ON THE TWO STATES OF CYG X-1 AND RELATED SOURCES

DEMOSTHENES KAZANAS

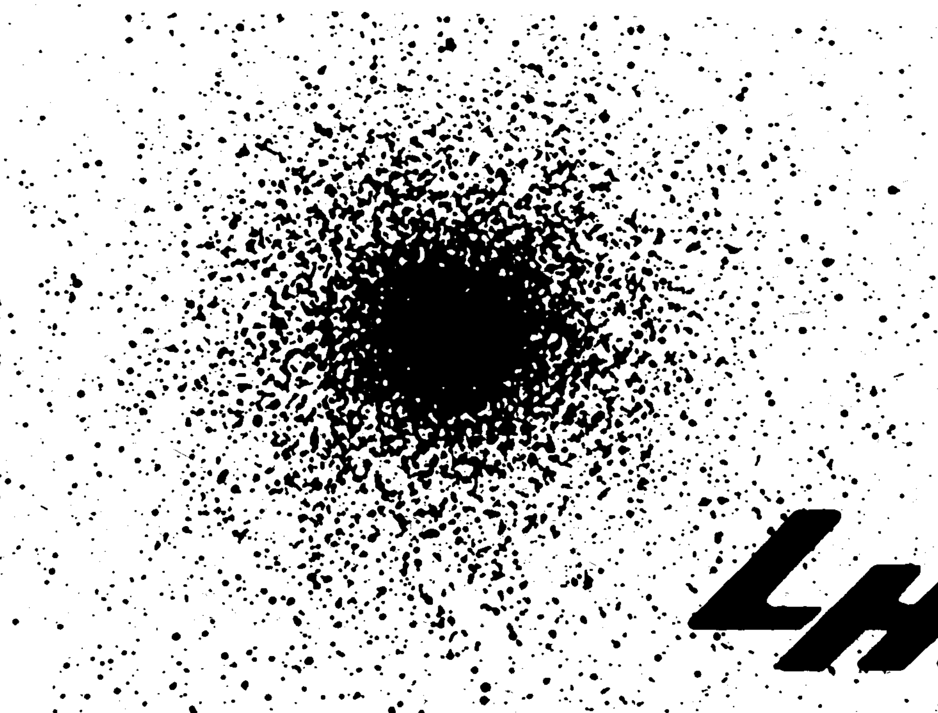
LABORATORY FOR HIGH ENERGY ASTROPHYSICS

(NASA-TM-89265) ON THE TWO STATES OF CYG X-1 AND RELATED SOURCES (NASA) 15 p CSCL 03B

N87-13367

Unclas

G3/90 44637



National Aeronautics And Space Administration

Goddard Space Flight Center

Greenbelt, Maryland 20771

86-014

ON THE TWO STATES OF CYG X-1 AND RELATED SOURCES

Demosthenes Kazanas

JULY 1986

To appear in Astronomy and Astrophysics Letters

On the two states of Cyg X-1 and related sources

Demosthenes Kazanas

Labaratory for High Energy Astrophysics, NASA/Goddard Space Flight Center, Code 665, Greenbelt MD 20771, and
Department of Physics, University of Maryland, College Park MD 20742.

Summary. A model for active galactic nuclei (AGN) (Kazanas and Ellison 1986a; hereafter KE), employing quasi spherical accretion onto a black hole, when scaled down to solar mass objects, provides a straightforward account of the bimodal spectral behaviour of Cyg X-1 and the other galactic black hole candidates. It is argued that the change in the spectrum is due to the drastic increase of the source compactness (L/R) with the accretion rate $\dot{m}$ and the subsequent conversion of most of the energy released by accretion into $e^+ e^-$ pairs. It is also argued that similar changes may be observed in active galactic nuclei.

Key words : X-ray binaries - pair plasmas - AGNs - black holes

There are by now several accreting X-ray binaries thought to derive their luminosity by accretion onto a black hole rather than a neutron star. Such a conclusion rests solely on estimates of the mass of the compact object from optical observations of the dynamics and the spectral properties of the binary system. Thus it was deduced that Cyg X-1 (Webster and Murdin 1972; Bolton 1972), LMC X-3 (Cowley et al 1983) and possibly LMC X-1 (Hutchins Crampton and Cowley 1983) are black holes. A similar conclusion was recently reached for the transient source A0620-00 (McClintock and Remillard 1986) which had an outburst in 1975 (Elvis et al 1975). On the other hand, comparative studies of accreting binaries in the X-rays, have indicated certain similarities between the spectra of black hole candidates and several other accreting X-ray binaries, thus hinting a similar nature for the compact object. The most striking such spectral characteristic is the bimodal spectral behaviour exhibited by Cyg X-1 (Sanford et al 1975; Coe, Engel and Quenby 1976). It appears that Cyg X-1 emits in two distinct states that are associated with "high" and "low" intensity in the 1-10 keV band. The "low" state is characterized by a power law of energy index $\alpha \sim 0.5$ similar to that observed in AGN (Rothschild et al 1983) extending to 100 keV. Above that energy there appears to be a roll-off or a break in the spectrum (see fig. 4 of Liang and Nolan 1984 and refernces therein). The "high" state is characterized by an increased flux in the 1-10 keV band (by a factor 10-30) with a much softer spectrum, which can be fitted either with a power law of $\alpha \sim 2$-4 or a thermal specrtum of $kT \sim 1$-2 keV, with an indication of a power law tail extending to higher energies. It has been thus conjectured (White, Kaluzienski and Swank 1984; White and Marshall 1984; White Fabian and Mushotzky 1984, hereafter WFM) that several other sources which show this spectral bimodal behaviour are also related to accretion onto a black hole rather than a neutron star. Besides Cyg X-1 these sources are Cir X-1 (Jones 1977), and GX339-4 (Jones 1977; Ricketts 1982). In addition, the above mentioned X-ray transient A0620-00 (which is now known to be a black hole) and the transient A1524-62 (Kaluzienski et al 1975) show the same bimodal behaviour during their rising phase. At low fluxes their spectra are hard and similar to that of Cyg X-1 in the low state; as the luminosity increases they become significantly softer, quasi-thermal with temperature $kT \sim 1$-2 keV. The other black hole candidates LMC X-1 and LMC X-3, showing always a soft spectrum, appear to be permanently locked in the high state.

To date no satisfactory theoretical interpretation of such a behaviour exists, in the sense that it is associated in a coherent fashion to the dynamics responsible for the energy release and the emission of radiation. In the framework of unsaturated Comptonization (Shapiro, Lightman and Eardley 1976; Katz 1976; Liang 1980) it is attributed to variations in the Comptonization parameter caused either by changes in the accretion disk (Shapiro and Lightman 1976), or variations in the flux of soft photons (Guilbert and Fabian 1982). More recently, WFM have suggested that the change in the spectrum may be due to production of $e^+ e^-$ pairs in the source. In this note we present a simple, straightforward model which produces this bimodal spectral behaviour as a result of the dynamics of accretion onto a black hole and an ensuing $e^+ e^-$ cascade as in WFM. This model has been proposed originally for the central engine of AGN (KE); however, since in this model, the luminosity, the source size and the accretion rate scale linearly with the mass of the black hole, the same model should be applicable to solar mass black holes in accreting binaries. At present we will review the basic concepts and features of the model, referring the reader to KE for more details.

The model considers (quasi)spherical accretion and a collisionless shock as a means for randomizing the inflowing kinetic energy. At the same time the shock converts, via the first order Fermi shock acceleration, a substantial fraction of the energy flux into relativistic protons. The latter, since their kinetic energy is much larger than the gravitational potential, resist their prompt accretion into the black hole and their energy density can increase significantly to self-consistently sustain the shock. This is possible if the nuclear collision time scale t_{pp} (the main energy loss for relativistic protons) is longer than the local free-fall time t_{ff}. One is thus led to the picture of a "relativistic proton radiative shock", whereby the accumulated relativistic protons at the base of the flow create a shock at a radius $R \sim v_{ff}\, t_{pp}$, in analogy with accretion shocks on white dwarfs. The dissipation mechanism in this model is therefore well understood and the corresponding time scale, t_{pp}, is inversely proportional to the ambient density. This latter fact assures that the luminosity, source size and accretion rate scale linearly with the mass of the black hole and allows the same model to be applied both to QSOs and accreting binary sources (Rees 1984). The basic relations of the model are shown graphically in figures 4a and 4b of KE where the luminosity, F, (in units of L_{Edd}) is given as a function of the radius, x_1 (in units of $R_S = 2GM/c^2$) and as a function of $\dot{m}/M_9$. (The two curves $T_8 = 4$ or 1 refer to the upstream temperatures in units of 10^8 K, a typical value as argued in KE). The model also provides the electrons necessary to produce the observed radiation as secondaries of nuclear collisions, resulting from the decay of $\pi^{\pm}$'s. Therefore, since most of the energy is injected well above $m_e c^2$ the model offers the possibility of vigorous $e^+ e^-$ pair production, due to photon-photon collisions in the source. This is a most important process in determining the spectrum as it will be discussed later. The optical depth of a photon with $E \sim m_e c^2$ to this process depends mainly on the ratio L/R (called the compactness) of the source and it becomes greater than unity for L/R greater than the critical value $(L/R)_c \sim 10^{29}$ erg/s/cm. In the dimensionless units of the model the condition on L/R reads $(F/x_1)_c \sim 2\ 10^{-4}$. One can now show the combined effects of the dynamics, i.e. F, and of pair production, i.e. F/x_1 , as a function of the dimensionless accretion rate $\dot{m}/\dot{m}_E$ (or $\dot{m}/M_9$) . These relations are shown in figs. 1a and 1b. The most important feature of these figures is the strong dependence of F and especially of F/x_1 on $\dot{m}$. Changes of only 10%-50% in the accretion rate can change the luminosity F by a factor 10 - 30 and the compactness F/x_1 by a factor ~ 10^3. As explained in KE this is due to the sensitivity of the shock acceleration efficiency on the Mach number of the shock. This sensitivity on $\dot{m}$ can then account for the change in luminosity from the "low" to the "high" states. Similarly, one can see that it is easy to "lock" the source in the high state if $\dot{m}/\dot{m}_E > 4$. Assuming the canonical value for the mass of the black hole (~ 10 M_0 ; Webster and Murdin 1972; Bolton 1972), the luminosity of 10^{38} erg/s in the high state suggests F ~ 0.1, while for the "low" state F ~ 3 10^{-3} - 10^{-2}. These values are shown by the the appropriate horizontal lines in fig. 1. The value for the low state is very similar to that corresponding to the Seyfert-1 galaxy NGC 4151 as determined in KE. The size of the source in the "low" state (see fig 1a) is then R ~ 100 R_S ~ 3 10^8 cm which implies a variability time scale of 10^{-2} s in good agreement with observation (Meekins et al 1984). Variability down to 3 ms has been reported for Cyg X-1 by the latter authors and also by Rothschild et al (1974). However, the variability on these time scales is statistically much less significant, and also most of the power is in the ~ 10 ms time scales (fig. 3 of Meekins et al 1984), which is most likely associated with the large scale dynamics of the source. Furthermore, it is interesting to note that the shot noise model of Meekins et al (1984) with decay times between 3.5 ms and 300 ms is in agreement with the proton acceleration model of KE, if these time scales are identified with the p-p collision time scales (eq. (4) of KE) at R ~ 5 R_S and R ~ 100 R_S respectively. Therefore, in this simple picture, the variability is accounted for by the continuous, impulsive injection of relativistic protons, (the shots); the decay time of the shots is indicative of the density of matter in which these protons interact (producing the radiation through hadronic collisions), while most of the power is emitted on time scales characteristic of the size of the shock (~10 ms). Given the simplicity of the assumptions involved, such an interpretation appears quite satisfactory. Finally, the absense of apparent variability in the "high" state may simply mean that a soft comparatively constant component may mask the variability of this state (Meekins et al. 1984).

The spectra of the source in the two different states can also be accounted for in terms of the non-thermal particle injection (through pion decay) suggested by the model and the ensuing $e^+ e^-$ pair cascade induced in the source. For the "low" state, $F/x_1<(F/x_1)_C$. This situation has been discussed in detail in connection with the spectra of AGNs in the framework of non-thermal particle injection (Kazanas 1984; Zdziarski and Lightman 1985). As shown in these references, the spectra in the X-ray band should be power laws with energy index $\alpha \sim 0.5$ breaking at higher energies to $\alpha \sim 1$. This is consistent with the observed spectrum of Cyg X-1 and the similarity of the spectra of Cyg X-1 and NGC 4151 is naturaly attributed in the similar values of F/x_1 in the two objects. In the "high" state, the increase in the luminosity by tenfold, results in an increase of the compactness by ~ 10^3 (fig. 1a). Because the injection of energetic electrons takes place through the hadronic channel, it is unaffected by the increase in the luminosity of the source. The drastic change in the spectrum is due to change in the steady state distribution function, resulting from the increase in the source compactness. Since in this case $F/x_1>(F/x_1)_C$, the optical depth of a photon of energy $E \sim m_e c^2$ to γ-γ pair production is much greater than one. As a result, virtually all the power (injected at $E \gg m_e c^2$) is converted into $e^+ e^-$ pairs. The cooling time of the pairs is shorter than their annihilation time scale and the pair distribution develops a pronounced peak at energies ~ kT <<

$m_e c^2$, while it is a power law of index $p = -3$ at higher energies (Kazanas 1984). The cool pair density can then be determined by the balance between injection and removal of the cool pairs by annihilation (the fastest removal process). The Thomson optical depth of the resulting plasma can then be determined from the above considerations and is found to be

$$\tau_T \simeq 2\ (2\ r)^{0.5} \qquad (1)$$

where r is the ratio of the compactness parameter of the source to the critical one. As seen in fig 1 , $r \sim 100$ for the high state and hence the resulting Thomson optical depth should be of the order of 20 - 40. This is several orders of magnitude larger than the optical depth of the ambient accreting plasma, assuming spherical accretion. The temperature of this plasma is determined by demanding that a given luminosity is emitted from a source of a given size (determined by the model) and a given optical depth (~ 20 - 40). For the values corresponding to Cyg X-1 the temperature is $kT \sim 2$ keV, in agreement with the observations of the "high" state spectrum. The change in the spectrum is therefore due to the conversion of essentially all the luminosity into a cool, optically thick pair plasma whose temperature, because of the large optical depth, is determined by the thermodynamics of the source. This temperature is much lower than that found by Svensson (1984; fig. 6a) for plasmas in pair equilibrium. The lowest temperatures derived in the latter reference were $kT \sim 0.1\ m_e c^2 \sim 50$ keV, while in the present case there does not seem to be a lower limit in the temperature of the resulting pair plasma. The reason of this, perhaps counterintuitive at first sight, result lies in the different energization processes of the plasmas in the two cases. While in the thermal plasmas the temperature is assumed fixed and the pairs are produced by the high energy tails of the thermal electron distributions, in the present case the pairs are injected as such through the hadronic channel continuously; the resulting temperature is hence allowed to be $kT << m_e c^2$. The pairs eventually annihilate and the annihilation photons quickly downscatter in energy in the optically thick plasma. Most of the energy is eventually emitted in the low energy "hump" observed in the "high" state of Cyg X-1. The observation of an annihilation feature from this plasma is an interesting question. WFM argue that such a feature could be observed only if the pairs escape slow down and annihilate in a cool medium. Annihilation in the thick plasma is smeared by Compton scattering, while annihilation in a wind is smeared by Doppler broadening. The conditions for an observable feature might be more favourable in the "low" state, since the optical depth is much smaller; in this case of however the annihilation emissivity is also smaller and the width of the line would be larger because of the higher temperature, thus making the detection of such a feature more difficult. (However, Nolan and Matteson (1983) suggest that such a broad feature has been observed in the HEAO-1 data). Finally, one should bear in mind that a narrow feature has indeed been observed from the direction of the galactic center (Leventhal et al 1978).

The thermodynamics of the source, which determine the temperature of the pair plasma, as argued above, scale differently than the dynamics responsible for the radiation emission (i.e. proportionally to the mass as shown in KE). It is therefore expected that the temperature of the plasma will depend on the mass of the black hole. Although no precise scaling laws can be derived, in general, the temperature will be smaller for the higher mass sources, despite the fact that the optical depth to pairs, given by eq (1), depends only on the value of F at which the particular source is operating. The dependence of the mass is rather weak (e.g. fixing F/x_1 and assuming black body emission $T \propto M^{-1/4}$). Estimates of the temperature for sources emitting 10^{45} - 10^{47} erg/s indicate values ranging from 10^5 to 10^7 degrees, depending on the compactness of the source. If therefore a similar pair plasma exists in AGN it should manifest itself as a quasi thermal component of temperature ~ 10^5 to 10^7 K. Such a feature is known to exist in AGN and QSO. It has actually been attributed to emission from the accretion disk of these objects and is known as the "UV excess" or "blue bump" (Malkan 1983; Malkan and Sargent 1982). In the simple, unified picture presented here, this feature should be due to e^+ e^- pairs and objects with pronounced "bumps" should indicate that these objects are locked in the equivalent of the "high" state of Cyg X-1, just like LMC X-3 (there exists of course the possibility that both the e^+ e^- pair plasma and an accretion disk contribute to the "bump" emission). This view is different from that of WFM who suggest that the equivalent of the "high" state in AGN is the BL Lac phenomenon. In principle one should be able to decide between the two interpretations (if a single one is correct at all) by monitoring the long term spectral behaviour of AGNs. According to the WFM conjecture, a increase (decrease) in the luminosity of a Seyfert-1 (BL Lac) could possibly turn it into a BL Lac object (Seyfert-1). To the best of the aurthor's knowledge no such transitions have been recorded so far. According to the conjecture proposed in the present note, and in accordance with the KE model, an increase (decrease) in the luminosity should cause the appearance (disappearance) of the quasi thermal feature referred to as the "blue bump". The author is aware of at least one object (the Seyfert-1 galaxy Fairall 9) which behaves qualitatively in the way outlined by the present model. Extensive monitoring of this object (Morini et al 1986) has shown that its luminosity has been decreasing steadily over the past three years. An apparent "blue bump" component was identified in the UV and soft X-rays, along with power law components in the optical and hard X-rays. During the overall luminosity decrease, the "blue bump" component decreased by a substantially larger factor than the non-thermal power law components, as suggested by the model (decrease in the bolometric luminosity causes a corresponding decrease in the pair content and hence the importance of the quasi thermal feature; since however the importance of pairs, which depends on F/x_1 decreases faster than F, which reflects the bolometric luminosity, the importance of the "bump" relative to the power law should decrease).

A final point concerns the application of the model in accreting binaries involving neutron stars rather than black holes. Clearly, the dynamics should be different because of the existence of the star's magnetosphere, which now sets the position of the shock; therefore one would not expect the same sensitive dependence of F/x_1 on $\dot{m}$ and hence the bimodal spectral behaviour (a fact consistent with observations). However, the non-thermal energy injection arguments and the resulting pair equilibrium should still be valid. Interestingly, the spectra of these sources require the existence of two components, a black body one (associated with emission from the neutron star) and a thin thermal bremsstrahlung component of emission measure $n^2 R^3 \sim 10^{60}$ cm^{-3} (Swank and Serlemitsos; White et al 1986). It is interesting that eq (1), despite the simplicity of the arguments used in deriving it, indicates an emission measure $\sim 10^{58}$ cm^{-3}, without any further assumptions. This value is actually compatible with that given above, if one considers that in the optically thick plasmas considered here the emissivity is higher than the bremsstrahlung one and hence the required emission measure should be smaller. Clearly, further discussion on this point, though important, is beyond the scope of the present note. Finally, the discovery of VHE (10^{12} eV) and UHE

(10^{15} eV) emission from these objects, are in accord with the arguments of acceleration and non-thermal injection presented here and elswhere (e.g. Kazanas and Ellison 1986b and references therein).

It is a pleasure to acknowledge discussions with T. Kallman, R. Kelley, F. Marshall, J. Meekins and J. Swank.

REFERENCES

Bolton, C. T., 1975, Ap. J., 200, 269.

Coe, M. J., Engel, A. R. and Quenby, 1976, Nature, 259, 544.

Cowley, A. P.,Crampton, D., Hutchings, J. B., Remillard, R., and Penfold, J. E., 1983, Ap. J., 272, 118.

Elvis, M. et al, 1975, Nature, 257, 656.

Guilbert, P. W., and Fabian, A. C., 1982, Nature, 296, 226.

Jones, C., 1977, Ap. J., 258, 335.

Hutchins, J. B., Crampton, D. and Cowley, A. P., 1983, Ap. J.(Lett.), 275, L43.

Kaluzienski, L. J., Holt, S. S., Boldt, E. A., Serlemitsos, P. J., Eadie, G., Pounds, K. A., Ricketts, M. J. Watson, M., 1975, Ap. J., 201, L121.

Katz, J. I., 1976, Ap. J., 206, 910.

Kazanas, D., 1984, Ap. J., 287, 112.

Kazanas, D. and Ellison, D. C., 1986a, Ap. J., 304, 178.

Kazanas, D. and Ellison, D. C., 1986b, Nature, 319, 380.

Leventhal, M., MacCallum, C. J. and Stang, P., 1978, Ap. J.(Lett.), 225, L11.

Liang, E. P. T., 1980, Nature, 283, 642.

Liang, E. P. T. and Nolan P. L., 1984, Space Sci. Rev., 38, 353.

Malkan, M. A. and Sargent, W. L. W., 1982, Ap. J.,254, 22.

Malkan, M. A., 1983, Ap. J., 268, 582.

McClintock, J. E. and Remillard, R. A., 1986, Ap. J., 308, (in press).

Meekins, J. F., Wood, K. S., Hedler, R. L., Byram, E. T., Yentis, D. J., Chubb, T. A., and Friedman, H., 1984, Ap. J., 278, 288.

Morini, M. et al, 1986, Ap. J., in press.

Nolan, P. L. and Matteson, J. L., 1983, Ap. J., 265, 389.

Ricketts, M. J., 1982, Astr. Ap., 118, L3.

Rothschild, R. E., Boldt, E. A., Holt, S. S. and Serlemitsos, P. J., 1974, Ap. J. (Lett.), 189, L13.

Rothschild, R., Mushotzky, R. F., Baity, W. A., Gruber, D. E., Matteson, J. L. and Peterson, L. E., 1983, Ap. J., 269, 423.

Sanford, P. W., Ives, J. C., Bell-Burnell, S. J., Mason, K. O., Murdin, P., 1975, Nature, 256, 109.

Shapiro, S. L., Lightman, A. P., and Eardley, D. M., 1976, Ap. J., 204, 187.

Shapiro, S. L., and Lightman, A. P., 1976, Ap. J., 204, 555.

Svensson, R., 1984, M.N.R.A.S., 209, 175.

White, N. E., Kaluzienski, L. and Swank, J. E., 1984, in "High Energy Transients in Astrophysics", AIP Conf. Proc. 115, p.31. S. Woosley editor.

White, N. E. and Marshall, F. E., 1984, Ap. J.,281, 354.

White, N. E., Fabian, A. C. and Mushotzky, R. E, 1984, Astr. Ap.,133, L9.

White, N.E., Peacock, A., Hasinger, G., Mason, K. O., Taylor, B. G. and Branduardi-Raymont, G., 1986, M.N.R.A.S., 218, 129.

Webster, B. L. and Murdin, P., 1972, Nature, 235, 37.

Zdziarski A. A. and Lightman, A. P., 1985, Ap. J.(Lett.), 294, L79.

Figure 1. The dimensionless compactness parameter, F/x_1 , versus (a) the Eddington efficiency F and (b) the dimensionless accretion rate $\dot{m}/\dot{m}_E$. The dot-dash line indicates the critical compactness, while the dashed lines the "high" and "low" states of Cyg X-1 and NGC 4151.