QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY
Q. J. R. Meteorol. Soc. (2009)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.519

Canonical Hamiltonian representation of pseudoenergy in shear flows using counter-propagating Rossby waves

E. Heifetz*, N. Harnik and T. Tamarin
Department of Geophysics, Tel-Aviv University, Israel

ABSTRACT: Pseudoenergy serves as a non-canonical Eulerian Hamiltonian of linearized shear flow systems. It is non-canonical in the sense that the canonical Hamilton equations cannot be written when the dynamical variable is taken as the (potential) vorticity. Here we apply the counter-propagating Rossby wave kernel (KRW) perspective to obtain a compact form of the pseudoenergy as a domain integral of the local KRW pseudomomentum carried by the instantaneous KRW phase speed in the mean flow frame of reference. Written this way, with the generalized momenta taken as the KRW pseudomomenta and the generalized coordinates as the instantaneous KRW locations, canonical Hamilton equations can be derived both in their continuous (using functional derivatives) and discrete (using function derivatives) forms.

As a simple example of the insight such a formulation can yield, we reexamine the classical stability transition from Rayleigh to Couette flow. In this transition the instability is lost even though the classical necessary conditions of Fjørtoft and Rayleigh are still satisfied. The pseudoenergy-KRW formulation allows to interpret the stabilization both as an inability of the KRWs to phase lock constructively, and in terms of the pseudoenergy becoming negative. These two apparent different rationalizations are shown to be essentially one and the same. Copyright © 2009 Royal Meteorological Society

KEY WORDS pseudoenergy; Eulerian canonical Hamiltonian; Rossby wave interaction; shear flow

Received 23 May 2008; Revised 24 July 2009; Accepted 2 September 2009

1. Introduction

It is common practice in many physical disciplines, to try to describe conserved dynamical systems in a canonical Hamiltonian form (e.g. Goldstein, 1969; Peskin and Schroede, 1995). The generalized momenta and coordinates of the canonical representation are considered to be the natural physical framework to describe the system. In geophysical fluid dynamics, such a representation is possible in a particle-following Lagrangian framework, where the Hamiltonian is the total integrated energy of all the fluid particles, and the particles' positions and velocities are the generalized coordinates and momenta. Since a Lagrangian framework is usually too complex for practical use, much effort in developing a Hamiltonian geophysical fluid dynamics theory has been directed towards its representation in the simpler Eulerian framework. However, in the transition to the Eulerian framework, the phase space is degenerated due to the inherent 'particle-relabelling symmetry' along surfaces of constant (potential) vorticity (hereafter generally PV), which is Lagrangianly conserved (e.g. Shepherd, 1990; Salmon, 1988). Under this symmetry, the Hamiltonian becomes invariant under translation of fluid particles along PV surfaces. This reduction in phase space prevents a canonical representation of the dynamics in the Eulerian framework, so that only a non-canonical representation can be obtained (Shepherd, 1990; Salmon, 1988; Chapter 7 of Salmon 1998 gives a comprehensive review of non-canonical Hamiltonian fluid dynamics). In this paper, we will show that, for linear dynamics, we can nonetheless reformulate the Hamilton equations in a canonical form, by using a counter-propagating Rossby wave perspective.

Shear flows with basic states that are constant in time and in the zonal direction conserve both pseudoenergy and pseudomomentum, which are exact wave activity invariants of the nonlinear dynamics (they are defined up to the Casimirs of the flow which can be the integrals of any function of the PV). For the linearized dynamics, the pseudoenergy can be shown to become the non-canonical Hamiltonian, whereas the pseudomomentum becomes a conserved Noether current. In the context of linear instability, the conservation of pseudomomentum and pseudoenergy yield the two necessary conditions for modal shear instability – the Rayleigh (1880) and Fjørtoft (1950) conditions, respectively. A mechanistic interpretation of these conditions is obtained in terms of a mutual interaction of two oppositely propagating Rossby waves, which phase lock and reinforce each other in the presence of shear, due to the action-at-a-distance nature of PV anomalies (e.g. Hoskins *et al.*, 1985; Heifetz *et al.*, 2004a,b). This is best illustrated for the simple case where the mean flow PV gradients are concentrated in two localized regions, each supporting a Rossby wave which propagates to the left of the local mean PV gradient. Though the waves are PV localized, each induces a non-local velocity field which affects the other wave by advecting the mean PV gradient. The Rayleigh condition

*Correspondence to: E. Heifetz, Department of Geophysics, Tel-Aviv University, Israel, 69978. E-mail: eyalh@cyclone.tau.ac.il

– that the mean PV gradient changes sign between the two jumps – enables the two Rossby waves to align in a way in which they interact constructively to yield mutual instantaneous amplification. When the Fjørtoft condition is satisfied – that the mean PV gradient is globally positively correlated with the mean wind – the waves propagate counter the local mean wind (hence denoted by Bretherton (1966) as 'counter-propagating Rossby wave', CRW). In a shear flow, such a counter-propagation will reduce the relative phase speed between the two. Under the right conditions, the CRW interaction will allow a phase-locked coherent propagation in a mutually reinforcing configuration, leading to modal growth.

Heifetz and Methven (2005) generalized this two-wave interaction formulation to one based on a multiple interaction of an infinite number of localized Rossby wave 'kernels' (referred to, hereafter, as a kernel Rossby wave, KRW), each propagating on its local PV gradient, and inducing a non-local velocity field which affects all other kernels. Thus, each KRW changes its amplitude and phase due to advection of the mean PV in its own layer, where the advecting velocity is attributable to all other kernels. The growth, propagation and interaction mechanism are the same as for the two-wave example mentioned above, except that here each KRW affects, and is being affected by, an infinite number of other kernels.

Formulating the PV evolution in terms of a two-CRW interaction, Heifetz *et al.* (2004a) also obtained a pair of canonical equations, in which pseudoenergy is the Hamiltonian, and the CRW positions and pseudomomentum are the generalized coordinates and momenta respectively. In this paper, we wish to generalize this two-wave canonical formulation to the more general multiple KRW system, and to make use of the mechanistic understanding that the KRW framework provides to better understand the Hamiltonian perspective. By choosing the generalized coordinates to be the KRWs' positions, and the generalized momenta to be the local contribution of the KRWs to the pseudomomentum, pseudoenergy obtains a compact form from which the canonical Hamilton equations are derived.

We present the mathematical formulation in section 2, and then (in section 3) apply it to understand why modal instability can be lost even when the two necessary conditions for instability of Rayleigh and Fjørtoft are satisfied. We consider one of the simplest cases of a stability transition – that of a single shear layer bounded by approaching boundaries. At the vicinity of the transition zone, both the conditions are satisfied however as the boundaries become too close to the shear layer, the pseudoenergy integral becomes negative, and instability is lost. The KRW–Hamiltonian formulation allows us to interpret the stabilization both as an inability of the KRWs to phase lock in a growing configuration, and as a decrease of the KRW energy–enstrophy ratio, which yields a negative pseudoenergy. Thus, the two seemingly different rationalizations, KRW phase-locking and the vanishing of the pseudoenergy constraint, are essentially one and the same. We conclude with a short discussion in section 4.

2. Canonical KRW Hamilton equations

2.1. KRW equations

In this section we present the linear dynamical equations in terms of the amplitude and phase of a set of KRWs, which are mutually interacting. We consider a basic shear zonal flow U whose perturbation is assumed to be described by the linearized horizontal advection of PV:

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right) q' = -v'\overline{q}_y\,, \tag{1}$$

where q' is the perturbation PV, $\overline{q}_y$ is the mean meridional PV gradient and v' is the perturbation meridional velocity. This equation can be applied when the PV is either the Ertel–Rossby one evaluated on isentropic surfaces or the quasi-geostrophic PV on horizontal surfaces, or to barotropic instability on horizontal surfaces where the PV becomes the vorticity. Here, for the sake of simple demonstration, we refer to the latter case where both U and $\overline{q}_y$ are functions of y only and $q' = (\partial v'/\partial x) - (\partial u'/\partial y)$. We apply a zonal Fourier decomposition on the perturbation

$$q'(x, y, t) = \int_{k=0}^{\infty} q_k(y, t)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k\,, \tag{2a}$$

$$v'(x, y, t) = -\mathrm{i}\int_{k=0}^{\infty} v_k(y, t)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k\,, \tag{2b}$$

(the –i on the RHS of (2b) indicates that a positive PV anomaly induces positive meridional velocity a quarter of wavelength to the east of it) and assume that if $\mathcal{L}_y[v_k(y,t)] = q_k(y,t)$, then by inversion

$$v_k(y, t) = \int_{y'=y_{\min}}^{y_{\max}} q_k(y', t)G(y, y')\,\mathrm{d}y', \tag{3}$$

where $G(y, y')$ is the positive definite Green function satisfying $\mathcal{L}_y[G(y, y')] = \delta(y - y')$ and the associated boundary conditions at $y = (y_{\min};\ y_{\max})$. Writing the PV perturbation in terms of amplitude and phase, $q_k(y,t) = Q_k(y,t)\mathrm{e}^{\mathrm{i}\epsilon_k(y,t)}$, substituting into (1) using (2) and (3), and taking the real and imaginary parts of (1) we get

$$\dot{Q}(y) = \overline{q}_y(y)\int_{y'=y_{\min}}^{y_{\max}} Q(y')G(y, y')\sin[\epsilon(y, y')]\mathrm{d}y', \tag{4a}$$

$$\dot{\epsilon}(y) = -kU(y) + \frac{\overline{q}_y(y)}{Q(y)}\int_{y'=y_{\min}}^{y_{\max}} Q(y')G(y, y')\cos[\epsilon(y, y')]\,\mathrm{d}y', \tag{4b}$$

where the subscript k has been dropped and $\epsilon(y, y') \equiv \epsilon(y) - \epsilon(y')$. Here we refer only to the cases where PV anomalies result solely from mean PV advection, so that (4b) is not singular in regions where $\overline{q}_y(y) = 0$, since $q(y)$ also vanishes there. As described in Heifetz and Methven (2005), these equations can be considered as

DOI: 10.1002/qj

the continuous generalization of the CRW-pair equations described in Heifetz *et al.* (2004a). Equation pair (4) indicates that each KRW changes its amplitude and phase due to meridional advection of the mean PV in its own layer, where the meridional velocity is attributable to all other kernels and attenuated according to the Green function $G(y, y')$ and the relative phase $\epsilon(y, y')$. Hence, the mechanism of amplitude growth and counter-propagation is the same as for a CRW pair, except that here each KRW affects, and is being affected by, an infinite number of other kernels.

2.2. KRW representation of pseudoenergy and pseudomomentum

The KRW formulation allows us to obtain a compact expression for the pseudoenergy integral. We first write the domain-averaged energy E in terms of v_k and q, for a given wavelength $\lambda = 2\pi/k$, and width $D = (y_{\max} - y_{\min})$:

$$E = \left\langle \frac{v_k q}{2k} \right\rangle = \frac{1}{\lambda D} \int_{x=0}^{x=\lambda} \int_{y=y_{\min}}^{y_{\max}} \frac{1}{4k} Re(v_k^* q)\, \mathrm{d}x \mathrm{d}y. \quad (5)$$

Substituting (3) into (5) yields:

$$E = \frac{1}{4kD} \int \int_{y'=y_{\min}}^{y_{\max}} Q(y)Q(y')G(y, y') \cos[\epsilon(y, y')]\, \mathrm{d}y' \mathrm{d}y. \quad (6)$$

Taking the time derivative of E, using (4) to express the time derivative of Q and ϵ, we get after some algebra:

$$\dot{E} = \frac{1}{4D} \int \int_{y'=y_{\min}}^{y_{\max}} Q(y)Q(y')G(y, y') \times \sin[\epsilon(y, y')]\,[U(y) - U(y')]\, \mathrm{d}y' \mathrm{d}y. \quad (7a)$$

As indicated in Harnik and Heifetz (2007), the absence of the mean PV gradient from the energy growth expression results from an exact cancellation by two opposing effects of the mean PV advection on the energy growth. When the KRW action-at-a-distance interaction acts to increase the enstrophy (the KRW amplitudes), it also acts to weaken the correlation between v_k and q (by increasing the phase between the KRWs) and vice versa. As is evident from the RHS, only in the presence of shear and when the KRW are tilted against it (i.e. the correlation between $\sin[\epsilon(y, y')]$ and $[U(y) - U(y')]$ is positive), we get energy growth. This was referred to by Harnik and Heifetz as the CRW perspective of the Orr (1907) mechanism.

Flipping the dummy variables (y, y') in the RHS of (7a), taking into account that $G(y, y') = G(y', y)$ and $\sin[\epsilon(y, y')] = -\sin[\epsilon(y', y)]$, yields then a simpler expression for the energy tendency:

$$\dot{E} = \frac{1}{2D} \int \int_{y'=y_{\min}}^{y_{\max}} Q(y)Q(y')G(y, y') \sin[\epsilon(y, y')]U(y)\, \mathrm{d}y' \mathrm{d}y. \quad (7b)$$

The time-invariant pseudoenergy integral is $H \equiv E + S$, where

$$S = \left\langle -\frac{U(y)}{\overline{q}_y(y)} \frac{q^2}{2} \right\rangle = -\frac{1}{4D} \int_{y=y_{\min}}^{y_{\max}} \frac{U(y)}{\overline{q}_y(y)} Q^2(y)\, \mathrm{d}y. \quad (8)$$

H is indeed conserved with time, as can be verified by taking the time derivative of S, using (4a) and comparing it with (7b). Then writing H in terms of (6) and (8), using (4b) to express the integral over y' in (6), we obtain

$$\begin{aligned} H &= \frac{1}{4D} \int_{y=y_{\min}}^{y_{\max}} \frac{Q^2(y)}{\overline{q}_y(y)} \frac{\dot{\epsilon}(y)}{k}\, \mathrm{d}y \\ &= \frac{1}{4D} \int_{y=y_{\min}}^{y_{\max}} \frac{Q^2(y)}{\overline{q}_y(y)} \dot{\chi}(y)\, \mathrm{d}y = \text{const}, \end{aligned} \quad (9)$$

where $\chi(y) \equiv \epsilon(y)/k$ is defined as the 'KRW location'. (9) can be related to the time-independent pseudomomentum integral (P):

$$\begin{aligned} P &= \left\langle \frac{q^2}{2\overline{q}_y(y)} \right\rangle = \frac{1}{4D} \int_{y=y_{\min}}^{y_{\max}} \frac{Q^2(y)}{\overline{q}_y(y)}\, \mathrm{d}y \\ &= \int_{y=y_{\min}}^{y_{\max}} p(y)\, \mathrm{d}y, \end{aligned} \quad (10)$$

where we have defined $p(y) \equiv Q^2(y)/4D\overline{q}_y(y)$ to be the pseudomomentum density of the local (in y) KRW contribution to the pseudomomentum integral P. We note that while P is conserved in time†, $p(y)$ is generally not. Using this definition of the localized contribution to pseudomomentum yields the compact expression for the pseudoenergy integral:

$$H = \int_{y=y_{\min}}^{y_{\max}} p(y, t)\, \dot{\chi}(y, t)\, \mathrm{d}y. \quad (11)$$

We see that pseudoenergy is an integral over minus the product of local instantaneous pseudomomentum and local instantaneous phase speed. This relation suggests the definition of pseudoenergy density, $h(y, t) \equiv p(y, t)\dot{\chi}(y, t)$. An elegant interpretation of the local pseudomomentum is the amount of PV flux which crossed the local latitude line over the lifetime of the perturbation (Held, 2000), which is essentially the amount of PV-material 'stored' in the local KRW. Thus, the local pseudoenergy is minus an instantaneous *zonal* flux of

†The conservation of P, which stems from the domain-integrated PV flux vanishing, can be obtained directly from the KRW formulation when multiplying (4a) by $Q(y)/\overline{q}_y(y)$ and integrating by y to obtain

$$\dot{P} = \frac{1}{4D} \int \int_{y'=y_{\min}}^{y_{\max}} Q(y)Q(y')G(y, y') \sin[\epsilon(y, y')]\, \mathrm{d}y' \mathrm{d}y.$$

The RHS vanishes immediately when flipping the dummy variables (y, y'). Essentially this results from the symmetry in the interaction (represented by $G(y, y')$) between each KRW pair – the PV fluxes induced by the meridional flow of one kernel on the other are equal and oppositely signed.

DOI: 10.1002/qj

'stored' KRW PV, carried by the wave in a frame of rest[‡]. For the case of normal modes, for which by definition $\dot{\chi} = c_{nm}$ =const, (11) reduces to Equation (2.18) of Held (1985). Moreover, the pseudoenergy condition for normal mode instability – that it has to vanish – is trivially obtained in this case.

2.3. Canonical Hamiltonian KRW equations

A continuous system in the general form of (11), i.e.

$$\mathcal{H} = \int_{y=y_{\min}}^{y_{\max}} h[p(y,t), \chi(y,t)]\,\mathrm{d}y = \text{const}, \qquad (12)$$

where p and χ are the continuous generalized momenta and coordinates[§], is considered canonical if

$$\frac{\delta\mathcal{H}}{\delta p} = \dot{\chi}\ ; \qquad \frac{\delta\mathcal{H}}{\delta\chi} = -\dot{p}\,, \qquad (13a,b)$$

where the functional derivative $\delta\mathcal{F}/\delta f$ is defined for infinitesimal variation δ by (e.g. Salmon, 1998)

$$\delta\mathcal{F}[f(y)] \equiv \mathcal{F}[f+\delta f] - \mathcal{F}[f] \equiv \int_{y=y_{\min}}^{y_{\max}} \frac{\delta\mathcal{F}}{\delta f(y')}\delta f(y')\,\mathrm{d}y'. \qquad (14)$$

Alternatively, applying the Poisson bracket formulation

$$\{\mathcal{F},\mathcal{G}\} \equiv \int_{y=y_{\min}}^{y_{\max}} \left[\frac{\delta\mathcal{F}}{\delta\chi(y')}\frac{\delta\mathcal{G}}{\delta p(y')} - \frac{\delta\mathcal{F}}{\delta p(y')}\frac{\delta\mathcal{G}}{\delta\chi(y')}\right]\mathrm{d}y', \qquad (15)$$

the Hamilton equations become

$$\{\chi,\mathcal{H}\} = \dot{\chi}\ ; \quad \{p,\mathcal{H}\} = \dot{p}\,. \qquad (16a,b)$$

Using these definitions, it is straightforward to verify that H satisfies (13a) and (16a). (13b) or (16b) are obtained after some algebra when (4b) is substituted into (11) and (14) and (15) are applied[¶].

The fact that (13) or (16) are the KRW amplitude and phase evolution equations suggests, along with expression (11), that the amplitude–phase evolution and phase locking are at the heart of basic conserved quantities like pseudomomentum and pseudoenergy. In the next section we show this explicitly, using a classic problem in which a stability transition occurs even when the classical necessary conditions of Fjørtoft (1953) and Rayleigh (1880) are still satisfied.

[‡]A similar minus sign issue arises for the definition of pseudomomentum. Andrews and McIntyre (1976) defined the pseudomomentum as $-p$ (minus the amount of PV material which flows through a latitude circle). In their definition, a wave carrying positive pseudomomentum will accelerate the zonal mean flow once this pseudomomentum is deposited in it.

[§]We avoid the standard notation of q for the generalized coordinate, since the latter denotes PV.

[¶]Note however that this Hamiltonian is different from the familiar canonical continuous representation,

$$H = \int p(y)\dot{\chi}(y)\,\mathrm{d}y - L(\chi,\dot{\chi},y,t),$$

where the Lagrangian L satisfies $\delta L/\delta\dot{\chi} = p$. L cannot be properly defined since the KRW equation set (4) does not provide a Legendre transformation from the basis $(\chi,\dot{\chi})$ to (χ,p). Hence, this system has no direct translation to the Euler–Lagrange formulation.

3. The transition from Rayleigh to Couette flow

We consider one of the simplest set-ups for which the Rayleigh and Fjortoft conditions are met: a 2D, inviscid, incompressible shear layer, bounded by two finite regions of constant flow (Figure 1(a)). We denote the shear layer boundary location as $y = \pm b$, and the wall locations at $y = \pm a$, with a varying from infinity to b. In this set-up, the mean meridional vorticity gradient, which is independent of a, is simply a positive delta function at the northern edge of the shear layer, and a negative one at the southern edge (Heifetz *et al.*, 1999, give details of the infinite domain set-up) dynamics. Since the PV gradient is concentrated into two δ functions, there is not a distinction between the globally defined CRWs, and the localized KRWs – both are the two localized PV δ functions. Thus we will use the term CRW (in line with traditional discussions of this problem), but the analysis also applies to KRWs.

The location of the boundaries strongly affects the stability of the system (e.g. Drazin, 2002), as can be discerned from considering the two extremes. When the boundaries are located at $y \to \pm\infty$, the unstable Rayleigh shear layer is obtained. When the boundaries

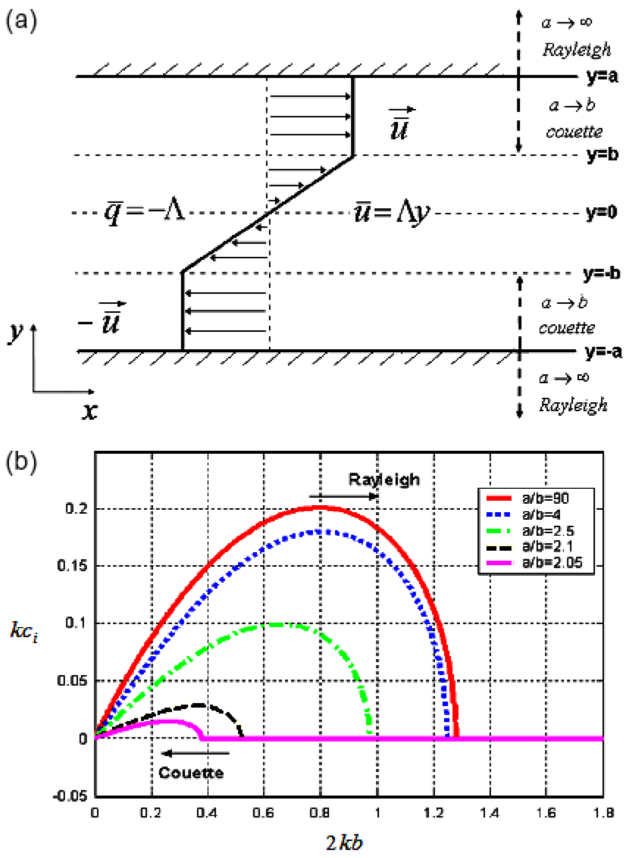

Figure 1. (a) Schematic illustration of the transition of the Rayleigh problem to the plane Couette one. When the boundaries at $y = \pm a$ go to $\pm\infty$, the Rayleigh problem is obtained while, where $a = b$, the Couette problem is obtained. ($y = \pm b$ are the edges of the shear layer.) (b) The growth rate, normalized by the shear Λ, as a function of the normalized wavenumber $K = 2kb$, for different values of the ratio a/b. Where $a/b >> 1$, the Rayleigh problem limit is obtained whereas, when $a/b = 2$, the instability disappears. This figure is available in colour online at www.interscience.wiley.com/journal/qj

are located exactly at the edges of the shear layer, stable Couette flow is established. In between, as the boundaries approach monotonically from $\pm\infty$ to the edges of the shear layer, the instability is gradually reduced, and it disappears when the boundaries reach a distance from the shear layer which equals half the shear layer width. Figure 1(b) shows the normal mode growth rates for this transition. As the boundaries approach the shear layer, the growth rate gradually decreases and the short-wave cut-off, as well as the most unstable mode, are both shifted to smaller wavenumbers.

The most direct effect of the boundaries is through the requirement that the meridional velocity vanish there. When the boundaries are close enough, this strongly reduces the meridional velocity induced by a PV kernel. Physically, this can be understood from the electric charge mirror imaging analogy for PV (Thorpe and Bishop, 1995) – an anti-phased CRW mirror image should be placed at the opposite side of the boundary to cancel the meridional velocity there[‖].

Mathematically, this is expressed in the Green function for this set-up:

$$G(y, y') = \frac{1}{\sinh(2ka)} \begin{cases} \sinh[k(a+y')]\sinh[k(a-y)] \\ \qquad \text{for } y' \le y < a\,, \\ \sinh[k(a-y')]\sinh[k(a+y)] \\ \qquad \text{for } -a < y \le y'\,. \end{cases} \tag{17}$$

In the Rayleigh limit ($a \to \infty$), $G(y, \pm b)$ reduces to the open flow form ($\mathrm{e}^{-k|y\mp b|}/2$) while in the Couette limit ($a \to b$), $G(y, \pm b)$ vanishes. This is consistent with the PV gradient vanishing in this problem, thus there are no CRW kernels.

Note that, since there are only two KRWs in this problem (at $y = \pm b$), we only need to consider the Green functions at these locations. There are thus only two types of Green function – 'self' ($G_s = G(\pm b, \pm b)$) and 'induced' ($G_i = G(\pm b, \mp b)$). G_s affects the self counter-propagation rate whereas G_i affects the strength of CRW interaction. Both of these are reduced as the boundaries approach the shear layer, which explains the loss of instability.[**]

The loss of instability can also be understood from the perspective of pseudoenergy conservation. McIntyre and Shepherd (1987) discussed generally how, due to the nature of the inversion of PV, decreasing the domain size reduces the domain-integrated energy contribution E, relative to the domain-integrated enstrophy $|S|$. Since for modal instability any conserved quantity involving the eddies must vanish, $H = 0$ implies $E = |S|$. If E becomes too small, as the boundaries become too close, the pseudoenergy becomes negative definite and modal instability is ruled out regardless of whether the Rayleigh and Fjørtoft conditions are met.

The Hamiltonian–KRW formulation of the previous section allows us to tie these two approaches to the problem together. To do this, we apply the discrete KRW version suitable to the case where the basic state is discretized to a PV staircase (cf. Appendix). Using (A.4) the two parts of the pseudoenergy $H = E + S$ in terms of the two CRWS become:

$$E = \frac{1}{8ak}\left\{(\widehat{Q}^2_{+b} + \widehat{Q}^2_{-b})G_s + 2\widehat{Q}_{+b}\widehat{Q}_{-b}G_i\cos\epsilon\right\}, \tag{18a}$$

$$S = -\frac{b}{8a}\left\{\widehat{Q}^2_{+b} + \widehat{Q}^2_{-b}\right\}, \tag{18b}$$

where $\epsilon \equiv \epsilon_{+b} - \epsilon_{-b}$.

For this simple symmetric set-up, modal growth is obtained when $\widehat{Q}_{+b} = \widehat{Q}_{-b} \equiv \widehat{Q}$. Under this condition, the ratio between the energy and the absolute value of the enstrophy term is:

$$\frac{E}{|S|} = \frac{1}{kb}\left(G_s + G_i\cos\epsilon\right)\,. \tag{19}$$

When the boundaries approach the shear layer, this ratio vanishes, due to the Green function vanishing. The pseudoenergy condition for normal mode instability, that $H = 0$ ($E/|S| = 1$) yields:

$$\cos\epsilon = \frac{kb - G_s}{G_i}. \tag{20}$$

It is easy to verify, using (A.2b) for this set-up, that this is exactly the condition for phase locking ($\dot{\epsilon} = 0$). Hence, the condition $H = 0$ yields the condition for unstable modal phase locking.

We can obtain a physical picture of the effect of the approaching boundaries on the decrease of E by examining how the velocity fields of the two CRWs superpose (i.e. $\mathbf{v}(y) = \mathbf{v}(y, b) + \mathbf{v}(y, -b)$):

$$E = \frac{1}{2}\left\langle\left[u_b(y) + u_{-b}(y)\right]^2 + \left[v_b(y) + v_{-b}(y)\right]^2\right\rangle. \tag{21}$$

For a given non-dimensional wavenumber kb, the energy–enstrophy ratio of (19) is maximized when $\epsilon = 0$. This is due to the energy being maximized when the two CRWs are exactly in phase (Figure 2). Instability is allowed as long as $E_{\max} > |S|$, and the stability cut-off (Figure 1(b)) occurs when $E = E_{\max} = |S|$.

Looking at the CRW configuration for $E_{\max}$, shown in Figure 2, we see that the meridional velocities superpose constructively in the entire domain. The zonal perturbation velocities, however, superpose constructively outside the shear layer, and superpose destructively inside it. Thus, when the outer region shrinks, the superposition of u tends to decrease the integrated value of u^2, and the overall energy integral becomes too small to match the enstrophy term. Consequently H becomes negative and the instability is lost.

[‖] When two boundaries exist, an infinite series of mirror images is actually needed to cancel the velocity at both boundaries, but the leading-order effect is accounted for by a single mirror image.

[**] McIntyre (2005) actually alludes to this by noting that 'the Rossby-wave propagation mechanism does not have room to operate sufficiently strongly to hold a phase-locked configuration', in the presence of close boundaries.

DOI: 10.1002/qj

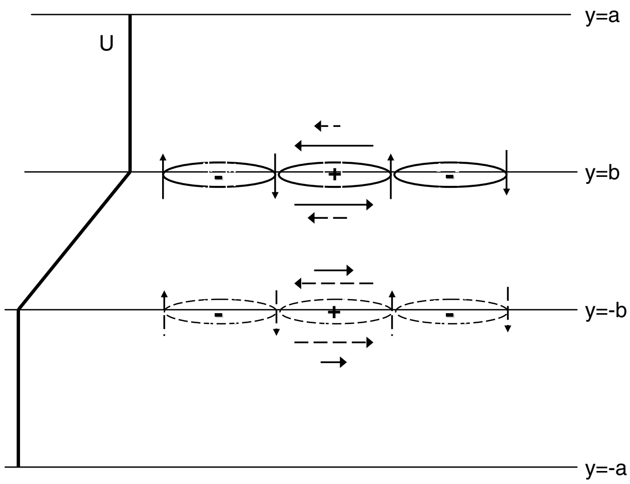

Figure 2. Schematic illustration of the CRW configuration for $E_{\max}$. The velocity fields induced by the CRW at the northern and southern shear edges are represented by the solid and dashed arrows, respectively. While the meridional velocities superpose constructively in the entire domain, the zonal perturbation velocities superpose constructively outside the shear layer ($a > y > b$, and $-b > y > -a$), and superpose destructively inside it ($b > y > -b$).

4. Conclusions

Geophysical shear flows that Lagrangianly conserve PV can be often presented as non-canonical Eulerian Hamiltonian systems (e.g. Shepherd, 1990). They are non-canonical in the sense that, although a Hamiltonian can be well defined, the Hamilton equations cannot be written in their canonical form. Here, in a much more modest attempt, pertaining to linear shear flow dynamics, by taking advantage of the fact that laminar shear sets a preferred direction in the flow, we manage to write the canonical Hamilton equations in terms of multiple action-at-a-distance interactions between localized PV anomalies, propagating as Rossby waves. This formulation however is peculiar in the sense that, while the canonical Hamilton equations can be derived (both in a continuous and discrete forms), a proper Lagrangian cannot be defined due to the lack of a proper Legendre transformation. The resulting kernel Rossby wave (KRW) formulation leads to a very compact form of the pseudoenergy – as the product of pseudomomentum and the local instantaneous disturbance phase speed. This relation has been noted before for normal modes (Held, 1985), but we show here that it holds locally and instantaneously, and therefore can be applied also to non-modal transient growth.

As an example of the insight such a formulation can yield, we revisited the classical stability transition from Rayleigh to Couette flow. In this problem, the loss of normal mode instability occurs when the pseudoenergy integral becomes negative, even though the Rayleigh and Fjørtoft necessary conditions for instability are still satisfied. The KRW formulation explicitly shows that pseudoenergy is maximized when the flow is marginally stable. This is a kinematic action-at-a-distance effect – when all KRWs are in phase the ratio of eddy energy to eddy enstrophy is maximized. At the same time, when the KRWs are in phase, they fully help each other to maintain phase locking against the shear. The vanishing of pseudoenergy occurs when the ability of the KRWs to phase lock in a mutually growing configuration is lost due to a significant weakening of the action-at-a-distance by the approaching boundaries. This occurs because the constraint of vanishing flow at the boundaries considerably weakens the meridional flow which a given PV kernel induces. Since the KRW description is general to shear flows, we expect the conclusions from this simple problem of the relation between pseudoenergy and phase locking to hold for more complex shear instability set-ups.

Since pseudoenergy is an exact invariant for finite amplitude perturbations dynamics, this suggests the KRW formulation might be extendable to finite amplitude (as seems to be the case for nonlinear eddy life cycles, cf. Methven *et al.*, 2005). We are currently investigating this possibility in a semi-Lagrangian perspective.

Acknowledgements

This work is supported by the Israel Science Foundation (ISF, grant 1084/06) and by the Bi-National Israel–USA Science Foundation (BSF, grant 2004087). Part of the work by NH was supported by a European Union Marie Curie International reintegration grant (MIRG-CT-2005-016835). EH is grateful to Wim Verkley, Orkan Umurhan and Ely Kovetz for fruitful discussion.

Appendix

Discrete canonical Hamiltonian KRW formulation

Discretizing the basic state to a piecewise PV staircase (with arbitrarily small but finite jumps):

$$\overline{q}_y(y_n) = (\Delta\overline{q})_n \delta(y - y_n)\,, \qquad \text{(A.1)}$$

where $n = 1, 2, ..., N$ is the number of PV jumps and $(\Delta\overline{q})_n = \overline{q}_{n+1} - \overline{q}_n$ is the difference of the mean PV across the jump n. From (1) it is clear that the PV perturbations are also concentrated as δ-functions on the mean PV interfaces, i.e.

$$q_k(y_n, t) = \widehat{Q}_k(y, t)\mathrm{e}^{\mathrm{i}\epsilon_k(y,t)}\delta(y - y_n) \equiv \widehat{Q}_n \mathrm{e}^{\mathrm{i}\epsilon_n}\delta(y - y_n).$$

For such discretization, equation set (4) can be rewritten as:

$$\dot{\widehat{Q}}_n = (\Delta\overline{q})_n \sum_{j=1}^{N} \widehat{Q}_j G(y_n, y_j) \sin\left[\epsilon(y_n, y_j)\right], \qquad \text{(A.2a)}$$

$$\dot{\epsilon}_n = -kU(y_n) + \frac{(\Delta\overline{q})_n}{\widehat{Q}_n} \sum_{j=1}^{N} \widehat{Q}_j G(y_n, y_j) \cos\left[\epsilon(y_n, y_j)\right], \qquad \text{(A.2b)}$$

DOI: 10.1002/qj

and the pseudomomentum takes the form:

$$P = \frac{1}{4D}\sum_{n=1}^{N}\frac{\widehat{Q}_n^2}{(\Delta\overline{q})_n} \equiv \sum_{n=1}^{N}\widehat{p}_n\,. \qquad \text{(A.3)}$$

Using (6), (8), (11) and (A.3), the pseudoenergy becomes:

$$\begin{aligned} H &= E + S \\ &= \frac{1}{4Dk}\sum_{l=1}^{N}\sum_{m=1}^{N}\widehat{Q}_l\widehat{Q}_m G(y_l, y_m)\cos[\epsilon(y_l, y_m)] \\ &\quad + \sum_{n=1}^{N} -U(y_n)\widehat{p}_n \qquad \text{(A.4)} \\ &= \sum_{n=1}^{N}\widehat{p}_n\dot{\chi}_n\,. \end{aligned}$$

We can now take the derivative of H with respect to χ_n and $\widehat{p}_n$. Using equation set (A.2) and recalling that $\widehat{p}_n = \widehat{Q}_n^2/\{4D(\Delta\overline{q})_n\}$, and that the derivation of quantities of index n have cross-term contributions from the multiple summations when $l = m = n$, we get the discrete version of (13a,b):

$$\frac{\partial H}{\partial \chi_n} = -\dot{\widehat{p}}_n\ ; \quad \frac{\partial H}{\partial \widehat{p}_n} = \dot{\chi}_n\,. \qquad \text{(A.5a,b)}$$

Applying the discrete Poisson bracket formulation:

$$\{\mathcal{F}, \mathcal{G}\} \equiv \sum_{n=1}^{N}\left(\frac{\partial \mathcal{F}}{\partial \chi}\frac{\partial \mathcal{G}}{\partial \widehat{p}_n} - \frac{\partial \mathcal{F}}{\partial \widehat{p}_n}\frac{\partial \mathcal{G}}{\partial \chi_n}\right), \qquad \text{(A.6)}$$

we obtain the discrete version of (16a,b):

$$\{\chi_n, H\} = \dot{\chi}_n\ ; \quad \{p_n, H\} = \dot{p}_n\,. \qquad \text{(A.7a,b)}$$

References

Andrews DG, McIntyre ME. 1976. Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.* **33**: 2031–2048.

Bretherton FP. 1966. Baroclinic instability and the short wave cut-off in terms of potential vorticity. *Q. J. R. Meteorol. Soc.* **92**: 335–345.

Charney JG, Stern ME. 1962. On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.* **19**: 159–172.

Drazin PG. 2002. *Introduction to Hydrodynamic Stability*. Cambridge University Press: Cambridge, UK.

Fjørtoft R. 1950. Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex. *Geofys. Publ.* **17**: 6, 1–52.

Goldstein H. 1969. *Classical Mechanics*. Addison-Wesley: Reading, Mass, USA.

Harnik N, Heifetz E. 2007. Relating over-reflection and wave geometry to the counter-propagating Rossby wave perspective: Toward a deeper mechanistic understanding of shear instability. *J. Atmos. Sci.* **64**: 2238–2261.

Heifetz E, Methven J. 2005. Relating optimal growth to counter-propagating Rossby waves in shear instability. *Phys. Fluids* **17**: 064107.

Heifetz E, Bishop CH, Alpert P. 1999. Counter-propagating Rossby waves in the barotropic Rayleigh model of shear instability. *Q. J. R. Meteorol. Soc.* **125**: 2835–2853.

Heifetz E, Bishop CH, Hoskins BJ, Methven J. 2004a. The counter-propagating Rossby wave perspective on baroclinic instability. I: Mathematical basis. *Q. J. R. Meteorol. Soc.* **130**: 211–232.

Heifetz E, Methven J, Hoskins BJ, Bishop CH. 2004b. The counter-propagating Rossby wave perspective on baroclinic instability. II: Application to the Charney model. *Q. J. R. Meteorol. Soc.* **130**: 233–258.

Held IM. 1985. Pseudomomentum and the orthogonality of modes in shear flows. *J. Atmos. Sci.* **42**: 2280–2288.

Held IM. 2000. *'Summer-school notes'*. GFDL: Princeton, NJ, USA. http://www.gfdl.noaa.gov/ih/.

Hoskins BJ, McIntyre ME, Robertson AW. 1985. On the use and significance of isentropic potential vorticity maps. *Q. J. R. Meteorol. Soc.* **111**: 877–946.

McIntyre ME. 2005. 'Rossby-wave propagation and shear instability'. In *GEFD summer school notes*. http://www.atm.damtp.cam.ac.uk/people/mem/gefd-supplem-material.html.

McIntyre ME, Shepherd TG. 1987. An exact local conservation theorem for finite-amplitude disturbances to non-parallel shear flows, with remarks on Hamiltonian structure and on Arnold's stability theorems. *J. Fluid. Mech.* **181**: 527–565.

Methven J, Hoskins BJ, Heifetz E, Bishop CH. 2005. The counter-propagating Rossby wave perspective on baroclinic instability. IV: Nonlinear life cycles. *Q. J. R. Meteorol. Soc.* **131**: 1425–1440.

Orr WMF. 1907. Stability or instability of the steady motions of a perfect liquid and of a viscous liquid. *Proc. R. Irish Acad.* **A27**: 9–138.

Peskin ME, Schroeder DV. 1995. *An Introduction to Quantum Field Theory*. Westview Press: USA.

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

Salmon R. 1988. Hamiltonian fluid mechanics. *Ann. Rev. Fluid Mech.* **20**: 225–256.

Salmon R. 1998. *Lectures on Geophysical Fluid Dynamics*. Oxford University Press: Oxford, UK.

Shepherd TG. 1990. Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Adv. Geophys.* **32**: 287–338.

Thorpe AJ, Bishop AH. 1995. Potential vorticity and the electrostatics analogy: Ertel-Rossby formulation. *Q. J. R. Meteorol. Soc.* **121**: 1477–1495.

DOI: 10.1002/qj