Rotating shallow water dynamics: Extra invariant and the formation of zonal jets

Alexander M. Balk, Francois van Heerden, and Peter B. Weichman

Phys. Rev. E **83**, 046320 — Published 28 April 2011

DOI: 10.1103/PhysRevE.83.046320

Rotating Shallow Water Dynamics: Extra Invariant and the Formation of Zonal Jets.

Alexander M. Balk,[1] Francois van Heerden,[2] and Peter B. Weichman[3]

[1] *Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, Utah 84112, USA*

[2] *South African Nuclear Energy Corporation, Reactor and Radiation Theory, Building P-1900, POBox 582, Pretoria 0001, South Africa*

[3] *BAE Systems, Advanced Information Technologies, 6 New England Executive Park, Burlington, Massachusetts 01803, USA*

We show that rotating shallow water dynamics possesses an approximate (adiabatic-type) positive quadratic invariant, which exists not only at mid-latitudes (where its analogue in the quasigeostrophic equation has been previously investigated), but near the equator as well (where the quasigeostrophic equation is inapplicable). Deriving the extra invariant, we find "small denominators" of two kinds: (1) due to the triad resonances (as in the case of the quasigeostrophic equation) and (2) due to the equatorial limit, when the Rossby radius of deformation becomes infinite. We show that the "small denominators" of *both* kinds can be canceled. The presence of the extra invariant can lead to the generation of zonal jets. We find that this tendency should be especially pronounced near the equator. Similar invariant occurs in magnetically confined fusion plasmas and can lead to the emergence of zonal flows.

Keywords: Rossby waves; Drift waves; Triad resonance; Shallow water; Conservation; Adiabatic invariants; Zonal jets; Zonal flows in plasmas

I. INTRODUCTION

To introduce the topic of the present paper, let us start with two different physical situations, which are known.

The first is the inverse cascade of energy in two-dimensional hydrodynamics [1, 2]. The inverse cascade is related to the presence of an additional (compared to the 3D hydrodynamics) positive quadratic invariant — *enstrophy*.

The second situation [3] is the appearance of longer waves in sea-wave turbulence: The length of typical waves on the sea surface is often much bigger than those generated directly by the wind and increases with time (a process known as *wave aging*). The sea waves generated directly by wind produce—via nonlinear interaction—longer waves, the latter produce even longer waves, and so on. This process is a manifestation of the inverse cascade [4]. As in the first situation, the inverse cascade is related to the presence of another (in addition to the energy) positive quadratic invariant, in this case the *wave action*. The later invariant holds because the gravity wave dispersion forbids 3-wave interactions, and so the main resonance interaction involves 4 waves and conserves the total wave action. The conservation of the wave action is similar to the conservation of the total number of molecules in rarefied gas (when the main interactions are binary collisions). However, unlike the number of molecules, the wave action is only an approximate invariant, whose conservation fails in higher order interactions (e.g., 5-wave interactions are possible, and they fail to conserve the wave action).

The present paper considers wave dynamics in rapidly rotating geophysical fluids. An integral part of this dynamics is the emergence of zonal jets [5], see also the collection of papers [6]; the stripes on Jupiter make a famous example [7]. Zonal jets are, as well, observed in the dynamics of magnetized plasmas (which is mathematically similar to geophysical fluid dynamics); they appear to act as transport barriers in tokamaks [8]. Plasma regimes with zonal jets become an integral part of modern controlled nuclear fusion installations, in particular, ITER (International Thermonuclear Experimental Reactor).

It is interesting to see if the emergence of zonal jets can be related to the existence of an additional invariant (similar to the two examples given above). This is the main motivation of the present paper.

It is believed that zonal jets emerge as a result of Rossby wave interactions, see [9, 10]. The nonlinear dynamics of Rossby waves in the beta plane is often modeled by the quasigeostrophic equation (see e.g. [11])

$$(\Delta\psi - \alpha^2\psi)_t + \beta\psi_x + \psi_y\Delta\psi_x - \psi_x\Delta\psi_y = 0 \tag{I.1}$$

for the stream function $\psi(x, y, t)$ of the horizontal fluid velocity $(u, v) = (\psi_y, -\psi_x)$. Here, α is the inverse Rossby radius of deformation, and β is the beta parameter characterizing the variation of the Coriolis force. The subscripts x, y, t denote partial derivatives, and Δ is the two-dimensional Laplacian.

It has been shown that the quasigeostrophic equation (I.1) indeed possesses an approximate (adiabatic-type) quadratic invariant, which requires the inverse cascade to transfer energy not just to large scales but specifically towards zonal flow [12–16].

However, in several physical situations (including Jupiter) zonal jets are well pronounced near the equator, while equation (I.1) is not applicable there. It is a major challenge to see if the approach to zonal jets based on the extra invariant works near the equator as well.

To account for the equatorial region, we consider the rotating shallow water dynamics in the beta-plane (equatorial or mid-latitudinal)

$$u_t + u\,u_x + v\,u_y - f(y)\,v = -g\,H_x\,, \tag{I.2a}$$

$$v_t + u\,v_x + v\,v_y + f(y)\,u = -g\,H_y\,, \tag{I.2b}$$

$$H_t + (H\,u)_x + (H\,v)_y = 0\,, \tag{I.2c}$$

e.g. [17, 18]. This system of equations describes the evolution of the horizontal fluid velocity (u, v) and the fluid height H (flat bottom is assumed). The function $f(y)$ is the Coriolis parameter, and g is the acceleration due to gravity.

Considering perturbation expansions (see below) for the system (I.2), we find "small denominators" not only related to the resonance triads [like in the case of (I.1)], but also related to the equatorial limit ($f \to 0$). We show that "small denominators" of both kinds can be canceled. The possibility of such cancelation is a remarkable property of the rotating shallow water system. We are unaware of any other system with similar attributes, even remotely.

Once we pass from a single equation (I.1) with constant coefficients to a system (I.2) of three equations with y-dependent coefficients, we also face two other problems:

1. There could be resonant interactions and energy transfer between the Rossby waves and the inertia-gravity waves.

2. The translational symmetry is broken.

The first problem is resolved due to a general fact [19] observed in a variety of rotating fluid systems: The coupling constant in the triad interaction between a slow mode and two fast modes vanishes in the equation for the slow mode (in our case, the slow is the Rossby mode, and the fast is the inertia-gravity mode). The second problem makes perturbation expansions significantly harder; in particular, the Rossby mode needs to be refined (Section IV A).

We show that in the limit of weak nonlinearity, the system (I.2) possesses an additional approximate (adiabatic-type) invariant, which is described in Sec. II. Before its formal derivation in Sec. IV, we demonstrate how the presence of this invariant makes the inverse cascade *anisotropic* and steers energy toward zonal flow (Sec. III A). Even more specific features, observed in some experiments, are in agreement with the proposed picture:

- Near the equator, the emergence of zonal jets is more pronounced than it is at mid-latitudes (Sec. III B).
- In the opposite limit (when typical length of waves excited by forcing is much greater than the Rossby radius of deformation), the extra invariant just says that the energy should transfer into the sector of wave vectors $\mathbf{k}$ with polar angles $> 60°$ (Sec. III C).

During the last half century, several ideas were proposed to explain the emergence of zonal jets through the dynamics of weakly nonlinear Rossby waves, e.g., random wave closures [9], wave kinetic equation [20], modulational instability [21], and almost resonant interactions [22]. Since these approaches consider the weakly nonlinear regime, we believe, they should be intimately related to the presence of the extra invariant. For the reasons discussed above, it is crucial to see that this invariant is present not only in the quasigeostrophic equation, but in the shallow water system as well.

To explain often powerful equatorial zonal jets, a deep approach was developed over the past 40 years; it derives the formation of zonal jets from the instability of the equatorial mixed Rossby-gravity waves (see [23, 24] and references therein). Since this approach is also based on small nonlinearity, the presence of the extra invariant remains relevant.

Some proposed scenarios are based on a *direct, nonlocal* transfer of energy from small scales to large scale zonal flows. Our theory is applicable in both cases—when the cascade is local or nonlocal; however, the relation between the invariant and zonal jets is more important when the cascade is local. Then the invariant provides crucial bookkeeping of the energy transfer towards zonal jets. In this case it is the weak turbulent inverse cascade, not the particular form of the initial instability, that controls the geometry of any emergent large-scale feature. Our theory then provides a continuously operating dynamical mechanism specific to the emergence of zonal jets, as observed in numerical simulations.

Concerning *Earth's* equatorial oceans and atmospere, we mention the following issue that merits future investigation. Very near the equator (within a couple of degrees), there is evidence [24] that a significant role is played by the dynamical terms neglected in the "traditional approximation" (in which the horizontal component of Earth's angular velocity is ignored, producing the standard beta plane approximation; see [25]). Our theory ignores such terms, so it is presently unknown if the extra invariant still exists under these conditions. The present work, at minimum, extends the domain of validity of the extra invariant from mid-latitudes to immediate vicinity of the equator, where the quasigeostrophic equation is already inapplicable, but the "traditional approximation" still remains valid.

We should also mention several mechanisms that connect zonal jets with strong nonlinearity, when nonlinear terms (in dynamic equations) are similar in magnitude to linear terms (see, e.g., [26–28]). Note that situations where linear terms are *negligible* compared to the nonlinear terms are unrelated to the formation of zonal jets because only the linear terms are anisotropic (the phenomenon of spontaneous emergence of anisotropy is not relevant here because the emerging jets are always observed to be parallel to the equator). As long as the linear terms are significant, even if not dominant, the extra invariant should still play an important role.

II. EXTRA INVARIANT

It is well known that (I.2) conserves the following quantities:

- ◇ The energy, see equation (IV.8) below.
- ◇ The infinite set of potential vorticity integrals

$$\int F\left[\frac{v_x - u_y + f(y)}{H}\right] H\, dxdy, \tag{II.1}$$

  where F is an arbitrary function of a single variable. These conservation laws are related to the advective conservation of potential vorticity [17].

In particular, for $F \equiv 1$ one obtains the total mass $\int H\, dxdy$; its conservation implies the existence of *time-independent* space-averaged fluid height $\bar{H}$, such that

$$\int (H - \bar{H})\, dxdy = 0. \tag{II.2}$$

◇ The x-momentum

$$\int u\, H\, dxdy, \tag{II.3}$$

which is related to translational symmetry in zonal direction.

We will see that the dynamics (I.2) *adiabatically* conserves three more quantities.

Before we describe them, let us eliminate some dimensional parameters by rewriting (I.2) in terms of the fractional relative height

$$h(x, y, t) = \frac{H - \bar{H}}{\bar{H}}; \tag{II.4}$$

and rescaling

$$ct \to t, \quad (u/c, v/c) \to (u, v), \quad f(y)/c \to f(y), \qquad \text{where} \quad c = \sqrt{g\bar{H}}. \tag{II.5}$$

Then the shallow water dynamics (I.2) takes the form

$$u_t + uu_x + vu_y - f(y)v = -h_x, \tag{II.6a}$$
$$v_t + uv_x + vv_y + f(y)u = -h_y, \tag{II.6b}$$
$$h_t + u_x + v_y + (hu)_x + (hv)_y = 0, \tag{II.6c}$$

where u, v, h are dimensionless, while x, y, t and $1/f$ have dimension of length.

Consider the *linearized perturbational potential vorticity* (see [29])

$$\mathcal{Q} = v_x - u_y - f(y)\, h\,, \tag{II.7}$$

which, according to (II.6), obeys the equation of motion

$$\mathcal{Q}_t \ + \ (u\, \mathcal{Q})_x \ + \ (v\, \mathcal{Q})_y \ = \ -\beta\, v\, (1 + h), \qquad \text{where} \quad \beta = f'(y). \tag{II.8}$$

Since our goal is to describe the energy transfer in Fourier space, we consider the Fourier transform $\mathcal{Q}_{\mathbf{k}}$ of the field $\mathcal{Q}$:

$$\mathcal{Q}(x, y, t) = \int \mathcal{Q}_{\mathbf{k}}(t)\, e^{i(px+qy)}\, dpdq \quad [\mathbf{k} = (p, q),\ k^2 = p^2 + q^2]. \tag{II.9}$$

We will show that the shallow water dynamics (II.6) adiabatically conserves three quantities of the form

$$I = \frac{1}{2} \int X_{\mathbf{k}}\, \mathcal{Q}_{\mathbf{k}}\, \mathcal{Q}_{-\mathbf{k}}\, dp\, dq. \tag{II.10}$$

The first is the energy of the Rossby waves (the inertia-gravity component is excluded); it has

$$X_{\mathbf{k}}^{\text{energy}} = \frac{1}{f^2 + k^2}. \tag{II.11a}$$

The second is the enstrophy of the Rossby component; it has

$$X_{\mathbf{k}}^{\text{enstrophy}} = 1. \tag{II.11b}$$

In addition to these two, there is an extra invariant with

$$X_{\mathbf{k}}^{\text{extra}} = \frac{1}{f^5 p} \left[\arctan \frac{f(q + p\sqrt{3})}{k^2} \ - \ \arctan \frac{f(q - p\sqrt{3})}{k^2} \ - \ \frac{2\sqrt{3} f p}{f^2 + k^2} \right]. \tag{II.11c}$$

This expression is nonsingular as $f \to 0$

$$X_{\mathbf{k}}^{\text{extra}} \simeq 8\sqrt{3}\,p^2\,\frac{p^2+5q^2}{5k^{10}} \;-\; 8\sqrt{3}\,p^2\,\frac{5p^4+42p^2q^2+21q^4}{7k^{14}}\,f^2 \;+\; O(f^4). \tag{II.12}$$

The notion of *adiabatic conservation* here is similar to that in the theory of dynamical systems [30]: The adiabatic invariants are conserved *approximately* over *long time*. However, here adiabatic conservation is due not to the slowness of parameter change in time, but to the slowness of *spatial* change and to the *smallness of the wave amplitudes*. This adiabatic conservation is due to the presence of *two* small parameters (see Section IV B): First is the strength of nonlinearity, compared to the beta effect [see (IV.10)], and second is the degree of spatial inhomogeneity, i.e., the slowness of the dependence of the Coriolis force on the latitudinal coordinate y, compared to the length scale L of field variation and the Rossby radius of deformation [see (IV.11)].

In the present paper, we derive adiabatic conservation of the above integrals in the lowest possible non-trivial—leading—orders. We aim here only to establish the fact of adiabatic conservation at minimal accuracy (although the actual conservation accuracy might be higher, or the conservation time interval might be longer). Our derivation is formal asymptotic, but we take special care that no secular terms appear.

In the present paper we consider the simplest possible case, making two simplifications:

1. We use the beta-plane approximation (disregarding complications of spherical geometry).

2. We assume the fields (u, v, h) vanish at infinity, i.e., at the periphery of the beta plane. The same assumption is made for the quasigeostrophic equation (I.1) when considering its invariants. Without this assumption, we need to account for the boundary terms. These can be dealt with, but their account leads to heavy mathematical calculations, which will not be presented here.

The central result of our derivation—which allows us to establish the extra invariant near the equator—is the possibility to cancel "small denominators" at $f \to 0$; see equations (IV.23), (IV.24).

We derive the extra invariant in Sec. IV, but first we demonstrate the connection between the invariant and zonal jets.

III. THE EMERGENCE OF ZONAL JETS

A. Why the extra invariant implies the emergence of zonal jets

The approximate conservation of the energy and enstrophy, contained in the Rossby component (see Sec. II) implies the inverse cascade of Rossby wave energy. At the same time (as we will see now), the presence of the extra invariant ensures the anisotropy of the inverse cascade: The energy is transported not just towards the origin, but specifically to the region of the $\mathbf{k} = (p, q)$-plane around the q-axis ($|p| \ll |q|$), which corresponds to zonal jets.

Indeed, the extra invariant can be written in the form

$$I = \int \phi_{\mathbf{k}}\, \varepsilon_{\mathbf{k}}\, dp\, dq\,, \tag{III.1}$$

where $\varepsilon_{\mathbf{k}}$ is the Rossby wave energy spectrum, and $\phi_{\mathbf{k}}$ is the ratio of the extra invariant spectral density to the energy spectral density

$$\phi_{\mathbf{k}} = \frac{f^2+k^2}{f^5\,p}\left[\arctan\frac{f(q+p\sqrt{3})}{k^2} \;-\; \arctan\frac{f(q-p\sqrt{3})}{k^2} \;-\; \frac{2\sqrt{3}\,fp}{f^2+k^2}\right], \tag{III.2}$$

see (II.11c) and (II.11a). Figure 1 shows a contour plot of the values of the ratio $\phi_{\mathbf{k}}$ (on a logarithmic scale) vs. $\mathbf{k}$. We pose the following question: Is it possible for the energy from the region A in Fig. 1 be transferred (via the inverse cascade) into the region B? The value of the ratio ϕ in the region B is about 7×10^3 times greater than its value in the region A. So, if the transfer $A \to B$ did occur, the value I of the extra invariant (III.1) would significantly increase. The only way for the inverse cascade to transfer the energy towards the origin would be to transport the energy (on average) along the level lines of the function $\phi_{\mathbf{k}}$. Thus, the dynamics must display anisotropic "Bose condensation": Spatial anisotropy, which is only weakly broken in small scale dynamics, becomes ever more strongly broken on large scales.

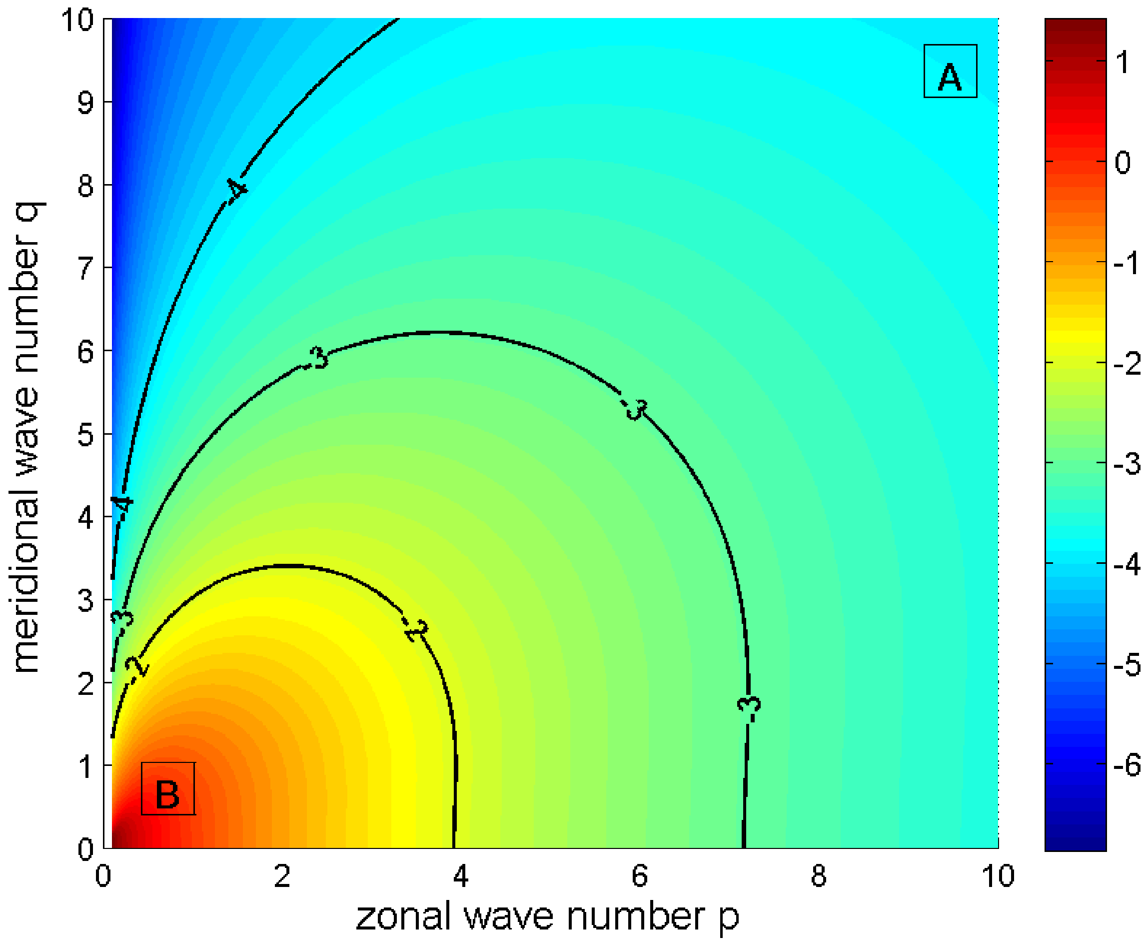

FIG. 1. (Color online) Contour plot of $\log_{10}(\phi_{\mathbf{k}})$; the ratio $\phi_{\mathbf{k}}$ measures how much extra invariant is carried per unit energy by an excitation with wave vector $\mathbf{k} = (p, q)$; see (III.1)–(III.2). The plot spans the range $0.1 \leq p \leq 10,\ 0 \leq q \leq 10$, while $f = 1$. The values of the ratio $\phi_{\mathbf{k}}$ at the centers of boxes A and B differ by roughly a factor of 7 000. Therefore, the only way for the energy to transfer towards the origin (via the inverse cascade—be it local or nonlocal) is for it to 'squeeze' around the q-axis.

B. Why zonal jets should be more clearly observed near the equator

The difference between the values of ϕ in the regions A and B increases as $|f|$ decreases ($f \approx 0$ near the equator). Figure 2 shows the values of the ratio $\phi_{\mathbf{k}}$ when $f = 0.03$. The value of ϕ in the center of region B is now about 3×10^4 times greater than its value in region A. Therefore, close to equator, when $f \approx 0$, the inverse cascade is forced to transfer energy even tighter towards the q-axis.

The more pronounced formation of zonal jets near the equator can be seen quantitatively from Fig. 3, which shows the dependence (on a log-log scale) of the ratio ϕ vs. the wave number k at fixed polar angles θ (due to symmetries, we need consider polar angles only in the range $0^\circ \leq \theta \leq 90^\circ$). The curves shown in Fig. 3 are steeper for large k/f than for small k/f: For large k/f they vary as k^{-4}, while for small k/f they vary as k^{-1} if $\theta \leq 60^\circ$ and as k^0 if $\theta > 60^\circ$. Therefore, during the inverse cascade, the ratio ϕ increases more significantly with decreasing k if $f \approx 0$ (near the equator).

For example, if the energy originated in the region $k/f > 20$, then the inverse cascade must transfer this energy (on average) into the sector $89.9^\circ < \theta < 90^\circ$. [Indeed, $f^4\phi$ at ($k/f > 20$ and all θ) is less than $f^4\phi$ at ($\theta = 89.9^\circ$ and $k/f \to 0$).] Such 'tight squeezing' of energy around the q-axis hardly can be accounted for by the relative decrease of the nonlinearity as $f \to 0$ (which might be expected in some situations).

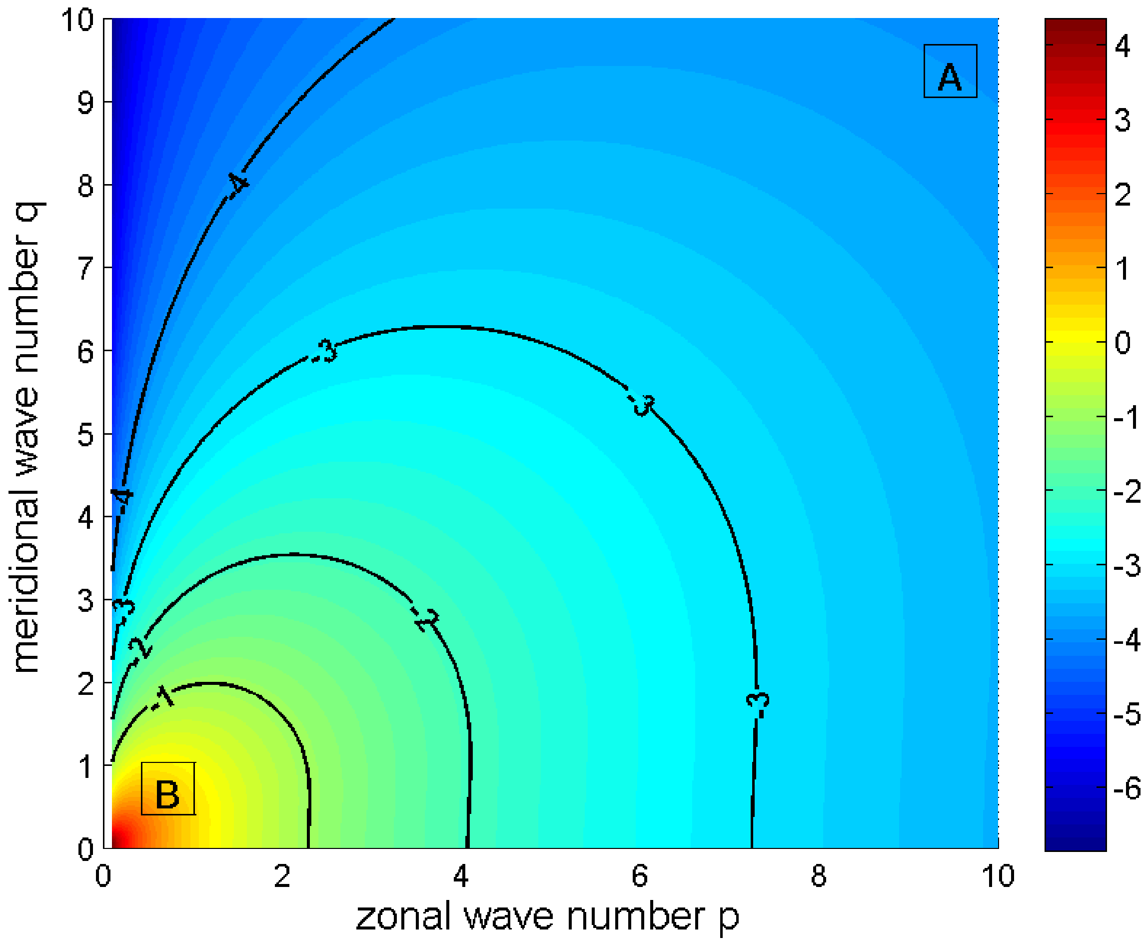

FIG. 2. (Color online) Same as Fig. 1, but close to the equator, i.e., for small f (for this particular figure, $f = 0.03$). Comparison of Fig. 1 and Fig. 2 demonstrates that the energy transfer towards zonal jets should be more pronounced near the equator than at mid-latitudes. The values of the ratio $\phi_{\mathbf{k}}$ at the centers of boxes A and B now differ by roughly a factor of 30 000.

C. Long-wavelength limit — polar angle $60°$

Now let us consider the opposite limit where k/f is small. According to Fig. 3, the inverse cascade can now transfer energy anywhere into the sector

$$60° < \theta < 90°. \tag{III.3}$$

This is exactly the sector that was found [31] on the basis of satellite altimeter observations of the spectra of very long mid-latitude Rossby waves (with periods of several years). The sector (III.3) is clearly visible in the contour plot of $\log_{10}(\phi_{\mathbf{k}})$ for small k/f—see Fig. 4; it shows the values of the ratio $\phi_{\mathbf{k}}$ when $f = 30$. The magnitude of $\phi_{\mathbf{k}}$ drops sharply when the polar angle θ increases beyond $60°$; it is clear that, following any level curve beginning at larger k, one may approach the origin only through the sector (III.3).

We see that if the energy is generated at large scales (much greater than the Rossby radius of deformation) then the balance argument, based on the extra invariant, does not require the inverse cascade to accumulate energy in zonal flows. This conclusion agrees with the investigation [32], which reported "suppression of the Rhines effect" for large f.

To conclude this Section, we note that the existence of the extra invariant and the balance argument (described in Secs. III A, III B, III C) holds for a wide class of wave systems *with Rossby dispersion law*. When the *nonlinearity*

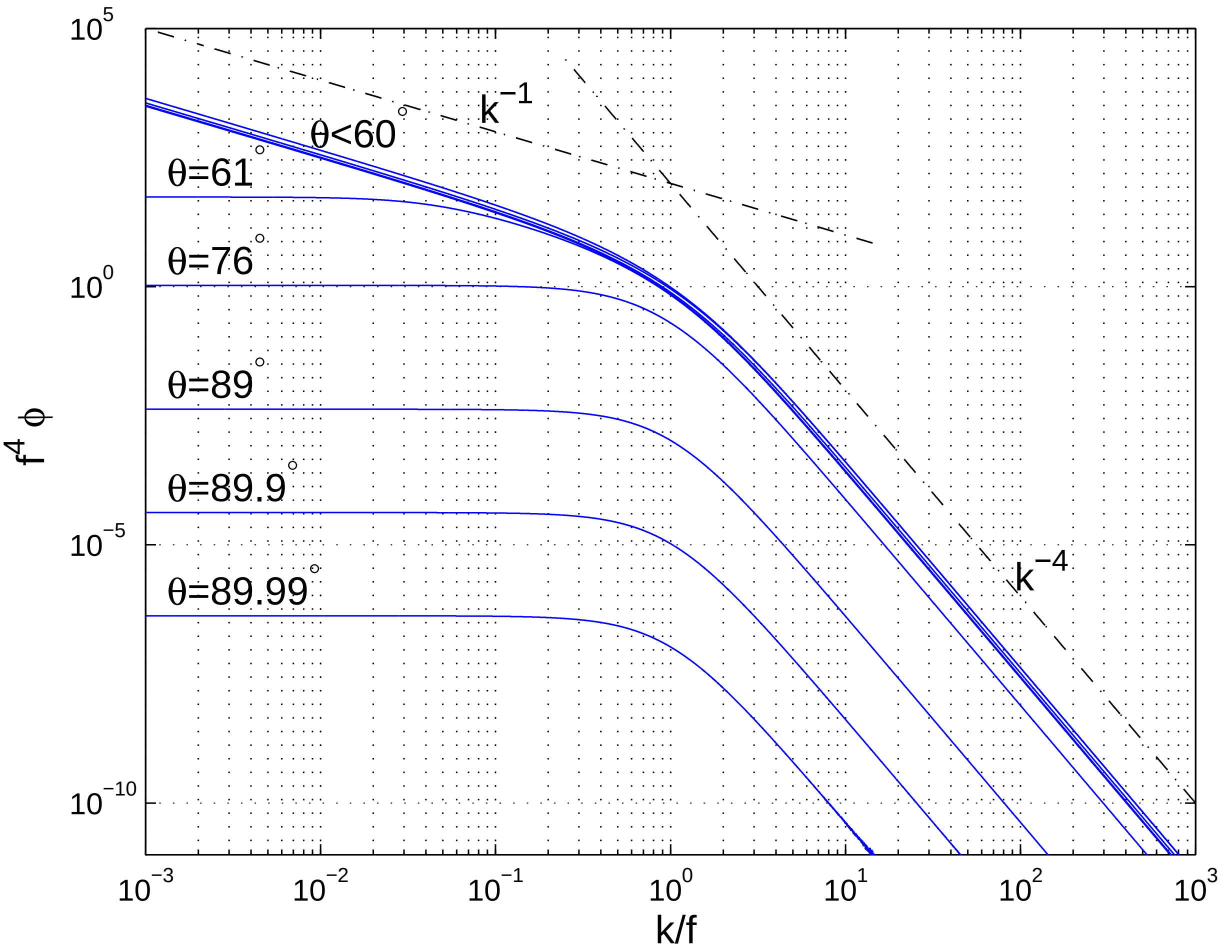

FIG. 3. (Color online) The quantity $f^4\phi$ is a function of k/f while the polar angle θ is held fixed. The curves correspond to the ten rays with $\theta = 0^\circ, 15^\circ, 30^\circ, 45^\circ, 60^\circ, 61^\circ, 76^\circ, 89^\circ, 89.9^\circ, 89.99^\circ$. Curves with $\theta \leq 60^\circ$ (including $\theta = 60^\circ$) are almost indistinguishable and marked by one label $\theta < 60^\circ$. If $\theta > 60^\circ$, the curves become horizontal when $k \ll f$.

is taken into account, then for some special forcing, the energy can still concentrate in zonal flows, even with large f, see [33]. In that paper it was also found that, in the *short wave* case (or *near the equator*, Sec. III B), specially arranged forcing can accelerate the formation of zonal jets.

The presented balance argument for the emergence of zonal jets has the appeal that it is based on a (previously unnoticed) conservation law. However, this argument crucially relies on the assumption of weak nonlinearity. Whether the nonlinearity is weak depends both on the forcing strength and on the location of sources and sinks in Fourier space. Physical examples often show that the turbulence is weak in the large-scale part of the inertial range, in spite of the fact that the energy spectrum becomes infinite when $\mathbf{k} \to 0$; e.g., consider the sea wave turbulence [3]. In the case of geostrophic turbulence, the ratio of the magnitude of nonlinear terms to the magnitude of linear terms in the quasigeostrophic equation is the Rhines number $\epsilon = A/(\beta L^2)$. [For simplicity, we consider here the short-wave limit, when the Rossby radius of deformation is effectively infinite; more refined estimates will be given in Sec. IV B.] During the inverse cascade the length scale L increases, while the velocity scale A stays roughly constant (determined by the energy), and so, $\epsilon \to 0$.

IV. DERIVATION OF THE ADIABATIC INVARIANTS

In this Section, we demonstrate approximate conservation of the quadratic invariants (II.10) with the kernels (II.11).

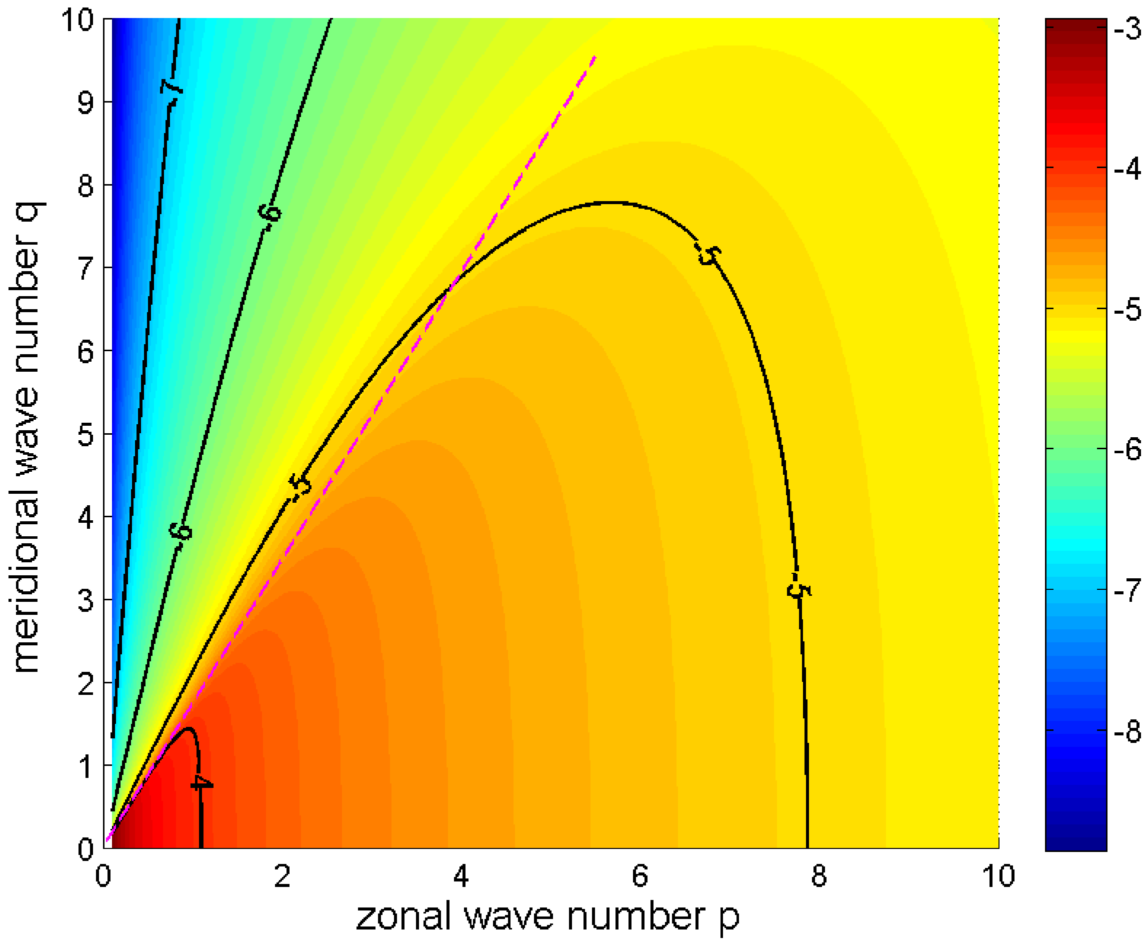

FIG. 4. (Color on line) Contour plot of $\log_{10}(\phi_{\mathbf{k}})$ for large f (for this particular figure, $f = 30$). When $k \ll f$, the extra invariant forces energy to accumulate in the sector $60^\circ < \theta < 90^\circ$ (cf. Fig. 1). The dashed ray marks polar angle 60°.

A. Refining the Rossby mode

Dropping the nonlinear terms in (II.6) leads to the linearized system

$$u_t = -fv - h_x, \tag{IV.1a}$$
$$v_t = fu - h_y, \tag{IV.1b}$$
$$h_t = -u_x - v_y, \tag{IV.1c}$$

while the linearized perturbational potential vorticity (II.7) obeys the equation

$$\mathcal{Q}_t = -\beta\, v. \tag{IV.2}$$

Because of the y-dependence of coefficients in (IV.1), we need to refine the Rossby mode. Let us add to $\mathcal{Q}$ a correction $\mathcal{R}$ (to be determined below) that is of higher order with respect to the parameter β, to obtain a new field

$$s = \mathcal{Q} + \mathcal{R} \tag{IV.3}$$

such that in the linear approximation (IV.1) the derivative s_t will be determined by s alone (not by u, v, h taken separately or in any other combination, besides s). Calculations show that we need to construct $\mathcal{R}$ such that

$$(f^2 - \Delta)\mathcal{R} = \beta(fu + h_y). \tag{IV.4}$$

Indeed, if $\mathcal{R}$ is determined by (IV.4) then

$$(f^2-\Delta)s_t = (f^2-\Delta)(\mathcal{Q}_t+\mathcal{R}_t) = -(f^2-\Delta)(\beta v) + \beta(fu+h_y)_t, \tag{IV.5}$$

and, according to the dynamics (IV.1),

$$(f^2-\Delta)s_t = \beta s_x + 2\beta' v_y + \beta'' v. \tag{IV.6}$$

Here the right hand side is $\beta s_x + O(\beta^2)$, so that the non-s terms are indeed pushed to higher order. Neglecting the higher order terms, we see that the s-mode has the Rossby wave dispersion

$$\Omega_{\mathbf{k}} = -\frac{\beta p}{f^2+k^2}\,. \tag{IV.7}$$

B. Small parameters

The energy of the system (II.6) is

$$E = \frac{1}{2}\,\bar{H}c^2 \int [(u^2+v^2)(1+h) \;+\; h^2]\, dx\, dy. \tag{IV.8}$$

In the weakly nonlinear limit, the integrand of the energy (IV.8) reduces to $u^2+v^2+h^2$, and we assume u, v, h to have the same magnitude A.

As mentioned above, we exploit *two small parameters*: First, the field magnitude A should be "small", compared to the beta-effect [see (IV.10)]. Second, the Coriolis parameter $f(y)$ must be a "slow" function of y, so that $\beta(y) \equiv f'(y)$ is "small" [see (IV.11)]. We first define the two *non-dimensional* small parameters when the field variations are characterized by a single length scale L, being the same in x and y directions.

The magnitude of the linearized potential vorticity is given by

$$\mathcal{Q} \propto \frac{A}{l}, \quad \text{where} \quad \frac{1}{l} = \frac{1}{L} + f. \tag{IV.9}$$

To measure the degree of nonlinearity we consider the ratio ϵ of the magnitude of the nonlinear convective terms in (II.8) to the magnitude of the linear term:

$$\epsilon = \frac{A}{\beta l L}. \tag{IV.10}$$

Near the equator (where $f \ll L^{-1}$), the nonlinearity degree becomes the Rhines number: $\epsilon = A/\beta L^2$.

Small inhomogeneity means that f changes little over the length scale L. The change is $\Delta f \approx \beta L$. Away from the equator, Δf should be compared to f; near the equator it should be compared to L^{-1}. Thus, to quantify the degree of spatial inhomogeneity we use the parameter

$$b = \beta l L, \quad \text{and so,} \quad A = \epsilon b. \tag{IV.11}$$

In more general situations the length scales in x and y directions can be different (which is especially relevant when considering zonal jets). Moreover, the dynamics can be characterized by a wide range of length scales, and they can change in time (they can easily change by an order of magnitude during the inverse cascade). To account for different situations, we will just keep track of powers of A and β ($A \to 0, \beta \to 0$). To maintain the condition that the field be small in comparison to the beta-effect, we will assume the existence of a small parameter ϵ, such that $A \propto \epsilon\beta$. [In general, ϵ and b will have a more complex dependence on physical scales than (IV.10) and (IV.11).]

When there is a single length scale L, then the $\mathcal{R}$-correction in (IV.3) is $O(AL^{-1}b)$ and is proportional to β. However, in a general situation, with many length scales, we can only guarantee that $\mathcal{R} \propto \sqrt{\beta}$. Indeed, for states almost constant in the zonal direction ($\partial/\partial x = 0$) and near the equator ($f = \beta y$) equation (IV.4) becomes

$$\beta^2 y^2 \mathcal{R} - \frac{\partial^2 \mathcal{R}}{\partial y^2} = \beta(\beta y u + \frac{\partial h}{\partial y}), \tag{IV.12}$$

which is reduced by rescaling $\tilde{y} = \sqrt{\beta} y$ to the form

$$\tilde{y}^2 \mathcal{R} - \frac{\partial^2 \mathcal{R}}{\partial \tilde{y}^2} = \sqrt{\beta}(\tilde{y} u + \frac{\partial h}{\partial \tilde{y}}), \tag{IV.13}$$

exhibiting explicitly the $\sqrt{\beta}$ scale of $\mathcal{R}$.

Since the difference between fields $\mathcal{Q}$ and s is small (proportional to $\sqrt{\beta}$), we replace $\mathcal{Q}$ in (II.10) by s:

$$I \approx I^\star = \frac{1}{2} \int X_{\mathbf{k}}\, s_{\mathbf{k}}\, s_{-\mathbf{k}}\, dp\, dq. \tag{IV.14}$$

C. Supplementing the quadratic extra invariant with cubic terms

Our central claim is that the increment $\Delta I^\star \equiv I^\star(t) - I^\star(0)$ remains small over long times t. However, this does not necessarily mean that $\dot{I}^\star$ is small: $I^\star(t)$ can oscillate in time, similar to the behavior of adiabatic invariants in the theory of dynamical systems. So, we use the approach [34] and supplement the quadratic integral (IV.14) with a cubic part

$$I^{\text{suppl}} = I^\star \; + \; I^{\text{cubic}}; \tag{IV.15}$$

then require $\dot{I}^{\text{suppl}}$ to vanish to leading order. The general form of the cubic correction is

$$\begin{aligned} I^{\text{cubic}} &= \frac{1}{6}\int \left[Y^{uuu}_{123}\, u_1 u_2 u_3 \; + \; Y^{vvv}_{123}\, v_1 v_2 v_3 \; + \; Y^{hhh}_{123}\, h_1 h_2 h_3\right] d_{123} \\ &+ \frac{1}{2}\int \left[Y^{uuv}_{123}\, u_1 u_2 v_3 \; + \; Y^{uuh}_{123}\, u_1 u_2 h_3 \; + \; Y^{vvu}_{123}\, v_1 v_2 u_3 \; + \; Y^{vvh}_{123}\, v_1 v_2 h_3 \right. \\ &\qquad \left. + \; Y^{hhu}_{123}\, h_1 h_2 u_3 \; + \; Y^{hhv}_{123}\, h_1 h_2 v_3\right] d_{123} \; + \; \int Y^{uvh}_{123}\, u_1 v_2 h_3 \, d_{123} \end{aligned} \tag{IV.16}$$

with 10 kernels $Y^{uuu}, Y^{vvv}, \ldots$. Here and throughout the rest of this paper a subscript j stands for the wave vector $\mathbf{k}_j = (p_j, q_j)$ $(j = 1, 2, 3)$; e.g. $u_1 = u_{\mathbf{k}_1}$, likewise, $Y_{123} = Y(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ for any kernel Y, and $\delta_{123} = \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$, $d_{123} = d\mathbf{k}_1\, d\mathbf{k}_2\, d\mathbf{k}_3$. In addition, a subscript $-j$ will denote $-\mathbf{k}_j$, in particular, $Y_{-123} = Y(-\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$.

The form of the shallow water system allows us to consider the following more simple form of the cubic correction

$$I^{\text{cubic}} = \frac{1}{2}\int s_1\, s_2\, \left[M_{123}\, u_3 \; + \; N_{123}\, v_3 \; + \; T_{123}\, h_3\right] d_{123} + \frac{1}{6}\int Y_{123}\, s_1 s_2 s_3\, d_{123} \tag{IV.17}$$

with only 4 kernels M, N, T, and Y instead of the ten kernels in (IV.16). The general form (IV.16) and the simplified form (IV.17) lead to the same final result. A much longer calculation demonstrates that the kernels in (IV.16) must be related in such a way that the terms may be collected in the form (IV.17).

When calculating $\dot{I}^{\text{suppl}}$, we will have contributions of different nonlinearity orders

$$\dot{I}^{\text{suppl}} = \dot{I}^\star \quad \text{due to the } \textit{linear} \text{ terms in the equations} \tag{IV.18a}$$

$$+\; \dot{I}^\star \quad \text{due to the } \textit{quadratic} \text{ terms in the equations} \tag{IV.18b}$$

$$+\; \dot{I}^{\text{cubic}} \quad \text{due to the } \textit{linear} \text{ terms in the equations} \tag{IV.18c}$$

$$+\; \dot{I}^{\text{cubic}} \quad \text{due to the } \textit{quadratic} \text{ terms in the equations} \tag{IV.18d}$$

We will see that the first contribution (IV.18a) vanishes automatically. Our goal is to show that it is possible to find the cubic correction I^{cubic}—with non-singular (uniformly bounded) kernels—such that the next two contributions (IV.18b) and (IV.18c) exactly cancel each other, implying that, indeed, $\dot{I}^{\text{suppl}}$ is determined by only higher order terms (IV.18d). Formally, one can always achieve such cancellation for any wave system, but the corresponding kernels will generally be singular. The possibility to escape these singularities takes place only for *very few* systems [34, 35]. Significantly, our results demonstrate that the rotating shallow water system is among them.

D. The time derivative $\dot{I}^{\text{suppl}}$

We can always assume the obvious symmetries:

$$X_{\mathbf{k}} = X_{-\mathbf{k}}, \;\; M_{123} = M_{213}, \;\; N_{123} = N_{213}, \;\; T_{123} = T_{213}, \;\; Y_{123} = Y_{213} = Y_{321}. \tag{IV.19}$$

According to the rotating shallow water dynamics (II.6), (II.8), along with the definition (IV.3), we have to leading orders

$$\begin{aligned} \dot{I}^{\text{suppl}} &= \int X_1\, s_{-1} \left[-i\Omega_1\, s_1 + \int (-ip_1 u_3 - iq_1 v_3)\, s_2\, \delta_{-123}\, d_{23}\right] d_1 \\ &+ \frac{1}{2}\int s_1\, s_2\, \left[M_{123}(f v_3 - ip_3 h_3) \; + \; N_{123}(-f u_3 - iq_3 h_3) \right. \\ &\qquad\qquad \left. + \; T_{123}(-ip_3 u_3 - iq_3 v_3)\right] d_{123} \\ &+ \frac{1}{2}\int Y_{123}\, s_1\, s_2\, (-i\Omega_3\, s_3)\, d_{123} \end{aligned} \tag{IV.20}$$

[in accordance with our notations (introduced in Sec. IV C), $\delta_{-123} = \delta(-\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$]. The equation (IV.20) explicitly displays the contributions summarized in (IV.18abc). We will determine the kernels X, M, N, T, Y from the requirement that the right hand side of this expression vanish, and that they be nonsingular.

First, the integral $\int X_1\, \Omega_1\, s_{-1} s_1\, d_1$ vanishes automatically since $X_{\mathbf{k}}$ is even, and $\Omega_{\mathbf{k}}$ is odd.

Using (II.7) and (IV.3), we substitute

$$u_3 = \frac{ip_3 v_3 - f h_3 - s_3}{iq_3} \tag{IV.21}$$

into (IV.20) and collect terms into three groups: those containing (1) ssv, (2) ssh, and (3) sss:

$$\begin{aligned}
\dot{I}^{\rm suppl} &= \frac{1}{2} \int d_{123}\, s_1\, s_2 \\
&\times \left\{ \frac{v_3}{q_3} \left[f q_3 M_{123} - f p_3 N_{123} - k_3^2 T_{123} + i(p_3 p_1 X_1 + p_3 p_2 X_2 + q_3 q_1 X_1 + q_3 q_2 X_2)\delta_{123} \right] \right. \\
&\quad + \frac{h_3}{q_3} \left[i p_3 q_3 M_{123} p_3 q_3 + i(f^2 + q_3^2) N_{123} - f p_3 T_{123} + f(p_1 X_1 + p_2 X_2)\delta_{123} \right] \\
&\quad \left. + \frac{s_3}{q_3} \left[i f N_{123} - p_3 T_{123} + (p_1 X_1 + p_2 X_2)\delta_{123} \right] \right\} \\
&\qquad - \frac{i}{6} \int Y_{123}\, s_1\, s_2\, s_3\, (\Omega_1 + \Omega_2 + \Omega_3)\, d_{123}
\end{aligned} \tag{IV.22}$$

E. Canceling "small denominators" which are due to the equatorial limit ($f \ll k$)

Equating to zero the coefficients of v_3 and h_3 produces a system of two linear algebraic equations, which we solve for the kernels M and N:

$$\begin{aligned}
M_{123} &= \frac{iq_3 T_{123}}{f} - \\
&\quad \frac{iq_3 \left[(p_1 p_3 + q_1 q_3) X_1 + (p_2 p_3 + q_2 q_3) X_2 \right] + i f^2 (q_1 X_1 + q_2 X_2)}{f(f^2 + k_3^2)}\, \delta_{123},
\end{aligned} \tag{IV.23a}$$

$$\begin{aligned}
N_{123} &= -\frac{ip_3 T_{123}}{f} + \\
&\quad \frac{ip_3 \left[(p_1 p_3 + q_1 q_3) X_1 + (p_2 p_3 + q_2 q_3) X_2 \right] + i f^2 (p_1 X_1 + p_2 X_2)}{f(f^2 + k_3^2)}\, \delta_{123}
\end{aligned} \tag{IV.23b}$$

These expressions have an apparent singularity when $f \to 0$, which would invalidate our perturbational expansion (e.g., making the cubic correction larger than the main quadratic part). However, it is possible to choose the kernel T in such a way as to eliminate the singularities in M and N. *Both* expressions (IV.23) become non-singular as $f \to 0$ if we take

$$T_{123} = \frac{p_3(p_1 X_1 + p_2 X_2) + q_3(q_1 X_1 + q_2 X_2)}{f^2 + k_3^2}\, \delta_{123}. \tag{IV.24a}$$

This cancels all terms proportional to $1/f$, and produces

$$M_{123} = -i \frac{f(q_1 X_1 + q_2 X_2)}{f^2 + k_3^2}\, \delta_{123}, \tag{IV.24b}$$

$$N_{123} = i \frac{f(p_1 X_1 + p_2 X_2)}{f^2 + k_3^2}\, \delta_{123}. \tag{IV.24c}$$

The denominators in (IV.24) still appear to be singular when $f \to 0$, $k_3 \to 0$ simultaneously. However, due to the presence of the delta functions $\delta_{123} \equiv \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$, the condition $k_3 \to 0$ implies $\mathbf{k}_1 + \mathbf{k}_2 \to 0$, and so, the expressions $p_1 X_1 + p_2 X_2$ and $q_1 X_1 + q_2 X_2$ in the numerators (IV.24) are linear in p_3, q_3 when $k_3 \to 0$. Therefore, the numerators in (IV.24) are quadratic in f, p_3, q_3, and the expressions (IV.24) are bounded.

Substituting (IV.24) into (IV.22), we find

$$\dot{I}^{\rm suppl} = \frac{1}{2}\int \left[\frac{p_3q_1 - p_1q_3}{f^2+k_3^2}X_1 + \frac{p_3q_2 - p_2q_3}{f^2+k_3^2}X_2\right] s_1\, s_2\, s_3\, \delta_{123}\, d_{123}$$
$$- \frac{i}{6}\int Y_{123}\, s_1\, s_2\, s_3\, (\Omega_1 + \Omega_2 + \Omega_3)\, d_{123} \qquad \text{(IV.25)}$$

F. Canceling the triad resonance "small denominators"

Note the obvious identity

$$\begin{vmatrix} p_1 & p_2 \\ q_1 & q_2 \end{vmatrix} = \begin{vmatrix} p_2 & p_3 \\ q_2 & q_3 \end{vmatrix} = \begin{vmatrix} p_3 & p_1 \\ q_3 & q_1 \end{vmatrix} \quad \text{when } \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0 \qquad \text{(IV.26)}$$

(to see, e.g., the first equality, substitute $p_1 = -p_2 - p_3,\ q_1 = -q_2 - q_3$ from the last equation). Because of (IV.26), equation (IV.25) reduces to

$$\dot{I}^{\rm suppl} = \frac{1}{2}\int \frac{p_3q_1 - p_1q_3}{f^2+k_3^2}(X_1 - X_2)\, s_1\, s_2\, s_3\, \delta_{123}\, d_{123}$$
$$- \frac{i}{6}\int Y_{123}\, s_1\, s_2\, s_3\, (\Omega_1 + \Omega_2 + \Omega_3)\, d_{123}. \qquad \text{(IV.27)}$$

Symmetrizing the first term on the right hand side (over all permutations of the indices 1,2,3), and using again the identity (IV.26), we see that (IV.27) vanishes if

$$Y_{123} = \frac{p_1q_2 - p_2q_1}{i(\Omega_1 + \Omega_2 + \Omega_3)}\, \delta_{123} \left[\frac{X_2 - X_1}{f^2+k_3^2} + \frac{X_3 - X_2}{f^2+k_1^2} + \frac{X_1 - X_3}{f^2+k_2^2}\right]. \qquad \text{(IV.28)}$$

It is apparent that this expression is singular at the points $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ satisfying the resonance relations

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0, \qquad \text{(IV.29a)}$$
$$\Omega_{\mathbf{k}_1} + \Omega_{\mathbf{k}_2} + \Omega_{\mathbf{k}_3} = 0, \qquad \text{(IV.29b)}$$

unless the expression in square brackets vanishes at these points. We will see now that the latter is indeed the case.

On the resonance manifold (IV.29), the bracketed expression in (IV.28) may be put in the form

$$[\ldots] = \frac{p_1\Omega_2 - p_2\Omega_1}{\beta\, p_1p_2p_3}\,(p_1X_1 + p_2X_2 + p_3X_3)\,. \qquad \text{(IV.30)}$$

To obtain this, the dispersion relation (IV.7) must be used along with the identities

$$\begin{vmatrix} p_1 & p_2 \\ \Omega_1 & \Omega_2 \end{vmatrix} = \begin{vmatrix} p_2 & p_3 \\ \Omega_2 & \Omega_3 \end{vmatrix} = \begin{vmatrix} p_3 & p_1 \\ \Omega_3 & \Omega_1 \end{vmatrix} \quad \text{when} \begin{cases} p_1 + p_2 + p_3 = 0 \\ \Omega_1 + \Omega_2 + \Omega_3 = 0, \end{cases} \qquad \text{(IV.31)}$$

which are similar to (IV.26). Thus, to obtain a non-singular form for Y_{123}, we have to require that the function $pX_\mathbf{k}$ be conserved in the triad resonance interactions, i.e., that the equation

$$p_1X_1 + p_2X_2 + p_3X_3 = 0 \qquad \text{(IV.32)}$$

must hold in all points $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ of the resonance manifold (IV.29).

G. Kernel X

The requirement (IV.32) is satisfied for the following five functions:

$$pX_\mathbf{k} = \Omega_\mathbf{k}, \qquad \text{(IV.33a)}$$
$$pX_\mathbf{k} = p, \qquad \text{(IV.33b)}$$
$$pX_\mathbf{k} = q, \qquad \text{(IV.33c)}$$
$$pX_\mathbf{k} = \xi_\mathbf{k} \stackrel{\rm def}{=} \frac{1}{2}\ln\left[f^2(q + p\sqrt{3})^2 + k^4\right] - \frac{1}{2}\ln\left[f^2(q - p\sqrt{3})^2 + k^4\right] \qquad \text{(IV.33d)}$$
$$pX_\mathbf{k} = \eta_\mathbf{k} \stackrel{\rm def}{=} \arctan\left[f(q + p\sqrt{3})/k^2\right] - \arctan\left[f(q - p\sqrt{3})/k^2\right]. \qquad \text{(IV.33e)}$$

For the functions $\xi_{\mathbf{k}}$ and $\eta_{\mathbf{k}}$ see [13]; their physical meaning remains unclear, let alone their possible relation to some continuous symmetries.

For the function (IV.33a), the integral I is the *energy* of the Rossby component.

The function (IV.33b) corresponds to the *enstrophy*. More precisely, in this case the integral I is the zonal (East-West) momentum, which is a linear combination of the energy and enstrophy.

The function (IV.33c) corresponds to the North-South momentum [36]. However, this choice fails to give a physically meaningful quantity in real (coordinate) space because the corresponding function $X_{\mathbf{k}}$ is singular (when $p \to 0$). This singularity means that respective invariant in real space

$$I = \frac{1}{2} \int X(\mathbf{r}_1, \mathbf{r}_2) s(\mathbf{r}_1, t) s(\mathbf{r}_2, t) d\mathbf{r}_1 d\mathbf{r}_2 \tag{IV.34}$$

has kernel $X(\mathbf{r}_1, \mathbf{r}_2)$ which does not vanish at large separation $\mathbf{r}_1 - \mathbf{r}_2$; see [37] for a detailed discussion.

The function (IV.33d) fails to produce an invariant either. This is because $\xi_{\mathbf{k}}$ is even in $\mathbf{k}$, and so, $X_{\mathbf{k}}$ is odd, contradicting the symmetry (IV.19).

Unlike $\xi_{\mathbf{k}}$, the function $\eta_{\mathbf{k}}$ is odd, and the corresponding kernel $X_{\mathbf{k}}$ determines an extra invariant for rotating shallow water dynamics. The previously described kernel (II.11c) is a linear combination of the functions (IV.33a) and (IV.33e).

The proof of the fact that the functions $\xi_{\mathbf{k}}$ and $\eta_{\mathbf{k}}$ are conserved in triad resonance interactions has recently been significantly simplified. The new proof is more straightforward and can be accomplished with the aid of symbolic algebra software. Indeed,

$$\xi_{\mathbf{k}} + i\eta_{\mathbf{k}} = \ln Z_{\mathbf{k}} \quad \text{where} \quad Z_{\mathbf{k}} = \frac{if(q + p\sqrt{3}) + k^2}{if(q - p\sqrt{3}) + k^2}, \tag{IV.35}$$

(ln denotes the principal branch of the complex logarithm, with argument between $-\pi$ and π), and the required conservation equation

$$(\xi_1 + i\eta_1) + (\xi_2 + i\eta_2) + (\xi_3 + i\eta_3) = 0 \tag{IV.36}$$

implies

$$Z_1\, Z_2\, Z_3 = 1. \tag{IV.37}$$

Now, using (IV.29a), substitute $p_3 = -p_1 - p_2$, $q_3 = -q_1 - q_2$ into (IV.29b) and (IV.37). These equations may then be reduced to two polynomial equations of degree 5 in p_1, q_1, p_2, q_2. It is easy to check (e.g., with Mathematica software) that these two polynomials are identical up to a constant factor. It follows immediately that the resonance equations (IV.29) imply (IV.37), and hence that

$$\ln Z_1 \,+\, \ln Z_2 \,+\, \ln Z_3 \,=\, 2\pi m i, \quad \text{where} \quad m = 0, \pm 1, \pm 2, \dots. \tag{IV.38}$$

Continuity considerations require $m = 0$ [13], and the conservation (IV.36) then follows.

Thus, there are three invariants:

- *the energy of the Rossby component* [corresponding to (IV.33a)],
- *the enstrophy* [corresponding to (IV.33b)],
- *the extra invariant* [corresponding to (IV.33e)].

H. Dropping cubic terms

The cubic terms I^{cubic} have served their purpose in the proof, and can now be dropped, similar to the argument [38] for the quasigeostrophic equation. To see this, first, note that the β^2-terms in (IV.6) can be neglected over a time interval of length at most of order $\beta^{-\nu}$ with $\nu < 2$. We also need to consider time intervals containing many wave periods, and so $\nu > 1$. For specificity, we choose $\nu = 3/2$.

Considering (IV.20), we have neglected terms $\propto A^3\beta^{1/2} \propto A^2\epsilon\beta^{3/2}$ (such terms come from neglecting $\mathcal{R}$-correction in the nonlinear terms of the shallow water equations). Therefore, over a time $t \propto \beta^{-3/2}$, the error can accumulate at most up to a total error $\propto A^2\epsilon$. The M, N, and T-corrections in (IV.17) have the order $A^3 \propto A^2\epsilon\beta$; the Y-correction in (IV.17) has the order $A^3/\beta \propto A^2\epsilon$ [the kernel Y is proportional to $1/\beta$, while the kernels M, N, T are $O(1)$, $\beta \to 0$]. So, all cubic corrections are within the total conservation error $\propto A^2\epsilon$ and can be safely dropped. As alluded to earlier, these corrections were needed in the derivation only to control oscillatory terms; their amplitude is now seen to be small, but their time derivative is large (has lower order).

V. REMARKS

A. *Unique* invariant

The existence of an extra invariant motivates a natural question: Do there exist other invariants in the shallow water system? The answer appears to be "No", although rigorous investigation of this question has not been attempted. To elaborate, if such an invariant did exist, then the resonance triad interaction (IV.29) would seem to have another conserved quantity, besides (IV.33). The latter, however, is known to be untrue [39]. This was established by the connection [40] between invariants of wave interactions and *Web geometry* [41]. It has not been ruled out, however, that the shallow water system (I.2) has several invariants, which collapse into a single invariant for the quasigeostrophic equation (I.1); though this seems unlikely.

The connection to the Web geometry also shows that the dispersion laws that admit extra invariants are extremely rare. We are aware of only one other physical system (besides Rossby waves) that possesses extra invariants. This is the generalization of the Korteweg-de Vries (KdV) equation for two spatial dimensions

$$(\psi_t + \psi\psi_x + \psi_{xxx})_x = \psi_{yy}; \tag{V.1}$$

it has dispersion law

$$\Omega(p,q) = -p^3 - \frac{q^2}{p}. \tag{V.2}$$

Equation (V.1) is integrable via the inverse scattering method and has infinitely many extra invariants [35]. The system (V.1) is called the Kadomtsev-Petviashvili equation of the first kind (KP1); the Kadomtsev-Petviashvili equation of the second kind (KP2) has a minus sign in front of the term on the right of (V.1); because of this, triad resonances do not exist at all for KP2.

Unlike to the KP1 case, the Rossby wave triad resonance admits only one extra invariant; and moreover, it is impossible to extend this invariant to the next nonlinearity order [38]. So, the extra invariant of the shallow water dynamics (I.2) is an attribute of weak nonlinearity.

The triad resonances that admit finite number of extra invariants are even more rare than the ones with infinitely many invariants (see [39]): The former constitute a several parameter family among all functions depending on two variables; and moreover, most of the members of this family are not even elementary functions and hardly can be dispersion laws of physical systems.

Unlike the potential vorticity integrals (II.1) [with the exception of constant and linear F], the extra invariant is conserved in the Galerkin approximation with a finite number of Fuorier modes. The invariant can be important for the dynamics of wave clusters, formed by the triad resonances of wave vectors on a lattice, [42].

B. The impact of the extra invariant on statistical equilibrium

The existence of the extra invariant may provide barriers to statistical equilibration. The equilibrium theory for the quasigeostrophic ([43] and references therein) and shallow water systems [44] were derived by enforcing only the exact conservation laws (energy, momentum, and the potential vorticity hierarchy). Since the latter fully define the equilibrium state (under the ergodic hypothesis), the adiabatic conservation laws will generally be violated. Given that true equilibration is an infinite time property, the presence of an adiabatic invariant does not lead to any mathematical contradiction here. However, there are practical issues since the extra invariant could greatly increase the equilibration time scale. This issue needs to be investigated.

We should also note a parallel between the existence of the extra invariant, determined only by the Rossby component, and the equilibrium theory. In the latter it is found that although the inertia-gravity waves do remove some of the initial energy to small scale surface ripples, they do not inhibit the inverse cascade of the remaining energy to form large-scale vortex equilibria.

C. Using perturbational potential vorticity instead of its linearization

Since the extra conservation holds only in the weakly nonlinear limit, to the same accuracy we are free to write the invariants in terms of the perturbational potential vorticity

$$\tilde{Q} = \frac{v_x - u_y + f(y)}{H} - \frac{f(y)}{\bar{H}} \tag{V.3}$$

instead of the *linearized* perturbational potential vorticity $\mathcal{Q}$, equation (II.7). We have $\tilde{Q} \approx \bar{H}\mathcal{Q}$ (with the error due to nonlinear terms), and instead of (II.10),

$$I = \frac{1}{2\bar{H}^2} \int X_{\mathbf{k}}\, \tilde{Q}_{\mathbf{k}}\, \tilde{Q}_{-\mathbf{k}}\, dp\, dq, \tag{V.4}$$

where $\tilde{Q}_{\mathbf{k}}$ is the Fourier transform of the field $\tilde{Q}$.

D. Can rotating shallow water dynamics be approximated by a single equation?

There is a question whether the shallow water system (I.2) can be approximated *near the equator* by a single equation. Certainly, in the rigid lid approximation (H = const) the system (I.2) is reduced to the equation of 2D hydrodynamics with beta-effect. However, the shallow water dynamics contain three independent variables u, v, H, and accordingly the system (I.2) contains 3 time derivatives. We allow significant deviations of H from its average value $\bar{H}$.

We have attempted to approximate the equatorial shallow water dynamics by a single equation

$$\dot{a}_1 = \Omega_1 a_1 + \int W_{-123}\; a_2\, a_3\; \delta_{-123}\; d_{23} \tag{V.5}$$

for the Fourier transform $a_{\mathbf{k}}(t)$ of the stream function or some other variable (the notation is defined in Sec. IV C). However, we found that such an equation would have insufficient accuracy to establish the extra conservation. More specifically, the formula for the kernel W would lack one more cancellation in equations similar to (IV.24) [numerator in W, instead of being quadratic, would be linear in f, p_3, q_3], and so, the kernel W would be singular.

E. Possible fast dependence on the y-coordinate

The extra conservation holds if the coefficients in the shallow water system (II.6) additionally contain fast, but small amplitude, dependence on the y-coordinate. Such inhomogeneity may be considered at lowest order as a resonant triad interaction between two Rossby waves, with dispersion law (IV.7), and one inhomogeneity wave, with zero dispersion law:

$$p_1 = p_2 + p_3, \tag{V.6a}$$
$$q_1 = q_2 + q_3, \tag{V.6b}$$
$$\Omega(p_1, q_1) = \Omega(p_2, q_2) + 0; \tag{V.6c}$$

here $\mathbf{k}_1 = (p_1, q_1)$ and $\mathbf{k}_2 = (p_2, q_2)$ are the Rossby wave vectors, and $\mathbf{k}_3 = (p_3, q_3)$ is the inhomogeneity wave vector.

If translation symmetry is still maintained in the x-coordinate, one has $p_3 \equiv 0$. For this case, one can readily see that an arbitrary function $\varphi(p, q)$ that is even in q satisfies

$$\varphi(p_1, q_1) = \varphi(p_2, q_2) + 0 \tag{V.6d}$$

at each point of the resonance manifold (V.6abc). Indeed, (V.6a), with $p_3 \equiv 0$, and (V.6c) imply $p_1 = p_2$ and $|q_1| = |q_2|$. In particular, the function (II.11c) is even in q, and the conservation (V.6d) holds for $\varphi \equiv \eta$. Thus, the function (II.11c) is conserved in triad resonant interactions of Rossby waves with the inhomogeneity waves.

Actually, the function (II.11c) is conserved in resonant interactions of any order n ($n \geq 3$), which involve 2 Rossby waves and $n - 2$ inhomogeneity waves:

$$p_1 = p_2 + \underbrace{0 + 0 + \ldots + 0}_{n-2}, \tag{V.7a}$$
$$q_1 = q_2 + q_3 + \ldots + q_n, \tag{V.7b}$$
$$\Omega(p_1, q_1) = \Omega(p_2, q_2) + \underbrace{0 + 0 + \ldots + 0}_{n-2}; \tag{V.7c}$$

Indeed (V.7a) and (V.7c) imply $p_1 = p_2$ and $|q_1| = |q_2|$, and therefore,

$$\varphi(p_1, q_1) = \varphi(p_2, q_2) + \underbrace{0 + 0 + \ldots + 0}_{n-2} \tag{V.7d}$$

for any function $\varphi(p, q)$ which is even in q.

VI. CONCLUSION

The Rossby waves have been known [12, 13] to possess a rare property: Their triad resonance admits an extra conserved quantity:

$$\left.\begin{array}{l}\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=0,\\ \Omega(\mathbf{k}_1)+\Omega(\mathbf{k}_2)+\Omega(\mathbf{k}_3)=0\end{array}\right\} \quad\Rightarrow\quad \eta(\mathbf{k}_1)+\eta(\mathbf{k}_2)+\eta(\mathbf{k}_3)=0 \tag{VI.1}$$

where

$$\mathbf{k}=(p,q) \quad (k^2=p^2+q^2), \qquad \Omega(\mathbf{k})=\frac{\beta p}{f^2+k^2}\ , \tag{VI.2}$$

$$\eta(\mathbf{k})=\arctan\left(f\frac{q+p\sqrt{3}}{k^2}\right)-\arctan\left(f\frac{q-p\sqrt{3}}{k^2}\right) . \tag{VI.3}$$

Despite of the implication (VI.1), the extra invariant I is actually independent of the energy and momentum (enstrophy) because the integrals (II.10) or (III.1) contain field variables $\mathcal{Q}_{\mathbf{k}}(t)$ or $\varepsilon_{\mathbf{k}}(t)$ respectively. [Recall that $X_{\mathbf{k}}=\eta(\mathbf{k})/p,\ \phi_{\mathbf{k}}=\eta(\mathbf{k})/\Omega(\mathbf{k})$.]

In the present paper, we have established two key results:

- The Rossby wave extra invariant can be extended to the shallow water dynamics in spite of the presence of inertia-gravity waves and in spite of the explicit inhomogeneity (the y-dependence of the Coriolis parameter f).
- The shallow water dynamics possesses an extra invariant in the equatorial limit (when $f\to 0$, but the derivative f' stays away from zero). This limit also leads to small denominators, but different from those related to the triad resonance. We have shown that it is possible to cancel these small denominators.

We have also found that for weakly nonlinear shallow water dynamics, the presence of the extra invariant constrains the inverse cascade energy transfer to be from small scale eddies to large scale zonal flow. The results are also in agreement with some more specific experimental features: more pronounced zonal jets near the equator, when $f\to 0$, and suppression of zonal jets and the 60° polar angle in the energy spectrum when $f\to\infty$ (see Sec. III). We have seen that the formation of zonal jets is a basic phenomenon that can be related to the set of invariants of the rotating shallow water dynamics.

For future work, it would be crucial to see whether the theoretical predictions agree with experimental observations quantitatively, and whether the effects of the extra invariant can be clearly resolved from other mechanisms in the plethora of zonal jets phenomena. In particular, we believe it important to develop our results for the dynamics of magnetized plasmas; it would be very interesting to examine the effects of the extra invariant on the formation of internal transport barriers in fusion plasmas.

ACKNOWLEDGMENTS

A.B. and P.W. wish to thank the Kavli Institute for Theoretical Physics, University of California Santa Barbara, for hospitality and support via the National Science Foundation Grant PHY05-51164; we are grateful for beneficial discussions with the participants in two Programs *Physics of Climate Change* and *Dynamo Theory* held concurrently at the Institute. This paper is based upon work partially supported by the National Science Foundation under Grant DMS-0405905.

[1] G. Batchelor, *Homogeneous Turbulence* (Cambridge University Press, 1953).
[2] R. Fjørtoft, Tellus **5**, 225 (1953).
[3] V. E. Zakharov, in *Breaking Waves IUTAM Symposium*, edited by M. L. Banner and R. H. J. Grimshaw (Sydney, Australia, 1991) pp. 69–91.
[4] V. E. Zakharov, Zh. Prikl. Mekh. Tekh. Fiz. **9**, 86 (1968), [Inglish Transl.: J. Appl. Mech. Tech. Phys. 9, 190-194 (1968)].
[5] P. B. Rhines, Annu. Rev. Fluid Mech. **11**, 401 (1979).
[6] *Jets and Annular Structures in Geophysical Fluids* (2007) proceedings of the Chapman conference held in Savannah, GA, January, 2006, published in a special issue of the J. Atmos. Sci.
[7] see NASA images, "http://photojournal.jpl.nasa.gov," (2001-2007).

[8] P. H. Diamond, S.-I. Itoh, K. Itoh, and T. S. Hahm, Plasma Phys. Control. Fusion **47**, R35 (2005).
[9] A. C. Newell, J. Fluid Mech. **35**, 255 (1969).
[10] P. B. Rhines, J. Fluid Mech. **69**, 417 (1975).
[11] R. Salmon, *Lectures on Geophysical Fluid Dynamics* (Oxford University Press, 1998).
[12] A. M. Balk, S. V. Nazarenko, and V. E. Zakharov, Phys. Lett. A **152**, 276 (1991).
[13] A. M. Balk, Phys. Lett. A **155**, 20 (1991).
[14] A. M. Balk, Phys. Lett. A **345**, 154 (2005).
[15] S. Nazarenko and B. Quinn, Phys. Rev. Lett. **103**, 118501 (2009).
[16] S. Nazarenko and B. Quinn, in *Turbulence in the Atmosphere and Oceans, Proceedings of Workshop held 8–12 December 2008 at Isaac Newton Institute* (Springer-Verlag, 2009).
[17] G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics. Fundamentals and Large-scale Circulation* (Cambridge, 2006).
[18] J. C. McWilliams, *Fundamentals of Geophysical Fluid Dynamics* (Cambridge, 2006).
[19] Q. Chen, S. Chen, G. L. Eyink, and D. D. Holm, J. Fluid Mech. **542**, 139 (2005).
[20] G. M. Reznik and T. E. Soomere, Okeanologiya **24**, 558 (1984).
[21] P. H. Diamond, M. N. Rosenbluth, F. L. Hinton, M. Malkov, J. Fleischer, and A. Smolyakov, in *17th IAEA Fusion Energy Conference, Yokohama, Japan* (International Atomic Energy Agency, Vienna, 1998) pp. IAEA–CN–69/TH3/1, 8 pages.
[22] Y. Lee and L. M. Smith, J. Fluid Mech. **576**, 405 (2007).
[23] M. D. Fruman and T. G. Shepherd, J. Atmos. Sci. **65**, 1927 (2008).
[24] M. D. Fruman, B. L. Hua, and R. Schopp, J. Atmos. Sci. **66**, 2600 (2009).
[25] P. H. LeBlond and L. A. Mysak, *Waves in the Ocean* (Elsevier, Amsterdam, 1978).
[26] O. Bühler and M. E. McIntyre, J. Fluid Mech. **492**, 207 (2003).
[27] S. Sukoriansky, N. Dikovskaya, and B. Galperin, Phys. Rev. Lett. **101**, 178501 (2008).
[28] P. Berloff, I. Kamenkovich, and J. Pedlosky, J. Fluid. Mech. **628**, 395 (2009).
[29] A. E. Gill, *Atmosphere — Ocean Dynamics* (Academic Press, New York, 1982).
[30] L. D. Landau and E. M. Lifshitz, *Mechanics*, 3rd ed., Course of theoretical physics, v. 1 (Butterworth-Heinemann, New York, 1976).
[31] R. E. Glazman and P. B. Weichman, Dynamics of Atmospheres and Oceans **38**, 173 (2005).
[32] A. Okuno and A. Masuda, Phys. Fluids **15**, 56 (2002).
[33] A. M. Balk and V. E. Zakharov, Phys. Lett. A **373**, 4049 (2009).
[34] V. E. Zakharov and E. I. Schulman, Physica D **29**, 283 (1988).
[35] V. E. Zakharov and E. I. Schulman, Physica D **10**, 192 (1980).
[36] M. S. Longuet-Higgins and A. E. Gill, Proc. Roy. Soc. London, Ser. A **299**, 120 (1967).
[37] A. M. Balk and T. Yoshikawa, Physica D **238**, 384 (2009).
[38] A. M. Balk and F. van Heerden, Physica D **223**, 109 (2006).
[39] A. M. Balk and E. V. Ferapontov, in *Nonlinear waves and weak turbulence*, edited by V. E. Zakharov (Amer. Math. Soc. Trans. Ser. 2, vol 182, 1998) pp. 1–30.
[40] E. V. Ferapontov, "Web geometry and mathematical physics," (Kluwer Academic Publishers, Norwell, MA, 1992) pp. 310–323, appendix in the book "Gemetry and Algebra of Multidimensional Three-Webs" by M. A. Akivis and A. M. Shelekhov.
[41] W. Blaschke, *Topological differential geometry* (Chicago University Press, 1932).
[42] E. Kartashova and V. S. L'vov, Europhys. Let. **83**, 50012 (2008).
[43] P. B. Weichman, Phys. Rev. E **73**, 036313 (2006).
[44] P. B. Weichman and D. M. Petrich, Phys. Rev. Lett. **86**, 1761 (2001).