin\epsilon(y, y', t) \, dy' = 0. \tag{A1b}$$

In addition, via Eq. (12b) we verify that the wave indeed propagates with a constant phase speed c, which satisfies the dispersion relation 14. Using $\dot{\epsilon} = -kc$, the definitions of Q, ϵ and $G(y', y)$, and rearranging Eq. (12b), this boils down to showing that

$$\frac{U - c}{\beta} = \frac{1}{k} \int_{-\infty}^{\infty} \frac{Q(y')}{Q(y)} G(y, y') \cos\epsilon(y, y', t) \, dy'$$
$$= \frac{1}{k^2 + n^2}, \tag{A2a}$$

which is straightforward to verify.

Finally, the near KRW interaction within $y \pm \Delta y$ is obtained when the integral is taken between these two limits:

$$\frac{U - c}{\beta} = \frac{1}{k^2 + n^2}\left\{ 1 - e^{-k\Delta y}\left[\cos(n\Delta y) - \frac{n}{k} \sin(n\Delta y) \right] \right\}. \tag{A2b}$$

Integrating over a full meridional wavelength (i.e., $\Delta y = \pi/n$) yields

$$\frac{U - c}{\beta} = \frac{1 + e^{-\pi k/n}}{k^2 + n^2}. \tag{A2c}$$

Taking, for example, a unit aspect ratio, we get a phase speed that is within about 4% of c, implying the overall contribution of all the remote KRWs, to the phase propagation of $y \pm \pi/n$ can be neglected.

b. Wave evanescence

1) THE ANOMALOUS CASE ($\psi \propto q$) OF STRONG EVANESCENCE

From the KRW perspective, the simple evanescent structure of Eqs. (25a) and (25b), which is centered on $y = y_0$ (for simplicity we take $y_0 = 0$), is composed of continuous KRWs that are in phase and a KRW δ function at $y = 0$ which is antiphased with them.[14] According to Eq. (12a), this structure is neutral since all the KRWs are in phase or anti phased. Writing Eq. (25b) as $q(y) = Q(y) = [(m^2 - k^2)e^{-m|y|} - 2m\delta(y)]\Psi$, and recalling that $(m^2 - k^2) = -\beta/(U - c) > 0$, Eq. (12b) yields after some manipulation:

$$\frac{(U - c)}{\beta} = \frac{e^{m|y|}}{(m^2 - k^2)}\left[\frac{(m^2 - k^2)}{2k} \int_{-\infty}^{\infty} e^{-m|y'|}e^{-k|y-y'|} \, dy' - \frac{m}{k} e^{-k|y|} \right] = \frac{-1}{m^2 - k^2}, \quad \text{for} \quad y \neq 0. \tag{A3}$$

We see that this structure indeed assumes a single phase speed, which is consistent with the dispersion relation (14). Moreover, we get that $(U - c)/\beta$ is indeed negative, implying an anomalous ψ–q relation. The first term in the brackets, which represents the overall contribution from the in-phase column of KRW kernels is positive (propropagation), while the second term, which represents the contribution from the initial PV δ function kernel is negative (counterpropagation). The fact that the sum of these two terms is negative means that the influence of the δ kernel overwhelms the overall effect of the PV column, thus allowing the anomalous ψ–q relation.

---

[14] To avoid confusion, note that according to Eq. (25b), the KRWs $\tilde{q}(y') = q(y')\delta(y' - y)$, which have units of PV per length, have a finite amplitude $q(y') = |\beta/(U - c)|e^{-m|y'|}\Psi$ at $y \neq 0$ and an infinite value $q(y' = 0) = -2m\delta(y')\Psi$ at $y = 0$.

For $y = 0$, Eq. (A3) does not hold since in the case of finite mean PV gradient and finite meridional velocity, one cannot obtain an infinite PV anomaly from linearized PV dynamics. Hence, such a δ function PV kernel must be forced on the solution from an external source.[15]

2) THE NORMAL CASE ($\psi \propto -q$) OF WEAK EVANESCENCE

We have already mentioned above, that a necessary condition for a wave signal to be able to propagate away from a wave source of zonal wavenumber k and phase speed c, is that the basic state be able to support a self-sustained structure moving at that phase speed. If the forcing kernel is needed to maintain the phase speed, no signal can propagate away, and the only possible steady solution is a perturbation which decays away from the wave source, so that the source can indeed sustain it (i.e., wave evanescence). The case of the strong anomalous evanescence is easy to understand in this way—the KRWs are propropagating, hence for any anomaly, the integrated self advection can never balance the relative mean flow advection $U - c$ since they are both in the same direction.

The case of weak normal evanescence, however, is more subtle because the self-induced propagation is counter the mean flow advection, but it is too weak to fight it. This can be shown by examining the evolution of a single-wavelength anomaly in the y direction ($\lambda = 2\pi/n$). Equations (A1b) and (A2b) indicate that such an anomaly is a self-sustained structure that does not grow in amplitude and propagates with a phase speed c_{pack} given by Eq. (A2c). According to this equation, the self-induced phase speed of this single-wavelength wavepacket is $c_{pack} - U = -\beta(1 + e^{-\pi k/n})/(k^2 + n^2)$. We also see from Eq. (A2c) that for a given k, the self-propagation speed will be largest at $n = 0$, with an absolute value of β/k^2. This makes physical sense since in this configuration all KRWs are in phase so their mutual interaction is optimal. This upper limit on the possible self-sustained phase speed is exactly the condition for a basic state to support wave propagation: $(U - c)/\beta < 1/k^2$. From this viewpoint, when we turn on a wave source with a phase speed of c, the meridional wavenumber n, which the anomaly assumes is that for which the self-induced phase speed exactly equals that of the wave source (provided of course that $\bar{q}_y/(U - c) > 0$).

The wave evanescence limit can also be obtained from Eq. (A3), by removing the contribution from the PV delta function at $y = 0$:

$$c = U - \frac{\beta}{(k^2 - m^2)}\left[1 - \frac{m}{k}e^{-(k^2 - m^2)|y|}\right]. \quad (A4)$$

We see first of all, that we cannot have a coherent self-sustained evanescent structure ($m \neq 0$) of the form described by Eq. (A4), since the phase speed c is a function of y. In fact, only $m = 0$ can yield a coherent structure, for which the self propagation can resist the mean flow advection. This is easily seen when we substitute $k^2 - m^2 = \beta/(U - c) > 0$ in Eq. (A.4). The case of $m = 0$ is the simplest zonal Rossby wave structure where $c = U - \beta/k^2$.

c. Full reflection

For the steady-state configuration of full reflection ($R = |R|e^{i\theta}, |R| = 1$), Eq. (19) can be rewritten (taking $y_0 = 0$) as

$\psi(y)$

$$= \begin{cases} e^{-iny} + e^{i\theta}e^{iny} = 2\cos\left(ny + \dfrac{\theta}{2}\right)e^{i(\theta/2)} & \text{for} \quad y < 0 \\[2mm] (1 + e^{i\theta})e^{-my} = 2\cos\dfrac{\theta}{2}e^{i(\theta/2)}e^{-my} & \text{for} \quad y > 0 \end{cases},$$

(A5a)

whose PV structure is

$$q(y) = \begin{cases} -(n^2 + k^2)\psi = -\dfrac{\beta_0}{|U - c|}\psi & \text{for} \quad y < 0 \\[2mm] (m^2 - k^2)\psi = \dfrac{\beta_0}{|U - c|}\psi & \text{for} \quad y > 0 \end{cases}.$$

(A5b)

Since $(n^2 + k^2) = (m^2 - k^2) = \beta_0/|U - c| > 0$, the PV jumps and changes sign at $y = 0$, whereas the streamfunction is continuous.

The structure is untilted for all y and hence according to Eq. (12a) is neutral. This means that if we consider, for instance, the KRW just below $y = 0$, then the influence on its growth by the incident wave is canceled out exactly by the decaying influence of the reflecting wave (and is not affected by the antiphased PV column above it at $y > 0$).

The phase difference θ, between the incident and the reflecting waves can be obtained by substituting Eq. (A5b) in Eq. (12b) [with $\dot{\epsilon} = -kc$, and $G(y, y') = e^{-k|y-y'|}/2$] to get

---

[15] A similar evanescent structure does, however, represent the lower CRW in the Charney (1947) model, since the surface has an infinite mean PV gradient in opposite sign to the finite mean PV gradient above it in the interior (Heifetz et al. 2004b).

$$-k\frac{(U-c)}{\beta_0} = \int_{-\infty}^{0}\frac{\cos\left(ny'+\dfrac{\theta}{2}\right)}{2\cos\dfrac{\theta}{2}}e^{ky'}\,dy'$$

$$-\frac{1}{2}\int_{0}^{\infty}e^{-(k+m)y'}\,dy', \tag{A6}$$

yielding after some algebra that $\tan(\theta/2) = m/n$.

APPENDIX B

Numerical Calculations

We use a discretized KRW model to verify some of our results. These numerical calculations involve discretizing Eqs. (12a) and (12b), using Eq. (11), and forward integrating in time, from a specified initial PV distribution. The numerically discretized Green functions are defined as in appendix A, but with a $1/\Delta z$ factor to represent the vertical integration. We use a third-order Adams–Bashford method (we follow Durran 1991) for the time integration.

A KRW source with phase speed c_r is implemented by initializing the model with an amplitude and phase of zero everywhere, except at a source level where $Q(\mathrm{src}) = 1$ and $\epsilon(\mathrm{src}) = -\pi/2$, and setting $\dot{\epsilon}(\mathrm{src}) = -kc_r$. The initial $-\pi/2$ phase is necessary, otherwise no anomaly will develop beyond the source level [an initial phase of zero would yield $\dot{Q} = 0$ at all levels besides the source; cf. Eq. (12a)]. We force the phase speed either for the duration of the integration, or for a time period of $\tau = (\lambda_y/C_{g_y}) = (2\pi/nC_{g_y})$ [with C_{g_y} and n defined using Eqs. (14) and (15a)–(15b)], which is the time required for a full wavelength to propagate away from the source. In the latter case, a single wavelength is formed and continues to propagate with an approximate phase speed c_r, consistent with Eq. (A2c). While the two versions yield qualitatively similar results, the latter is cleaner.

To avoid the complications of a back reflection from the southern boundary, we also initialize the model by specifying the initial PV amplitude and phase to be: $Q(y_{\mathrm{in}}) = 1$, $\epsilon(y_{\mathrm{in}}) = ny$, and zero otherwise, where y_{in} is a region of single wavelength width: $y(\mathrm{src}) - \lambda/2 \le y_{\mathrm{in}} \le y(\mathrm{src}) + \lambda/2$. This corresponds to the single wavelength anomaly of Eq. (A2c), which propagates northward (recall that $\beta < 0$ in the source region), and zonally with a phase speed of c_r. This initialization yields the cleanest results since the initial anomaly propagates only northward. This was used to create Figs. 6 and 7.

We tested the model by comparing to a known analytical solution, as well as by comparing to results of the normal-mode eigenvalue solver of Harnik and Lindzen (1998). We find the solutions converge for a reasonable vertical resolution [e.g., the Charney (1947) normal modes have been fully recovered with 100 layers].

APPENDIX C

KRW Description for Energy Growth

For vanishing meridional velocity at the boundaries, Eqs. (26), (4), and (11) indicate that

$$E = -\frac{1}{2}\int_{y_s}^{y_n}(\overline{\psi q})\,dy = \frac{1}{4k}\int_{y=y_s}^{y_n}Q(y)$$

$$\times\left[\int_{y'=y_s}^{y_n}Q(y')\cos\epsilon(y,y')G(y,y')\,dy'\right]dy. \tag{C1}$$

Therefore

$$\dot{E} = \mathrm{I} + \mathrm{II} \tag{C2a}$$

where

$$\mathrm{I} = \frac{1}{4k}\int_{y}\int_{y'}[\dot{Q}(y)Q(y') + \dot{Q}(y')Q(y)]\cos\epsilon(y,y')G(y',y)\,dy'\,dy = \frac{1}{2k}\int_{y}\int_{y'}\dot{Q}(y)Q(y')\cos\epsilon(y,y')G(y,y')\,dy'\,dy, \tag{C2b}$$

since $G(y', y) = G(y, y')$ and $\cos\epsilon(y, y') = \cos\epsilon(y', y)$, and the second integral is

$$\mathrm{II} = -\frac{1}{4k}\int_{y}\int_{y'}\dot{\epsilon}Q(y)Q(y')\sin\epsilon(y,y')G(y',y)\,dy'\,dy. \tag{C2c}$$

Using Eq. (12a), integral I can be rewritten as

$$\mathrm{I} = \frac{1}{2k}\int_{y}\int_{y'}\int_{y''}\overline{q}_y(y)Q(y')Q(y'')\cos\epsilon(y,y')\sin\epsilon(y,y'')G(y',y)G(y'',y)\,dy''\,dy'\,dy. \tag{C3}$$

Using Eq. (12b), integral II can be rewritten as

$$II = III + IV + V, \tag{C4a}$$

where

$$III = \frac{1}{4} \int_y \int_{y'} Q(y)Q(y') \sin \epsilon(y, y') G(y', y)[U(y) - U(y')] \, dy' \, dy, \tag{C4b}$$

$$IV = -\frac{1}{4k} \int_y \int_{y'} \int_{y''} \overline{q}_y(y)Q(y')Q(y'') \cos \epsilon(y, y'') \sin \epsilon(y, y') G(y', y) G(y'', y) \, dy'' \, dy' \, dy, \tag{C4c}$$

$$V = \frac{1}{4k} \int_y \int_{y'} \int_{y''} \overline{q}_y(y')Q(y)Q(y'') \cos \epsilon(y', y'') \sin \epsilon(y, y') G(y', y) G(y'', y') \, dy'' \, dy' \, dy . \tag{C4d}$$

Please note that since $\sin \epsilon (y, y') = -\sin \epsilon (y', y)$ exchanging between the dummy indices (y, y') in integral V gives: $V = IV$. Also note that exchanging between the indices (y', y'') in IV yields $I = -2 \, IV$ and consequently

$$\dot{E} = I + II = I + III + IV + V = I + III + 2IV$$
$$= I + III - I = III, \tag{C5}$$

and indeed integral III is the rhs of Eq. (31).

REFERENCES

Bishop, C. H., and E. Heifetz, 2000: Apparent absolute instability and the continuous spectrum. *J. Atmos. Sci.,* **57,** 3592–3608.

Bretherton, F. P., 1966: Baroclinic instability, the short wave cut-off in terms of potential vorticity. *Quart. J. Roy. Meteor. Soc.,* **92,** 335–345.

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.,* **4,** 135–163.

——, and M. E. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.,* **19,** 159–172.

Davies, H. C., and C. H. Bishop, 1994: Eady edge waves and rapid development. *J. Atmos. Sci.,* **51,** 1930–1946.

de Vries, H., and J. D. Opsteegh, 2005: Optimal perturbations in the Eady model: Resonance versus PV unshielding. *J. Atmos. Sci.,* **62,** 492–505.

——, and ——, 2006: Dynamics of singular vectors in the semi-infinite Eady model: Nonzero β but zero mean PV gradient. *J. Atmos. Sci.,* **63,** 547–564.

——, and ——, 2007a: Resonance in optimal perturbation evolution. Part I: Two-layer eady model. *J. Atmos. Sci.,* **64,** 673–694.

——, and ——, 2007b: Resonance in optimal perturbation evolution. Part II: Effects of a nonzero mean PV gradient. *J. Atmos. Sci.,* **64,** 695–710.

Dirren, S., and H. C. Davies, 2004: Combined dynamics of boundary and interior perturbations in the Eady setting. *J. Atmos. Sci.,* **61,** 1549–1565.

Drazin, P. G., and W. H. Reid, 1981: *Hydrodynamic Stability.* Cambridge University Press, 527 pp.

Durran, D. R., 1991: The third-order Adams–Bashforth method: An attractive alternative to leapfrog time differencing. *Mon. Wea. Rev.,* **119,** 702–720.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus,* **1,** 33–52.

Eliassen, A., and E. Palm, 1961: On the transfer of energy in stationary mountain waves. *Geofys. Publ.,* **22,** 1–23.

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.

Harnik, N., and R. S. Lindzen, 1998: The effect of basic-state potential vorticity gradients on the growth of baroclinic waves and the height of the tropopause. *J. Atmos. Sci.,* **55,** 344–360.

Heifetz, E., and J. Methven, 2005: Relating optimal growth to counterpropagating Rossby waves in shear instability. *Phys. Fluids,* **17,** 064 107, doi:10.1063/1.1937064.

——, C. H. Bishop, and P. Alpert, 1999: Counter-propagating Rossby waves in the barotropic Rayleigh model of shear instability. *Quart. J. Roy. Meteor. Soc.,* **125,** 2835–2853.

——, C. B. B. Hoskins, and J. Methven, 2004a: The counter-propagating Rossby-wave perspective on baroclinic instability. I: Mathematical basis. *Quart. J. Roy. Meteor. Soc.,* **130,** 211–232.

——, J. Methven, C. Bishop, and B. Hoskins, 2004b: The counter-propagating Rossby-wave perspective on baroclinic instability. II: Application to the Charney model. *Quart. J. Roy. Meteor. Soc.,* **130,** 233–258.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877–946.

Kim, H. M., and M. C. Morgan, 2002: Dependence of singular vector structure and evolution on the choice of norm. *J. Atmos. Sci.,* **59,** 3099–3116.

Kuo, H. L., 1949: Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Atmos. Sci.,* **6,** 105–122.

Lindzen, R. S., 1988: Instability of plane parallel shear-flow (toward a mechanistic picture of how it works). *Pure Appl. Geophys.,* **126,** 103–121.

——, and K. K. Tung, 1978: Wave overreflection and shear instability. *J. Atmos. Sci.,* **35,** 1626–1632.

——, and A. J. Rosenthal, 1981: A WKB asymptotic analysis of baroclinic instability. *J. Atmos. Sci.,* **38,** 619–629.

——, and J. W. Barker, 1985: Instability and wave over-reflection in stably stratified shear flow. *J. Fluid Mech.,* **151,** 189–217.

——, B. Farrell, and K. K. Tung, 1980: Concept of wave overre-

flection and its application to baroclinic instability. *J. Atmos. Sci.,* **37,** 44–63.

Morgan, M. C., 2001: A potential vorticity and wave activity diagnosis of optimal perturbation evolution. *J. Atmos. Sci.,* **58,** 2518–2544.

——, and C.-C. Chen, 2002: Diagnosis of optimal perturbation evolution in the Eady model. *J. Atmos. Sci.,* **59,** 169–185.

Orr, W. M. F., 1907: Stability or instability of the steady motions of a perfect liquid and of a viscous liquid. *Proc. Roy. Irish Acad.,* **A27,** 9–138.

Rayleigh, L., 1880: On the stability, or instability, of certain fluid motions. *Proc. London Math. Soc.,* **9,** 57–70.

Robinson, W. A., 1989: On the structure of potential vorticity in baroclinic instability. *Tellus,* **41A,** 275–284.

Rossby, C. G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *J. Mar. Res.,* **2,** 38–55.

Waugh, D. W., and D. G. Dritschel, 1991: The stability of filamentary vorticity in two dimentional geophysical vortex-dynamics models. *J. Fluid Mech.,* **231,** 575–598.