THE ASTROPHYSICAL JOURNAL, 341:685–691, 1989 June 15

EVOLUTION OF SELF-GRAVITATING ACCRETION DISKS IN ACTIVE GALACTIC NUCLEI

ISAAC SHLOSMAN[1] AND MITCHELL C. BEGELMAN[2,3]
Joint Institute for Laboratory Astrophysics, University of Colorado, and National Institute of Standards and Technology
Received 1988 August 29; accepted 1988 November 15

ABSTRACT

We investigate the evolution of self-gravitating gaseous disks in active galactic nuclei on scales of ~ 10–10^3 pc. Star formation is a plausible outcome of the Jeans instability operating in a disk which violates the criterion for local stability. Even a low efficiency of star formation would deplete the gaseous disk on a short time scale and create a flat stellar system. These systems can evolve (sphericalize) secularly by means of stellar encounters but this process appears to be too slow to be important. Such flattened stellar systems may be common in the circumnuclear regions of disk galaxies. Conventional viscosities are inefficient in building anew the accretion process even in a cosmological time. Strongly self-gravitating disks are unstable to global nonaxisymmetric modes, which can induce radial inflow of gas in a short dynamical time. The latter effect is studied in a separate paper.

Subject headings: black holes — galaxies: nuclei — stars: formation

I. INTRODUCTION

The wide range of luminosities encountered in active galactic nuclei (hereafter AGNs) is presumably a direct consequence of a correspondingly broad range of accretion rates onto central supermassive black holes (SBHs). Unfortunately, very little is known about the source of this accreting material. The major sources suggested so far to feed the accretion in AGNs lie either in the nucleus itself, e.g., tidal disruption of stars by the SBH (Hills 1975, 1978; Young, Shields, and Wheeler 1977; Shields and Wheeler 1978), stellar collisions (Sanders 1970; Begelman and Rees 1978), stellar ablation (Shlosman 1988) and the second-order tidal heating effect (Allen and Hughes 1987), or in the host galaxy (Begelman, Blandford, and Rees 1984 and references therein). Assuming that the AGN is fueled by the host galaxy via an accretion disk, we have investigated the possible structure of the outer parts of this disk (beyond 1 pc) for different accretion rates. In an earlier paper (Shlosman and Begelman 1987, hereafter Paper I), we showed that such a disk, if fueled by cold and dusty interstellar gas, tends to preserve its low temperature, $T < 1000$ K, even when its energetics is dominated by backscattered hard radiation from the AGN. We also gave a necessary condition for disk fragmentation. The present paper elaborates on the consequencies of local Jeans instability in accretion disks of AGNs. We consider star formation as a possible outcome of local gravitational instability and investigate its impact on disk structure and evolution, as well as the effect it has on the accretion rate onto the SBH.

For a vertically homogeneous and isothermal disk with a scale height corresponding to a temperature $20T_{20}$ K, accretion rates in excess of $\sim \alpha c_s{}^3/G \sim 1.5 \times 10^{-5} \alpha T_{20}{}^{3/2}\ M_\odot$ yr^{-1} will cause the disk to become self-gravitating in the vertical (z) direction, while accretion rates in excess of $\sim 2\alpha c_s{}^2 v_\phi/G \sim 8 \times 10^{-3} \alpha T_{20} v_{\phi 2}\ M_\odot$ yr^{-1} will result in a globally (radially) self-gravitating disk. Here, for simplicity, we have adopted the α-scaling law of the viscous stress with the gas thermal pressure (Shakura and Sunyaev 1973), $v_{\phi 2} \equiv v_\phi/100$ km s^{-1} is the local Keplerian speed due to the external (stellar) potential and c_s is the velocity of sound in the gas. Both of the above accretion rates fall within the envelope of accretion rates generally assumed to be characteristic of AGNs. It is mainly for this reason that it is so important to understand the way disk structure and evolution are affected by self-gravity. The above accretion rates define three regimes under which disk accretion could be operating in astrophysical objects: (1) negligible self-gravity, relevant for cataclysmic variables, X-ray binaries, and low-luminosity AGNs; (2) dominant vertical self-gravity, which we discuss in §§ II and III of this paper, and (3) dominant radial (or global) self-gravity, discussed in a subsequent paper (Shlosman, Frank, and Begelman 1989).

II. VERTICALLY SELF-GRAVITATING GASEOUS DISKS

a) Jeans Instability and Fragmentation

We will use the term "vertically self-gravitating" to refer to disks which are subject to local Jeans instability, but which are not globally unstable. For an isothermal uniformly rotating disk with density ρ and characteristic thickness $h = (\int \rho\, dz)^2/\int \rho^2\, dz$, the criterion for an axisymmetric instability with wavelength λ is (Goldreich and Lynden-Bell 1965; Binney and Tremaine 1987)

$$Q < Q_{\rm crit}(r) = \frac{v_{\rm rs}\,\kappa}{\gamma G \Sigma} \sim 0.7\ ; \quad \gamma = \begin{cases} \pi & \text{gaseous disk} \\ 3.36 & \text{stellar disk} \end{cases}, \qquad (1)$$

where $v_{\rm rs}$ is the *radial* dispersion velocity in the disk, κ is the epicyclic frequency, and Σ is the surface density of the disk. In fact, the vertical conditions in the disk do not appear explicitly in the stability criterion: $Q_{\rm crit} = 1$ for infinitely thin disks (Toomre 1964). The fact that only radial dispersion velocity enters the stability criterion (1) may lead to an important difference between purely stellar and gaseous disks. The velocity dispersion in a gaseous disk is isotropic, while in a collisionless stellar disk the vertical velocity dispersion—an adiabatic invariant—seem to be decoupled from the velocity dispersion in the plane of the disk. We shall return to this point later on.

We first analyze the effect of the Jeans instability under the

[1] Now at Theoretical Astrophysics, Caltech.
[2] Presidential Young Investigator and Alfred P. Sloan Foundation Research Fellow.
[3] Also at Department of Astrophysical, Planetary and Atmospheric Sciences, University of Colorado, Boulder.

assumption that the stellar component responsible for the external potential is not perturbed itself. For a disk in solid body rotation, condition (1) can be expressed by the threshold surface density in the disk

$$\Sigma_{\rm crit} \simeq 0.2 T_{20}^{1/2} v_{\phi 2} r_{10}^{-1} \text{ g cm}^{-2} , \tag{2}$$

where r_{10} is given in units of 10 pc. Alternatively, using the α formalism, the critical accretion rate in a disk with Keplerian shear is given by

$$\dot{M}_{\rm crit,1} \simeq 5.4 \times 10^{-5} \alpha T_{20}^{3/2} M_\odot \text{ yr}^{-1} . \tag{3}$$

For accretion in the range $\dot{M}_{\rm crit,1} < \dot{M} < \dot{M}_{\rm crit,2}$ where

$$\dot{M}_{\rm crit,2} \simeq 8 \times 10^{-3} \alpha T_{20} v_{\phi 2} M_\odot \text{ yr}^{-1} \tag{4}$$

is the rate for the disk to become globally self-gravitating, the isothermal disk will become clumpy on the characteristic scale of the most rapidly growing wavelength $\sim h$, although for $\hat{\Sigma} \equiv \Sigma/\Sigma_{\rm crit} > 1$ a range of wavelengths $h \lesssim \lambda \lesssim \hat{\Sigma}^2 h$ will be unstable (Goldreich and Lynden-Bell 1965) with growth times $\sim t_{\rm orb} \hat{\Sigma}^{-1} (\lambda/h)^{1/2}$. Here $h \simeq 1.3 \times 10^{17} \hat{\Sigma}^{-1} T_{20}^{1/2} v_{\phi 2}^{-1} r_{10}$ cm, and $t_{\rm orb}$ is the local orbital time in the plane of the disk. Two possible situations can be realized in the disk as the outcome of Jeans instability: (1) the instability saturates before the individual fragments are able to contract sufficiently, leaving large cross-sections for direct two-body interactions between clumps (Paczyński 1978; Lin and Pringle 1987; Lin, Pringle, and Rees 1988); and (2) the fragments cool down on a sufficiently short time scale, collapse, and cease to interact hydrodynamically (Paper I). In the second case a quasi-continuous gaseous disk approximation becomes invalid and the accretion rate onto the central SBH is strongly affected. The ability of these self-gravitating clumps to contract further and eventually to form stars is the major topic we address below.

For a gaseous disk with a local viscosity as the cause of both internal heating and radial inflow, the cooling time scale must be equal to the viscous dissipation time scale, $t_{\rm cool} \sim t_{\rm diss}$. In a non–self-gravitating disk both of the above time scales are longer than the orbital time, $t_{\rm orb}$, by a factor of α^{-1}, but are much shorter than the viscous inflow time, $t_{\rm in} \simeq 6 \times 10^9 \alpha^{-1} T_{20}^{-1} v_{\phi 2} r_{10}$ yr. The energetics of the outer regions of realistic accretion disks at $r > 1$–10 pc, however, will probably be dominated *not* by internal viscous dissipation but by external sources, particularly by backscattered hard radiation from the immediate vicinity of the SBH or by radiation from a jet above the disk, as we argued in Paper I, and by diluted starlight. In either case, if the disk becomes vertically self-gravitating, a necessary condition for disk fragmentation which is independent of α implies that

$$\frac{t_{\rm cool}}{t_{\rm orb}} < \hat{\Sigma}^{-1} . \tag{5}$$

For disks whose temperature is regulated either by molecular cooling or by dust, condition (5) can be satisfied easily at distances larger than roughly a few parsecs from the SBH. Beyond this point, therefore, the disk may fragment.

In Paper I we assumed that the accretion rate onto the SBH even at a distance of $\sim$10–100 pc is a good measure of AGN luminosity, i.e., $L_{\rm AGN} \propto \dot{M}$. Here we relax the above condition and do not require such a correlation, at least not on time scales shorter than the inflow time $t_{\rm in}$.

Assuming that $10^{-2}\eta_{-2}$ is the backscattering efficiency of the central AGN luminosity L, we define a fiducial luminosity $\hat{L} = L/10^{43}$ ergs s^{-1} which corresponds approximately to $\dot{M}_{\rm crit,2}$ when $v_{\phi 2} = T_{20} = 1$. The incoming flux then is given by $F_{\rm back} \sim 8\eta_{-2}\hat{L} r_{10}^{-2}$ ergs cm^{-2} s^{-1}. Because the radiation enters symmetrically from both sides of the disk, the energy density is uniform in the disk and is given by $U_{\rm rad} \sim F_{\rm back}/c$. For disks with the usual ISM abundance of dust ($\sim$0.01 by mass) and a surface density $\hat{\Sigma} > 1$, the external flux of optical-to-soft X-ray radiation will be effectively absorbed by the surface layers and reradiated in the infrared. In order to estimate the disk temperature and the optical depth to its IR radiation, we have assumed graphite core grains, which may be a primary constituent of interstellar dust (Draine and Lee 1984), and whose properties are relatively well known (e.g., Kellman and Gaustad 1969; Gilman 1974; Draine 1981).

The temperature of a graphite grain of size a on the disk surface can be estimated from

$$\frac{1}{\pi} \int_0^\infty Q_{\rm abs}(\lambda, a) F_{\rm back}(\lambda) d\lambda = \int_0^\infty Q_{\rm abs}(\lambda, a) B_\lambda(T_{\rm gr,surf}) d\lambda , \tag{6}$$

where $B_\lambda(T_{\rm gr,surf})$ is the Planck distribution at the grain temperature and $Q_{\rm abs}(\lambda, a)$ is the efficiency factor for pure absorption (e.g., Spitzer 1978). The shape of the AGN spectrum has been approximated by $F_{\rm back}(\lambda) \propto \lambda^{-1}$. The temperature of both the surface grains and those inside the disk is insensitive to the grain size and $\sim 0.1 a_{-1}$ μm was adopted for a reference value. The revised version of $Q_{\rm abs} \approx 2.5 a \lambda^{-1}$ (Draine 1981; Draine and Lee 1984) was used for long wavelengths $\lambda > 1$ μm, and the approximation by Jura (1982) was used for $\lambda \simeq 0.01$ μm–1 μm. The grains become transparent at shorter wavelengths. Under these conditions the surface grain temperature is given by equation (6):

$$T_{\rm gr,surf} \simeq 45 (\eta_{-2} \hat{L}/a_{-1})^{1/5} r_{10}^{-2/5} \text{ K} . \tag{7}$$

The temperature of the interior grains will generally be lower than $T_{\rm gr,\,surf}$ because they are shielded from the backscattered UV and optical radiation. However, as a result of low grain emissivity in the infrared, the grain temperature in the disk will be higher than the effective blackbody temperature of the incoming nonthermal flux,

$$T_{\rm BB} \simeq 20 (\eta_{-2} \hat{L})^{1/4} r_{10}^{-1/2} \text{ K} . \tag{8}$$

When the disk is optically thin to the radiation emitted both by the surface layer and by the interior grains, the interior grain temperature is given by the solution of the equation $\sigma T_{\rm gr}^4 Q_P(T_{\rm gr}) \sim c U_{\rm rad} Q_P(T_{\rm gr,\,surf})$, where $Q_P = 6.7 \times 10^{-5} a_{-1} T$ is the Planck-mean radiative efficiency of the grains in the temperature range $5 \lesssim T \lesssim 10^3$ K (Draine 1981):

$$T_{\rm gr} \simeq 23 (\eta_{-2} \hat{L})^{6/25} a_{-1}^{-1/25} r_{10}^{-12/25} \text{ K} . \tag{9}$$

Even if backscattered AGN radiation is absent, the diluted stellar radiation in the inner parts of a galaxy will maintain the grain temperature at $T_{\rm gr,20} \gtrsim 1$ (Jura 1982). Knowing the interior and the surface temperatures, it is possible to estimate the optical depth in the disk to its own radiation, assuming a "normal" dust-to-gas ratio (e.g., Shull and Woods 1985).

$$\tau_{\rm gr}(T_{\rm rad}) \simeq 3 \times 10^{-2} \Sigma(r) T_{\rm rad}(r) , \tag{10}$$

where $T_{\rm rad}$ is equal to $T_{\rm gr}$ or $T_{\rm gr,\,surf}$, respectively, and $\Sigma(r)$ is given in units of g cm^{-2}. The disk can be found in the optically thin regime, i.e., $\tau_{\rm gr}(T_{\rm gr,\,surf}) < 1$ and $\tau_{\rm gr}(T_{\rm gr}) < 1$ or in the optically thick regime, i.e., both $\tau_{\rm gr}(T) > 1$. The intermediate situation, when the disk is thick to the radiation emitted by the

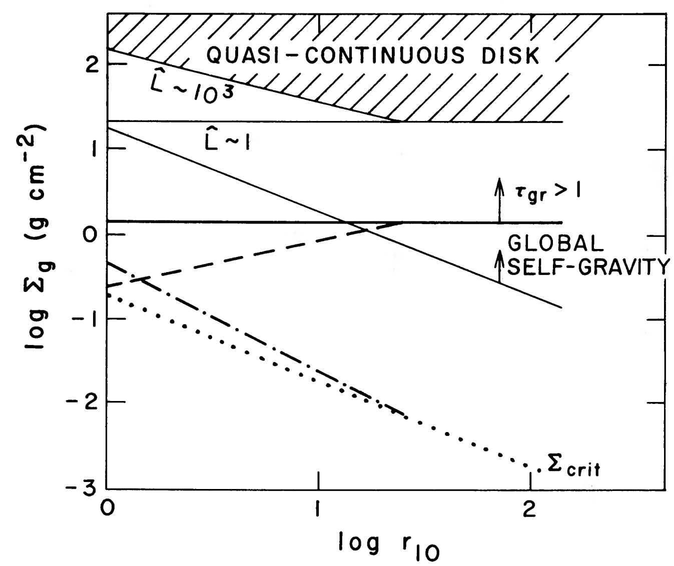

FIG. 1.—Optically thin vs. optically thick regime in the gaseous disk, with $v_{\phi 2} = 1$. The optical depth $\tau_{gr} = 1$ is given for $\hat{L} = 1$ (*heavy solid line*) and for $\hat{L} = 10^3$ (*dashed line*). The shaded area is defined by the condition for fragmentation (5) for $\hat{L} = 1$, 10^3. The boundary for dominant vertical self-gravity in the disk, Σ_{crit}, is given by dotted line ($\hat{L} = 1$) and dash-dotted line ($\hat{L} = 10^3$). Disks with Σ_g above the "global self-gravity" line are expected to be strongly unstable to bar modes.

surface but thin to the radiation from its interior, occurs for a very narrow range of parameters and therefore will not be considered. Although densities $n(H_2) \gtrsim 10^5$ cm^{-3} are required for gas-grain thermal coupling in the interstellar medium (e.g., Goldreich and Kwan 1974; Goldreich and Langer 1978), there are strong indications that densities $\sim 10^4$ cm^{-3} or even smaller are enough to ensure it in molecular clouds (Takahashi, Hollenbach, and Silk 1983). For the heating rates and densities we estimate for the interiors of self-gravitating disks, we expect the gas temperature to follow closely the grain temperature.

Figure 1 displays the two regimes in the disk: optically thick versus optically thin. A disk with a surface density in excess of $\sim \hat{L}^{-1/4} r_{10}^{1/2}$ g cm^{-2} will have $\tau_{gr} > 1$, which has an important implication for the development of Jeans instability. Namely, in the optically thick regime there is only one generation of fragments (Low and Lynden-Bell 1976) and the intial Jeans mass in the disk therefore reflects also the final mass of a newly formed star, i.e., further fragmentation in the disk is suppressed.

As was stated above, the evolution of a self-gravitating disk depends largely upon the ability of individual clumps to contract on a short time scale, i.e., to reduce their cross sections for direct and dissipative physical collisions—a problem of primary importance also in the theory of star formation for determining the initial stellar mass spectrum. Clearly the local Jeans mass in the disk prior to its fragmentation can provide us with a clue to possibly understanding the consequences of the gravitational instability. We shall cast our discussion of characteristic mass scales in the context of other possible differences and similarities between the conditions in molecular clouds—the major sites of star formation—and in the disk.

The maximum mass of a fragment in the disk at a radius r depends upon two factors: the surface density of the disk and the incoming flux. The actual mass of the fragment, in the optically thin regime, depends of course on the dynamics of the collapse, i.e., fragmentation versus coalescence. The local Jeans mass in the disk, $\sim \Sigma(r)h^2$, is given by

$$M_J \simeq 0.3 T_{gr,20}{}^2 \Sigma(r)^{-1} \, M_\odot \, . \quad (11)$$

Notice that M_J is a *decreasing* function of the disk surface density. Disks residing in the cores of galaxies hosting bright AGNs are capable of forming massive stars only in their outer parts, $r_{10} >$ few $\times$ 10, and only if the surface density is low, $\sim \Sigma_{crit}$.

Recent observations have shown that massive stars form preferentially in cores of giant molecular cloud complexes (GMCs), $M \simeq 10$–$10^3 \, M_\odot$, with a low efficiency of the order of $\sim 0.1\%$ per free-fall time $\sim 10^7$ yr (Blitz and Thaddeus 1980; Larson 1982; Duerr, Imhoff, and Lada 1982; Silk 1985, 1987). Dark molecular cloud cores (DMCs), $M \sim 1 \, M_\odot$, form almost exclusively low-mass stars with a much higher efficiency, $\sim 5\%$ per free-fall time $\sim 10^6$ yr (Wilking and Lada 1983; Myers 1985; Silk 1985, 1987). The GMC cores have higher temperatures by a factor of 3–10 compared to the smaller and colder DMC cores with $T \sim 10$ K (e.g., Myers 1985). This is important because the entire self-gravitating disk may be the site of star formation.

It follows from Figure 2, however, that disks with $\Sigma(r_{10} \gtrsim 1) > 30$–$100$ g cm^{-2} violate the necessary condition for disk fragmentation (eq. [5] and Paper I). The explanation lies in the high opacity of these disks that increases the cooling time, effectively damping the instability. We would expect the gaseous disk to be marginally stable in this regime (Paczyński 1978) and star formation to be suppressed. The further increase in the surface density may lead to a global self-gravitational instability and formation of a bar.

Low-mass stars are able to form throughout the unstable part of the disk, but high-mass stars potentially have the largest effect on the thermal balance in the disk. High surface densities, very much greater than Σ_{crit}, exclude massive stars from forming in the disk. Low-mass stars can heat the disk, if at all, only during the T Tauri phase, by means of their powerful winds. This kind of heating was found to be important for

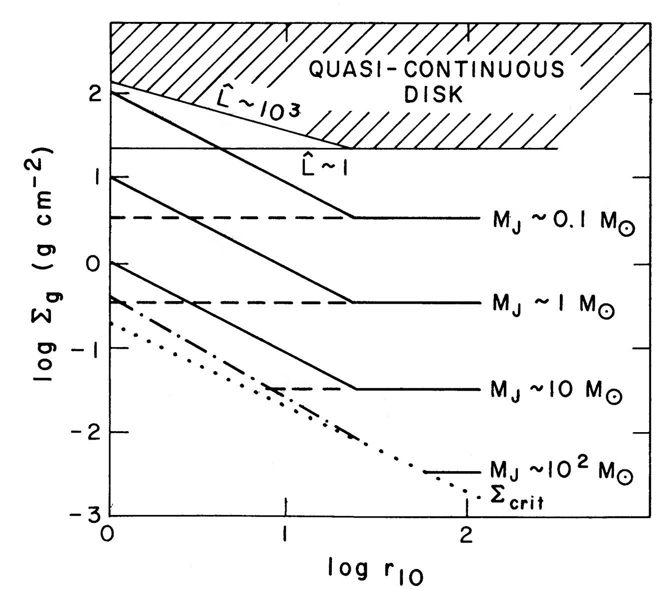

FIG. 2.—The distribution of Jeans masses in the gaseous disk at the onset of fragmentation for $\hat{L} = 1$ (*dashed line*) and for $\hat{L} = 10^3$ (*heavy solid line*). Other parameters are the same as in Fig. 1.

DMCs and GMCs (Norman and Silk 1980). We therefore consider this heating mechanism for the disk, along with heating mechanisms associated with high-mass stars. Because the disk fragments initially into clumps $\sim M_J$, rather than $\gg M_J$, it is likely that the clumps will *not* fragment subsequently even if they are optically thin, but will rather go into gravitational collapse and form single objects. This is related to the fact that density perturbations are smeared out on scales $\lesssim \lambda_J$ by the thermal pressure (Larson 1978; Silk 1980). In view of the latter, we shall assume that each clump collapses without fragmentation, forming a single stellar object.

b) *Effects of Star Formation on the Disk*

We assume that the onset of star formation in the disk leads to a gradual conversion of the gas into stars on a time scale of $t_* \sim 10^{1-3} t_{ff}$, where t_{ff} is the local dynamical time, which is of order $t_{orb}/\hat{\Sigma}$ in the disk and is generally shorter than the free-fall time in DMCs and GMCs. In the absence of any replenishment process, the disk would be depleted of gas on a time scale of a few t_*, and the SBH would be fed at the rate not exceeding $\dot{M}_{crit,1}$. Viscous dissipation operates on a time scale of $t_{in} \sim 6 \times 10^9 \alpha^{-1} T_{20}^{-1} v_{\phi 2} r_{10}$ yr (Paper I), which is much longer than t_*. For $\alpha \lesssim 1$ and at these radii this is a very inefficient process for bringing in the new material. Presumably it is superceded by a different mechanism for angular momentum transfer which operates on a much shorter time scale, if galactic material from $r_{10} \sim 1$–100 is to participate in the creation and feeding of the SBH. The dominant mechanism could be a global gravitational instability, for example. The latter possibility, which means effectively that $\alpha \gg 1$, is treated in a separate paper (Shlosman, Frank, and Begelman 1989). In this paper we deal with two limiting cases: (1) there is a counterbalancing process to star formation that keeps the disk surface density of the gas, Σ_g, constant in time (e.g., infall from high latitudes onto the disk); and (2) the total surface density in the disk (stars plus gas) is conserved over time scales of interest.

A low-mass star has its greatest effect on the surrounding ISM during its T Tauri phase, which lasts approximately the time it takes the star to reach the main sequence, $t_T \simeq 10^7 (M_T/M_\odot)^{-2}$ yr (e.g., Kashlinsky and Rees 1983). This phase in stellar evolution is characterized by a substantial outflow, $\dot{M}_T \simeq 10^{-8}\ M_\odot\ \mathrm{yr}^{-1}$, with asymptotic velocity $v_{w2} \simeq v_w/100$ km s^{-1} ~ 1. T Tauri stars are known to be formed in DMC cores. Their mechanical luminosities are $\sim 10\%$ of their bolometric luminosities (e.g., DeCampli 1981) and the momentum in the wind exceeds that of the radiation by two orders of magnitude (Bally and Lada 1983; Pudritz 1986).

To estimate the extent to which the T Tauri winds are capable of heating up the disk, we assume that their mass-loss rates scale linearly with the stellar mass, i.e. $\dot{M}_{T8} \equiv \dot{M}_T/10^{-8}\ M_\odot\ \mathrm{yr}^{-1} \simeq M_T/M_\odot$. The T Tauri wind acts as a piston sweeping up the disk gas and producing a cavity with a maximum radius that can be estimated by comparing the energy density in the cavity with the thermal pressure in the ambient medium (Castor, McCray, and Weaver 1975; Norman and Silk 1980). If the shocked wind inside the cavity behaves adiabatically, then this maximum radius is given by

$$R_{wT} \simeq (\dot{M}_T/4\rho c_s)^{1/2} v_w/c_s \simeq 7 \times 10^{17} \dot{M}_{T8}^{1/2} v_{\phi 2}^{-1} v_{w2} T_{20}^{-3/4} \hat{\Sigma}^{-1} r_{10}\ \mathrm{cm}\ , \quad (12)$$

where c_s is the velocity of sound and ρ is the unshocked density in the disk. The ratio R_{wT}/h may be somewhat larger than one, indicating that some of the pressure in the hot cavities may be discharged above and below the disk. This would reduce the effectiveness of heating by T Tauri winds. The shocked wind cools radiatively if $v_w \lesssim 100$ km s^{-1} (Norman and Silk 1980). Despite the wide range of possible densities in the disk, say $n \sim 10^3$–10^{10} cm^{-3}, the value of n will have a negligible effect on the thermal balance in the cavity (Castor, McCray, and Weaver 1975; Weaver *et al.* 1977; Norman and Silk 1980). In this case, most of the energy in the wind is radiated away; R_{wT} is smaller than in the adiabatic case by a factor $\sim (c_s/v_w)^{1/2}$, and $R_{wT}/h \lesssim 1$ throughout the disk. If all of the wind energy were available for heating, and denoting the surface density of the T Tauri stars in the disk by Σ_{*T}, we can estimate the volume heating rate in the disk:

$$\Gamma_T = \frac{1}{2} \frac{\dot{M}_T v_w^2}{M_T h} \Sigma_{*T} \simeq 2.4 \times 10^{-20} \beta_T v_{\phi 2}^2 \times v_{w2}^2 \hat{\Sigma}^2 r_{10}^{-2}\ \mathrm{ergs\ cm^{-3}\ s^{-1}} \quad (13)$$

versus the cooling rate for an optically thin disk:

$$\Lambda \simeq \frac{\tau_{gr} \sigma_{SB} T^4}{h} \simeq 8.4 \times 10^{-18} T_{20}^5 \times v_{\phi 2}^2 \hat{\Sigma}^2 r_{10}^{-2}\ \mathrm{ergs\ cm^{-3}\ s^{-1}}\ , \quad (14)$$

where $\beta_T = \Sigma_{*T}/\Sigma_g$ and σ_{SB} is the Stefan-Boltzmann constant. The heating rate Γ_T does not depend upon the grain radiative efficiency Q_P, because the stellar winds deposit energy primarily in the gas. The grains will gain this energy collisionally. Note that $\beta_T < 1$ means that there is less than one T Tauri star per $\sim h^3$ cell, and the cavities produced by the winds do not overlap. The condition $\beta_T \sim 1$ is reached after $t_* \sim \epsilon^{-1} t_{ff}$ (if the cavities remain open for this time), ϵ being the efficiency of star formation per $t_{ff} \simeq 5 \times 10^5 v_{\phi 2}^{-1} \hat{\Sigma}^{-1} r_{10}$ yr. We take $\epsilon \sim 0.05$ for the low-mass stars and $\epsilon \sim 10^{-3}$ for the massive stars, in accordance with the efficiencies observed in DMCs and GMCs, respectively (see § II*a*).

The equilibrium temperature in the disk due to heating by T Tauri winds can be estimated crudely from equations (13) and (14), assuming that the energy is spread uniformly through the disk:

$$T_{20}^{eq} \simeq 0.3 \beta_T^{1/5} v_{w2}^{2/5}\ \text{for}\ \beta_T \gtrsim 1\ . \quad (15)$$

It follows from equation (15) that T Tauri winds cannot substantially alter the energy balance in the disk, even if the surface density of T Tauri stars satisfies $\beta_T > 1$.

An interesting effect follows from the picture of colliding shells of swept-up material in the disk for $\beta_T \gg 1$ (provided the cavities excavated by T Tauri winds remain open for a sufficiently long time). According to the Norman and Silk (1980) scenario for molecular clouds, dense clumps will form at these intersections with initial velocities, v_{cl}, of order the relative velocities of the shells. As β_T in the disk increases, shells with smaller radii and progressively higher relative velocities will overlap, increasing v_{cl}. Unlike the situation in molecular clouds where clump kinetic energy is quickly dissipated via drag and collisions, a simple geometrical model shows that most of the clumps would escape from the disk up to a distance of $\sim v_w t_s \sim 1$ pc above the disk, where t_s is the sound crossing time of the clump and the typical clump size was taken to be $\sim 0.1h$. It was implicitly assumed in the above estimate that there is no confining medium above the disk. The consequences of this effect will be discussed elsewhere.

Massive stars with $M_* > 10\ M_\odot$ can form in the disk in a narrow range of surface densities only (Fig. 2):

$$0.2\eta_{-2}^{1/8} v_{\phi 2} r_{10}^{-1} < \Sigma(r) < 3 \times 10^{-2} \eta_{-2}^{1/2}\ \text{g cm}^{-2} \quad \text{for } r_{10} > \hat{L}^{1/2} \quad (16)$$

$$0.2(\eta_{-2}\hat{L})^{1/8} v_{\phi 2} r_{10}^{-6/5} < \Sigma(r) < 3 \times 10^{-2}(\eta_{-2}\hat{L})^{1/2} r_{10}^{-1}\ \text{g cm}^{-2} \quad \text{for } r_{10} < \hat{L}^{1/2}\ .$$

One consequence of inequality (16) is that a high incoming flux, $\hat{L} > 100$, in addition to low surface densities, is necessary for massive stars to form in the inner 100 pc. The outermost region of the disk is free from the above condition on $\hat{L}$, but still depends upon the fine tuning of $\Sigma(r)$, as can be seen also from Figure 2.

High-mass stars can have a three-fold effect on the cool disk. The radius of the H II region around an early-type star is given by

$$R_{\text{H II}} \simeq 66.9(S_{49}/n^2)^{1/3}\ \text{pc} \simeq 7 \times 10^{-3} S_{49}^{1/3} v_{\phi 2}^{-4/3} \hat{\Sigma}^{-4/3} r_{10}^{4/3}\ \text{pc} \quad (17)$$

(e.g., McKee 1986), where n is the volume density in the disk and $S_{49} \equiv S/10^{49}$ is the number of ionizing photons per second, e.g., $S_{49} \simeq 0.04$ for a B0 star, $S_{49} \simeq 1$ for an O6.5 star and $S_{49} \simeq 8$ for an O4 star. The presence of dust will decrease $R_{\text{H II}}$ even more. Therefore, in a dense self-gravitating disk $R_{\text{H II}}/h \ll 1$. These compact H II regions will have only local effects on the disk.

In addition to UV fluxes, OB stars produce line-driven winds (LDWs), $v_{\text{LDW}} \sim$ few $\times 10^3$ km s^{-1}. The effect of these winds on the disk is similar in many ways to that of the T Tauri winds. The interior of an OB bubble remains hot at $\sim 10^{6-7}(v_{\text{LDW}}/10^3\ \text{km s}^{-1})^2$ K, and the terminal radius (calculated as in eq. [12]) is always larger than the disk thickness. The bubble, hence, will discharge its internal pressure above the disk (McCray and Kafatos 1987), and the assumption of a uniform density ambient medium is incorrect. In fact, the cold shell surrounding the bubble will accelerate, become Rayleigh-Taylor unstable and break into filaments (Weaver *et al.* 1977). The expansion into the disk will be driven only by the intercepted fraction of wind *momentum*; the sideways expansion will stop well before $R_{\text{OB}}/h \sim 10$. A hot corona may form above the disk with a pressure scale height very much greater than h, and cold filaments will be injected into it with velocities $\sim v_{\text{LDW}}$. Because the main-sequence lifetime of an OB star, t_{OB}, is much shorter than the local star formation e-folding time t_*, the surface density of the OB stars in the disk will be low at all times, $\Sigma_{\text{OB}}/\Sigma_g \sim t_{\text{OB}}/t_* \sim 10^{(-2)-(-3)}$. The mass supply rate to the corona by these stars can be estimated as $\sim 10^{-4}\ M_\odot$ yr^{-1} inside ~ 100 pc for $\Sigma \sim \Sigma_{\text{crit}}$. Consequently, the effect on the energy balance within the disk is likely to be small.

Stars with $M > 7\ M_\odot$ become Type II supernovae (Trimble 1982). The effect of supernovae on the disk will be greatly diminished by the fact that most of the energy will be deposited above the disk, similarly to the case for OB stars. Once the supernova breaks through the disk, only the momentum of the ejecta will play a role in excavating the disk. For plausible values of Σ, we expect $R_{\text{SN}} \lesssim 10h$. A small amount of disk material may be ejected at high speeds, perhaps becoming line-emitting clouds or absorption-line systems, but the net effect on the energetics of the disk or on the accretion rate should be small.

III. EVOLUTION OF THE DISK OF STARS

The hydrodynamic approximation is invalid in the stellar disk. A vertically self-gravitating disk of stars will evolve in response to both collective effects and two-body encounters.

a) *Collective Effects*

In the limit of insignificant infall, $\dot{\Sigma}_g \ll \Sigma_g/t_*$, the disk can be assumed to consist of a single stellar fluid after a time $\sim t_*$. Hence, in the absence of efficient cooling mechanisms in the stellar disk, it will be heated by gravitational instability up to a point where $Q \sim 1$ (or $Q \sim Q_{\text{crit}} < 1$, if corrected for a finite disk thickness). While the Toomre parameter $Q \gtrsim Q_{\text{crit}}$ provides the condition against axisymmetric instabilities in the stellar or gaseous disk, a somewhat higher value of $Q \sim 3$ is needed to ensure that nonaxisymmetric instabilities will not grow with time (e.g., Binney and Tremaine 1987). The latter appear in the form of spiral waves whose characteristic range of wavelengths depends on the relative amounts of gas and stars (Lubow 1988 and references therein). Stellar spiral waves of small amplitude ($\sim 15\%$) can enhance the star formation process (as was shown for a case of a galactic disk by Lubow, Balbus, and Cowie 1986) and at the same time will heat up the stellar fluid, increasing Q and damping the instability. Large-scale nonaxisymmetric instabilities, particularly the mode $m = 2$, may result in the radial transfer of mass in the disk, and is treated separately (Shlosman, Frank, and Begelman 1989).

Any increase in surface density thereafter (e.g., through infall and subsequent star formation) will increase the stellar radial velocity dispersion, v_{rs}, while the vertical dispersion, v_{zs}, is an adiabatic invariant and, therefore, is difficult to affect (E. S. Phinney 1987, private communication). The resulting dispersion anisotropy cannot increase indefinitely, however. Hosepipe instability will operate on a dynamical time scale, if v_{zs}/v_{rs} becomes smaller than $\sim \frac{1}{3}$, buckling the disk and mixing the motion along the z- and r-axes (Kulsrud, Mark, and Caruso 1971; Polyachenko and Shukman 1977; A. Toomre 1987, private communication). Hence, it can be assumed that the increase in surface density results in the thickening of the stellar disk, simultaneously.

Star formation in the gaseous disk is triggered by local gravitational instability, which creates inhomogeneities and leads to disk fragmentation. The finite thickness of a stellar disk has an important effect on ongoing star formation in the embedded gaseous disk. As was shown by Toomre (1964) this finite thickness results in the effective reduction of the gravitational potential in the equatorial plane, which is equivalent to a reduction of Σ by a factor $(1 - e^{-kh})/kh$ in equation (1) where $k = 2\pi/\lambda$. Once the stellar disk is formed, its scale height will be $h_* > h$, increasing with time due to the lack of a cooling mechanism. The heating of the stellar component is due ultimately to the continuous gravitational instability in the gas. The dispersion relation for such a two-fluid system was studied extensively by Jog and Solomon (1984*a*, *b*) for the case of a galactic disk. Neglecting the thickness of the gaseous disk, the system becomes gravitationally unstable when

$$\frac{2\pi k G \Sigma_g}{\kappa^2 + k^2 c_s^2} + \frac{2\pi k G \Sigma_*}{\kappa^2 + k^2 v_{rs}^2} \frac{1}{k h_*} > 1\ , \quad (18)$$

where Σ_* is the surface density of the stellar disk. Condition (1) follows from equation (18) in the one-fluid limit.

Starting from a purely gaseous and gravitationally unstable disk, and on the assumption that there is no infall, i.e.,

$\Sigma_g + \Sigma_* \sim$ const., the final product of star formation is a stellar disk which is marginally stable. If $\hat{\Sigma} > 1$ initially, this final state requires an increase in the stellar dispersion velocity, v_{rs}, as a result of collective heating. The finite thickness of the stellar disk is also a stabilizing factor, decreasing the second term in equation (18) while the first term becomes very much less than 1 with the depletion of the gas.

Significant infall onto the disk, e.g., preserving $\Sigma_g \sim$ const (and $c_s \sim$ const), will *not* stabilize the system by means of star formation as long as $h_*/h > \Sigma_*/\Sigma_g$ or until the stars start to alter the external gravitational potential. We notice that the first term in equation (18) is already greater than one, and the second term is always positive. This result has an important consequence on the evolution of the whole system, ensuring that star formation will continue as long as the gaseous disk is gravitationally unstable, regardless of the stellar component. The newly formed stellar disk, however, will have an effect on the range of unstable wavelengths. As was shown numerically by Jog and Solomon (1984*a*, *b*) when $\Sigma_*/\Sigma_g >$ few, this range will broaden to $h \lesssim \lambda_J \leq h_*$.

b) Stellar Encounters

In a differentially rotating disk of stars the dispersion velocities can grow with time at the expense of the rotational energy. If collective effects in the fragmented disk rapidly increase the dispersion velocities up to $Q \gtrsim Q_{\rm crit}$, two-body stellar encounters take over at this point and operate on a secular (two-body relaxation) time scale, which can be estimated initially as

$$t_{\rm rel}{}^0 \sim 3 \times 10^6 \hat{\Sigma}^4 v_{\phi 2}{}^{-1} r_{10}\ {\rm yr}\ , \qquad (19)$$

and which may result in additional thickening of the disk. In the long run, $t_{\rm rel}$ itself is a function of stellar velocity dispersion, v_*, and the volume density of stars, n_*, i.e., $t_{\rm rel} \propto v_*{}^3/n_*$.

If the surface density of the disk remains constant (no infall), then $n_* \propto h_*{}^{-1} \propto v_*{}^{-1}$ and $t_{\rm rel} \propto v_*{}^4$. The characteristic time to double the disk thickness is $\sim 16 t_{\rm rel}$. If mass is added to the disk and the fragments are constantly created, so that $n_* \sim$ const, then $t_{\rm rel} \propto v_*{}^3$ and the disk doubles its thickness during $\sim 8 t_{\rm rel}$. In both limiting cases the system sphericalizes on a time scale which is long compared to the Hubble time. We notice from equation (19) that the innermost part of the system, $r \sim$ few parsecs, may achieve dispersion velocities $\sim 0.1 v_\phi$. It would appear then as a rapidly rotating and highly flattened ellipsoid.

The interaction between the disk of stars and any of the galactic spheroidal components will occur on the relaxation time scale of the hotter component, which is very long. Therefore such an energy exchange will be ignored here.

We therefore conclude that a vertically self-gravitating disk, even if it fragments, persists as a flattened configuration.

IV. DISCUSSION

We have argued that gravitational instability in a vertically self-gravitating disk will trigger star formation on scales $\sim$10–few × 100 pc. Unless there is substantial infall the gaseous disk will be depleted in a short time, much less than the inflow time required to rebuild the disk by conventional viscous processes. Heating of the gaseous disk by stellar and other processes appears to be inefficient. However, supernovae and OB stars in the disk may inject significant quantities of energy into the interstellar medium above and below the disk. Star formation should limit the mass supply to the central engine through the disk at $\sim \dot{M}_{\rm crit,1} \sim 10^{-4}\ M_\odot\ {\rm yr}^{-1}$, or even prevent the central black hole from forming at all.

In the framework of fueling an AGN with cold and dusty interstellar gas from the host galaxy, local viscous processes seem to be incapable of replenishing the gaseous disk, even on a Hubble time. Global gravitational instability that can induce radial flow of gas on a short dynamical time is a promising alternative mass supply mechanism for AGNs (Shlosman, Frank, and Begelman 1989).

Flattened and rapidly rotating stellar subsystems can be expected in the circumnuclear regions of disk galaxies—relics of the gravitationally unstable gaseous disks that once existed there. Recent observational results on the possible existence of SBHs in nearby galaxies (Dressler and Richstone 1988; Kormendy 1988) provide us with some insight concerning the stellar dynamical configuration there. We find it very intriguing that both of the above works stress the importance of rotation in the circumnuclear regions of the observed galaxies. Our estimates of stellar relaxation times (see § III*b*), indicate that rotationally supported stellar subsystems on scales greater than roughly a few pc will remain highly flattened to the present time.

We are grateful to our colleagues and especially to Juhan Frank, Sterl Phinney, Jacob Shaham, and Alar Toomre for illuminating discussions. This research was supported in part by NSF grant AST83-51997, NASA Astrophysical Theory Center grant NAGW-766 and grants from Ball Aerospace Systems Division, Rockwell International Corporation and the Alfred P. Sloan Foundation. We thank the director of the Institute for Theoretical Physics, University of California, Santa Barbara, where this work was initiated. Research at ITP is supported in part by NSF grant PHY82-17853, supplemented by funds from NASA.

REFERENCES

Allen, A. J., and Hughes, P. A. 1987, *Ap. J.*, **313**, 152.
Bally, J., and Lada, C. J. 1983, *Ap. J.*, **265**, 824.
Begelman, M. C., Blandford, R. D., and Rees, M. J. 1984, *Rev. Mod. Phys.*, **56**, 255.
Begelman, M. C., and Rees, M. J. 1978, *M.N.R.A.S.*, **188**, 847.
Binney, J., and Tremaine, S. 1987, in *Galactic Dynamics* (Princeton: Princeton University Press).
Blitz, L., and Thaddeus, P. 1980, *Ap. J.*, **241**, 676.
Castor, J., McCray, R., and Weaver, R. 1975, *Ap. J.* (*Letters*), **200**, L107.
DeCampli, W. M. 1981, *Ap. J.*, **244**, 124.
Draine, B. T. 1981, *Ap. J.*, **245**, 880.
Draine, B. T., and Lee, H. M. 1984, *Ap. J.*, **285**, 89.
Dressler, A., and Richstone, D. O. 1988, *Ap. J.*, **324**, 701.
Duerr, R., Imhoff, C. L., and Lada, C. T. 1982, *Ap. J.*, **261**, 135.
Gilman, R. C. 1974, *Ap. J. Suppl.*, **28**, 397.
Goldreich, P., and Kwan, J. 1974, *Ap. J.*, **189**, 441.
Goldreich, P., and Langer, W. D. 1978, *Ap. J.*, **222**, 881.
Goldreich, P., and Lynden-Bell, D. 1965, *M.N.R.A.S.*, **130**, 97.
Hills, J. G. 1975, *Nature*, **254**, 295.
———. 1978, *M.N.R.A.S.*, **182**, 517.
Jog, C. J., and Solomon, P. M. 1984*a*, *Ap. J.*, **276**, 114.
———. 1984*b*, *Ap. J.*, **276**, 127.
Jura, M. 1982, *Ap. J.*, **254**, 70.
Kashlinsky, A., and Rees, M. J. 1983, *M.N.R.A.S.*, **205**, 955.
Kellman, S. A., and Gaustad, J. E. 1969, *Ap. J.*, **157**, 1465.
Kormendy, J. 1988, *Ap. J.*, **325**, 128.
Kulsrud, R. M., Mark, J. W.-K., and Caruso, A. 1971, *Ap. Space Sci.*, **14**, 52.
Larson, R. B. 1978, *M.N.R.A.S.*, **184**, 69.
———. 1982, *M.N.R.A.S.*, **200**, 159.
Lin, D. N. C., and Pringle, J. E. 1987, *M.N.R.A.S.*, **225**, 607.

Lin, D. N. C., Pringle, J. E., and Rees, M. J. 1988, *Ap. J.*, **328**, 103.
Low, C., and Lynden-Bell, D. 1976, *M.N.R.A.S.*, **176**, 367.
Lubow, S. H. 1988, *Ap. J. (Letters)*, **328**, L3.
Lubow, S. H., Balbus, S. A., and Cowie, L. L. 1986, *Ap. J.*, **309**, 496.
McCray, R., and Kafatos, M. 1987, *Ap. J.*, **317**, 190.
McKee, C. F. 1986, *Ap. Space Sci.*, **118**, 383.
Myers, P. C. 1985, in *Protostars and Planets II*, ed. D. C. Black and M. Shapley-Matthews (Tucson: University of Arizona Press), p. 81.
Norman, C., and Silk, J. 1980, *Ap. J.*, **238**, 158.
Paczyński, B. 1978, *Acta. Astr.*, **28**, 91.
Polyachenko, V. L., and Shukman, I. G. 1977, *Soviet Astr. Letters*, **3**, 134.
Pudritz, R. E. 1986, *Publ. Astr. Soc. Pacific*, **98**, 709.
Sanders, R. H. 1970, *Ap. J.*, **162**, 784.
Shakura, N. I., and Sunyaev, R. A. 1973, *Astr. Ap.*, **24**, 337.
Shields, G. A., and Wheeler, J. C. 1978, *Ap. J.*, **222**, 667.
Shlosman, I. 1988, in *Supermassive Black Holes, Proc. 3rd George Mason Workshop*, ed. M. Kafatos (Cambridge: Cambridge University Press), p. 231.
Shlosman, I., and Begelman, M. C. 1987, *Nature*, **329**, 810 (Paper I).
Shlosman, I., Frank, J., and Begelman, M. C. 1989, *Nature*, **338**, 45.
Shull, J. M., and Woods, D. T. 1985, *Ap. J.*, **288**, 50.
Silk, J. 1980, in *Advanced Course of Swiss Society Astr. Ap., Star Formation*, ed. A. Maeder and L. Martinet (Sauverny: Geneva Observatory).
———. 1985, *Ap. J. (Letters)*, **292**, L71.
———. 1987, in *IAU Symposium 115, Star Forming Regions*, ed. M. Peimbert and J. Jugaku (Dordrecht: Reidel), p. 663.
Spitzer, L. P., Jr. 1978, in *Physical Processes in the Interstellar Medium* (New York: Wiley).
Takahashi, T., Hollenbach, D. J., and Silk, J. 1983, *Ap. J.*, **275**, 175.
Toomre, A. 1964, *Ap. J.*, **139**, 1217.
Trimble, V. 1982, *Rev. Mod. Phys.*, **54**, 1183.
Young, P. J., Shields, G. A., and Wheeler, J. C. 1977, *Ap. J.*, **212**, 367.
Weaver, R., McCray, R., Castor, J., Shapiro, P., and Moore, R. 1977, *Ap. J.*, **218**, 377.
Wilking, J. A., and Lada, C. J. 1983, *Ap. J.*, **274**, 698.

MITCHELL C. BEGELMAN: Joint Institute for Laboratory Astrophysics, University of Colorado, Boulder, CO 80309-0440

ISAAC SHLOSMAN: Theoretical Astrophysics 130–33, California Institute of Technology, Pasadena, CA 91125