1. INTRODUCTION

Cai & Shu (2002) put forward the hypothesis that relativistic disks may ultimately play as important a role in astrophysics as their spherical counterparts. There is evidence that supermassive black holes in quasars were formed during the early stages of galaxy formation, when the major contributor to the mass-energy density was gas rather than stars. The dynamics of gas is highly dissipative, so it is easier for gas to lose significant amounts of energy to reach the desired compactness for general relativistic effects to be important. More problematic is how to get rid of excess angular momentum if it is present initially.

The pioneering work of Bardeen & Wagoner (1971) on uniformly rotating disks showed that with some assumed angular momentum loss, a Kerr black hole may result in the collapse of such disks. However, as Mestel (1963) pointed out in a Newtonian context, there are astrophysical reasons to think that a disk specified by constant linear rotation velocity \( v \) is more realistic than one that has constant angular velocity. Through bar formation and spiral density waves (for a review, see, e.g., Shu et al. 2000), such differentially rotating disks possess natural mechanisms for the outward transport of angular momentum and the inward transport of mass that would promote, in the relativistic regime, blackhole formation at the center.

The first step in a systematic theoretical study of this possibility is to construct fully relativistic, self-similar, rotating, flattened solutions. Lynden-Bell & Pineault (1978) performed such a pioneering analysis, but only in the cold limit. To suppress well-known axisymmetric instabilities that would fragment the disk into rings, it is necessary to include partial support from isothermal pressure. If the pressure is exerted isotropically in three directions, the result (Cai & Shu 2003) is the relativistic generalization of the singular isothermal toroids (SITs) found by Toomre (1982) and Hayashi et al. (1982). A useful approximation when such states become sufficiently flattened by rotation is to ignore the thermal dispersive speed in the vertical direction, while retaining it in the horizontal directions. In this approximation SITs become completely flattened, singular isothermal disks (SIDs), whose equilibrium properties in the relativistic regime were studied by Cai & Shu (2002). These solutions are infinite in extent, possess infinite total mass, and contain a naked singularity at the origin (a “baby black hole” with vanishingly small mass).

The formal approximation of highly flattened configurations is satisfied for SITs only when the Mach number \( M \equiv v / \sqrt{\gamma} \gg 1 \). Nevertheless, even if \( M \sim 1 \), Cai & Shu (2003) found that the critical condition under which the sequence of equilibria terminates as a function of \( M \) is nearly identical, whether the pressure is exerted in three dimensions (SITs) or two (SIDs). The insensitivity of crucial properties of the equilibria to the assumption of infinitesimal thickness will hopefully carry over to the analysis of their stability.

On dimensional grounds, if the disk becomes gravitationally unstable to overall gravitational collapse (a possibility if the disk is rotating sufficiently slowly), the mass of the baby black hole will grow linearly in time as a result of axisymmetric collapse. In the analogous problem of the collapse of relativistic singular isothermal sphere (SIS), M. J. Cai & F. H. Shu (2004, in preparation) have shown that the growth of a black hole with finite mass introduces a (spherically symmetric) horizon that covers up the singularity. It is intriguing to ask whether such a singularity in the case of a collapsing, relativistic SID will remain naked if the requirement of axial symmetry is relaxed. In order to answer this question, one must first construct relativistic SIDs that are in nonaxisymmetric states of equilibria. One of the goals of the present paper is to make a start on this problem, by finding the points of nonaxisymmetric bifurcation along the sequence of axisymmetric SIDs, thereby generalizing the Newtonian work of Syer & Tremaine (1996) and Galli et al. (2001).

A feature with no Newtonian analog appears with the completion of the analysis: the appearance of a secular instability that afflicts all relativistically rotating disks. This instability, associated classically with the name Rossby (so the corresponding perturbations are called R-modes), arises because general relativity admits the radiation of angular momentum (and energy) by gravitational waves (Chandrasekhar 1970a, 1970b). Basically, if the rotation of the underlying axisymmetric state is rapid enough, a counterrotating disturbance appears corotating to an inertial observer. Such disturbances have negative angular momentum density in the local rest frame of the disk. Since gravitational radiation carries away positive angular momentum, the amplitude of the R-mode perturbation grows in time.
In some sense, the phenomenon renders all astrophysically rotating systems potentially unstable to spin-down on a gravitational radiation timescale. Whether the R-mode instability competes with the gravitational torques associated with barlike deformations or spiral density waves remains a problem for future study. The self-similar models and techniques used in the present paper are capable of determining only the criterion for the onset of secular instabilities and not their growth rates and evolution into the nonlinear regime. Thus, modifications are still required for application to astrophysically realistic circumstances, where the origin does not contain a singularity from the start and spacetime at infinity is flat.

In this paper, we restrict our study of the stability of relativistic SIDs to nonaxisymmetric perturbations with the same scale-free character as the equilibrium state (i.e., with the same power-law radial dependence). In the nomenclature of Syer & Tremaine (1996), we consider only aligned perturbations and no spiral disturbances. In § 2, we review the basic properties of axisymmetric SIDs. In § 3, we develop the mathematical formulation of the stability analysis, including the metric and matter perturbations. In § 4, the equations are solved in the Newtonian limit, and the result is compared to that of Shu et al. (2000). In § 5, the perturbation equations are solved in the full relativistic context, and we offer physical interpretation of the results.

2. REVIEW OF AXISYMMETRIC DISK SOLUTION

Start out with a self-similar axisymmetric metric,

\[ ds^2 = -r^{2n}e^N dt^2 + r^2 e^{2P-N} (d\phi - r^{n-1} e^{N-P} Q dt)^2 + e^{2Z-N} (dr^2 + r^2 d\theta^2), \]

where \( N, P, Q, \) and \( Z \) are functions of \( \theta \) and \( n \) is a constant measuring the strength of the gravitational field. For numerical convenience, we have chosen the equatorial plane to be at some polar angle \( \theta_0 \), which is determined as an eigenvalue of the problem. The locally nonrotating observer (LNRO) defines an orthonormal tetrad frame analogous to the inertial frame:

\[ e_0^n = \left( r^{-n} e^{-N/2}, r^{-1} Q e^{N/2-P}, 0, 0 \right), \]
\[ e_1^n = \left( 0, r^{-1} e^{N/2-P}, 0, 0 \right), \]
\[ e_2^n = \left( 0, 0, e^{(N-Z)/2}, 0 \right), \]
\[ e_3^n = \left( 0, 0, r^{-1} e^{(N-Z)/2} \right). \]

We look for solutions to the Einstein field equations with a disk matter source described by a constant linear rotation velocity and a two-dimensional isotropic pressure. In the frame of an LNRO, the stress-energy tensor is taken to be

\[ T_{(0)(0)} = \frac{\varepsilon + p_\phi v^2}{1 - v^2}, \]
\[ T_{(0)(1)} = -\frac{\varepsilon + p_\phi v}{1 - v^2}, \]
\[ T_{(1)(1)} = \frac{p_\phi + \varepsilon v}{1 - v^2}, \]
\[ T_{(2)(2)} = p_r, \]

where \( \varepsilon \propto \delta(\theta - \theta_0) \) and \( p_\phi = p_r = \gamma \varepsilon \). We define

\[ \Theta = (1 + n)\theta, \quad \xi = 8\pi \frac{\varepsilon}{1 + n} r^2 e^{-N_0}, \quad \Delta = \delta(\Theta - \Theta_0). \]

After some algebra, a part of the Einstein equations are cast into a set of dynamic equations:

\[ N'' = -N'P' - \frac{2n}{1+n} + Q^2 F^2 + Q^2 \left( \frac{1-n}{1+n} \right)^2 + \xi \frac{1 + 2\gamma + v^2}{1 - v^2} \Delta, \]
\[ P'' = -P^2 - 1 + \xi \gamma \Delta, \]
\[ Q'' = -Q' P' - \xi (1 + \gamma) \frac{2n + Q + Q v^2}{1 - v^2} \Delta + Q \left( 1 - Q^2 \right) \left( \frac{1-n}{1+n} \right)^2 + (N' - P')^2 - Q^2 F^2, \]
\[ Z'' = -Z' P' + 2Q^2 \left( \frac{1-n}{1+n} \right)^2 - 4\frac{n^2}{(1+n)^2} + Q^2 F^2 + 2\xi \frac{v^2 + \gamma}{1 - v^2} \Delta, \]

where \( F = N' + \log Q' - P' \) and a prime denotes differentiation with respect to \( \Theta \). The rest of the field equations and the equation of motion (EOM) form a set of constraint equations:

\[ Q(\Theta_0) v(1 - n)(1 + \gamma) + \gamma + v^2 - n(1 + \gamma) = 0, \]
\[ Z' - \frac{2n}{1+n} N' - Q^2 \frac{1-n}{1+n} F = 0, \]
\[ Q^2 \left[ F^2 - \left( \frac{1-n}{1+n} \right)^2 \right] + 2 \cot \Theta Z' - N'^2 + \frac{4n^2}{(1+n)} = 0. \]
3. PERTURBED CONFIGURATION

We use the Eulerian description for the nonaxisymmetric modes. The perturbation in the metric is
\[ \delta g_{\mu\nu} = h_{\mu\nu}, \quad \left| \frac{h_{\mu\nu}}{\eta_{\alpha\beta}} \right| \ll 1. \]

We define the change in the contravariant components of the metric as
\[ \delta g^{\mu\nu} = -h^{\mu\nu} = -h_{\alpha\beta}g^{\mu\nu}g^{\beta\alpha}, \]
so that\(^1\)
\[ (g^{\mu\nu} + \delta g^{\mu\nu})(g_{\nu\rho} + \delta g_{\nu\rho}) = \delta^{\mu}_{\rho} + O(h^2). \]

Note that \(h_{\mu\nu}\) is not a tensor with respect to the unperturbed metric, and its projection onto an LNRO frame is not a Lorentz scalar. Thus, the directional derivatives of \(h_{(a)(b)}\) involve more than the usual Ricci rotation coefficients, which destroys the simplicity of the tetrad formalism. As a result, we compute in the coordinate frame whenever derivatives are involved. However, the tetrad frame defined by equation (2.2) does offer a clean separation of \(r\) from the other coordinates, so we project our results onto an LNRO frame after the derivatives have been taken.

The change in the Ricci tensor reads (see, e.g., Wald 1984)
\[ \delta R_{\mu\nu} = \frac{1}{2} \left( \delta T^{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) = \frac{1}{2} \left( h^{\mu\nu} + h^\alpha_{\mu\alpha} - h^\alpha_{\nu\alpha} - h_{\mu\nu} \right), \quad (3.1) \]
where the raising and lowering of indices are done with the unperturbed metric. Instead of computing the Einstein tensor, we work out the trace-reversed stress-energy tensor. Taking the direct variation of the stress-energy tensor, we have
\[ \delta \left( T_{\mu\nu} \right) - \frac{1}{2} g_{\mu\nu} T = \frac{1}{2} h_{\mu\nu} T - \frac{1}{2} g_{\mu\nu} \delta T^{\alpha\beta} g_{\alpha\beta} - \frac{1}{2} g_{\mu\nu} T^{\alpha\beta} h_{\alpha\beta}. \]

There is one subtlety in writing down this expression, because of the nontensorial nature of the metric variations \(h_{\mu\nu}\). Explicitly,
\[ \delta T_{\mu\nu} = \delta T^{\nu\beta} g_{\alpha\beta} g_{3\nu} + T^{\alpha}_{\mu} h_{\alpha\nu} + T^{\alpha}_{\nu} h_{\alpha\mu} - \delta T^{\nu\beta} g_{\alpha\beta} g_{3\nu}. \]

To avoid such confusion, we adopt the convention that only contravariant components \(T^{\alpha\beta}\) are varied.\(^2\) So the Einstein field equation reads
\[ h_{\mu\nu} + h^\alpha_{\nu\alpha} - h^\alpha_{\mu\alpha} - h_{\mu\nu} = 8\pi \left( 2\delta T^{\nu\beta} g_{\alpha\beta} g_{3\nu} + 2T^{\alpha}_{\mu} h_{\alpha\nu} + 2T^{\alpha}_{\nu} h_{\alpha\mu} - h_{\mu\nu} T - g_{\mu\nu} \delta T^{\alpha\beta} g_{\alpha\beta} - g_{\mu\nu} T^{\alpha\beta} h_{\alpha\beta} \right). \quad (3.2) \]

3.1. Gauge Choice

Let us consider perturbations with time and angular dependence \(e^{i\omega t - \omega \theta}\). On dimensional grounds, a scale-free disk cannot support modes with \(\omega \neq 0\). The limiting case \(\omega = 0\) signals the onset of bifurcation or marginal stability of a particular mode. The most general form of \(h_{(a)(b)}\) may be written as
\[ h_{(a)(b)} = \begin{pmatrix} h_{00} & h_{01} & h_{02} & h_{03} \\
 h_{01} & h_{11} & h_{12} & h_{13} \\
 h_{02} & h_{12} & h_{22} & h_{23} \\
 h_{03} & h_{13} & h_{23} & h_{33} \end{pmatrix} e^{i\omega \theta}, \]
where the 10 \(h\)-entries are functions of \(\theta\) only. Geometrically, the metric coefficients are the inner products of the basis vectors,
\[ g_{\mu\nu} = \frac{\partial}{\partial \chi^\mu}, \frac{\partial}{\partial \chi^\nu}. \]

\(^1\) It is unfortunate that we have two \(\delta\) symbols here—one denoting Eulerian change and the other denoting the Kronecker-Delta function. There should be little confusion in context, however.

\(^2\) This is not entirely unfamiliar. Recall that in the super-Hamiltonian formalism, the conjugate momentum to \(x^\mu\) is \(p_\mu\), which is what we vary, not \(p^\mu\).
Since we are only considering polar perturbation, the system is symmetric about the equator, and the metric is invariant under the diffeomorphism \( \theta \to 2\theta_0 - \theta \). This implies the boundary condition

\[
h_{\theta \theta} = h_{\phi \phi} = h_{r \theta} = 0 \Rightarrow h_{\mu \lambda} = 0 \quad \text{for } \mu \neq 3
\]

on the disk.

We proceed as follows. Project the left-hand side of equation (3.2) onto the LNRO frame, and write the result as \( l_{(a)(b)}(1 + n)^2e^{im\phi}e^{N-Z}/r^2 \). Expand \( l_{(a)(b)} \), and replace the second derivatives of zeroth-order metric coefficients with the unperturbed Einstein equations (eq. [2.4]). This will introduce singular terms on the disk. Since we require the metric to be continuous, all singular terms must balance for the first-order equations in \( l_{(a)(b)} \), which are \( l_{(a)3} \) (the second-order equations are acceptable, since the first derivatives will in general have a jump there). Miraculously, with the condition in equation (3.3), all first-order equations are regular.

To proceed further, we need to choose a gauge. Consider an infinitesimal coordinate transformation \( x^\mu \to y^\mu = x^\mu + \xi^\mu(x) \), where \( \xi^\mu \) is of the same magnitude as \( h_{\mu \nu} \). This induces a transformation on the metric in the usual way,

\[
g_{\alpha \beta}(x) = \frac{\partial y^\alpha}{\partial x^\alpha} \frac{\partial y^\beta}{\partial x^\beta} g_{\mu \nu}(x + \xi) = \left( \xi^\mu_{,\alpha} + \xi^\mu_{,\beta} \right) \left( \xi^\nu_{,\alpha} + \xi^\nu_{,\beta} \right) \left( g_{\mu \nu} + \xi^\nu g_{\mu \nu,\rho} \right) = g_{\alpha \beta} + g_{\mu \nu} \xi^\mu_{,\beta} + \xi^\nu g_{\alpha \beta,\rho} = g_{\alpha \beta} + \left( C \xi g^\mu \right)_{\alpha \beta} = g_{\alpha \beta} + 2\xi_{\alpha \beta,\rho}.
\]

Thus, the coordinate freedom we have in general relativity corresponds to the gauge freedom \( h_{\mu \nu} \to h_{\mu \nu} + 2\xi_{\mu \nu,\rho} \). As suggested by the boundary condition on the disk, we promote equation (3.3) to a gauge condition. There is one more degree of freedom, which we fix here. The total gauge thus reads

\[
h_{03} = h_{13} = h_{23} = 0, \quad h_{11} = h_{33}.
\]

The last condition resembles the Regge-Wheeler gauge in spherical symmetry. With the gauge condition, we may write

\[
h_{tt} = r^{2n}e^{N}(a + Q^2b - 2Qd)e^{im\phi}, \quad h_{\phi \phi} = r^{n+1}e^{P(d - Qb)}e^{im\phi},
\]

\[
h_{rr} = ir^ae^{Z/2}(f - Qj)e^{im\phi}, \quad h_{\phi \phi} = r^2e^{2P-N}be^{im\phi} = e^{2P-Z}h_{\theta \theta},
\]

\[
h_{\phi \phi} = ire^{P+Z/2-N}je^{im\phi}, \quad h_{rr} = e^{Z-N}ce^{im\phi},
\]

which corresponds to

\[
l_{(a)(b)} = e^{im\phi} \begin{pmatrix} a & d & if & 0 \\ d & b & ij & 0 \\ if & ij & c & 0 \\ 0 & 0 & 0 & b \end{pmatrix}.
\]

The left-hand side of equation (3.2) now reads

\[
l_{00} = -a'' - \left( \frac{1}{2}N' + P' \right)a' + \frac{1}{2}N'e' + 2QFd' + \left[ -Q^2 \frac{1 - n}{1 + n} \right] a + Q^2 \frac{1}{1 + n} - Q^2 F^2 - \bar{\epsilon} \Delta \frac{2 + 2\gamma + v^2}{1 - v^2} + e^{Z-2P} \frac{m^2}{(1 + n)^2} \] a
\]

\[
+ \left[ -2Q^2 \frac{1 - n}{1 + n} \right] a + Q^2 \left( 1 - n \right)^2 - Q^2 F^2 + 2 \frac{2n}{1 + n} + 2e^{Z-2P} Q^2 \frac{m^2}{(1 + n)^2} - \frac{2 + 2\gamma + v^2}{1 - v^2} \] b
\]

\[
+ \left[ Q^2 \frac{1 - n}{1 + n} \right] c + 2 \left\{ Q \left[ \left( \frac{1 - n}{1 + n} \right)^2 - e^{Z-2P} \frac{m^2}{(1 + n)^2} + \left( N' - P' \right) F \right] - 2v(1 + \gamma) \right\} d
\]

\[
= e^{Z/2-P} 2Qm(1 + 2n) \frac{m}{(1 + n)^2} + e^{Z/2-P} 2m(n + nQ^2 - Q^2) f.
\]
\[ l_{11} = -b'' + \left( P' - \frac{1}{2} N' \right) a' - P'b' + \left( \frac{1}{2} N' - P' \right) c' + 2QF d' + \begin{cases} \left\{ -e^{Z-2P} \frac{m^2}{(1+n)^2} - Q^2 \left[ \left( \frac{1-n}{1+n} \right) + P^2 \right] \right\} a \\ + e^{Z-2P} \frac{(1-Q^2)m^2}{(1+n)^2} - Q^2 \left[ 2 \left( \frac{1-n}{1+n} \right) + P^2 \right] - 2 \left( \frac{1}{1+n} \right) b \\ + \left[ e^{Z-2P} \frac{m^2}{(1+n)^2} + Q^2 \left( \frac{1-n}{1+n} \right)^2 + 2 \left( \frac{1}{1+n} \right) \right] c \\ + 2Q \left[ e^{Z-2P} \frac{m^2}{(1+n)^2} + \frac{1-n}{1+n} \right] (N' - P') \right\} d \\ + 2Qe^{Z/2-p} \frac{m}{(1+n)^2} f - 2e^{Z/2-p} \frac{m}{(1+n)^2} \left[ Q^2 (1-n) + n + 2 \right] j, \end{cases} \]

\[ l_{22} = -c'' + \frac{1}{2} (Z' - N') a' + \frac{1}{2} (N' - Z' - 2P') c' + Q^2 \left( \frac{1-n}{1+n} \right)^2 a + \begin{cases} \left[ 2Q^2 \left( \frac{1-n}{1+n} \right)^2 + 2n(1-n) - \tilde{\varepsilon} \Delta \right] b - 2Q \left( \frac{1-n}{1+n} \right)^2 d \\ + \left[ (1-Q^2) e^{Z-2P} \frac{m^2}{(1+n)^2} + \tilde{\varepsilon} \Delta - Q^2 \left( \frac{1-n}{1+n} \right)^2 - 2n(1-n) \right] c \\ + Qe^{Z/2-p} \frac{2mn}{(1+n)^2} f + e^{Z/2-p} \frac{2m}{(1+n)^2} \left( Q^2 - 1 - nQ^2 \right) j, \end{cases} \]

\[ l_{01} = -d'' + \frac{1}{2} QF a' + QF b' + \frac{1}{2} QF c' - P'd' + \begin{cases} Q \left[ e^{Z-2P} \frac{m^2}{(1+n)^2} - 2 \frac{1-n}{(1+n)^2} \right] b + Q \left[ e^{Z-2P} \frac{m^2}{(1+n)^2} + 2 \frac{1-n}{(1+n)^2} \right] c \\ + \left[ (1-Q^2) \left( \frac{1-n}{1+n} \right)^2 - \tilde{\varepsilon} \Delta \gamma + (N' - P')^2 - Q^2 F^2 \right] d \\ - e^{Z/2-p} \frac{2mn}{(1+n)^2} f + 2Qe^{Z/2-p} \frac{m(n-2)}{(1+n)^2} j, \end{cases} \]

\[ i l_{02} = f'' + P'f' - QF j' + \begin{cases} \left[ Qe^{Z-2P} \frac{m^2}{(1+n)^2} + \left( P' - \frac{1}{2} Z' \right) \right] QF + 2 \left( \frac{1 + \gamma}{1 - v^2} \right) \tilde{\varepsilon} \Delta \right] j \\ + e^{Z/2-p} \frac{m}{(1+n)^2} \left[ \frac{1}{2} Q(1-n)a + Q(1-n)b - \frac{1}{2} (n+3)Qc - (1-n)d \right] \\ + \left[ -e^{Z-2P} \frac{m^2}{(1+n)^2} \left( N' - \frac{1}{2} Z' \right)^2 + 2n(1-n)^2 + \frac{1}{2} Q^2 F^2 + \frac{v^2 + \gamma}{1 - v^2} \tilde{\varepsilon} \Delta \right] f, \end{cases} \]

\[ i l_{12} = j'' - QFj' + P'j' + \begin{cases} \left[ -e^{Z-2P} \frac{m^2}{(1+n)^2} - 2 \frac{1-n}{(1+n)^2} + \left( \frac{1}{2} Z' - N' \right) \right] F \\ + e^{Z/2-p} \frac{m}{(1+n)^2} \left[ (1-n)\left( \frac{1}{2} Q^2 \right) a + 2(1-n)Q^2 b \\ - \left[ 1 + n \frac{1}{2} Q^2 (1-n) \right] c - 3(1-n)Qd \right] \end{cases} \]
Recall that the disk is made of a two-dimensional perfect fluid. Explicitly, if we choose the equation of state $p = \gamma e$, the unperturbed stress-energy tensor may be written as

$$
T^\mu{}\!{}_{\nu} = \varepsilon[(1 + \gamma)u^\mu{}u^\nu + \gamma g^\mu{}\!{}_{\nu}],
\mu, \nu = t, \phi, r \text{ and } T_{\rho\theta} = 0.
\tag{3.7}
$$

In the presence of a perturbation, we still need to impose the condition that momentum flux and stress in the vertical direction vanish. Hence, the first-order change in the stress-energy tensor is only for the top left 3x3 block:

$$
\delta T^\mu{}\!{}_{\nu} = \delta \varepsilon[(1 + \gamma)u^\mu{}u^\nu + \gamma g^\mu{}\!{}_{\nu}] + \varepsilon[(1 + \gamma)(\delta u^\mu{}u^\nu + u^\mu{}\delta u^\nu) - \gamma h^\mu{}\!{}_{\nu}].
$$

As usual, the four-velocity is normalized,

$$(u^\mu{} + \delta u^\mu{})(u^\nu{} + \delta u^\nu)(g_{\mu\nu} + h_{\mu\nu}) = -1 \Rightarrow \delta u^\mu u_\mu = -\frac{1}{2} u^\mu u^\nu h_{\mu\nu}.
$$

Projecting onto the LNRO frame, we have

$$
\delta T_{(a)\!(b)} = \delta \varepsilon[(1 + \gamma)u_{(a)}u_{(b)} + \gamma \tilde{e} a_{(a)\!(b)}] + \varepsilon[(1 + \gamma)(\delta u_{(a)}u_{(b)} + u_{(a)}\delta u_{(b)}) - \gamma w_{(a)\!(b)}]
\tag{3.8}
$$

and

$$
\delta u_{(a)}u_{(a)} = -u_{(a)}u_{(b)}w_{(a)\!(b)},
$$

where

$$
u(u_{(a)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{\sqrt{1 - v^2}} & \frac{v}{\sqrt{1 - v^2}} & 0 & 0 \end{pmatrix},
\delta u_{(a)} = \begin{pmatrix} xe^{im_\phi} & ye^{im_\phi} & ze^{im_\phi} & 0 \\ \frac{x}{\sqrt{1 - v^2}} & \frac{y}{\sqrt{1 - v^2}} & \frac{z}{\sqrt{1 - v^2}} & 0 \end{pmatrix}.
$$

With this parameterization, the normalization condition reads

$$
x = (y + d)v + \frac{1}{2}(a + v^2 b).
\tag{3.9}
$$

Next, we work out the EOM for the perturbed quantities. Although it is not needed if we solve all six unknown metric perturbations directly, the EOM provides a consistency check for the algebraic mess that is notorious in general relativity. Furthermore, as we see below, for the self-similar disk, the EOMs are all algebraic and thus are much easier to handle than the full Einstein equations. A direct variation of $T^\mu{}\!{}_{\nu} = 0$ reads

$$
\delta \Gamma^\mu{}\!{}_{\rho\nu} + \delta \Gamma^\mu{}\!{}_{\rho\nu} T^\rho{}\!{}_{\nu} + \delta \Gamma^\nu{}\!{}_{\rho\nu} T^\rho{}\!{}_{\mu} = 0,
$$

where

$$
\delta \Gamma^\mu{}\!{}_{\rho\nu} = \frac{1}{2} g^{\rho\nu}(h_{\alpha\nu\rho} + h_{\alpha\rho\nu} - h_{\rho\nu\alpha}).
$$

Just as in the unperturbed EOM, the $\mu = \theta$ component needs to be satisfied identically. Physically, this component gives the fluid evolution in the $\theta$-direction, which is trivial in this case. If we assume that the metric coefficients are even about the equatorial plane, then the first derivatives are odd and vanish upon integrating through the plane. The resulting equation only contains $h_{\rho\theta}$, where $\nu \neq 3$. This is another reason why the full gauge has to satisfy the boundary condition (eq. [3.3]). Next, we consider the density perturbations. The zeroth-order density that is self-similar may be written as

$$
\varepsilon = \frac{4}{r} \delta(\theta - \theta_0),
$$

where
where $A = \bar{e} e^{N_0 - Z_0}/8\pi$ is some constant. Conventionally, the equatorial plane is located at $\theta = \pi/2$ in spherical polar coordinates. However, to reduce eigenvalues of the problem, we rescaled $\theta$ so that the disk is located at $\theta_0$. When the geometry is perturbed, this “hidden” eigenvalue will in general change; hence, we need to vary $\theta_0$ as well. Thus, we obtain

$$
\delta \varepsilon = \frac{\delta A e^{im\phi}}{r^2} \delta(\theta - \theta_0) + \frac{A}{r^2} \left[ \delta(\theta - \theta_0 - \theta_1 e^{im\phi}) - \delta(\theta - \theta_0) \right]
$$

$$
= \frac{e^{im\phi}}{r^2} \left[ \delta A \delta(\theta - \theta_0) - A \delta'(\theta - \theta_0) \theta_1 \right].
$$

(3.10)

If we project the right-hand side of equation (3.2) onto the LNRO frame and write the result as $H_{(\omega)(\eta)} (1 + n)^2 e^{i m \phi} e^{N_0 - Z}/r^2$, the nonzero components are

$$
H_{00} = \frac{1 + 2 \gamma + \nu^2}{1 - \nu^2} \zeta \Delta + \left[ - \frac{1 + \gamma \nu^2 + Q \nu (1 + \gamma)}{1 - \nu^2} a + \frac{v(2\nu - Q)(1 + \gamma)}{1 - \nu^2} b 
+ \frac{(\nu - Q)(1 + \gamma) - Q \nu^2 (1 + \gamma)}{1 - \nu^2} d + 4\nu \frac{1 + \gamma}{1 - \nu^2} \right] \epsilon \Delta,
$$

$$
H_{01} = -\frac{2\nu (1 + \gamma)}{1 - \nu^2} \zeta \Delta + \left[ - \frac{(2 + \nu^2)(1 + \gamma)}{1 - \nu^2} b - \frac{2(1 + \gamma \nu^2) + \gamma + \nu^2}{1 - \nu^2} d 
- 2 \left( \frac{1 + \gamma}{1 - \nu^2} + \nu \right) \right] \epsilon \Delta,
$$

$$
i H_{02} = \epsilon \Delta \left( \frac{\gamma + \gamma \nu^2 + 2}{1 - \nu^2} f + 2\nu \frac{1 + \gamma}{1 - \nu^2} j + 2 \frac{1 + \gamma}{1 - \nu^2} \right),
$$

$$
H_{11} = \frac{1 + \gamma + 2 \nu^2}{1 - \nu^2} \zeta \Delta + \left[ \frac{(\gamma + 3 \nu^2 + 2 \gamma \nu^2) - 3 \nu \frac{1 + \gamma}{1 - \nu^2} d + 4 \nu \frac{1 + \gamma}{1 - \nu^2}}{1 - \nu^2} \right] \epsilon \Delta,
$$

$$
i H_{12} = -\epsilon \Delta \left( 2\nu \frac{1 + \gamma}{1 - \nu^2} f + \frac{\gamma + \gamma \nu^2 + 2 \nu^2}{1 - \nu^2} j + 2 \nu \frac{1 + \gamma}{1 - \nu^2} \right),
$$

$$
H_{22} = \zeta \Delta + \gamma e \epsilon \Delta,
$$

(3.11)

where

$$
\zeta \Delta = 8\pi \delta \varepsilon \frac{r^2 e^{2^{Z_0 - N_0}}}{1 + n} e^{-im\phi}.
$$

If we use the definition of $\delta \varepsilon$, the above expression simplifies to

$$
\zeta \Delta = \tilde{\varepsilon} \Delta - \tilde{\varepsilon} e^{N_0 - Z_0} e^{Z_0 - N_0} \Delta \Theta_1 = \tilde{\varepsilon} \Delta + \tilde{\varepsilon} (Z'_0 - N'_0) \Theta_1 \Delta \Rightarrow \zeta = \tilde{\varepsilon} + \tilde{\varepsilon} (Z'_0 - N'_0) \Theta_1,
$$

where we integrated by parts on the last term. In fact, the explicit form of $\zeta$ is not required here, since $\tilde{\varepsilon}$ and $\Theta_1$ never appear independently in the equations. This observation suggests that $\Theta_1$ is a second-order effect.

3.3. Boundary Conditions

From symmetry, all the nonaxisymmetric metric components should vanish on the axis (where $\Theta = 0$). Thus, we can impose the conditions

$$
a = b = c = d = f = j = 0.
$$

(3.12)

On the disk, the delta functions give rise to a jump in the derivatives of the perturbation functions using the second-order equations. Integrating across the disk, we have

$$
2a' = \tilde{\varepsilon} \left[ \frac{\nu^2 + 2 \gamma - \gamma \nu^2 - Q \nu (1 + \gamma)}{1 - \nu^2} a + \frac{1 + 2 \gamma + 3 \nu^2 + 2 \gamma \nu^2 - Q \nu (1 + \gamma)}{1 - \nu^2} b 
+ \frac{(1 + \gamma)(5\nu - Q \nu^2)}{1 - \nu^2} d + \frac{4\nu (1 + \gamma)}{1 - \nu^2} \right] + 2 \frac{1 + 2 \gamma + \nu^2}{1 - \nu^2} \zeta,
$$

$$
2b' = \tilde{\varepsilon} \left[ \frac{2 \gamma \nu^2 + 3 \nu^2 + \gamma}{1 - \nu^2} b + \frac{v(1 + \gamma)}{1 - \nu^2} (3d + 4y) \right] + \frac{1 + \nu^2 + 2 \gamma \nu^2}{1 - \nu^2} \zeta,
$$

$$
2c' = \tilde{\varepsilon} [b - (1 - \gamma)c] + \zeta,
$$

3 Actually $H_{33}$ is nonzero, too, but it contains exactly the singular terms from $l_{33}$. 


The equation for \( N \) reads
\[
N'' + N' \cot \Theta + 2n = \tilde{\varepsilon} \Delta,
\]
which has the solution
\[
N' = -2n \tan \frac{\Theta}{2}, \quad 4n \tan \frac{\Theta_0}{2} = \tilde{\varepsilon}.
\]
Since \( \Theta_0 \approx \pi/2 \), we have \( \tilde{\varepsilon} \approx 4n \), which means that it is also small. Thus, the equation for \( P \) becomes
\[
0 = 2 \cot \Theta_0 \Rightarrow \Theta_0 \equiv \frac{\pi}{2}.
\]
and $\Theta \equiv \theta$. Here $Z$ may be most directly computed through the second constraint equation, relating it to $N'$,

$$Z' = 2nN' = -4n^2 \tan \frac{\theta}{2} = O(v^4).$$

Finally, the last constraint equation may be solved order by order. The $O(v^4)$ terms are identically zero by our solution of $N'$. The next order is $O(v^6)$, which reads

$$(Q \cot \theta - Q')^2 - Q^2 - 2Q \cot \theta(Q \cot \theta - Q') = 0$$

$$\Rightarrow \log Q' = \csc \theta$$

$$\Rightarrow Q = C \tan \frac{\theta}{2}.$$  

Integrating the dynamic equation for $Q$, we obtain a jump condition, which in this limit reads

$$C = 2n(C + 2v) \Rightarrow C = 4nv.$$ 

When we put everything together, the limiting solution is

$$n = v^2 + \gamma, \quad \tilde{v} = 4n, \quad P = \log \sin \theta,$$

$$N' = -2n \cot \frac{\theta}{2}, \quad Q = 4nv \tan \frac{\theta}{2}, \quad Z' = -4n^2 \tan \frac{\theta}{2}.$$  

(4.3)

Of course, in the purely Newtonian case, $Q$ and $Z$ are taken to be $0$, since they are of higher order. A simple integration yields

$$N = -2n \log(1 - \cos \theta)/2 \Rightarrow N = -\left(v^2 + \gamma\right) \log \left[2 \sin \frac{\theta}{2} \left(1 - \cos \theta\right)\right];$$

which is the correct result for a hot Mestel disk.

In the presence of a perturbation, we still demand that the Newtonian limit be valid. Thus, the only nontrivial metric perturbation, which is $a$, is in $g_{00}$, and

$$g_{00} = -r^{2n}e^N(1 - ae^{im\phi}).$$

When we expand everything to leading order in $v$, the Einstein equations reduce to

$$-a'' - \cot \theta a' + \frac{m^2}{\sin^2 \theta} a = \zeta \Delta.$$  

(4.4)

This is nothing more than the Poisson equation for the perturbed potential $\Phi = -a/2$.

Next, we derive the Newtonian version of the EOMs. From the Poisson equation, we know that $a$ is of order $\zeta$, which is in turn of order $v^2$, while $y$ and $z$ are both of order $v$. It is worthwhile to point out that we are doing a two-parameter expansion, one in $v$ and one in the perturbation. In this limit, the EOM becomes

$$y + \frac{v \zeta}{\tilde{v}} = 0,$$

$$\frac{a_0}{2} - 2\nu y - \frac{v}{m} z - (v^2 + \gamma) \frac{\zeta}{\tilde{v}} = 0,$$

$$2y + mz = 0.$$  

(4.5)

where $a_0$ is evaluated on the disk, of course. A less trivial calculation by Galli et al. (2001) shows that $a_0 = \zeta/2m$. Combining all three equations, we have [recall that $\tilde{v} = 4(v^2 + \gamma)$]

$$\frac{v^2 + \gamma}{m} + v^2 - \frac{2v^2}{m^2} - 2v^4/m^2 - \gamma = 0 \Rightarrow \frac{v^2}{\gamma} = \frac{m^2}{m + 2}, \text{ or } m = 1.$$ 

This is the Newtonian bifurcation point obtained by Shu et al. (2000).  

5. NUMERICAL IMPLEMENTATION

Equation (3.2) is linear in the metric perturbations. Therefore, it is ideal for finite differencing, which transforms the differential equations into a matrix equation. Consider a vector

$$V = a \oplus b \oplus c \oplus d \oplus e \oplus f \oplus j \oplus x,$$  

(5.1)

The conclusion that eccentric $m = 1$ bifurcations can occur at any level of disk rotation is flawed, as are the analyses of Syer & Tremaine (1996), Shu et al. (2000), and Galli et al. (2001) on this point (A. Toomre 2002, personal communication; F. H. Shu et al. 2004, in preparation). Except for one special rotation rate (the true bifurcation value), the stress tensor is nonzero at the origin (implying a physically unrealistic steady flow of momentum from the origin to infinity), even though it has zero divergence.
where
\[ a = (a_1, a_2, \ldots, a_N), \quad a_i = a(\Theta_i), \quad \Theta_N = \Theta_{\text{max}}, \]
etc., and
\[ x = (\tilde{z}, y, iz). \]

We do not include the values of the perturbation on the pole, since they all vanish by the boundary conditions (eq. [3.12]). For simplicity, let us use a uniform grid, as we did in the unperturbed solution. Thus,
\[ \Theta_i = i h, \quad h = \Theta_{\text{max}} / N. \]

With these definitions, the differential operators may be written as
\[ a_i' = \frac{a_{i+1} - a_{i-1}}{2h}, \quad a_i'' = \frac{a_{i+1} - 2a_i + a_{i-1}}{h^2}. \quad \text{(5.2)} \]

It is easy to check that these expressions are accurate to second order. On the boundary, the situation is a little bit trickier. To find the correct differencing scheme, let us expand the function on the axis to second order:
\[ a_1 = a_0 + ha_0' + \frac{1}{2} h^2 a_0'', \]
\[ a_2 = a_0 + 2ha_0' + 2h^2 a_0''. \]

Taking the proper linear combinations, we have
\[ a_0' = \frac{4a_1 - 3a_0 - a_2}{2h}, \quad a_0'' = \frac{a_2 - 2a_1 + a_0}{h^2}. \quad \text{(5.3)} \]

Similarly, on the disk,
\[ a_{N-1} = a_N - ha_N' + \frac{1}{2} h^2 a_N'', \]
\[ a_{N-2} = a_N - 2ha_N' + 2h^2 a_N''. \]

Thus,
\[ a_N' = \frac{3a_N - 4a_{N-1} + a_{N-2}}{2h}, \quad a_N'' = \frac{a_{N-2} - 2a_{N-1} + a_N}{h^2}. \quad \text{(5.4)} \]

Equation (3.2) may now be written as a matrix equation,
\[ M \cdot V = 0. \quad \text{(5.5)} \]

It is not too hard to convince oneself that the matrix \( M \) is \((6N + 3) \times (6N + 3)\). After finite differencing, each component of \( l_i \) is represented by a \((6N + 3) \times N\) submatrix, where the left-hand side is evaluated at \( \Theta = h, 2h, \ldots, (N - 1)h \). The last row in this submatrix is replaced by the boundary condition of equation (3.13). We thus fill the first \( 6N \) rows of \( M \). The very last three rows are the remaining boundary conditions of equation (3.12).

In order for equation (5.5) to have nontrivial solutions \( V \), the determinant of \( M \) must vanish. Schematically, the elements of \( M \) are nonlinear functions of the azimuthal quantum number \( m \) and unperturbed metric coefficients (which are, in turn, functions of \( v \) and \( \gamma \)). Thus, for a given value of \( m \) and \( \gamma \), the problem of stability analysis is reduced to the root finding of the equation
\[ |M_{m,\gamma}(v)| = 0. \quad \text{(5.6)} \]

6. RESULTS

We scan the entire solution space looking for the solution to equation (5.6). As expected, the bifurcation points form two sets of tracks in the \( \gamma-v^2 \) space. The behaviors of these two tracks are fundamentally different, since they are caused by two different mechanisms. We discuss them separately.

6.1. Radiation-Driven Neutral Modes

The first set of tracks is shown in Figure 1. These modes are believed to be the analog of Rossby modes, first discovered by Chandrasekhar in 1970 and subsequently studied extensively in the context of neutron stars. Even though our self-similar disk
geometry is in some sense infinitely different from the finite spherical geometry of neutron stars, the underlying mechanism for these neutral modes can still be understood if one is comfortable with the idea of gravitational radiation with infinite wavelength. To make a bad situation even worse, since our disks are formally infinite in size, one is never able to reach the radiation zone to study the gravitational wave. However, since the original argument of Chandrasekhar did not rely crucially on the asymptotic flatness of spacetime, the qualitative result still holds in this pathological case. Consider a nonaxisymmetric disturbance in the disk, moving at a velocity $v_1 < v$. As a result, the total angular momentum (or more appropriately, the specific angular momentum) decreases. These perturbations will in general radiate because of nonaxisymmetry. If in the LNRO frame $v_1 < 0$, then gravitational radiation carries away negative angular momentum, which damps the amplitude of perturbation, and these modes are stable. On the other hand, if $v_1 > 0$, gravitational radiation carries away positive angular momentum, and thus the amplitude of perturbation has to grow to make the total angular momentum more negative.

For a given equation of state specified by $\gamma$, the radiation-driven neutral modes occur at lower Mach number for increasing $m$. This is expected from the analysis of Friedman & Schutz (1975, 1978a, 1978b). In fact, the onset of instability for $m \to \infty$ occurs at zero rotational velocity. However, for a realistic system, the strength of a particular unstable mode is intimately related to the magnitude of the imaginary part of its frequency. For the $m \to \infty$ mode, even though it is formally unstable at zero rotation, the characteristic growth timescale is infinite. This is due to the fact that multipole radiation is exceedingly weak for higher values of $m$. If we truncate the self-similar disk, the strongest unstable modes—that is, the modes with shortest growth time—are still the ones with small $m$. In the absence of viscosity, these modes will grow until the nonlinear effects set in and limit the final rotation speed of the full disk.

In addition to the R-mode tracks, the $Q = 1$ curve is also plotted in Figure 1. We would like to remind the readers that these tracks represent models in which the characteristic frequency of a given mode becomes purely real, i.e., modes that are marginally stable. It is not too difficult to convince oneself that every mode in the background of an ergoregion is unstable. Using the simplistic picture of retrograde disturbances, we see that in the ergoregion, every mode has to propagate in the direction of the underlying disk as seen by an LNRO, and gravitational radiation will drive them unstable. Furthermore, as Friedman (1978) demonstrated, a spacetime with an ergoregion is unstable even under scalar and vector perturbations. From Figure 1, the $m = 2$ track is entirely above the $Q = 1$ curve, and so is part of the $m = 3$ track. In the presence of perturbing matter fields other than those in the disk itself, the ergoregion is likely to put a more stringent limit on the maximum rotation rate for a disk.

The question arises whether the growth of R-modes might be suppressed by viscous torques due to, e.g., a magnetorotational instability (MRI) acting in the disk (Balbus & Hawley 1991). Since the viscous effects must act to erase nonaxial symmetries of length scale $\sim r/m$, the damping time associated with linear global R-modes must be the diffusion timescale $t_D \sim r^2/(\nu m^2)$, where $\nu$ is the effective MRI kinematic viscosity. If relativistic disks are even weakly magnetized and electrically conducting, as their Newtonian counterparts are believed to be, for $m \sim 1$, the timescale $t_D$ may be only 2 orders of magnitude longer than the dynamical timescale $r/\nu$. R-mode spin-down for such disks may then be effectively suppressed by the MRI viscosity, but calculations of non–self-similar relativistic disks are needed to answer definitively whether the growth rate of R-modes (here zero) can overcome the viscous damping.

### 6.2. Newtonian Bifurcation Track

The second set of tracks is the generalization of Newtonian bifurcation computed by Shu et al. (2000). In Figure 2, we plotted these extended “Newtonian” tracks for $m = 2, 3, 4, 5, \infty$. The finite $m$ values are the numerical results of solving equation (5.6).
When $m$ becomes large, we may approximate it by a continuous variable and use it as a parameter in the asymptotic expansion. Assume that all the perturbation amplitudes remain infinitesimal in the large-$m$ limit; then the coefficients of each power of $m$ need to vanish independently. This requirement translates to

$$
a + 2Q^2 b + Q^2 c - 2d = 0, \quad Q(1+2n)f = (n+nQ^2 - Q^2)j, \\
-a + (1-Q^2) b + c + 2Qd = 0, \quad Qf = [Q^2(1-n) + n + 2]j, \\
c = 0, \quad nQf = -(Q^2 - 1 - nQ^2)j, \\
Qb + Qc = 0, \quad nf = Q(n-2)j, \\
\frac{1}{2}Q(1-n)a + Q(1-n)b - \frac{1}{2}(n+3)Qc - (1-n)d = 0, \quad Qj = f, \quad Qf = Q^2j, \\
(1-n)\left(1 + \frac{1}{2}Q^2\right)a + 2(1-n)Q^2b - \left[1 + n + \frac{1}{2}Q^2(1-n)\right]c - 3(1-n)Qd. \quad (6.1)
$$

Regardless of whether the disk is rotating, these (linear, homogeneous) equations only have trivial solutions. This is the familiar Cowling approximation, where the metric perturbations vanish in the large-$m$ limit. These Cowling modes may be understood in the following schematic way. Recall that in the Newtonian limit, we can invert Poisson’s equation via a Green’s function and obtain an integral representation of the gravitational field. This procedure can also be done for the full Einstein equations in principle, although not analytically. Mathematically, the integral representation of metric perturbations is effectively an average of the matter perturbation times the Green’s function. Therefore, in the $m \to \infty$ limit, the gravitational field is indifferent to the matter perturbation, since it averages to the axisymmetric equilibrium over any finite angular integration. This line of argument may be made mathematically rigorous with some more thought, but it is not necessary here.

In the absence of a metric perturbation, the EOMs now simplify to

$$(1 + v^2 + 2Qv)y + \left(Q + \frac{1+\gamma v^2}{1+\gamma} + v\right)\frac{\zeta}{\tilde{y}} = 0,$$

$$(Q + Qv^2 + 2v)\dot{y} + \left(Q + \frac{v^2 + \gamma}{1+\gamma}\right)\frac{\zeta}{\tilde{y}} = 0.$$

Eliminating the factor $\zeta/\tilde{y}$, we can combine these equations to give

$$(1 + v^2 + 2Qv)(Qv + Qv\gamma + v^2 + \gamma) = (Q + Q\gamma v^2 + v + v\gamma)(Q + Qv^2 + 2v).$$
The last equation implicitly defines a surface of neutral modes \( Q(\theta_0) = f(\gamma, \nu) \). Recall that the solution space of axisymmetric disks can also be viewed as a two-dimensional surface defined by \( Q(\theta_0) = g(\gamma, \nu) \). Thus, the intersection of these two surfaces gives the bifurcation curve in the \( \gamma-\nu \) space. This curve is also plotted in Figure 2.

Near the origin, these bifurcations tracks recover the Newtonian result, where the bifurcation point is located at

\[
\frac{v^2}{\gamma} = \frac{m}{m + 2}, \quad v, \gamma \to 0.
\]

As \( \gamma \) increases, the relativistic effects become important. Intriguingly, the point of bifurcation occurs at a lower Mach number in a relativistic disk than in its Newtonian counterpart. The curve defined by \( Q = 1 \) on the disk is again plotted. As seen in Figure 2, these tracks are confined in the portion of solution space where the timelike Killing vector remains timelike. In fact, the only place where a bifurcating neutral mode is allowed in the ergoregion is where \( m = \infty \) and \( \gamma = 1 \). Even then, the spacetime is only marginally ergo-like, in the sense that \( Q \to 1^{-} \) on the disk. To understand this phenomenon, we examine the velocity perturbation corresponding to the bifurcating neutral modes. In the Newtonian limit, we can evaluate it analytically (see eq. [4.5], or eq. [21] of Shu et al. 2000):

\[
\delta v = y e^{i m \phi} = -\frac{v}{4(v^2 + \gamma)} \delta \hat{v}.
\]

In general, this component of the eigenvector needs to be computed numerically. Actually, the exact form, or even the magnitude of this velocity perturbation, is not important. It suffices to know that \( \delta v \) for this mode always has the sign opposite that of \( \delta \hat{v} \). In the linear perturbation regime, where we can still apply the superposition principle, this observation has the following physical picture.

The full disk solution has two components. One is the axisymmetric equilibrium rotating at \( v \) in the \( +\hat{\phi} \) direction, with energy density given by \( \varepsilon = \delta \hat{v}(\theta - \theta_0)/v^2 \). The second component is the nonaxisymmetric perturbation, which is a disk of infinitesimal energy density \( \delta \hat{v} \) and infinitesimal velocity field \( \delta \hat{v} \). Along the bifurcation track, this perturbation disk is always counterrotating. 

The empirical result of Figure 2 leads us to postulate that the existence of counterrotating nonaxisymmetric disturbance is another necessary condition for bifurcation, at least for the disk geometry. A corollary is that the full nonaxisymmetric disk spacetime cannot have a stable ergoregion. This conjecture leads naturally to the confinement of bifurcation tracks in the portion of the solution space without ergoregions. For models with \( Q > 1 \) on the disk, the timelike Killing vector becomes spacelike, and thus no counterrotating trajectory is allowed.

Can we elevate this conjecture to apply to a more generic relativistic nonaxisymmetric equilibrium? The answer is a cautious yes. Without digressing too much into the mathematical structure of Riemannian geometry, we would like to offer the following plausibility argument. Suppose we are able to construct a fully nonlinear, nonaxisymmetric stationary solution to the Einstein field equations. It cannot have an event horizon, since black holes cannot have “hair” (in this case, “hair” refers to mass multipole moments). Furthermore, in the absence of a spacelike Killing vector \( \partial \phi \), the nonvanishing component of \( g_{tt} \) will generate a time-dependent quadrupole moment, as seen by an inertial observer. As a result, gravitational radiation will continue to carry away angular momentum and energy until the system is either axisymmetric or static. In this aspect, the relativistic nonaxisymmetric equilibrium are analogous to the Dedekind ellipsoids, where the figure axes are static in an inertial frame, and the configuration is supported by pressure and internal motion. It can be shown (see, e.g., Chandrasekhar 1983) that a static metric can always be brought to the diagonal form after appropriate coordinate transformations. With a diagonal metric, the ergoregion defined by \( g_{tt} = 0 \) coincides with the event horizon. Therefore, the absence of an event horizon means that a nonaxisymmetric static solution does not have an ergoregion.

7. SUMMARY AND DISCUSSION

We performed a linear stability analysis of the relativistic self-similar disk against nonaxisymmetric perturbations. For simplicity, we restricted the class of perturbation under consideration to be self-similar and polar. Mathematically, this means that the scaling law is preserved and that the metric is symmetric about the midplane even in the presence of perturbation.

As expected, the Newtonian bifurcations found by Shu et al. (2000) and Galli et al. (2001) have extensions into the fully relativistic regime. These tracks seem to exist only in models that do not admit an ergoregion. The corresponding velocity perturbation is strictly negative for any positive energy density increase. We thus hypothesize that in addition to the existence of neutral modes, retrograde disturbance may also be a necessary condition for bifurcation to nonaxisymmetric disk equilibria. This line of argument leads us to speculate that the nonaxisymmetric equilibrium solutions in general cannot have ergoregions. We have no proof that this behavior is generic, whereas the bifurcation of rapidly rotating axisymmetric equilibria to nonaxisymmetric forms probably is.

In addition, we have discovered the onset of R-mode instability, which is driven by gravitational radiation. The marginal stability tracks follow the qualitative behavior first discussed by Friedman & Schutz (1978b), and the result here is probably also generic. For a self-similar disk, the entire \( m = 2 \) neutral-mode tracks and part of the \( m = 3 \) occur in models with ergocones. We believe that in general for a fixed equation of state, the onset of instability occurs either at the ergoregion formation or on the tracks we computed, whichever has lower velocity.

These studies are the first step toward constructing fully nonaxisymmetric relativistic equilibria. For a given value of \( \gamma \), if the axisymmetric state contracts quasi-statically by shedding angular momentum, the linear rotational velocity will increase. This evolution represents a vertical line in Figures 1 and 2. Eventually, the velocity will reach a value where a nonaxisymmetric mode becomes unstable. If SIDs undergo gravitational collapse, the secular instability will most likely survive over many dynamic timescales (which, in the purely self-similar case, are infinite). Therefore, the collapse will be fundamentally nonaxisymmetric. For
simplicity, we have only considered each Fourier component independently. In the linear perturbation regime, the effect of a general nonaxisymmetric disturbance can always be decomposed into its Fourier components, with each component decoupled from others. When the amplitudes become finite, our Fourier series will fail to converge, and a more detailed analysis is required.

The current work serves as a springboard to one of the ultimate challenges in numerical relativity—the fully nonlinear numerical simulation of a nonaxisymmetric collapse. Only then can we answer questions such as whether the central singularities of objects like relativistic SIDs remain naked.

We thank the referee for valuable comments. This work has been supported by Academia Sinica, through the Distinguished Postdoctoral Research Fellowship awarded to M. J. C., and by National Science Council, through grants NSC 91-2112-M-007-29 and NSC 92-2112-M-001-062 awarded to F. H. S.

REFERENCES

Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214
Bardeen, J. M., & Wagoner, R. V. 1971, ApJ, 167, 359
Cai, M. J., & Shu, F. H. 2002, ApJ, 567, 477
———. 2003, ApJ, 583, 391
Chandrasekhar, S. 1970a, ApJ, 161, 561
———. 1970b, ApJ, 161, 571
———. 1983, The Mathematical Theory of Black Holes (Oxford: Oxford Univ. Press)
Friedman, J. 1978, Commun. Math. Phys., 63, 243
Friedman, J., & Schutz, B. 1975, ApJ, 200, 204
Friedman, J., & Schutz, B. 1978a, ApJ, 221, 937
———. 1978b, ApJ, 222, 281
Galli, D., Shu, F. H., Laughlin, G., & Lizano, S. 2001, ApJ, 551, 367
Hayashi, C., Narita, S., & Miyama, S. M. 1982, Prog. Theor. Phys., 68, 1949
Lynden-Bell, D., & Pineault, S. 1978, MNRAS, 185, 685
Mestel, L. 1963, MNRAS, 126, 553
Shu, F. H., Laughlin, G., Lizano, S., & Galli, D. 2000, ApJ, 535, 190
Syer, D., & Tremaine, S. 1996, MNRAS, 281, 925
Toomre, A. 1982, ApJ, 259, 535
Wald, R. M. 1984, General Relativity (Chicago: Univ. Chicago Press)