Vortex dynamics and zonal flows*

P. S. Marcus[†] and T. Kundu
Department of Mechanical Engineering, University of California, Berkeley, California 94720

Changhoon Lee
Department of Mechanical Engineering, University of Seoul, 90 Jeonnong-dong, Dongdaemun-gu, Seoul, 130-743, Korea

Two-dimensional vortex dynamics have been studied in plasmas by exploiting the analogy between fluid velocity and the $\mathbf{E}\times\mathbf{B}$ drift velocity. The analogy extends to geophysical flows by including physics that mimic zonal flows, dissipation and the β-effect due to the variation in the Coriolis parameter. Vortices with the same sign as the ambient zonal shear are stable, while opposite-signed vortices fragment. Rules for vortex merger derived by maximizing entropy or minimizing enstrophy do not work for vortices embedded in zonal flows. New rules based on the minimization of energy hold. When zonal flows are not imposed, and the flow is forced at small scales, large, coherent jet streams or eddies form that co-exist with turbulence. Their sizes are determined by an energy balance, not the length scales of the forcing or boundaries. The motivation for this work is to understand atmospheric and ocean vortices: Gulf stream meanders and eddies, the Antarctic ozone hole, the jet streams of Earth and Jupiter, and the Jovian Great Red Spot and White Ovals. © *2000 American Institute of Physics.* [S1070-664X(00)96805-8]

I. INTRODUCTION AND MOTIVATION

There is an analogy between two-dimensional (2-D), inviscid, incompressible fluid dynamics and pure-electron plasmas: the fluid velocity acts like the ($\mathbf{E}\times\mathbf{B}$) drift velocity, the vorticity like the electron density, the stream function like the potential, the fluid Poisson equation like the electrical Poisson equation, and the conservation of circulation like the conservation of charge. Exploiting this analogy, 2-D vortex dynamics have been studied in plasmas confined in Malmberg–Penning traps including vortex merger,[1] Diocotron instabilities (known as Kelvin–Helmholtz instabilities in the hydrodynamics literature)[2-11] and the equilibration and stability of arrays of vortices.[12-14] Fluid flows are rarely 2-D unless strong rotation or stratification is present. These conditions prevail in atmospheres and oceans, so it is not surprising that 2-D flow is mostly of concern to geophysical fluid dynamicists. Jovian vortices, such as the Great Red Spot, the White Ovals, and the numerous Kármán vortex streets, a primary interest of the authors, usually interact strongly with the planetary zonal (east–west) jet streams or are strongly influenced by the β-effect due to the north–south gradient of the Coriolis force (see Ref. 15 and see Sec. II) or by dissipation. The analog of zonal flows can be created in plasma experiments by imposing a radial potential with a charged axial wire. The β-effect can also be created in a plasma trap,[16] and the analog of a Rayleigh fluid dissipation (see Sec. IV for definition) can possibly be induced in a plasma by introducing a gas into the trap with a high electron affinity such as sulfur hexaflouride.

With these modifications, 2-D geophysical fluid dynamics can be studied in a plasma. In contrast 2-D fluid dynamics, with or without these additional effects, are difficult to study in laboratory experiments with real fluids. To keep the flow 2-D, rapid rotation is necessary, but rotation creates an Ekman dissipation due to the drag of the top and bottom boundary layers.[15] To be within the parameter regime where the rotation is strong enough to keep the flow 2-D but weak enough so that the Ekman dissipation time is much greater than a vortex turn-around time requires that the diameter of a rotating tank be greater (and usually much greater) than a meter, which is both complicated and expensive.[17] Moreover, in the rotating tank experiments, it is difficult to impose arbitrary initial conditions, so they almost always begin with solid-body rotation. In addition, it is difficult to force the flow in a precise manner. Generally, it is forced by jets in either the side or bottom boundaries or by oscillating grids. In contrast, a plasma in a Malmberg-Penning trap can be forced by creating or destroying vorticity very precisely both as a function of space and time, and almost arbitrary initial conditions can be created. (This includes the creation and forcing of vortices with both signs by using a photocathode source; see Ref. 18 and Sec. II.)

Two-dimensional fluid flow can be numerically simulated using contour dynamics or spectral methods, but simulation works best for initial-value codes with short integration times and for steady-state finders where the effects of numerical dissipation are unimportant. (These are the types of numerical results reviewed here.) Typically, flows computed with contour dynamics become unstable after short times unless "contour surgery" is used to remove the small filaments that form.[19] After long integration times, the accuracy of flows computed with contour surgery (or with numerical dissipation in spectral or finite-difference methods)

---

decreases and changes by order unity the values of some of the flow's ''conserved quantities'' thereby making the results dubious. For these reasons plasma experments may offer the best method of studying the long-time evolution of 2-D flows relevant to geophysics.

The purpose of this review paper is to provide an overview of the new physics that could be explored if zonal flows, Rayleigh dissipation, and/or a β-effect were included in the plasma experiments. In Sec. II we present the equations of motion and conserved quantities. In Sec. III we provide a tutorial on how vortex dynamics is altered by a zonal flow. More detailed reviews can be found in Refs. 20 and 21. These results require neither a β-effect nor a Rayleigh dissipation—only a zonal flow imposed by a radial potential created by a charged axial wire. We show how the zonal flow breaks the degeneracy between positive and negative vortices and how vortex merger is changed substantially. In Sec. IV we show how a turbulent, but coherent, long-lived, zonal flow forms spontaneously from an initial condition at rest when a β-effect, a random small-scale forcing and a Rayleigh dissipation are included (with, of course, no imposed zonal flow). Using a theory based on the inverse cascade of energy from small to large scales, the strength and widths of the resulting zonal jet streams are correctly predicted.[22] It is shown that the widths are independent of the length-scales of the forcing and boundaries. It is determined by balancing the energy input and dissipation rates. Thus, jets with arbitrary widths can be created by adjusting the magnitude of the forcing or dissipation. Our Discussion is in Sec. V.

II. EQUATIONS

The governing equations for 2-D fluid flow can be expressed in terms of the potential vorticity q

$$\frac{Dq}{Dt} \equiv \left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)q = F + D \tag{1}$$

where D/Dt is the advective derivative, $\mathbf{v}(r,\phi,t)$ is the 2-D velocity, F is the forcing, and D is the dissipation. The potential vorticity is defined as

$$q(r,\phi,t) \equiv \omega(r,\phi,t) + \beta r \tag{2}$$

where $\omega(r,\phi,t) \equiv \mathbf{z}\cdot\nabla\times\mathbf{v}$ is the vorticity, $\mathbf{z}$ is a unit vector, and β is the gradient in latitude of the Coriolis force (in the laboratory, the topographic β-effect is due to the radially-sloped bottom boundary—see Ref. 15). In a plane-parallel geometry $q = \omega + \beta y$ where y is the local north–south coordinate. In the absence of β, Eq. (1) reduces to the Navier–Stokes equation when $D = \nu\nabla^2\mathbf{v}$ where ν is the kinematic viscosity.

To more easily understand the effect of a zonal flow, we decompose the velocity into a time-independent, axisymmetric zonal component $\bar{\mathbf{v}}(r) \equiv \bar{v}_\phi(r)\ \boldsymbol{\phi}$ (where $\boldsymbol{\phi}$ is the unit vector in the azimuthal direction) and the remainder $\tilde{\mathbf{v}}$

$$\mathbf{v}(r,\phi,t) \equiv \bar{\mathbf{v}}(r) + \tilde{\mathbf{v}}(r,\phi,t) \tag{3}$$

with similar decompositions for the vorticity and potential vorticity

$$\omega(r,\phi,t) \equiv \bar{\omega}(r) + \tilde{\omega}(r,\phi,t), \tag{4}$$

$$q(r,\phi,t) \equiv \bar{q} + \tilde{q}(r,\phi,t), \tag{5}$$

where $\bar{\omega}(r) = r^{-1}d(r\bar{v}_\phi)/dr = \mathbf{z}\cdot\nabla\times\bar{\mathbf{v}}$. The zonal component of the potential vorticity is defined

$$\bar{q} \equiv \bar{\omega}(r) + \beta r. \tag{6}$$

We require $\bar{q}$ to be constant, which makes $\bar{v}_\phi(r) = c/r + \bar{q}r/2 - \beta r^2/3$ where c is a constant. Note that $\tilde{q}(r,\phi,t)$ is homogeneous in $\tilde{\mathbf{v}}$ and

$$\tilde{q}(r,\phi,t) = \tilde{\omega}(r,\phi,t) = \mathbf{z}\cdot\nabla\times\tilde{\mathbf{v}}. \tag{7}$$

In the absence of forcing and dissipation

$$\frac{D\tilde{q}}{Dt} = \left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\tilde{q}. \tag{8}$$

The advective velocity $\mathbf{v}$ in Eq. (8) contains both the zonal and non-zonal components, whereas only the non-zonal component of q appears. Thus $\tilde{q}$ advects with the fluid velocity.

Unlike 3-D flows, incompressible 2-D flows are invariant under rotation, so without loss of generality we can set $\bar{q} = 0$. Note that due to this invariance any constant may be added to ω without changing the dynamics. Thus if only one sign of vorticity (or charge in the plasma trap) can be created, a compact vortex of the opposite sign may be effectively created by creating a space-filling vortex with a compact ''hole.'' We define the zonal shear as $\bar{\sigma}(r) \equiv rd(\bar{v}_\phi/r)/dr = -2c/r^2 - \beta r/3$. If $\beta = 0$ a central charged wire in a trap produces a zonal flow with $\bar{\mathbf{v}} \propto 1/r$ and a non-zero zonal shear. Equation (8) along with the computation of $\tilde{\mathbf{v}}$ from $\tilde{q}$ by using Eq. (7) and the Biot-Savart law form a compete set of equations for determining the flow. At the radial boundaries $\tilde{\mathbf{v}} = 0$ for viscous flow and $\tilde{v}_r = 0$ for inviscid flow. If the origin is included in the domain (i.e., there is no central wire) analyticity of $\tilde{\mathbf{v}}$ must be imposed there.

Equation (8) with inviscid boundary conditions conserves circulation $\tilde{\Gamma} \equiv \int \tilde{q}dA$; all moments of the enstrophy $\int \tilde{q}^n dA$ for integers $n > 1$; energy (see Sec. III C); and angular momentum or, equivalently,

$$L \equiv \int \tilde{q}(r,\phi,t)r^2 dA. \tag{9}$$

When $\bar{\mathbf{v}} = 0$ and when the domain is unbounded, angular momentum around *any* point, not just the coordinate origin, is conserved, i.e.,

$$L(\mathbf{r}') \equiv \int \tilde{q}(\mathbf{r},t)|\mathbf{r}-\mathbf{r}'|^2 dA \tag{10}$$

is conserved for all $\mathbf{r}'$.

III. EFFECTS OF THE ZONAL FLOW

A. Breaking sign degeneracy

Vortices with different signs of $\tilde{q}$ behave as mirror images when $\bar{\mathbf{v}} = 0$ but quite differently when $\bar{\mathbf{v}} \neq 0$. When a compact region of potential vorticity has the sign of its $\tilde{q}$ the same sign as the shear $\bar{\sigma}$ of its ambient zonal flow, it is defined as *prograde*; if the signs are opposite it is *adverse*. The breaking of the degeneracy between prograde and ad-
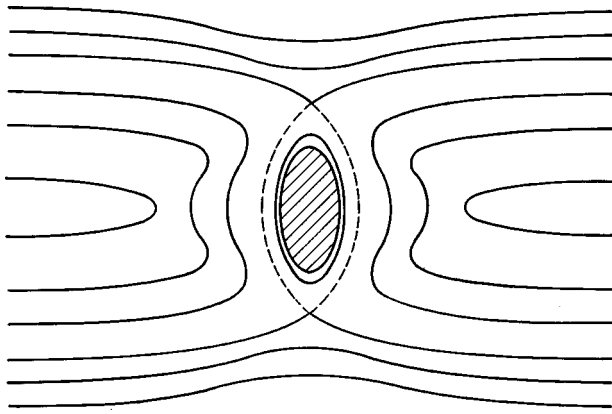
FIG. 1. Sketch of streamlines for an adverse vortex (shaded). Two stagnation points outside the vortex are joined by the last closed streamline (broken curve) that circumscribes the vortex and acts as a separatrix.

verse vortices is important; for example, one needs to understand why all of the long-lived jovian vortices are prograde. When $\bar{\mathbf{v}}=0$, steady vortices are round shaped, but when embedded in a zonal flow, prograde vortices become nearly elliptically shaped with their major axis aligned in the direction of the zonal flow. In fact, for compact vortices with spatially uniform $\tilde{q}$ (which is a good approximation for many geophysical vortices[21,23]), embedded in Cartesian zonal flows (i.e., with $\bar{\mathbf{v}}$ a function of y and in the x direction) with a constant $\bar{\sigma}$, the steady-state equilibria are exactly elliptical in shape with aspect ratio λ (maximum extent in x divided by extent in y) a function of $\bar{\sigma}/\tilde{q}$[24]

$$\lambda = \left(1 + \frac{\bar{\sigma}}{\tilde{q}}\right) M(\bar{\sigma}/\tilde{q}) \tag{11}$$

where

$$M(\chi) \equiv \frac{1 + \chi \pm \sqrt{1 + 6\chi + \chi^2}}{2 + 2\chi}. \tag{12}$$

Because $M(\chi) \simeq 1$ for $\chi > 0$, the λ of prograde vortices is nearly proportional to $(1 + \bar{\sigma}/\tilde{q})$. There are no steady solutions for adverse vortices with $\bar{\sigma}/\tilde{q} < (2\sqrt{2}-3) \simeq -0.17$. The equilibrium shape of vortices in cylindrical zonal flows and/or with non-uniform $\tilde{q}$ can be computed with contour dynamics[25,26] and show similar behavior with their λ a monotonic function of $\bar{\sigma}/\tilde{q}$ and no steady equilibria for adverse vortices with $-\bar{\sigma}/\tilde{q} \gtrsim O(1)$.

Numerical simulations show adverse vortices to be unstable to small, finite-amplitude perturbations. This can be understood from the schematic of Fig. 1. Adverse vortices have two stagnation points close-by connected by the last closed streamline that circumscribes the vortex which acts like a separatrix. Streamlines between the separatrix and the edge of the vortex also circumscribe the vortex. Those outside do not, and fluid on them is carried far from the vortex. Small perturbations keep the vortex's $\tilde{q}$ within the separatrix and close to the vortex indefinitely. Large perturbations push the vortex's perturbed $\tilde{q}$ outside the separatrix where it gets carried away and is unlikely to rejoin it. Numerical simulations verify that adverse vortices are finite-amplitude unstable to Lagrangian perturbations with displacements bigger than the distance between the vortex edge and the separatrix. In contrast, all streamlines circumscribe a prograde vortex and there is no separatrix. Numerical simulations show that even when prograde vortices are perturbed so that they break into two or more pieces, the pieces often merge to reform the vortex. Figure 2 shows an initial condition of a prograde and adverse vortex. The prograde vortex relaxes to its equilibrium shape while the adverse vortex is stretched by $\bar{\sigma}$ and destroyed. [The filaments are broken apart by a Kelvin-Helmholtz instability in Figs. 2(d)–2(f).] Laboratory experiments show that prograde vortices are stable even when the ambient zonal flow is turbulent.[17] In our initial-value calculations prograde vortices arise from a variety of initial conditions including unstable vortex rings, random vortex fields, and solid-body rotation with random small-scale forcing;
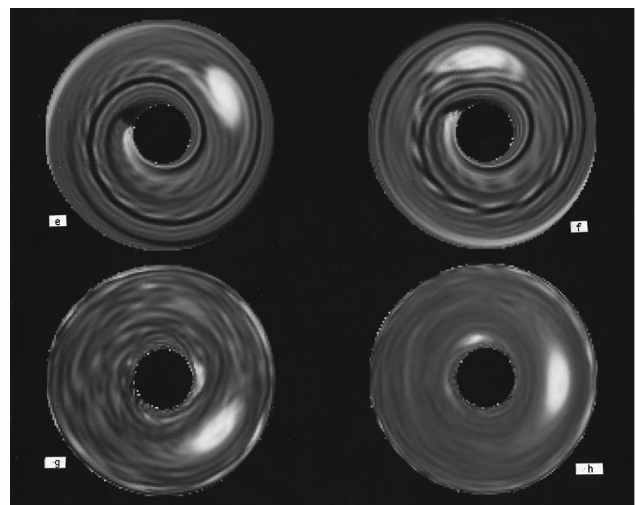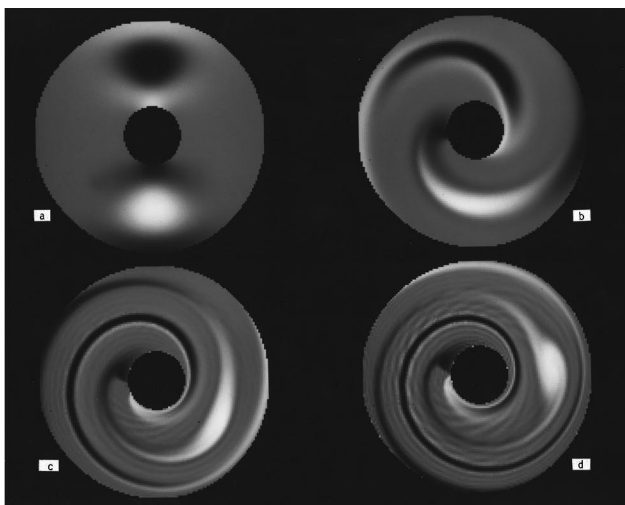

FIG. 2. Simulation of a flow in an annulus with prograde (adverse) $\tilde{q}$ light (dark) at 8 times. Only the prograde $\tilde{q}$ survives. The adverse vorticity is stretched by the zonal $\bar{\mathbf{v}}$ into small enough filaments that it is destroyed by numerical dissipation. Kelvin–Helmholtz instabilities are seen in the light spirals in Figs. d–f.

while we never see adverse vortices form unless they are included as part of an initial condition and have $-\bar{\sigma}/\tilde{q} \ll 1$.

Zonal flows alter vortex dynamics by preventing vortices from having internal degrees of freedom. With no zonal flow, a steady-state round vortex can be perturbed into a time-dependent state that oscillates or has waves on its boundaries. The different temporal behaviors allow vortices to have different energies although they have the same circulations, enstrophies and momenta as the steady states. Vortices embedded in zonal flows with $\bar{\sigma}/\tilde{q} \gtrsim O(1)$ also have time-dependent equilibria, but our numerical simulations show that they are finite-amplitude unstable; they quickly relax back to their (approximate) steady states (with very weak time dependence) by shedding vorticity and remerging with it. During the relaxation, energy is exchanged with the ambient flow (see Sec. III C). Thus when $\bar{\mathbf{v}} \neq 0$ a vortex's late-time shape and energy are determined uniquely by its circulation, momentum and ambient value of $\bar{\sigma}/\tilde{q}$; whereas when $\bar{\mathbf{v}} = 0$ a late-time vortex can be time-dependent and have a range of energies. This distinction will be important in determining the rules of vortex merger in Sec. III C.

B. Expulsion of adverse vorticity

Another reason adverse vortices with $-\bar{\sigma}/\tilde{q} \gtrsim O(1)$ are not observed is that they are stretched into filaments by $\bar{\mathbf{v}}$. If the zonal flow alternates in direction so that $\bar{\sigma}$ changes sign as a function of r (as it does for Jupiter and Saturn) then adverse $\tilde{q}$ is expelled from its ambient zone into its neighboring prograde zone where it rolls up and forms a stable vortex. Figure 3 shows a numerical simulation that illustrates this in an annular, rather than cylindrical, geometry. Here $\bar{v}_\phi(r) = (-(r^2/6)(2+(r_0/r)^3)+0.48r)$, so $\bar{\sigma}(r) = (r/3)((r_0/r)^3 - 1)$ where r_0 is midway between the inner and outer radial boundaries. Note $\bar{\sigma}(r) > 0$ for $r < r_0$, and $\bar{\sigma}(r) < 0$ for $r > r_0$. Initially both the inner and outer zones each contain one prograde and one adverse vortex, with light regions corresponding to $\tilde{q} > 0$ and dark to $\tilde{q} < 0$. In each case the adverse vorticity is expelled in a spiral. Part is stretched to the small scales where it is destroyed (by numerical dissipation) and part of the adverse $\tilde{q}$ crosses $r = r_0$ where it becomes prograde and rolls up (see middle sequence of Fig. 3), eventually settling into a stable equilibrium. In contrast the two initial prograde vortices barely change.

To understand this, consider the schematic in Fig. 4 which shows part of the first frame of Fig. 3. The dark, adverse, potential vortex with $\tilde{q} < 0$ is drawn shaded, and it lies in a shear with $\bar{\sigma} > 0$. The $\bar{\mathbf{v}}(r)$ is represented by heavy arrows. The figure is drawn in the rotating frame where the center of potential vorticity of the shaded vortex is at rest, so $\bar{\mathbf{v}}(r)$ is approximately zero at its center. An infinitesimal piece of $\tilde{q}$ labeled $\mathbf{A}$ moves with $\mathbf{v}(\mathbf{r},t) \equiv \tilde{\mathbf{v}}(\mathbf{r},t) + \bar{\mathbf{v}}(r)$. The Biot–Savart law gives $\tilde{\mathbf{v}}$ (shown with thin solid arrows); it is clockwise around the vortex, and at $\mathbf{A}$ $\tilde{v}_\phi \simeq -\tilde{q}R_r/2$ where R_r is the vortex semi-radius in r. Taylor expansion of $\bar{\mathbf{v}}$ around the vortex center gives $\bar{\mathbf{v}}$; at $\mathbf{A}$, $\bar{v}_\phi \simeq R_r\bar{\sigma}$. Thus if $|\bar{\sigma}/\tilde{q}| \gtrsim O(1)$, fluid element $\mathbf{A}$ is dragged to the right and $\mathbf{B}$ to the left (shown with broken arrows). The clockwise motion of $\tilde{\mathbf{v}}$ then pushes $\mathbf{A}$ downward and outward to the ad-
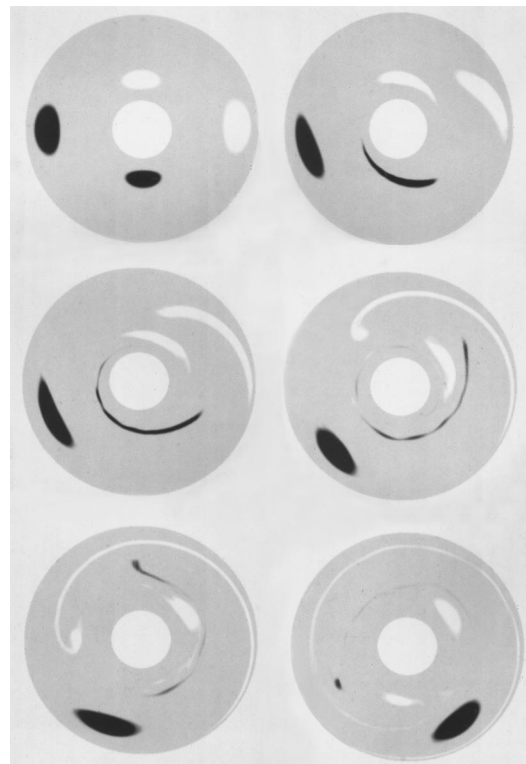

FIG. 3. Six different times in the evolution of four potential vortices in an annulus. The shear $\bar{\sigma}(r)$ of the zonal flow changes sign at a radius r_0 halfway between the inner and outer radial boundaries. The $\bar{\sigma}(r)$ is negative in the outer region. The light (dark) vortices have $\tilde{q} > 0$ ($\tilde{q} < 0$), so light vortices are prograde in the inner half and adverse in the outer half of the annulus. The pieces of vorticity that are initially adverse but cross over $r = r_0$ to become prograde develop tadpole-like heads, roll up, and survive; while those left behind continue to stretch and eventually decay due to numerical dissipation.

joining zone and $\mathbf{B}$ upward towards the annulus' boundary. If the sign of the shaded vortex were reversed so that it were prograde, then both $\bar{\mathbf{v}}$ and $\tilde{\mathbf{v}}$ initially move $\mathbf{A}$ to the right and $\mathbf{B}$ to the left. Then the counter-clockwise $\tilde{\mathbf{v}}$ would pull both $\mathbf{A}$ and $\mathbf{B}$ away from the boundaries towards the center of the prograde zone. Thus for $\bar{\sigma}/\tilde{q} = O(1)$, prograde $\tilde{q}$ is drawn in towards the center of a prograde zone while adverse $\tilde{q}$ is expelled.
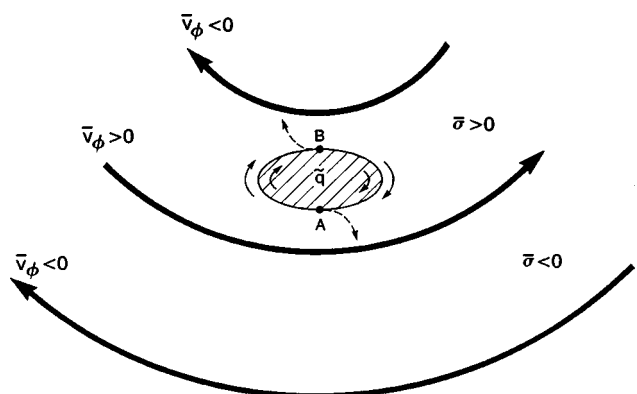

FIG. 4. Schematic of Fig. 3 showing adverse $\tilde{q}$ being expelled. The adverse $\tilde{q} < 0$ is shaded; the zonal velocity $\bar{\mathbf{v}}$ is indicated with heavy arrows, the self-induced $\tilde{\mathbf{v}}$ computed with the Biot–Savart law with thin arrows, and the total velocity $\mathbf{v}$ which advects the $\tilde{q}$ at $\mathbf{A}$ and $\mathbf{B}$ with broken arrows.

C. Vortex merger

Only vortices with like signs merge, but merger is quite different and much more common when $\bar{\mathbf{v}} \neq 0$. It is important to understand why this happens. For example in the jovian atmosphere, which is dominated by zonal winds, the long-lived vortices survive by constantly merging with smaller vortices that are continually created by the local "weather."[20] If merger were difficult, Jupiter's "long-lived" vortices would disappear.

When $\bar{\mathbf{v}} = 0$ vortices must be within approximately a diameter of each other to merge. When they do, they shed large filaments of $\tilde{q}$ outward far from their center of vorticity, so the resulting merged vortex has much less area and circulation than that of the two initial vortices. Two point vortices (delta functions of q) can never merge. They rotate around their common center of vorticity (and interact with the boundaries if there are any). As shown below it takes energy to push two like-signed vortices together, so in the absence of $\bar{\mathbf{v}}$, it is necessary to throw large amounts of q outward in order to push together pieces of two different vortices in order to conserve energy. In the absence of boundaries, angular momentum, which is conserved about every point including the center of vorticity [see Eq. (10)], also prevents vortices from merging unless a significant amount of q is thrown outward.

Vortex merger is more common in the presence of zonal flows for several reasons. Angular momentum is conserved only with respect to the origin, and $\bar{\mathbf{v}}$ pushes vortices together. So except at the origin, angular momentum conservation is not a barrier to merger. Energy is still conserved, but now includes an interaction term E_{int} between $\tilde{\mathbf{v}}$ and $\bar{\mathbf{v}}$ which can supply the energy needed to push two like-signed vortices together. To see this write the energy E as a sum of two pieces: E_{self} the self-energy and E_{int}.

$$E_{\text{self}} \equiv -\frac{1}{4\pi} \int \int \tilde{q}(\mathbf{r},t) \tilde{q}(\mathbf{r}',t) \ln|\mathbf{r} - \mathbf{r}'| dA dA' \quad (13)$$

$$E_{\text{int}} \equiv -\int \tilde{q}(\mathbf{r},t) \bar{\psi}(r) dA \quad (14)$$

where the integrals are over area(s) of the domain and $\bar{\psi}(r)$ is the stream function (electrical potential in the plasma) of the zonal component of the flow: $\bar{v}_\phi(r) = d\bar{\psi}/dr$. Not surprisingly E_{self} is of the same form it would be for 2-D patches of electric charge. Two charges or patches of vorticity with the same sign increase their energy (and therefore require it from some external source) if they are brought together. We can approximate the changes in E_{self} and E_{int} due to the merger of two identical vortices with potential vorticities $\tilde{q}$, circulations $\tilde{\Gamma} \equiv \int \tilde{q} dA$ and average radii R. Let the vortices be embedded in $\bar{\mathbf{v}} = c/r$, and assume that the length over which $\bar{\sigma}$ varies is large compared to R and that $R/D \ll 1$ where D is the initial separation between the vortices. Let r_0 be the mean radial location of the vortices: $r_0^2 \equiv \int r^2 \tilde{q} dA/\tilde{\Gamma}$, which is a constant of the motion. Assuming that the two vortices merge, we can predict everything about the merged vortex: (1) its value of $\tilde{q}$ is the same as that of the initial vortices; (2) it is located at radius r_0; (3) it is approxi-

mately steady in time (see Sec. III A) and embedded in a shear $\bar{\sigma}(r_0)$ with a nearly elliptical shape given by Eq. (11); and (4) its area is $2A$ and circulation is $2\tilde{\Gamma}$ because they are conserved quantities of the total flow, and numerical simulations show that even though circulation can be stripped from the vortices during violent mergers, most of it eventually re-attaches (unlike mergers with $\bar{\mathbf{v}} = 0$). To leading order, the increase in E_{self} after merger is

$$\Delta E_{\text{self}} = \frac{\tilde{\Gamma}^2}{2\pi} \ln(D/R) > 0. \quad (15)$$

By Taylor expanding $\bar{\sigma}(r)$ about r_0 and writing $r \equiv r_0 + y$, we find to leading order in y/r_0

$$E_{\text{int}}(t) = -\frac{\bar{\sigma}(r_0)}{2} \int \tilde{q}(r,\phi,t) y^2 dA \quad (16)$$

where we have ignored an additive constant. The quantity $|\int \tilde{q}(r,\phi,t) y^2 dA|$ is a measure of the radial width of the vortex or vortices. If the two initial vortices are located at the same radius, then the merged vortex has the same shape, but twice their area, so its radial width is larger. Thus from Eq. (16), $\Delta E_{\text{int}}(t)$ is negative (positive) after the merger of prograde (adverse) vortices. The total change in energy ΔE is positive for the merger of adverse vortices but could be of either sign for prograde vortices. If the two initial vortices have a radial separation b, then the change in $|\int \tilde{q}(r,\phi,t) y^2 dA|$ due to merger is a monotonically decreasing function of b. For prograde vortices there is a critical value $b_{\text{crit}} \simeq 2R$ such that $\Delta E = 0$.[27] A standard argument used in fluid dynamics is that a flow will undergo a kinematically allowable transition if the energy of the final state is less than the initial.[28] As applied to the merger of vortices it means that if $\Delta E < 0$, a merger will occur and the energy $|\Delta E|$ is transferred from the large-scale coherent vortices to the turbulent, small-scale component of the flow. Merger would only be possible for $\Delta E > 0$ if energy could be contrived to go from the turbulence into the coherent vortices, and this is not observed. Numerical experiments have confirmed to within 5% that prograde vortices with $b < b_{\text{crit}}$ merge, and those with greater values do not.[27] Two adverse vortices with large b could lower their energy by merging, but clearly it is necessary for the vortices to approach each other before they could merge. Because the vortices move with $\mathbf{v}$, it is not obvious how they could do so and overcome their large initial separation in radius. We know of no numerical or observational examples of adverse vortex merger with $|\bar{\sigma}/\tilde{q}| > 0.1$. In fact, it is energetically favorable for a single adverse vortex to fission which explains, in part, their finite-amplitude instability.

Various hypotheses based on entropy maximization,[29–33] enstrophy minimization,[34,35] and statistical mechanics[36–38] have been developed to predict vortex merger (and the equilibria of arrays of vortices). However, these predictions do not appear to work for $\bar{\mathbf{v}} \neq 0$. This may be due to the fact that the theories depend on the flows' ergodicity. Numerical experiments with for $\bar{\mathbf{v}} \neq 0$ show the flows are not ergodic and have strong "memories" of their initial conditions.[27]

Phys. Plasmas, Vol. 7, No. 5, May 2000

Vortex dynamics and zonal flows   1635

IV. SPONTANEOUS GENERATION OF ZONAL FLOWS

A. Physical and mathematical description

Zonal winds and jet streams occur not only on Earth but also on all of our solar system's large gaseous planets (Jupiter, Saturn, Uranus and Neptune). It is still an open question as to what created and maintains them for long periods of time. Jupiter's zonal jet streams (or more accurately, the cloud patterns that we associate with them) have been observed for more than 350 years.

When a zonal flow is not imposed on the fluid and when the fluid starts at rest, a small-scale forcing can lead to the creation of large coherent structures such as zonal winds *via* an inverse energy cascade—*under the right conditions*. In this section we consider the effects of a β term and large-scale dissipation on the cascade and the resulting coherent structures. To simplify the presentation we use Cartesian coordinates with y acting like the radial direction and x like the azimuthal. We do not use the decomposition of variables into their zonal and nonzonal components.

We begin with Eq. (1) but now explicitly consider the roles of the forcing $F(x,y)$ and dissipation D. The forcing in a rotating atmosphere is $F=2\Omega\,\partial v_z(x,y)/\partial z$, where Ω is the angular velocity of the rotating fluid layer and z is the vertical coordinate.[15] An atmosphere, such as the weather layer of Jupiter (which contains the visible clouds), would be expected to have large (negative) values of $\partial v_z(x,y)/\partial z$ at the locations where vertical plumes of fluid rise from the underlying convective layer and abruptly stop at the tropopause (which we take to be the upper 'lid' of our Jovian atmospheric model, where the atmosphere becomes strongly vertically stratified and severely inhibiting to vertical motion).[20] Both theory and the limited observations of plumes indicate that they occur on small scales (less than 500 km—compared with zonal widths of order 10 000 km) over a narrow range of length scale. Thus we model F with forcing at a pair of high-wavenumber $\mathbf{k_f}$ Fourier modes: $F=(A/2)\times[e^{i(k_f/\sqrt{2})x}e^{i(k_f/\sqrt{2})y}+e^{i(k_f/\sqrt{2})x}e^{-i(k_f/\sqrt{2})y}+\text{c.c.}]$. The dissipation $D=-\omega/\tau+\nu^{\text{hyper}}\nabla^{12}\omega$ in Eq. (1) has two terms. One is hyperviscosity whose purpose is to prevent a numerical instability and whose numerical coefficient $\nu^{\text{hyper}}\nabla^{12}$ is self-normalizing (with no pre-assigned values) and designed to prevent enstrophy from piling up at the smallest numerically resolvable scale. All 2-D calculations, even 'run-down' experiments, (defined as those with $F\equiv0$, $\tau\to\infty$ and nonzero initial conditions) include this term or an equivalent 'turbulent viscosity' to prevent instability (see Ref. 21). The other term in D is a Rayleigh friction or Ekman dissipation. It models any phenomena that dissipates potential circulation, or momentum in the layer on time scale τ (which dissipation terms that can be written as a horizontal divergence cannot do) such as oblique Rossby waves, secondary circulations with a vertical component, or turbulent upper or lower boundaries.[22] Unlike molecular viscosity or hyperviscosity, this term mostly dissipates energy at large scales if the energy spectrum decreases with wavenumber.

Thus in this section we are interested in solutions to

$$\frac{\partial\omega}{\partial t}=-\mathbf{v}\cdot\nabla\omega-\beta v_y-\frac{\omega}{\tau}+(A/2)[e^{i(k_f/\sqrt{2})x}e^{i(k_f/\sqrt{2})y}$$
$$+e^{i(k_f/\sqrt{2})x}e^{-i(k_f/\sqrt{2})y}+\text{c.c.}]+\nu^{\text{hyper}}\nabla^{12}\omega \qquad (17)$$

over a computational domain of $L\times L$. Ideally, we would like to solve this flow over an infinite domain, but as long as L is much larger than the largest-scale features of the system, box-size effects are unimportant. When the largest features approach the size of the computational domain, finite-area effects become important (see Ref. 21).

For all flows, the initial velocity is machine zero. There is a 5-dimensional (3-nondimensional) parameter space to explore: β, L, τ, k_f, and A, though we postpone choosing non-dimensional units until Sec. IV E.

B. Energy spectra

The arguments for the Kolmogorov energy spectrum are the same for 2- and 3-D flows and independent of whether the energy forward cascades from large to small scales or inverse cascades.[39] The argument depends only on the assumption that the energy transfer in Fourier space ($\mathbf{k}$) is local and conservative

$$E(k)=c_{kg}(\dot{E}_{up})^{2/3}k^{-5/3} \qquad (18)$$

where the kinetic energy per unit mass is $E\equiv\int_0^\infty E(k)dk$, and $\dot{E}_{up}$ is the rate of transfer up (or down) the spectrum. The dimensionless Kolmogorov constant c_{kg} in 3-D flow is approximately 1.5.[39] It is convenient to use the notion of an 'eddy' as the coherent component of the velocity made up of a band of Fourier modes with wavenumbers between $|\mathbf{k}|/2$ and $|\mathbf{k}|$ with energy

$$\varepsilon_{eddy}(k)=\int_{k/2}^k E(k)dk. \qquad (19)$$

C. Overview of numerical results

Unlike 'run-down' experiments in which all initial conditions and all parameter values (with $\beta\neq0$) that we tried produced zonal flows, very few choices of parameters for Eq. (17) did. For an inverse cascade to occur, certain inequalities of the five parameters must be met. Discussion of these conditions will follow. Only when these conditions are met do our calculations produce coherent structures. Most regions within our 5-dimensional parameter space do not satisfy the necessary conditions for a nonlinear cascade, so most of our calculations never produced large-scale features and resulted instead in energy spectra confined to a delta function at k_f.

When an inverse cascade of energy occurred, it produced a final, statistically steady state that had east–west zonal jets or eddies. For runs which formed east–west jets, there was a wide range in their number and size. For runs which formed eddies, there was also a wide range, though for each numerical calculation in which there was an inverse cascade there was always one well-defined size.

Some examples of runs which establish an inverse cascade and produce coherent features are given in Figs. 5–7. In parts (a) of Figs. 5 and 6, we have plotted the east–west (x)
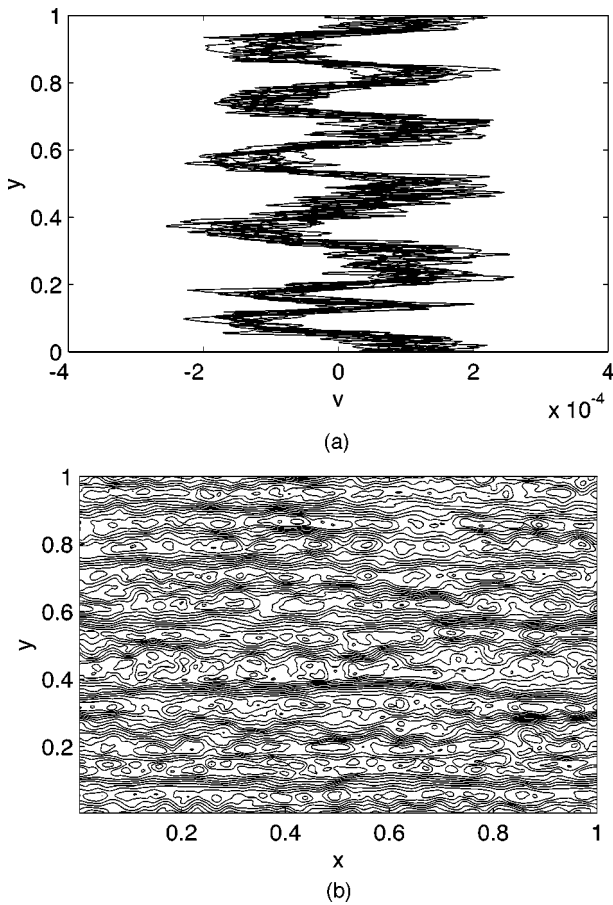
(a)

(b)

FIG. 5. (a) The east–west velocity as a function of latitude y at 10 different values of longitude; i.e., $v_x(x_i,y)$ at $x_i = (i/10)L$ for $i = 1,2,\ldots,10$; and (b) stream function $\psi(x,y)$ for a numerical simulation with $\tau^* = 73.2$, $k_f^* = 1.17$, and $L^* = 273$. (See Sec IV. for our definition of dimensionless units indicated by a $*$.) These figures are computed at a late time ($t \gg \tau$) after the flow has come to a statistically steady state.

velocity as a function of latitude (y) at 10 different longitudes (x). Note that in Figs. 5(a) and 6(a) the east–west extrema of the jets correlate remarkably well even though the flow is highly turbulent, indicating that these flows have become zonal. It is important to note that even with the same values of β these flows can have a different number of zones (five for Fig. 5 but ten for Fig. 6). The zonal jets can vary in appearance; they are very laminar-looking in Fig. 6(a) but much more turbulent-looking in Fig. 5(a). This contrast in appearance can also be seen in the stream functions plotted in Figs. 5(b) and 6(b). The flow in Fig. 5(b) consists of much more turbulence, waves, and eddies, and although this is taken at one moment in time, it is representative of the features of the system once the statistically-steady equilibrium has been established.

Figure 7 represents a run where zones are never established. Large-scale, isotropic eddies are embedded within a turbulent flow, but there is no evidence of east–west jets. For Fig. 7 in particular, the wavenumber associated with large-scale coherent features is $k \simeq 58$, but other numerical experiments yielded eddies of different scales as well.
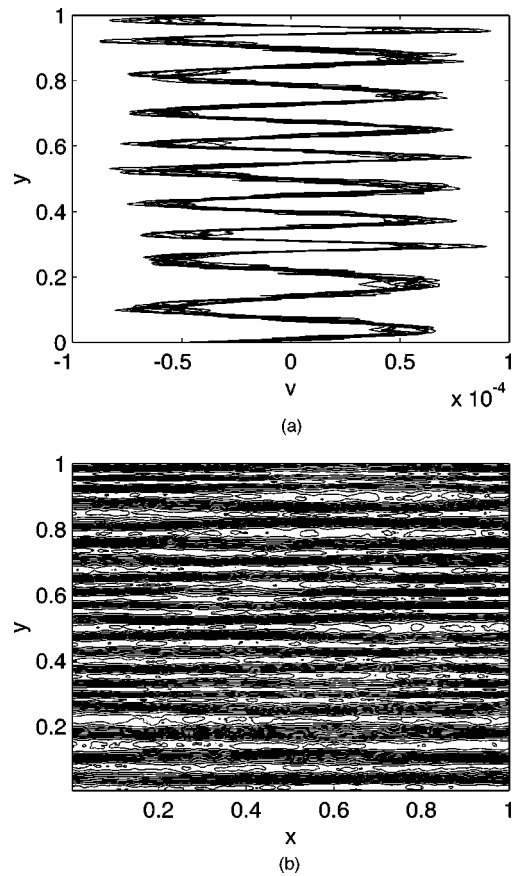

(a)

(b)

FIG. 6. Same as Fig. 5, but for $\tau^* = 44.0$, $k_f^* = 195$, and $L^* = 454$. Here the zonal flow is much more regular and there are many more zones.

D. Determination of the size of coherent features

It is possible to produce flows with *any* number of zones. (This is not just due to our re-scaling the length—all of the runs were computed on domains with $L = 1.0$.) In this section we consider balance between the energy input and dissipation rates in Eq. (17) and from it derive the conditions under which zones or other large-scale coherent features form and their characteristic wavenumbers $k_{\min}$. We shall show that dissipation-forcing balance sets the scale (not β, since the theory works even with $\beta = 0$).

Multiplying both sides of Eq. (17) by the stream function ψ (where $\mathbf{v} \equiv -\nabla \times \psi\mathbf{z}$) and integrating over the domain gives an equation for E:

$$\frac{\partial E}{\partial t} = \frac{-2E}{\tau} - \dot{E}_{\mathcal{H}} + \dot{E}_{\mathrm{in}} \tag{20}$$

where

$$\dot{E}_{\mathrm{in}}(t) \equiv -2\Pi^2 A[\,\hat{\psi}(k_f/\sqrt{2}, k_f/\sqrt{2}) + \hat{\psi}(k_f\sqrt{2}, -k_f/\sqrt{2}) + \mathrm{c.c.}\,] \tag{21}$$

where a 'hat' over a quantity means the discrete Fourier transform of that quantity,[22] and where $\dot{E}_{\mathcal{H}}$ is the energy loss rate per unit mass due to hyper-viscosity. In general we find that the numerical calculations have little loss from hyper-viscosity (typically, $\dot{E}_{\mathcal{H}} < 0.08\dot{E}_{\mathrm{in}}$) because energy inverse
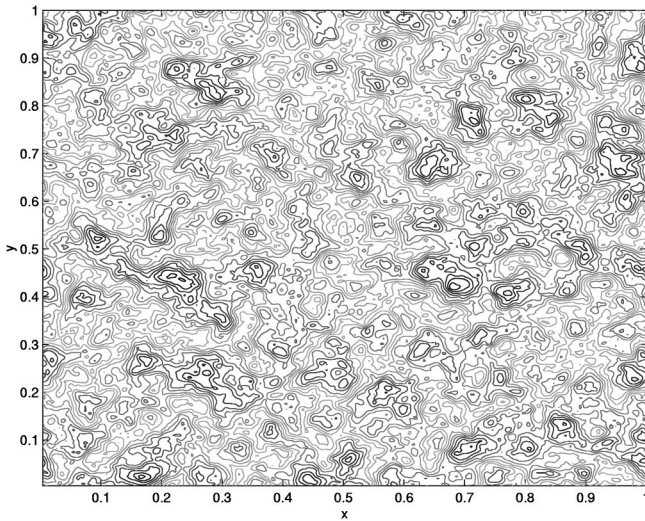
FIG. 7. Stream function $\psi(x,y)$ for a numerical simulation with $\tau^* = 1484$, $k_f^* = 1.37$, and $L^* = 936$. Here the flow shows a well-defined prominent wavenumber ($k_{pk} \approx 58$) but the flow produces an isotropic eddy field rather than east–west zones.

cascades to large scales rather than forward cascading to the small scales where hyper-viscosity is effective. Thus, to a first approximation the upward energy transfer rate in Eq. (18) is

$$\dot{E}_{up} = \frac{2E}{\tau} \approx \dot{E}_{in}. \tag{22}$$

We consider the energy balance at $\mathbf{k} = \mathbf{k_f}$ by examining the $\mathbf{k_f}$-Fourier component of Eq. (17):

$$\partial \hat{\omega}(\mathbf{k_f})/\partial t = -[\widehat{(v \cdot \nabla)\omega}]|_{\mathbf{k_f}} - \beta \hat{v}_y(\mathbf{k_f}) + A/2 + \hat{\mathcal{H}}|_{\mathbf{k_f}}$$
$$- \hat{\omega}(\mathbf{k_f})/\tau \tag{23}$$

where a vertical bar with subscript indicates the wavevector of the transformed quantity. When Eq. (23) is multiplied by $-\hat{\psi}|_{-\mathbf{k_f}}$ it determines the time evolution of the energy in the $\mathbf{k_f}$-Fourier mode. We would expect, and our numerical calculations verify, that there is a dominant balance between two terms in Eq. (23) for the parameter range where the solutions form large-scale flows. The forcing due to the $A/2$ term in Eq. (23) not only supplies the energy to the $\mathbf{k_f}$-Fourier mode but also to the entire flow. For the flow not to run down, this must be one of the two dominant terms. The linear terms in Eq. (23) cannot balance the forcing term, e.g., if either of the two dissipation terms (i.e., $-\hat{\omega}/\tau$ or $\hat{\mathcal{H}}$) did, then the energy would go into the $\mathbf{k_f}$ mode, dissipate there, and be incapable of creating a zonal or other large spatial-scale flow. Nor can the β term balance $A/2$ because then there would be no dissipation, and the flow would create Rossby waves[15] with wavevector $\mathbf{k_f}$ that would build in amplitude until there was a numerical blow-up. The balance must be by the nonlinear term because this is the only term that can transfer energy out of the $\mathbf{k_f}$ mode to other (larger) modes. This argues that three inequalities and one equation are necessary conditions to create zonal or any other type of large scale flow:

$$|A| \gg |\beta \hat{v}(\mathbf{k_f})| \tag{24}$$

$$|A| \gg |\tau^{-1} \hat{\omega}(\mathbf{k_f})| \tag{25}$$

$$|A| \gg |\hat{\mathcal{H}}|_{\mathbf{k_f}}| \tag{26}$$

$$|A| \simeq |[\widehat{(v \cdot \nabla)\omega}]|_{\mathbf{k_f}}|. \tag{27}$$

These relations have been verified numerically.

By writing the right-hand side of Eq. (27) as a convolution sum in $\mathbf{k}$ it can be shown[22] that it is approximately equal to $k_f^2 \varepsilon_{eddy}(k_f)$. Using this approximation with Eqs. (19) and (27) allows us to write $\dot{E}_{up}$ in terms of the control parameters

$$\dot{E}_{up} \equiv c_p A^{3/2} k_f^{-2} \tag{28}$$

where c_p is a dimensionless constant of order unity. Our numerical experiments show that c_p is universal like c_{kg}, in the sense that it is nearly independent of the values of β, L, τ, k_f, and A.

The constraint inequalities (24) and (25) can now be written in terms of the control parameters:[22]

$$A^{1/2} k_f \gg \beta \tag{29}$$

and

$$A \gg \tau^2. \tag{30}$$

Inequality (26) is a constraint on the numerics, not the physics, so we disregard it in our discussion. All of the numerical results presented in Sec. IV B that form large-scale flows, either zones or eddy fields (e.g., the flows in Figs. 5–7), satisfy the constraint inequalities (29) and (30), and all numerical calculations that we carried out for which either inequality was not satisfied failed to produce large-scales flows and the energy remained in Fourier modes with $k \gtrsim k_f$.

If we ignore the effect of β, then according to Eq. (22), the kinetic energy inverse cascades from k_f to larger length scales until it reaches wavenumber $k_{min}^\dagger$ where

$$E = \int_{k_{min}^\dagger}^\infty E(k) dk = \tau \dot{E}_{up}/2. \tag{31}$$

If the Kolmogorov scaling (18) is valid for $k < k_f$, then

$$\dot{E}_{up} \simeq \frac{2}{\tau} \int_{k_{min}^\dagger}^{k_f} c_{kg} (\dot{E}_{up})^{2/3} k^{-5/3} dk \tag{32}$$

$$\simeq \frac{3}{\tau} c_{kg} (\dot{E}_{up})^{2/3} (k_{min}^\dagger)^{-2/3} \left[1 - \left(\frac{k_{min}^\dagger}{k_f} \right)^{2/3} \right] \tag{33}$$

$$\simeq \frac{3}{\tau} c_{kg} (\dot{E}_{up})^{2/3} (k_{min}^\dagger)^{-2/3}. \tag{34}$$

The large coherent structures have wavenumber $k_{min}^\dagger$

$$k_{min}^\dagger \simeq \left(\frac{\tau}{3} \right)^{-3/2} c_{kg}^{3/2} \dot{E}_{up}^{-1/2} = \sqrt{27 c_{kg}^3 / c_p} \, \tau^{-3/2} k_f A^{-3/4} \tag{35}$$

where we used Eqs. (22), (28), and (34). This predicts the dominant scale for eddies in the case of non-zonal flows. Equation (35) can only make sense if $k_{min}^\dagger \lesssim k_f$ so we might

consider this to be a constraint inequality in addition to those already listed. However, it is trivially satisfied if inequality (30) is satisfied.

For small values of β, once the value of the dimensionless, order-unity constants c_p and c_{kg} are determined, our scaling laws correctly predict all of the gross features of the numerical simulations in terms of the input parameters β, τ k_f, and A: $\dot{E}_{up}$ [from Eq. (28)], $E = c_p A^{(+3/2)} \tau/2 k_f^2$ [from Eq. (22)], and $k_{min}^{\dagger}$ [from Eq. (35)]. Note that the results are independent of L (as we argued in Sec. IV A on the assumption that $2\pi/k_{min}^{\dagger} \ll L$—although when $2\pi/k_{min}^{\dagger} \simeq L$, the numerical results depend on L[21]). The expressions for E, $\dot{E}_{up}$ and $k_{min}^{\dagger}$ depend only on τ and the combination of parameters $(A^{3/2} k_f^{-2})$ but not on A or k_f independently. This is due to the fact that the value of $(A^{3/2} k_f^{-2})$ determines $\dot{E}_{up}$. Whether the energy spectrum $E(k)$ begins at k_f or some other value has very little effect on E, or $k_{min}^{\dagger}$ if $k_f \gg k_{min}^{\dagger}$. [This is mathematically equivalent to ignoring the $(k_{min}^{\dagger}/k_f)^{2/3}$ term on the right-most side of Eq. (33).] It may define the range over which the spectrum extends, but the $k_{min}^{\dagger}$ depends primarily on the values of $\dot{E}_{up}$ and τ. Note though that constraint inequalities *do* depend separately on the values of k_f and A. Based on our numerical findings that c_p is approximately constant, we nondimensionalize our results (indicated by an asterisk) by choosing units of length and time such that $c_p A^{3/2} k_f^{-2} \equiv \dot{E}_{up} = 1$ and $\beta = 1$. The input variables can then be written as τ^*, L^*, and k_f^*, so *if* an inverse cascades forms, the gross features will depend only on τ^*, but the constraint inequalities will depend also on k_f^* and L^*.

E. Effect of β

The preceding analysis of the size of coherent structures was independent of β and appears to be in conflict with the generally accepted belief that it does depend on β.[40] To resolve this, define a band (analogous to an eddy) as a coherent feature made of Fourier components with $k_x \equiv 0$ where $k/2 < k_y \leq k$. Modes with $k_x = 0$ have velocities only in the x direction, so they contribute to the zonal component of the flow and not to eddies. Consider the component $E_{zone}(k)$ of the total energy spectrum $E(k)$ due only to Fourier modes with $k_x \equiv 0$, and define $E_{eddy}(k) \equiv E(k) - E_{zone}(k)$. Then the energy of a band of wavenumber k is

$$\varepsilon_{zone}(k) = \int_{k/2}^{k} E_{zone}(k)dk. \tag{36}$$

We found numerically that the scaling of the amplitudes of the Fourier modes changes when β is included in the calculation. For large (and 'large' will be defined below) the scaling of $E_{eddy}(k)$ remains unchanged and proportional to $k^{-5/3}$. However, our numerical calculations show that the scaling of $E_{zone}(k)$ changes when β is included and yields

$$E_{zone}(k) = c_\beta \beta^2 k^{-5}, \tag{37}$$

or equivalently

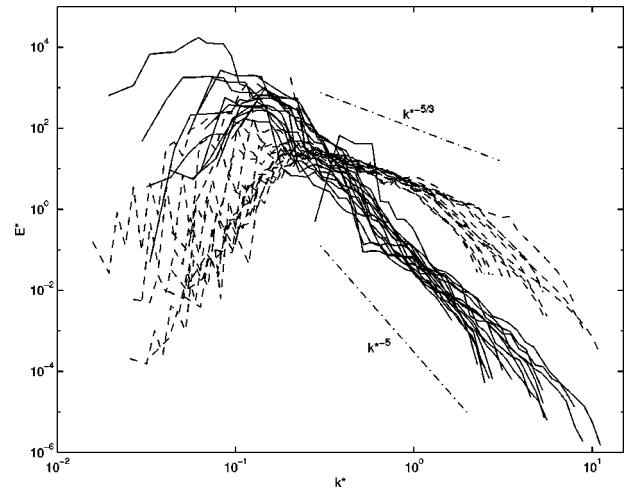$$\varepsilon_{zone}(k) = (15/4) c_\beta \beta^2 k^{-4} \tag{38}$$

FIG. 8. Composite showing the spectra $E_{zone}^*(k^*)$ and $E_{eddy}^*(k^*)$ for all runs that produced zonal flows. For all of the runs, the inertial ranges of $E_{zone}^*(k^*)$ (solid lines) and $E_{eddy}^*(k^*)$ (broken lines) collapse to the same two curves, though the endpoints of the inertial ranges, e.g., k_f^* and k_{min}^*, where the spectra sharply fall off, differ. See text for details.

where c_β is a dimensionless constant of order-unity. Our numerically computed value is $c_\beta = 0.08 \pm 0.015$. The k^{-5} scaling in Eq. (37) is illustrated in Fig. 8. Heuristically this can be explained by noting that the term proportional to β and the nonlinear term in Eq. (23) scale differently with k. Multiplying the equation by $\hat{\psi}(k)$ and averaging over all wave numbers between k and $k/2$ makes the β term approximately $\beta \varepsilon_{eddy}(k)$ and the nonlinear term approximately $k^2 \varepsilon_{eddy}(k) \varepsilon_{zone}^{1/2}(k)$. Balancing these two terms gives the scaling in Eq. (38).

Numerical simulations (see Fig. 8) show that $E_{eddy}(k)$ obeys Kolmogorov $k^{-5/3}$ scaling over a wide range of k, while $E_{zone}(k)$ simultaneously obeys Eq. (37). The total energy spectrum is the sum of $E_{eddy}(k)$ and $E_{zone}(k)$, or

$$E(k) = c_{kg}(\dot{E}_{up})^{2/3} k^{-5/3} + c_\beta \beta^2 k^{-5}. \tag{39}$$

For $k < k_b$ the second term dominates and vice versa, where

$$k_b \equiv c_{kg}^{-3/10} c_\beta^{3/10} \dot{E}_{up}^{-1/5} \beta^{3/5} = c_{kg}^{-3/10} c_\beta^{3/10} c_p^{-1/5} k_f^{2/5} A^{-3/10} \beta^{3/5}. \tag{40}$$

For choices of input parameters such that $k_b < k_{min}^{\dagger}$, the Kolmogorov scaling law is a good approximation to Eq. (39), but for $k_b > k_{min}^{\dagger}$, the correct approximation is

$$E(k) \simeq \begin{cases} c_{kg}(\dot{E}_{up})^{2/3} k^{-5/3} & k_b < k < k_f \\ c_\beta \beta^2 k^{-5} & k_{min} \leq k < k_b. \end{cases} \tag{41}$$

In Eq. (41), k_{min} is the smallest wavenumber of the flow and is determined as in Eq. (31) by balancing the energy dissipation rate with the input rate

$$\dot{E}_{\text{up}} = (2/\tau)E \tag{42}$$

$$\simeq \frac{2}{\tau} \int_{k_{\min}}^{k_{\text{b}}} c_{\beta}\beta^2 k^{-5} dk + \frac{2}{\tau} \int_{k_{\text{b}}}^{k_{\text{f}}} c_{\text{kg}}(\dot{E}_{\text{up}})^{2/3} k^{-5/3} dk \tag{43}$$

$$\simeq \frac{1}{2\tau} c_{\beta}\beta^2 k_{\min}^{-4}, \tag{44}$$

or using Eq. (28)

$$k_{\min} \simeq (c_{\beta}/2c_{\text{p}})^{1/4} \tau^{-1/4} \beta^{1/2} k_{\text{f}}^{1/2} A^{-3/8} \tag{45}$$

where we have assumed that $k_{\text{f}} \gg k_{\text{b}}$ and $k_{\text{b}} \gg k_{\min}$. In many of our calculations the former is true but not the latter, in which case

$$k_{\min} \simeq [k_{\text{b}}^{-4} + (2\tau c_p A^{3/2} k_{\text{f}}^{-2} - 6c_{\text{kg}}c_p^{2/3} A k_{\text{f}}^{4/3} k_{\text{b}}^{-2/3})/(c_{\beta}\beta^2)]^{-1/4}.$$

According to our theory, zones will form when $k_{\min}^{\dagger} \leqslant k_{\text{b}}$, or

$$3^{-5} c_{\text{kg}}^{-6} c_{\beta} c_{\text{p}} \beta^2 \tau^5 A^{3/2} k_{\text{f}}^{-2} > 1. \tag{47}$$

In terms of nondimensional quantities, zones form when

$$\tau^* > 3 c_{\text{kg}}^{6/5} c_{\beta}^{-1/5} \equiv \tau_{\text{crit}}^* \approx 26, \tag{48}$$

which is good agreement with the numerical simulations.

Figure 8 is a composite figure displaying $E_{\text{eddy}}^*(k^*)$, $E_{\text{zone}}^*(k^*)$, and $E^*(k^*)$ for all of the runs which produced zones. Note how well the inertial ranges of $E_{\text{zone}}^*(k^*)$ and $E_{\text{eddy}}^*(k^*)$ correlate with k^{*-5} and $k^{*-5/3}$ scaling, respectively. In contrast, for runs which produced eddy fields, the spectra $E_{\text{zone}}^*(k^*)$ did not exhibit k^{*-5} scaling. In fact, for eddy fields, both $E_{\text{zone}}^*(k^*)$ and $E_{\text{eddy}}^*(k^*)$ scaled as $k^{*-5/3}$, i.e., the presence of β was never felt. It is also important to note how well the inertial ranges of $E_{\text{zone}}^*(k^*)$ collapse to single curve for all the runs. The same is true for $E_{\text{eddy}}^*(k^*)$. This is extraordinary when one considers that some runs are very turbulent in appearance while others look fairly laminar. Note the endpoints of the inertial ranges are different for each run. Beyond the endpoints the spectra fall off rapidly. The right endpoint at large k^* is $k^* = k_{\text{f}}^*$ which is different for each run. For $E_{\text{zone}}^*(k^*)$ the left endpoint is $k^* = (k_{\min}^{\dagger})^*$ which is also different for each run. For $E_{\text{eddy}}^*(k^*)$ the left endpoint is $k^* = k_{\text{b}}^*$. Figure 8 can be used to evaluate the universal constant $k_{\text{b}}^* \approx 0.30$.

In summary, our analyses and numerical simulations show that if either of the two inequalities (29) and (30) are not satisfied, then large, coherent structures do not form. (If β is large, Rossby waves with length scales equal to the forcing length dominate the flow.) If both inequalities are satisfied but inequality (47) is not satisfied, then nearly isotropic eddies form. If all three inequalities are satisfied, then large coherent zonal flows form. In any case, the length scale of the large structures is set by the energy balance in Eq. (31).

V. DISCUSSION

It is difficult to study nearly-inviscid, turbulent, 2-D fluid flows in the laboratory and it is hard to numerically simulate them for long times due to numerical instabilities and dissi-

pation. Because there is an analogy between 2-D fluid flow and the **E×B** drift velocity in a plasma, there is an opportunity of using plasma traps as an ''analog computer'' for examining these flows. We have shown here that 2-D vortex dynamics are different when zonal flows, large-scale (Rayleigh-like) dissipation, or a β-effect are included. These are some of the leading physical processes in jet streams, ozone holes, and the Great Red Spot. All of these effects are plausible additions to existing plasma experiments that mimic 2-D fluid flow. Because these additional effects are important to geophysical flows, which represent the majority of 2-D flows of interest to hydrodynamicists, it would be useful to include these effects in future generations of plasma experiments. When zonal flows are included, adverse and prograde vortices behave differently. It is energetically favorable for adverse vortices (those with opposite sign of the ambient zonal shear) to fragment. If the adverse vortices are embedded in zonal flows that alternate in direction (such as the winds of Jupiter and Saturn), the fragments are expelled into neighboring regions where the zonal shear is prograde (same sign) with respect to the vorticity. When the zonal shear and the potential vorticity of the vortices are of the same order, there are no known equilibrium solutions of adverse vortices. Rules governing vortex merger and the relaxation to equilibrium change when zonal flows are present. They do not appear to be governed by the principle of maximizing the flow's entropy or minimizing its enstrophy. Instead, merger and relaxation to new equilibria occur if they allow the energy of the coherent vortices to decrease (giving their excess energy to the turbulent component of the velocity).

When zonal flows are not imposed, but the flow is randomly forced at small scales, dissipated at large scales, and a β-effect is included, an inverse cascade of energy from small to large scales is set up. When the appropriate conditions outlined in Sec. IV. are satisfied, large-scale structures form and co-exist with turbulence. The size of the structures is determined by a balance of energy input and dissipation rates. This is unusual because in most cases where coherent structures co-exist with turbulence (e.g., the Great Red Spot, solar granulation, turbulent Couette-Taylor vortices, ocean eddies) the sizes are determined by the length-scale of the forcing, the boundaries, or the Rhines length.[40] By varying the dissipation time or forcing rate in the numerical simulations (and hopefully in the experiments) the size of the coherent structures can be continuously varied from the length of the small-scale forcing to the size of the entire domain. The magnitude of β only has a secondary role in setting the lengths of the coherent structures (by changing the shape of the energy spectrum). The value of β primarily determines whether the forcing creates an isotropic eddy field, jet streams, or Rossby waves. Our analytic theories for the criteria for what types of structures form, for their length scales and velocity scales, and for the rules of merger and relaxation were shown to be verified by numerical results. However the numerical results, which are only formally valid for short time integrations, have not been rigorously tested by laboratory experiments. Plasma traps may well be the best way to test them, as well as their application to the vortices

and jet streams of the earth and Jupiter, environments that are difficult to reproduce in the laboratory but which are the authors' motivation for our study of vortex dynamics. It is our hope that this Review provides encouragement and guidance to build geophysically relevant plasma experiments.

ACKNOWLEDGMENTS

[1] K. S. Fine, C. F. Driscoll, J. H. Malmberg, and T. B. Mitchell, Phys. Rev. Lett. **67**, 588 (1991).

[2] K. S. Fine, Ph.D. thesis, University of California, San Diego, 1988.

[3] K. S. Fine, C. F. Driscoll, and J. H. Malmberg, Phys. Rev. Lett. **63**, 2232 (1989).

[4] C. F. Driscoll, Phys. Rev. Lett. **64**, 645 (1990).

[5] K. S. Fine, Bull. Am. Phys. Soc. **36**, 2331 (1991).

[6] R. Chu, J. Wurtele, J. Notte, and J. Fajans, in 1992 International Conference on Plasma Physics (European Physical Society, 1992), pp. III–1811.

[7] A. Peurrung, J. Notte, and J. Fajans, Bull. Am. Phys. Soc. **37**, 1803 (1992).

[8] A. J. Peurrung and J. Fajans, Phys. Fluids A **5**, 493 (1993).

[9] J. Fajans and L. Friedland, Bull. Am. Phys. Soc. **43**, 1927 (1998).

[10] K. S. Fine and C. F. Driscoll, Phys. Plasmas **5**, 601 (1998).

[11] J. Fajans, E. Gilson, and L. Friedland, Phys. Rev. Lett. **82**, 4444 (1999).

[12] K. Fine, A. Cass, W. Flynn, and C. Driscoll, Phys. Rev. Lett. **75**, 3277 (1995).

[13] D. A. Schecter, D. H. E. Dubin, K. S. Fine, and C. F. Driscoll, Phys. Fluids **11**, 905 (1999).

[14] D. Durkin and J. Fajans, Phys. Fluids **12**, 298 (2000).

[15] J. Pedlosky, *Geophysical Fluid Dynamics* (Springer-Verlag, New York, 1979).

[16] J. M. Finn, D. del Castillo-Negrete, and D. C. Barnes, Phys. Plasmas **6**, 3744 (1999).

[17] J. Sommeria, S. D. Meyers, and H. L. Swinney, Nature (London) **331**, 689 (1988).

[18] D. Durkin and J. Fajans, Rev. Sci. Instrum. **70**, 4539 (1999).

[19] D. G. Dritschel, Comput. Phys. Rep. **10**, 77 (1989).

[20] P. S. Marcus, Annu. Rev. Astron. Astrophys. **31**, 523 (1993).

[21] P. S. Marcus and C. Lee, Chaos **4**, 269 (1994).

[22] T. Kundu, P. S. Marcus, and C. Lee, J. Fluid Mech., submitted.

[23] P. S. Marcus, Nature (London) **331**, 693 (1988).

[24] D. W. Moore and P. G. Saffman, *Aircraft Wake Turbulence* (Plenum, New York, 1971), pp. 339–354.

[25] N. J. Zabusky, M. H. Hughes, and K. V. Roberts, J. Comput. Phys. **30**, 96 (1979).

[26] R. D. V. Buskirk and P. S. Marcus, J. Comp. Physics **115**, 302 (1994).

[27] P. S. Marcus, J. Fluid Mech. **215**, 393 (1990).

[28] P. G. Drazin and W. H. Reid, *Hydrodynamic Stability* (Cambridge University Press, Cambridge, 1981), pp. 424–435.

[29] R. Robert, J. Stat. Phys. **65**, 531 (1991).

[30] N. Whitaker and B. Turkington, Phys. Fluids **6**, 3963 (1994).

[31] D. Jin and D. Dubin, Bull. Am. Phys. Soc. **41**, 1605 (1996).

[32] D. Z. Jin and D. H. E. Dubin, Phys. Rev. Lett. **80**, 4434 (1998).

[33] H. Brands, P. H. Chavanis, R. Pasmanter, and J. Sommeria, Phys. Fluids **11**, 3465 (1999).

[34] C. E. Leith, Phys. Fluids **27**, 1388 (1984).

[35] X. P. Huang and C. F. Driscoll, Phys. Rev. Lett. **72**, 2187 (1994).

[36] J. Miller, Phys. Rev. Lett. **65**, 2137 (1990).

[37] R. Robert and J. Sommeria, J. Fluid Mech. **229**, 291 (1991).

[38] J. Miller, P. Weichman, and M. C. Cross, Phys. Rev. A **45**, 2328 (1992).

[39] H. Tennekes and J. L. Lumley, *A First Course in Turbulence* (Massachusetts Institute of Technology Press, Cambridge, 1972).

[40] P. B. Rhines, J. Fluid Mech. **69**, 417 (1975).