EVOLUTION OF AN ACCRETION DISK IN AN ACTIVE GALACTIC NUCLEUS

ANETA SIEMIGINOWSKA, 1,2 BOŻENA CZERNY,2 AND VADIM KOSTYUNIN 3 Received 1995 April 17; accepted 1995 August 22

ABSTRACT

We present results of the modeling of thermal-viscous ionization instabilities in accretion disks in active galactic nuclei (AGNs). This instability, analogous to the instability which drives outbursts in cataclysmic variables or X-ray transients, is expected also to occur in accretion disks around supermassive black holes unless self-gravity effects or irradiation prevents the existence of the partial ionization zone (which causes the instability). We assume that viscosity scales with the gas pressure (β -viscosity). We discuss in detail evolution of the accretion disk around an $M_{\rm bh}=10^8~M_{\odot}$ black hole mass, assuming a range of accretion rates corresponding to luminosities between 10^{-5} and 1, in Eddington units.

We show that the ionization instability does operate in such disks. Depending on the viscosity, the instability can develop in a very narrow unstable zone, thus causing a flickering, or can propagate over the entire disk, resulting in the large-amplitude outbursts on timescales of order 10^3-10^5 yr. As a consequence: (1) AGNs are not always in their active state, and (2) the emitted optical-UV spectra are quite different from the spectrum of a thin, stationary accretion disk. External irradiation is shown to have a large effect on the instability and on the emitted optical-UV spectra. We briefly consider observational consequences.

Subject headings: accretion, accretion disks — galaxies: active — quasars: general

1. INTRODUCTION

The accretion process that is the leading model for powering quasars (e.g., Rees 1984) is frequently described by the model of a stationary thin disk (Lynden-Bell 1969; Shakura & Sunyaev 1973). However, there are both observational and theoretical arguments indicating that the time-dependent effects are of extreme importance.

Observationally, the evidence for global evolutionary effects is compelling in a number of accretion disks around Galactic X-ray sources. Outbursts (by factors greater than 10⁴) of cataclysmic variables (CVs) or X-ray novae last for weeks or months and happen with frequencies ranging from every few months up to a number of years. The outbursts are essentially caused by the disk thermal instability in the partial ionization zone (Meyer & Meyer-Hoffmeister 1982; Smak 1982; see also Cannizzo 1993b and Mineshige & Kusunose 1993 for a recent review), although some secondary effects due to variability in the outflow rate from the companion star are also observed (e.g., the secondary maximum in X-ray nova outbursts; Chen, Livio, & Gehrels 1993). The conclusion from the study of transient Galactic sources is that neither the gas supply rate (which determines the instantaneous accretion rate) nor a stationary disk model can account for the observed disk properties.

There is a strong similarity between Galactic X-ray sources and AGNs, both in spectral behavior and in overall variability (Fiore & Elvis 1994; Tanaka & Lewin 1995), which leads us to expect similar accretion disk behavior in AGNs. However, because the characteristic timescales scale roughly with the central mass, the expected variability takes thousands to millions of years in AGNs. This is necessarily much longer than the observed variability timescales, and so the time-dependent

¹ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, M/S 4, Cambridge, MA 02138.

² Copernicus Astronomical Center, Bartycka 18, 00-716 Warsaw, Poland.

disk models have not been widely applied to AGNs. However, such models need to be investigated, since the emitted spectrum predicted by these models can differ strongly from that of stationary disk models. Moreover, physical models for the observed evolution of the quasar luminosity function (e.g., Cavaliere & Padovani 1988) will be affected if quasars undergo recurrent strong outbursts.

Theoretically, accretion disks around the massive black holes in AGNs are expected to have a partial ionization zone, as in Galactic binaries, and therefore to be subject to the same instability. This zone forms at a distance of a few hundred Schwarzschild radii ($\sim 0.01~M_8$ pc) from the center (Lin & Shields 1986; Clarke 1987) unless self-gravity effects or irradiation of the outer parts prevents its existence.

Current models of time evolution of accretion disks in AGNs have confirmed the presence of disk eruptions, but they appeared to be of minor amplitude (Clarke & Shields 1989). In these models the viscosity parameter α (Shakura & Sunyaev 1973) for hot and cold gas was assumed to be the same. However, stronger outbursts can be produced if the viscosity of the cold gas is considerably lower than the viscosity of the hot gas (Mineshige & Shields 1990), as seems to be the case for CV disks (Smak 1984a, b).

The ratio of the hot and cold state viscosity parameter (α_{hot} , α_{cold}) between these two states depends on the source and may be as high as 5 (e.g., in SS Cyg; Cannizzo 1993a). Therefore, the use of a variable viscosity in AGN disk models seems to be worth exploring.

The models of accretion disks in AGNs studied by Mineshige & Shields (1990) did not cover the range of luminosities appropriate for bright Seyfert galaxies and quasars because of numerical difficulties (Mineshige 1990). The parameters of these models were set for low-luminosity objects with a small central black hole mass. The case of a massive black hole (10⁸ M_{\odot}) was considered only for a very low accretion rate ($L/L_{\rm Edd} \sim 10^{-4}$ to 10^{-5}). Also the effects of irradiation were completely neglected, although irradiation strongly modifies the predictions. This can be seen from stationary models, as

³ Institute of Astronomy, Russian Academy of Sciences, 48 Pyatnitskaya Street, Moscow 109017, Russia.

well as from the time evolution of a single disk zone (Tuchman, Mineshige, & Wheeler 1990; Mineshige, Tuchman, & Wheeler 1990).

In this paper we use the approximate but very efficient method of Smak (1984a) to calculate the evolution of disks for the relevant range of parameters. We study a broad range of external accretion rates corresponding to luminosities ranging from values as low as 10^{-5} up to 1, in Eddington units. The lowest values are characteristic of starving disks, like the one postulated at the center of our Galaxy (Falcke & Heinrich 1994), while higher values are appropriate for luminous guasars. We confine ourselves to β -disks, i.e., models in which the viscous torque is proportional only to the gas pressure (as in Clarke & Shields 1989 and Mineshige & Shields 1990). This avoids the Lightman-Eardley (1974) instability in the inner, radiation-dominated parts of α-disks (i.e., models in which the viscous torque is proportional to the total [gas plus radiation] pressure). This instability has not been observationally confirmed, nor is it theoretically understood. However, if it is present, it certainly operates on much shorter timescales. Luminous AGNs are known to vary on short timescales, but there is no clear connection between this variability and the Lightman-Eardley instability. In this paper we will consider only the ionization instability, which operates on timescales of $\sim 10^5 \text{ yr.}$

We show that the ionization instability does operate in accretion disks surrounding supermassive black holes. Depending on the viscosity, the instability can develop in a very narrow unstable zone, or can propagate over the entire disk with the following effects: (1) large-amplitude outbursts can result (greater than 10⁴); as a consequence (2) accretion disks around supermassive black holes at the center of AGNs are not always in their active state, so many low-luminosity AGNs may contain an accretion disk in quiescence; and (3) the emitted optical-UV spectra are quite different from the spectrum of a thin stationary accretion disk. Finally, external irradiation is shown to have a large effect on the instability and on the emitted optical-UV spectra. We briefly consider some further observational consequences.

We discuss the local disk structure and the scaling laws in § 2; § 3 contains the details of the method; and in § 4 we present our results. We discuss irradiation effects in § 5. Section 6 contains a discussion of some observational consequences.

2. LOCAL STRUCTURE OF AN ACCRETION DISK AROUND A MASSIVE BLACK HOLE

2.1. Scaling Law for the Σ - T_{eff} S-Curve

We consider a geometrically thin Keplerian accretion disk around a supermassive black hole. The local structure for such disks (assuming that viscosity is proportional to gas pressure) is described by standard equations (e.g., Frank, King, & Raine 1992). We integrate these equations at a given radius using a procedure described by Pojmański (1984, 1986). We included convection in our vertical structure calculations; however, we do not consider vertical advection (Kley, Papaloizou, & Lin 1993). We assumed that the density at the surface of the disk is low $[\rho(H) = 10^{-15} \text{ g cm}^{-3}]$. Opacities were taken from Cox & Stewart (1970) for log $T_{\rm eff} > 4.0$ and from Alexander (1975) for lower temperatures.

There is a local equilibrium relationship between a steady state accretion rate \dot{M} (or effective temperature $T_{\rm eff}$) and a surface density Σ for a given central black hole mass $M_{\rm bh}$,

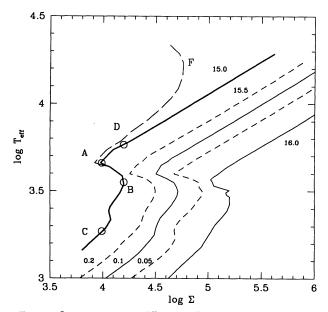


Fig. 1.—S-curves (Σ - $T_{\rm eff}$) at different radii (log r=15.0, 15.5, and 16.0 cm) and $\alpha=0.1$ are plotted with solid lines. Short-dashed lines indicate the S-curves at log r=15.5 and different $\alpha(0.2, 0.1, 0.05)$. Central black hole mass $M_{\rm bh}=10^8~M_{\odot}$ and $\tau_{r\phi}=\alpha P_{\rm gas}$. The critical points A and B are indicated for minimum and maximum surface densities. Points C and D represent the maximum and minimum effective temperatures reached by the transitions between two stable states when the surface density remains constant. The long-dashed line shows the S-curve when $\tau_{r\phi}=\alpha P_{\rm total}$.

viscosity parameter α , and distance r from the center. This relation has a characteristic S-curve shape for an optically thick, geometrically thin accretion disk. Three characteristic regions on the S-curve describe different physical conditions in the disk.4 The lower branch is thermally stable; cool and neutral hydrogen dominates the chemical composition of the gas, and molecules strongly contribute to the opacities. A disk on the upper branch is also thermally stable but is hot, and the hydrogen is fully ionized; bound-free transitions in heavy metals as well as free-free transitions and electron scattering generate the opacity. The middle branch corresponds to a partially ionized disk, which is thermally unstable due to the rapid increase of opacity with temperature (Bath & Pringle 1982; Faulkner, Lin, & Papaloizou 1983). Small perturbations in the cooling-heating balance cause cooling or heating of the disk until a stable condition on the upper or lower branch is achieved. This thermally unstable situation has been analyzed by many authors for cataclysmic variables, low-mass X-raybright variables, and soft X-ray transients (e.g., Huang & Wheeler 1989; Mineshige & Wheeler 1989), and in less detail also for AGNs (Lin & Shields 1986; Clarke 1987; Clarke & Shields 1989; Mineshige & Shields 1990).

We can calculate a local vertical structure of the disk at every radius and so derive the S-curve for each location. The curves in Figure 1 show the local disk structure (Σ vs. $T_{\rm eff}$) at log r(cm) equal to 15.0, 15.5, and 16.0 for a black hole mass of $10^8~M_{\odot}$. The dependence on the viscosity parameter α is also shown.

Smak (1984a) parameterized disk solutions for dwarf novae using a standard S-curve (Σ - $T_{\rm eff}$) and the position of the point

⁴ For models with $\tau_{r\phi} = \alpha P_{\text{total}}$ there exists a fourth, thermally unstable, region where $P_{\text{rad}} \gg P_{\text{gas}}$.

on the curve having the minimum surface density (point A in Fig. 2). Siemiginowska (1991) adopted this parameterization for accretion disks around supermassive black holes:

$$\log T_{\text{eff}}^{r} = \log T_{\text{eff}} - 0.12 \log r_{15} + 0.03 \log \alpha_{0.1} , \qquad (1)$$

$$\log \Sigma = f(\log T_{\text{eff}}^r) + 1.05 \log r_{15} - 0.69 \log \alpha_{0.1}, \quad (2)$$

where $\alpha_{0.1} = \alpha/0.1$, $r_{15} = r/10^{15}$, and $T_{\rm eff}^r =$ effective temperature reduced to log r = 15 and $\alpha = 0.1$. The coefficients were determined numerically.

We have estimated the accuracy of this parameterization by means of a comparison of the surface density profile calculated from the vertical structure equations with the profile calculated by scaling. An example is shown in Figure 2 (for particular values of the black hole mass and accretion rate). As expected, the parameterization works well ($\gtrsim 5\%$) in the middle parts of the disk where there is no influence of the inner boundary condition and the dependence of all the parameters on radius are monotonic and well approximated by power laws. Close to the marginally stable orbit the representation of the disk structure by scaling laws is accurate only to $\gtrsim 10\%$. Departures from an exact scaling are also visible (at $\sim 20\%$) in the outer parts of the disk, far from the radius used to define the S-curve (see Fig. 2).

2.2. Radial Extent of the Instability Strip

The thermally stable upper branch of the S-curve ("hot" phase) is characterized by an accretion rate higher than the external stationary value, while the stable lower branch ("cold" phase) has a lower than stationary accretion rate, and so accumulates gas. For a given external accretion rate there is usually a range of radii where neither of these stationary solutions is stable and in which the cyclic instability is active.

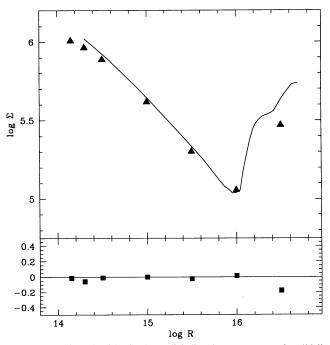


Fig. 2.—Surface densities in the radial direction. *Upper panel*: solid line represents the surface densities calculated using the parameterization (eqs. [1] and [2]); filled triangles indicate the surface density calculated directly from the standard vertical structure equations. *Lower panel*: filled squares show differences in surface density values calculated using the parameterization and the direct vertical structure models.

We can estimate the size and location of the unstable zone by finding inner and outer radii, r_A and r_B , of the zone. Using the parameterization of the radius versus effective temperature relation (eqs. [1] and [2]), we find the radius r_A described by the local values of Σ_A and T_A :

$$r_A \sim 1.6 \times 10^{16} M_8^{0.4} \dot{M}^{0.4} \alpha_{0.1}^{-0.05} \text{ cm}$$
, (3)

where $M_8 = M_{\rm bh}/10^8~M_{\odot}$, and \dot{M} is the accretion rate in M_{\odot} yr⁻¹. The temperature T_A on the standard S-curve is $\log T_A = 3.64$. The scaling relations then give us the radius for any (stationary) effective temperature T:

$$\log (r/r_A) \approx -1.4 \log (T/T_A). \tag{4}$$

This expression gives the radius of the turning point B (Fig. 1, $\log T_B = 3.54$) to be $r_B \sim 1.4 r_A$, similar to the value given by Lin & Shields (1986).

However, the entire region between points C and D on the S-curve is affected by the limit cycle to some extent. A better estimate of the position and the width of the instability strip is given by r_D and r_C (see Fig. 1). On the standard curve (log $T_C = 3.23$, log $T_D = 3.76$) these are

$$r_D \sim 0.68 r_A$$
 and $r_C \sim 5.5 r_D \sim 3.8 r_A$. (5)

The location of the instability zone depends on the accretion rate (eq. [3]). The unstable region is located close to the black hole for low accretion rates and far away when the accretion rates are high. Since the bulk of the energy release in the disk is at small radii, we expect much stronger luminosity variations when the accretion rates are low. The locations of the inner and outer edges of the instability strip for the assumed model parameters are given in Table 1.

The position and the extent of this region is also sensitive to the adopted shape of the standard S-curve (Mineshige & Shields 1990). In particular, if the viscosity parameter for the cold phase $\alpha_{\rm cold}=0.025$ (one-fourth that for the hot phase), then the temperatures T_C and T_D become more widely separated (Fig. 6, log $T_D=3.89$, log $T_C=3.01$), and the coefficient in equation (3) becomes equal to 0.41 close to the inner radius of the disk (see Table 1). In such models the instability develops over the entire disk, instead of being constrained to an instability belt, as for a constant viscosity parameter and moderate accretion rates.

The characteristic temperatures also give a rough estimate of the ratio, τ , of the time spent by the disk in the high state, t_{high} ,

TABLE 1 Models

Model	\dot{M} $(M_{\odot} \text{ yr}^{-1})$	$L/L_{ m Edd}$	$\alpha_{\rm cold}$	$r_D \ (r_g)$	$r_C \choose (r_g)$
A0	3×10^{-5}	10-5	0.1	6.06	61.7
A1	3×10^{-5}	10 ⁻⁵	0.1	4.39	106
B0	0.01	0.0034	0.1	90.9	665
B 1	0.01	0.0034	0.075	81.1	1120
B2	0.01	0.0034	0.050	55.1	1120
В3	0.01	0.0034	0.025	59.5	1500
C0	0.1	0.034	0.1	234	1670
C1	0.1	0.034	0.075	209	2800
C2	0.1	0.034	0.050	143	2800
C3	0.1	0.034	0.025	155	3760
D0	1.0	0.34	0.1	385	2730
D1	1.0	0.34	0.075	344	4570
D3	1.0	0.34	0.025	255	6130
E0	3.4	1.0	0.1	973	6850
E1	3.4	1.0	0.075	870	7930

relative to the time spent in the low state (quiescent), t_{low} ; τ is approximated by the ratio of the blackbody cooling times:

$$\tau = \frac{t_{\text{high}}}{t_{\text{low}}} \approx \left(\frac{T_B}{T_A}\right)^4. \tag{6}$$

This gives $\tau = 0.4$ in the constant- α model (with $\alpha = 0.1$), but $\tau = 0.03$ in the model with $\alpha_{cold} = 0.025$. So when $\alpha_{cold} = 0.025$ and $\alpha_{hot} = 0.1$, the disk remains in the outburst phase only for a few percent of the time. This has significant observational consequences, which we discuss in § 6. The accuracy of this estimate is good only to a factor of a few, but it shows a clear trend in the duty cycle and allows us to estimate the basic properties of the solution.

2.3. Timescales

The surface density and the local accretion rate evolve on the viscous timescale (Lightman 1974) $t_{\text{visc}} \sim r/v_r$ (where v_r is the radial velocity). For stationary α -disk models with parameters scaled to AGN values.

$$t_{\rm visc} \sim 2 \times 10^5 \alpha_{0.1}^{-0.8} \dot{M}^{-0.3} M_8^{0.25} r_{16}^{1.25} f^{1.2} \text{ yr},$$
 (7)

where $r_{16} = r/10^{16}$ cm, $f = 1 - (3r_g/r)^{0.5}$. Thus the timescale on which the global luminosity variations caused by the "ionization instability" can occur is $\sim 10^5$ yr for a 10^8 M_{\odot} central black hole.

The local thermal timescale is described by $t_{\rm th} \sim 1/\alpha\Omega_{\rm K}$, and in our case it is

$$t_{\rm th} \sim 2.7 \alpha_{0.1}^{-1} M_8^{-0.5} r_{16}^{1.5} \text{ yr}$$
 (8)

So

$$t_{\rm visc}/t_{\rm th} \sim 10^5 \alpha_{0.1}^{0.2} \dot{M}^{-0.3} M_8^{0.75} r_{1.6}^{-0.25} f^{1.2}$$
 (9)

for standard accretion disk models (e.g., Frank et al. 1992). In CVs the ratio $t_{\rm visc}/t_{\rm th}\sim 10^2$, 10^3 times smaller than for AGNs. This difference has important consequences, since changes in AGN disks which happen on the thermal timescales will not then propagate over the entire disk as fast as in Galactic sources. We can expect to observe variations on the thermal timescales in a quasar, while it is impossible to observe global disk evolution on the viscous timescale. Instead the global evolution of a nuclear disk must be studied through statistical departures from stationary accretion in a population of quasars.

3. CONSTRUCTION OF TIME-DEPENDENT SOLUTIONS

3.1. Local Disk Structure: Time Evolution under Thermal Instability

When solutions for the local vertical structure are defined by points on the local S-curve (Fig. 1) the disk is locally in thermal equilibrium. We assume that solutions which are not in thermal equilibrium are also described by the standard disk equations (e.g., Frank et al. 1992), but with an extra term (Smak 1984a):

$$\frac{dF_0}{dz} = \frac{dF_\eta}{dz} + \frac{dF_{\rm th}}{dz} \,, \tag{10}$$

where F_0 is the total radiation flux; F_n is the standard energy flux generated by viscous forces;

$$\frac{dF_{\eta}}{dz} = \frac{3}{2} \alpha \Omega_{K} P_{gas} ; \qquad (11)$$

and F_{th} describes the extra flux due to vertical expansion or contraction of the disk.

In the simple case of a perfect gas and homologous expansion and contraction the second term in equation (10) can be transformed to

$$\frac{dF_{\rm th}}{dz} = -\left(\frac{3}{2}\frac{d\ln T_e}{dt} + \frac{d\ln H}{dt}\right)P_{\rm gas}, \qquad (12)$$

where T_a is the temperature in the equatorial plane and H is the disk thickness. Therefore it can be represented as

$$\frac{dF_{\rm th}}{dz} = \frac{3}{2} \alpha_{\rm th} \Omega_{\rm K} P_{\rm gas} \tag{13}$$

(Smak 1984a), where α_{th} describes thermal effects ($\alpha_{th} < 0$ in the heating case and $\alpha_{th} > 0$ for cooling). The parameter α_{th} is a convenient parameterization of deviations from thermal equilibrium such that models with $\alpha_{th} = 0$ are in thermal equilibrium and others are out of thermal equilibrium. This parameterization allows us to define the quantity α_{eff} = $\alpha + \alpha_{th}$. The total flux from the disk surface can then be defined by

$$F_0 = \sigma T_{\text{eff}}^4 = F_n + F_{\text{th}} \,, \tag{14}$$

where

$$F_{\eta} = \frac{\alpha}{\alpha_{\text{eff}}} F_0 , \qquad (15)$$

$$F_{\rm th} = \frac{\alpha_{\rm th}}{\alpha_{\rm rs}} F_0 \ . \tag{16}$$

3.2. The Temperature Equation

The partial flux resulting from contraction plus cooling or expansion plus heating, can be related to the total energy changes (including the internal energy and the potential energy which depends on the height of the disk) using

$$-\frac{2F_{\rm th}}{\Sigma} = \frac{dE}{dt} = \frac{\partial E}{\partial \Sigma} \frac{d\Sigma}{dt} + \frac{\partial E}{\partial T_{\rm eff}} \frac{dT_{\rm eff}}{dt} , \qquad (17)$$

where $E = E(\Sigma, T_{eff}) = E(T_{eff}^r)$, where $T_{eff}^r(\Sigma, T_{eff})$ is given by equation (1). Equation (17) can be rewritten as

$$\frac{\partial \ln T_{\rm eff}}{\partial t} = -\left(\frac{2F_{\rm th}}{E} + \frac{\partial \ln E}{\partial \ln \Sigma} \frac{\partial \Sigma}{\partial t}\right) / \left(\Sigma \frac{\partial \ln E}{\partial \ln T_{\rm eff}}\right) - v_r \frac{\partial \ln T_{\rm eff}}{\partial r} . \tag{18}$$

Cooling by advection (in which energy is carried inward with the accreting gas as internal energy) was not included in our calculations, although it may be important when the radial temperature gradient is very steep. The importance of the advection was analyzed by Abramowicz et al. (1988) for high accretion rates [also recently by Narayan & Yi (1994, 1995) for very low accretion rates] with viscosity scaled by total pressure in an accretion disk around a supermassive black hole. We can expect that at the boundary of the unstable region the temperature gradient will be quite large, and that a radial cooling flow may occur. This effect may help to smooth out the discontinuity at the edge of the unstable region, and may affect the timescale on which the instability travels through the disk.

3.3. The Surface Density Equation and the Boundary Conditions

In the case of Keplerian motion the surface density evolution may be written (Pringle 1981; Smak 1984a)

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{r} \frac{\partial}{\partial r} \left[r^{1/2} \frac{\partial}{\partial r} (v \Sigma r^{1/2}) \right] + \frac{1}{2\pi r} \frac{\partial \dot{M}}{\partial r} + \frac{1}{\pi r} \frac{\partial}{\partial r} \left[(r^{1/2} - r_{\text{out}}^{1/2}) \frac{\partial \dot{M}}{\partial r} \right], \tag{19}$$

where $r_{\rm out}$ is the outer radius of the disk and ν is the kinematic viscosity coefficient given in α -disks by $\nu \Sigma = 8F_{\eta} \Omega_{\rm K}^{-2}/9$.

The local accretion rate is given by

$$\dot{M}(r) = 6\pi r^{1/2} \frac{\partial}{\partial r} \left(r^{1/2} v \Sigma \right). \tag{20}$$

Equation (19) describes a disk which is continuously supplied not only with matter but also with angular momentum. Only a fraction of the angular momentum is transported to the central object (black hole) because the angular momentum of the matter at the marginally stable orbit is approximately equal to $l_0 = (GMr_{\rm in})^{1/2}$. The inner edge of the disk, $r_{\rm in}$, is situated at $3r_g$, and below this the matter is freely falling onto the black hole.

In the case of binary systems, the excess of angular momentum is subtracted from the disk by the tidal forces due to the presence of the companion. However, in AGN accretion disks the supply mechanism is unknown. It seems most probable that global dynamical instabilities are responsible for transporting the gas available from distances of a few parsecs or kiloparsecs to the nucleus, but there are no actual predictions for the resulting conditions at the outer edge of the disk. Therefore we simply postulate that the excess of the angular momentum is subtracted at the outer radius in such a way as to produce stationarity of the disk outside the instability strip. The amount of angular momentum "outflow" can be determined from the global balance between the inflow rate $(\dot{M}l_{\rm out})$ and the storage rate in the stationary case $(\dot{M}l_0)$.

If we incorporate this stationarity requirement at the outer edge, equation (19) is changed into

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{r} \frac{\partial}{\partial r} \left[r^{1/2} \frac{\partial}{\partial r} \left(v \Sigma r^{1/2} \right) \right] + \frac{1}{2\pi r} \frac{\partial \dot{M}}{\partial r} + \frac{1}{\pi r} \frac{\partial}{\partial r} \left[\left(r^{1/2} - r_{\rm in}^{1/2} \right) \frac{\partial \dot{M}}{\partial r} \right]. \tag{21}$$

To solve the equations conveniently, we use variables $x = 2r^{1/2}$ and $S = x\Sigma$ (Bath & Pringle 1981).

3.4. Method

Two equations determine the time evolution of an accretion disk: equations (18) and (21) give directly the surface density and the effective temperature, respectively. Scaling law (eqs. [1] and [2]) determines the new viscosity coefficient $\alpha_{\rm eff}$, which allows us to calculate the viscous flux (eqs. [14] and [15]) and the departure of the model from thermal equilibrium (eq. [16]). The local accretion rate is given by equation (20), and the radial velocity results from the standard continuity equation.

Using these equations, we have developed a numerical code based on that of Smak (1984a) and have applied it to the case of accretion disks around supermassive black holes. The disk was divided into N (45–100) annuli in the radial direction. For each annulus the S-curve was defined by the parameterization of the standard S-curve (eqs. [1] and [2]). This parameterization allows the code to run on widely available computers and still complete a disk structure calculation in a short time (usually about 1 hr on IBM 58H or several hours on a SPARC IPX station). Mineshige & Shields (1990) calculated their model using detailed vertical structure calculations at each step separately, which is a more accurate method. However, this method requires the use of supercomputers, takes more time, and for some parameters, especially large central masses $(10^8\ M_{\odot})$, leads to numerical problems (Mineshige 1990).

4. RESULTS OF THE EVOLUTIONARY CALCULATIONS

Two different kind of models can be distinguished for the assumed character of the viscosity parameter. The first type has constant parameter α , while the second type has the viscosity dependent on the phase of the disk, and is different at the hot and cold phases. We always assume that $\alpha_{\rm hot}=0.1$ and vary only $\alpha_{\rm cold}$. All computations were run for a central black hole mass of $10^8~M_{\odot}$. Table 1 contains the values of accretion rates and $\alpha_{\rm cold}$ for sets of models we have calculated. The boundaries of the unstable zone (r_C, r_D) are also shown for each model. We discuss the two classes of models in the next sections.

4.1. Models with Constant Viscosity Parameter ($\alpha = 0.1$)

This family of models shows only relatively weak variations of the total luminosity, as was noticed for low accretion rates by Mineshige & Shields (1990). We discuss results of our calculations for the model CO. We also discuss other cases briefly. We compare the two models with the lowest accretion in Table 1 (models A0 and A1) directly with the Mineshige & Shields models

The luminosity variations due to thermal and viscous instabilities developing in the accretion disk (model C0: $\dot{M}=0.1$ M_{\odot} yr⁻¹, and constant viscosity parameter $\alpha=0.1$) are shown in Figure 3a. The luminosity variations are very small (± 0.1 in log L). They are not regular and show some clear substructure. Effectively, the changes in the luminosity happen on timescales of order of 10^4 yr, shorter than estimated by equation (9), although the major outbursts take rather 10^5 yr as predicted.

The instability does not propagate strongly beyond the unstable region. This is illustrated in Figure 4c by variations in the local accretion rate. Major changes can be seen inside the instability strip, and only small fluctuations occur outside that region. This gives only small variability in the effective accretion rate and in the bolometric luminosity.

The surface density and effective temperature for the hot and cold phases are plotted for model C0 in Figures 4a and 4b. The radial surface density distribution in the high state generally evolves toward lower values thus approaching the limiting values for the onset of the cooling front given by $\log \Sigma_A(r)$ in Figure 4a. Because the distribution is not a monotonic function of a radius, the region close to the minimum becomes unstable first. The cooling front starts to propagate in both directions. In the meantime mass accumulates in the disk, and the surface density increases until it reaches the limit given by Σ_B [a straight line parallel to $\log \Sigma_A(r)$ but shifted up]. The outbursts start at the radius at which the surface density is equal to the critical local Σ_B (see Fig. 1). This may happen before the cooling front completes its travel through the entire

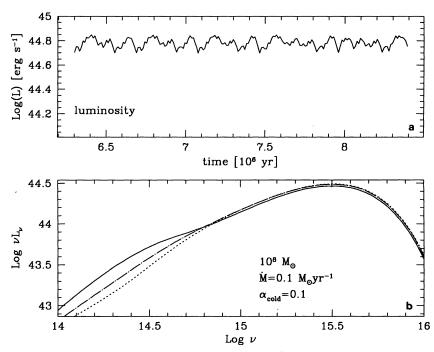


Fig. 3.—(a) Luminosity variations for the accretion disk around a $10^8~M_{\odot}$ black hole with $\dot{M}=0.1~M_{\odot}~\rm yr^{-1}$, $\alpha=0.1~(\alpha_{\rm cold}=\alpha_{\rm hol})$. (b) Spectrum in the high state (solid line) and low state (dotted line) of the disk. A separation in time between the two states is equal to $2\times10^4~\rm yr$.

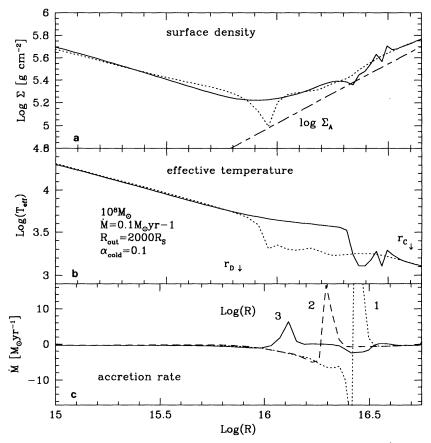


Fig. 4.—High and low states of an accretion disk around a black hole of $M_{\rm bh}=10^8~M_{\odot}$, with accretion rate $\dot{M}=0.1~M_{\odot}~{\rm yr}^{-1}$ and parameter $\alpha=0.1~(\alpha_{\rm cold}=\alpha_{\rm hol})$. A separation in time between the two states is $2\times10^4~{\rm yr}$. Panels show (a) surface density, (b) effective temperature, and (c) evolution of a local accretion rate. The major changes occur in the unstable zone. The arrows indicate the boundaries of the instability strip.

unstable region. Therefore, the disk never actually reaches the quiescent state. This is reflected in the shape of the light curve, which shows the flickering.

The disk eruptions start somewhere in the middle of the instability strip and show significant irregularity. This is related to the fact that the value of Σ_A is very close to the Σ_B value. As a result, the two timescales—the viscous timescale (for stationary accretion) at the outer edge of the instability strip and the viscous timescale (in the cold state) at the inner edge of the strip—are almost equal. The first timescale determines a supply of the mass, and the second one the buildup of the mass in the disk. If either one is considerably shorter than the other, then the outbursts always start either at the inner or at the outer edge of the instability strip, while if the two are similar the outburst starts in the middle of the instability strip.

The location of the unstable region for a given central mass depends on the outer accretion rate (eq. [3]). For very low accretion rates the unstable region is situated closer to the center and influences the regions of strong gravitational potential and significant energy generation. In Figures 5a and 5b we show the luminosity and local accretion rate variations for model A0 (which has an accretion rate $\sim 1/3000$ that of model C0). The inner edge of the instability strip is now situated very close to the inner edge of the disk, and the entire part of the disk is strongly affected by the instability. This is reflected in the amplitude of variations ($\Delta \log L = 0.9$). This model corre-

sponds to the model F0 of Mineshige & Shields. The amplitudes of the luminosity variations are of the same order, and the timescales compare closely. However, light curves do not follow each other closely, since the scaling approach to the $\Sigma(r)$ distribution in our models predicts the flickering mode even close to the marginally stable orbit. These models are not clearly applicable to quasars, because the average bolometric luminosity is at the extreme low end of the luminosity function of observed AGNs (see Tables 2 and 3). There is also a theoretically based argument that such low accretion rates may produce instead an ion torus (e.g., Begelman 1985) rather than a geometrically thin disk.

4.1.1. Spectra

The amplitude of luminosity variations is not significant for models with constant viscosity parameter. However, as the changes in the structure of the outer cold parts of a disk are considerable, they are reflected in significant modifications of the disk spectrum in the IR-optical band. We calculate these spectra assuming local blackbody emission. The spectral energy distribution for the two states, hot and cold, for model C0 is shown in Figure 3b. In the hot state a bump in the spectrum can be seen due to the additional emission from the regions in the high state. However, the UV luminosity as well as the bolometric luminosity may be higher in the low state. The inner regions which contribute the most to the UV emis-

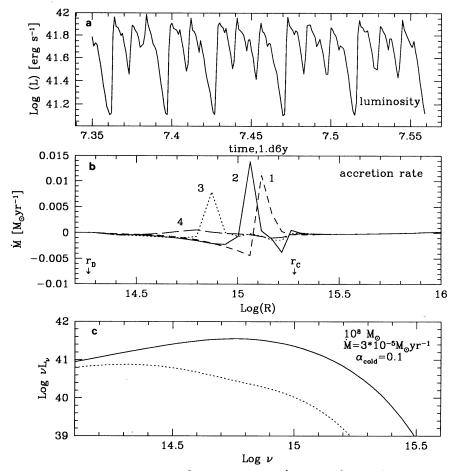


Fig. 5.—(a) Luminosity variations for the accretion disk around a $10^8~M_\odot$ black hole with $\dot{M}=3\times10^{-5}~M_\odot~yr^{-1}$, $\alpha=0.1~(\alpha_{cold}=\alpha_{hot})$. (b) Evolution of a local accretion rate (time separation $2\times10^3~yr$). (c) Spectrum in the high state (solid line) and low state (dotted line) of the disk.

TABLE 2
LUMINOSITY AND SPECTRAL INDEX

Model	$\langle L \rangle$	$\Delta L/L$	Duty Cycle ^a	α_{med}^{b}	α_{mean}	σ	α_{high}	α_{low}	α_{high}^{irr}	α_{low}^{irr}
			<u> </u>	α_{cold}	= 0.1					
A0	41.70	0.41	55.5	-1.05	-1.41	1.15	-1.33		-1.25	
B 0	43.88	0.09	99.9	0.02	-0.05	0.24	-0.45	0.05	-0.65	-0.27
C0	44.78	0.07	99.9	0.15	0.11	0.18	-0.08	0.23	-0.57	-0.34
D0	45.73	0.08	100	0.18	0.16	0.14	-0.12	0.34	-0.63	-0.44
E0	46.2	0.08	100	0.23	0.22	0.13	0.01	0.43	-0.66	-0.55
				α _{cold} =	= 0.075					
A1	41.78	1.18	22.0	-0.11	-0.12	0.03	-1.67		-1.65	
B1	44.11	0.80	30.1	0.23	0.14	0.26	0.08	0.16	-0.44	0.01
C1	44.75	0.71	29.8	0.27	0.17	0.32	0.25	0.28	-0.49	0.06
D1	45.28	0.81	27.0	0.31	0.15	0.42	0.21	0.33	-0.57	0.01
E1	46.18	0.87	20.6	0.36	0.18	0.41	-0.70	0.40	-0.76	0.19
				$\alpha_{\rm cold}$	= 0.05					
B2	44.15	0.63	18.1	0.14	-0.06	0.49	-0.14	0.20	-0.53	-0.08
C2	44.82	0.75	18.8	0.24	0.09	0.44	0.14	0.07	-0.55	-0.21
D2	45.19	1.01	11.0	0.28	0.19	0.29	-0.12	0.11	-0.64	-0.21
				α _{cold} =	= 0.025					
В3	43.88	2.63	6.0	0.31°	0.13°	0.54°	-0.28		-0.58	
C2	44.30	3.01	3.1	0.45°	0.09°	0.86°	-0.15		-0.63	
D2	45.21	3.98	3.0	0.51°	0.05°	0.84^{c}	-1.15		-0.7	

^a A fraction of time when the luminosity is larger than the half of the maximum value (in %).

sion are not affected strongly by the instability in the unstable zone except in model A0 (with the lowest accretion rate; see Fig. 5c). The changes in the local accretion rates in the inner regions are seen a long time after the unstable zone has gone through the hot state (the viscous timescale governs the evolution of the inner stable parts, while in the unstable zone the characteristic timescale is directly related to the instabilities continuously developing there). Spectral changes in the optical and UV are therefore essentially independent and uncorrelated.

A good representation of the output spectrum is the mean spectral index $(F_{\nu} \sim \nu^{2})$ in the frequency range $10^{14.5}$ – $10^{15.0}$ Hz (3000–9500 Å). This provides a straightforward comparison with the available AGN samples (see Table 3). We selected four independent AGN samples for which spectral slopes are available (Table 3). The spectral slopes, the range of the optical luminosities, and the average luminosity for each sample are given in Table 3.

The mean index calculated from model C0 over the entire cycle is 0.11 ± 0.18 , which fits the observed spectral index of

 -0.2 ± 0.4 in bright quasars (Neugebauer et al. 1987; Elvis et al. 1994; Fiore et al. 1995) better than the 0.33 index given by stationary models. Nevertheless, it is not a good fit. Values for higher accretion rates fare no better (Table 2). Because the instability zone moves farther out as \dot{M} is increased, the spectrum is determined mainly by the steady, hot inner regions, and so the index approaches the standard stationary disk model value. For the model with the lowest accretion rate the spectral index is closer to the observed value. However, as we have already mentioned, for such low accretion rates the luminosities are much too low, and the geometrically thin disk approximation may not be valid.

The range of the observed luminosities (43.5 < log L < 47.0) is higher than the range of averaged luminosities covered by our studies (41.70 < log L < 46.20; see Table 2). Our calculations have been carried out for only one central mass (10⁸ M_{\odot}), so the average luminosity depends only on the accretion rate in our runs, while in general the luminosity also depends on the black hole mass. Adopting a range of masses, we may easily model the entire observed luminosity range.

TABLE 3
OBSERVATIONS

Sample	Spectral Range	α_{mean}	α_{med}	$\Delta \alpha$	$\logL_{\rm min}$	$\logL_{\rm max}$	Reference
Neugebauer 1987	14.5–15.0		-0.2	0.8	43.6	47.1	1
Neugebauer 1987m	14.5-15.0		-0.16		45.0	47.1	1
Francis	14.7-15.3		-0.32	0.2	44.1a	•••	2
Francis	14.7-15.0		-0.02				2
Christiani	14.7-15.4	-0.70			43.5	47.5	3
Elvis	14.5-15.2		-0.1	0.3	44.8	47.0	4

^a There have been 30 objects in the sample of 718 with log L between 39.9 and 44.1.

^b Spectral index defined as $F_{\nu} \sim \nu^{\alpha}$.

^c Calculated adopting the lower limit of 42.00 for the luminosity $\log (\nu L_{\nu})$ at $\log \nu = 14.8$.

REFERENCES FOR SAMPLES.—(1) Bright quasars of Neugebauer et al. 1987 (here Neugebauer 1987) and the same sample constrained to objects brighter than 11.5 in solar units (here Neugebauer 1987m; (2) bright quasars of Francis et al. 1991; (3) Cristiani & Vio 1990; Elvis et al. 1994.

The choice of mass also determines the location of the peak of the spectrum for a given luminosity. If the spectrum peaks near or below the optical-UV band, the calculated UV slopes may become very steep and would not apply to the observed values. We can clearly see that our models brighter than log $L \sim 44$ tend to have spectra rising too strongly in the UV to agree with the data (Table 3). It is possible in principle to obtain more luminous models with a spectral peak in the UV, but only by raising the mass by 2 orders of magnitude or more.

4.1.2. Summary

We conclude that in accretion disks with constant viscosity parameter α around supermassive black holes, instabilities develop inside a well-defined unstable zone. The instability results in periodic changes of the optical-IR tail of the disk spectrum with negligible effect on the UV and on the bolometric luminosity. The amplitude of bolometric luminosity variations is very small. The mean optical-UV spectra generated from these models (assuming local blackbody emission) are significantly steeper than the spectra predicted by stationary models, but are still not steep enough (for all the models except the one with the lowest accretion rate) to match the observed spectral distribution. The situation is very different for two-valued viscosity disks.

4.2. Instability in Accretion Disks with a Two-valued Viscosity Parameter

The origin and the nature of the viscosity in accretion disks are unknown. However, detailed studies of some Galactic systems show that the viscosity must be different in the hot and cold states. For example, in SS Cygni α up to 4–5 times smaller in the cold state than in the high state (Cannizzo 1993a). Similar variable viscosities may well apply to accretion disks in AGNs. This is at least consistent with the limits on α based on the thermal timescale in the radiation-pressure-dominated region (Siemiginowska & Czerny 1989; Czerny & Czerny 1986). Two types of radial relation for the viscosity parameter α have been developed: (1) discrete values of α for the disk in the hot (α_{hot}) and cold (α_{cold}) states (Smak 1984a; Cannizzo 1993a); (2) a continuous change with the effective temperature, $\alpha \sim$ $\alpha_0(h/r)^n$ (Meyer & Meyer-Hoffmeister 1984). Observations of binaries do not clearly discriminate between these two cases. Mineshige & Shields (1990) used the second relation in their models. Since h/r is smaller for accretion disks in AGNs than in CVs, there should be a systematic difference between the viscosity parameters in the two cases. Mineshige & Shields multiplied their equation (2.12) by a factor of 10³ to avoid such differences. We adopt the first prescription in our modeling and assume the ratio of $\alpha_{hot}/\alpha_{cold}$ from the CV studies. The effective viscosities should be the same in both cases.

We discuss models assuming that the viscosity on the lower branch is 0.025, 0.05, or 0.075, while keeping the viscosity coefficient on the upper branch always equal to 0.1. The standard S-curves for these models are different from the $\alpha_{\rm cold} = \alpha_{\rm hot}$ case (Fig. 6). The effective temperatures and surface densities of the unstable region cover an increasing range as the $\alpha_{\rm hot}/\alpha_{\rm cold}$ ratio increases, and the outer radii of the unstable zone are located farther out, relative to the case of constant α .

A ratio between the two extreme values of the surface densities (Σ_A/Σ_B) characterizes the outbursts. By decreasing the value of the viscosity in the low state, this ratio is increased. Smak (1984a, b) noticed that the quiescent state was not reached in any of his models when α was the same in hot and

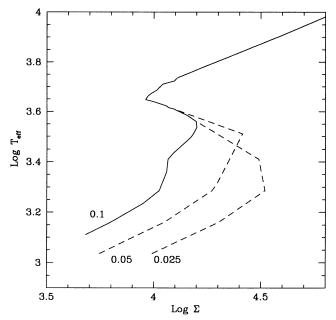


Fig. 6.—Modified S-curves (log r=15.0 cm): $\alpha_{hot}=0.1$ and $\alpha_{cold}=0.05$, 0.025.

cold states. The instability was triggered again before the low state was reached by the disk after the previous outbursts. As a result the heating and cooling transition fronts are always present in such disk. However, a pure quiescent state after an outburst was achieved if the ratio between the two critical surface densities was higher then 2–3 (Cannizzo 1993b). Our models reach quiescence only if $\alpha_{\rm cold} \leq 0.025$ (then $\alpha_{\rm hot}/\alpha_{\rm cold} \geq$ 4), and only in these models do we obtain strong outbursts and global changes in the disk structure corresponding to strong and well-separated outbursts.

In Figure 7 we compare the bolometric luminosity variations when the instability develops in the disk around a 10^8 M_{\odot} black hole and an accretion rate 0.1 M_{\odot} yr⁻¹ (models C0, C2, C3) for the three considered values of $\alpha_{\rm cold}$. Only very small luminosity fluctuations are present when $\alpha_{\rm cold}=0.1$. In the case of $\alpha_{\rm cold}=0.05$ the modulations become stronger (± 0.55 in $\log L$), but for the low state of the disk is never reached. Only in the last case, $\alpha_{\rm cold}=0.025$, the large luminosity variations, well separated in time, are present.

The repeated total luminosity variations accompanying the developing instability are shown in Figure 8 (model B3), The large amplitude ($\Delta \log L \sim 4$) on a timescale of about 10^6 yr can be seen. The period is much longer than in the previous model (§ 4.1), which is related to the wider unstable zone, and to the fact that the instability travels through the whole disk.

The evolution of the effective temperatures and the surface densities for this model (B3) are plotted in Figures 9a and 9b and Figures 10a and 10b. The mechanism of the developing outbursts is exactly the same as in the case of the constant viscosity. However, there are significant qualitative differences. In the cold, low state the surface density increases with radius. The matter flows slowly through the disk and accumulates at the inner parts. At a certain point the critical value of surface density is reached at the radius located inside the instability strip $[\Sigma_B(r),$ a straight line shifted up from $\Sigma_A(r)$; see Fig. 9a and the thermal instability starts. It propagates in both directions. Since the surface density gradient is now steeper than in the constant-viscosity case, infall is favored and there is a high

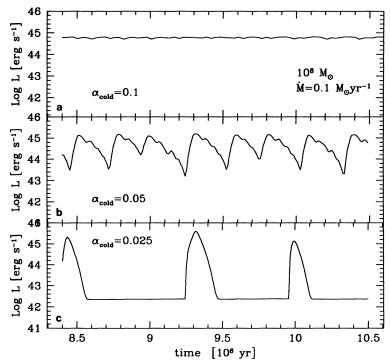


Fig. 7.—Luminosity variations due to instabilities in an accretion disk around a $10^8\,M_\odot$ black hole with $\dot{M}=0.1\,M_\odot$ yr $^{-1}$, $\alpha_{\rm hot}=0.1$, and $\alpha_{\rm cold}$ assumed to be (a) 0.1, (b) 0.05, (c) 0.025.

rate of matter transfer inward. However, the peak of the surface density also travels outward, and the rapid infall starts after it has passed through a given radius (see Figs. 9a and 10a). The peak of the surface density propagates up to the end of the instability strip, where there is no more matter supply and the cooling starts. The transfer into the cold state happens more gradually (see Figs. 9b and 10b). The cooling starts at the outer edge and propagates into the inner regions until the entire disk reaches the cold, quiescent state. Once in the cold state, the disk accumulates matter up to the point when the local conditions at a given radius put it into the high state and the instability begins once again.

During the hot state the matter flows through the disk very quickly and the local accretion rate is high. Variations of the local accretion rate are presented in Figures 11a and 11b. The most dramatic changes happen inside the instability strip, where a large discontinuity forms and locally matter can also flow outward. However, there is always infall beyond the insta-

bility strip, and changes in the local accretion rate propagate over the entire disk, generating large luminosity variations. The average luminosity and the dispersion for each model are given in Table 2. While the dispersion is very small in all the models with $\alpha_{\rm hot}/\alpha_{\rm cold} < 4$, it is significantly larger in the models with a larger ratio. All models with $\alpha_{\rm cold} = 0.025$ show changes of the order of a few in log L. The amplitudes increases with the increase in the accretion rate, since the size of the unstable region grows with the accretion rate too.

The outbursts discussed by Mineshige & Shields (1990) start always at the outer edge of the instability strip and move inward, opposite to the situation in most of our models. The reason is that the dynamics of an outburst depends on the local shape of the S-curve, and in particular on the mass-supply timescale at the outer edge and the accumulation (viscous) timescale in the low state at the inner edge, as discussed in § 4.1. When the viscous timescale is longer than the accretion timescale, the outburst starts inside the instability zone, while in the

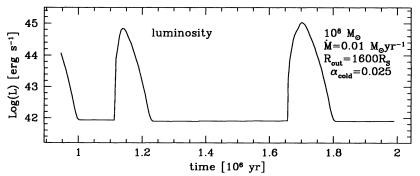


Fig. 8.—Variations of the luminosity for a model with black hole mass $10^8~M_{\odot}$ and accretion rate $0.01~M_{\odot}~{\rm yr}^{-1}$; the viscosity parameter is different in the high and low states: $\alpha_{\rm hot}=0.1$ and $\alpha_{\rm cold}=0.025$.

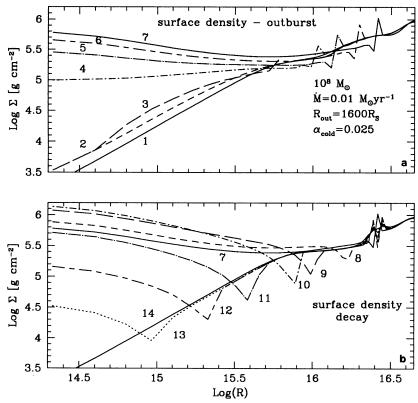


FIG. 9.—Evolution of the surface density of an accretion disk around a black hole of mass $10^8 M_{\odot}$ with accretion rate $0.01 M_{\odot}$ yr⁻¹; the viscosity parameter is different in the high and low states: $\alpha_{\text{hot}} = 0.1$ and $\alpha_{\text{cold}} = 0.025$. (a) Outbursts. (b) Decay. Time between the curves in years: curves 1-2, 2.82×10^4 ; 2-3, 700; 3-4, 600; 4-5, 600; 5-6, 700; 6-7, 600; 7-8, 600; 8-9, 1200; 9-10, 700; 10-11, 5200; 11-12, 2600; 12-13, 2600; 13-14, 7800.

opposite case it always starts at the outer edge. The definition of viscosity assumed by Mineshige & Shields (1990) forces the second situation.

4.2.1. Spectra

The large variations in the local accretion rates affect the emitted spectrum significantly. The evolution of the spectral energy distribution for the outburst and decay phases (model B3) is plotted in Figures 12a and 12b. In the low state the accretion rate is very low, and the peak of the emission moves into the IR part, while the luminosity drops below the average values observed in bright AGNs. The disk in the hot state contributes to the optical-UV part of the spectrum, and the spectral index varies over the outburst and decay phases.

We compute the distribution of the spectral slopes in the optical-UV range. However, in this case we do not include the entire cycle, as the luminosity at the low state is frequently below the level which makes the AGN signature detectable, and such objects would not be included in an AGN sample. Therefore, we adopted an arbitrary limit of $\log(\nu L_{\nu}) = 42.0$ at $\log \nu = 14.8$. We checked that the results do not depend significantly on this value. We also exclude the spectra for which the luminosity peaks below 3000 Å, since there are no data to support such a spectral shape for bright AGNs. These values are quoted in Table 2. The mean spectral index for one cycle in model B3 ($\dot{M}=0.01~M_{\odot}~\rm yr^{-1}$) is equal to 0.13 ± 0.54 , and it is smaller for models with higher accretion rates. The values of the spectral index are between -0.28 (model B3) and -1.15 (model D3) for the disk in the high state, in which an outburst

extends over the entire instability strip (curve 7 in Figs. 9, 10, and 12).

As the ratio of the two values of the viscosity coefficient increases, the median value for the spectral slope increases (e.g., from 0.02 to 0.31 for the B models). This reflects the fact that during certain stages (see curves 8–10 in Figs. 9, 10, and 12) the innermost part of the disk dominates and the spectrum is closer to a single-temperature blackbody than to a stationary disk. Such spectra are even more discrepant from the observational data than single-viscosity models. However, when the ratio of the two viscosity parameters is 4, the outbursts are so strong that any luminosity-limited sample has to be biased toward high-state models. Therefore, we should perhaps be looking at the high-state spectral index only. If so, these models are quite promising, particularly when compared with the samples of Neugebauer et al. (1987) and Francis et al. (1991) (see Tables 2 and 3).

4.2.2. Summary

For the models with $\alpha_{\rm hot}/\alpha_{\rm cold}$ the instability develops over the entire disk, causing luminosity variations with the amplitude as large as $\Delta \log L \sim 4$ for the models with $\alpha_{\rm hot}/\alpha_{\rm cold} = 4$. The light curves show outbursts clearly separated in time. The disk remains in the hot state for only a small percentage of the entire cycle. The luminosity in the hot state is close to the Eddington value. The optical spectral index varies a lot over the entire cycle. In the very high state the spectrum is even harder than that of a stationary disk, close to that of a single blackbody, while in the intermediate state it becomes close to

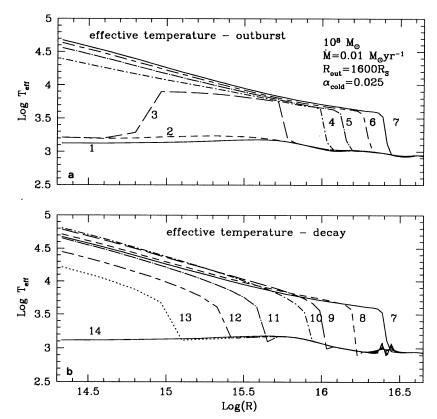


Fig. 10.—Evolution of the effective temperature of an accretion disk around a black hole of $10^8\,M_\odot$ with accretion rate $0.01\,M_\odot$ yr $^{-1}$; the viscosity parameter is different in the high and low states: $\alpha_{\rm hot}=0.1$ and $\alpha_{\rm cold}=0.025$. (a) Outbursts. (b) Decay. Time between the curves in years: 1-2, 2.82×10^4 ; 2-3, 700; 3-4, 600; 4-5, 600; 5-6, 700; 6-7, 600; 7-8, 600; 8-9, 1200; 9-10, 700; 10-11, 5200; 11-12, 2600; 12-13, 2600; 13-14, 7800.

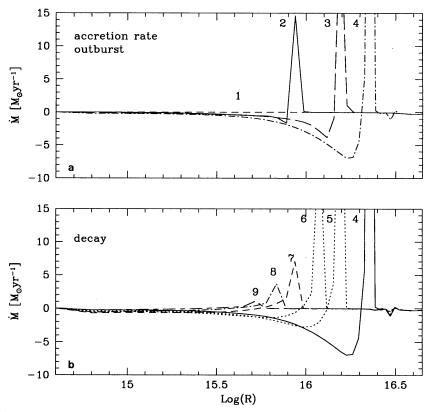


Fig. 11.—Evolution of a local accretion rate during (a) outburst and (b) decay, for the model with $M_{\rm bh}=10^8~M_{\odot}$ and $\dot{M}=0.01~M_{\odot}~\rm yr^{-1}$, with the viscosity parameter different in the high and low states: $\alpha_{\rm hot}=0.1$ and $\alpha_{\rm cold}=0.025$. (a) Outbursts. (b) Decay. Time between the curves in years: 1–2, 400; 2–3, 1100; 3–4, 1000; 4–5, 1100; 5–6, 700; 6–7, 1200; 7–8, 1200; 8–9, 1600.

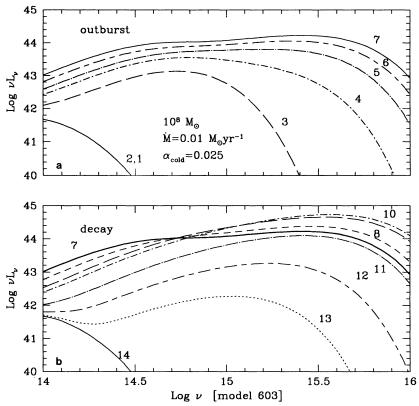


FIG. 12.—Evolution of the spectrum emitted by an accretion disk around a black hole of mass $10^8\,M_\odot$ with accretion rate $0.01\,M_\odot$ yr $^{-1}$; the viscosity parameter is different in the high and low states: $\alpha_{hot}=0.1$ and $\alpha_{cold}=0.025$. We assumed that the disk emits locally as a blackbody. (a) Outbursts. (b) Decay. Time between the curves in years: 1-2, 2.82×10^4 ; 2-3, 700; 3-4, 600; 4-5, 600; 5-6, 700; 6-7, 600; 7-8, 600; 8-9, 1200; 9-10, 700; 10-11, 5200; 11-12, 2600; 12-13, 2600; 13-14, 7800.

the observed spectrum. In the low state the disk does not contribute to the optical-UV region but is relatively bright in the IR.

5. THE EFFECT OF CORONAL IRRADIATION

Observations show clearly that a considerable fraction of the energy of the accreting gas is released close to the black hole in the form of X-ray emission (thermal or nonthermal) and that a substantial fraction of the X-ray emission suffers Compton reflection from a "cold" medium (Mushotzky, Done, & Pounds 1993). (There is no clear connection between this emission and an accretion disk, although a number of models have been suggested.) In the disk models this "reflector" is the disk, which is then strongly irradiated.

Fully self-consistent calculations should include this irradiation when computing the disk structure, as in the one-zone model by Mineshige et al. (1990). The irradiation not only modifies the disk spectrum (Malkan 1991; Siemiginowska et al. 1994) but also influences its time evolution. We are unable to follow such an evolution with our code because irradiation changes the shape of the S-curve, making the scaling technique inapplicable. We can, however, estimate the importance of irradiation by calculating the increase in effective temperature due to different levels of irradiation.

If a fraction of the X-ray photons originating close to the black hole hits the disk surface, these photons are either scattered or absorbed in an upper layer of Thomson optical depth of the order of a few. The scattered, or Compton-reflected, component predicted by Lightman & White (1988) is observed in many Seyfert galaxies (e.g., Pounds et al. 1990; Nandra &

Pounds 1994) and contains 15%-30% of the total luminosity (e.g., Życki et al. 1994), as expected if one half of the primary X-rays is reprocessed and the other half goes directly to an observer. The rest of the energy goes into heating the absorber and eventually is reemitted in the form of lines, recombination continua, and free-free emission (Dumont & Collin-Souffrin 1990a, b; Ross & Fabian 1993). As the disk is optically thick, most of the reemission forms a thermal component (e.g., Hurè et al. 1994).

For simplicity, we assume that the X-ray radiation flux intercepted by the disk surface is thermalized with 100% efficiency.

The irradiation may be either direct, due to the exposure of the disk surface to the X-ray radiation, or indirect, due to the scattering of the X-rays by extended fully ionized corona. In this paper we confine ourselves to the second possibility. At every radius r we assume that 10% of the total X-ray flux is intercepted by the disk $(L_X/4\pi r^2)$, where L_X is the X-ray bolometric luminosity assumed to be connected to the accretion rate at the inner disk radius through the relation $L_{\rm X}=0.1$ $\dot{M}c^2$). Such a description of the scattering of X-rays in the corona gives roughly the order of magnitude for both the flux and its radial distribution. Accurate computations show a strong dependence of the scattered flux on the ratio of luminosity to the bolometric luminosity, as well as on the corona temperature, and give a somewhat steeper decrease of the flux with the radius (e.g., Kurpiewski 1994). This simple description allows us to estimate the effect of irradiation on the evolving accretion disk spectra. An example of the change of the effective temperature distribution and the resultant spectra due to irradiation in the high state and in the low state in models C0

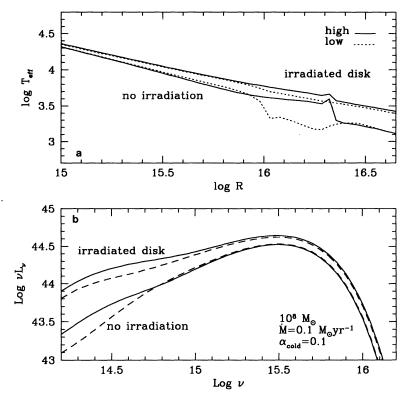


Fig. 13.—(a) Effective temperature in the high and low states of the disk ($10^8 M_{\odot}$, $\dot{M} = 0.1 M_{\odot}$ yr⁻¹, $\alpha_{\rm cold} = 0.1$), which is irradiated by an external hard source (the luminosity of the external source is assumed to be 10% of the total luminosity generated in the disk). (b) Spectrum of the irradiated and nonirradiated disk in the high and low states.

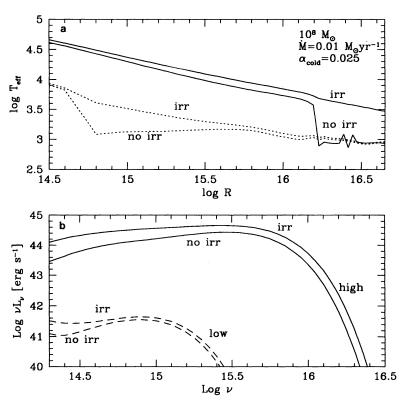


Fig. 14.—(a) Effective temperature in the high and low states of the disk ($10^8\,M_\odot$, $\dot{M}=0.01\,M_\odot$ yr $^{-1}$, $\alpha_{\rm cold}=0.025$), which is irradiated by an external hard source (the luminosity of the external source is assumed to be 10% of the total luminosity generated in the disk). (b) Spectrum of the irradiated and nonirradiated disk in the high and low states.

and B3 are shown in Figures 13 and 14. The optical spectral indices for all models are given in Table 2.

We can see that strong irradiation causes a profound change in the effective temperature distribution (a factor of 2 in the instability strip) and results in a different spectral shape. As the irradiating flux decreases more slowly with radius than the flux generated in the disk, the enhancement of the disk radiation flux is stronger in the outer parts, leading to significant steepening of the spectra in low states and, even more strongly, in high states. These models fit well the spectral range given by the data (spectral index can be of the order of -0.7 as in the Cristiani sample; see Table 3). Detailed conclusions, however, can be made only after full time-dependent calculations of the irradiated disk are done.

6. DISCUSSION

6.1. Have Low-State Disks Been Detected?

Since the amplitude of the bolometric luminosity variations can reach 4 orders of magnitude, it is interesting to ask whether the "activity" of AGNs is connected with its accretion disks being in the high state, while the "inactivity" of normal galaxies or faint AGNs is due to their having quasars with a disk in the low state.

There are just a few normal, or weakly active, galaxies that have had detailed studies of their nuclei. Fabbiano (1988) presented a spectral energy distribution (SED) from radio to X-rays for the low-luminosity nucleus of the spiral galaxy M81, which shows UV, radio, and X-ray activity. Relative to the rest of the SEDs, the UV luminosity is lower by an order of magnitude than in regular AGNs. Moreover, the stellar contribution to this UV luminosity is rather high, and the true nonstellar component could be significantly lower. This abnormally weak UV bump strength may indicate that we are observing an accretion disk in a low activity state in M81 (cf. Fig. 12). In the low state there is no contribution from the disk emission into the UV, while in the high state the ratio between the UV and IR fluxes is ~ 2 . Scaling the observed UV-IR emission of the M81 nucleus to that appropriate for high-state disk values produces a UV bump of normal strength.

To determine whether the nuclei of low-luminosity objects are indeed in the low state, we need to constrain $L/L_{\rm Edd}$, which requires estimation of the mass of the central black hole. We cannot constrain the mass using the optical-UV spectra when the disk instability is present, so other methods are needed. Surprisingly, the best method so far comes from recent observations of water masers in NGC 4258. These water masers are located at distances only a fraction of a parsec away from the central mass (Miyoshi et al. 1995). This gives a good measurement for the central mass, and so central black hole mass, of $3.6 \times 10^7~M_{\odot}$. The nuclear X-ray flux from ROSAT HRI observations is $\sim 3 \times 10^{39}~{\rm ergs~s^{-1}}$ (Cecil, Wilson, & De Pree 1995). Combining these implies $L_{\rm X} \sim 5 \times 10^{-7} L_{\rm Edd}$. If we relate the X-ray emission to an accretion rate at the inner edge of an accretion disk (as seems to be the case in the X-ray transient sources; Huang & Wheeler 1989; McClintock, Horne, & Remillard 1995), then this luminosity constrains the accretion rates to very low values ($\dot{M} \sim 10^{-6}~M_{\odot}~\rm{yr}^{-1}$), suggesting that we see the disk in its low state. This galaxy is a known radio source with an extraordinary "braided" jet (Cecil, Wilson, & Tully 1992). Further detailed study of the radio and X-ray structure may give more information about its activity in the past.

The double-peaked line profiles observed in some radio galaxies (Eracleous & Halpern 1994) may originate in an accretion disk. For these radio galaxies the starlight contribution to the continuum emission near $H\alpha$ is much higher than for single-peaked radio galaxies. About 9% of the total continuum emission in the objects with the single-peaked profiles comes from the starlight, while in the objects with the double-peaked profiles it is more than 50%. As a result, the central optical-UV continuum component is much weaker in the objects with double-peaked lines than in the other objects, possibly indicating the presence of a disk in the low state.

6.2. Long-Term Variability and Quasar Evolution

Cavaliere & Padovani (1988) have studied luminosity functions for different types of quasar evolution models. One possibility they consider is a recurrent model, in which a quasar has a short active phase related to the outer fuel supply-for example, from mergers. Our modeling suggests another intrinsic mechanism controlling the matter supply into the innermost central disk region—the ionization instability in α_{cold} < α_{hot} disks. In our models the luminosity varies, over the cycle, between $10^{-5}L_{\rm Edd}$ and $L_{\rm Edd}$, and objects remain in the low state for a large fraction of the time. Such objects would contribute to the low-luminosity end of the luminosity function, while the rarer, high-state objects would populate the highluminosity end. Since low-luminosity AGNs are much more common than high-luminosity AGNs, this model at least gives the correct sense of the quasar luminosity function. A detailed analysis of the luminosity function generated by a population of objects with disk-instability-driven activity is presented in Siemiginowska et al. (1995). We note only that in such an evolutionary model the mass of the central black hole does not increase significantly, since the average accretion rate is very small, even when the active state luminosity reaches Eddington values. As a consequence, central black holes can be present in normal galaxies and, at the same time, may not be very massive, even if they accrete matter continuously over hundreds of millions of years.

Since the characteristic timescales for the thermal-viscous instabilities are of the order of 10^4 – 10^5 yr, one method of studying the long-timescale activity of quasars can be to use the "proximity effect" (Kovner & Rees 1989; Dobrzycki & Bechtold 1991). The characteristic recombination time for the Ly α forest clouds is about 10^4 yr. If there is any significant change in the UV luminosity of a quasar, it must be reflected in the Ly α forest clouds at different distances from the quasar. We may be able to see the variations when observing the Ly α forest absorption lines in the spectra of apparent pairs of quasars, which are at different redshifts but close in angular separation (Bechtold 1995; Carswell 1995).

6.3. Location of the Unstable Zone versus the Characteristic Radii

There are some constraints on the location of the unstable zone that we have to consider. We can compare the inner radius of the instability strip r_D (eq. [5]) with the transition radius, r_{ab} , from the gas-dominated to the radiation-pressure-dominated unstable region in a standard α -disk (Shakura & Sunyaev 1973):

$$r_{ab} = 1.8 \times 10^{15} M_8^{1/3} \dot{M}^{16/21} \alpha_{0.1}^{2/21} \text{ cm}$$

 $\sim 0.11 r_A M_8^{-1/15} \dot{M}^{38/105} \alpha_{0.1}^{1/21}$. (22)

Thus both instability (Lightman-Eardley and ionization) will be present in α -disks in the same regions, and probably cannot be treated separately. The global disk evolution we have described for β -viscosity ($\tau_{r\phi} \sim P_{\rm gas}$) model may be invalid if the doubly unstable α -viscosity ($\tau_{r\phi} \sim P_{\rm total}$) is actually a better representation of disk properties. Clearly, spectral predictions for the ultraviolet region would be affected, since this part of the spectrum comes from the innermost region undergoing its own instabilities (in the α-viscosity case). Luminous AGNs are variable in the optical-UV band on timescales of days to years. Therefore, regardless of the origin of the observed variability, we cannot expect the disk to be stable (at least in the innermost regions) on short timescales either. There might be a connection between this variability and the Lightman-Eardley instability, but making such connections requires better theoretical understanding of this instability.

The second problem with the applicability of the model is related to the self-gravity effects which are estimated to be important beyond the radius (Toomre 1964; Clarke 1989):

$$r_{\rm sg} = 1.7 \times 10^{16} M_8^{1/3} \dot{M}^{4/9} \alpha_{0.1}^{2/9} \text{ cm}$$
 (23)

A significant part of the instability strip lies within this region (see Fig. 4). This may affect the appearance of the ionization instability and also change the overall evolution of an accretion disk. Gravitational instabilities have been examined as a source of the viscosity (Paczyński 1978). They can cause fragmentations of the outer parts of the disk (Sakimoto & Coroniti 1981) and also lead to star formation (Shlosman & Begelman 1989).

7. SUMMARY

The ionization instability operating in cataclysmic variables, and also most probably responsible for X-ray transients, may well operate in AGN accretion disks, causing evolutionary changes on timescales of hundreds of thousands of years. This variability has two basic consequences: (1) the massive black hole does not always have to be in an active stage, and (2) the spectrum of an accretion disk is not that of a stationary model.

The effects of the thermal-viscous instability depend on two parameters only: an accretion rate and assumed viscosity prescription. The accretion rate determines the location and dimension of the unstable region. If it is large, then the hydrogen ionization zone is far away from the center and the instability does not develop. On the other hand, for very low accretion rates the unstable region is situated close to the inner radius of the disk, causing significant luminosity variations even in the case of a constant viscosity parameter.

Viscosity plays an important role in the development of the instability in the disk. For the constant viscosity parameter α the instability propagates only inside a very narrow unstable zone. Therefore, the resulting luminosity variations depend critically on the location of this zone, and so on the accretion rate. The main contribution to the luminosity comes from the hottest, innermost regions which are stable for high accretion rates and unstable for very low accretion rates.

In the case of low accretion rates the unstable zone moves closer to the center and the influence of the instability on the disk emission is clearly visible. However, even in this case the amplitude is small (± 0.4 in log L), and the average luminosity is too low to be appropriate for AGNs. There is no full quiescence of the disk reached after the outburst which is reflected in the shape of the light curve.

Models with a nonconstant viscosity prescription have a much broader unstable zone covering a significant part of the disk, and therefore they show much more violent behavior. For the largest studied ratio between the viscosity in the hot and cold states the changes in the local accretion rates are very high. For these models ($\alpha_{cold} = 0.25\alpha_{hot}$) the variations are dramatic (± 4.0 in log L). The hot and cold phases of the disk are clearly separated in the light curve. After the outburst the disk reaches quiescence and accumulates matter.

We have shown that the effect of the disk evolution on the disk spectra also depends on the description of the viscosity. However, the typical effect is just a slight softening of the spectrum in the constant-viscosity case, not enough to make models compatible with the observed spectral distribution, or even a hardening of the spectrum in the case of nonconstant viscosity.

A much stronger effect on disk spectra than the evolution itself is expected, due to irradiation if some fraction of the accretion energy is converted into hard X-rays, as observed, and the X-ray flux irradiates the disk surface via scattering in the extended corona. Irradiated disk spectra compare well with the observational samples. This approach is therefore worth pursuing. Future models should include the effect of the irradiation on the disk structure (and not just on disk spectra), since the feedback may well change the evolutionary pattern

We did not include advection in our evolutionary modeling. Since it may affect the overall appearance of instability, it should be considered in the future modeling. Also, self-gravity, which affects the outer regions of the disk, should be included in overall studies of accretion disks in AGNs. An interesting problem comes from the possibility that there may be two coexisting unstable zones in supermassive accretion disks: one related to the ionization instability and the other to the instability of the radiation-pressure-dominated regions. The presence of both instabilities will result in different variability patterns which we cannot predict from the present studies.

This work was supported in part by grant 2P30401004 of the Polish State Committee for Scientific Research and NASA grant NAGW-2201. Part of the computations have been done in Osservatorio Astronomico di Roma in Monteporzio (Italy) during A. S.'s visit. The numerical code to calculate the disk evolution was developed from the original dwarf nova code kindly provided by Professor J. I. Smak. We thank Jill Bechtold, John Cannizzo, Martin Elvis, Pepi Fabbiano, and J. I. Smak for many helpful discussions. We thank the anonymous referee for helpful comments.

REFERENCES

Abramowicz, M., Czerny, B., Lasota, J.-P., & Szuszkiewicz, E. 1988, ApJ, 332, Alexander, D. R. 1975, ApJS, 29, 363

Begelman, M. 1985, Astrophysics of Active Galaxies and Quasi-stellar Objects, ed. J. S. Miller (Mill Valley: University Science Books), 141 Cannizzo, J. K. 1993a, ApJ, 419, 318

1993b, in Accretion Disks in Compact Stellar Systems, ed. J. C. Wheeler (Singapore: World Scientific), 6 Carswell, R. F. 1995, in Proc. ESO Workshop on QSO Absorption Lines, ed.

G. Meylan (Heidelberg: Springer), 313

Editions Frontières), 165

McClintock, J. E., Horne, K., & Remillard, R. A. 1995, ApJ, 442, 358

—. 1984, Astr. Ap., 132, 143

Mineshige, S., 1990, private communication

Mineshige, S., & Shields, G. A. 1990, ApJ, 351, 47

Mineshige, S., Tuchman, Y., & Wheeler, J. C. 1990, ApJ, 359, 176

Mineshige, S., & Wheeler, J. C. 1989, ApJ, 343, 241 Miyoshi, M., Moaran, J., Herrnstein, J., Greenhill, L., Nakai, N., Diamond, P., & Inoue, M. 1995, Nature, 373, 127 Mushotzky, R. F., Done, C., & Pounds, K. A. 1993, ARA&A, 31, 717 Nandra, P., & Pounds, K. A. 1994, MNRAS, 268, 405 Nandra, P., & Pounds, K. A. 1994, MNRAS, 268, 405

Narayan, R., & Yi, I. 1994, ApJ, L13

——. 1995, ApJ, 444, 231

Neugebauer, G., Green, R. F., Matthews, K., Schmidt, M., Soifer, B. T., & Bennet, J. 1987, ApJS, 63, 515

Paczyński, B. 1978, Acta Astron., 28, 91

Pojmański, G. 1984, Ph.D. thesis, Warsaw Univ.

——. 1986, Acta Astron., 36, 69

Paczyński, B. 1978, Acta Astron., 36, 69

Paczyński, G. 1984, Ph.D. thesis, Warsaw Univ. Pounds, K. A., Nandra, P., Steward, G. C., George, I. M., & Fabian, A. C. 1990, Nature, 344, 132 Pringle, J. E. 1981, ARA&A, 19, 137 Rees, M. Y. 1984, ARA&A, 22, 471 Ross, R. R., & Fabian, A. C. 1993, MNRAS, 261, 74 Ross, R. R., & Fabiai, A. C. 1973, MAINTON, 201, 77
Sakimoto, P. J., & Coroniti, F. V. 1981, ApJ, 247, 19
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Shlosman, I., & Begelman, M. 1989, ApJ, 341, 685
Siemiginowska, A. 1991, in IAU Colloq. 129, Structure and Emission Properties of Accretion Disks, ed. C. Bertout, S. Collin-Souffrin, J.-P. Lasota, & J. Tran Thanh Van (Gif-sur-Yvette: Editions Frontières), 519 Siemiginowska, A., & Czerny, B. 1989, MNRAS, 239, 289 Siemiginowska, A., Kuhn, O., Elvis, M., Fiore, F., McDowell, J. C., & Wilkes, B. J. 1995, ApJ, 454, 77 Siemiginowska, A., et al. 1995, in preparation Smak, J. I. 1982, Acta Astron., 32, 199 1984a, Acta Astron., 34, 161 1984b, PASP, 96, 5

Tanaka, Y., & Lewin, W. H. G. 1995, in X-Ray Binaries, ed. W. H. G. Lewin, J. van Paradijs, & E. P. J. van den Heuvel (Cambridge: Cambridge Univ. Press), in press Toomre, A. 1964, ApJ, 139, 1217 Tuchman, Y., Mineshige, S., & Wheeler, J. C. 1990, ApJ, 359, 164 Życki, P. T., Krolik, J. H., Zdziarski, A. A., & Kallman, T. R. 1994, ApJ, 437,

Meyer, F., & Meyer-Hoffmeister, E. 1982, A&A, 106, 34