JM PAPERS

Lower bounds on zonal enstrophy

H. Aibara[1],† and Z. Yoshida[1]

[1]Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan

An analytic estimate of the lower bounds on zonal enstrophy has been studied for the beta-plane model of two-dimensional barotropic flow. The estimate provides exact lower bounds on the zonal enstrophy, which hence must be satisfied regardless of the dynamics. The energy, impulse, circulation as well as the total enstrophy are invoked as constraints for the minimization of the zonal enstrophy. The corresponding variational principle has an unusual mathematical structure (primarily because the target functional is not a coercive form), by which the constraints work out in an interesting way. A discrete set of zonal enstrophy levels is generated by the energy constraint; each level is specified by an eigenvalue that represents the meridional mode number of zonal flow. However, the value itself of the zonal enstrophy level is a function of only impulse and circulation, being independent of the energy (and total enstrophy). Instead, the energy works in selecting the 'level' (eigenvalue) of the relaxed state. The relaxation occurs by emitting small-scale wavy enstrophy, and continues as far as the nonlinear effect, scaled by the energy, can create wavy enstrophy. Comparison to numerical simulations shows that the theory gives a consistent estimate of the zonal enstrophy in the relaxed state.

Key words: pattern formation, variational methods, atmospheric flows

1. Introduction

The creation of zonal flow in the planetary atmosphere is a spectacular example of the self-organization in physical systems (Charney 1971; Hasegawa 1985). There is a strong analogy between the geostrophic turbulence and the electrostatic turbulence of a magnetized plasma in the plane perpendicular to an ambient magnetic field. Because the generation of zonal flow (coherent structure) affects the turbulent transport in magnetized plasmas, how strong it can be is of great interest in the context of plasma confinement (Diamond *et al.* 2005). The aim of this work is to estimate exact lower bounds on the 'zonal enstrophy', which hence must be satisfied regardless of the dynamics, and elucidate

† Email address for correspondence: aibara.hiroto19@ae.k.u-tokyo.ac.jp

how the lower bounds are determined; lower bounds on the zonal enstrophy indicate that the zonal flow must be stronger than the given value.

The inverse-cascade model explains the essence of the self-organization process. Because of the approximate two-dimensional geometry (due to the scale separation between the shallow vertical direction and wide horizontal directions), the vortex dynamics is free from the stretching effect. Then, the energy of flow velocity tends to accumulate into large-scale vortices, while the enstrophy (the norm of vorticity) cascades to small scales (Kraichnan 1967). On a rotating sphere, the gradient of the Coriolis force yields the Rossby-wave term in the vortex dynamics equation, which brings about latitude/longitude anisotropy, and the large-scale vorticity forms zonal flow (Charney 1971). The nonlinear term driving the inverse cascade becomes comparable to the linear Rossby-wave term at the Rhines scale, which gives a crude estimate of the latitudinal size of the zonal flow (Rhines 1975).

While the inverse-cascade model illustrates the general tendency of a nonlinear process, the underlying mechanism requires more detailed analysis. The modulational instability plays an essential role in exciting the energy transfer in the wavenumber space (Lorentz 1972; Gill 1974; Connaughton *et al.* 2010). In addition to the energy and enstrophy, another quadratic integral is known to be an adiabatic invariant (only changes by fourth order of perturbations) in the zonal-flow domain of wavenumber space, restricting the energy transfer there (Balk, Nazarenko & Zakharov 1991; Balk 1991, 2005). Various numerical simulations have been done to demonstrate the creation of zonal flow. Two different categories of models must be distinguished; one is the unforced, free decaying turbulence, and the other is the forced, quasi-stationary turbulence. In the latter case, the interaction between the mean flow and the turbulence (Farrell & Ioannou 2007; Bakas & Ioannou 2011; Srinivasan & Young 2012) or inhomogeneous vorticity mixing (Dritschel & McIntyre 2008; Scott & Dritschel 2012) has been found to be a causal mechanism of zonal-flow generation. For these forced, quasi-stationary cases, one has to include some dissipation mechanism for large-scale flows in order to remove the energy accumulating in the large-scale regime by the inverse cascade. The usual viscosity only works for short-scale flows, so something like 'friction' is added to the model (however, which mechanism works in a realistic planetary system is still controversial). For the free decaying case, early simulation results (Vallis & Maltrud 1993; Yoden & Yamada 1993; Yoden *et al.* 1999) demonstrated the self-organization of zonal flow, and found that the scale of the zonal flow is similar to that of Rhines' estimate. However, the quantitative comparison between the Rhines scale and the zonal-flow scale was left unclear. On the other hand, in the forced turbulence case, a more complex relation has been found, because of the influence of the dissipation mechanism for large-scale flows (see Williams 1978; Danilov & Gurarie 2002; Sukoriansky, Dikovskaya & Galperin 2007).

In parallel with simulation studies, there have been theoretical attempts to nail down the 'target' of the spontaneous process, i.e. formulating a variational principle that reveals what the dynamics tends to reach. This can be done by identifying the target functional to be minimized (or maximized) as well as the constraints that restrict admissible candidates. A well-known example is the entropy maximization in a microcanonical ensemble (the constraints are total particle number and total energy), which gives the Gibbs distribution. In the application to field theories, where we have to deal with infinite-dimensional phase spaces, we encounter the problem of ultraviolet catastrophe (which must be removed by appropriate quantization, see Ito & Yoshida (1996)). Suspending such subtle problems, formal calculations have been made to obtain the thermal equilibrium distribution of flow fields. In the context of the planetary atmosphere, the statistical equilibrium state in the

two-dimensional incompressible Euler system was studied in Miller (1990) and Robert & Sommeria (1991). In Turkington *et al.* (2001), the maximum entropy distribution over the ensemble constrained by total energy and circulation was compared with the large-scale vortex structures observed on Jupiter. However, because of the essential non-equilibrium property of turbulence (as the cascade model is based on 'dissipation' in the Kolmogorov micro-scales, one has to assume a 'driving force' to maintain the (quasi-) stationary state, or consider a transient process of free decay), the entropy may not be an effective tool to dictate the self-organization.

There is a different type of approach guided by the notion of selective dissipation (Hasegawa 1985). The Taylor state of a magneto-fluid (Taylor 1974, 1986) is the prototype of such a model of self-organization, which minimizes the magnetic field energy ($E = (1/2) \int |\boldsymbol{B}|^2 \, \mathrm{d}^3 x = (1/2) \int |\boldsymbol{\nabla} \times \boldsymbol{A}|^2 \, \mathrm{d}^3 x$, where $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ is the magnetic field and $\boldsymbol{A}$ is the vector potential) under the constraint on the magnetic helicity ($H = (1/2) \int \boldsymbol{A} \cdot \boldsymbol{B} \, \mathrm{d}^3 x$). The reason why E is more fragile than H is because E includes another differential operator curl in the integrand. This model explains the relaxed states of magnetized plasmas in various systems, ranging from laboratory experiments to astronomical objects. Regarding the two-dimensional turbulence in the planetary atmosphere, the minimization of the generalized enstrophy (see Proposition 2.1) under the constraint on the energy has been studied to show that the solution of the minimization problem predicts a steady state with streamlines parallel to contours of the topography (Bretherton & Haidvogel 1976). Although these two stories, i.e. the energy–helicity relation in the magneto-fluid and the enstrophy–energy relation in the two-dimensional (2-D) fluid appear to be parallel (as Hasegawa (1985) describes in the unified vision), there is a fundamental difference when viewed from their Hamiltonian structures, and the latter needs a careful interpretation. In both systems, the ideal constants (the helicity in magneto-fluids and the enstrophy in 2-D fluids) are Casimirs, by which the orbits are constrained on the level sets of these constants (Morrison 1998). In the magneto-fluid phase space, the orbits converge into the equilibrium point as the energy diminishes; the minimum energy (Hamiltonian) state, on each level set of the helicity, gives an equilibrium point. In the 2-D fluid system, however, the level set of the enstrophy is not embedded as a smooth submanifold in the topology of the energy norm (because the enstrophy is a fragile quantity, its level set looks like a fractal set; see Appendix A). Hence, we have to reverse the role of the Hamiltonian (energy) and the Casimir (enstrophy), and minimize the enstrophy for a given energy. Then, the critical point is not necessarily an equilibrium point. In this specific problem, however, it happens to be so, because it is the 'maximum point' of the energy. Notice that the minimization of the enstrophy under a constrained energy is equivalent to the maximization of the energy under a constrained enstrophy (see Appendix A). The maximization of the energy appears to be consistent with the inverse-cascade story. However, the simultaneous process, i.e. the forward cascade of the enstrophy, violates the constancy of the enstrophy. The dual aspects of the 2-D turbulence pose a paradox in the mechanical interpretation of the selective dissipation.

The target of this study is totally different. Whereas we formulate a variational principle using the list of ideal constants of motion, the target functional is not selected from them. We estimate the minimum of the enstrophy possibly given to the zonal component (which we call the zonal enstrophy). Knowing how strong the zonal flow must be and how it is controlled is an important issue in the study of turbulent transport. While the total enstrophy is an ideal constant of motion, the zonal part alone is not. We are not proposing that the zonal enstrophy is selectively dissipated; we never provide the target functional with the role of dictating dissipation process. Our target functional is simply what we

want to estimate. We derive an *a priori* estimate of the zonal enstrophy, which must apply to every possible dynamics under a set of prescribed conditions; the ideal invariants are used as such constraints (we do not include the adiabatic invariant, because it needs the wavenumber information that is not amenable to our formulation). The actual dynamics is the second subject to be explored, which will be the task of § 5. The analogy of quantum mechanical energy levels may be helpful to explain our perspective. When we want to estimate the energy of an orbital electron, the variational principle to find the critical values of energy, for a fixed total probability, leads us to the eigenvalue problem for the Hamiltonian. The actual energy level that a particular electron will take is determined, for example, by the de-excitation process of emitting photons. We will find a similar picture for the 2-D-fluid turbulence; the zonal enstrophy has discrete levels of critical values (local minima); by emitting wavy enstrophy, the zonal enstrophy relaxes into lower levels.

If there is no constraint on partitioning, the zonal enstrophy can be minimized to zero (even if the total enstrophy is kept at a non-zero constant). But some constraints prevent this occurring. We will identify the 'key constraints' that determine a reasonable estimate of the zonal enstrophy.

The reciprocal problem, which maximizes the complementary wavy enstrophy (= total enstrophy − zonal enstrophy), was first studied by Shepherd (Shepherd 1988) with a different motivation, i.e. to estimate upper bounds on instabilities in the nonlinear regime. This is seemingly equivalent to the minimization of the zonal component, however, the effective 'constraints' may differ (see Appendix A). The conservation of the pseudo-momentum was invoked as the essential constraint. Improved estimates have been proposed by taking into account more general set of invariants which are known as Casimirs (Ishioka & Yoden 1996). In the present study of the minimization of the zonal component, however, we invoke a different constant of motion, the energy, as the principal constraint (in addition to other ones such as the impulse). The physical reason is clear because the self-organization is a spontaneous process in which the redistribution of the enstrophy between the zonal and wavy components can occur only if the energetics admits. Moreover, the energy constraint imparts a mathematically peculiar property to the variational principle, which is the other incentive of this study.

In the next section, we will start by reviewing the basic formulation and preliminaries. Section 4 describes the main result. We will derive discrete levels of the minimum zonal enstrophy. We will propose the notion of de-excitation to lower enstrophy levels (in analogy to energy levels of quantum states); the relaxation into lower levels corresponds to the inverse cascade. According to the conjecture of the Rhines scale (Rhines 1975), the inverse cascade continues until the linear Rossby-wave term overcomes the nonlinear term. In § 5, we will study the relaxation process by numerical simulation. The conventional Rhines scale will be revisited to give an improved estimate of the relaxed zonal enstrophy level. Section 6 concludes this paper.

2. Governing equations

2.1. *Vortex dynamics on a beta plane*

We consider a barotropic fluid on a beta-plane

$$M = \{\boldsymbol{\xi} = (x, y)^{\mathrm{T}}; \ x \in [0, 1), y \in (0, 1)\}. \tag{2.1}$$

Here, $\boldsymbol{\xi}$ is the coordinate while x is the azimuthal coordinate (longitude) and y is the meridional coordinate (latitude). The aspect ratio of the domain does not influence the results of Theorems 4.1 and 4.2 (see § 5.1), so we consider a unit square domain.

We are using the standard single-layer beta-channel model of geophysical fluid dynamics on the unit square, with periodic boundary conditions in the zonal x-direction and no-flux boundary conditions in the meridional y-direction. Without loss of generality, we work in a reference frame in which the net zonal mass flux vanishes. We will denote the standard L^2 inner product by $\langle f, g \rangle$

$$\langle f, g \rangle = \int_M f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) \, d^2\xi, \tag{2.2}$$

and the L^2 norm by $\|f\| = \langle f, f \rangle^{1/2}$. Here, f and g are arbitrary functions.

The state vector is the fluid vorticity $\omega \in L^2(M)$. We define the streamfunction (or Gauss potential) ψ by

$$-\Delta \psi = \omega, \tag{2.3}$$

where $\Delta = \partial_x^2 + \partial_y^2$. The flow velocity is given by

$$\boldsymbol{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \boldsymbol{\nabla}_\perp \psi = \begin{pmatrix} \partial_y \psi \\ -\partial_x \psi \end{pmatrix}. \tag{2.4}$$

Taking into account the Coriolis force, the governing equation of ω is

$$\partial_t \omega + \{\omega + \beta y, \psi\} = 0, \tag{2.5}$$

where $\{f, g\} = (\partial_x f)(\partial_y g) - (\partial_x g)(\partial_y f)$, and β is a real constant number measuring the meridional variation of the Coriolis force. When $\beta = 0$, (2.5) reduces into the standard vorticity equation. A finite β introduces anisotropy to the system, resulting in the creation of zonal flow. The Rhines scale (Rhines 1975) speaks of the balance of the two terms $\{\omega, \psi\}$ and $\{\beta y, \psi\}$, by which we obtain the typical scale length of the zonal flow (see § 4.5).

Inverting (2.3) by $\mathcal{K} = (-\Delta)^{-1}$, we may rewrite (2.5) as

$$\partial_t \omega + \{\omega + \beta y, \mathcal{K}\omega\} = 0. \tag{2.6}$$

We call

$$\omega_t := \omega + \beta y \tag{2.7}$$

the total vorticity, which is the sum of the fluid part ω and the ambient part βy (the latter is due to the rotation of the system).

The following identity will be useful in the later calculations:

$$\langle f, \{g, h\} \rangle = \langle g, \{h, f\} \rangle, \tag{2.8}$$

where f, g and h are C^1-class functions in M, and either f or g satisfy the boundary conditions.

2.2. *Conservation laws and symmetries*

PROPOSITION 2.1 (CONSTANTS OF MOTION). *The following functionals are constants of motion of the evolution equation* (2.6):

(i) *Energy*

$$E(\omega) := \tfrac{1}{2} \langle \omega, \mathcal{K}\omega \rangle. \tag{2.9}$$

By rewriting

$$E = \tfrac{1}{2} \langle (-\Delta \psi), \psi \rangle = \tfrac{1}{2} \int_M |\boldsymbol{\nabla} \psi|^2 \, d^2\xi = \tfrac{1}{2} \int_M |\boldsymbol{\nabla}_\perp \psi|^2 \, d^2\xi = \tfrac{1}{2} \int_M |\boldsymbol{v}|^2 \, d^2\xi, \tag{2.10}$$

we find that E evaluates the kinetic energy of the flow $\boldsymbol{v}$.

(ii) *Longitudinal momentum*

$$P(\omega) := \int_M \partial_y(\mathcal{K}\omega)\, \mathrm{d}^2\xi. \tag{2.11}$$

We may rewrite

$$P = \int_M \partial_y\psi\, \mathrm{d}^2\xi = \int_M v_x\, \mathrm{d}^2\xi, \tag{2.12}$$

to see that P is the integral of the longitudinal momentum. Considering the boundary conditions, P must be constantly zero.

(iii) *Circulation*

$$F(\omega) := \langle 1, \omega \rangle. \tag{2.13}$$

Integrating by parts, we may write

$$F = \int_0^1 [v_x]_{y=0}^{y=1}\, \mathrm{d}x, \tag{2.14}$$

which evaluates the circulation of the flow $\boldsymbol{v}$ along the boundary ∂M.

(iv) *Impulse*

$$L(\omega) := \langle y, \omega \rangle. \tag{2.15}$$

Integrating by parts and using the boundary conditions, we may rewrite

$$L = \int_M y(\partial_x v_y - \partial_y v_x)\, \mathrm{d}^2\xi = \int_M v_x\, \mathrm{d}^2\xi - \int_0^1 \left[y v_x \right]_{y=0}^{y=1}\, \mathrm{d}x. \tag{2.16}$$

The first term on the right-hand side is P, which vanishes by the boundary condition. Hence, L corresponds to the angular momentum $\boldsymbol{\xi} \times \boldsymbol{v}$ averaged over the boundary.

(v) *Generalized enstrophy*

$$G_\beta(\omega) := \int_M f(\omega + \beta y)\, \mathrm{d}^2\xi, \tag{2.17}$$

where f is an arbitrary C^1-class function, and the argument $\omega + \beta y$ is the total vorticity including the ambient term βy. For $f(u) = u^2/2$, $G_\beta(\omega)$ is the conventional enstrophy of the total vorticity.

(vi) *Fluid enstrophy*

$$Q(\omega) := \tfrac{1}{2}\|\omega\|^2. \tag{2.18}$$

3. Zonal and wavy components

The phase space of the vorticity ω is

$$V = L^2(M). \tag{3.1}$$

We say that ω is zonal when $\partial_x\omega \equiv 0$ in M. The totality of zonal flows defines a closed subspace $V_z \subset V$. The zonal average

$$\mathcal{P}_z\omega := \int_0^1 \omega(x, y)\, \mathrm{d}x \tag{3.2}$$

may be regarded as a projection from V onto V_z. By the orthogonal decomposition $V = V_z \oplus V_w$, we define the orthogonal complement V_w, i.e. $\omega_w \in V_w$, iff $\langle \omega_w, \omega_z \rangle = 0$

for all $\omega_z \in V_z$. We call $\omega_w \in V_w$ a wavy component, which has zero zonal average: $\mathcal{P}_z\omega_w = 0$. We will denote

$$\mathcal{P}_w = I - \mathcal{P}_z, \tag{3.3}$$

which is the projector onto V_w. Now we may write

$$V = V_z \oplus V_w = (\mathcal{P}_zV) \oplus (\mathcal{P}_wV). \tag{3.4}$$

The following basic properties may be known to the reader, but we summarize them as Lemmas for the convenience of the analysis in § 4:

LEMMA 3.1 (PARTITION LAWS). *Let us decompose* $\omega = \omega_z + \omega_w(\omega_z = \mathcal{P}_z\omega \in V_z, \omega_w = \mathcal{P}_w\omega \in V_w)$.

   (i) *The circulation and the impulse are occupied by the zonal component* ω_z, *i.e.*

$$F(\omega) = F(\omega_z), \tag{3.5}$$

$$L(\omega) = L(\omega_z). \tag{3.6}$$

   (ii) *Quadratic invariants such as the fluid enstrophy and the energy are separated without any mixture terms as*

$$Q(\omega) = Q(\omega_z) + Q(\omega_w), \tag{3.7}$$

$$E(\omega) = E(\omega_z) + E(\omega_w). \tag{3.8}$$

4. Estimate of zonal enstrophy

4.1. *Zonal enstrophy versus wavy enstrophy*

The aim of this work is to find the minimum of the zonal enstrophy defined by

$$Z(\omega) := \tfrac{1}{2}\|\mathcal{P}_z\omega\|^2. \tag{4.1}$$

The complementary wavy enstrophy is $W(\omega) = (1/2)\|\mathcal{P}_w\omega\|^2$. By (3.7), the total enstrophy is

$$Q(\omega) = Q(\mathcal{P}_z\omega) + Q(\mathcal{P}_w\omega) = Z(\omega) + W(\omega). \tag{4.2}$$

When the total enstrophy $Q(\omega)$ is conserved (see Proposition 2.1(v)), the minimum of $Z(\omega)$ gives the maximum of $W(\omega)$.

The simplest version of the minimization problem is to find the minimum $Z(\omega)$ under the constraint of $Q(\omega) = C_Q (\neq 0)$. Introducing a Lagrange multiplier ν, we minimize

$$Z(\omega) - \nu Q(\omega). \tag{4.3}$$

Using the self-adjointness of $\mathcal{P}_z$, we obtain the Euler–Lagrange equation

$$\mathcal{P}_z\omega - \nu\omega = 0. \tag{4.4}$$

Operating $\mathcal{P}_z$ on (4.4) yields

$$(1 - \nu)\mathcal{P}_z\omega = 0. \tag{4.5}$$

On the other hand, operating $\mathcal{P}_w$ yields

$$\nu\mathcal{P}_w\omega = 0. \tag{4.6}$$

There are two possibilities for solving these simultaneous equations.

(i) If $\nu = 0$ then, $\mathcal{P}_z\omega = \omega_z = 0$ and $\mathcal{P}_w\omega = \omega_w$ is an arbitrary function satisfying $Q(\omega_w) = C_Q$; hence, $\min Z(\omega) = 0$. (This simple exercise reveals an unusual aspect of the present variational principle, which is caused by the non-coerciveness of the functional $Z(\omega)$ to be minimized. Notice that the minimizer is not unique, because $\mathcal{P}_z$ has non-trivial kernel, i.e. $\mathrm{Ker}(\mathcal{P}_z) = V_w$).

(ii) If $\nu = 1$ then, $\mathcal{P}_w\omega = \omega_w = 0$ and $\mathcal{P}_z\omega = \omega_z$ is an arbitrary function satisfying $Q(\omega_z) = Z(\omega) = C_Q$; hence, this solution gives the 'maximum' of $Z(\omega)$.

As we mentioned above, the minimizer is not unique here. To obtain a non-trivial estimate of the minimum $Z(\omega)$, we have to take into account 'constraints' posed on the dynamics of redistributing enstrophy. Guided by Proposition 2.1, we start with some simple ones.

4.2. *Constraints by circulation and impulse*

Let us consider the circulation and impulse as constraints.

THEOREM 4.1. *The minimizer of the zonal enstrophy $Z(\omega)$ under the constraints on the circulation $F(\omega) = C_F$, the impulse $L(\omega) = C_L$, as well as the total enstrophy $Q(\omega) = C_Q$ is a vorticity ω such that*

$$\mathcal{P}_z\omega = a + by, \quad (a = 4C_F - 6C_L, \; b = 12C_L - 6C_F), \tag{4.7}$$

which gives

$$Z_0 := \min Z(\omega) = 2C_F^2 - 6C_F C_L + 6C_L^2. \tag{4.8}$$

Proof. Let us minimize

$$Z(\omega) - \nu Q(\omega) - \mu_0 F(\omega) - \mu_1 L(\omega). \tag{4.9}$$

The Euler–Lagrange equation is

$$\mathcal{P}_z\omega - \nu\omega = \mu_0 + \mu_1 y. \tag{4.10}$$

Operating $\mathcal{P}_z$ on both sides of (4.10) yields

$$(1 - \nu)\mathcal{P}_z\omega = \mu_0 + \mu_1 y. \tag{4.11}$$

On the other hand, operating $\mathcal{P}_w$ yields

$$\nu\mathcal{P}_w\omega = 0. \tag{4.12}$$

First, assume that $1 - \nu \neq 0$. Inserting $\mathcal{P}_z\omega$ of (4.11) into the definition of $F(\omega) = F(\mathcal{P}_z\omega)$ and $L(\omega) = L(\mathcal{P}_z\omega)$ (see Lemma 3.1(i) and (ii)), we determine μ_0 and μ_1 to match the constraint $\langle 1, \omega \rangle = C_F$ and $\langle y, \omega \rangle = C_L$; we obtain $a := \mu_0/(1 - \nu) = 4C_F - 6C_L$, and $b := \mu_1/(1 - \nu) = 12C_L - 6C_F$. Inserting this $\omega_z = a + by$ into $Z(\omega)$, we obtain the minimum (4.8). On the other hand, (4.12) is satisfied by $\nu = 0$ (consistent with the forgoing assumption $1 - \nu \neq 0$) and an arbitrary $\omega_w = \mathcal{P}_w\omega$ such that

$$\tfrac{1}{2}\|\omega_w\|^2 = C_Q - (2C_F^2 - 6C_F C_L + 6C_L^2). \tag{4.13}$$

The right-hand side is non-negative, if the constraints $F(\omega) = C_F$, $L(\omega) = C_L$ and $Q(\omega) = C_Q$ are consistent. It is only when the constants C_F, C_L and C_Q are given so that the right-hand side of (4.13) is zero, that the other assumption $1 - \nu = 0$ applies; then, the unique solution $\mathcal{P}_w\omega = 0$ (hence, $\omega = \mathcal{P}_z\omega$) is obtained. ∎

Notice that the minimizer is still non-unique (excepting the special case mentioned in the proof); every $a + by + \omega_w$ ($\forall\omega_w \in V_w$ such that (4.13) holds) satisfies (4.7). However, the minimum value (4.8) is uniquely determined.

4.3. *Constraint by energy*

The situation changes dramatically when we include the energy constraint $E(\omega) = C_E$; a laminated vorticity distribution, epitomizing the structure of zonal flow, is created by the energy constraint. The meridional mode number of the zonal flow is identified by the 'eigenvalue' of the Euler–Lagrange equation, which specifies the 'level' of the zonal enstrophy (in analogy to the quantum number of discrete energy in quantum mechanics). To highlight the role of the energy constraint, we first omit the constraints on the circulation and impulse.

Taking into account the energy and total enstrophy constraint, we seek the critical points of

$$Z(\omega) - \nu Q(\omega) - \mu_2 E(\omega). \tag{4.14}$$

The Euler–Lagrange equation is

$$\mathcal{P}_z \omega - \nu \omega - \mu_2 \mathcal{K}\omega = 0. \tag{4.15}$$

Operating $\mathcal{P}_z$ yields (denoting $\omega_z = \mathcal{P}_z \omega$)

$$\omega_z - \nu \omega_z - \mu_2 \mathcal{K}\omega_z = 0. \tag{4.16}$$

On the other hand, $\omega_w = \mathcal{P}_w \omega$ must satisfy

$$\nu \omega_w + \mu_2 \mathcal{K}\omega_w = 0. \tag{4.17}$$

Putting $\omega_z = -\partial_y^2 \psi_z(y)$ in (4.16), we obtain

$$\partial_y^2 \psi_z + \lambda^2 \psi_z = 0, \quad \lambda^2 = \frac{\mu_2}{1 - \nu}. \tag{4.18a,b}$$

The solution satisfying the boundary conditions $\psi_z(0) = \psi_z(1) = 0$ is

$$\psi_z = A \sin \lambda y, \tag{4.19}$$

with eigenvalues

$$\lambda = n_1 \pi \ (n_1 \in \mathbb{Z}). \tag{4.20}$$

The corresponding zonal vorticity is

$$\omega_z = A\lambda^2 \sin \lambda y. \tag{4.21}$$

On the other hand, putting $\omega_w = -\Delta \psi_w$, (4.17) reads

$$\Delta \psi_w + k^2 \psi_w = 0, \quad k^2 = -\frac{\mu_2}{\nu}. \tag{4.22a,b}$$

The solution satisfying the boundary conditions $\psi_w(x, 0) = \psi_w(x, 1) = 0$, as well as the periodicity in x, is given by

$$\psi_w = B \sin k_x x \ \sin k_y y, \quad k^2 = k_x^2 + k_y^2 \tag{4.23a,b}$$

with eigenvalues

$$k_x = 2n_2 \pi, \quad k_y = n_3 \pi \ (n_2, \ n_3 \in \mathbb{Z}). \tag{4.24a,b}$$

The corresponding wavy vorticity is

$$\omega_w = Bk^2 \sin k_x x \ \sin k_y y. \tag{4.25}$$
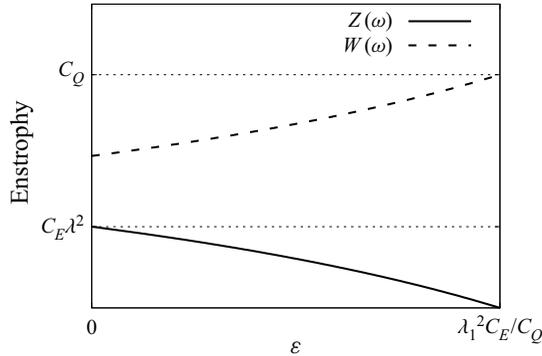
Figure 1. The graph of $Z(\omega)$ and $W(\omega)$ given in (4.30) and (4.31).

Summing the zonal and wavy components, we obtain

$$\psi = A \sin \lambda y + B \sin k_x x \, \sin k_y y, \tag{4.26}$$

$$\omega = A\lambda^2 \sin \lambda y + Bk^2 \sin k_x x \, \sin k_y y. \tag{4.27}$$

The two amplitudes A and B are determined by the constraints $E(\omega) = C_E$ and $Q(\omega) = C_Q$; inserting (4.26) and (4.27) into the definitions of $E(\omega)$ and $Q(\omega)$, we obtain

$$C_E = \frac{A^2\lambda^2}{4} + \frac{B^2k^2}{8}, \tag{4.28}$$

$$C_Q = \frac{A^2\lambda^4}{4} + \frac{B^2k^4}{8}. \tag{4.29}$$

Solving (4.28) and (4.29) for A and B, and inserting the solution into the zonal enstrophy $Z(\omega)$ and wavy enstrophy $W(\omega)$, we obtain the critical values

$$Z_{\lambda,\epsilon} = \frac{\lambda^2 C_E - \epsilon C_Q}{1 - \epsilon}, \tag{4.30}$$

$$W_{\lambda,\epsilon} = \frac{C_Q - \lambda^2 C_E}{1 - \epsilon}, \tag{4.31}$$

where $\epsilon = \lambda^2/k^2$, scaling the ratio of the wavelength of the zonal components to that of the wavy components. For $Z_{\lambda,\epsilon} \geq 0$ and $W_{\lambda,\epsilon} \geq 0$, there are two possibilities: $\epsilon \leq (\lambda^2 C_E)/C_Q \leq 1$ or $\epsilon \geq (\lambda^2 C_E)/C_Q \geq 1$. Here, the former regime of ϵ is relevant, because we assume that the wavy components have smaller scales in comparison with the zonal component (i.e. $\epsilon < 1$). Then, $Z_{\lambda,\epsilon}$ of (4.30) increases monotonically as ϵ decreases (or k^2 increases; see figure 1), and we have

$$\lim_{\epsilon \to 0} Z_{\lambda,\epsilon} = \lambda^2 C_E. \tag{4.32}$$

Notice that this limit gives the upper bound for $Z(\omega)$ of the corresponding eigenvalue λ, which is achieved when the wavy component has the smallest scale $\epsilon \to 0$. For actual wavy components, $Z(\omega)$ takes a smaller value than $\lambda^2 C_E$, i.e.

$$Z(\omega) \leq \lambda^2 C_E. \tag{4.33}$$

919 A33-10

4.4. *Constraints by energy, circulation, impulse and total enstrophy*

Now we study the minimum of the zonal enstrophy $Z(\omega)$ under all constraints of energy, circulation, impulse and total enstrophy. In contrast to the observation of § 4.3 (where the minimum of $Z(\omega)$ is not determined by the energy C_E), we will find that the minimum of $Z(\omega)$ is determined by the circulation C_F and impulse C_L. In comparison with the result of § 4.2, however, we have a discrete set of enstrophy levels (each of which corresponds to a different mode number). Whereas they are due to the energy constraint, $Z(\omega)$ itself does not depend on the values of the energy C_E.

Introducing Lagrange multipliers, we seek the minimizer of

$$Z(\omega) - \nu Q(\omega) - \mu_0 F(\omega) - \mu_1 L(\omega) - \mu_2 E(\omega). \tag{4.34}$$

The Euler–Lagrange equation is

$$\mathcal{P}_z\omega - \nu\omega - \mu_0 - \mu_1 y - \mu_2 \mathcal{K}\omega = 0. \tag{4.35}$$

The solution satisfying the boundary conditions $\psi(x, 0) = \psi(x, 1) = 0$, as well as the periodicity in x, is $\psi = \psi_z + \psi_w$ with

$$\psi_z = A_1 \cos \lambda y + A_2 \sin \lambda y - \frac{\mu_0 + \mu_1 y}{\mu_2}, \tag{4.36}$$

$$\psi_w = B \sin k_x x \, \sin k_y y, \tag{4.37}$$

where

$$\lambda = \sqrt{\frac{\mu_2}{1 - \nu}}, \quad k^2 = k_x^2 + k_y^2 = -\frac{\mu_2}{\nu}, \tag{4.38a,b}$$

and

$$k_x = 2n_2\pi, \quad k_y = n_3\pi \ (n_2, \ n_3 \in \mathbb{Z}). \tag{4.39a,b}$$

The corresponding vorticities are

$$\omega_z = A_1\lambda^2 \cos \lambda y + A_2\lambda^2 \sin \lambda y. \tag{4.40}$$

$$\omega_w = Bk^2 \sin k_x x \, \sin k_y y. \tag{4.41}$$

The zonal enstrophy $Z(\omega)$ of the minimizer is

$$Z(\omega) = \frac{A_1^2\lambda^3}{8}(2\lambda + \sin 2\lambda) + \frac{A_2^2\lambda^3}{8}(2\lambda - \sin 2\lambda)$$

$$+ \frac{A_1 A_2\lambda^3}{4}(1 - \cos 2\lambda). \tag{4.42}$$

We have yet to determine the eigenvalue λ and the coefficients A_1, A_2 and B. Inserting $\psi = \psi_z + \psi_w$ and $\omega = \omega_z + \omega_w$ into the constraints $F(\omega) = C_F, L(\omega) = C_L, E(\omega) = C_E,$

and $Q(\omega) = C_Q$, we obtain

$$C_F = A_1 \lambda \sin \lambda + A_2 \lambda (1 - \cos \lambda), \tag{4.43}$$

$$C_L = A_1 (\lambda \sin \lambda + \cos \lambda - 1) + A_2 (\sin \lambda - \lambda \cos \lambda), \tag{4.44}$$

$$
\begin{aligned}
C_E = {} & \frac{A_1^2 \lambda}{8}(2\lambda + \sin 2\lambda) + \frac{A_2^2 \lambda}{8}(2\lambda - \sin 2\lambda) \\
& + \frac{A_1 A_2 \lambda}{4}(1 - \cos 2\lambda) - \frac{A_1 C_F}{2} \\
& - \frac{[A_1(\cos \lambda - 1) + A_2 \sin \lambda]C_L}{2} + \frac{B^2 k^2}{8},
\end{aligned} \tag{4.45}
$$

$$
\begin{aligned}
C_Q = {} & \frac{A_1^2 \lambda^3}{8}(2\lambda + \sin 2\lambda) + \frac{A_2^2 \lambda^3}{8}(2\lambda - \sin 2\lambda) \\
& + \frac{A_1 A_2 \lambda^3}{4}(1 - \cos 2\lambda) + \frac{B^2 k^4}{8}.
\end{aligned} \tag{4.46}
$$

We may write (4.43) and (4.44) as

$$\begin{pmatrix} C_F \\ C_L \end{pmatrix} = \boldsymbol{D}(\lambda) \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. \tag{4.47}$$

with

$$\boldsymbol{D}(\lambda) := \begin{pmatrix} \lambda \sin \lambda & \lambda(1 - \cos \lambda) \\ \lambda \sin \lambda + \cos \lambda - 1 & \sin \lambda - \lambda \cos \lambda \end{pmatrix}. \tag{4.48}$$

For given C_F and C_L, we solve (4.47) to determine the amplitudes of zonal vorticity

$$A_1 = \frac{C_F(\sin \lambda - \lambda \cos \lambda) + C_L(-\lambda + \lambda \cos \lambda)}{\det D(\lambda)}, \tag{4.49}$$

$$A_2 = \frac{C_F(-\lambda \sin \lambda - \cos \lambda + 1) + C_L \lambda \sin \lambda}{\det D(\lambda)}, \tag{4.50}$$

where $\det D(\lambda) = \lambda(2 - \lambda \sin \lambda - 2 \cos \lambda)$. Inserting (4.49) and (4.50) into (4.42), we obtain the zonal enstrophy evaluated as a function of λ, which we denote by Z_λ. The critical points (local minima) of $Z(\omega)$, given by

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} Z_\lambda = 0, \tag{4.51}$$

determine the eigenvalues λ characterizing the enstrophy levels.

Instead of displaying the lengthy expression of Z_λ, we will show its graphs for typical choices of the parameters C_F and C_L. Notice that Z_λ depends only on C_F (circulation) and C_L (impulse); it does not contain C_E (energy) and C_Q (enstrophy) as parameters. First, we pay attention to the singularities given by $\det D(\lambda) = 0$, where $A_1 \to \infty$ and $A_0 \to \infty$, hence $Z_\lambda \to \infty$ (there is an exception, as discussed later). We show the graph of $\det D(\lambda)$ in figure 2.

919 A33-12
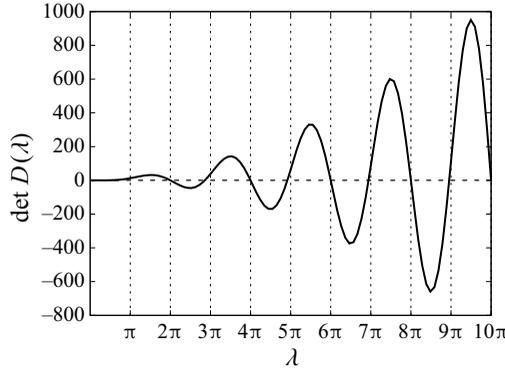
Figure 2. The graph of det$D$.
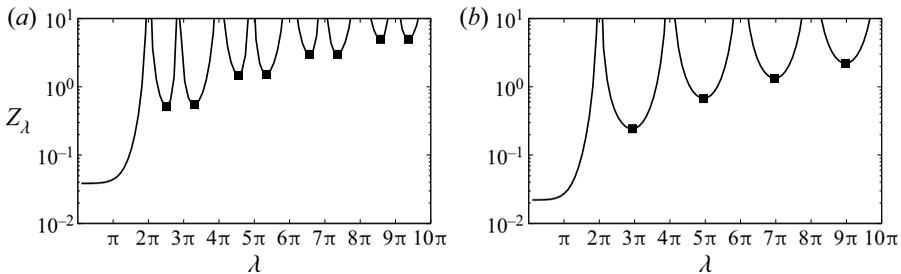
Figure 3. The graph of the critical zonal enstrophy Z_λ as a function of λ. The minima of Z_λ determine the eigenvalues of λ. The points on the curve indicate the eigenvalues. We assume (*a*) $C_F = 0.21$ and $C_L = 0.0525$, and (*b*) $C_F = 0.21$ and $C_L = 0.105$.

There are two types of solutions

$$\lambda = \begin{cases} \Lambda_{2n} = 2n\pi, \\ \Lambda_{2n+1} = (2n+1)\pi - \delta_n, \end{cases} \quad (n = 0, 1, \ldots), \tag{4.52}$$

where each δ_n is a small positive number such that $\delta_n \to 0$ as $n \to \infty$. The minima of Z_λ appear in every interval $(\Lambda_{2n}, \Lambda_{2n+1})$. However, if $C_F = 2C_L$, Z_λ remains finite at $\lambda = \Lambda_{2n+1}$. In this special case, the minima of Z_λ appear in intervals $(\Lambda_{2n}, \Lambda_{2n+2})$.

In figure 3, we show examples of Z_λ calculated for (3*a*) $C_F = 0.28$ and $C_L = 0.07$, (3*b*) $C_F = 0.28$ and $C_L = 0.14$ ($C_F = 2C_L$).

At $\lambda = 0$, Z_λ reproduces the result of Theorem 4.1, i.e.

$$\lim_{\lambda \to 0} Z_\lambda = Z_0 = 2C_F^2 - 6C_F C_L + 6C_L^2, \tag{4.53}$$

which is the absolute minimum of the zonal enstrophy under the constraints on the circulation $F(\omega) = C_F$, the impulse $L(\omega) = C_L$ and the total enstrophy $Q(\omega) = C_Q$.

The role of the energy constraint $E(\omega) = C_E$ is to create eigenvalues of λ at which Z_λ takes local minimum values. However, the value of C_E does not influence the value of Z_λ directly. As we have seen in (4.33), it poses a constraint on the maximum

$$Z(\omega) \leq \lambda^2 C_E, \tag{4.54}$$

in addition to the other implicit constraint $Z(\omega) \leq C_Q$. Instead of the zonal component ω_z of (4.40), C_E and C_Q work for determining the complementary wavy component ω_w

of (4.41). By (4.45) and (4.46), we obtain

$$k^2 = \frac{C_Q - Z_\lambda}{C_E - E_{z,\lambda}}, \tag{4.55}$$

$$B^2 = \frac{8(C_E - E_{z,\lambda})^2}{C_Q - Z_\lambda}, \tag{4.56}$$

where $E_{z,\lambda}$ is the energy of the zonal component ω_z evaluated at the eigenvalue λ. Notice that $k^2 B^2$ ($\sim$ energy of the wavy component) is determined only by C_E and $E_{z,\lambda}$. So, the role of the total enstrophy constraint is to determine the wavenumber k of the wavy component.

For the special case of $C_F = 0$ and $C_L = 0$, a laminated zonal flow ($A_1 \neq 0$ and/or $A_2 \neq 0$) can occur only if

$$\det D(\lambda) = \lambda(2 - \lambda \sin \lambda - 2 \cos \lambda) = 0. \tag{4.57}$$

Then, the eigenvalues are $\lambda = \Lambda_{2n}$ and Λ_{2n+1} ($n = 0, 1, 2, \ldots$), the previous singular points; see figure 2. For $\lambda = \Lambda_{2n}$ ($\lambda = 0$ gives the trivial solution $\omega_z = 0$),

$$D(\lambda) = \begin{pmatrix} 0 & 0 \\ 0 & -\lambda \end{pmatrix}, \tag{4.58}$$

hence, $A_2 = 0$. On the other hand, for $\lambda = \Lambda_{2n+1}$,

$$D(\lambda) = \frac{1}{4} \lambda \sin \lambda \begin{pmatrix} 4 & 2\lambda \\ 2 & \lambda \end{pmatrix}, \tag{4.59}$$

and then $A_2 = -2A_1/\Lambda_{2n}$. In both cases, A_1 is arbitrary, so we cannot determine the amplitude of the zonal vorticity ω_z. Therefore, the trivial conditions $C_F = 0$ and $C_L = 0$ reproduce the situation of 'no constraint' discussed in § 4.3. We only have the estimate of the maximum (4.33).

The forgoing results are summarized as:

THEOREM 4.2. *For a given set of constants $F(\omega) = C_F$, $L(\omega) = C_L$, $E(\omega) = C_E$ and $Q(\omega) = C_Q$, the zonal enstrophy $Z(\omega)$ has a discrete set of critical (local minimum) values quantized by the eigenvalue λ measuring the mode number of the zonal vorticity.*

  (i) *When $C_F \neq 0$ or $C_L \neq 0$, the eigenvalue λ is given by (4.51) as a function of C_F and C_L. The corresponding eigenfunction ω_z, and the critical value of $Z(\omega)$ are determined by C_F and C_L; see (4.40), (4.42), (4.49) and (4.50). The other constants C_E and C_Q determine upper bounds $C_E \lambda^2 \geq Z(\omega)$ and $C_Q \geq Z(\omega)$.*
  (ii) *For the special values $C_F = C_L = 0$, additional eigenvalues $\lambda = 2n\pi$ and $\lambda = \Lambda_n$ ($n = 1, 2, \ldots$) occur. However, the eigenfunctions ω_z and the critical values of $Z(\omega)$ are no longer determined by such C_F and C_L; we only have estimates of upper bounds $C_E \lambda^2 \geq Z(\omega)$ and $C_Q \geq Z(\omega)$.*

4.5. *Determination of the zonal enstrophy level*

To apply Theorem 4.2 to the estimation of attainable zonal enstrophy, we have to determine the eigenvalue λ that identifies the zonal enstrophy level. Here, we suggest the following method (which we will examine and improve in § 5).

The self-organization of zonal flow can be seen as a relaxation process of the zonal enstrophy level, which parallels the inverse cascade in the meridional wavenumber space.

Just as the transition of the quantum energy level is caused by photon emission, the relaxation of the zonal enstrophy level is due to the emission of wavy vorticity, which is driven by the nonlinear coupling of the zonal and wavy components. Therefore, the relaxation can proceed as far as the nonlinear term $\{\omega, \psi\}$ dominates the evolution equation (2.5). Relative to the concomitant linear term $\beta\{y, \psi\}$, the nonlinear term becomes weaker as the length scale increases (i.e. the inverse cascade proceeds). On the Rhines scale (Rhines 1975)

$$L_R = \sqrt{\frac{2U}{\beta}}, \tag{4.60}$$

the linear and nonlinear terms have comparable magnitudes, where U is the representative magnitude of the zonal-flow velocity.

Since the energy is conserved, we may estimate $U = \sqrt{2C_E}$. Hence, we have an *a priori* estimate

$$\lambda \sim \frac{\pi}{L_R} = \pi\sqrt{\frac{\beta}{2\sqrt{2C_E}}}. \tag{4.61}$$

Notice the influence of the energy C_E on the eigenvalue λ. Although each value of the zonal enstrophy level is independent of C_E, the selection of the level is made by C_E.

In the next section, we will examine the theoretical estimates by comparing numerical simulation results.

5. Comparison with numerical simulations

5.1. *Simulation model*

In this section, we compare the forgoing theoretical estimates with numerical simulation results. Here, we consider a square domain as with § 4. With a system size L and a rotation period T, we normalize the variables as

$$\check{x} = \frac{x}{L}, \quad \check{y} = \frac{y}{L}, \quad \check{t} = \frac{t}{T}, \quad \check{\omega} = \omega T, \quad \check{\psi} = \frac{\psi T}{L^2}, \tag{5.1a–e}$$

by which the vorticity equation reads

$$\partial_{\check{t}}\check{\omega} + \{\check{\omega} + \beta\check{y}, \check{\psi}\} = \nu\Delta\check{\omega}, \tag{5.2}$$

where ν represents the viscosity (reciprocal Reynolds number). We remark that, although the dynamics of the system changes under rescaling of the coordinates with a general aspect ratio of the domain, the statements of Theorems 4.1 and 4.2 remain unchanged.

For simplicity, we will omit the normalization symbol ˇ in the following description. Whereas our theoretical analysis is based on the dissipation-free model (2.5), we add a finite viscosity ν for numerical stability (typically, we put $\nu = 1.0 \times 10^{-6}$). A finite viscosity is also indispensable for the self-organization process, because the ideal (zero viscosity) dynamics is constrained by infinite number of Casimirs (local circulations), preventing changes in streamline topology. The theoretical model, however, ignores the dissipation by assuming the robustness of the invariants that are used as constraints (see Proposition 2.1). The influence of dissipation will be examined carefully when we compare the theory and numerical simulation.
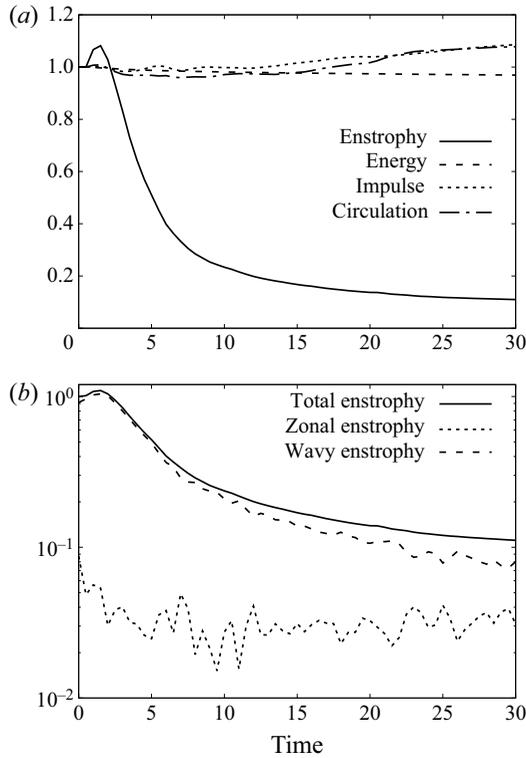
Figure 4. (*a*) The evolution of the 'ideal' constants in the simulation. Each value is normalized by the corresponding initial value. (*b*) The partition of total enstrophy between zonal and wavy components.

In the following simulation, we assume parameters comparable to the Jovian atmosphere, where $L = 4.4 \times 10^8$ m, $T = 8.6 \times 10^5$ s. The parameter β is determined as

$$\beta = \frac{2\Omega}{R}(\cos\theta)LT, \tag{5.3}$$

where Ω is the angular vorticity of rotating frame, R is the radius and θ is the latitude. For $L \sim 2\pi R$ and $\theta \sim 0$, we obtain $\beta \sim 10^2$. The jet velocity reaches $U \sim 1 \times 10^2$ m s^{-1}, which yields $C_F \sim 4 \times 10^{-1}$ and $C_L \sim 2 \times 10^{-1}$ if the jet achieves the maximum opposite velocities on both north and south boundaries. Here, we assume moderate values $C_F \sim 10^{-1}$ and $C_L \sim 10^{-1}$.

5.2. *Self-organized zonal flow*

As we have seen, the theoretical estimate of the minimum $Z(\omega)$ changes dramatically depending on whether C_F and C_L are finite or not (§ 4.4). Here, we study the general case where both C_F and C_L are finite. We assume an initial condition such that

$$\omega|_{t=0} = 5.0 \sin 15\pi y + \sum_{m,n} \alpha_{mn} e^{imx} \sin n\pi y, \tag{5.4}$$

with random $\alpha_{mn}(|\alpha_{mn}| \in [0, 50)$ for $5 \leq m, n \leq 10)$, which yields $C_E = 3.6 \times 10^{-2}$, $C_F = 0.21$ and $C_L = 0.105$.

919 A33-16
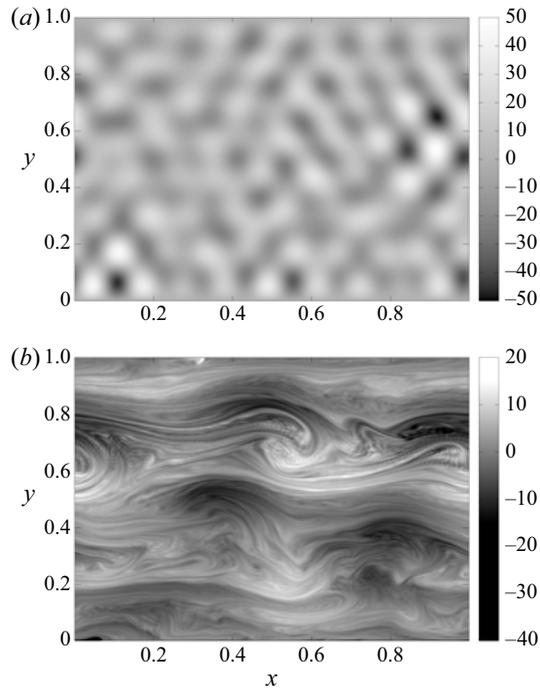
Figure 5. Self-organization of zonal flow (grey level represents to the local value of ω). (a) Initial condition with finite circulation $C_F = 0.21$ and impulse $C_L = 0.11$. (b) Creation of zonal flow observed at $t = 20$.

In figure 4, we show the evolution of the 'ideal' constants. The total energy C_E is well conserved. The changes in C_F and C_L are also tolerable. Because of a finite viscosity ($\nu = 1.0 \times 10^{-6}$), however, the total enstrophy C_Q changes significantly. But it is not essential for the present purpose of comparison, because the theoretical estimate of minimum $Z(\omega)$ is independent of the C_Q. As noted after (4.55)–(4.56), the total enstrophy $Q(\omega) = C_Q$ only contributes to estimating the wavenumber k of the wavy component ω_w. As the simulation shows, the 'dissipation' of the total enstrophy is the signature of the relaxation, when we consider a finite viscosity. We may interpret the dissipation as the scale separation between the visible scale and micro-scale; the latter is separated from the vortex dynamics model by suppressing the amplitudes of micro-scale vortices. This scenario is consistent with the local interaction model; the nonlinear dynamics is dominated by interactions among similarly sized vortices (i.e. local in the Fourier space) within the inertial range, so it is not influenced by vortices of far smaller scales.

Figure 5 shows the self-organized state ($t = 20$), where an appreciable zonal component manifests. In figure 6, we compare the Fourier spectrum of the zonal component $\omega_z = \mathcal{P}_z \omega$ in the initial and self-organized states. We find the redistribution of the spectrum into lower λ modes (i.e. inverse cascade). A comparison with the Rhines scale will be described later.

To make a comparison with the theoretical estimate of zonal enstrophy, we plot Z_λ (the theoretical minimum of zonal enstrophy) and $C_E \lambda^2$ (the theoretical maximum of zonal enstrophy), evaluated for the parameters determined by the given initial condition, in figure 7. As $\lambda = 5\pi$ is the dominant mode (figure 6), we obtain $Z_\lambda = 0.69$ and $C_E \lambda^2 = 8.8$. In figure 8, we compare the simulation results and the theoretical estimates, demonstrating that the actual zonal enstrophy $Z(\omega)$ stays between the theoretical minimum and maximum; the estimate of the lower bound is reasonably accurate.
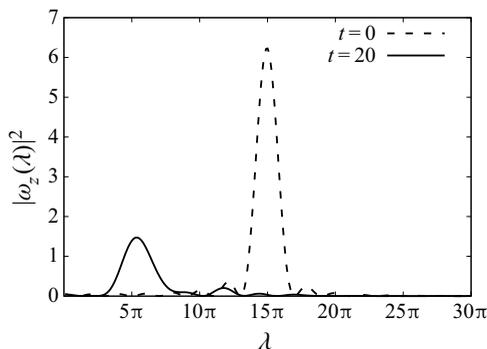
Figure 6. The Fourier spectrum of the zonal vorticity $\omega_z = \mathcal{P}_z \omega$ in the self-organized state ($t = 20$). The eigenvalue $\lambda \sim 5\pi$ is dominant.
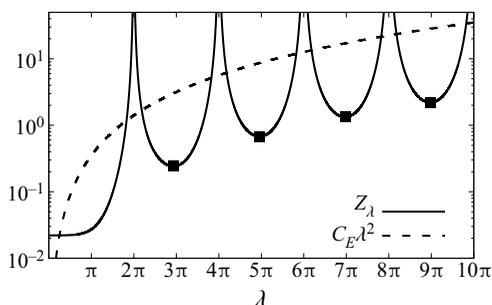
Figure 7. The graphs of Z_λ and $C_E \lambda^2$, evaluated for the parameters corresponding to the simulation of figure 5. The points on the curve indicate the eigenvalues.
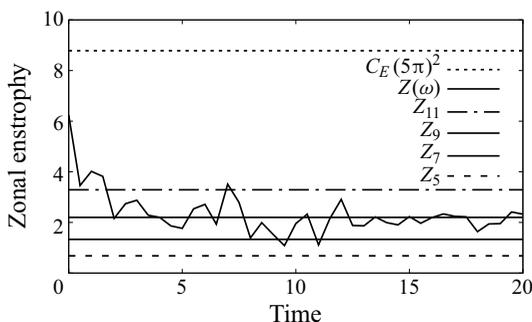
Figure 8. Evolution of the zonal enstrophy $Z(\omega)$, and its comparison with the theoretical minimum Z_λ and the maximum $C_E \lambda^2$ evaluated for the self-organized state $\lambda \sim 5\pi$. To demonstrate the sensitivity of the minimum value, we also show the theoretical minimum Z_λ evaluated for $\lambda \sim 5\pi$, 7π, 9π and 11π.

5.3. *Improved Rhines scale*

The forgoing discussion depends on the *a posteriori* estimate of the eigenvalue λ. As discussed in § 4.5, however, we need an *a priori* estimate of λ to make the theory useful. While the Rhines scale L_R has been proposed to estimate $\lambda \sim \pi/L_R$, it turns out to be too crude. Here, we propose an improved Rhines scale to make more accurate estimates. Figure 9 compares the dominant scale in the final state obtained by simulation and the
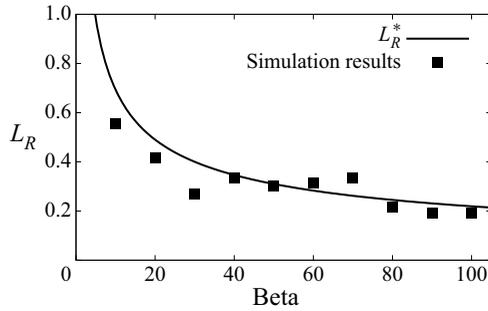
Figure 9. The comparison between the dominant scale in the self-organized state and the Rhines scale for different values of β.

Rhines scale for different values of β. It is shown that the dominant scale is approximately 3 times of the Rhines scale.

The Rhines scale (4.60) is the length scale L_R at which the magnitudes of the nonlinear term $\{\omega, \psi\}$ and the linear term $\beta\{y, \psi\}$ become comparable. However, it seems that the function of the nonlinear term, that derives the relaxation of the enstrophy level λ, does not end immediately at L_R; the numerical experiment shows that the relaxation continues up to $\sim 3 \times L_R$, where the magnitude of the nonlinear term becomes approximately one eighth of the linear term. Therefore, we propose to use $L_R^* = 3L_R$ for the *a priori* estimate $\lambda = \pi/L_R^*$; modifying (4.61), we estimate

$$\lambda \sim \frac{\pi}{3}\sqrt{\frac{\beta}{2\sqrt{2C_E}}}. \tag{5.5}$$

6. Conclusion

We have found a discrete set of zonal enstrophy levels that are quantized by the eigenvalue λ measuring the mode number (= system size in the latitude direction/lamination number). As shown in figure 3, a finite circulation C_F and/or impulse C_L brings about symmetry breaking in the eigenstates (minimizers), inhibiting even mode numbers. In the actual situation, however, the mixed state may include spectra of even mode numbers (see figure 6). By comparing to simulation results, we verified that the theoretical value Z_λ gives a reasonable estimate of the zonal enstrophy, if we choose the relevant mode number. We note that the enstrophy levels are determined independently of the selection mechanism. Just as the quantum energy level of an orbital electron is lowered by photon emission (see figure 10), the relaxation of the zonal enstrophy level proceeds by the emission of short-scale wavy vorticity. The relaxation process can be viewed as the forward cascade of enstrophy (creation of short-scale wavy vortices) and the simultaneous inverse cascade of the energy spectrum (de-excitation to lower zonal enstrophy states). The energy constraint plays an essential role in selecting the level; the relaxation continues as far as the nonlinear term, measured by the energy, dominates the evolution. The Rhines scale estimates the balance point of the magnitudes of the nonlinear term and the linear Rossby-wave term, but we found that the nonlinear effect continues to work until it becomes approximately one order of magnitude smaller than the linear term, so we propose an improved Rhines scale.

Comparing Theorems 4.1 and 4.2, we find that the energy constraint $E(\omega) = C_E$ plays an essential role in creating the discrete zonal enstrophy levels Z_λ. Interestingly, the value
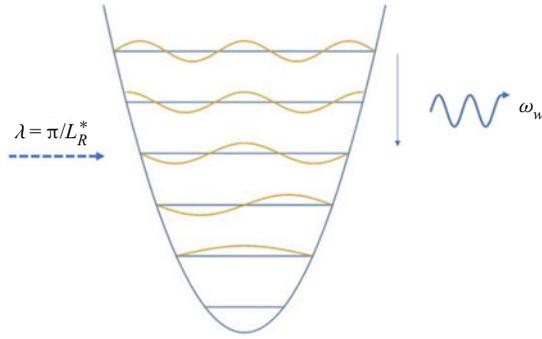
Figure 10. Analogy of quantum energy levels and 'de-excitation' by emitting small-scale wavy enstrophy, which parallels the forward cascade of enstrophy.

of C_E does not influence the value of each zonal enstrophy Z_λ, which is determined only by the other constants C_F (circulation) and C_L (impulse). However, in the absence of the energy constraint, we only have the 'ground state' $\lambda = 0$ as given in Theorem 4.1. In the eigenstate of ω_z (belonging to the eigenvalue λ), the zonal enstrophy $Z(\omega_z)$ and the zonal energy $E(\omega_z)$ are related by $Z(\omega_z) = \lambda^2 E(\omega_z)$. Under the energy constraint (and a fixed λ), therefore, $Z(\omega_z)$ may take a smaller value when the wavy component ω_w shares a larger energy $E(\omega_w)$. The simultaneous total enstrophy constraint contributes to determining the wavenumber k of the wavy component ω_w. Without the symmetry breaking constraints by the circulation $F(\omega)$ and the impulse $L(\omega)$, $E(\omega_z)$ can minimize to zero (see figure 1), and then, $k \to \infty$ (independently of the specific value of C_Q). Finite symmetry breaking by C_F and/or C_L brings about a non-trivial minimum $E(\omega_z)$ (and the corresponding $Z(\omega_z) = \lambda^2 E(\omega_z)$). Then, the partition of the energy to the wavy component is determined as $E(\omega_w) = C_E - E(\omega_z)$, and the wavenumber k is determined by $k^2 E(\omega_w) = C_Q - Z(\omega_z)$. Interestingly, we may not remove the total enstrophy constraint from the variational principle, in order to retain a finite wavy component as the complementary to the zonal component, while its role is limited to characterizing only the wavy component. This unusual phenomenon in the variational principle is caused by the non-coerciveness of the target functional $Z(\omega)$ with respect to the norm $\|\omega\|$.

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.
ⓘ H. Aibara https://orcid.org/0000-0002-0783-0838;
ⓘ Z. Yoshida https://orcid.org/0000-0002-9165-6616.

Appendix A. The ABC of the variational principle

To see the mathematical non-triviality of the variational principle for the zonal enstrophy, we review the standard relation between the target functional and constraint.

A.1. *Target functional and constraint*

We start with a textbook example. The isoperimetric problem is to (i) maximize the surface area S with a constraint on the periphery length L, or (ii) minimize the peripheral length L with a constraint on the surface area S. Both problems have the same solution, i.e. a circular disk or its periphery. Notice that reversing the target and constraint in each setting results in an ill-posed problem; one can make L infinitely long without changing S, or one can make S infinitely small without changing L. Let us concentrate on minimization problems. For a variational principle to be well posed, the target (L) must be 'fragile' and the constraint (S) must be robust. Here, the fragility speaks of the sensitivity to small-scale perturbations. Suppose that we make pleats on a periphery; then L is increased, but S is not necessarily changed. In analytical formalism, a fragile functional includes a larger number of differentiations – derivatives are sensitive to small-scale perturbations. In the forgoing example, we may formally write

$$S = \int_{\mathbb{R}^2} \mathbb{I}_M \mathrm{d}^2 x, \quad L = \int_{\mathbb{R}^2} |\nabla \mathbb{I}_M| \mathrm{d}^2 x, \quad \mathbb{I}_M(x) = \begin{cases} 1 & \text{if } \mathbf{x} \in M \\ 0 & \text{if } \mathbf{x} \notin M, \end{cases} \tag{A1a–c}$$

where M is a simply connected domain $\subset \mathbb{R}^2$ that should be optimized to minimize L for some given value of S. Including ∇ in the integrand, L is fragile.

A.2. *Coerciveness and continuity*

To make the argument more precise, we introduce the notion of coercive functionals; cf. Yoshida & Mahajan (2002); Lions & Magenes (2012). Let u be a real-valued function (state vector) belonging to a function space (phase space) V, which is a Banach space with a norm $\|u\|$. A real-valued functional $G(u)$ is said to be coercive if

$$\|u\|^2 \le c\, G(u), \tag{A2}$$

where c is some positive constant. On the other hand, a real-valued functional $H(u)$ is continuous if

$$|H(u + \delta) - H(u)| \to 0 \quad (\|\delta\| \to 0). \tag{A3}$$

We can formulate a well-posed minimization problem with a coercive target functional $G(u)$ and a continuous constraining functional $H(u)$ (we may also consider multiple constraints with continuous functionals).

To see how the coerciveness and continuity influence variational principles, let us consider an example with two functionals

$$G(u) = \int_M |\nabla u(\mathbf{x})|^2 \, \mathrm{d}^n x, \quad H(u) = \int_M |u(\mathbf{x})|^2 \mathrm{d}^n x, \tag{A4a,b}$$

where u is a scalar function defined in a smoothly bounded open set $M \subset \mathbb{R}^n$. We assume that $u = 0$ on the boundary ∂M. Notice that $H(u)^{1/2}$ is the L^2 norm $\|u\|$. Therefore, $H(u)$ is a continuous functional on the function space $V = L^2(M)$. By the Poincaré inequality, we have

$$\|u\|^2 \le c\|\nabla u\|^2 = cG(u), \tag{A5}$$

with a positive constant c. Therefore, $G(u)$ is a coercive functional.

First, we seek a minimizer of $G(u)$ with the constraint $H(u) = 1$. This is a well-posed problem. The minimizer is found by the variational principle

$$\delta[G(u) - \lambda H(u)] = 0, \tag{A6}$$

where λ is a Lagrange multiplier. The Euler–Lagrange equation

$$- \Delta u = \lambda u, \tag{A7}$$

together with the above-mentioned boundary condition, constitute an eigenvalue problem. We can easily show that every eigenvalue λ is positive. Let λ_j be an eigenvalue and φ_j be the corresponding normalized eigenfunction ($\|\varphi_j\|^2 = 1$). Setting $u = a\varphi_j$ and demanding $H(u) = 1$, we obtain $a = 1$ and $G(u) = \lambda_j$. The smallest λ_j, then, yields the minimum $G(u)$.

The reverse problem of finding a minimizer of $H(u)$ with the restriction $G(u) = 1$ is ill posed, because the constraint is posed by a functional $G(u)$ that is not continuous in the topology of $L^2(M)$. Let us elucidate the pathology. The variational principle $\delta[H(u) - \mu G(u)] = 0$ (μ is a Lagrange multiplier) yields the Euler–Lagrange equation $-\Delta u = \mu^{-1} u$. Let $\mu^{-1} = \lambda_j$ (an eigenvalue of $-\Delta$), and $u = a\varphi_j$. The condition $G(u) = 1$ yields $a = \lambda_j^{-1/2}$, and $H(u) = 1/\lambda_j$. Hence, the minimum of $H(u)$ is achieved by the largest eigenvalue that is unbounded, viz., $\inf H(u) = 0$ and the minimizer $\lim_{\lambda_j \to \infty} \lambda_j^{-1/2} \varphi_j = 0$ is nothing but the minimizer of $H(u)$ without any restriction. The constraint $G(u) = 1$ plays no role in this minimization problem.

A.3. *Non-coercive target functional*

Let us modify the target functional of (A6) to a non-coercive functional. Let $V_k = \mathrm{span}\,\{\varphi_1, \ldots, \varphi_k\}$, which is a closed (finite-dimensional) subspace of $V = L^2(M)$. We denote the orthogonal complement by V', i.e. we decompose $V = V_k \oplus V'$. Let $\mathcal{P}$ be the orthogonal projector $V \to V'$. Consider

$$G'(u) = \|\nabla (\mathcal{P}u)\|^2 = \int_M (-\Delta \mathcal{P}u)(\mathcal{P}u)\,\mathrm{d}^n x = \sum_{j>k}^{\infty} \lambda_j (u, \varphi_j)^2, \tag{A8}$$

where $(f, g) = \int_M f(x)\,g(x)\,\mathrm{d}^n x$ is the inner product of $L^2(M)$. Evidently, $G'(u)$ is not coercive. The modified variational principle

$$\delta[G'(u) - \lambda H(u)] = 0 \tag{A9}$$

yields the Euler–Lagrange equation that reads, after expanding with eigenfunctions,

$$\lambda_j'(u, \varphi_j) = \lambda(u, \varphi_j) \quad (j = 1, 2, \ldots), \tag{A10}$$

where the 'modified eigenvalues' are

$$\lambda_j' = \begin{cases} 0, & (j = 1, \ldots, k), \\ \lambda_j, & (j > k). \end{cases} \tag{A11}$$

The minimizer of $G'(u)$ is a solution of (A10) such that $\lambda = 0$ and

$$u = \sum_{j=1}^{k} a_j \varphi_j, \tag{A12}$$

where constants $a_1, \ldots, a_k$ can be arbitrarily chosen provided that $\sum_{j=1}^{k} |a_j|^2 = 1$ in order to satisfy the constraint $H(u) = 1$. We obtain $\min G' = 0$, but the minimizer is not a unique function.

This prototypical example elucidates the essence of the pathology created in a variational principle with non-coercive target functional. We encounter a similar non-uniqueness (degeneracy) problem in § 4.1. Interestingly, however, the energy constraint brings about a dramatical change in the mathematical structure, and removes the degeneracy; § 4.3.

REFERENCES

BAKAS, N.A. & IOANNOU, P.J. 2011 Structural stability theory of two-dimensional fluid flow under stochastic forcing. *J. Fluid Mech.* **682**, 332–361.

BALK, A.M. 1991 New invariant for Rossby wave systems. *Phys. Lett.* A **155**, 20–24.

BALK, A.M. 2005 Angular distribution of Rossby wave energy. *Phys. Lett.* A **345**, 154–160.

BALK, A.M., NAZARENKO, S.V. & ZAKHAROV, V.E. 1991 New invariant for drift turbulence. *Phys. Lett.* A **152**, 276–280.

BRETHERTON, F.P. & HAIDVOGEL, D.B. 1976 Two-dimensional turbulence above topography. *J. Fluid Mech.* **78** (1), 129–154.

CHARNEY, J.G. 1971 Geostrophic turbulence. *J. Atmos. Sci.* **28**, 1087–1095.

CONNAUGHTON, C., NADIGA, B., NAZARENKO, S. & QUINN, B. 2010 Modulational instability of Rossby and drift waves and generation of zonal jets. *J. Fluid Mech.* **654**, 207–231.

DANILOV, S. & GURARIE, D. 2002 Rhines scale and spectra of the beta-plane turbulence with bottom drag. *Phys. Rev.* E **65**, 067301.

DIAMOND, P.H., ITOH, S.-I., ITOH, K. & HAHM, T.S. 2005 Zonal flows in plasma – a review. *Plasma Phys. Control. Fusion.* **47**, R35–R161.

DRITSCHEL, D.G. & MCINTYRE, M.E. 2008 Multiple jets as PV staircases: the phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.* **65**, 855–874.

FARRELL, B.F. & IOANNOU, P.J. 2007 Structure and spacing of jets in barotropic turbulence. *J. Atmos. Sci.* **64**, 3652–3665.

GILL, A.E. 1974 The stability of planetary waves on an infinite beta-plane. *Geophys. Fluid Dyn.* **6**, 29–47.

HASEGAWA, A. 1985 Self-organization processes in continuous media. *Ad. Phys.* **34**, 1–42.

ISHIOKA, K. & YODEN, S. 1996 Numerical methods of estimating bounds on the non-linear saturation of barotropic instability. *J. Met. Soc. Japan* **74**, 167–174.

ITO, N. & YOSHIDA, Z. 1996 Statistical mechanics of magnetohydrodynamics. *Phys. Rev.* E **53**, 5200–5206.

KRAICHNAN, R.H. 1967 Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10**, 1417–1423.

LIONS, J.L. & MAGENES, E. 2012 *Non-Homogeneous Boundary Value Problems and Applications*, vol. 1. Springer Science & Business Media.

LORENTZ, E.N. 1972 Barotropic instability of Rossby wave motion. *J. Atmos. Sci.* **29**, 258–265.

MILLER, J. 1990 Statistical mechanics of Euler equations in two dimensions. *Phys. Rev. Lett.* **65** (17), 2137–2140.

MORRISON, P.J. 1998 Hamiltonian description of the ideal fluid. *Rev. Mod. Phys.* **70**, 467–521.

RHINES, P.B. 1975 Waves and turbulence on a beta-plane. *J. Fluid Mech.* **69**, 417–443.

ROBERT, R. & SOMMERIA, J. 1991 Statistical equilibrium states for two-dimensional flows. *J. Fluid Mech.* **229**, 291–310.

SCOTT, R.K. & DRITSCHEL, D.G. 2012 The structure of zonal jets in geostrophic turbulence. *J. Fluid Mech.* **711**, 576–598.

SHEPHERD, T. 1988 Rigorous bounds on the nonlinear saturation of instabilities to parallel shear flows. *J. Fluid Mech.* **196**, 291–322.

SRINIVASAN, K. & YOUNG, W.R. 2012 Zonostrophic instability. *J. Atmos. Sci.* **69**, 1633–1656.

SUKORIANSKY, S., DIKOVSKAYA, N. & GALPERIN, B. 2007 On the arrest of inverse energy cascade and the Rhines scale. *J. Atmos. Sci.* **64**, 3312–3327.

TAYLOR, J.B. 1974 Relaxation of toroidal plasma and generation of reverse magnetic fields. *Phys. Rev. Lett.* **33**, 1139–1141.

TAYLOR, J.B. 1986 Relaxation and magnetic reconnection in plasmas. *Rev. Mod. Phys.* **58**, 741–763.

TURKINGTON, B., MAJDA, A., HAVEN, K. & DIBATTISTA, M. 2001 Statistical equilibrium predictions of jets and spots on jupiter. *Proc. Natl Acad. Sci.* **98**, 12346–12350.

VALLIS, G.K. & MALTRUD, M.E. 1993 Generation of mean flows and jets on a beta plane and over topography. *J. Phys. Oceanogr.* **23**, 1346–1362.

WILLIAMS, G.P. 1978 Planetary circulations: 1. Barotropic representation of jovian and terrestrial turbulence. *J. Atmos. Sci.* **35**, 1399–1426.

YODEN, S., IHIOKA, K., HAYASHI, Y.-Y. & YAMADA, M. 1999 A further experiment on two-dimensional decaying turbulence on a rotating sphere. *Il Nuovo Cimento* **22**, 803–812.

YODEN, S. & YAMADA, M. 1993 A numerical experiment on two-dimensional decaying turbulence on a rotating sphere. *J. Atmos. Sci.* **50**, 631–644.

YOSHIDA, Z. & MAHAJAN, S.M. 2002 Variational principles and self-organization in two-fluid plasmas. *Phys. Rev. Lett.* **88**, 095001.