THE ASTROPHYSICAL JOURNAL, 400:610–621, 1992 December 1

A POWERFUL LOCAL SHEAR INSTABILITY IN WEAKLY MAGNETIZED DISKS. IV. NONAXISYMMETRIC PERTURBATIONS

STEVEN A. BALBUS AND JOHN F. HAWLEY
Virginia Institute for Theoretical Astronomy, Department of Astronomy, University of Virginia, Charlottesville, VA 22903-0818
Received 1992 February 28; accepted 1992 June 4

ABSTRACT

In this paper, we continue our study of a powerful axisymmetric MHD instability recently put forward by the authors as the underlying cause of anomalous transport in accretion disks. The theory of local nonaxisymmetric perturbations in weakly magnetized disks is presented. Such disturbances are of interest for several reasons, most notably in the context of a dynamo magnetic field amplification scheme. The most volatile disturbances are those associated with the presence of a poloidal field, which grow by tens of orders of magnitude with e-folding times measured in fractions of an orbital period. We also examine the stability of a purely azimuthal field configuration and find that nonaxisymmetric instability is present, but with a growth time measured in tens of orbital periods. In general, the most rapid growth occurs for very small radial and azimuthal wavenumbers, leading to coherent magnetic field structure in planes parallel to the disk. We suggest that this instability will prove to be a key ingredient for the generation of magnetic fields in disks.

Subject headings: accretion, accretion disks — instabilities — MHD

1. INTRODUCTION

Recently, Balbus & Hawley (1991, hereafter BH; 1992) and Hawley & Balbus (1991, hereafter HB) have identified a powerful and generic linear axisymmetric MHD instability in differentially rotating systems. The remarkable nature of this instability stems from its rapid growth rate (in a Keplerian disk, the most rapidly growing wavenumbers have an e-folding time equal to $2/3\pi$ times the rotation period), and its complete indifference to many properties of the magnetic field. Provided only that some poloidal component of the field is present and the total field strength is subthermal, the maximal growth rate is independent of both the magnetic field strength and the field geometry. It has been argued that the growth rate associated with this instability is likely to be the most rapid that any process feeding off the differential shear can attain (Balbus & Hawley 1992). This mechanism thus seems to be an extremely promising candidate for the physical basis of the "anomalous transport" long conjectured to be present in accretion disks. In this paper, we extend our original study by examining the stability of weakly magnetized disks to nonaxisymmetric disturbances.

The investigation of finite azimuthal wavenumbers allows the case of a purely toroidal field to be considered, whereas the orthogonality of the field and the wavenumber vectors precludes an axisymmetric treatment of the field geometry. We find that such a field configuration is also essentially unstable (more accurately, it has an extended exponential growth phase), but with a smaller growth rate than is shown by fields which also have a poloidal component. Another technical issue of interest is a simple by-product of the calculation. In BH is was stated, but not explicitly shown, that axisymmetric perturbations develop in disks in a simple exponential or oscillatory manner even when shear causes the background azimuthal field to grow linearly with time. In the analysis we present here, the time dependence of the perturbations is not assumed a priori. Under this more general assumption, we show that the axisymmetric limit of our equations corresponds to the general dispersion formula presented in BH, thereby confirming the original claim.

But the fate of nonaxisymmetric disturbances is a more far reaching issue, for it bears on another well-known problem of astrophysical theory: the origin of cosmic magnetic fields. Does the instability presented in BH lead to large-scale magnetic field amplification in differentially rotating disks? A traditional approach to field amplification has been to postulate fluid velocity fields which lead to turbulent dynamo amplification (Moffatt 1978; Parker 1979; Cowling 1981; Zel'dovich, Ruzmaikin, & Sokoloff 1983). In such schemes, much is made of the need to close the dynamo equations. By this it is meant that whereas differential rotation generates a toroidal field component from a poloidal one, it is not so straightforward to close the feedback loop and obtain the sustained generation of poloidal field from toroidal. In classical dynamo theory, this issue is addressed by appealing to background turbulence with nonvanishing mean helicity (Zel'dovich et al. 1983; Cowling 1981), giving rise to the so-called "α" process in the $\alpha\omega$ dynamo. (This should not be confused with the turbulent viscosity α parameter.) Taken at face value, however, the results of BH imply that to obtain exponential growth of a weak disk field requires little more than differential rotation. "Little more" means linear perturbations. Poloidal field seems to beget not only toroidal field, but through the instability, additional poloidal field as well.

The nonlinear behavior of a small, weakly magnetized meridonal slice of a Keplerian disk has recently been studied numerically by Hawley & Balbus (1992, hereafter Paper III), the companion paper to this article. With the imposition of periodic boundary conditions, a constant gas density, and a nonvanishing dipole moment for the magnetic field, the simulations showed persistent and dramatic field growth. In many cases, no indication of saturation was evident after many rotation periods. A common nonlinear resolution seen in the simulations appears as a sort of streaming motion. Two channels of fluid, separated along the vertical axis and consisting of oppositely directed high and low angular momentum material, flow radially with a steadily increasing fluid velocity. J. Goodman (1991) has made the important observation that the exact, nonlinear solution for incompressible radial displacements in a vertical magnetic field with a shearing-sheet geometry is identical to the linear exponential solution. Is the channel solution a dynamo? Does the existence of this solution "close the dynamo equations" without the need to postulate an ad hoc turbulent background?

One difficulty with this conclusion is that it is openly contemptuous of a famous theorem: that it is impossible to sustain isolated dynamo activity by axisymmetric fluid motions alone (Moffatt 1978). Classical dynamo theory confronts this by arguing that it is only on average that axisymmetry may prevail, and that the true motion of the fluid turbulence is considerably more complicated on microscopic scales. But the key word in the statement of the theorem is "isolated." The critical step in the proof requires that surface integrals at infinity involving magnetic field lines vanish (e.g., Paper III.) For dipole or higher order fields, this procedure can be justified. For an accretion disk threaded by a field that may be Alfvénically coupled to an interstellar ambient field or to a central star, it may be more problematic to ignore these surface integrals. Also, on scales large enough to be regarded as "infinity," the assumption of axisymmetry itself must break down, another reason to question the direct applicability of the theorem to accretion disks.

It is only the restrictive symmetry imposed by the assumption of axisymmetric fluid motions that allows the induction equation to be cast in a conserved current form, which is critical to obtain the anti-dynamo result. More general fluid behavior expected in physical systems has no such global constraints. Whatever the difficulties of understanding the nature of local axisymmetric field amplification in systems connected globally to a more complex ambient magnetic topology, local *nonaxisymmetric* disturbances present no problems of this sort. Thus, the fate of this class of perturbation is of particular interest in the study of the generation of magnetic fields in both accretion and galactic disks.

Focusing attention on formal closure of the dynamo equations seems to us to be very much the wrong approach to understanding field amplification in disks. One obvious difficulty is posed by the existence of the channeling solutions. More generally, the presence of the instability renders untenable the assumption of independent magnetic field and velocity fluctuations, which is always used as the starting point of kinematical dynamo theory (Zel'dovich et al. 1983). However, before the instability of BH can be regarded as a viable large-scale magnetic field generator, it needs to be shown that nonaxisymmetric disturbances evolve productively: that they do not, for example, ultimately break the field down into uncorrelated parcels. This would be inconsistent with such observations as the large-scale correlation of magnetic fields seen in radio observations of disk galaxies (e.g., Mathewson, van der Kruit, & Brouw 1972). The suggestion sometimes voiced in the literature that large wavelengths are associated with the maximal linear growth rate of the axisymmetric instability, and therefore that only local uncorrelated structures are to be expected, is not in itself compelling. In fact, the maximal growth perturbations have zero radial wavenumber, and execute extended radial displacements. In planes of constant z, linearly growing fields show a regular geometry. As the field grows in strength, progressively larger vertical wavelengths grow more rapidly. There is nothing in the axisymmetric linear theory that precludes the growth of ordered fields.

It is the possibility that something unusual may be associated with nonaxisymmetric disturbances that begs attention. Unfortunately, the technical issues involved in this sort of undertaking are frustratingly complex, and the literature is characterized by a lack of consensus among experts on the relative importance of fundamental processes. Genuine theoretical progress is likely to come only with painstaking numerical work, but a start can be made by carefully examining the linear problem. Toward these ends, we present here an analysis of nonaxisymmetric disturbances in a rotationally supported disk. Our basic finding is that large azimuthal wavenumbers show a much smaller growth rate than is shown by small azimuthal wavenumbers, so that at least a linear nonaxisymmetric calculation is consistent with large scale magnetic field coherence. The plan of this paper follows.

The analysis is wholly contained in § 2. We study the evolution of Eulerian perturbations in a thin disk in comoving, local Lagrangian coordinates (Goldreich & Lynden-Bell 1965), details of which are presented in § 2.1. After deriving two coupled equations for the evolution of the poloidal field components in § 2.2, we show that the physically interesting case of vanishing buoyant frequency allows for a reduction to a single fourth-order equation in § 2.3. When the ratio of the vertical to azimuthal wavenumber is large, the equation has WKB solutions. We present a qualitative discussion of the behavior of these solutions, before returning to more general numerical solutions in § 2.4. The case of a purely azimuthal field is studied, as is the more general case when a poloidal field is also present. Finally, § 3 summarizes our findings and conclusions.

2. LOCAL NONAXISYMMETRIC DISTURBANCES

2.1. *Preliminaries*

Our starting point is similar to the one in BH: we consider a weakly magnetized axisymmetric accretion disk of finite vertical extent. But in contrast to the presentation in BH, we assume at the outset that the disk is thin, and apart from the velocity shear, we ignore local radial structure in the unperturbed disk. We employ a standard (R, ϕ, z) cylindrical coordinate system, and assume that the angular velocity $\Omega(R)$ is constant on cylinders. All other flow variables may depend upon z. A weak magnetic field is present in the disk with an arbitrary local geometry. The field is weak in the sense that the Alfvén speed is asymptotically small compared with sound speed, and hence with the rotation velocity as well. We denote the azimuthal field component $B_\phi \hat{\boldsymbol{\phi}}$, and the poloidal components $B_z \hat{\boldsymbol{z}}$ and $B_R \hat{\boldsymbol{r}}$. (The notation $\hat{\boldsymbol{\phi}}$, etc. is used to denote a unit vector.) The basic dynamical equations are

$$\frac{d \ln \rho}{dt} + \nabla \cdot \boldsymbol{v} = 0 , \tag{2.1a}$$

$$\frac{d\boldsymbol{v}}{dt} + \frac{1}{\rho} \nabla\left(P + \frac{B^2}{8\pi}\right) - \frac{1}{4\pi\rho} (\boldsymbol{B} \cdot \nabla)\boldsymbol{B} + \nabla\Phi = 0 , \tag{2.1b}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} - \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) = 0 . \tag{2.1c}$$

The notation d/dt indicates the Lagrangian derivative, and Φ is the external gravitational potential. (This may include a massive halo.) Other symbols have their usual meanings.

We consider the local response of the disk to nonaxisymmetric perturbations. As is well known, such disturbances cannot have a simple plane waveform because of the effect of the shearing background on the wave crests (Goldreich & Lynden-Bell 1965). One adopts Lagrangian shearing coordinates

$$R' = R \, , \tag{2.2a}$$

$$\phi' = \phi - \Omega(R)t \, , \tag{2.2b}$$

$$z' = z \, , \tag{2.2c}$$

comoving with the unperturbed flow. Then

$$\frac{\partial}{\partial R} = \frac{\partial}{\partial R'} - t\,\frac{d\Omega}{dR}\,\frac{\partial}{\partial \phi'} \, , \tag{2.3a}$$

$$\frac{\partial}{\partial \phi'} = \frac{\partial}{\partial \phi} \, , \tag{2.3b}$$

$$\frac{\partial}{\partial z'} = \frac{\partial}{\partial z} \, . \tag{2.3c}$$

Let us further note that

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \Omega(R)\,\frac{\partial}{\partial \phi} \, . \tag{2.4}$$

In local Largrangian variables, linear perturbations are assumed to have a space dependence $e^{i(k_R' R' + m\phi' + k_z' z')}$. The effect of this variable transformation is simply to replace a fixed radial wavenumber with a shearing one:

$$k_R \leftarrow k_R(t) = k_R' - mt\,\frac{d\Omega}{dR} \, . \tag{2.5}$$

No modification of ϕ and z variables are needed. Provided we now work with the Lagrangian time derivative, equation (2.5) summarizes the transformation.

The presence of shear causes the toroidal magnetic field to grow linearly with time:

$$B_\phi(t) = B_\phi(0)\left(1 + \frac{B_R}{B_\phi(0)}\,\frac{d\Omega}{d\ln R}\,t\right) , \tag{2.6}$$

where $B_\phi(0)$ is the initial azimuthal field. But the combination of equations (2.5) and (2.6) shows that the Alfvén parameter $\boldsymbol{k} \cdot \boldsymbol{B}$ is constant with time, despite the fact that neither the Eulerian wavenumber $\boldsymbol{k} = [k_R(t), m/R, k_z]$ nor the magnetic field vector $\boldsymbol{B}$ is individually constant:

$$\boldsymbol{k} \cdot \boldsymbol{B} = k_R' B_R + \frac{mB_\phi(0)}{R} + k_z B_z \, . \tag{2.7}$$

Evidently, the wavevector and unperturbed magnetic field become increasingly more orthogonal in just the right proportion to cancel the effect of the growth of their magnitudes and maintain constant $\boldsymbol{k} \cdot \boldsymbol{B}$.

2.2. *Linear Disturbances*

Consider the fate of linear perturbations in a shearing disk with a weak magnetic field. The time-dependent Fourier amplitudes of Eulerian perturbations are denoted $\delta\rho$, δP, etc. Following BH we use the Boussinesq approximation, and as mentioned above, except for the presence of shear flow, we ignore local radial structure. Then the linearized dynamical equations expand to

$$k_R\,\delta v_R + \frac{m}{R}\,\delta v_\phi + k_z\,\delta v_z = 0 \, , \tag{2.8a}$$

$$\frac{d\delta v_R}{dt} - 2\Omega\delta v_\phi + ik_R\left(\frac{\delta P}{\rho} + \frac{\boldsymbol{B} \cdot \delta\boldsymbol{B}}{4\pi\rho}\right) - i\,\frac{\boldsymbol{k} \cdot \boldsymbol{B}}{4\pi\rho}\,\delta B_R = 0 \, , \tag{2.8b}$$

$$\frac{d\delta v_z}{dt} + ik_z\left(\frac{\delta P}{\rho} + \frac{\boldsymbol{B} \cdot \delta\boldsymbol{B}}{4\pi\rho}\right) - i\,\frac{\boldsymbol{k} \cdot \boldsymbol{B}}{4\pi\rho}\,\delta B_z - \frac{\delta\rho}{\rho^2}\,\frac{\partial P}{\partial z} = 0 \, , \tag{2.8c}$$

$$\frac{d\delta v_\phi}{dt} + \frac{\kappa^2}{2\Omega}\,\delta v_R + \frac{im}{R}\left(\frac{\delta P}{\rho} + \frac{\boldsymbol{B} \cdot \delta\boldsymbol{B}}{4\pi\rho}\right) - i\,\frac{\boldsymbol{k} \cdot \boldsymbol{B}}{4\pi\rho}\,\delta B_\phi = 0 \, , \tag{2.8d}$$

$$\frac{d\delta B_R}{dt} = i(\boldsymbol{k} \cdot \boldsymbol{B})\delta v_R , \tag{2.8e}$$

$$\frac{d\delta B_z}{dt} = i(\boldsymbol{k} \cdot \boldsymbol{B})\delta v_z , \tag{2.8f}$$

$$\frac{d\delta B_\phi}{dt} - \frac{d\Omega}{d\ln R}\,\delta B_R = i(\boldsymbol{k} \cdot \boldsymbol{B})\delta v_\phi . \tag{2.8g}$$

(We have suppressed the explicit time-dependent notation in k_R.) Since $dk_R/dt = -m d\Omega/dR$, equations (2.8e)–(2.8g) together with equation (2.8a) guarantee the divergence free condition $d(\boldsymbol{k} \cdot \delta\boldsymbol{B})/dt = 0$. The entropy equation for adiabatic perturbations is

$$-\frac{5}{3}\frac{d}{dt}\frac{\delta\rho}{\rho} + \delta v_z \frac{\partial \ln P\rho^{-5/3}}{\partial z} = 0 . \tag{2.8h}$$

Using equation (2.8f), this may be integrated to yield

$$\frac{\delta\rho}{\rho} = \frac{3\delta B_z}{5i(\boldsymbol{k} \cdot \boldsymbol{B})}\frac{\partial \ln P\rho^{-5/3}}{\partial z} . \tag{2.9}$$

Using this result in equation (2.8c) leads to

$$\frac{\boldsymbol{B} \cdot \delta\boldsymbol{B}}{4\pi\rho} + \frac{\delta P}{\rho} = \frac{1}{k_z \boldsymbol{k} \cdot \boldsymbol{B}}\frac{d^2\delta B_z}{dt^2} + \left[\frac{N^2}{k_z \boldsymbol{k} \cdot \boldsymbol{B}} + \frac{\boldsymbol{k} \cdot \boldsymbol{B}}{k_z 4\pi\rho}\right]\delta B_z , \tag{2.10}$$

where

$$N^2 = -\frac{3}{5}\frac{1}{\rho}\frac{\partial P}{\partial z}\frac{\partial \ln P\rho^{-5/3}}{\partial z} \tag{2.11}$$

in the Brunt-Väisälä frequency.

To proceed further, we note that equation (2.8a), (2.8e), and (2.8f) give

$$\delta v_\phi = -\frac{R}{mi(\boldsymbol{k} \cdot \boldsymbol{B})}\left(k_R \frac{d\delta B_R}{dt} + k_z \frac{d\delta B_z}{dt}\right) . \tag{2.12}$$

We now use equations (2.8e), (2.10), and (2.12) in equation (2.8b) to obtain

$$\frac{d^2\delta B_R}{dt^2} + \frac{2\Omega R}{m}\left(k_R \frac{d\delta B_R}{dt} + k_z \frac{d\delta B_z}{dt}\right) + (\boldsymbol{k} \cdot \boldsymbol{v}_A)^2\left(\delta B_R - \frac{k_R}{k_z}\,\delta B_z\right) - \frac{k_R}{k_z}\left(\frac{d^2\delta B_z}{dt^2} + N^2\delta B_z\right) = 0 , \tag{2.13}$$

where the Alfvén velocity is written

$$\boldsymbol{v}_A = \frac{\boldsymbol{B}}{(4\pi\rho)^{1/2}} . \tag{2.14}$$

We need another independent differential equation coupling δB_R and δB_z. Using $\boldsymbol{k} \cdot \delta\boldsymbol{B} = 0$, first obtain

$$\delta B_\phi = -\frac{R}{m}(k_R \delta B_R + k_z \delta B_z) . \tag{2.15}$$

Next, differentiating equation (2.12) gives

$$\frac{d\delta v_\phi}{dt} = -\frac{R}{mi(\boldsymbol{k} \cdot \boldsymbol{B})}\left(k_R \frac{d^2\delta B_R}{dt^2} + k_z \frac{d^2\delta B_z}{dt^2}\right) + \frac{1}{i(\boldsymbol{k} \cdot \boldsymbol{B})}\frac{d\Omega}{d\ln R}\frac{d\delta B_R}{dt} . \tag{2.16}$$

Multiply equation (2.8d) by k_R and equation (2.8b) by m/R, then subtract one from the other:

$$k_R \frac{d\delta v_\phi}{dt} + \frac{\kappa^2}{2\Omega}k_R \delta v_R - \frac{i(\boldsymbol{k} \cdot \boldsymbol{B})}{4\pi\rho}\left(k_R \delta B_\phi - \frac{m}{R}\,\delta B_R\right) - \frac{m}{R}\frac{d\delta v_R}{dt} + 2\Omega\frac{m}{R}\,\delta v_\phi = 0 . \tag{2.17}$$

Using equations (2.12), (2.15), and (2.16) in equation (2.17) leads after simplification to

$$\left[1 + \frac{m^2}{(k_R R)^2}\right]\frac{d^2\delta B_R}{dt^2} + \frac{k_z}{k_R}\frac{d^2\delta B_z}{dt^2} - \frac{2m}{k_R R}\frac{d\Omega}{d\ln R}\frac{d\delta B_R}{dt} + \frac{k_z}{k_R}\frac{2m\Omega}{k_R R}\frac{d\delta B_z}{dt} + (\boldsymbol{k} \cdot \boldsymbol{v}_A)^2\left[\delta B_R\left(1 + \frac{m^2}{R^2 k_R^2}\right) + \frac{k_z}{k_R}\,\delta B_z\right] = 0 . \tag{2.18}$$

Equations (2.13) and (2.18) are the two coupled differential equations in δB_R and δB_z that form the cornerstone of the analysis. It can be shown that in the limit $m \to 0$ the axisymmetric dispersion formula is obtained, but the calculation is rather involved and is deferred to the Appendix. As stated earlier, the result is not without interest because it shows explicitly that local axisymmetric instabilities are always described by simple exponential solutions, even when an azimuthal field changes with time.

For numerical work, it is convenient to isolate the second-order time derivatives. Equations (2.13) and (2.18) can be recombined to yield

$$\frac{d^2\delta B_z}{dt^2} = \frac{2\Omega R}{m}\frac{k_z k_R}{k^2}\left(k_R\frac{d\delta B_R}{dt} + k_z\frac{d\delta B_z}{dt}\right) + \frac{2mk_z}{Rk^2}\frac{d(R\Omega)}{dR}\frac{d\delta B_R}{dt} - \left[(\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2 + \frac{k^2-k_z^2}{k^2}\right]\delta B_z\,, \tag{2.19}$$

$$\frac{d^2 B_R}{dt^2} = -\frac{2\Omega R}{m}\frac{k_z^2}{k^2}\left(k_R\frac{d\delta B_R}{dt} + k_z\frac{d\delta B_z}{dt}\right) + \frac{2mk_R}{k^2}\frac{d\Omega}{dR}\frac{d\delta B_R}{dt} - \frac{2mk_z\Omega}{k^2 R}\frac{d\delta B_z}{dt} - \left[(\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\delta B_R + \frac{k_R k_z}{k^2}N^2\delta B_z\right], \tag{2.20}$$

where

$$k^2 \equiv k_R^2 + \frac{m^2}{R^2} + k_z^2\,, \tag{2.21}$$

and the time dependence of k_R is given in equation (2.5).

2.3. *Qualitative Behavior*

Equations (2.19) and (2.20) are complicated, and to understand their physical content it helps to combine them into a single fourth-order equation. Since the most unstable disturbances are associated with $N^2 = 0$, and the calculations become considerably more manageable in this limit, we restrict ourselves to this case in the section. Nothing fundamental to the instability is lost by working in this limit, which is in fact exact for the disk midplane $z = 0$. Even with this simplification, lengthy manipulations are involved, but the final equation is relatively simple. Only an outline is given.

With $N^2 = 0$, solve equation (2.20) for $d\delta B_z/dt$:

$$\frac{d\delta B_z}{dt} = -\mu\frac{k^2}{k_z^2}\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]\delta B_R - \nu\frac{k_R}{k_z}\frac{d\delta B_R}{dt}\,, \tag{2.22}$$

where we have introduced

$$\mu = \left[\frac{2\Omega k_z R}{m}\left(1 + \frac{m^2}{k_z^2 R^2}\right)\right]^{-1}, \tag{2.23a}$$

and

$$\nu = \left(1 - \frac{m^2}{k_z^2 R^2}\frac{d\ln\Omega}{d\ln R}\right)\Big/\left(1 + \frac{m^2}{k_z^2 R^2}\right). \tag{2.23b}$$

The idea now is to use equation (2.22) to obtain expressions for $d^2\delta B_z/dt^2$ and $d^3\delta B_z/dt^3$, and to insert these terms into the once-differentiated form of equation (2.19). It is a messy business, and it seems best only to write down the final result:

$$\frac{k^2}{k_z^2}\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]^2\delta B_R + \left[\kappa^2 + 6\frac{m^2}{k_z^2 R^2}\left(\frac{d\Omega}{d\ln R}\right)^2\right]\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]\delta B_R$$
$$- \left[1 + \frac{m^2}{k_z^2 R^2}\left(\frac{d\ln\Omega}{d\ln R}\right)^2\right]4\Omega^2(\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\delta B_R - 6\frac{mk_R}{k_z^2 R}\frac{d\Omega}{d\ln R}\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]\frac{d\delta B_R}{dt} = 0\,. \tag{2.24}$$

The case of greatest physical interest is "low" m, $m/k_z R \ll 1$. This still leaves a great deal of asymptotic space for m, since $k_z R \gg k_z H \gg 1$, where H is the disk scale height. The existence of the large parameter $k_z R/m$ immediately suggests a WKB approach. Physically, the WKB parameter represents the ratio of the time for $\boldsymbol{k}$ to change significantly to the shearing time $(d\Omega/d\ln R)^{-1}$. When $m/k_z R \ll 1$,

$$\frac{k^2}{k_z^2}\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]^2\delta B_R + \kappa^2\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]\delta B_R - 4\Omega^2(\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\delta B_R - 6\frac{mk_R}{k_z^2 R}\frac{d\Omega}{d\ln R}\left[\frac{d^2}{dt^2} + (\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2\right]\frac{d\delta B_R}{dt} = 0\,, \tag{2.25}$$

which is similar to the dispersion relation (2.17) in BH, and clearly goes over to it as m vanishes. One difference is the final term in equation (2.25). It arises from the interplay of nonaxisymmetry and shear, and its importance in a WKB treatment is as an amplitude modifier on longer time scales. It does not affect the essentially oscillatory or exponentially behavior of the solution. The other difference between equation (2.25) and a simple dispersion formula is of course that k_R^2 and k^2 are now time-dependent.

Equation (2.25) is a fourth-order equation; explicit WKB solutions are unwieldy and less than transparent. But the *existence* of WKB solutions means that the qualitative nature of the behavior is easily understood by analogy to the axisymmetric solution. The behavior of the ratio k/k_z is important, as shown in Figure 1. Initially large, it drops to a minimum value near unity when k_R passes through zero, and then rises again at later times. Since $\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A}$ is constant, the wavevector evolution traces (and retraces) a horizontal line in the plane of Figure 1. As $t \to -\infty$, the tightly wrapped wave is a superposition of inertial and Alfvénic modes. If the constant $(\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2/\Omega^2$ is above the critical value $(d\ln\Omega^2/d\ln R)$, there is no dynamical amplitude growth at any time. If $(\boldsymbol{k}\cdot\boldsymbol{v}_{\rm A})^2/\Omega^2$ is below

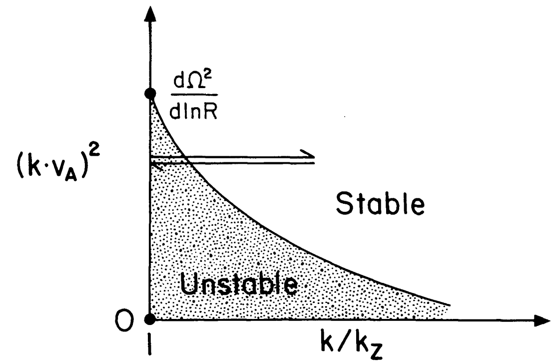

FIG. 1.—Region of stability and instability in $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2$–$k/k_z$ plane. As discussed in the text, the quantity $\boldsymbol{k} \cdot \boldsymbol{v}_A$ is constant for a shearing wavevector. Initially, k/k_z is large, and the point defining the wavevector moves to the left in the plane on a constant $\boldsymbol{k} \cdot \boldsymbol{v}_A$ line. The minimum value of k/k_z is unity (to order $m^2/k_z^2 R^2$.) After attaining its minimum, the wavevector point retraces its path to the right. If $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 < d\Omega^2/d\ln R$, then a finite portion of time is spent in the unstable region, and substantial growth may occur.

this value, then there is a period when $(k/k_z)^2$ is small enough that a finite portion of the evolution is spent in an exponential growth phase. Call the WKB growth rate σ. For values of $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2/\Omega^2$ below critical, but of order unity, both the duration of the exponential growth phase and the magnitude of σ grow larger with decreasing $(\boldsymbol{k} \cdot \boldsymbol{v}_A)$. The maximum amplification is achieved when $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2/\Omega^2$ is near $1 - \kappa^4/16\Omega^4$. Finally, when $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2/\Omega^2 \ll 1$, then $\sigma \simeq (3)^{1/2}(\boldsymbol{k} \cdot \boldsymbol{v}_A)$ (Balbus & Hawley 1992). When $\boldsymbol{k} \cdot \boldsymbol{v}_A/\Omega$ becomes comparable to the (small) WKB parameter $m/k_z R$, the time scale for the perturbation to evolve is no longer rapid compared with the long $\boldsymbol{k}$-changing time scale, WKB methods break down, and numerical integration is required for even a qualitative assessment. We shall find that exponential amplitude growth does not disappear, but takes place over time scales long compared with an orbital period. For all solutions, as $t \to \infty$, wavelike behavior is regained when the wavevector evolution emerges from the unstable region and moves rightward in Figure 1.

2.4. *Field Geometry and Stability*

We return to the general solution of equations (2.19) and (2.20). Let us parameterize the rotation curve by introducing an index p defined by

$$\Omega^2(R) \sim R^{-p} . \tag{2.26}$$

Thus, a Keplerian disk is characterized by $p = 3$, a galactic disk by $p = 2$. Next, following lines similar to those developed in Goldreich & Lynden-Bell (1965), we incorporate the radial wavenumber into a new independent time variable τ:

$$\tau \equiv k_R R = k'_R R - m \frac{d\Omega}{d\ln R} t , \tag{2.27}$$

so that $k^2R^2 = \tau^2 + m^2 + k_z^2 R^2$.

Using equations (2.26) and (2.27), equations (2.19) and (2.20) become

$$\frac{d^2\delta B_R}{d\tau^2} = -\frac{k_z^2}{k^2m^2}\left(\frac{4}{p} + \frac{2m^2}{R^2k_z^2}\right)\left(\tau \frac{d\delta B_R}{d\tau}\right) + \frac{4}{p^2m^2}\left(\frac{N^2}{\Omega^2}\frac{k_z \tau}{k^2R}\delta B_z - \frac{(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2}{\Omega^2}\delta B_R\right) - \frac{4}{p}\frac{k_z}{k^2R}\left(1 + \frac{R^2k_z^2}{m^2}\right)\frac{d\delta B_z}{d\tau} , \tag{2.28}$$

$$\frac{d^2\delta B_z}{d\tau^2} = \frac{2}{p}\frac{k_z}{k^2R}\left(\frac{2\tau^2}{m^2} - p + 2\right)\frac{d\delta B_R}{d\tau} + \frac{4}{pm^2}\frac{k_z^2}{k^2}\left(\tau\frac{d\delta B_z}{d\tau}\right) - \frac{4}{p^2m^2}\left(\frac{(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2}{\Omega^2} + \frac{k^2 - k_z^2}{k^2}\frac{N^2}{\Omega^2}\right)\delta B_z . \tag{2.29}$$

Equations (2.28a) and (2.28b) are the form of the evolutionary equations that are used in the numerical integrations. In this work, we shall concentrate on the Keplerian $p = 3$ case, retaining the above for the sake of completeness. Nothing essential is lost by this restriction.

One of the more interesting cases to examine for nonaxisymmetric instability is the purely azimuthal field configuration, since in the absence of a poloidal component, the Alfvén wave branch is inaccessible by an axisymmetric analysis. Of course, the field geometry enters equations (2.28) and (2.29) only through the combination $\boldsymbol{k} \cdot \boldsymbol{v}_A = mB_\phi/[(4\pi\rho)^{1/2}R]$. Large m values are required to generate a perturbed magnetic tension $\boldsymbol{k} \cdot \boldsymbol{v}_A \sim \Omega$. But going to large azimuthal wavenumbers has its price. The problem is equation (2.27). As m becomes large, there is less time before the radial wavenumber of the disturbance becomes large, which leads to tight wrapping of the wave crest and stabilization of the disturbance. The question then is whether significant growth can be achieved when m is not large and $\boldsymbol{k} \cdot \boldsymbol{v}_A \ll 1$. The answer appears to be yes, although the growth rates are relatively small compared to the Ω time scale we are used to seeing when a poloidal component is present—typically a few percent of the orbital frequency, corresponding to growth times of tens of orbital periods. The most rapidly growing disturbances seem to arise for values of m of order $(k_z R)^{1/2}$.

For an azimuthal field, we may write

$$\frac{\boldsymbol{k} \cdot \boldsymbol{v}_A}{\Omega} = \left(\frac{m}{k_z R}\right)\left(\frac{k_z v_{A\phi}}{\Omega}\right) . \tag{2.30}$$

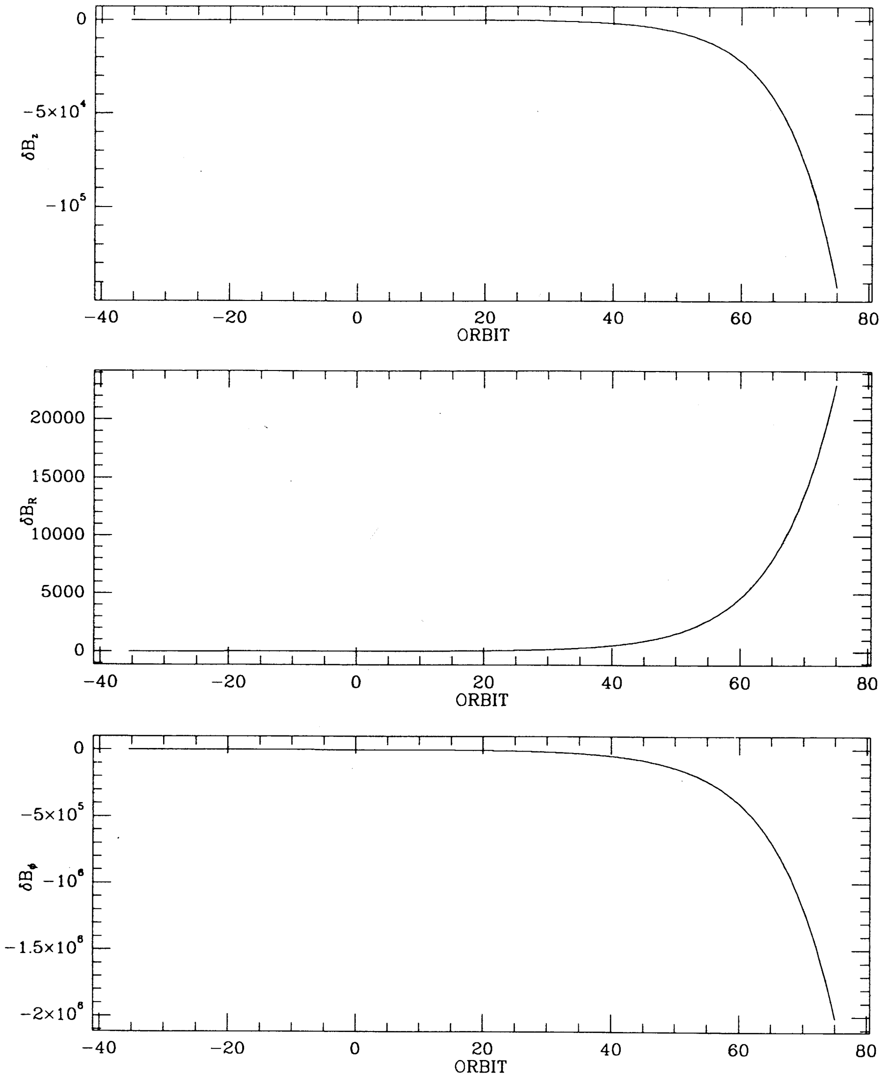

FIG. 2.—Evolution of δB_z, δB_R, and δB_ϕ. The initial amplitudes are $\delta B_z = 1$, $\delta B_R = 0.3$, and $\delta B_\phi = 0.3$. We select $k_z R = 100$, $m = 1$, and allow the initial value of $\tau = k_R R$ to be determined by $\nabla \cdot \delta \boldsymbol{B} = 0$. N^2 has been set to zero. $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 10^{-4}\Omega^{-2}$, appropriate to an initial toroidal field configuration with the selected wavenumbers.

It seems reasonable to adopt a value near unity for the second factor on the right, since the z wavenumbers would then be of the same order as those of interest when a poloidal field is present of comparable strength to the assumed azimuthal field. Thus, we find that the small Alfvén parameter on the left-hand side of equation (2.30) is of order the WKB parameter $m/k_z R$, which is the most natural scaling assumption. Note that it also means WKB methods are marginal.

In Figures 2, 3, and 4, we plot the evolution of δB_z, δB_R, and δB_ϕ for three different m values: $m = 1, 10, 100$. The initial conditions for the displayed run are $\delta B_z = 1$, $\delta B_R = 0.3$, and the initial τ value is chosen to correspond to $\delta B_\phi = 0.3$; the initial time derivatives are zero. We set $k_z R = 100$ throughout, and for the displayed runs, we have chosen the midplane value $N^2 = 0$. (Positive N^2 runs are slightly more stable.) The $m = 1$ and $m = 10$ runs both show growth, with $m = 10$ displaying by far the more vigorous behavior. But increasing m leads to diminishing returns, as shown in the $m = 100$ run. This is quite stable, even though it corresponds to what would be a highly unstable value of $\boldsymbol{k} \cdot \boldsymbol{v}_A$ for an axisymmetric disturbance. The difficulty is that k_R grows too rapidly for large m values, and a wavelike response becomes the lone possibility.

In common with the axisymmetric disturbances, the greatest response for nonaxisymmetric perturbations occurs when a poloidal field is present. Then, we can always adjust $\boldsymbol{k} \cdot \boldsymbol{v}_A$ to be near Ω for maximum growth, without incurring the large m penalty of shortening the wrapping time of the wavevector. We show here the response of $m = 1$ disturbances for a few different values of $\boldsymbol{k} \cdot \boldsymbol{v}_A$.

The Keplerian critical value of $\boldsymbol{k} \cdot \boldsymbol{v}_A$ is $3^{1/2}\Omega$, and to illustrate the onset of instability, we plot results at, and just below this value. Further decreasing $\boldsymbol{k} \cdot \boldsymbol{v}_A$ leads to behavior indistinguishable from an exponential instability. In general, there are few surprises. Growth occurs when k/k_z is near unity, corresponding to τ and the orbit number equal to 0. The WKB amplification factor during the growth phase is just exp $(\int |\omega| dt)$, integrated over the time that k/k_z is in the unstable region shown in Figure 1. This rapidly becomes an enormous factor, leading, as already emphasized, to unstable growth.

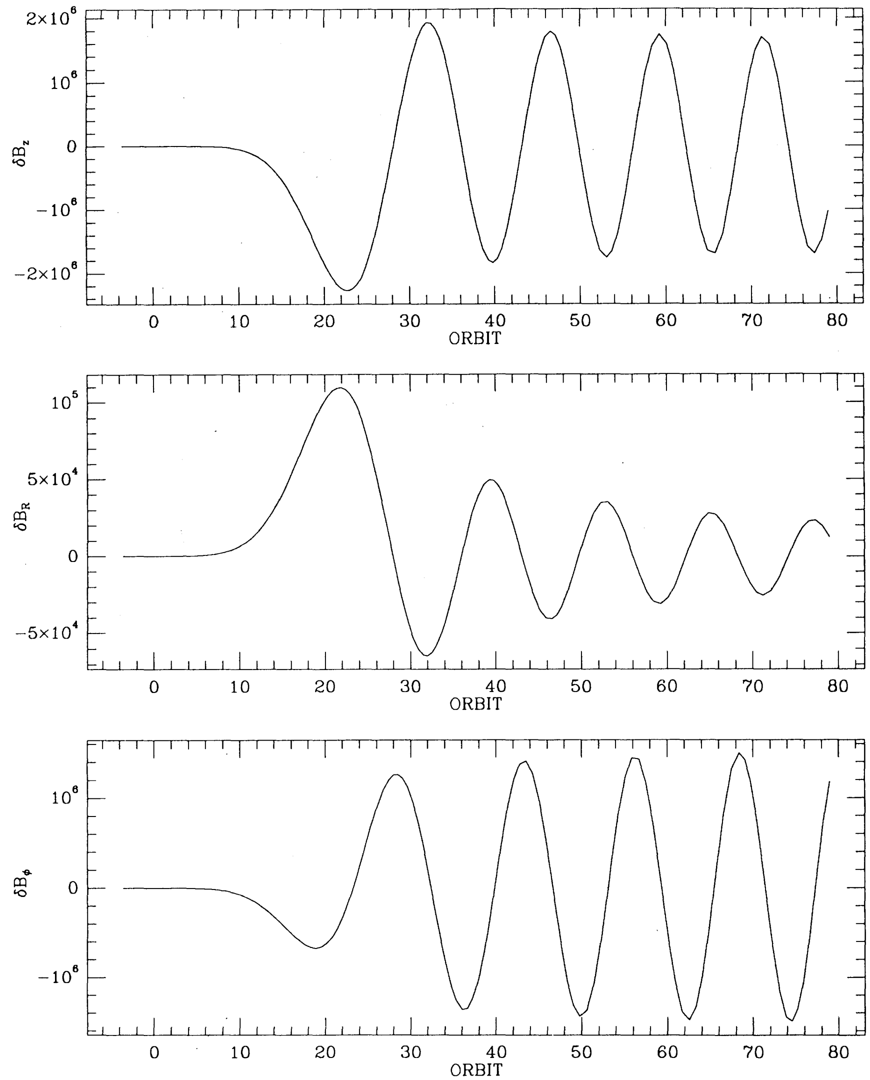

FIG. 3.—Same as in Fig. 2, with $m = 10, (\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 10^{-2}\Omega^{-2}$

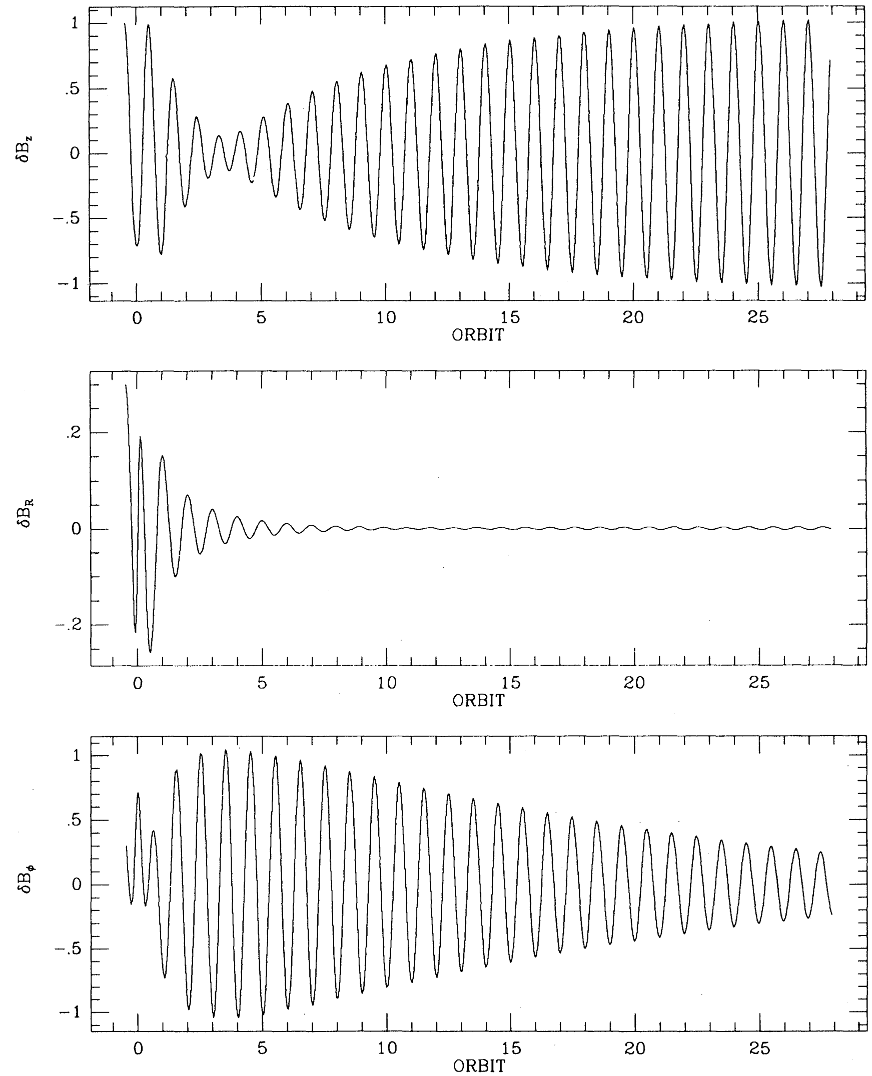

FIG. 4.—Same as in Fig. 2, with $m = 100, (\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = \Omega^2$

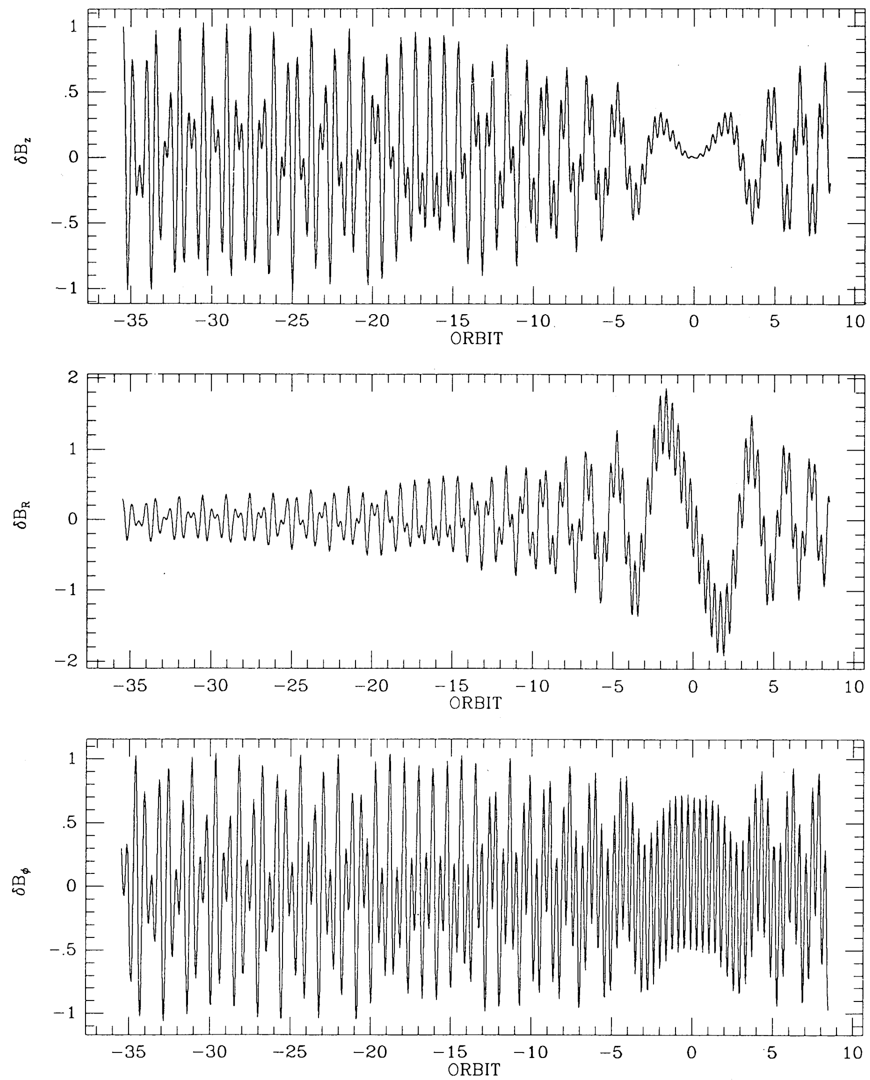

FIG. 5.—Same as in Fig. 2, with $m = 1, (\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 3\Omega^2$. These parameters are natural for a configuration with a weak poloidal field.

In Figure 5, δB_z, δB_R, and δB_ϕ are shown for $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 3\Omega^2$. There is no growth, and the generally choppy nature of the solution corresponds to a superposition of two different WKB waves with a nonsteady phase relation. Figure 6 shows the case $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 2.9\Omega^2$, just into the unstable regime. The same initial conditions give rise to a much different result. The same choppy superposition is evident a times $t < 0$, but the amplification "filter" picks out only one of the WKB waves for growth. A smooth oscillatory behavior is evident in the stable $t > 0$ portion of the run. Further reduction of $\boldsymbol{k} \cdot \boldsymbol{v}_A$ yields explosive growth, as shown for example in Figure 7. For the sake of completeness, we show in Figure 8 a run with nonvanishing $N^2 = 0.8\Omega^2$, and rotation curve parameter $p = 2$ (flat rotation curve.) The behavior is qualitatively similar to the earlier runs.

A final point worth touching upon is the relative growth between the three components of the magnetic field. Equation (2.15) is

$$\delta B_\phi = -\frac{1}{m}(\tau \delta B_R + k_z R \delta B_z) . \tag{2.31}$$

Since τ increases monotonically, either δB_ϕ will eventually dominate δB_R and δB_z, or $\delta B_R \simeq -(k_z R/\tau)\delta B_z$ and decline with time. Starting with initial conditions in which the three components are comparable, the latter outcome is evident in the numerical solutions. It is unlikely however that this behavior will be retained in the nonlinear resolution of the instability, because the explosive growth phase occurs for relatively small τ-values. For this reason it seems more likely that the dominant poloidal component of the field will be radial (as in the axisymmetric limit), but with a comparable or larger toroidal component growing with the ever present differential shear.

3. SUMMARY

In this work, we have investigated the stability of weakly magnetized thin accretion disks to nonaxisymmetric instabilities. We refer to the disturbances as "instabilities," even though they are technically transient amplifications, because the amplification

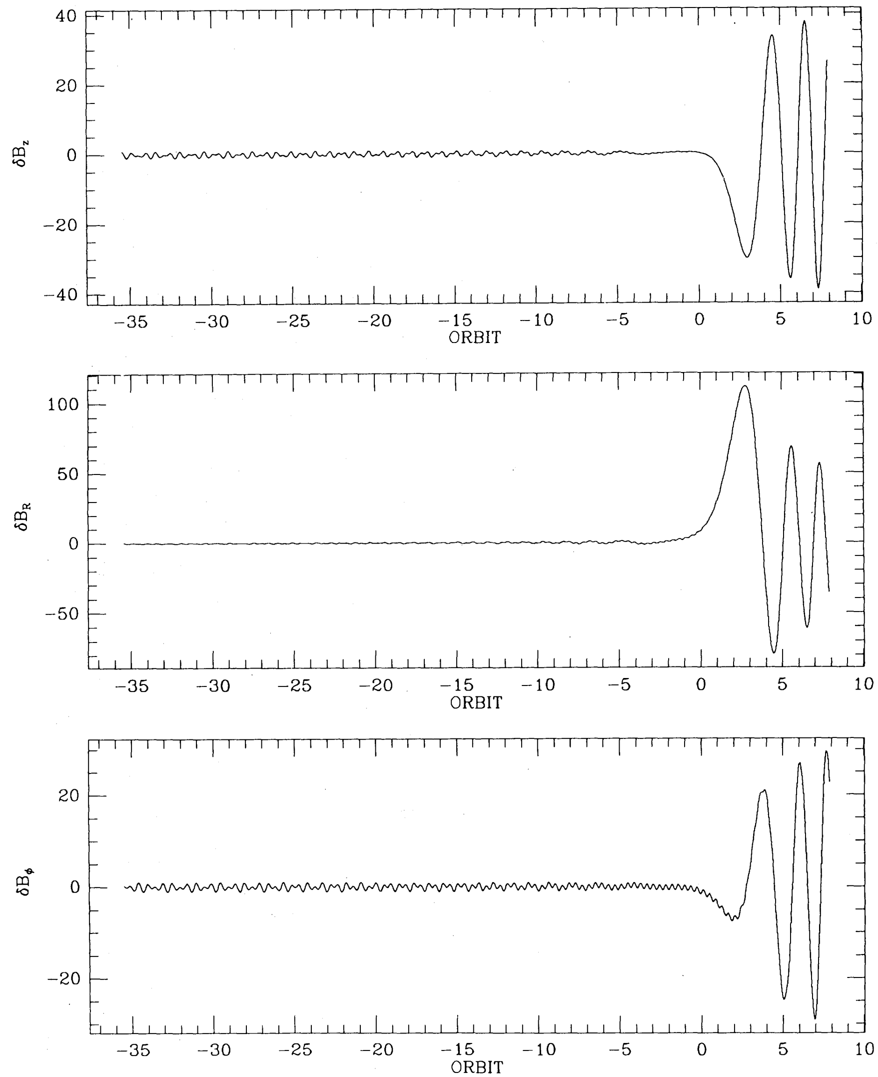

FIG. 6

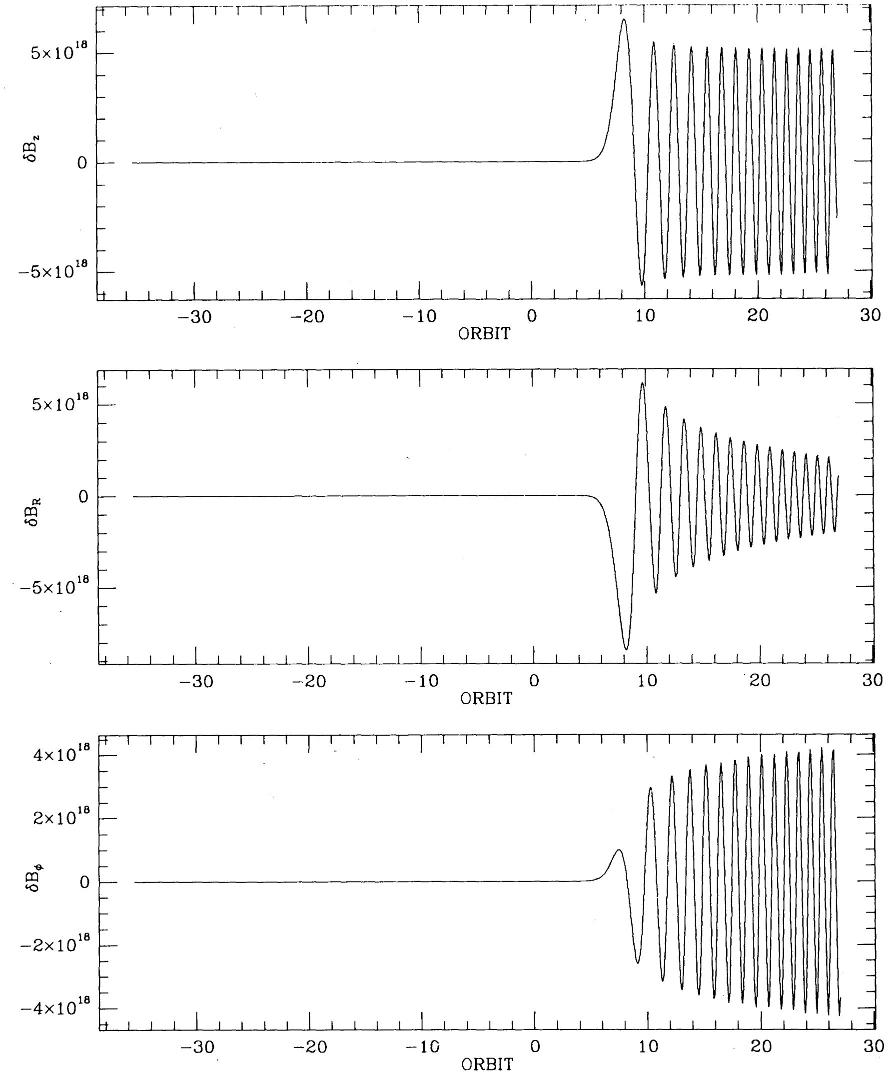

FIG. 7

FIG. 6.—Same as in Fig. 5, with $m = 1$, $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 2.9\Omega^2$. Note the selection of one WKB wave for amplification from the two initially present.
FIG. 7.—Same as in Fig. 5, with $m = 1$, $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2 = 2\Omega^2$. The growth phase is explosive and physically indistinguishable from an exponential instability. Note the ordinate scale.

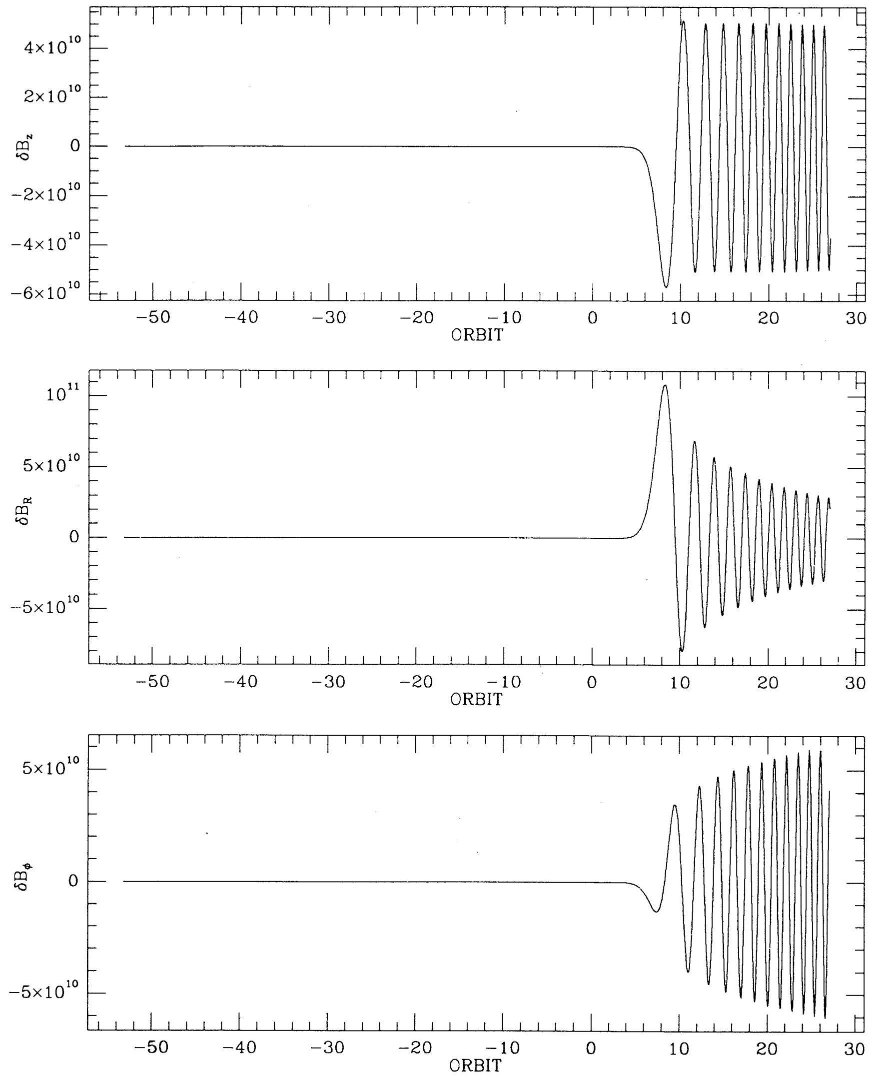

FIG. 8.—Evolution of the field components in a $p = 2$ (flat rotation curve) disk, with initial conditions as in Fig. 2. The critical value of $(\boldsymbol{k} \cdot \boldsymbol{v}_A)^2$ is $2\Omega^2$, and we have taken here a value of $1.5\Omega^2$. Also, we have $N^2 = 0.8\Omega^2$. The runs show the same recurring pattern seen in the previous figures.

factors can be tens of orders of magnitude. The linear time dependence of the unperturbed azimuthal component of the magnetic field in no way affects the essential oscillatory/exponential behavior of either axisymmetric or nonaxisymmetric local disturbances. To the extent that nonaxisymmetric instabilities may be a prerequisite to true magnetic field amplification, our results are very encouraging. The presence of a poloidal field and small m perturbations is sufficient to ensure vigorous exponential growth on time scales a fraction of a rotational period, and even a purely azimuthal field offers no respite from an instability whose growth rate is measured in tens of orbital periods. There is no hint from linear theory that the correlated, large-scale magnetic features seen in the two-dimensional numerical simulations break down when nonaxisymmetric structure is allowed.

We believe that the nonlinear development of weak field shearing instabilities in accretion disks is the most promising generic explanation for the presence of anomalous viscosity, and will also prove to be a powerful magnetic field amplifier. The time is passed that we may look to gas dynamics alone to understand nonspherical accretion processes. Transport mechanisms premised on the fiat that accretion disk magnetic fields are perpetually weak, or that the fields exhibit no large-scale coherence or are otherwise inconsequential, receive support neither from linear MHD analyses nor from numerical simulations. The linear theories presented in this paper and in BH suggest that large-scale magnetic fields in the disk plane are seeded by low k_R, low m, rapidly growing perturbations. The calculations challenge fundamental assumptions of kinematic dynamo theory when applied to disk systems. The $m = 0$ disturbances studied in detail in Paper III vividly show the rapid large-scale growth of initially weak azimuthal and radial field components, as do the simulations of Stone & Norman (1992). We see nothing in the linear nonaxisymmetric work presented here to deter the notion that three-dimensional simulations will show the same correlated field growth and hefty angular momentum transport. But this prediction need not be taken on faith. Such confidence will very soon be tested. Little more can be done with linear theory in weakly magnetized, single-fluid accretion disks. Two- and three-dimensional MHD simulations of stratified disks are the essential next steps.

We acknowledge useful discussions with Omer Blaes, Roger Blandford, Peter Goldreich, Jeremy Goodman, Sterl Phinney, and Scott Tremaine. We particularly thank Sterl Phinney and the TAPIR group at Caltech for their hospitality and support during our visit when a portion of this work was completed. This work is partially supported by NSF grants PHY-9018251 and AST-9021348, and NASA grants NAGW-1510 and NAGW-2376.

APPENDIX

It is not altogether straightforward to obtain the axisymmetric dispersion relation as $m \to 0$ from equations (2.19) and (2.20). A straightforward limiting procedure gives to leading order $\nabla \cdot \delta \boldsymbol{B} = 0$, not the axisymmetric dispersion formula. The dispersion relation emerges only in the next higher order in a small m expansion. Thus we need to expand the δ-amplitudes as a power series in m:

$$\delta B_R = \delta B_R^0 + m\delta B_R^1 + \cdots \qquad \delta B_z = \delta B_z^0 + m\delta B_z^1 + \cdots \tag{A1a}$$

Only terms through order m are needed. We must also make use of the (exact) expression for k_R:

$$k_R = k_R^0 - m\frac{d\Omega}{dR}t . \tag{A1b}$$

Henceforth, in this section only, we suppress the 0 superscript in k_R^0. The zeroth-order amplitudes have time dependence $e^{-i\omega t}$, and satisfy

$$\frac{d\delta B_j^0}{dt} = -i\omega\delta B_j^0 , \tag{A2a}$$

where j may stand for either R or z, and

$$k_R\delta B_R^0 + k_z\delta B_z^0 = 0 . \tag{A2b}$$

To recover the axisymmetric dispersion relation, it is slightly more convenient to start with the equations in the forms of equations (2.13) and (2.18). If we substitute the expansions (A1a) in equation (2.13), the leading order terms give (as stated above) $\nabla \cdot \delta \boldsymbol{B} = 0$, while the next order terms require

$$k_R\frac{d\delta B_R^1}{dt} + k_z\frac{d\delta B_z^1}{dt} = -i\omega t\frac{d\Omega}{dR}\delta B_z^0 - \frac{k_R}{k_z}\left(\frac{1}{2\Omega R}\right)\left(\frac{k^2}{k_R^2}\tilde{\omega}^2 - N^2\right)\delta B_z , \tag{A3a}$$

where

$$\tilde{\omega}^2 \equiv \omega^2 - (\boldsymbol{k}\cdot\boldsymbol{v}_A)^2 . \tag{A3b}$$

We shall also need the twice-differentiated form of equation (A3a):

$$k_R\frac{d^3\delta B_R^1}{dt^3} + k_z\frac{d^3\delta B_z^1}{dt^3} = \left(i\omega^3 t\frac{d\Omega}{dR} - 2\omega^2\frac{d\Omega}{dR}\right)\delta B_R^0 + \frac{k_R}{k_z}\frac{\omega^2}{2\Omega R}\left(\frac{k^2}{k_R^2}\tilde{\omega}^2 - N^2\right)\delta B_z^0 . \tag{A4}$$

The last ingredient we shall need is a relation for the order m terms of the once-differentiated form of equation (2.18):

$$\left(\frac{d^3\delta B_R^1}{dt^3} + \frac{k_z}{k_R}\frac{d^3\delta B_z^1}{dt^3}\right) - \frac{k_z}{k_R^2}\omega^2\frac{d\Omega}{dR}(1 - i\omega t)\delta B_z^0 + \frac{2\Omega\omega^2}{k_R}\left(\frac{d\ln\Omega}{dR}\delta B_R^0 - \frac{k_z}{k_R R}\delta B_z^0\right)$$
$$+ (\boldsymbol{k}\cdot\boldsymbol{v}_A)^2\left[\frac{d\delta B_R^1}{dt} + \frac{k_z}{k_R}\frac{d\delta B_z^1}{dt} + \frac{k_z}{k_R^2}\frac{d\Omega}{dR}(1 - i\omega t)\delta B_z^0\right] = 0 . \tag{A5}$$

By substituting equations (A3a) and (A4) into equation (A5) and using (A2b), after considerable simplification the axisymmetric dispersion equation emerges:

$$\tilde{\omega}^4 - \tilde{\omega}^2\left(\frac{k_z^2}{k^2}\kappa^2 + \frac{k_R^2}{k^2}N^2\right) - 4\Omega^2\frac{k_z^2}{k^2}(\boldsymbol{k}\cdot\boldsymbol{v}_A)^2 = 0 , \tag{A6}$$

which is our desired result.

REFERENCES

Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214 (BH)
———. 1992, ApJ, 392, 662
Cowling, T. G. 1981, ARA&A, 19, 115
Goldreich, P., & Lynden-Bell, D. 1965, MNRAS, 130, 125
Goodman, J. 1991, private communication
Hawley, J. F., & Balbus, S. A. 1991, ApJ, 376, 223 (HB)
———. 1992, ApJ, 400, in press (Paper III)
Mathewson, D. S., van der Kruit, P. C., & Brouw, W. N. 1972, A&A, 17, 468
Moffatt, H. K. 1978, Magnetic Field Generation in Electrically Conducting Fluids (Cambridge: Cambridge Univ. Press)
Parker, E. N. 1979, Cosmic Magnetic Fields (Oxford: Oxford Univ. Press)
Stone, J., & Norman, M. 1992, preprint
Zel'dovich, Ya. B., Ruzmaikin, A. A., & Sokoloff, D. D. 1983, Magnetic Fields in Astrophysics (NY: Gordon & Breach)

Note added in proof.—The stability of strong, vertical magnetic fields in disks to nonaxisymmetric disturbances has recently been considered by Tagger, Pellat, & Coroniti (1992) (Tagger, M., Pellat, R., & Coroniti, F. V., ApJ, 393, 708 [1992]).