ACCRETION DISK MODELS OF ACTIVE GALACTIC NUCLEI: A CRUCIAL ROLE FOR REPROCESSING?

Mitchell C. Begelman and Martijn de Kool

Joint Institute for Laboratory Astrophysics
University of Colorado
Boulder, Colorado 80309-0440

ABSTRACT

GINGA and EXOSAT data on the spectra and X-ray variability of AGN indicate that a large fraction of the primary X-ray emission is reprocessed by cold gas located close to the central black hole. However, it is not clear that this gas forms the inner part of a "standard" accretion disk. In particular, the observed behavior of the iron Kα fluorescence line is difficult to explain in terms of the reprocessing of hard X-rays by a standard accretion disk, unless most of the X-ray luminosity is generated in local regions close to the disk, with a size smaller than R_g. We outline some alternatives to the standard disk model, both with and without a disk-like geometry. We briefly discuss the possible importance of reprocessing by a disk at large radii. (Based on talk delivered by M. Begelman.)

1. INTRODUCTION

Listening to the talks at this meeting, I have learned that the temperature distributions of disks in cataclysmic variables and protostellar systems do not seem to fit the simple Shakura–Sunyaev theory. Instead of declining $\propto R^{-3/4}$, the effective temperatures seems to fall off more slowly, suggesting either that the energy generation in the disk is nonlocal or that the outer parts are irradiated by inner parts. Those of us studying AGN are not so fortunate as to have such clean diagnostics of disk properties; indeed, the evidence that there are disks at all in the central regions of AGN is weak at best. Unlike CVs, we never observe eclipses of AGN disks, and have little handle on the geometry. In contrast to protostellar systems, the central regions of an AGN are far too small to resolve spatially. Model-dependent fits of the "big UV bump" [1] suggest the plausibility of an accretion disk model for AGN, although Guilbert and Rees [2] have stressed that the UV-emitting gas need not take the form of a disk. More recently, X-ray observations by GINGA and EXOSAT have given us the strongest evidence yet that there is "cold" ($\sim 10^5$ K) gas close to the central black hole. These observations also indicate that the radiation we observe is not simply that released locally by the infalling gas, but is heavily modified by reprocessing. For much of this talk, I will focus on the implications of these X-ray results for disk models of the innermost regions of AGN, later commenting briefly on the possible role of reprocessing on larger scales.

Let me begin by contrasting the conditions inferred to exist in the central regions of AGN with those thought to obtain in another type of system where accretion disks may be present. The closest relative to an AGN is an X-ray binary, where the central object is also compact (either a neutron star or a stellar-mass black hole) but is many orders of magnitude less massive than the black hole in an AGN. Table 1 contrasts the two cases in black hole mass, luminosity normalized to the Eddington limit, disk temperature, ratio of radiation to gas pressure, and disk surface gravity. Although L/L_{Edd} may be similar, the observable properties of the disks will be very different in the two cases. First, radiation from an AGN disk will peak in the extreme ultraviolet to soft X-ray band, making it more difficult to observe directly (due to absorption by heavy elements). Second, an AGN disk will have a much higher ratio of radiation pressure to gas pressure, making its surface layers more susceptible to runaway heating by nonthermal X-rays formed in a disk corona; we will discuss this in more detail below. Third, the density and surface gravity in an AGN disk will be lower, implying that the photosphere will be geometrically thicker and the radiation further from LTE than in the X-ray binary case.

TABLE 1

XRB		AGN
$1 - 10M_{\odot}$	M	$10^6 - 10^9 M_{\odot}$
$10^{-2} - \sim 1(?)$	L/L_{Edd}	$10^{-2} - \sim 1(?)$
$10^7 - 10^8$ K	$T \propto M^{-1/4}$	$10^5 - 10^6$ K
$1 - 100$	$p_{rad}/p_{gas} \propto M^{1/4}$	$100 - 10^4$
$\sim 10^8 g_{\odot}$	$g \propto M^{-1}$	$\sim g_{\odot}$

2. EVIDENCE OF X-RAY REPROCESSING

New X-ray data from GINGA and EXOSAT have shown that the hard X- ray spectra of many AGN are quite complex [3-6]. Instead of a "canonical" $\alpha_X \approx 0.7$ power-law [7], the spectra are better represented by the combination of a steeper ($\alpha_X \sim 0.9$) "nonthermal" continuum, a ~ 6.4 keV iron fluorescence line, a 7–8 keV absorption edge due also to iron, and a broad hump at $\sim 10 - 30$ keV interpreted as being due to "Compton reflection" of the hard continuum by cold gas [2,6,8]. The hard X-ray continua of a number of AGN have been observed to vary rapidly [9]. To my knowledge, the record-holder for variability is the Seyfert galaxy NGC 6814, which has been observed to change its apparent luminosity by about 10^{43} erg s^{-1} in 50 seconds [10]. The observed change was a substantial fraction (about half) of the apparent luminosity at that time. Perhaps even more remarkable is the observation that the iron fluorescence line in NGC 6814 seems to vary more or less in phase with the 2–20 keV flux (with a time delay of less than 250 seconds). The 6.4 keV energy of the line indicates that it is produced by relatively neutral iron, meaning that the K-shell seems to be quite complete. Its equivalent width is so large that it is almost certainly due to fluorescence by cold ($\lesssim$ few $\times 10^6$ K) gas which is intercepting and reprocessing a large fraction of the very hard X-ray flux [3,6,11]. The rapid variability then indicates that this reprocessing is occurring very locally.

These observations are the strongest evidence to date that there really is cold gas in the central engine. But it does not mean that there is an accretion disk in the usual sense. Below I will discuss some of the implications of these observations, and suggest some alternatives to the "standard" model.

3. THERMAL STATE OF GAS IN THE CENTRAL ENGINE

It is not too surprising to find cold gas in the central engine, because optically thick gas will tend to reprocess the ambient radiation (energy density U) and equilibrate to the effective black body temperature $T_{bb} = (U/a)^{1/4}$. In terms of the parameters which appear to be typical of NGC 6814:

$$T_{bb} > \left(\frac{\Delta L}{4\pi a c^3 \Delta t^2}\right)^{\frac{1}{4}} \sim 2\times 10^5 \left(\frac{\Delta L_{43}}{\Delta t_{50}^2}\right)^{\frac{1}{4}} \ \mathrm{K},$$

where $\Delta L_{43} \equiv \Delta L/10^{43}$ erg s^{-1} is the luminosity produced in the varying region and $\Delta t_{50} \equiv \Delta t/50$ sec is the variability timescale. When I talk about the varying regions, I am not assuming at this point that the whole "central engine" of the AGN is varying. I think that this is still an open question, to which I will return later. The cold gas intercepts an amount of hard X-rays similar to the observed luminosity, so it must cover a large part of the sky if viewed from the source of the hard X-ray flux. To explain the fluorescence line, the cold gas must also absorb most of the incident flux at energies just above the iron K-edge. This implies that the optical depth of the cold gas to electron scattering cannot be very much smaller than one.

The rapid variability indicates that the "compactness" ℓ of the varying regions is high,

$$\ell > \frac{\Delta L \sigma_T}{4\pi m_e c^4 \Delta t} \sim 14 \frac{\Delta L_{43}}{\Delta t_{50}} > 1.$$

This implies that any electrons that are hotter than the inverse Compton temperature, i.e., relativistic electrons and hot thermal electrons of T greater than 10^7 K or 10^8 K, will cool down to the inverse Compton temperature T_{IC} on a time scale that is short compared to the light crossing time across the varying region ($t_{cool} \sim \Delta R/c\ell$). Consequently, it is unlikely that a large fraction of the energy in these objects at any one time resides in relativistic or hot thermal electrons. Any energy that is put into hot or relativistic electrons would be rapidly transferred to radiation, and lost (unless the radiation is trapped by a large optical depth). In a highly variable source, the distribution of hot electrons is likely to be time dependent and inhomogeneous on small scales.

In order to remain cold despite X-ray heating, the irradiated gas must be sufficiently dense [2,12]. Since the observed hard X-ray spectrum is not very steep, the inverse Compton temperature associated with the hard X-rays is probably of order 10^8 K. When combined with the soft radiation that is present (e.g., reprocessed X-rays and primary UV radiation from an accretion disk), it is probably still larger than 10^7 K because a significant fraction of the flux is in hard X-rays. Gas exposed to that kind of spectrum tends to have a bimodal behavior. If it is dense enough, it will be able to process the radiation and equilibrate at a temperature close to T_{bb} ($\sim 2 \times 10^5$ K). But if it is not dense enough, the two-body interactions between particles needed to cool the gas will not be able to counteract the heating effect of the X-rays and the gas will tend to heat up to the inverse Compton temperature.

It turns out that the transition between these two temperatures occurs rather suddenly. Ferland and Rees [12] plotted temperature versus density, which gives a very smooth transition, but this is a bit misleading. If you instead plot temperature versus pressure of the gas, or versus radiation pressure divided by gas pressure, which is sometimes called the ionization parameter, the transition is very steep, and may even "double back" (Fig. 1). Thus, there will be a tendency for most

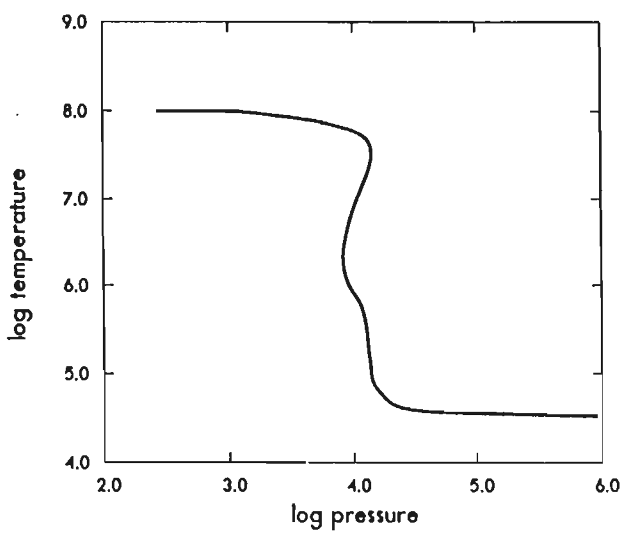

Fig. 1. The equilibrium temperature as a function of gas pressure for an optically thin gas irradiated by a spectrum $F_\nu \propto \nu^{-0.7}$, with a total flux appropriate to the central regions of an AGN (after [12]).

of the gas either to be close to T_{bb} or to heat up to T_{IC}, with very little gas at intervening temperatures. If the p vs. T curve doubles back, then gas at intervening temperatures is thermally unstable, and there is a range of pressures over which the hot and cold phases can coexist at fixed pressure [13]. The exact conditions under which the phase transition occurs are not very sensitive to the spectrum of the radiation, and correspond to a situation where the radiation pressure and the gas pressure are comparable. Thus, if the radiation pressure is larger than the gas pressure by a significant amount then the gas will be heated up to the inverse Compton temperature. Conversely, the density required to keep the gas down at the cold temperature is approximately

$$n > 10^{16} \left(\frac{T_{bb}}{10^5 K} \right)^3 \quad \mathrm{cm}^{-3} \quad .$$

What is the distribution and dynamical state of the cold gas in an AGN? Some dynamical support is necessary, because 10^5 K is orders of magnitude below the virial temperature in the central engine. One possibility is that the gas is supported by angular momentum, in which case the most likely geometric configuration would be a disk. Confinement is then by the vertical component of gravity. Alternatively, the gas might be in the form of filaments which are both confined and prevented from settling into a thin disk by very powerful magnetic fields [2,14]. It is unlikely that the cold gas could be confined by the pressure of a hot intercloud medium in thermal equilibrium, since T_{IC} is also well below the virial temperature. Any significant contribution to gas pressure confinement would presumably have to have to come from a "two-temperature" plasma [15], with most of the pressure being provided by ions.

4. IMPLICATIONS FOR ACCRETION DISK MODELS

Could the the cold gas responsible for reprocessing the X-rays comprise the inner part of an accretion disk? If we assume that the very short time scale variability in NGC 6814 reflects global fluctuations of the inner parts of the disk, we can infer that the observed luminosity is at least a tenth of the Eddington limit [10]. We would then infer that the inner parts of the disk are supported by radiation pressure. This means that the radiation pressure in the surface layers of the disk, as well as everywhere in the interior, has to be much greater than the gas pressure. Roughly,

$$\frac{P_{rad}}{P_{gas}} \sim 10^4 \left(\frac{L}{L_{Edd}} \right)^2$$

at the photosphere. Recall that, in order for gas exposed to hard X-rays to remain cold, the gas pressure has to be comparable to or larger than the radiation pressure; here, the ratio of radiation pressure to gas pressure is greater than 100.

Thus, if our lower bound on L/L_{Edd} is correct, then we expect the irradiated surface layers of the disk to heat up to the inverse Compton temperature (or higher, if there are other local sources of heating such as magnetic reconnection). At such temperatures the iron K-shell will be largely stripped, hence the gas will not produce the observed fluorescence line at 6.4 keV when irradiated by hard X-rays. Given that the disk is clad by an X-ray heated corona, what are the prospects for passing the hard X-rays through the corona to induce fluorescence in the layers below? To obtain a significant fluorescence yield requires a region in which iron is in a sufficiently low ionization stage *and* hard X-ray photons are present. To avoid too high an ionization state due to collisional ionization the gas temperature has to be below $\sim 3 \times 10^6$ K. If the radiation pressure in ionizing photons is significantly larger than the gas pressure photoionization is likely to be more important than collisional ionization, and the requirement is that the "color temperature" T_X of the hard radiation field be less than $(4-8) \times 10^6$ K. How deep does one have to go into the disk before these conditions are met? First we note that if the X-ray luminosity of the object is comparable to the blackbody luminosity, as seems to be the case, the radiation pressure in hard, ionizing photons is much higher than the gas pressure, and hence both the gas temperature and T_X have to be below a few times 10^6K. The gas temperature will be equal to the Compton temperature of the total radiation field:

$$T_{gas} = T_{IC} \sim \frac{J_{bb}T_{bb} + J_X T_X}{J_{bb} + J_X}$$

where J is the average intensity and the subscript bb refers to thermal radiation from inside the disk. From a simple gray atmosphere model we know that J_{bb} increases linearly with optical depth, whereas J_X in first approximation remains constant or decreases with optical depth. Thus, the inverse Compton and gas temperatures decrease at least as fast as τ^{-1}, and will become less than a few times 10^6K at an optical depth of order 10 or less. The hard photons diffusing into the disk will lose energy by Compton downscattering. Assuming that the layer is cold ($T_{gas} \ll T_X$), a photon of energy kT_X will be strongly downscattered where the optical depth satisfies the condition

$$\frac{4kT_X}{m_e c^2}\tau^2 = 1.$$

Thus, photons with an initial energy above $(4-8) \times 10^6$ K will have lost their excess energy at optical depths greater than $(14-20)$ and there will be a cutoff in the spectrum. This constraint turns out to be more restrictive than the one on the gas temperature, since the ionization state of iron is more sensitive to the color temperature of the hard radiation than to the kinetic temperature of the gas.

This result does not bode well for producing the iron fluorescence line in the sub-coronal layers of the disk, because nearly all of the photons that are capable of causing fluorescence are downscattered below the K-edge before reaching the depth at which the K-shell recombines. A crude estimate (assuming that the hard photon distribution evolves into a Wien spectrum) suggests that only a fraction of order $10^{-3} - 10^{-5}$ of the incident X-ray photons will be available to cause fluorescence at these depths. Furthermore, if any Kα photons were produced, they would be downscattered before they escaped and would not be observed as line photons.

To review what I think are some of the problems or uncertainties in applying the standard accretion disk model in the light of recent X-ray observations: If the luminosity of the disk is greater than a small fraction (~ 0.03) of the Eddington limit, it seems hard for reprocessing by the disk to produce the observed Kα fluorescence line locally so that it can vary in phase with the hard X-rays, because the irradiated layers are too hot and optically thick. The assumption that the luminosity is close to the Eddington limit comes from assuming that the entire inner region of the disk is varying. If you drop that assumption, and assume instead that the disk is very much less luminous than the Eddington limit, then the Kα fluorescence line could be produced locally in the surface layers of a thin accretion disk. In this case the observed rapid variability would not result from global fluctuations of the disk, but might be associated with flares or flashes occurring over a very much smaller scale.

5. ALTERNATIVES TO ACCRETION DISKS

Another possibility is that the standard accretion disk theory does not apply in the central regions of AGN. A thermal component in the spectrum, with roughly the right temperature for an accretion disk, does not necessarily mean that a disk is present. It simply means that there is probably some optically thick gas, which is able to thermalize and reradiate whatever radiation is present. Even if the gas in the center of these sources forms a disk, it may have a structure very different from that of a standard α disk. If we demand that the gas pressure be comparable to or larger than the radiation pressure in order to produce the fluorescence line, then we will be seeking a model in which the energy released by accretion is not dissipated into radiation inside the cold gas itself. In this limit the cold gas will be isothermal, at a temperature close to the local black body temperature.

A disk which is supported by gas pressure and confined by the vertical component of the central black hole's gravity would be very thin compared to a standard radiation pressure-dominated accretion disk. Its density would have to be larger than about $10^{16}\mathrm{cm}^{-3}$ in order to stay cool, but a much higher density inside the disk is guaranteed anyway if the primary energy source is disk accretion. How could a disk provide this luminosity without becoming radiation pressure supported? To construct a model with this property, we would have to drop one of our most cherished assumptions about accretion disks, namely that the energy that is released by gravitational infall is dissipated inside the disk and radiated. The energy is presumably stored in some other form, perhaps in twisted loops of magnetic fields which ultimately pop out of the disk, reconnect, and release the energy in the form of heat. Comptonization of soft photons in these heated regions then produces X-rays and gamma rays, generating a hard flux that shines back down on the disk and is reprocessed. In other words, the dissipation of the gravitational binding energy is occurring externally. The advantage of having a disk reprocess the hard flux is that it automatically has a large optical depth and a large covering factor, so it is easy to get the required equivalent width of the iron fluorescence line as well as the other signatures of X-ray reflection. An interesting feature of this kind of model is that there is no Eddington limit on the energy output of such a disk, because the radiation is shining down on the disk from the outside.

Alternatively, there need not be a disk at all. Rees and collaborators [2] have proposed that the cold gas forms a chaotic array of filaments in random motion. A configuration like this would probably have to be confined and supported by the energy density of very strong magnetic fields, since confinement by the thermal pressure of hot gas is difficult so close to the central black hole (see above). Confinement of individual filaments would require a magnetic energy density of at least 10^5 erg cm^{-3}, comparable to the local radiation energy density. Dynamical support against gravity would demand a magnetic energy density at least as large as the mean kinetic energy density contained in the motion of the filaments under the influence of gravity. In order for the same magnetic field strength to both confine and dynamically support the filaments, the filling factor of filaments must be quite small, of order $T_{bb}/T_{vir} \sim 10^{-5}$, where T_{vir} is the local virial temperature. To produce the observed Kα fluorescence line, this model would have to satisfy a couple of additional constraints, i.e, the filaments would have to intercept a large fraction of the hard X-ray flux and they would have to have a significant optical depth to Thomson scattering. It is completely unknown whether these constraints would be satisfied "naturally," or would require "fine-tuning".

Our search for possible alternatives to a standard accretion disk model has been motivated by the hypothesis that the observed rapid and large amplitude X-ray variability represents a global fluctuation of the central engine. Yet there is no compelling evidence that this is the case. In fact, the are no strong indications that the minimum values of Δt observed to date reflect any characteristic scale of the central engine. When we observe with higher time resolution, we may well detect even faster variability! This suggests that we should also consider models in which the variability is localized. Since the amplitude of variability is so high, the energy of a fluctuation probably has to be stored up in a small region, and released in a very intense form on a short time scale. If local flaring were the cause of variability, then we could say little about L/L_{Edd} or the mass of the black hole. Flare energy could be stored, for example, in intense magnetic flux tubes or in the kinetic energy of filaments.

6. REPROCESSING AT LARGER RADII

While the evidence for X-ray reprocessing in the inner regions of AGN seems relatively secure (but does not guarantee the presence of a disk), arguments favoring disk models for the emission at larger radii are less direct, and rely crucially on the assumption that reprocessing occurs. Shields was the first to suggest that an irradiated disk can produce the broad emission lines of an AGN

[16]. Netzer and Collin-Souffrin independently explored these ideas further [17,18]; they have been developed in the greatest detail to date by Collin-Souffrin and Dumont [19-21]. The biggest uncertainty is the extent to which the outer disk is irradiated by hard continuum from the inner regions, which can depend sensitively on the geometry (e.g., the flaring of the disk [22,23]). It is not hard for the reprocessed flux to dominate over the energy dissipated locally in the disk: all that is required is that the "reprocessing fraction" not decrease faster than $\propto R^{-1}$. However, to explain the broad emission lines the reprocessing fraction must be quite large, ~ 0.1, which places stringent demands on the geometry. Additional arguments for efficient reprocessing come from observations of the *lack* of infrared and optical variability in X-ray variable sources [24] and the IR-to-far-IR spectral shape of AGN [25].

While the most immediate implication of reprocessing is its effect on the observed AGN spectrum, it is important to keep in mind the possible *dynamical* consequences of reprocessing. These include the generation of X-ray heated winds, which has been studied in some detail [22], and the inflation of the disk's surface layers by the pressure of trapped Lyα photons, which has been explored for broad emission line clouds [26] but not for disks. The latter might conceivably enhance the efficiency of reprocessing by increasing the geometrical cross section of the disk to radiation from the central source.

7. SUMMARY

While recent X-ray observations indicate the presence of "cold" gas in the central engines of AGN, they do not tell us whether this gas is part of an accretion disk. The interpretation of rapid X-ray variability in AGN is still far from clear. If we believe that the observed fluctuations are global, then the central engines are fairly luminous compared to the Eddington limit. In this case it is difficult to see how a standard accretion disk could be responsible for producing the variable iron Kα flux by fluorescence, because the density of the disk's surface layers would be so low that they would heat up to the inverse Compton temperature. To explain the Kα flux by reprocessing requires denser gas, which is presumably not producing radiation internally. Denser gas also seems to be required to explain the absence of Lyman edges in QSO spectra [12,27]. I suggested a couple of forms that this gas might take. What is clear is that reprocessing strongly modifies the radiation that we observe, certainly in the X-ray but very likely in the UV, optical and infrared bands as well.

This work was partially supported by NSF grant AST88-16140, NASA Astrophysical Theory Program grant NAGW-766, and the Alfred P. Sloan Foundation.

REFERENCES

[1] Sun, W. H. and Malkan, M. A. 1989, ApJ, 346, 68.
[2] Guilbert, P. W. and Rees, M. J. 1988, MNRAS, 233, 475.
[3] Pounds, K. A., Nandra, K., Stewart, G. C., and Leighly, K. 1989, MNRAS, 240, 769.
[4] Nandra, K., Pounds, K. A., and Stewart, G. C. 1990, MNRAS, 242, 660.
[5] Matsuoka, M., Piro, L., Yamauchi, M., and Murakami, T. 1990, ApJ, 361, 440.
[6] Pounds, K. A., Nandra, K. A., Stewart, G. C., George, I. M. and Fabian, A. C. 1990, Nature, 344, 132.
[7] Rothschild, R. E., Mushotzky, R. F., Baity, W. A., Gruber, D. E., Matteson, J. L., and Peterson, L. E. 1983, ApJ, 269, 423.
[8] Lightman, A. P. and White, T. R. 1988, ApJ, 335, 57.
[9] Turner, T. J. and Pounds, K. A. 1988, MNRAS, 232, 463.
[10] Kunieda, H., Turner, T. J., Awaki, H., Koyama, K., Mushotzky, R., and Tsusaka, Y. 1990, Nature, 345, 786.
[11] Nandra, K. A., Pounds, K. A., Stewart, G. C., Fabian, A. C. and Rees, M. J. 1989, MNRAS, 236, 39P.
[12] Ferland, G. J. and Rees, M. J. 1988, ApJ, 332, 141.
[13] Krolik, J. H., McKee, C. F. and Tarter C. B. 1981, ApJ, 249, 422.
[14] Rees, M. J. 1987, MNRAS, 228, 47P.
[15] Rees, M. J., Begelman, M. C., Blandford, R. D. and Phinney, E. S. 1982, Nature, 295, 17.
[16] Shields, G. A. 1977, ApLett, 18, 119.
[17] Netzer, H. 1987, MNRAS, 225, 55.
[18] Collin-Souffrin, S. 1987, AstrAp, 179, 60.
[19] Collin-Souffrin, S. and Dumont, A. M. 1990, AstrAp, 229, 292.
[20] Dumont, A. M. and Collin-Souffrin, S. 1990, AstrAp, 229, 302.
[21] Dumont, A. M. and Collin-Souffrin, S. 1990, AstrAp, 229, 313.
[22] Begelman, M. C., McKee, C. F., and Shields, G. A. 1983, ApJ, 271, 70.
[23] Begelman, M. C. and McKee, C. F. 1983, ApJ, 271, 89.
[24] Done, C., Ward, M. J., Fabian, A. C., Kunieda, H., Tsuruta, S., Lawrence, A., Smith, M. G., and Wamsteker, W. 1990, MNRAS, 243, 713.
[25] Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., and Matthews, K. 1989, ApJ, 347, 29.
[26] Elitzur, M. and Ferland, G. J. 1986, ApJ, 305, 35.
[27] Czerny, B. and Pojmański G. 1990 MNRAS, 245, 1P.