Relating optimal growth to counterpropagating Rossby waves in shear instability

Eyal Heifetz[a)]
Department of Geophysics and Planetary Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel

John Methven[b)]
Department of Meteorology, University of Reading, Reading, United Kingdom

(Received 27 August 2004; accepted 28 April 2005; published online 3 June 2005)

The optimal dynamics of conservative disturbances to plane parallel shear flows is interpreted in terms of the propagation and mutual interaction of components called counterpropagating Rossby waves (CRWs). Pairs of CRWs were originally used by Bretherton to provide a mechanistic explanation for unstable normal modes in the barotropic Rayleigh model and baroclinic two-layer model. One CRW has large amplitude in regions of positive mean cross-stream potential vorticity (PV) gradient, while the second CRW has large amplitude in regions of negative PV gradient. Each CRW propagates to the left of the mean PV gradient vector, parallel to the mean flow. If the mean flow is more positive where the PV gradient is positive, the intrinsic phase speeds of the two CRWs will be similar. The CRWs interact because the PV anomalies of one CRW induce cross-stream velocity at the location of the other CRW, thus advecting the mean PV. Although a single Rossby wave is neutral, their interaction can result in phase locking and mutual growth. Here the general initial value problem for disturbances to shear flow is analyzed in terms of CRWs. For the discrete spectrum (which could alternatively be described using normal modes), the singular value decomposition of the dynamical propagator can be obtained analytically in terms of the CRW interaction coefficient and the intrinsic CRW phase speeds. Using this formalism, optimal perturbations, the disturbances which grow fastest in a given norm over a specified time interval, can readily be found. The most natural norm for CRWs is related to air parcel displacements or enstrophy. However, if an energy norm is taken, it is shown to grow due to both mutual amplification of air parcel displacements and the untilting of PV structures (the Orr mechanism) associated with decreasing phase difference between the CRWs. A generalization of the CRW description to the optimal dynamics of the complete spectrum solution is outlined. Although the dynamics then involves the interaction between an infinite number of "CRW kernels," the form of the simple interaction between any two CRW kernels is the same as in the discrete case.
 [DOI: 10.1063/1.1937064]

I. INTRODUCTION

The concept of counterpropagating Rossby waves (CRWs) was developed originally by Bretherton[1] to explain the baroclinic instability mechanism for cyclone growth in the atmosphere. He described the instability in terms of interaction between two edge waves, one located at the earth's surface and the other on the atmospheric tropopause. Each wave by itself is neutral and propagates zonally (i.e., along a latitude circle), via the Rossby[2] mechanism. At the ground the wave propagates eastward and the mean flow is weak while at the tropopause there is a strong eastward jet but the wave propagates westward, counter to the flow. The waves interact in an "action at a distance" fashion, by inducing meridional wind (i.e., along a longitude circle) that advects the basic state potential vorticity (PV) at the other level. Depending on their phase difference, the additional meridional displacements of PV contours modify the zonal propagation rate of the other CRW and can also promote its growth. Growing normal modes are obtained if a suitable phase difference exists where interaction makes the CRW phase speeds equal so that they phase lock, and also results in mutual growth. The mechanism also applies to the barotropic Rayleigh[3] model where CRWs propagate along the edges of an isolated strip of uniform vorticity (cf. Heifetz *et al.*[4]).

Hoskins *et al.*[5] revised and clarified the CRW concept, and Heifetz *et al.*[6] generalized the CRW description to cover conservative quasigeostrophic three-dimensional (3D) disturbances to all plane parallel shear flows which are linearly unstable. They showed that CRW evolution equations can be derived from the Hamiltonian equations where the Hamiltonian, generalized momenta and coordinates are the quasigeostrophic disturbance pseudoenergy, the CRW pseudomomenta and phases. This generalization also rationalized the necessary criterion for instability of Rayleigh[3] (and its baroclinic analog of Charney and Stern[7]) and the Fjørtoft[8] criterion. Methven *et al.*[9] have further extended the CRW theory to general zonally symmetric basic states on a sphere where

[a)]Author to whom correspondence should be addressed. Electronic mail: eyalh@cyclone.tau.ac.il
[b)]Electronic mail: J.Methven@reading.ac.uk

Ertel PV is conserved following air parcels along isentropic surfaces. Methven *et al.*[10] have shown that the CRW properties derived from linear theory are remarkably robust and can be used to predict some features of the nonlinear evolution of baroclinic eddies on realistic midlatitude jets.

As discussed by Heifetz *et al.*,[6] the interaction of a CRW pair describes the discrete spectrum evolution for any problem where the initial conditions are a linear superposition of a growing normal mode and its complex conjugate. The focus of the companion papers (Heifetz *et al.*,[11] Methven *et al.*[9,10]) has been on interpreting growing normal mode structures and dispersion relations in terms of CRW interaction for more realistic representations of the midlatitude atmosphere. However, the extensive work by Farrell,[12–15] Trefethen *et al.*,[16] Schmid and Henningson,[17] Le Dizes,[18] and others showed how the fundamental nature of shear allows rapid nonmodal transient growth due to nonorthogonal interaction between the modes. Farrell and Ioannou[19] (hereafter FI96) developed the generalized stability theory (GST) for linear dynamical systems and showed how the optimal nonmodal growth can be obtained from a singular value decomposition (SVD) of the propagator matrix of the dynamical system.

In this paper the aim is to demonstrate how the CRW perspective might be helpful in understanding the physical mechanism behind optimal growth in shear flows. In Sec. II, a short review of the CRW perspective is presented using the Rayleigh[3] model for illustration. This example was chosen because the dynamical propagator is a simple 2×2 matrix whose SVD can be derived analytically. Then, in Sec. III, the GST is applied to this model and the optimal growth is interpreted in terms of CRW interaction. Based on this analysis the generalization of the CRW perspective to optimal growth on more general shear flows is outlined in the Appendix. The conclusions are summarized in Sec. IV.

II. CRW INTERACTION ILLUSTRATED USING THE RAYLEIGH MODEL

Heifetz *et al.*[4] discussed CRW interaction in the Rayleigh[3] model, focusing mainly on their phase-locking behavior into the growing normal mode configuration. The analysis is summarized here before examining optimal perturbations.

Consider a 2D, inviscid, incompressible flow whose basic state velocity and vorticity profiles are

$$\overline{u}=\begin{cases}\Lambda b & \text{for } y\geqslant b\\ \Lambda y & \text{for } -b<y<b\\ -\Lambda b & \text{for } y\leqslant -b,\end{cases} \tag{1a}$$

$$\overline{q}=\begin{cases}0 & \text{for } y\geqslant b\\ -\Lambda & \text{for } -b<y<b\\ 0 & \text{for } y\leqslant -b.\end{cases} \tag{1b}$$

Here (x,y) are the streamwise and cross-stream Cartesian coordinates associated with the velocity $\mathbf{v}=(u,v)=(\overline{u}+u',v')$, where overbars and primes indicate the basic state and perturbation, respectively. Linearizing the vorticity equation $Dq/Dt=0$ with respect to the basic state of (1) ($D/Dt=\partial/\partial t+\mathbf{v}\cdot\boldsymbol{\nabla}$, $q=\hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\times\mathbf{v}$, and $\hat{\mathbf{z}}$ is the vertical unit vector) gives

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)q'=-v'\frac{\partial\overline{q}}{\partial y}=v'\Lambda[\delta(y+b)-\delta(y-b)]. \tag{2}$$

The discrete spectrum solution of (2) for wavenumber k can be written as

$$q'=[q_1(k,t)\delta(y+b)+q_2(k,t)\delta(y-b)]e^{ikx}, \tag{3a}$$

$$\psi'=-\frac{1}{2k}[q_1(k,t)e^{-k|y+b|}+q_2(k,t)e^{-k|y-b|}]e^{ikx}, \tag{3b}$$

(the continuous spectrum of the solution is discussed in the Appendix). The perturbation streamfunction ψ' satisfies $u'=-\partial\psi'/\partial y$, $v'=\partial\psi'/\partial x$, and $q'=\nabla^2\psi'$. Writing then

$$q_1=Q_1(k,t)e^{i\epsilon_1(k,t)}, \tag{4a}$$

$$q_2=Q_2(k,t)e^{i\epsilon_2(k,t)}, \tag{4b}$$

the perturbation can be regarded as two vorticity edge waves with amplitudes Q_1, Q_2 and phases ϵ_1, ϵ_2. By substituting (3) and (4) into (2) and taking the real and imaginary parts at $y=\pm b$, the CRW evolution equations are obtained,

$$\dot{Q}_1=\sigma Q_2\sin\epsilon, \tag{5a}$$

$$\dot{Q}_2=\sigma Q_1\sin\epsilon, \tag{5b}$$

$$\dot{\epsilon}_1=-kc_1^1-\sigma\frac{Q_2}{Q_1}\cos\epsilon, \tag{6a}$$

$$\dot{\epsilon}_2=-kc_2^2+\sigma\frac{Q_1}{Q_2}\cos\epsilon, \tag{6b}$$

with

$$\sigma=\frac{\Lambda}{2}e^{-K}, \tag{7a}$$

$$c_1^1=\overline{u}(-b)\left(1-\frac{1}{K}\right), \tag{7b}$$

$$c_2^2=\overline{u}(b)\left(1-\frac{1}{K}\right), \tag{7c}$$

where $K=2bk$ is the normalized wavenumber and $\epsilon=(\epsilon_2-\epsilon_1)$ is the CRW phase difference.

Equation (5) indicates that each wave can only grow through interaction with the other and the strength depends upon the interaction coefficient σ and the CRW phase difference ϵ. Equation (6) indicates that in the absence of interaction the waves would propagate with their intrinsic phase speeds c_1^1, c_2^2 given by (7b) and (7c). The intrinsic phase speed comprises two terms: a Doppler shift due to the basic state flow at the edge and a propagation counter to this flow that is proportional to the wavelength. The latter term results from Rossby[2] wave propagation to the left of the basic state vorticity gradient vector.

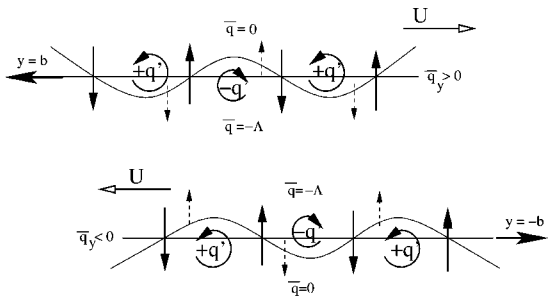

FIG. 1. Schematic illustration of the counterpropagating Rossby waves (CRWs) in the Rayleigh model in the configuration of the most unstable normal mode (wavenumber $K\approx 0.8$). Blank arrows indicate the basic state velocity $\bar{u}$ on the edges ($y=\pm b$). Bold horizontal arrows represent the CRWs' propagation direction. The edge positive and negative vorticity anomalies are indicated by $\pm q'$ and the circulation they induce is illustrated by the circled arrows and by the cross-stream solid arrows located $\mp\pi/2$ out of phase of $\pm q'$. The undulating solid lines illustrate the CRWs' cross-stream displacement, where the basic state vorticity $\bar{q}$ at the vicinity of the edges is given by (1b). The two CRWs are phase-locked with phase difference $\epsilon_+ \approx 0.65\pi$. The dashed arrows indicate the velocity induced by each CRW on the opposite edge [attenuated by e^{-K}, (3b)]. Since the phase difference $\pi/2<\epsilon_+<\pi$ each CRW advects the basic state vorticity on the opposite edge in a way that makes other CRW grow and hinders its natural propagation.

If the relative phases of the two waves are known then the sense of the interaction between the CRWs can be ascertained following the arguments in Hoskins *et al.*[5] In Fig. 1 the CRWs are shown with the phase difference ($\epsilon\approx 0.65\pi$) corresponding to the phase-locked configuration of the fastest growing normal mode. "CRW-1" exists on the negative vorticity gradient at the southern edge and "CRW-2" exists on a positive vorticity gradient at the northern edge. The strongest northward flow and hence negative vorticity tendency induced by the CRW-1 at the northern edge (dashed arrow) lies between 0 and $\pi/2$ to the east of the negative vorticity extremum of CRW-2. It therefore amplifies CRW-2 but reduces its propagation rate westwards. Similarly, the northward flow and hence positive vorticity tendency induced by CRW-2 at the southern edge is between 0 and $\pi/2$ to the west of the positive vorticity of CRW-1. This leads to growth but slows the eastward propagation of CRW-1. The interaction of the two CRWs leads to the growth of each but slows the propagation of each against the mean flow at their respective latitudes (referred to as "home bases" by Heifetz *et al.*[6]). As each wave slows the other in its counterpropagation, this is referred to as a "hindering" configuration.

For a phase difference between 0 and $\pi/2$, the negative vorticity tendency induced by CRW-1 at the northern edge is between $\pi/2$ and 0 to the west of the negative PV of CRW-2. There is again growth but this time the westward propagation of CRW-2 is enhanced by the interaction. Similarly, it can be argued that the CRW-1 grows and propagates more rapidly to the east as a consequence of its interaction with CRW-2. Thus, the interaction of the two CRWs again leads to the growth of each and "helps" their self-propagation (counteracting differential advection by the shear in flow).

Since the interaction acts to increase the westward phase shift when it is between 0 and $\pi/2$ and decrease it when it is between $\pi/2$ and π, if it is strong enough it can be expected to lead to the phase remaining in the range of $0\rightarrow\pi$ and continued growth. The equilibrium position for the phase would correspond to the growing normal mode. For eastward phase shifts, similar arguments apply and there is decay. However, in this case the hindering/helping mechanism tends to shift the phase from the decaying range of $-\pi\rightarrow 0$ toward the growing range of $0\rightarrow\pi$. This indicates the nonmodal evolution mechanism in the initial value problem which leads towards convergence to the phase-locking configuration of the growing normal mode.

Note that although the instantaneous growth is the fastest when the CRW phase difference is $\pi/2$, cf. (5), in this configuration the CRWs cannot affect each other's propagation speed, cf. (6), and therefore can remain phase locked only if their intrinsic phase speeds are equal, i.e., when $K=1$ [(7b) and (7c)]. The gravest normal mode is obtained when $\epsilon \approx 0.65\pi$ ($K\approx 0.8$), in a configuration that hinders the self-propagation of both CRWs.

The behavior of CRWs and normal modes in the Rayleigh problem is depicted in Fig. 2 as a function of wavenumber. The interaction coefficient (solid curve) decreases exponentially with wavenumber (7a) reflecting the scale dependence of the strength of cross-stream velocity induced by a CRW on the opposite edge of the shear region. The Rossby wave propagation rate is proportional to the wavelength and therefore as wavenumber increases the intrinsic phase speed of each CRW becomes dominated by advection by the mean flow at its home base (c_2^2 is shown by dashed curve). For high wavenumbers ($K>1.28$) the intrinsic counterpropagation of each CRW is weak relative to the shear and phase-locked states are only possible when there is no phase difference ($\cos\epsilon=1$) so that the interaction (6) lends maximum help to the propagation of each CRW counter to the flow at its home base. Also, the CRWs must have different amplitudes so that the interaction terms in (6) can counteract the difference in intrinsic phase speeds. The resulting pair of normal modes cannot grow because the CRWs are in phase (5). One is dominated by CRW-2 and advection on the northern edge, and therefore has a positive phase speed (crosses in Fig. 2), while the other is dominated by CRW-1 and westward advection on the southern edge (not shown).

At wavelengths just longer than the short-wave cutoff ($K\approx 1.28$), the counterpropagation rate of the CRWs is sufficiently strong relative to the shear that their intrinsic phase speeds are similar. Additionally, the interaction is strong enough to modify their phase speeds so that they can be made equal when the CRWs are not in phase ($\epsilon>0$). The CRW growth rates are only equal, when $\sin\epsilon\neq 0$, if their amplitudes are equal [$Q_1=Q_2$ in (5)]. The phase difference ϵ_+ in the locked configuration of the growing normal mode occurs when $\dot{\epsilon}_1=\dot{\epsilon}_2$ in (6), thus giving

$$\cos\epsilon_+ = \frac{k(c_2^2-c_1^1)}{2\sigma} \tag{8}$$

(indicated by circles). In the range $1<K<1.28$, counterpropagation is weaker than advection by the mean flow [$1-1/K>0$ in (7b) and (7c)] so that the CRWs lock with a small phase difference which helps counterpropagation (0

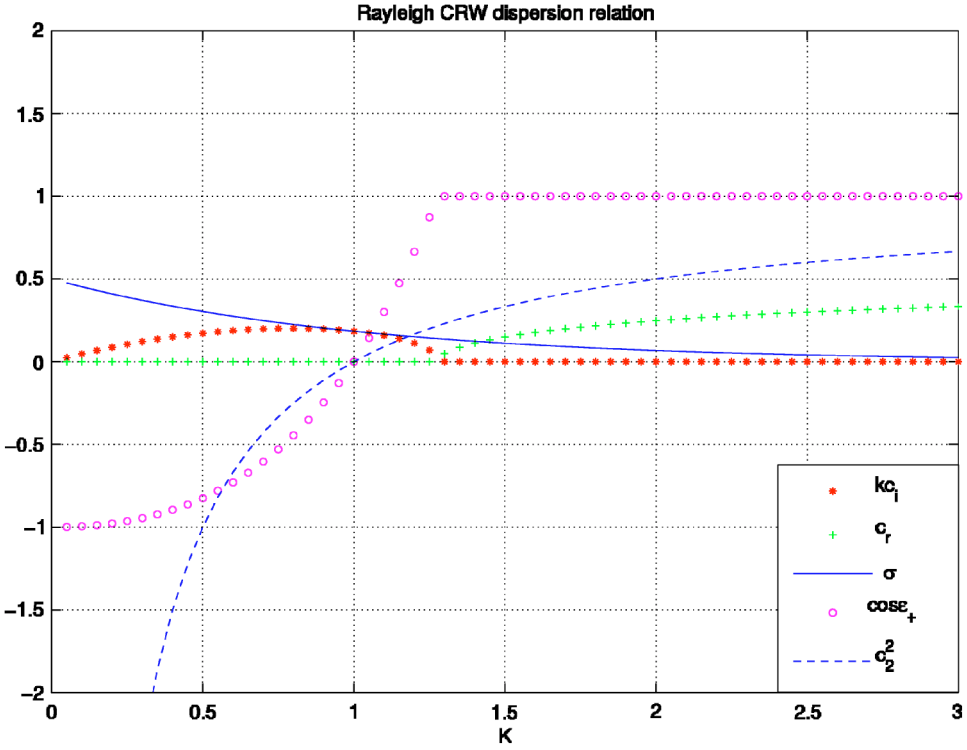

FIG. 2. The CRW parameters which determine the Rayleigh normal mode dispersion relation as functions of the normalized horizontal wavenumber K. The modal growth rate $kc_i=\sigma\sin\epsilon_+$, (11), normalized by the shear Λ, is indicated by asterisks. The positive branch of the modal phase speed, (14), normalized by $2b\Lambda$ is shown by the plus symbols. The cosine of the CRW phase difference in the growing mode configuration, $\cos\epsilon_+$, (8), is indicated by circles. The interaction coefficient σ, (7a), normalized by Λ, and the intrinsic upper CRW phase speed c_2^2, (7c), normalized by $2b\Lambda$, are indicated, respectively, by the solid and the dashed lines.

$<\cos\epsilon_+<1$) but enables mutual growth (the normal mode growth rate is shown by asterisks in Fig. 2). At longer wavelengths ($K<1$) the counterpropagation of each CRW is so strong that its intrinsic phase speed has a greater magnitude than the basic state flow at its home base (e.g., c_2^2 becomes negative). Phase locking can only occur when interaction hinders the propagation of each CRW, although enabling mutual growth ($-1<\cos\epsilon_+<0$). The maximum modal growth rate occurs at $K\approx 0.8$ where there is a trade-off between the exponential increase in interaction strength with wavelength and the requirement for an increasingly hindering configuration to enable phase locking (reducing the normal mode growth rate $\sigma\sin\epsilon_+$).

The illustration here is based on the Rayleigh model. However, the CRW evolution equations (5) and (6) can describe the evolution of a disturbance to any steady plane parallel flow $\overline{u}(y,z)$ that could alternatively be described by the superposition of a growing normal mode and its decaying complex conjugate (Heifetz *et al.*[6]). However, the expressions for interaction coefficient and CRW phase speeds depend on the basic state of the model and PV inversion relation, and (7) is specific to the Rayleigh model. It is clear from the CRW evolution equations (5) and (6) that the CRWs can gain the largest transient amplitude growth when the phase difference stays the longest in the vicinity of $\epsilon=\pi/2$. Next, we examine optimal transient growth from the CRW perspective.

III. GENERALIZED STABILITY THEORY IN TERMS OF CRWS

A. Formulation

The CRW evolution equations (5) and (6) are general and describe the discrete spectrum dynamics of any conservative, linearly unstable parallel shear flow. They can be rewritten with the amplitude and phase definition (4), in the matrix form as follows:

$$\dot{\mathbf{q}}=\mathbf{A}\mathbf{q}, \tag{9a}$$

where

$$\mathbf{q}=\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}, \tag{9b}$$

$$\mathbf{A}=-i\begin{pmatrix} kc_1^1 & \sigma \\ -\sigma & kc_2^2 \end{pmatrix}. \tag{9c}$$

The general solution to (9) can be written in the equivalent forms

$$\mathbf{q}(t)=e^{\mathbf{A}t}\mathbf{q}(0)=(\mathbf{P}e^{\mathbf{L}t}\mathbf{P}^{-1})\mathbf{q}(0)=(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\dagger})\mathbf{q}(0), \tag{10}$$

where $e^{\mathbf{A}t}$ is the propagator matrix of the linear dynamics. $\mathbf{L}$ is a diagonal matrix containing the complex eigenvalues of $\mathbf{A}$ [ordered by their real values, $\mathrm{Re}(\lambda_1)\geqslant\mathrm{Re}(\lambda_2)$] and $\mathbf{P}$ is the corresponding matrix of eigenvectors. Alternatively, $(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\dagger})$ is the SVD of the propagator matrix and both $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices (i.e., $\mathbf{U}\mathbf{U}^{\dagger}=\mathbf{I}$, where $\mathbf{U}^{\dagger}$ is the Hermitian conjugate of $\mathbf{U}$). $\boldsymbol{\Sigma}$ is a diagonal matrix which contains real positive entries, called singular values, ordered by magnitude along the diagonal ($\sigma_1\geqslant\sigma_2\geqslant 0$).

Only if the matrix $\mathbf{A}$ is Hermitian ($\mathbf{A}^{\dagger}=\mathbf{A}$) do the eigen decomposition and the SVD become identical. If the matrix $\mathbf{A}$ is normal but not Hermitian (i.e., commuting with its Hermitian conjugate, $\mathbf{A}\mathbf{A}^{\dagger}=\mathbf{A}^{\dagger}\mathbf{A}$) then its eigenvectors are orthogonal, $\sigma_j=e^{\mathrm{Re}(\lambda_j)t}$ and $\mathbf{U}=\mathbf{V}$.

When $\mathbf{A}$ is not normal, $\sigma_1>e^{\mathrm{Re}(\lambda_1)t}$. As shown by FI96, σ_1 is the largest possible growth that autonomous linear systems of the form of (9a) can achieve in a time interval t.

Since **U** and **V** are unitary, the initial optimal vector $\mathbf{q}(0)$ is the first column unit vector of **V**. The optimal vector at time t is found by amplifying $\mathbf{q}(0)$ by the factor σ_1 and then projecting it onto the first column vector of **U**. FI96 examined the optimal growth for the two target time limits of zero and infinity. The maximum instantaneous growth rate is equal to the maximum eigenvalue of $(\mathbf{A}+\mathbf{A}^\dagger)/2$ and its associated eigenvector yields the optimal instantaneous structure. For the target time infinity the optimal perturbation would eventually be projected onto the gravest mode $\mathbf{p}_1$. However its initial perturbation would be in the direction of the biorthogonal vector of the most unstable mode $\mathbf{r}_1$, defined by the first column vector of the matrix $\mathbf{R}=(\mathbf{P}^{-1})^\dagger$. At longer times $\mathbf{q}(t)\to|\mathbf{r}_1|e^{\mathrm{Re}(\lambda_1)t}\mathbf{p}_1$, where $|\mathbf{r}_1|>1$ if **A** is non-normal.

B. Normal modes in terms of CRWs

The eigenvalues of **A**, (9c), are

$$\lambda_{1,2}=-ik\overline{c}\pm\sigma\sqrt{1-f^2}, \tag{11}$$

where

$$\overline{c}=\tfrac{1}{2}(c_1^1+c_2^2), \tag{12a}$$

$$f=\frac{k(c_2^2-c_1^1)}{2\sigma}. \tag{12b}$$

If $f^2<1$ then pairs of growing and decaying normal modes exist. Inspection of (8) reveals that in this regime $f=\cos\epsilon_+$ and therefore $\lambda_{1,2}=-ik\overline{c}\pm\sigma\sin\epsilon_+$, where ϵ_+ is the CRW phase difference when they are locked into the growing normal mode configuration. In this case, $\overline{c}$ can be identified immediately as the normal mode phase speed and $\sigma\sin\epsilon_+$ is its growth rate. The eigenvectors of **A** are then

$$\mathbf{P}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ e^{i\epsilon_+} & e^{-i\epsilon_+}\end{pmatrix}. \tag{13}$$

If $f^2>1$, the eigenvalues are purely imaginary indicating that the normal modes are neutral. Their phase speeds are

$$c_{1,2}=\overline{c}\mp\frac{\sigma}{k}\sqrt{f^2-1}, \tag{14}$$

with the eigenvectors

$$\mathbf{P}=\frac{1}{\sqrt{\alpha^2+1}}\begin{pmatrix}\alpha & 1\\ 1 & \alpha\end{pmatrix}, \tag{15}$$

where $\alpha=f+\sqrt{f^2-1}$.

The matrix **A**, (9c), is normal only either when $c_1^1=c_2^2$ (occurring at $K=1$ where $c_1^1=0$), or in the limit case of $\sigma=0$ $(K\to\infty)$. When the CRW phase speeds are equal, $f=0$ (12b) and therefore (11) and (13) give

$$\mathbf{L}_{\mathrm{normal}}=\begin{pmatrix}\sigma & 0\\ 0 & -\sigma\end{pmatrix}, \tag{16a}$$

$$\mathbf{P}_{\mathrm{normal}}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix}. \tag{16b}$$

The CRW interpretation is straightforward. If the two CRWs have the same intrinsic phase speeds, phase locking is achieved when interaction cannot modify their propagation rates. This occurs when they are in quadrature, $\epsilon_+=\pi/2$. The CRWs cannot gain any additional growth by moving relative to each other and thus the maximal growth rate σ is achieved by the eigenvectors themselves. In the latter case of zero CRW interaction $(\sigma=0)$, no growth is possible and the two CRWs become two decoupled neutral edge waves which propagate with their own intrinsic phase speeds, cf. (14) and (15).

C. Optimal growth in the enstrophy norm

Aside from the special cases described above, the matrix **A** is non-normal and greater growth can be achieved by singular vectors than by normal modes. The SVD of $e^{\mathbf{A}t}$ can be calculated analytically (using, for instance, the symbolic math package of the software MATHEMATICA). **A** is given by (9c) and **U** and **V** are the eigenvectors of $e^{\mathbf{A}t}e^{\mathbf{A}^\dagger t}$ and $e^{\mathbf{A}^\dagger t}e^{\mathbf{A}t}$, respectively,

$$\mathbf{U}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -ie^{i\epsilon_0}\\ -e^{-i\epsilon_0} & -i\end{pmatrix}, \tag{17a}$$

$$\mathbf{V}=\frac{1}{\sqrt{2}}\begin{pmatrix}ie^{-i\epsilon_0} & 1\\ i & -e^{i\epsilon_0}\end{pmatrix}. \tag{17b}$$

The final optimal vector can be expressed in terms of the phase difference between the CRWs at the initial time ϵ_0 or final time ϵ_t by noting that the top entry corresponds to CRW-1 and the second to CRW-2,

$$\mathbf{u}_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -e^{-i\epsilon_0}\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ e^{i\epsilon_t}\end{pmatrix}, \tag{18}$$

which implies the relationship $\epsilon_t=\pi-\epsilon_0$. Thus, the optimal evolution is symmetric about the phase difference $\pi/2$. The change in phase difference over the target time interval t is

$$\Delta=\epsilon_t-\epsilon_0=\pi-2\epsilon_0. \tag{19}$$

Using this expression, the initial optimal vector can be written as

$$\mathbf{v}_1=\frac{1}{\sqrt{2}}\begin{pmatrix}ie^{-i\epsilon_0}\\ i\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}e^{i(0+\Delta/2)}\\ e^{i(\epsilon_t-\Delta/2)}\end{pmatrix}, \tag{20}$$

showing that the CRWs shift equal and opposite distances during optimal evolution.

These results can be understood by considering the evolution of enstrophy which can be expressed using (5) as

$$\frac{1}{2}\frac{\partial}{\partial t}(Q_1^2+Q_2^2)=2Q_1Q_2\sigma\sin\epsilon\leqslant(Q_1^2+Q_2^2)\sigma\sin\epsilon \tag{21}$$

[where $(Q_1-Q_2)^2\geqslant 0$ was used]. The maximal instantaneous growth rate is given by the interaction coefficient σ and is obtained when the growth is synchronous $(Q_1=Q_2=Q)$ and $\epsilon=\pi/2$. σ is indeed the largest eigenvalue of $(\mathbf{A}+\mathbf{A}^\dagger)/2$, and its associated eigenvector is indeed the first column vector of (16b) (as predicted by the general theory of FI96). The CRWs have equal amplitudes in the initial singular vector

(20) and (5) shows that the growth must continue synchronously. As indicated by (7a) and Fig. 2, when $K \to 0$, $\sigma \to \Lambda/2$, which is the upper theoretical limit for growth rate in the Rayleigh model (Drazin and Reid[20]). For finite target times the integration of (21) for synchronous growth yields

$$Q(t) = Q(0)\exp\left[\sigma \int_{t=0}^{t} \sin\epsilon(t)dt\right]. \tag{22}$$

Optimal perturbations will evolve so that the phase difference is symmetric in time about $\pi/2$, maximizing $\sin\epsilon$, as shown by the SVD analysis (19). For synchronous growth, subtracting (6a) and (6b) and substituting (12b) gives

$$\dot{\epsilon} = 2\sigma(\cos\epsilon - f). \tag{23}$$

Substituting (23) into the integrand of (22) yields

$$\frac{Q(t)}{Q(0)} = \exp\left[-\frac{1}{2}\int_{\cos\epsilon_0}^{-\cos\epsilon_0} \frac{d(\cos\epsilon)}{\cos\epsilon - f}\right], \tag{24}$$

where (19) is used to determine the integral boundaries for the optimal evolution. Integration gives the optimal growth factor

$$\frac{Q(t)}{Q(0)} = \sqrt{\frac{f - \cos\epsilon_0}{f + \cos\epsilon_0}}. \tag{25}$$

1. Optimal growth in the unstable modal regime ($f^2 < 1$)

The matrix of singular values, $\mathbf{\Sigma}$ in (10), can be defined by the eigenvalue matrix of $e^{\mathbf{A}t}e^{\mathbf{A}^\dagger t}$ (or $e^{\mathbf{A}^\dagger t}e^{\mathbf{A}t}$) which is equal to $\mathbf{\Sigma}^2$. In the regime where $f^2 < 1$,

$$\mathbf{\Sigma} = \mathbf{\Theta}e^{\mathbf{L}t}, \quad \mathbf{\Theta} = \begin{pmatrix} \theta & 0 \\ 0 & \theta^{-1} \end{pmatrix}, \tag{26}$$

where $\mathbf{L}$ is the diagonal eigenvalue matrix of $\mathbf{A}$ and θ is the factor by which the optimal growth exceeds the modal growth rate.

In this regime (8) shows that $f = \cos\epsilon_+$. If propagation is hindered by interaction in the phase-locked configuration, $\pi/2 < \epsilon_+ < \pi$, (23) shows that the CRW phase difference increases with time from any initial phase in the range $-\epsilon_+ < \epsilon_0 < \epsilon_+$. Since the optimal evolution must cross $\pi/2$, cf. (20), this yields $\epsilon_0 < \pi/2 < \epsilon_t$. Therefore, during the optimal evolution the perturbation structure becomes more tilted against the shear. By these same considerations, in the helping regime, $0 < \epsilon_+ < \pi/2$, the CRW phase difference decreases with time for any initial phase in the range $\epsilon_+ < \epsilon_0 < 2\pi - \epsilon_+$. Therefore $\epsilon_t < \pi/2 < \epsilon_0$, meaning that the perturbation structure untilts while optimally growing. The CRW optimal evolution is summarized in Fig. 3.

Optimal growth can also be related to modal growth (when $f^2 < 1$) by writing (22) as

$$\frac{Q(t)}{Q(0)} = \exp\left[\sigma \int_{t=0}^{t} (\sin\epsilon(t) - \sin\epsilon_+)dt\right] e^{\sigma \sin\epsilon_+ t} = \theta e^{\sigma \sin\epsilon_+ t}, \tag{27}$$

where θ is the factor by which the optimal growth exceeds the modal growth. Equation (23) can be used to write

$$\theta = \exp\left[-\frac{1}{2}\int_{\cos\epsilon_0}^{-\cos\epsilon_0} \cot\left(\frac{\epsilon + \epsilon_+}{2}\right)d\epsilon\right] = \frac{\sin\left(\frac{\epsilon_0 + \epsilon_+}{2}\right)}{\cos\left(\frac{\epsilon_0 - \epsilon_+}{2}\right)}. \tag{28}$$

This factor can also be expressed in terms of the CRW difference at the final time using (19), $\theta = \cos[(\epsilon_t - \epsilon_+)/2]/\sin[(\epsilon_t + \epsilon_+)/2]$. Figure 4 shows θ in the helping regime ($0 < \epsilon_+ < \pi/2$) as a function of $(\epsilon_t - \epsilon_+)/\pi$ and $(\epsilon_t + \epsilon_+)/2\pi$. The θ factor is the greatest, for given $(\epsilon_t + \epsilon_+)$, where $\epsilon_t = \epsilon_+$ (the phase difference eventually approaches the phase-locked configuration). It also increases as $(\epsilon_t + \epsilon_+) \to 0$ because the optimal perturbation spends more time with a phase difference such that $\sin\epsilon > \sin\epsilon_+$ (recall that in the helping regime the phase difference for the optimal perturbation decreases with time through $\pi/2$).

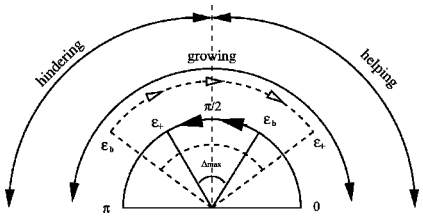

FIG. 3. The relative CRW phase angle ϵ in the growing regime of $0 < \epsilon < \pi$. ϵ_+ corresponds to the CRW phase difference in the growing normal mode configuration, while $\epsilon_b = (\pi - \epsilon_+)$ corresponds to its biorthogonal configuration. The maximal instantaneous growth occurs at $\epsilon = \pi/2$. For finite target time, (19) indicates that the initial and final optimal phases (ϵ_0, ϵ_t) are symmetric with respect to $\pi/2$. Hence for all target times, when ϵ_+ is in the hindering regime, ϵ increases during the optimal evolution (filled arrows), and $\Delta = (\epsilon_t - \epsilon_0)$ is positive. For target time infinity $(\epsilon_0, \epsilon_t) = (\epsilon_b, \epsilon_+)$.

2. Optimal growth in the neutral modal regime ($f^2 > 1$)

Equation (15) shows that the two CRWs forming the neutral normal modes for $f^2 > 1$ are in phase ($\epsilon = 0$) but with different amplitudes (such that $Q_2/Q_1 = \alpha$ or $1/\alpha$). However, the singular vectors of $e^{\mathbf{A}t}$, given by (17a) and (17b) for any f, are composed of CRWs with equal amplitudes. For $f > 1$ the singular value matrix becomes

$$\mathbf{\Sigma} = \begin{pmatrix} g & 0 \\ 0 & g^{-1} \end{pmatrix}, \tag{29a}$$

$$g = \sqrt{\frac{f - \cos\epsilon_0}{f + \cos\epsilon_0}}, \tag{29b}$$

and therefore the maximum possible growth G, called the global optimal by FI96, is obtained when $\epsilon_0 = \pi$ such that

$$G = \sqrt{\frac{f+1}{f-1}}. \tag{30a}$$

Integrating $1/\dot{\epsilon}$ with respect to ϵ from $\epsilon_0 = \pi$ to $\epsilon_t = 0$ [using (23)] gives the global optimal target time

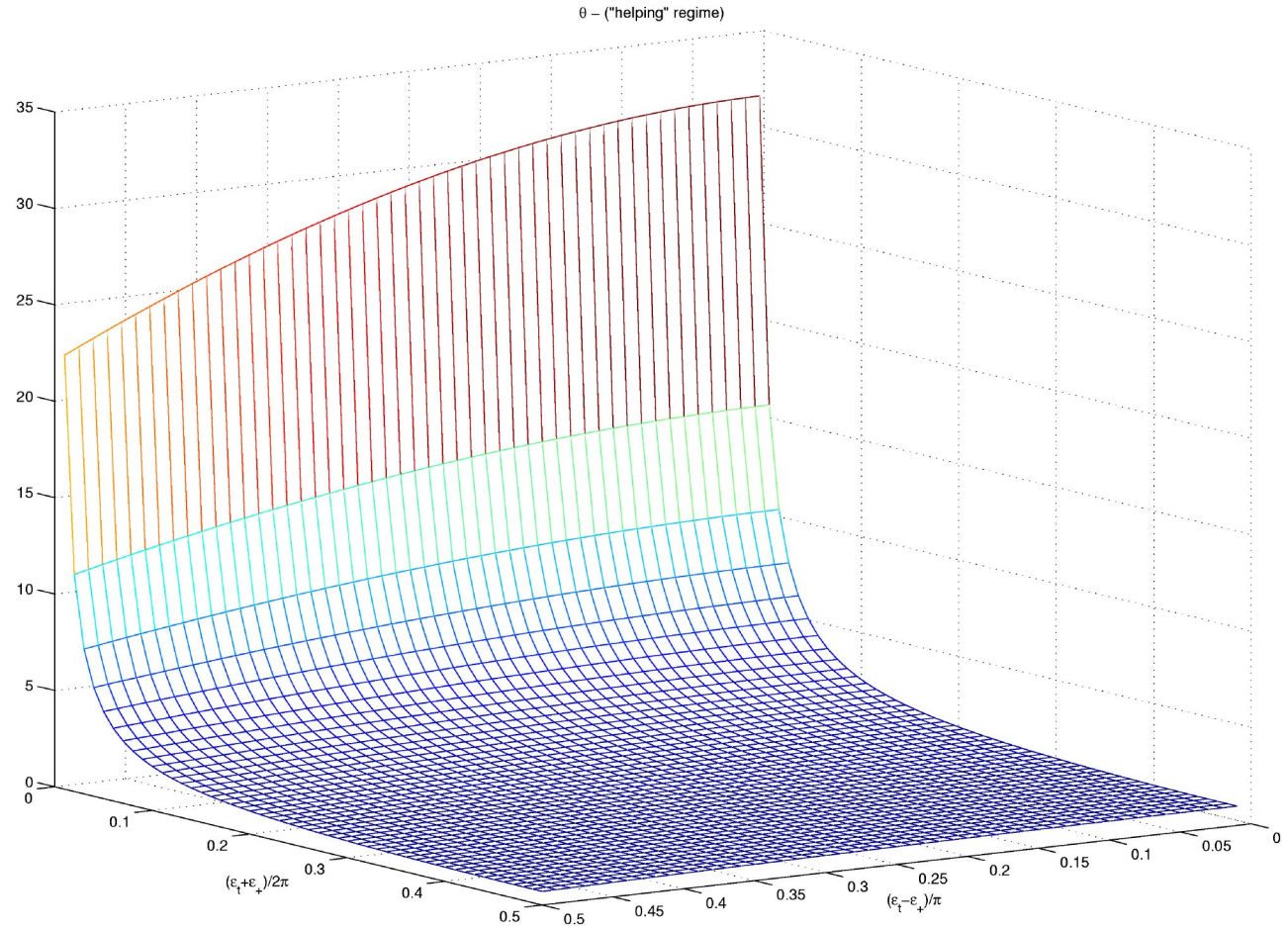

FIG. 4. The ratio between the optimal growth and the normal mode growth θ as a function of $(\epsilon_t-\epsilon_+)/\pi$ and $(\epsilon_t+\epsilon_+)/2\pi$, in the regime $0<\epsilon_+<\pi/2$ where each CRW must help the counterpropagation of the other to enable phase-locking. (ϵ_+,ϵ_t) are the CRW phase difference of the normal mode and the final phase difference of the optimal perturbation.

$$T=\left(n+\frac{1}{2}\right)\frac{\pi}{\sqrt{f^2-1}},\quad n=0,1,2\ldots. \tag{30b}$$

In Fig. 5 the example $f=1.1$ is chosen to illustrate the dependence of optimal growth and phase on normalized target time σt. The optimal growth is indicated in the upper panel by the solid line and the value of the global optimal, $G=4.58$, is indicated by the dashed line. The first global optimal is achieved after $\sigma t=\pi/2(f^2-1)^{-1/2}=3.43$. The optimal evolution is synchronous and the lower panel indicates that $(\epsilon_0+\epsilon_t)/2=\pm\pi/2$. The global optimal is achieved when $(\epsilon_0,\epsilon_t)=(\pi,0)$.

In the neutral regime $(f>1)$, if the CRWs are assumed to have equal amplitudes as is the case for the singular vectors (17), (23) shows that $\dot{\epsilon}<0$ implying that the CRWs are continually advected past each other by the shear. As discussed earlier, this occurs because the counterpropagation rates and interaction strength are too weak to attain a phase-locked state at high wavenumbers. Therefore, in order to obtain optimal growth the CRWs should cross $\epsilon=\pi/2$ so that $\epsilon_0>\epsilon_t$. For short target times, (ϵ_0,ϵ_t) are both close to $\pi/2$, but as the target time increases, the phase change increases until the global optimal configuration $(\epsilon_0,\epsilon_t)=(\pi,0)$ [Fig. 5(b)]. As the target time becomes slightly larger than the first global optimal, the CRWs must begin with a decaying configuration (tilted with the shear), pass through rapid growth at $\pi/2$ and then end up in a decaying configuration. Therefore, the optimal growth is smaller than the global optimal. Eventually, for target time $\pi(f^2-1)^{-1/2}$, the CRWs start at $\epsilon_0=-\pi/2$, the amount of decay exactly cancels the amount of growth and no net growth is obtained at the target time.

3. Wavelength dependence of optimal growth

The effective optimal growth rate $\gamma(t)$ can be defined as

$$\gamma(t)=\frac{1}{t}\ln\sigma_1(t), \tag{31}$$

where $\sigma_1(t)$ is the largest singular vector of the propagator matrix $e^{\mathbf{A}t}$, (9c). In Fig. 6(a) $\gamma(t)$ is shown as function of K for the target times $t=1/\Lambda,3/\Lambda,5/\Lambda$. These growth rates are compared with the maximal instantaneous growth rate σ and the normal mode growth rate $[\mathrm{Re}(\lambda_1)=kc_i,\ (11)]$. As expected $kc_i<\gamma<\sigma$ for all target times between zero and infinity, except at wavenumber $K=1$ where the system is normal and all growth rates are identical. The shorter the target time, the greater is the fraction of time spent within the vicinity of phase difference $\epsilon=\pi/2$ and therefore the effective

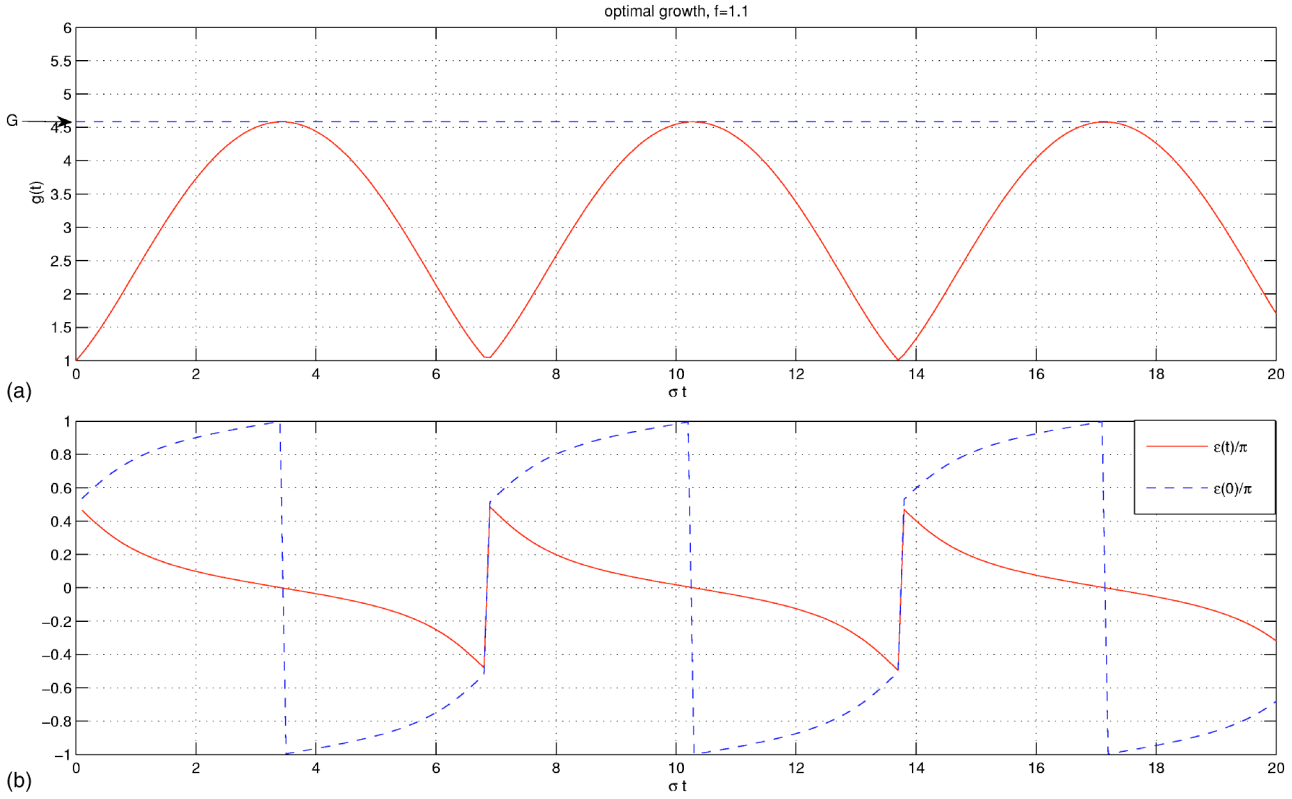

FIG. 5. Optimal evolution in the modal stable regime. In this example $f=1.1$, (12b), and the consecutive normalized target times $\sigma t=(0,20)$. (a) The optimal growth g (29b) is indicated by the solid line. The global optimal $G=4.58$ (30a) is indicated by the dashed line and is achieved first at the normalized time $T=\sigma t=3.43$. (b) The initial and final CRW optimal phase differences ϵ_0 and ϵ_t vs target time are indicated, respectively, by the dashed and the solid lines. Note that the phase difference always decreases with time during optimal evolution.

growth rate is larger. Since interaction strength σ decreases exponentially with K, (7a), γ decreases with wavenumber for all target times. The ratio between the optimal growth and the normal mode growth is equal either to θ, (28), when $f^2<1$ or to g, (29b), when $f^2>1$ (the modal short-wave cutoff where $f=1$ occurs at $K\approx 1.28$). In the unstable regime the maximal value of θ is equal to the magnitude of the biorthogonal vector of the most unstable mode $\sigma_{\max}=|\mathbf{r}_1|=1/\sin\epsilon_+$. In the stable regime the greatest possible amplification is given by the global optimal G, (30a).

The optimal change in the CRWs' relative phase Δ, (19), is plotted in Fig. 6(b) for the three target times. Δ is positive for $K<1$, where the CRWs phase lock in a hindering configuration, $\pi/2<\epsilon_+<\pi$, so that optimal growth is achieved when the CRWs start with a small phase difference and it increases, passing through the most rapid growth at $\pi/2$. Δ is negative for $1<K<1.28$ where the CRWs phase lock in a helping configuration, $0<\epsilon_+<\pi/2$. Maximal growth occurs for target time infinity in the unstable regime and is achieved by a change in CRW phase difference of $\Delta_{\max}=2\epsilon_+-\pi$. The longer the target time interval, the closer Δ approaches $\Delta_{\max}$. In the stable regime, optimal growth is achieved by CRWs with equal amplitudes but these cannot phase lock and their phase difference always decreases as they are advected past each other. Maximal growth can be achieved by a change in phase difference of $\Delta_{\max}=-\pi$ and occurs at the first global optimal time (30b) for that wavenumber. If the target time exceeds the first global optimal time for the given wavenumber then the CRW phase difference decreases by more than π and cycles of decay and growth are experienced, since the waves are periodic.

D. Optimal growth in the energy norm

While the normal mode growth rate is independent of the measured norm, the nonmodal optimal growth generally changes from one norm to another. The most suitable norms are "wave activities" that are globally conserved for disturbances to a specified basic state and are also quadratic in disturbance quantities. Pseudomomentum is conserved for waves on plane parallel shear flows and for Rossby waves is proportional to the cross-stream air-parcel displacement squared weighted by the mean cross-stream PV gradient. If the PV gradients at the two home bases of a CRW pair have the same magnitude (but opposite signs), as in the Rayleigh model, then the pseudomomentum of each CRW is proportional to its enstrophy. However, the total pseudomomentum of both waves tends to zero as the CRWs approach the phase-locked state of the growing normal mode (Held[21]). Therefore, it is natural to use total enstrophy as a positive definite measure of disturbance amplitude as in Sec. III C.

However, frequently eddy energy is used as a norm although it is not globally conserved for disturbance quantities alone (conservation is ensured for pseudoenergy). Here, the optimal evolution in the energy norm is examined from the CRW perspective.

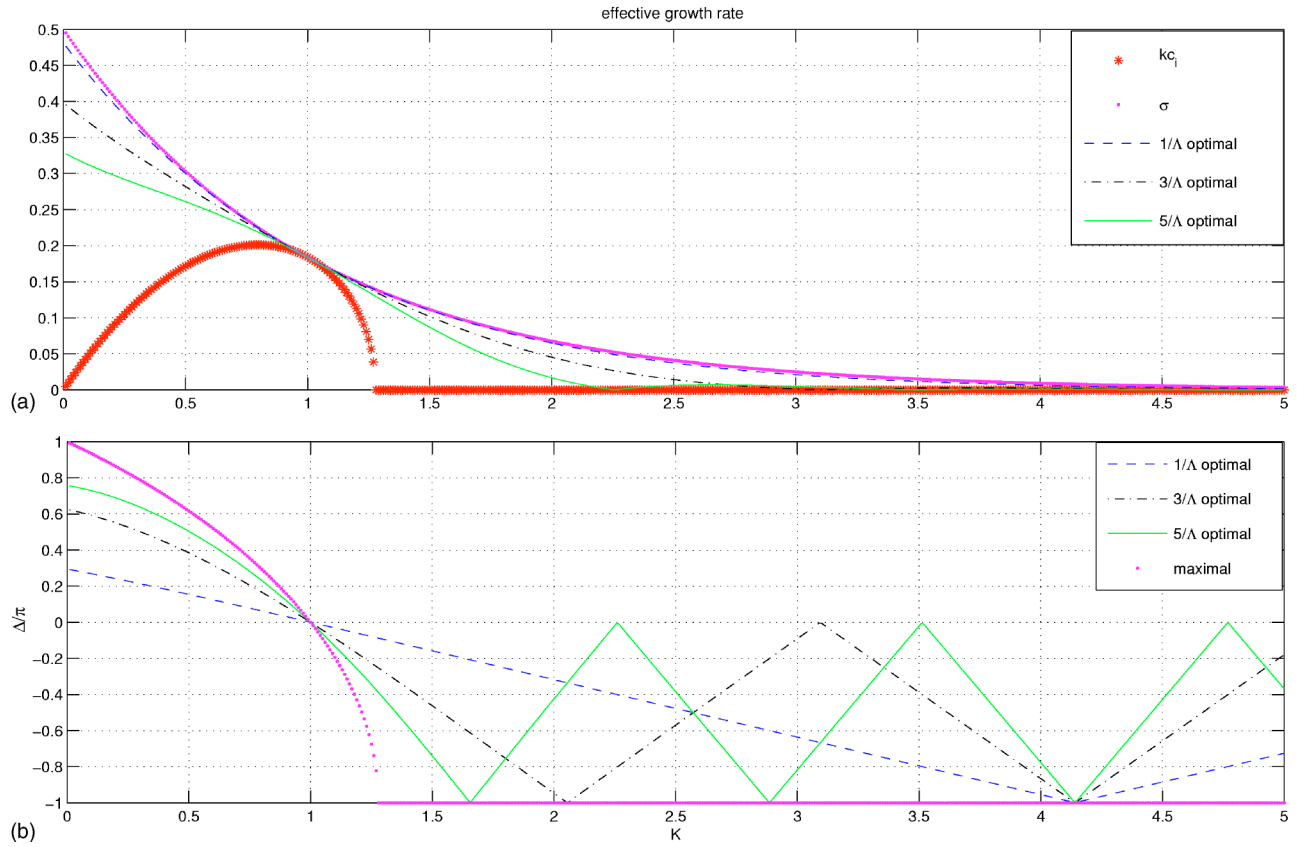

FIG. 6. (a) The normalized effective optimal growth rate $\gamma(t)$, (31), in the Rayleigh model, as a function of wavenumber, for the normalized target times $1/\Lambda$, $3/\Lambda$, $5/\Lambda$, indicated, respectively, by the dashed, dashed-dot, and solid lines. These are compared with the normalized instantaneous growth rate σ, (7a), (bold dots), and the normal mode growth rate kc_i (asterisks). (b) The optimal change in phase difference between CRWs, $\Delta=\epsilon_t-\epsilon_0$, as a function of K, for the three target times. The change in phase difference required for maximal growth is $\Delta_{\max}=\epsilon_b-\epsilon_+=\pi-2\epsilon_+$ and is achieved as target time tends to infinity.

The inviscid, incompressible kinetic energy perturbation can be written as

$$E=\tfrac{1}{2}\langle\nabla\psi,\nabla\psi\rangle=-\tfrac{1}{2}\langle q,\psi\rangle, \tag{32}$$

where $\langle f,g\rangle=\int f^{*}g\,dV$. Substituting (3) into (32), the energy norm is seen to be proportional to

$$E\propto\mathbf{q}^{\dagger}\mathbf{M}\mathbf{q}\equiv\mathbf{e}^{\dagger}\mathbf{e}, \tag{33}$$

where $\mathbf{e}=\mathbf{T}\mathbf{q}$ is the energy generalized coordinate vector, $\mathbf{T}=\mathbf{M}^{1/2}$, and $\mathbf{M}$ is the Hermitian matrix

$$\mathbf{M}=\frac{1}{K}\begin{pmatrix}1 & e^{-K}\\ e^{-K} & 1\end{pmatrix}. \tag{34}$$

Thus, in the energy norm (9a) becomes $\dot{\mathbf{e}}=\mathbf{D}\mathbf{e}$, where $\mathbf{D}=\mathbf{T}\mathbf{A}\mathbf{T}^{-1}$, and its optimal evolution is obtained by the SVD of the propagator matrix $e^{\mathbf{D}t}$ (while the eigenvalues of $\mathbf{A}$ and $\mathbf{D}$ are the same, their singular values are generally different).

In order to understand the different optimal dynamics in the two norms we write (32) explicitly in terms of (3), for synchronized growth, to obtain

$$E=\frac{Q^2}{K}(1+e^{-K}\cos\epsilon), \tag{35}$$

and therefore

$$\frac{dE}{dt}=\frac{d(Q^2)}{dt}\frac{1}{K}(1+e^{-K}\cos\epsilon)-\frac{d\epsilon}{dt}\frac{Q^2}{K}e^{-K}\sin\epsilon. \tag{36}$$

The first term on the right-hand side (RHS) of (36) is the energy growth attributable to the CRW displacement amplitude growth (enstrophy growth). The second term is associated with the Orr[22] mechanism, where kinetic energy can grow simply by changing the tilt of a PV wave without any change in enstrophy. If the CRW phase difference decreases ($\dot{\epsilon}<0$) while these are in the growing configuration ($\sin\epsilon>0$) then the velocity induced by CRW-1 at the home base of CRW-2 becomes more in phase with the self-induced velocity of CRW-2. As a result of this constructive superposition the net magnitude of the velocity vector increases, yielding growth in the kinetic energy.

In fact even without any basic state vorticity gradient, the Orr mechanism can yield growth. Consider, for instance, two "passive CRWs" located at $y=\pm b$ in an infinite constant shear layer $\overline{u}=\Lambda y$ $(-\infty<y<\infty)$. In this case, neither CRW amplitude growth nor counterpropagation is possible since both mechanisms involve the advection of basic state vorticity. As a result the two neutral CRWs are simply advected by the basic state flow at their home bases, $\overline{u}(\pm b)$. Therefore the phase difference decreases continually, $\dot{\epsilon}=-K\Lambda$, which leads to kinetic energy growth via the Orr mechanism, $\dot{E}$

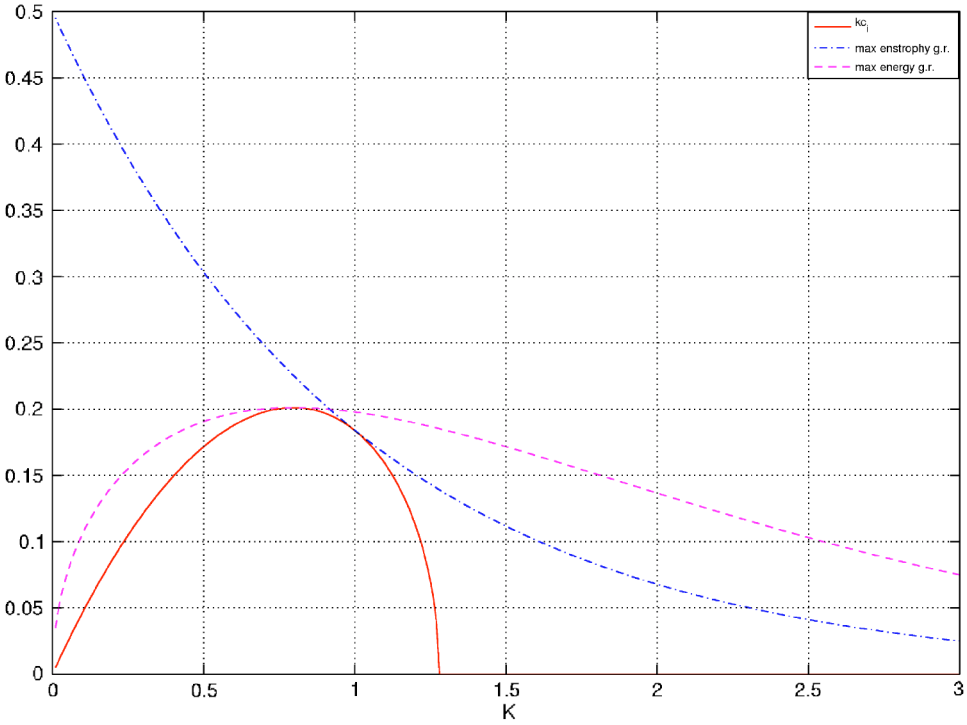

FIG. 7. The maximal instantaneous growth rate in the energy norm (dashed line) vs the maximal instantaneous growth rate in the enstrophy norm (dashed dot line) and the normal mode growth rate (solid line), as functions of wavenumber K.

$=\Lambda Q^2 e^{-K}\sin\epsilon$, when $0<\epsilon<\pi$. Such "passive CRWs" can be used to represent the continuous spectrum, as discussed in the Appendix.

It can be shown that the matrix **D** is normal for the wavenumber of fastest growing normal mode ($K\approx 0.8$), which implies that the maximal growth rate is modal in the energy norm [alternatively, by substituting (5) and (21) into (36) it can be shown that $\partial/\partial K(\dot{E}/E)=0$ is obtained for $K\approx 0.8$]. In the enstrophy norm the fastest growing normal mode does not achieve optimal growth because, although interaction increases with wavelength, so does the Rossby wave propagation rate creating the conflicting requirement that CRWs must hinder each other in order to phase lock. However, in the energy norm the fastest growing normal mode is the optimal configuration for growth. This results from the competition between the two processes attributable to energy growth on the RHS of (36). In Fig. 7 the maximal instantaneous growth rate in the energy norm (dashed line) is plotted as a function of wavenumber alongside maximal instantaneous enstrophy growth rate (dashed-dot line) and normal mode growth rate (solid line). The maximal energy growth rate is smaller than the maximal enstrophy growth rate in the regime where CRWs can phase lock in a hindering configuration ($K<1$) and generally larger in the helping regime ($K>1$). In the hindering regime the optimal evolution in both norms yields $\dot{\epsilon}>0$, and therefore the Orr mechanism reduces energy growth. In the helping regime, the optimal evolution yields $\dot{\epsilon}<0$ and thus both terms at the RHS of (36) act to promote energy growth.

The general stability analysis presented here applies not only to the Rayleigh model but to the discrete spectrum of any conservative plane parallel shear flow which supports modal instability (Heifetz *et al.*[7]). Nevertheless, even in the Rayleigh model the analysis does not take into account the role of the continuous spectrum in the optimal evolution. In the Appendix we suggest a scheme which generalizes the CRW analysis to include the complete spectrum solution, while preserving the principle understanding obtained from the interaction between pairs of CRWs.

IV. CONCLUDING REMARKS

The fundamental physical mechanisms behind the nonmodal linear dynamics of parallel shear flows have been explored in terms of counterpropagating Rossby waves (CRWs), focusing on the discrete spectrum dynamics. The emphasis of previous studies discussing CRWs (Refs. 1, 4, 6, 9–11, 23, and 24) was on understanding modal instability in terms of the mutual amplification of phase-locked CRWs. However, other authors[12–19] have shown that rapid growth of small disturbances on shear flows is dominated by the transient behavior described by nonmodal dynamics. The GST, established by FI96,[19] provides a powerful mathematical tool to compute the initial excitation, growth, and evolution of the perturbation which grows fastest in a given norm, over a given time interval. However, although predictions can be derived following the mathematical formalism of GST, few attempts have been made to describe these results in terms of interaction between components whose properties can be associated with physical mechanisms. The key advantage of the CRW perspective is that the physical mechanism responsible for the propagation of a Rossby wave is readily understood. Propagation results solely from advection of the basic state vorticity by velocities induced by the vorticity anomalies defining the CRW. Interactions between CRWs occur through the "action at a distance" property of the vorticity inversion; operator-induced velocities are felt outside the region where the vorticity anomalies exist. These far-field in-

duced velocities advect the basic state vorticity at the location of another CRW, thus affecting its propagation rate and amplitude.

A key element in the CRW interaction is that their optimal configuration for growth in the enstrophy norm, occurs when the CRWs have equal amplitudes and are tilted against the shear with a phase difference of $\pi/2$. In this configuration the CRWs amplify each other but cannot alter their propagation rates. However, unless the CRWs' intrinsic phase speeds are equal, they will not be able to maintain a phase difference of $\pi/2$ and, as a result, will move relative to each other, exhibiting nonmodal evolution. The system is therefore normal only in the special case of equal CRW intrinsic phase speeds, whereas for all other wavelengths the optimal evolution tends to maximize the duration at the vicinity of $\pi/2$.

If each CRW must hinder the other's counterpropagation rate in order that their phase speeds become equal and phase locking can occur, optimal growth involves an increasing phase difference passing through $\pi/2$ and consequently the growing perturbation evolves to be more tilted against the shear. If, on the other hand, phase locking requires mutual help in counterpropagation, the phase difference decreases during the optimal evolution and the growing structure becomes less tilted against the shear. For short target times the CRW phase difference does not stay far from $\pi/2$. For large target times the initial optimal phase tends to that of the biorthogonal vector of the unstable mode, $\pi-\epsilon_+$, where ϵ_+ is the phase difference in the normal mode configuration. The efficiency of the interaction is determined not only by the CRW phase difference but also by the strength of the CRW interaction coefficient, which is a function of wavelength.

CRWs are vorticity waves and thus describe enstrophy growth in a natural way. However, a simple transformation yields the nonnormal growth in the energy norm from the CRW perspective. Energy growth not only depends on CRW amplification (through increasing air parcel displacements and thus enstrophy growth), but also on a change in their phase difference. The latter results from the Orr[22] mechanism of kinetic energy growth involving the change in tilt of a vorticity disturbance, without a change in enstrophy. When the CRWs are in a growing configuration and their phase difference decreases, the velocities they induce on each other become more in phase. Consequently, the magnitude of the total velocity vector, and thus the kinetic energy, increase. For wavelengths where CRWs must help each other to counterpropagate in order to maintain modal phase locking, the nonmodal evolution is such that the CRW phase difference decreases with time while crossing $\pi/2$ (CRWs become more in phase). Therefore, energy grows due to both CRW amplitude amplification and the Orr mechanism. In cases where CRWs increase their phase difference during their optimal evolution, the growth in energy is usually smaller than the growth in enstrophy, since the Orr mechanism acts to decrease the velocity vector. In the Rayleigh[3] model of a single shear layer, the combination of the two mechanisms implies that the maximal energy growth is achieved by the most unstable normal mode.

The analysis presented here can be applied to the discrete spectrum optimal dynamics of any unstable conservative plane parallel shear flow. The Rayleigh model was used for illustration because it indicates how to generalize the CRW description in order to include continuous spectrum dynamics. In the Rayleigh model the basic state vorticity gradient is concentrated at the edges of the shear zone in two δ functions and, therefore, Rossby wave propagation is possible only at these locations. In a general shear profile we can use the Green function technique and represent the vorticity disturbance as composed of infinite number of vorticity δ functions described as "CRW kernels." These CRW kernels induce streamfunction anomalies of the same form as those of the Rayleigh edge waves, and thus the general linear dynamics can be regarded in terms of the interaction between infinite number of CRW kernels. Each CRW kernel affects the counterpropagation speeds and growth of all other CRWs and in turn is affected by them. The interaction depends on the phase differences and the interaction coefficients between the kernels.

If the flow is discretized in the cross-stream direction into a finite number of strips, the CRW kernel equations can be written into a matrix form which is the direct generalization of the two CRW interaction matrix of the Rayleigh model, and the SVD analysis of the propagator matrix yields the optimal evolution of the complete spectrum. Although the general optimal evolution might be complex, its interpretation remains relatively simple since the nature of interaction between any CRW pair is transparent.

ACKNOWLEDGMENTS

E.H. wishes to thank Michael McIntyre and Orkan Mehmet Umurhan for illuminating discussions. J.M. is grateful for an Advanced Fellowship sponsored jointly by the Natural Environment Research Council and the Environment Agency.

APPENDIX: CRW KERNEL DYNAMICS FOR A GENERAL SHEAR FLOW

Consider a general inviscid, incompressible, shear profile $\overline{u}(y)$ with a mean vorticity profile $\overline{q}(y)$. Then, writing all perturbation variables in the Fourier form $\eta(x,y,t) = \int_0^\infty \hat{\eta}(y,t,k)e^{ikx}dk$, the vorticity perturbation can be written as

$$\hat{q}(y,t,k) = \int_{y'=-\infty}^{\infty} [\hat{q}(y',t,k)\,\delta(y'-y)]dy' \equiv \int_{y'=-\infty}^{\infty} \tilde{q}(y',t,k)dy'. \tag{A1}$$

The "vorticity density kernel" $\tilde{q}(y',t,k)$ induces a streamfunction density $\tilde{\psi}(y,y',t,k)$ which must satisfy $\tilde{q}=-k^2\tilde{\psi}+\tilde{\psi}_{yy}$. Therefore,

$$\tilde{\psi}(y,y',t,k) = \hat{q}(y,t,k)G(y,y'), \tag{A2}$$

with the Green function

$$G(y,y') = -\frac{e^{-k|y-y'|}}{2k} \tag{A3a}$$

for unbounded flows and

$$G(y,y') = \frac{-1}{k\sinh(2kb)} \times \begin{cases} \sinh[k(b+y')]\sinh[k(b-y)] & \text{for } y' \leqslant y < b \\ \sinh[k(b-y')]\sinh[k(b+y)] & \text{for } -b < y \leqslant y' \end{cases} \tag{A3b}$$

for bounded flows at $y=\pm b$ [where $v=\psi_x(y\pm b)=0$]. Hence, for open flows, the inversion of (A1) can be written as

$$\hat{\psi}(y,t,k) = \int_{y'=-\infty}^{\infty} \hat{q}(y',t,k)G(y,y')dy'. \tag{A4}$$

Substituting (A1) and (A4) into the linearized vorticity equation (2),

$$\dot{\hat{q}} = -ik[\overline{u}(y)\hat{q} + \overline{q_y}(y)\hat{\psi}], \tag{A5}$$

and writing the vorticity in terms of amplitude and phase, $\hat{q}(y,t)=Q(y,t)e^{i\epsilon(y,t)}$, we obtain for the real and the imaginary parts in (A5),

$$\dot{Q}(y) = -k\overline{q_y}(y)\int_{y'=-\infty}^{\infty} Q(y')G(y,y')\sin\epsilon(y,y')dy', \tag{A6a}$$

$$\dot{\epsilon}(y) = -k\left\{\overline{u}(y) + \overline{q_y}(y) \times \int_{y'=-\infty}^{\infty} [Q(y')/Q(y)]G(y,y')\cos\epsilon(y,y')dy'\right\}. \tag{A6b}$$

(A6) indicates that each CRW kernel changes its amplitude and phase due to cross-stream advection of the mean vorticity in its own layer, where the cross-stream velocity is attributable to all other kernels and attenuated according to the Green function $G(y,y')$ and the relative phase $\epsilon(y,y') = \epsilon(y) - \epsilon(y')$. Hence, the mechanism of amplitude growth and counterpropagation is the same as for a CRW pair of the discrete spectrum, except that here each CRW kernel affects, and is being affected by, an infinite number of other kernels.

In the case of the Rayleigh model, where $\overline{q_y} = \Lambda[\delta(y-b) - \delta(y+b)]$, we can write

$$\hat{q}(y,t,k) = [Q_1(t)e^{i\epsilon_1(t)}\delta(y+b) + Q_2(t)e^{i\epsilon_2(t)}\delta(y-b)] + Q(y,t)e^{i\epsilon(y,t)}, \tag{A7}$$

where the two terms at the square brackets compose the discrete spectrum solution, (3a), and the latter term on the RHS represents the continuous spectrum solution at $y \neq \pm b$. Substituting (A7) and (A3a) into (A6), we obtain

$$\frac{\dot{Q}_1}{Q_1} = -\frac{\Lambda}{2}\left[\int_{y'=-\infty}^{\infty} \frac{Q(y')}{Q_1}e^{-k|y'+b|}\sin[\epsilon_1 - \epsilon(y')]dy' + \frac{Q_2}{Q_1}e^{-K}\sin[\epsilon_1 - \epsilon_2]\right], \tag{A8a}$$

$$\frac{\dot{Q}_2}{Q_2} = \frac{\Lambda}{2}\left[\int_{y'=-\infty}^{\infty} \frac{Q(y')}{Q_2}e^{-k|y'-b|}\sin[\epsilon_2 - \epsilon(y')]dy' + \frac{Q_1}{Q_2}e^{-K}\sin[\epsilon_2 - \epsilon_1]\right], \tag{A8b}$$

$$\dot{\epsilon}_1 = -kc_1^1 - \frac{\Lambda}{2}\left[\int_{y'=-\infty}^{\infty} \frac{Q(y')}{Q_1}e^{-k|y'+b|}\cos[\epsilon_1 - \epsilon(y')]dy' + \frac{Q_2}{Q_1}e^{-K}\cos[\epsilon_1 - \epsilon_2]\right], \tag{A9a}$$

$$\dot{\epsilon}_2 = -kc_2^2 + \frac{\Lambda}{2}\left[\int_{y'=-\infty}^{\infty} \frac{Q(y')}{Q_2}e^{-k|y'-b|}\cos[\epsilon_2 - \epsilon(y')]dy' + \frac{Q_1}{Q_2}e^{-K}\cos[\epsilon_2 - \epsilon_1]\right], \tag{A9b}$$

together with $\dot{Q}(y)=0$ and $\dot{\epsilon}(y)=-k\overline{u}(y)$ for $y \neq \pm b$. In the case of zero continuous spectrum (A8) and (A9) become (5) and (6). The continuous spectrum PV kernels are passive in the sense that, in the absence of local basic state vorticity gradient, they cannot grow or counterpropagate and therefore are simply advected by the basic state velocity. Nevertheless, as indicated by (A8) and (A9), the velocity they induce affects the discrete spectrum growth and phase (Bishop and Heifetz[24] showed how such interaction leads to a linear absolute baroclinic instability of the discrete spectrum in the semi-infinite Eady[25] model, whose barotropic analog is the semi-infinite Rayleigh model). Furthermore, as discussed in Sec. III D, even in the absence of a basic state vorticity gradient, the continuous spectrum can grow in the energy norm, without a change in enstrophy, via the Orr[22] mechanism.

Discretizing (A6) into N strips in the y direction and weighting the mean vorticity gradient by the strip width $\Delta y(j)=y(j+1)-y(j)$, we can then write (A6) in the matrix form

$$\dot{\hat{\mathbf{q}}} = \mathbf{A}\hat{\mathbf{q}}, \tag{A10a}$$

where

$$\mathbf{A} = -ik[\mathbf{U} + \mathbf{Q_y}\mathbf{G}] \tag{A10b}$$

and

$$\hat{\mathbf{q}} = \begin{pmatrix} \hat{q}_1 \\ \hat{q}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{q}_{N-1} \\ \hat{q}_N \end{pmatrix}, \tag{A10c}$$

$$\mathbf{U} = \begin{pmatrix} U_1 & 0 & \cdots & & 0 \\ 0 & U_2 & \cdots & & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & & \cdot \\ & & & U_{N-1} & \\ 0 & & \cdots & & U_N \end{pmatrix}, \tag{A10d}$$

$$\mathbf{Q_y} = \begin{pmatrix} \overline{q}_{y1} & 0 & \cdots & & 0 \\ 0 & \overline{q}_{y2} & \cdots & & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & & \cdot \\ & & & \overline{q}_{yN-1} & \\ 0 & & \cdots & & \overline{q}_{yN} \end{pmatrix}, \tag{A10e}$$

$$\mathbf{G} = \begin{pmatrix} G(1,1) & G(1,2) & \cdots & G(1,N-1) & G(1,N) \\ G(2,1) & G(2,2) & \cdots & G(2,N-1) & G(2,N) \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & & \cdot \\ G(N-1,1) & G(N-1,2) & \cdots & G(N-1,N-1) & G(N-1,N) \\ G(N,1) & G(N,2) & \cdots & G(N,N-1) & G(N,N) \end{pmatrix}, \tag{A10f}$$

where $\mathbf{G}$ is Hermitian [e.g., $G(i,j)=G(j,i)=-e^{-k|y(i)-y(j)|}/2k$ in open flow]. GST analysis can be applied in the enstrophy norm by exploring the SVD of the propagator matrix $e^{\mathbf{A}t}$, and in the energy norm by calculating the SVD of the transformed propagation matrix $e^{\mathbf{D}t}$, as described in Sec. III D. Discretization of (A4) immediately implies that the 2×2 matrix $\mathbf{M}$, (34), generalizes to $-\mathbf{G}$ and therefore in energy coordinates the evolution equation becomes $\dot{\hat{\mathbf{e}}}=\mathbf{D}\hat{\mathbf{e}}$, where $\mathbf{D}=\mathbf{G}^{1/2}\mathbf{A}\mathbf{G}^{-1/2}$.

[1] F. P. Bretherton, "Baroclinic instability and the short wave cut-off in terms of potential vorticity," Q. J. R. Meteorol. Soc. **92**, 335 (1966).

[2] C. G. Rossby, "Planetary flow patterns in the atmosphere," Q. J. R. Meteorol. Soc. **66**, 68 (1940).

[3] Lord Rayleigh, "On the stability, or instability, of certain fluid motions," Proc. London Math. Soc. **9**, 57 (1880).

[4] E. Heifetz, C. H. Bishop, and P. Alpert, "Counter-propagating Rossby waves in the barotropic Rayleigh model of shear instability," Q. J. R. Meteorol. Soc. **125**, 2835 (1999).

[5] B. J. Hoskins, M. E. McIntyre, and A. W. Robertson, "On the use and significance of isentropic potential vorticity maps," Q. J. R. Meteorol. Soc. **111**, 877 (1985).

[6] E. Heifetz, C. H. Bishop, B. J. Hoskins, and J. Methven, "The counter-propagating Rossby wave perspective on baroclinic instability. Part I: Mathematical basis," Q. J. R. Meteorol. Soc. **130**, 211 (2004).

[7] J. G. Charney and M. E. Stern, "On the stability of internal baroclinic jet in a rotating atmosphere," J. Atmos. Sci. **19**, 159 (1962).

[8] R. Fjørtoft, "Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex," Geofys. Publ. **17**, 1 (1950).

[9] J. Methven, E. Heifetz, B. J. Hoskins, and C. H. Bishop, "The counter-propagating Rossby wave perspective on baroclinic instability. Part III: Primitive equation disturbances on the sphere," Q. J. R. Meteorol. Soc. (in press).

[10] J. Methven, B. J. Hoskins, E. Heifetz, and C. H. Bishop, "The counter-propagating Rossby wave perspective on baroclinic instability. Part IV: Nonlinear life cycles," Q. J. R. Meteorol. Soc. (in press).

[11] E. Heifetz, J. Methven, B. J. Hoskins, and C. H. Bishop, "The counter-propagating Rossby wave perspective on baroclinic instability. Part II: Application to the Charney model," Q. J. R. Meteorol. Soc. **130**, 233 (2004).

[12]B. F. Farrell, "The initial growth of disturbances in a baroclinic flow," J. Atmos. Sci. **39**, 1663 (1982).
[13]B. F. Farrell, "Modal and non-modal baroclinic waves," J. Atmos. Sci. **41**, 668 (1984).
[14]B. F. Farrell, "Optimal excitation of neutral Rossby waves," J. Atmos. Sci. **45**, 163 (1988).
[15]B. F. Farrell, "Optimal excitation of baroclinic waves," J. Atmos. Sci. **46**, 3279 (1989).
[16]L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, "Hydrodynamic stability without eigenvalues," Science **261**, 578 (1993).
[17]P. J. Schmid and D. S. Henningson, *Stability and Transition in Shear Flows*, Applied Mathematical Sciences Vol. 142 (Springer, New York, 2001).
[18]S. Le Dizes, "Modal growth and non-modal growth in a stretched shear layer," Eur. J. Mech. B/Fluids **22**, 411 (2003).
[19]B. F. Farrell and P. J. Ioannou, "Generalized stability theory. Part I: Autonomous operators," J. Atmos. Sci. **53**, 2025 (1996).
[20]P. G. Drazin and W. H. Reid, *Hydrodynamic Stability* (Cambridge University Press, Cambridge, 1981).
[21]I. M. Held, "Pseudomomentum and the orthogonality of modes in shear flow," J. Atmos. Sci. **42**, 2280 (1985).
[22]W. McF. Orr, "Stability or instability of the steady motions of a perfect liquid," Proc. R. Ir. Acad., Sect. A **27**, 9 (1907).
[23]H. C. Davies and C. H. Bishop, "Eady edge waves and rapid development," J. Atmos. Sci. **51**, 1930 (1994).
[24]C. H. Bishop and E. Heifetz, "Apparent absolute instability and the continuous spectrum," J. Atmos. Sci. **57**, 3592 (2000).
[25]E. T. Eady, "Long waves and cyclone waves," Tellus **1**, 33 (1949).