

ROSSBY WAVE INSTABILITY OF KEPLERIAN ACCRETION DISKS

R. V. E. LOVELACE

Department of Astronomy, Cornell University, Ithaca, NY 14853; rvl1@cornell.edu

H. LI AND S. A. COLGATE

T-6, Los Alamos National Laboratory, Los Alamos, NM 87545; hli, colgate@lanl.gov

AND

A. F. NELSON

Department of Physics, The University of Arizona, Tucson, AZ 85721; andy@as.arizona.edu

Received 1998 March 19; accepted 1998 October 15

ABSTRACT

We find a linear instability of nonaxisymmetric Rossby waves in a thin nonmagnetized Keplerian disk when there is a local maximum in the radial profile of a key function $\mathcal{L}(r) \equiv \mathcal{F}(r)S^{2\Gamma}(r)$, where $\mathcal{F}^{-1} = \hat{z} \cdot (\nabla \times v)/\Sigma$ is the potential vorticity, $S = P/\Sigma^\Gamma$ is the entropy, Σ is the surface mass density, P is the vertically integrated pressure, and Γ is the adiabatic index. We consider in detail the special case where there is a local maximum in the disk entropy profile $S(r)$. This maximum acts to trap the waves in its vicinity if its height-to-width ratio $\max(S)/\Delta r$ is larger than a threshold value. The pressure gradient derived from this entropy variation provides the restoring force for the wave growth. We show that the trapped waves act to transport angular momentum outward. A plausible way to produce an entropy variation is when an accretion disk is starting from negligible mass and temperature, therefore, negligible entropy. As mass accumulates by either tidal torquing, magnetic torquing, or Roche-lobe overflow, confinement of heat will lead to an entropy maximum at the outer boundary of the disk. Possible nonlinear developments from this instability include the formation of Rossby vortices and the formation of spiral shocks. What remains to be determined from hydrodynamic simulations is whether or not Rossby wave packets (or vortices) “hold together” as they propagate radially inward.

Subject headings: accretion, accretion disks — hydrodynamics — instabilities — waves

1. INTRODUCTION

The central problem of accretion disk theory is how angular momentum is transported outward so that matter can accrete onto the central gravitating object. Ever since Shakura & Sunyaev (1973) and Pringle (1981) outlined the basic structure of accretion disk (see also Frank, King, & Raine 1985), the origin of enhanced (or α) viscosity has been the subject of hundreds of research papers. Classically, high Reynolds number turbulence has been invoked to explain the enhanced viscosity. The driving mechanism(s) of the turbulence in high-temperature (for example, active galactic nuclei) disk systems has been a long-standing problem because of the *apparent* absence of linear instabilities in nonmagnetized Keplerian disks (see the review by Papaloizou & Lin 1995). For this reason, magnetic fields and associated instabilities have been suggested as the origin of the enhanced viscosity in disks (Balbus & Hawley 1998). The widely cited magneto-rotational instability (Velikhov 1959; Chandrasekhar 1960) occurs only when the \mathbf{B} field is weak. Also, it has been studied only for small patches of a disk using unphysical boundary conditions on the magnetic field. The nature of the global angular momentum transport remains an open problem.

This paper reconsiders the question of the stability of nonmagnetized Keplerian disks. In particular, we find a linear instability of nonaxisymmetric perturbations for conditions where the disk quantities, such as surface density and entropy, have steep radial gradients. The conditions we consider are in general nonbarotropic, which distinguishes

our work from that of Papaloizou & Pringle (1984, 1985; see also Goldreich, Goodman, & Narayan 1986; Narayan, Goldreich, & Goodman 1987). Also, in contrast with the work of Papaloizou and Pringle, the modes we consider are trapped at least initially in a narrow range of radii and therefore do not depend on reflections from inner and outer radii of the disk (or tori).

The instability we find will give rise to Rossby vortices in the nonlinear limit. These vortices are well-known in thin, two dimensional planetary atmospheres, for example, Jupiter’s “Red Spot.” The persistence of such vortices would be crucial for the hydrodynamic transport of angular momentum in accretion disks. Here we present our findings of initial conditions that cause Rossby waves to grow and defer the nonlinear behavior of the vortices to future publications.

We emphasize that the astrophysical relevance of this instability is that in circumstances where accretion disks are important, mass at low temperature and entropy accumulates at some external radius to a thickness in such a way that a condition of a local maximum in entropy is created because of confinement of heat. Then a wave of nonlinear Rossby vortex excitation carries the mass and entropy maximum inward, exciting further vortices that transport the angular momentum outwards.

In § 2, we consider the general stability of a two-dimensional nonbarotropic disk and point out the role of a key function $\mathcal{L}(r)$ in determining the stability. We then study a case with a significant “bump” in the radial dis-

tribution of entropy (or temperature) and calculate its growth rate and dispersion relation. In § 3, we discuss the possible nonlinear evolution of Rossby waves, especially their role in angular momentum transport.

2. THEORY

We consider the stability of a thin (two-dimensional), nonmagnetized Keplerian accretion disk. In treating the disk as two-dimensional we assume that the vertical thickness $2h$ is $\ll r$ and perturbation wavelengths are much longer than $2h$. Using an inertial cylindrical (r, ϕ, z) coordinate system, we have a surface mass density $\Sigma(r) = \int_{-h}^h dz \rho(r, z)$ and a vertically integrated pressure $P(r) = \int_{-h}^h dz p(r, z)$. We consider the equilibrium disk to be steady ($\partial/\partial t = 0$) and axisymmetric ($\partial/\partial \phi = 0$), with the flow velocity $\mathbf{v} = v_\phi(r)\hat{\phi}$ (that is, the accretion velocity v_r is neglected), the gravitational potential $\Phi(r) \approx -GM/r$ (the self-gravity is neglected in this study), and the radial force equilibrium, $\Sigma v_\phi^2/r = dP/dr + \Sigma d\Phi/dr$, where M is the mass of the central object. The vertical hydrostatic equilibrium gives $h \approx (c_s/v_\phi)r$, where $c_s \equiv (\Gamma P/\Sigma)^{1/2}$ is the sound speed and Γ is the adiabatic index.

2.1. Basic Equations for a Nonbarotropic Disk

The perturbations considered are in the plane of the disk so that the perturbed surface mass density is $\tilde{\Sigma} = \Sigma + \delta\Sigma(r, \phi, t)$, the perturbed vertically integrated pressure is $\tilde{P} = P + \delta P(r, \phi, t)$, and the perturbed flow velocity is $\tilde{\mathbf{v}} = \mathbf{v} + \delta\mathbf{v}(r, \phi, t)$ with $\delta\mathbf{v} = (\delta v_r, \delta v_\phi, 0)$. The equations for the perturbed disk are

$$\frac{D\tilde{\Sigma}}{Dt} + \tilde{\Sigma} \nabla \cdot \tilde{\mathbf{v}} = 0, \quad (1a)$$

$$\frac{D\tilde{\mathbf{v}}}{Dt} = -\frac{1}{\tilde{\Sigma}} \nabla \tilde{P} - \nabla \Phi, \quad (1b)$$

$$\frac{D}{Dt} \left(\frac{\tilde{P}}{\tilde{\Sigma}^\Gamma} \right) = 0, \quad (1c)$$

where $D/Dt \equiv \partial/\partial t + \mathbf{v} \cdot \nabla$, and we refer to $S \equiv P/\Sigma^\Gamma$ as the entropy of the disk matter. Equation (1c) corresponds to the isentropic behavior of the disk matter. We use equation (1c) for simplicity rather than the rigorous relation $D(p/\rho^\gamma)/Dt = 0$ (see Papaloizou & Lin 1995).

We consider perturbations $\propto f(r) \exp(im\phi - i\omega t)$, where m is the azimuthal mode number and ω the angular frequency. From equation (1a), we have

$$i\Delta\omega\delta\Sigma = \nabla \cdot (\Sigma\delta\mathbf{v}), \quad (2)$$

where $\Delta\omega(r) \equiv \omega - m\Omega(r)$ and $\Omega = v_\phi/r \approx (GM/r^3)^{1/2}$. From equation (1b) we have

$$i\Delta\omega\delta v_r + 2\Omega\delta v_\phi = \frac{1}{\Sigma} \frac{\partial\delta P}{\partial r} - \frac{\delta\Sigma}{\Sigma^2} \frac{\partial P}{\partial r}, \quad (3a)$$

$$i\Delta\omega\delta v_\phi - \frac{\kappa^2}{2\Omega} \delta v_r = ik_\phi \frac{\delta P}{\Sigma}. \quad (3b)$$

Here $\kappa \equiv [r^{-3}d(r^4\Omega^2)/dr]^{1/2} \approx \Omega$ is the radial epicyclic frequency and $k_\phi = m/r$ is the azimuthal wavenumber. From equation (1c), we have

$$\delta P = c_s^2 \delta\Sigma - \frac{i\Sigma c_s^2}{\Delta\omega L_s} \delta v_r, \quad (4)$$

which is the perturbed equation of state, and

$$L_s \equiv \Gamma \left/ \left[\frac{d}{dr} \ln \left(\frac{P}{\Sigma^\Gamma} \right) \right] \right. \quad (5)$$

is the length scale of the entropy variation in the disk. The nonbarotropic nature of the disk renders L_s finite, whereas for a barotropic disk (i.e., P is a function of Σ only), $L_s \rightarrow \infty$.

To further simplify the momentum equation, we can write the right-hand side of equation (3a) as $(\delta P/\Sigma)' + (\delta P/\Sigma^2)\Sigma' - (\delta\Sigma/\Sigma^2)P'$, where the prime denotes $\partial/\partial r$. Next, equation (4) can be used to express this as $(\delta P/\Sigma)' - \delta P/(\Sigma L_s) - i\delta v_r c_s^2/(\Delta\omega L_s L_p)$, where

$$L_p \equiv \Gamma \left/ \left[\frac{d}{dr} \ln(P) \right] \right. \quad (6)$$

is the length-scale of the pressure variation. For definiteness, we assume $h < (|L_s|, |L_p|) \lesssim r$.

Using $\Psi \equiv \delta P/\Sigma$ as our key variable, we can now rewrite equations (3a)–(3b) as

$$i \left(\Delta\omega + \frac{c_s^2}{\Delta\omega L_s L_p} \right) \delta v_r + 2\Omega\delta v_\phi = \frac{\partial\Psi}{\partial r} - \frac{\Psi}{L_s}, \quad (7a)$$

$$i\Delta\omega\delta v_\phi - \frac{\kappa^2}{2\Omega} \delta v_r = ik_\phi \Psi. \quad (7b)$$

Solving equations (7a)–(7b) for $(\delta v_r, \delta v_\phi)$ gives

$$\Sigma\delta v_r = i\mathcal{F} \left[\frac{\Delta\omega}{\Omega} \left(\Psi' - \frac{\Psi}{L_s} \right) - 2k_\phi \Psi \right], \quad (8a)$$

$$\Sigma\delta v_\phi = \mathcal{F} \left[-k_\phi \left(\frac{\Delta\omega}{\Omega} + \frac{c_s^2}{\Delta\omega\Omega L_s L_p} \right) \Psi + \frac{\kappa^2}{2\Omega^2} \left(\Psi' - \frac{\Psi}{L_s} \right) \right], \quad (8b)$$

where

$$\mathcal{F}(r) \equiv \frac{\Sigma\Omega}{\kappa^2 - \Delta\omega^2 - c_s^2/(L_s L_p)}. \quad (9)$$

For ‘‘corotation modes’’ where $|\Delta\omega|^2$ is small compared with κ^2 and $c_s^2/|L_s L_p|$, we have $\mathcal{F} \approx \Sigma\Omega/\kappa^2 = \Sigma/2\omega_z$. Here $\omega_z \equiv \hat{z} \cdot (\nabla \times \mathbf{v})$ is the vorticity. The ratio $\omega_z/\Sigma \approx 1/(2\mathcal{F})$ is commonly referred to as the potential vorticity (see, for example, Drazin 1978).

We next use the continuity equation (2) and equations (4) and (8a)–(8b) to obtain

$$\frac{1}{r} \left(\frac{r\mathcal{F}}{\Omega} \Psi' \right)' - \frac{k_\phi^2 \mathcal{F}}{\Omega} \Psi = \frac{\Sigma\Psi}{c_s^2} + \frac{2k_\phi \mathcal{F}'}{\Delta\omega} \Psi + \left[\frac{\mathcal{F}}{\Omega L_s^2} + \frac{1}{r} \left(\frac{r\mathcal{F}}{\Omega L_s} \right)' + \frac{4k_\phi \mathcal{F}}{\Delta\omega L_s} + \frac{k_\phi^2 c_s^2 \mathcal{F}}{\Delta\omega^2 \Omega L_s L_p} \right] \Psi. \quad (10)$$

Equation (10) allows the determination of ω and identification of mode structure for general disk flows, stable or unstable. If the flow is barotropic (as has been studied extensively in the literature), all the terms on the lower line of this equation vanish, $|L_s| \rightarrow \infty$. In this paper, we explore additional effects that arise when $|L_s|$ is finite.

2.2. Instability Condition

A useful quadratic form can be obtained by multiplying equation (10) by Ψ^* and integrating over the disk.

Assuming $\Psi^* \Psi' r \mathcal{F} / \Omega \rightarrow 0$ for $r \rightarrow 0, \infty$, we obtain

$$\begin{aligned} & - \int d^2r \frac{\mathcal{F}}{\Omega} (|\Psi'|^2 + k_\phi^2 |\Psi|^2) \\ & = \int d^2r \frac{\Sigma}{c_s^2} |\Psi|^2 + \int d^2r \left[\frac{\mathcal{F}}{\Omega L_s^2} + \frac{1}{r} \left(\frac{r \mathcal{F}}{\Omega L_s} \right)' \right] |\Psi|^2 \\ & + 2 \int d^2r \left(\frac{k_\phi \mathcal{F}'}{\Delta \omega} + \frac{2k_\phi \mathcal{F}}{\Delta \omega L_s} \right) |\Psi|^2 + \int d^2r \frac{k_\phi^2 c_s^2 \mathcal{F}}{\Delta \omega^2 \Omega L_s L_p} |\Psi|^2. \end{aligned} \quad (11)$$

Note that only the last two integral terms on the right-hand side of this equation involve the complex variable $\Delta \omega(r)$.

Equation (10) can be solved by several methods (e.g., by a relaxation method or by shooting), but we defer the full discussion of its solutions to a future publication. In order to bring out the essential aspects of equation (10), we make simplifying assumptions that allow an analytic solution. First, when $|\Delta \omega|$ is a finite fraction of Ω and $c_s/v_\phi \ll 1$, we find that the integral involving $\Delta \omega^2$ is small compared to the integral with $\Delta \omega$, so that the last term in equation (11) is negligible. Furthermore, for the corotation modes where $|\Delta \omega|^2/\Omega^2 \ll 1$ and for conditions where the pressure variation with r is such that κ^2/Ω^2 is positive, the function \mathcal{F} is positive and real. It is physically possible to have κ^2 positive and yet have a finite $|L_s|$, because the pressure and entropy variations need not be the same. In fact, this is precisely a feature of nonbarotropic flow (see § 3). Thus the terms containing $\Delta \omega$ in equation (11) can be combined to give

$$2 \int d^2r \frac{k_\phi \mathcal{F}}{\Delta \omega} [\ln(\mathcal{F} S^{2/\Gamma})]' |\Psi|^2. \quad (12)$$

Putting equation (12) back into equation (11) and taking the imaginary part of equation (11), we find that a solution is possible only for conditions where the imaginary part of (12) vanishes (see also Lovelace & Hohlfield 1978; hereafter LH). Thus we find that an instability is possible only if $[\ln(\mathcal{F} S^{2/\Gamma})]'$ vanishes at some r ; in other words, the key function $\mathcal{L}(r) \equiv \mathcal{F}(r) S^{2/\Gamma}(r)$ has a maximum or minimum.

This newly defined function $\mathcal{L}(r)$ warrants further discussion. We believe that this expression is new and more general than those given in previous studies. With the above assumptions, we can write $\mathcal{L}(r)$ as

$$\mathcal{L}(r) = \mathcal{F} S^{2/\Gamma} \approx \frac{\Sigma \Omega S^{2/\Gamma}}{\kappa^2} \propto \frac{\Omega}{\kappa^2} \Sigma^{2/\Gamma-1} T^{2/\Gamma}, \quad (13)$$

where $\Omega^2(r)$ and $\kappa^2(r)$ are of course constrained by the radial force balance in the equilibrium.

An extremum of $\mathcal{L}(r)$ necessary for instability could come from several sources: (1) In the limit of incompressible and isothermal flow (see Staley & Gall 1979 for an application to flow in a tornado), an extremum in the radial profile of $\Omega(r)$ is required to have unstable vortices. (2) In the studies of LH and Sellwood & Kahn (1991), a cold disk was assumed, and $S(r)$ was effectively taken as a constant. In this case $\mathcal{L}(r)$ is simply $\mathcal{F} = \Sigma/\Omega$, which must have a maximum or a minimum for instability. Similarly, in the cases considered by Papaloizou & Pringle (1984, 1985), where a polytropic (barotropic) $P = P(\Sigma)$ equation of state is assumed, $S(r)$ is again a constant throughout the disk, so that a maximum or minimum of $\mathcal{L}(r) \propto \mathcal{F}$ is needed for instability. (3) In the present study, we treat the more general

case where $S(r)$ varies across the disk. An extremum in $\mathcal{L}(r)$ can result from having an extremum in $S(r)$ directly (that is, a local ‘‘bump’’ in the temperature profile of the disk) or by having a large enough gradient in $S(r)$ that causes an extremum in the epicyclic frequency κ^2 owing to the radial force balance. Furthermore, we find that instability occurs when a threshold in the variation of $\mathcal{L}(r)$ is exceeded. In view of equation (13), a variation of $T(r)$ is more important in giving a change in $\mathcal{L}(r)$ than a comparable fractional variation of $\Sigma(r)$.

2.3. Thermodynamic Driving of Vorticity

It is straightforward to obtain the vorticity equation (see Pedlosky 1987; Spiegel 1993) by taking the curl of equation (1b) and combining it with equation (1a),

$$\frac{D}{Dt} \left(\frac{\omega_z}{\Sigma} \right) = \frac{\nabla \Sigma \times \nabla P}{\Sigma^3}, \quad (14)$$

where again, $\omega_z = \hat{z} \cdot \nabla \times \mathbf{v}$ is the vorticity. The right-hand side of this equation represents the thermodynamic driving of the vorticity. For a barotropic flow, $P = P(\Sigma)$,

$$\frac{D}{Dt} \left(\frac{\omega_z}{\Sigma} \right) = 0. \quad (15)$$

This is the case for most previous studies (see Papaloizou & Pringle 1984, 1985; Goldreich et al. 1986; Narayan et al. 1987). Equation (15) says that each fluid element conserves its specific vorticity. In contrast, the term $\nabla \Sigma \times \nabla P \propto \nabla T \times \nabla S$ destroys this conservation and allows pressure forces to ‘‘produce’’ vorticity in the flow. In the nonbarotropic case, surfaces (or lines) of constant surface density and surfaces of constant specific entropy do not coincide.

2.4. Dispersion Relation

For the corotation modes ($|\Delta \omega|^2 \ll \Omega^2$) and $\mathcal{F}(r) \approx \Sigma/\Omega$, an approximate dispersion relation follows from equation (10) by taking the r -dependence $\Psi \propto \exp(ik_r r)$ with $(k_r r)^2 \gg 1$, $(k_r L_s)^2 \gg 1$, and $(k_r L_p)^2 \gg 1$. The dominant terms in equation (10) are

$$- \frac{k^2 \Sigma}{\Omega^2} \Psi = \frac{\Sigma}{c_s^2} \Psi + \frac{2k_\phi \Sigma}{\Delta \omega \Omega} (\ln \mathcal{L})' \Psi + \frac{k_\phi^2 c_s^2 \Sigma}{\Delta \omega^2 \Omega^2 L_s L_p} \Psi, \quad (16)$$

or

$$\Delta \omega = - \frac{k_\phi c_s^2 / \Omega}{1 + k^2 h^2} \left[(\ln \mathcal{L})' \pm \sqrt{[(\ln \mathcal{L})']^2 - \frac{1 + k^2 h^2}{L_s L_p}} \right], \quad (17)$$

where $k^2 \equiv k_r^2 + k_\phi^2$ and $h \equiv c_s/\Omega$. The dispersion relation (17) is the disk analog of the Rossby wave dispersion relation (see, for example, Brekhovskikh & Goncharov 1993). Note in particular that the group velocity of the waves in the corotating reference frame is in the direction opposite to that of the disk rotation. Instability is possible if $[(\ln \mathcal{L})']^2 < (1 + k^2 h^2)/(L_s L_p)$, which requires $L_s L_p > 0$.

From equation (17), the maximum growth rate occurs for $(\ln \mathcal{L})' = 0$, and it is

$$\frac{\max(\omega_i)}{\Omega} = \frac{|k_\phi| c_s^2 / \Omega^2}{(1 + k^2 h^2)^{1/2} (L_s L_p)^{1/2}} \approx \frac{|k_\phi| h}{(1 + k^2 h^2)^{1/2}} \left(\frac{c_s}{v_K} \right) \left(\frac{r^2}{L_s L_p} \right)^{1/2}, \quad (18)$$

where v_K is the Keplerian velocity and $\omega_i \equiv \text{Im}(\omega)$ is the growth rate. The maximum growth rate is $< \Omega$ for a thin disk where $h/r \approx c_s/v_\phi \ll 1$ and for $L_s L_p > h^2$. Note in particular the geometrical optics relation for a wave packet

$$\frac{dk_r}{dt} = -\frac{\partial \omega}{\partial r} \approx \frac{3}{2} k_\phi \Omega, \quad \frac{dk_\phi}{dt} = -\frac{\partial \omega}{\partial \phi} = 0, \quad (19)$$

so that $k_r \approx \frac{3}{2} k_\phi \Omega t + \text{const.}$ and $k_\phi = \text{const.}$ Thus a wave packet initially with $|k_r h| \ll 1$ is rapidly sheared by the differential rotation. In a time of order $t_{\text{shear}} = (\frac{3}{2} k_\phi h \Omega)^{-1}$, the wave evolves into having a very short radial-wavelength, $|k_r h| > 1$, trailing-spiral wave. This is discussed in detail by Goldreich & Lynden-Bell (1965). The maximum amplification a wave can have is $A = \exp(\int dt \omega_i) \sim \exp[\max(\omega_i) t_{\text{shear}}] \approx \exp[(2/3)(c_s/v_K)(r^2/L_s L_p)^{1/2}]$. The amplification is $A = \mathcal{O}(1)$ and is probably not significant for $L_s L_p > h^2$.

2.5. Trapped Modes

Here we consider conditions where $\mathcal{L} \approx (\Sigma/\Omega) S^{2/\Gamma}$ has a maximum or minimum as a function of r at r_0 of width $\Delta r \ll r_0$. The case where $\mathcal{F}(r) = \Sigma/\Omega$ has an extremum was treated earlier by LH. Therefore we consider the case where $\mathcal{F} \approx \text{const.}$ and where $S(r)$ has an extremum. The extremum of $S(r)$ is thus an extremum of $T(r)$. As mentioned earlier, a variation of $T(r)$ is more important in giving a change in $\mathcal{L}(r)$ than the variation of $\Sigma(r)$. For a maximum in $S(r)$, we find using equations (17) and (19) that the radial group velocity of the waves $v_{gr} = \partial \omega / \partial k_r$ is directed towards r_0 , whereas for a minimum the waves propagate away from r_0 . Thus a maximum of $S(r)$ can give trapping of the waves in the vicinity of r_0 . This trapping obviates the need for reflecting disk boundaries considered by Goldreich et al. (1986).

For definiteness we take

$$S(r) = S_1 + (S_2 - S_1) \exp[-(r - r_0)^2 / (2\Delta r^2)], \quad (20)$$

where S_1, S_2, r_0 , and Δr are constants. For simplicity we consider $\Delta r^2 \ll r_0^2$. In this limit, the real and imaginary parts of equation (10) for $\Psi = \delta P / \Sigma = \Psi_r + i\Psi_i$ can be written as

$$\begin{aligned} \Psi_r'' &= (K^2)_r \Psi_r - (K^2)_i \Psi_i, \\ \Psi_i'' &= (K^2)_r \Psi_i + (K^2)_i \Psi_r. \end{aligned} \quad (21)$$

Here we have omitted the detailed expressions for $(K^2)_r$ and $(K^2)_i$, both of which are real functions of $L_s, c_s/v_\phi, \Delta\omega/\Omega$, and m (see Figs. 1–3). It can be shown that $(K^2)_r$ and $(K^2)_i$ possess the following properties (assuming $\omega_r/[m\Omega(r_0)] = 1$): $(K^2)_i$ is an odd function of $r - r_0$ and is $\rightarrow 0$ when $(r - r_0)^2 \gg \Delta r^2$, but $(K^2)_r$ is an even function of $r - r_0$; $(K^2)_r$ is < 0 for $r \rightarrow r_0$, which corresponds to a classically allowed region of motion in a potential well, and it is > 0 for $(r - r_0)^2 \gg \Delta r^2$, which corresponds to a forbidden region. It is this potential well that gives rise to “bound states,” which determine the growth rates ω_i .

The existence of wave growth ($\omega_i > 0$) and its dependence on S_2/S_1 can be obtained by first considering the potential well for $\omega_i \rightarrow 0$. The first appearance of a bound state in the well is given by the Bohr-Sommerfeld quantization condition,

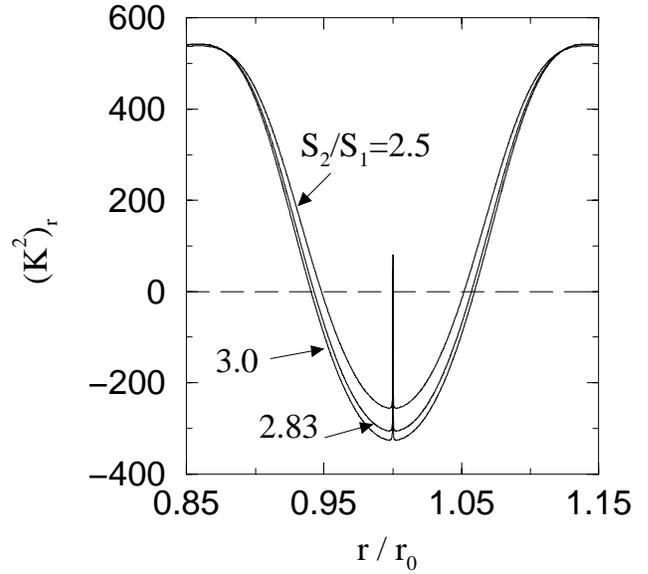


FIG. 1.—Sample profiles of $(K^2)_r$, in units of $1/r_0^2$ shown for different values of S_2/S_1 . For this plot $m = 5$, $\Delta r/r_0 = 0.05$, $\Gamma = \frac{5}{3}$, unperturbed $c_s/v_\phi = 0.045$, $L_s = L_p$, and $\omega_r/\Omega = 10^{-3}$. The spike at $r/r_0 = 1$ is due to the finite value of ω_i . For the intermediate curve, the Bohr-Sommerfeld condition (eq. [22]) is satisfied to a good approximation. For other values of m between 1 and 10, the critical value for the bound state is $S_2/S_1 \approx 2.7 + 0.83(m/10)^{5/2}$.

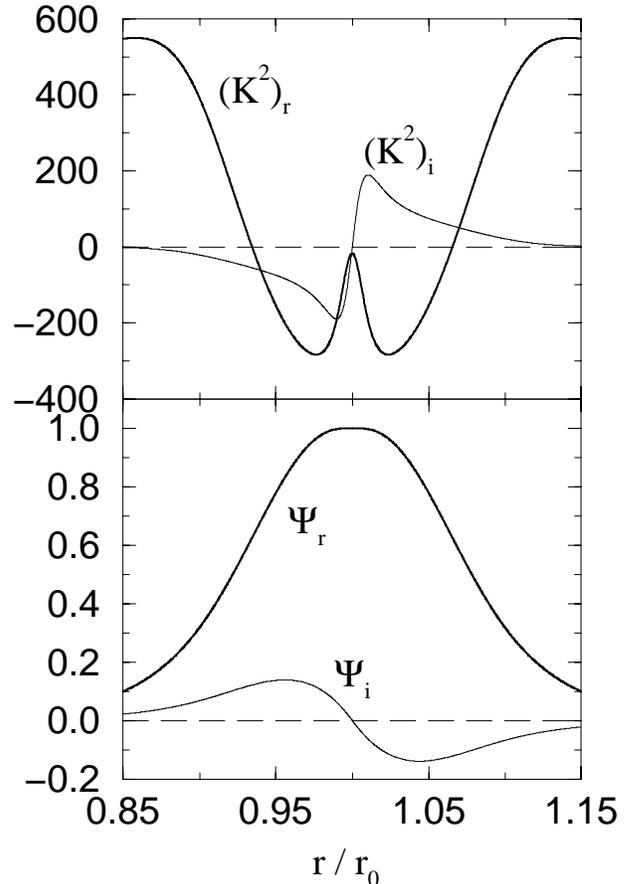


FIG. 2.—Sample profiles of $(K^2)_{r,i}$ and $\Psi_{r,i}$, obtained by a shooting method, are shown for $\omega_i/\Omega = 0.1$, $m = 5$, and $\Delta r/r_0 = 0.05$. For this case, $S_2/S_1 \approx 3.744$ and $\Psi_r(r_0) \approx -6.543$.

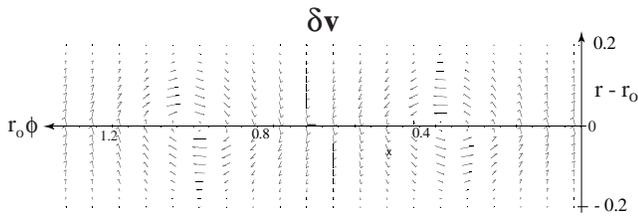


FIG. 3.—Velocity field of the unstable perturbation shown in Fig. 2, obtained from eqs. (8a)–(8b) and (21).

$$f(S_2/S_1) \equiv \oint dr (-K^2)_r^{1/2} = \pi. \quad (22)$$

This condition corresponds to a critical value of S_2/S_1 , so that there can be a bound state with $\omega_i/\Omega > 0$ (Fig. 1; *middle curve*). Physically this is the threshold required for the growth of Rossby waves. Again, we emphasize that the existence of wave growth depends critically on the non-barotropic nature of disks, because the main contribution to this potential well comes from $4k_\phi \mathcal{F} \Psi / (\Delta \omega L_s)$, which is always nonpositive since $\Re(1/\Delta \omega) \propto r - r_0$, whereas $1/L_s \propto r_0 - r$. This term vanishes when $L_s \rightarrow \infty$.

Figure 2 shows sample results for $(K^2)_{r,i}$ and $\Psi_{r,i}$ by solving equations (21) using a shooting method. We omit discussion of the detailed numerical method, but we only point out that one can make use of the symmetry of the potential well as well as the boundary conditions at $r = r_0$ and $r \rightarrow \infty$. The key physical variables are S_2/S_1 and ω_i/Ω . We find ω_i to have the approximate dependence $\omega_i/\Omega \approx k_1 [(S_2/S_1) - k_2]^{2/3}$ for $S_2/S_1 \geq k_2$, where k_2 is the mentioned critical value and k_1 is a constant.

Figure 3 shows the vortical nature of the velocity field for the same unstable case as Figure 2.

2.6. Angular Momentum Transport

To calculate the flux of angular momentum, we first note that, with the complex factors $\exp(im\phi - i\omega t)$ removed, $\delta v_r = \delta v_r^O + i\delta v_r^E$, $\delta v_\phi = \delta v_\phi^O + i\delta v_\phi^E$, and $\delta \Sigma = \delta \Sigma^E + i\delta \Sigma^O$, where the E, O superscripts denote even or odd *real* functions of $r - r_0$. Such properties are obtained by expressing them using equation (1a)–(1c) and $\Psi = \Psi_r^E + i\Psi_i^O$. Thus, across a circle of radius r_0 , we have

$$\begin{aligned} F_J(r_0) &= \int_0^{2\pi} d\phi r_0^2 \Re(\Sigma) \Re(\delta v_r) \Re(\delta v_\phi) = \pi r_0^2 \Sigma \delta v_r^E(r_0) \delta v_\phi^E(r_0) \\ &= \frac{2\pi m^2 \Sigma \omega_i}{\Omega^3} \left(\Psi_r + \frac{\omega_i r_0}{2\Omega m} \Psi_i' \right)_{r_0} \left(\Psi_r - \frac{\Omega r_0}{2\omega_i m} \Psi_i' \right)_{r_0}. \end{aligned} \quad (23)$$

We find in general that $m\Psi_i'(r_0) < 0$. Further, we find that each of the two factors in parentheses in equation (23) are positive. Therefore the trapped wave causes the *outward* flow of angular momentum across $r = r_0$. Note that the positive transport is from the nonlinear term $(\delta v_r, \delta v_\phi)$, since the first order term vanishes.

3. DISCUSSION

The astrophysical importance of this Rossby wave instability is the implication that in the nonlinear limit multiple Rossby vortices will form after these waves break and coalesce in a few revolutions. Hence they may provide a plausible explanation of the hydrodynamic transport of angular momentum in thin, Keplerian accretion disks. The

critical condition for the inception of this hydrodynamic transport, as derived in this linear case, is a minimum contrast (a factor of ~ 3) in the radial distribution of entropy. Thus it is important that heat (and thus entropy) be preserved against cooling by radiation in at least tens of revolutions. Roughly speaking, recognizing the large dynamic range and uncertainty of opacities, this implies a minimum mass “thickness” $\Sigma \sim 100\text{--}1000 \text{ g cm}^{-2}$. An important implication from this minimum mass thickness requirement is the prediction on the turn-on (active state) and turn-off (quiescent state) conditions for the angular momentum transport by the Rossby vortices. Specifically, we could envisage the following scenario: as mass accumulates in the disk by accretion, the resulting trapping of heat allows the temperature to rise by the energy input of accretion and an episode of Rossby vortex accretion ensues. Depending upon the heating, opacity conditions, and mass accretion rates, there will be a radius, a thickness, and therefore a mass where the Rossby instability initiates. This mass and mass accretion rate determine the event luminosity and episode rate. Finally we note that a given episode terminates when its mass is accreted, presumably rapidly, before additional mass is supplied. This understanding of the turn-on and turn-off conditions for Rossby vortex transport may prove relevant for several astrophysical systems, such as the planet formation after the solar nebula has cooled (Cameron 1962) and outbursts from cataclysmic variables and stellar mass black holes (Vishniac 1997).

Although a localized high-entropy region ($S_2/S_1 \sim 3$) may appear to be a large deviation in $S(r)$, the resultant pressure gradient force is still very small compared to the gravitational force. In the case represented by all the figures, we have used a background $c_s/v_\phi = 0.045$, which gives $c_s/v_\phi \approx 0.078$ at the peak of bump ($S_2/S_1 \sim 3$). But this deviation in $S(r)$ is sufficient to drive the growth of non-axisymmetric perturbations. A possible mechanism of formation of an entropy bump is from the interaction between the matter from Roche lobe overflow and the accretion disk. A spiral shock can form in this situation (Sawada, Matsuda, & Hachisu 1986).

Bracco et al. (1998) have studied *barotropic* disks with initial perturbations in the potential vorticity (ω_z/Σ) profile. They showed that vortices can have a rather long life time. Another possible nonlinear effect is the formation of spiral shocks that heat the disk and thereby maintain the bump against radiative cooling. In either case, requiring a finite entropy bump for the instability to grow implies a maximum cooling rate above which a bump that develops in the disk quickly decays away and such an enhanced viscosity model turns off. What remains to be determined is whether the bump “holds together” as it propagates radially inward. If these Rossby vortices are indeed playing a major role in angular momentum transport, then they are fundamentally different from the usual α viscosity prescription, which is based on homogeneous turbulence. In contrast, Rossby vortices are large-scale, coherent structures.

Preliminary numerical simulations using a high-resolution piecewise parabolic method code (Nelson et al. 1999) appear to confirm the existence of a threshold on S_2/S_1 for wave growth and the formation of anticyclonic vortices. Numerical studies are now underway to follow the nonlinear evolution of the vortices in nonbarotropic disks and to study the associated transport properties (Nelson et al. 1999).

We thank the referee for clarifications of many issues, which greatly improved this paper. The experimental work on the flow visualization for a liquid sodium dynamo by R. Ferrel and Van Romero is gratefully acknowledged in stimulating this analysis. Discussion with P. Yecko is also gratefully acknowledged. The work of R. V. E. L. was sup-

ported in part by NSF grant AST-9320068 and NASA grant NAG5-6311. H. L. gratefully acknowledges the support of an Oppenheimer Fellowship. The work of H. L. and S. A. C. was performed under the auspices of the US Department of Energy.

REFERENCES

- Balbus, S. A., & Hawley, J. F. 1998, *Rev. Mod. Phys.*, 70, 1
 Bracco, A., et al. 1998, in *Quasars and Accretion Disks*, ed. M. Abramowicz (Cambridge: Cambridge Univ. Press), in press
 Brekhovskikh, L. M., & Goncharov, V. 1993, *Mechanics of Continua and Wave Dynamics* (Berlin: Springer), 246
 Cameron, A. G. W. 1962, *Icarus*, 1, 13
 Chandrasekhar, S. 1960, *Proc. Natl. Acad. Sci.*, 46, 253
 Drazin, P. G. 1978, in *Rotating Fluids in Geophysics*, ed. P. H. Roberts & A. M. Soward (London: Academic Press), 139
 Frank, J., King, A. R., & Raine, D. J. 1985, *Accretion Power in Astrophysics* (Cambridge: Cambridge Univ. Press)
 Goldreich, P., Goodman, J., & Narayan, R. 1986, *MNRAS*, 221, 339
 Goldreich, P., & Lynden-Bell, D. 1965, *MNRAS*, 130, 125
 Lovelace, R. V. E., & Hohlfield, R. G. 1978, *ApJ*, 221, 51
 Narayan, R., Goldreich, P., & Goodman, J. 1987, *MNRAS*, 228, 1
 Nelson, A. F., Li, H., Colgate, S. A., & Lovelace, R. V. E. 1999, in preparation
 Papaloizou, J. C. B., & Lin, D. N. C. 1995, *ARA&A*, 33, 505
 Papaloizou, J. C. B., & Pringle, J. E. 1984, *MNRAS*, 208, 721
 ———. 1985, *MNRAS*, 213, 799
 Pedlosky, J. 1987, *Geophysical Fluid Dynamics* (Springer: New York), chap. 2
 Pringle, J. E. 1981, *ARA&A*, 19, 137
 Sawada, K., Matsuda, T., & Hachisu, I. 1986, *MNRAS*, 219, 75
 Sellwood, J. A., & Kahn, F. D. 1991, *MNRAS*, 250, 278
 Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, 24, 337
 Spiegel, E. A. 1993, *Astrophysical Fluid Dynamics*, ed. J.-P. Zahn & J. Zinn-Justin (Oxford: Elsevier Science Publishers), 1
 Staley, D. O., & Gall, R. L. 1979, *J. Atmos. Sci.*, 36(6), 973
 Velikhov, E. 1959, *Soviet Phys.-JETP Lett.*, 36, 1938
 Vishniac, E. T. 1997, *ApJ*, 482, 414