

Towards a Truly Unified Model of AGN: Aspect, Accretion and Evolution

Michael A. Dopita

Mt Stromlo and Siding Spring Observatories, Institute of Advanced Studies,
Australian National University, Private Bag, Weston Creek PO, ACT 2611, Australia
Michael.Dopita@anu.edu.au

Received 1997 May 1, accepted 1997 August 1

Abstract: The Unified Model holds that the aspect-dependent effects primarily determine the nature of the active galactic nucleus that we observe. In this paper, I argue that three parameters; aspect, accretion rate into the nuclear regions, and the evolutionary status of the central black hole hold the key to unification. The mystery of why the epoch of quasar formation occurred so early in the evolution of the Universe, why radio-loud QSOs represent only a small fraction of the general population of QSOs, and why ellipticals are invariably the hosts of radio-loud active galaxies could be explained if (a) the most rapid growth of black holes occurred in galactic merger events, and if (b) an excess in the rate of nuclear feeding was able to choke off the radio jets, producing radio quiet QSOs. In this paper, I develop the idea that rate of nuclear feeding plays a dominant role and that feeding at super-Eddington rates into the broad-line region (BLR) during merger events is the means whereby massive black holes are grown. In particular, I develop a toy model for the radio-loud, radio-quiet dichotomy based on the rate of nuclear feeding, suggest an electron scattering model for the ‘big blue bump’ and its relation to the BLR, and emphasise the important diagnostic capabilities offered by analyses of the narrow line regions based on shock excitation models.

Keywords: accretion — galaxies: active, evolution, interactions — quasars: general

1 Introduction

It is now universally believed that active galactic nuclei (AGN) derive their extraordinary luminosities from energy release by matter accreting towards, and falling into, a central supermassive black hole (BH). This process releases energy as UV, X-ray, and γ -ray continua, and material orbiting close to the BH is photoionised, producing the Doppler-broadened emission lines characteristic of the ‘broad line region’ (BLR). In the ‘radio-loud’ subclasses of AGN, relativistic jets of material are ejected in the poleward direction and may escape into intergalactic space as the classical double-lobed radio sources. Also lying generally in the poleward directions, but at a distance of up to several kpc, are found the ‘narrow line regions’ (NLR) excited either by photoionisation from the UV continuum of the central source (Koski 1978; Ferland & Netzer 1983; Stasinska 1984) or by shock excitation related to the jets (Dopita & Sutherland 1995). In addition, some AGN display low-ionisation nuclear emission line (LINER) nuclei which differ quantitatively from the NLR of Seyfert galaxies, and which, if photoionised, are excited by a much more dilute UV radiation field.

Whilst this gave a general understanding of the components which make up an AGN, the taxonomic

relationship between the various sub-classes of AGN remained very unclear and the ‘unified model’ was developed in response (Rowan-Robinson 1977; Lawrence & Elvis 1982; Antonucci & Miller 1985; Lawrence 1987). Our current understanding of the unified model has been summarised in two excellent reviews (Antonucci 1993; Urry & Padovani 1995) which represent essential reading in the field. To summarise, the basic idea behind the unified model is that orientation effects exercise a profound influence on the appearance of the AGN, and therefore determine into which particular sub-class it will be categorised. The principal cause of the orientation effect is believed to be due to a geometrically thick and optically-thick dusty molecular torus of material located in the accretion plane of the AGN. This is capable not only of obscuring the central source from direct view, but may also be optically-thick to the infrared photons produced by heating of the dust within it (Heisler, Lumsden & Bailey 1997).

A second orientation effect is found in the radio-loud categories, for which the unification concept was originally developed (Orr & Browne 1982; Barthel 1981). The classes of AGN which are observed close to the polar direction display apparently super-luminal motions, rapid variability, intense power-law continua, flat radio spectra and high

linear polarisation all due to relativistic beaming effects associated with the jet. The BL Lac objects, and the optically-violently variable (OVV) or flat-spectrum radio quasars (FSRQs) all seem to fit into this category (Blandford & Rees 1978; Blandford & Konigl 1979). It is not so clear to which class of radio-loud objects these various beamed objects belong. Urry & Padovani (1995) would argue that BL Lacs belong in the Fanaroff–Riley (FR) 1 class. However, recent work by Wurtz et al. (1997) shows that the clustering properties of BL Lac objects is like that of the radio-loud quasars and the (more luminous) FR 2 double lobe sources.

The ‘strong’ unification hypothesis, that all AGN classes are fundamentally the same phenomenon seen at different orientations, has not proven successful. However, in its more restricted form; that some classes of object transform into other classes as a function of viewing angle, it now seems secure. In particular, the following relationships between the Seyfert nuclei found in disk galaxy environments; the quasi-stellar objects (QSOs) found in more distant systems; and the weaker radio sources found in luminous elliptical galaxies are almost certainly correct:

Seyfert I $\Leftrightarrow$ Seyfert II

QSOs $\Leftrightarrow$ Ultraluminous IRAS narrow
line galaxies

BL Lac/Blazar $\Leftrightarrow$ Double lobe radio
elliptical galaxy.

Whilst this scheme is satisfying as far as it goes, it offers no way to fully unify the various tribes of AGN. In particular, there is no explanation of the radio-loud/radio-quiet or the Fanaroff–Riley class FR 1/FR 2 dichotomies. Clearly, a parameter other than orientation must play a role. It has been variously argued that radio loudness is related to host galaxy type (Smith et al. 1986), black hole spin (Blandford 1990; Wilson & Colbert 1995), or to differences in the rate of nuclear feeding (Rees et al. 1982; Baum, Zirbel & O’Dea 1995).

Nuclear feeding is certainly important in the evolution of AGNs. For example, Hutchings & Neff (1992), Boyce et al. (1996) and the work of Bahcall and his group using HST images (Bahcall et al. 1997) show that QSO host galaxies are, in the main, either interacting or photometrically disturbed post-merger systems. The luminous infrared galaxies (LIRGs) discovered by the IRAS satellite seem to represent the link between spiral and elliptical galaxies. It is now clear that they form as a result of collisions and mergers, and most of those objects which are well resolved display both double nuclei and tidal arms (Melnick & Mirabel 1990). The IR emission

may amount to 90% of the bolometric emission, and initially results from dust heated by the light from a massive burst of star formation induced by gas dynamical collisions. In the luminosity range $10 < \log(L_{IR}/L_{\odot}) < 11$, about 70% of all objects show line ratios typical of normal HII regions (Veilleux et al. 1995) and resolved objects display double nuclei with projected separation in the range 5–30 kpc. However, later on in the merger a greater and greater portion of the luminosity appears to be derived from energetic outflows powered by accretion to an active nucleus. These objects are found amongst the most luminous LIRGs (Saunders & Mirabel 1996). For $\log(L_{IR}/L_{\odot}) > 12$, less than 30% of objects show HII region-like characteristics in the optical, over 40% have some kind of Seyfert activity in the nucleus, and the balance show line ratios characteristic of a mixed excitation (see Armus, Heckman & Miley 1989, 1990; Veilleux et al. 1995; Kim 1995). Late starburst and Seyfert-like objects are typically at a late merger stage and may show the development of the $R^{-1/4}$ law of photometric profiles (Wright et al. 1990), multiple shells, evolved tidal arms, and (where still separate) projected distances between multiple nuclei < 1 kpc. These objects often appear to have warmer dust, as measured by $60\ \mu\text{m}/25\ \mu\text{m}$ and $60\ \mu\text{m}/100\ \mu\text{m}$ flux ratios. We therefore believe that warm IRAS galaxies represent merger systems in which the starburst is initially triggered by tidal interactions, is powered through the merger event by the direct collision of gas streams from the merging galaxies, and that in the late phases gas is accreted into the central regions to power the active nucleus or nuclei. The question of whether gas-rich mergers may lead to the formation of elliptical galaxies is not fully resolved. Nonetheless, van Albada (1982) showed that the $R^{-1/4}$ law is an inevitable result of violent relaxation, Scoville et al. (1991) demonstrated that the molecular gas in merging systems has reached surface densities $\geq 10^2 M_{\odot} \text{pc}^{-2}$ which would be necessary to form an elliptical-like core, and Kormendy & Sanders (1992) showed that a number of starbursts lie in the region of the ‘cooling diagram’ (the surface density versus velocity dispersion plot) occupied by ellipticals.

If mergers drive the formation of elliptical-like systems today, then the much higher rate of gas-rich mergers expected in the early Universe ($z \sim 2-3$) would have ensured that the major epoch of elliptical formation occurred at that time. Since statistical studies suggest that this was the epoch of the maximum space density of QSOs, it is tempting to associate QSO formation with this gas-rich merger epoch.

In this paper, I investigate the idea that rate of nuclear feeding plays the dominant role and that, moreover, feeding at super-Eddington rates into the BLR occurring in merger events is the means

whereby massive BHs are grown. In particular, I develop a simple parametric characterisation of the growth of a BH during merger events in order to establish the credibility of the hypothesis that the nuclear accretion rate is a key parameter in unification of AGN phenomena. I also present a model for the radiatively driven winds and use this to suggest a model for the ‘big blue bump’ in AGN spectra and its relation to the BLR, investigate the radio-loud, radio-quiet dichotomy based on the rate of nuclear feeding, and emphasise the important diagnostic capabilities offered by analyses of the NLR based on shock excitation models.

2 The Role of Accretion

Let us take as our basic ansatz that the growth of a black hole (BH) to produce a quasar, and subsequently, a massive black hole within an elliptical galaxy is the result of a merger of a pair of gas-rich systems, such as considered most recently by Barnes & Hernquist (1996). In addition, assume a detailed symbiosis between the properties of the accretion disk and the properties of the radio jet, inspired by the work of Falcke & Biermann (1995). In this model, the nature of the active galactic nucleus (AGN) that is produced depends critically upon the rate of accretion into the nuclear regions. This type of a model would then imply a number of consequences which we summarise in this section, and develop (to greater or lesser degree) in subsequent sections.

In the sub-Eddington accretion domain we may obtain a classic Shakura & Sunyaev (1973) thin disk. Such disks allow the free escape of the relativistic jet (assumed to be light; electrons+positrons) from the nuclear region, and may produce optical continuum by thermal emission, but are expected to be ineffective in producing either a broad-line or a narrow-line emission region. However, the radio jets in such systems are characterised by efficient relativistic boosting, and can therefore give rise to blazar phenomena. The sub-Eddington (and possibly, advective) nature of the accretion will ensure that such sources are of relatively low luminosity, and of low radio power, and such objects might be identifiable with the FR I radio sources as is indeed postulated in the unified models of radio sources (Urry & Padovani 1995).

At very high (super-Eddington) accretion rates into the nuclear regions, the accumulation of dusty matter in the nuclear regions tends to obscure the central engine. This can be identified with the dusty ‘thick’ torus of the unified models. Indeed, Dopita et al. (1997a) have demonstrated that since this thick dusty accretion torus is optically thick even at $25\ \mu\text{m}$, its properties can be investigated through IR colour-colour diagrams, at least for the hidden broad-line region (HBLR) Seyferts. In these objects the flow a few degrees outside the ionisation

cone is found to be highly super-Eddington. Dopita et al. (1997a) also point out the importance of the radiation pressure acting on the dust, since the opacity of dusty gas is so much higher than that due to free electrons. In this case, at least some of the inflowing material will be driven back out by radiation pressure. Nonetheless, at least some of the flow makes it past the sublimation point of the dust, thanks to the ram pressure of the inflow. If this flow into the broad-line region is also super-Eddington, much of the matter entering these regions has to be lost in the form of a thermal wind. Murray and his collaborators (Murray & Chiang 1995; Murray et al. 1995; Chiang & Murray 1996) have shown that this can be analogous to the radiatively-driven winds produced in hot stars, being optically thick to UV resonance line scattering, and therefore accelerated by radiation pressure. For high enough mass-loss rates, such a wind may become optically-thick to electron scattering as well, and a hot electron-scattering photosphere can be produced. Material below this photosphere serves to reprocess the hard radiation (EUV, X- and γ -rays) from the central BH and the innermost portions of the accretion disk, and this would serve to obscure the central engine from direct view except in the polar directions where the relativistic jets may escape. Such a toroidal reprocessing photosphere would provide enough absorption to explain the weakness of Seyfert 2 galaxies in X-rays (Mushotzky, Done & Pounds 1993, and references therein), and its inner surface could also be used for K- α scattering. Since multiple Compton scattering degrades harder photons to softer photons (analogous to the degradation of radioactive γ -rays in supernova fireballs), such processing can also make the photosphere an efficient source of UV photons (big blue bump?) which are then available to ionise the surrounding accretion disk and produce the broad-line region (BLR). In this model, an AGN seen in an intermediate angle (but outside the ‘thick’ dusty torus) would appear as either a QSO or as a Sy I galaxy, depending on the mass of the nuclear BH.

Models which include such a thermal radiation-driven wind allow the possibility of a direct interaction between the thermal and relativistic winds. Clearly, this interaction is favoured when the accretion into the BLR is highly super-Eddington with respect to the nuclear BH. When such interaction occurs below the reprocessing photosphere, it will lead to mass entrainment into the jet, and a slowing of the jet to highly sub-relativistic speeds. This is likely to be the cause of the radio-loud, radio-quiet dichotomy. If radio-loud galaxies can only be produced when the accretion rate to the BH is very much below the Eddington value, then radio-loud QSOs will only be produced in the later stages of such mergers, or when the massive BH subsequently swallows more matter

such as in a merger of the elliptical with a small gas-rich system, or when a extremely sub-Eddington cooling flow feeds down to the elliptical nucleus.

2.1 Nuclear Accretion Rates

There seems little doubt that galactic merger events can dump large amounts of gas ($\sim 10^{10} M_\odot$) into the nuclear regions (e.g. Solomon, Downes & Radford 1992). The process of gas dynamics in the case of a galactic merger has been considered by a number of authors, but most notably by Barnes & Hernquist (1991, 1992, 1996). In all of these computations, the effect of torques and dissipation in the gas is dramatic, and strong inflow towards the centre of the gravitational potential is produced. At the point where orbital support of the gas becomes important, a strong accretion shock is formed, which is likely to trigger a major nuclear starburst. In the simulations of Barnes & Hernquist (1991), about $5 \times 10^9 M_\odot$ of gas found its way to the central ~ 200 pc in a timescale of $\sim 10^8$ years. This process may continue to still smaller scales; see, for example, the simulations of Bekki (1995) where he considered the gas dynamics of the material within 1 kpc.

In the core region, the infalling gas is processed through an accretion shock and joins the already accreted material in a dense, rotating disk. Applying pressure balance across the accretion shock, we can set the ram pressure in the accretion flow to the gas pressure in the hot phase of the ISM; $P \sim 5 \times 10^{-7} \dot{M}_{200} v_{300} R_{300}^{-2} \omega_{in}^{-1}$ where R_{300} is the core radius in units of 300 pc, $\dot{M}_{200}$ is the rate of mass accretion in units of $200 M_\odot \text{ yr}^{-1}$, v_{300} is the velocity of infall in units of 300 km s^{-1} , and ω_{in} is the solid angle covered by the accretion flow. In this core region, we assume a two-phase medium is formed with a hot intercloud medium, and with dense self-gravitating molecular clouds limited by their tidal radius in the core potential. Thanks to their self-gravity, the molecular gas clouds have a mean internal density ρ_I , which is greater than their surface density P/c_I^2 , where c_I is their internal sound speed ($\sim 1 \text{ km s}^{-1}$). This translates to a H-number density in excess of 10^6 cm^{-3} .

Although the gas in the core is in a very compact region, it still needs to decrease its orbital angular momentum by a factor of $\sim 10^4$ relative to the BH if it is to be efficiently accreted. However, in post-merger galaxies this may be somewhat easier than at first appears. If a BH typical of those found in Seyfert galaxies ($\sim 10^7 M_\odot$) is present in the gas-rich region, initially its sphere of influence is quite small, and it will be attracted towards the most massive gaseous complex. Dynamical friction, which is high in the case of the dissipative accretion interactions the BH makes with the cloud, should then allow it to settle towards the dense cloud core

of the complex. In this way, the BH positions itself within a dynamical timescale ($\sim 10^6$ years) to sit at the densest point in the accretion flow where it can be optimally fed. The accretion rates for moving BHs have been derived both analytically (Petrich, Shapiro & Teukolsky 1988) and through supercomputer simulations (Petrich et al. 1989) to which the interested reader is referred. However, the basic physics is easily understood and was given by Bondi & Hoyle (1944). Inside a molecular cloud the accretion radius R_A and the Bondi–Hoyle mass accretion rate onto the BH is given by

$$R_A = GM_{BH}/v_{BH}^2, \\ \dot{M} = \pi R_A^2 \rho_I v_{BH}, \quad (1)$$

where v_{BH} is the relative velocity of the BH through the molecular cloud. This velocity will not be greater than the orbital velocity of the cloud around the core. Substituting numerical values for the time when the BH passes through a molecular cloud, we get $R_A \sim 0.5 M_7 v_{300}^{-2} \text{ pc}$, where M_7 is the mass of the BH in units of $10^7 M_\odot$, and v_{300} is now the relative velocity in units of 300 km s^{-1} . Therefore, given that the relative velocity is an upper limit, and the density ρ_I , defined above, is a lower limit, we get $\dot{M} > 7 M_7^2 v_{300}^{-3} n_6 M_\odot \text{ yr}^{-1}$, where n_6 is the H-atom number density in the molecular cloud in units of 10^6 cm^{-3} . This accretion rate is well above the Eddington accretion rate for any reasonable value of BH mass, and will increase rapidly as dynamical friction slows the relative velocities of cloud and BH. Although equation (1) suggests that the radius of attraction can become very large for small values of the relative velocity, in practice it is limited by the tidal radius, or radius of influence of the BH. For the case we are considering, this is ~ 25 pc between molecular clouds and ≤ 5 pc inside molecular clouds. It would therefore appear that the BH can move towards and ‘graze’ efficiently on the most massive molecular clouds in its vicinity. However, outside the molecular clouds, in the hot intercloud medium, the rate of accretion is several orders of magnitude lower. Therefore, super-Eddington accretion is only turned on when the BH enters a molecular cloud.

The feeding process in non-merger systems is likely to be somewhat different from this, and we would probably have to depend on material (stars or gas) which is either on a radial orbit or else has been driven into the nuclear regions as the consequence of the development of a bar-like potential (Simkin, Su & Schwartz 1980; Shlosman, Frank & Begelman 1989). However, the observational evidence in favour of bar feeding is not strong. Regardless of the details of the feeding processes, neither of these offer great potential for runaway growth of the BH, and this may be the reason that BHs in spiral galaxy hosts do not achieve the masses (or potential

luminosities) of BHs in elliptical hosts. However, should a molecular cloud happen to interact with the BH, the same considerations as developed in the previous paragraphs should apply. In Seyferts therefore, we might speculate that their duty cycle is low (it has to be of order 1% in order that the BH does not grow too seriously fat), but that during ‘active’ periods, it is shining at near-Eddington luminosity (as confirmed by reverberation mapping analyses—see Maoz 1994), and attempting to accrete molecular gas at super-Eddington rates.

During the accretion process, the molecular clouds are presumably being converted rapidly to stars in a massive nuclear starburst, so that the growth of the BH is determined by the competition between inflow, accretion, and star formation.

Consider a simple (toy) parametric form for the mass accretion rate into the nuclear core region;

$$\dot{M}(t) = \dot{M}_{av} \exp[-t/\tau_{acc}](1 - \exp[-t/\tau_{acc}]). \quad (2)$$

This is just about the simplest form that could be adopted, having only one characteristic timescale. It starts accreting at zero, reaching a maximum at $t \approx \tau_{acc}$ and then decreases exponentially at large t . If M_{tot} is the total mass accreted, then $\dot{M}_{av} = 2M_{tot}/\tau_{acc}$. This accretion timescale is related to the free-fall timescale from the point of origin of the infalling gas to the mass centre by a scaling factor:

$$\tau_{acc} = \lambda \tau_{ff} = \lambda R^{\frac{3}{2}} / (GM)^{\frac{1}{2}}; \quad (3)$$

with $R \approx 10$ kpc and $M \approx 10^{10} M_{\odot}$ then $\tau_{acc} \approx 1.5 \times 10^8 \lambda$ years.

Now, the accretion rate to the black hole during the period when the accretion rate into the nuclear regions is super-Eddington is given by

$$\dot{M}_{BH} = M_{BH} / \tau_{BH}, \quad (4)$$

where τ_{BH} is the growth timescale for the BH accreting at the Eddington limit. This can be calculated from the luminosity

$$L_{BH} = \phi \dot{M}_{BH} c^2 = \theta L_{Edd} = \theta 4\pi c G M_{BH} / \kappa, \quad (5)$$

where ϕ is the fraction of the rest mass energy radiated by matter falling into the BH. For a Schwarzschild BH we have $\phi = [1 - (2^{\frac{3}{2}}/3)]$ and typically $\phi \sim 0.1$. The factor θ is the fraction of the Eddington luminosity produced by this accretion, and is assumed to be ~ 1 . Thus, the growth timescale of the BH is

$$\begin{aligned} \tau_{BH} &= M_{BH} / \dot{M}_{BH} = \frac{\kappa c}{4\pi G} \left(\frac{\phi}{\theta} \right) \\ &\approx 4.4 \times 10^8 \left(\frac{\phi}{\theta} \right) \text{ years} \end{aligned} \quad (6)$$

and inserting numerical values gives $\tau_{BH} \approx 5 \times 10^7$ years. Thus, the growth timescale of the BH is of the same order as, or somewhat smaller than τ_{acc} , so that appreciable growth of the BH can occur during the merger. We discuss below whether it is possible that the BH grows from that typical of a Seyfert galaxy ($\approx 10^7 M_{\odot}$) to that typical of a BH in a giant elliptical galaxy ($\approx 10^9 M_{\odot}$) during the merger event.

3 Radiation-pressure Driven Wind Model

During the period of super-Eddington accretion and possibly beyond, the radiation pressure from the central BH would be sufficient to drive a wind. If this wind is driven by radiation pressure of the central source, then equations similar to those of hot stars apply. For the case of a magnetic wind externally illuminated by the central engine, the detailed theory has already been developed in a series of papers by Murray and his collaborators (Murray & Chiang 1995; Murray et al. 1995; Chiang & Murray 1996). However, in our case we would like to consider the possibility that the radiatively driven wind is both dense, so that it is not heated to the Compton temperature, and is optically thick to the escape of X- and γ -ray photons from the central source, so that the continuum observed is produced by a reprocessing photosphere dominated by electron scattering opacity. Such geometry may apply in the high accretion rate limit. In this case, the momentum flux in the wind is given by

$$\dot{M}_w v_w = \left(\frac{\eta}{c} \right) \left(\frac{\Omega}{4\pi} \right) L_{BH}, \quad (7)$$

where Ω is the solid angle covered by the optically-thick radiatively-driven wind, subtended at the BH, and η is the effective number of scatterings per photon. In stars, this factor is greater than unity; typically $\eta \sim 3-4$; however, in Wolf-Rayet stars it may rise even higher. In such radiatively-driven winds, the outflow velocity is similar to the escape velocity at the base of the outflow, which in this case will be at the inner edge of the region of super-Eddington flow:

$$v_w = \epsilon \left(\frac{G M_{BH}}{r_{in}} \right)^{\frac{1}{2}} \sim 3.6 \times 10^4 \epsilon M_7^{\frac{1}{2}} r_{14}^{-1/2} \text{ km s}^{-1}, \quad (8)$$

where M_{BH} the mass of the BH in units of $10^7 M_{\odot}$ and r_{14} is the inner radius in units of 10^{14} cm. In stars, we have $1 \leq \epsilon \leq 3$ typically. From equation (7) the number density in the wind is given by

$$n(r) = \left(\frac{\eta L_{BH}}{4\pi c v_w^2} \right) r^{-2}, \quad (9)$$

which corresponds to $\sim 4 \times 10^9 L_{45} r_{15}^{-2} v_4^{-2} \text{ cm}^{-3}$, where L_{45} is the luminosity of the BH in units of $10^{45} \text{ ergs s}^{-1}$, r_{15} is the radius in units of 10^{15} cm and v_4 is the velocity of the wind in units of 10^4 km s^{-1} . The photospheric radius is set by the point where the electron scattering optical depth in the outflowing wind is unity, which for a constant-velocity wind gives

$$r_{\text{phot}} = \kappa \left(\frac{\Omega}{4\pi} \right)^{-1} \left(\frac{\dot{M}_w}{4\pi v_w} \right) \sim 1.0 M_7 \text{ light days}, \quad (10)$$

where κ is the electron scattering opacity. Such photospheric radii are consistent with limits produced by variability studies and reverberation mapping analysis of nearby Sy I galaxies. The effective temperature of the photosphere is then determined from Stefan's Law, remembering that electron scattering photospheres have their radiation density diluted by a factor $\Psi \sim 10$ with respect to a black-body emitter:

$$T = \left(\frac{4\pi c^2 \Psi}{\sigma \kappa^2 \eta^2} \right)^{\frac{1}{4}} v_w L_{BH}^{-1/4} \\ \sim 37,000 \text{ K} \left[\frac{\Psi}{10} \right]^{\frac{1}{4}} \left[\frac{\eta}{3} \right]^{-1/2} v_4 L_{45}^{-1/4}. \quad (11)$$

Thus, when the wind velocity is of order $20,000 \text{ km s}^{-1}$ as implied by equation (8), the effective photospheric temperature is high enough ($\sim 10^5 \text{ K}$) to provide a 'big blue bump' in the continuum spectrum which, thanks to the high outflow velocity and the electron scattering, should provide only weak and broad photospheric lines. In addition, such electron scattering dominated extended atmospheres should show only a weak Lyman limit discontinuity. It should be noted that the spectral distribution of such an electron scattering photosphere is not characterised by a simple temperature; the temperature given in equation (11) is only representative of the hardness of the spectrum. In general, such atmospheres give a power-law below the peak in emergent flux, and roll off sharply above this peak. This is the result of the fact that the peak temperature, given roughly by equation (11) is seen in the line of sight to the centre of the structure where the deepest penetration occurs. The photosphere occurs further out in annuli around this point, and so is characterised by lower electron temperatures. The contribution from these annuli produce the power-law extension of the spectrum to lower frequencies. The slope of this power-law depends on the effective curvature of the atmosphere, i.e. on the ratio $\dot{M}_w/v_w$. This effect is very familiar to those who work on Wolf-Rayet stellar atmospheres (e.g. Abbott & Conti 1987).

In AGN, the value of this ratio is much higher than in Wolf-Rayet stars, so we expect them to be characterised by little curvature, and a rather flat power-law [$d(\log \nu F_\nu)/d(\log \nu) \leq 1$] at energies below the peak.

Since the photosphere is extended both above and below the accretion disk, it can illuminate and photoionise the surface layers of the accretion disk. This would give the broad-line region in our model. Since the effective temperature decreases for higher BH luminosities, we would expect that this photoionisation would provide lower ionisation conditions in the more luminous AGN. This is presumably the explanation of the Baldwin effect in QSOs (Baldwin 1977; Baldwin et al. 1978), where the C IV equivalent width and the ratio of C IV/Ly α is observed to decrease with increasing luminosity. This effect has been most comprehensively studied by Kinney, Rivolo & Koratkar (1990).

Equations (1), (2) and (3) imply together that

$$L_{BH} = \frac{4\pi \epsilon c G M_{BH}}{\kappa \eta} = \left(\frac{\epsilon}{\eta} \right) L_{Edd}; \quad (12)$$

consistency therefore requires that $\epsilon \sim \eta$ if the BH luminosity is to be of the same order as the Eddington limit. Equation (1) also implies that the mechanical energy flux in the wind is simply related to the bolometric luminosity of the central object:

$$L_{\text{mech}} = \left(\frac{\eta v_w}{2c} \right) \left(\frac{\Omega}{4\pi} \right) L_{BH} \\ = 0.05 \left[\frac{\eta}{3} \right] \left[\frac{\Omega}{4\pi} \right] v_4 L_{BH}. \quad (13)$$

Finally, we can relate the mass flux in the wind to the mass flux onto the BH, using equations (5) and (7):

$$\left(\frac{\dot{M}_w}{\dot{M}_{BH}} \right) = \left(\frac{\eta \phi c}{v_w} \right) \left(\frac{\Omega}{4\pi} \right) \\ = 9 \left[\frac{\phi}{0.1} \right] \left[\frac{\eta}{3} \right] \left[\frac{\Omega}{4\pi} \right] v_4^{-1}. \quad (14)$$

Thus, the mass flux into the radiatively driven wind will dominate over the mass flux into the BH until the accretion disk becomes thin ($\Omega/4\pi \sim 0.1$).

The outflowing radiative wind interacts directly with the accretion flow. The condition that the accretion flow is not seriously dynamically modified by this interaction is that the ram pressure in the wind is less than the ram pressure in the accretion flow. From equation (7) this implies that the BH luminosity (in units of $10^{46} \text{ ergs s}^{-1}$; L_{46}) is below $L_{46} \leq 5 \dot{M}_{200} v_{300} \omega_{in}^{-1}$ where $\dot{M}_{200}$ is the rate of mass accretion in units of $200 M_\odot \text{ yr}^{-1}$, v_{300} is the

velocity of infall in units of 300 km s^{-1} and ω_{in} is the solid angle covered by the accretion flow.

4 Relativistic Jet Model

In the spirit of our parametric analysis, let us adopt the main features of the jet model developed by Falke & Biermann (1995), and as fitted to the observational material by Falke, Malkan & Biermann (1995):

$$L_{jet} = \gamma_{jet} \dot{M}_{jet} c^2 \approx 0.3 L_{BH}; \quad 3 \leq \gamma_{jet} \leq 10. \quad (15)$$

It therefore follows immediately that

$$\left(\frac{\dot{M}_{jet}}{\dot{M}_{BH}} \right) = \left(\frac{\phi}{3\gamma_{jet}} \right). \quad (16)$$

Thus, the mass flux in the jet is only about 1% of the mass flux into the BH, and more importantly, is only about 0.1% of the mass flux in the radiatively-driven wind, when the BH is attempting to accrete at a super-Eddington rate. This very low mass flux compared to the wind then implies that mixing between the jet and the thermal wind will be critical in determining whether the jet can escape (radio-loud) or is trapped within the radio photosphere of the wind determined by its free-free opacity at radio frequencies.

4.1 Radio Loud or Radio Quiet?

The degree of mass-loading of the thermal wind material into the relativistic jet will determine the energy loss rate of the electrons in the jet. Energy losses from the relativistic plasma may either be through synchrotron radiation or by electron bremsstrahlung. For electron bremsstrahlung losses in a medium with ionic density n , the fractional energy loss per unit path length dx is given by

$$\frac{dE}{E} = \frac{16}{3} Z^2 n \left(\frac{e^2}{\hbar c} \right) r_0^2 \ln[\gamma] dx, \quad (17)$$

where r_0 is the classical electron radius. Thus for a jet with $dE/E \sim 5$ the stopping length, given by $dE/E \sim 1$, is $\sim 10^{26} n^{-1} \text{ cm}$. However, the density in the wind is given by equation (9). For a BH luminosity $L_{45} \sim 1$ (in units of $10^{45} \text{ erg s}^{-1}$) the radius of the photosphere is $r_{15} \sim 1$. Thus, in order to bury the loss region of the jet along with its associated bremsstrahlung photons below the electron scattering photosphere, the jet must interact with the wind at a radius of $\leq 2 \times 10^{14} \text{ cm}$; or about 100 gravitational radii, in this example. Since, to first order, the mass loss in the wind scales as the mass of the BH, this critical distance

of about 100 gravitational radii will hold true for different BH luminosities.

The radio frequency photons produced in the jet-wind interaction are blocked by the free-free optical depth in the ionised wind. At frequency ν the photospheric radius (defined as a free-free opacity of unity) is given by

$$\int_{r_{ff}}^{\infty} \tau(\nu, r) dr = 1, \quad \text{with} \quad \tau(\nu, r) = 1.1 \times 10^{-25} \nu_9^{-2.1} n_e^2(r) T_e(r)^{-1.35}, \quad (18)$$

where ν_9 is the frequency considered (GHz). For an electron temperature of 30,000 K, this implies $r_{ff} \sim 8 \times 10^{15} \nu_9^{-0.7} L_{45}^{\frac{2}{3}} \text{ cm}$, which is several times larger than the radius of the electron scattering photosphere. Presumably the radio flux emerging over this photosphere determines the brightness and frequency spectrum of the radio core component of the AGN.

As far as the electrons which escape to emerge into the lobe are concerned, Bicknell, Dopita & O'Dea (1997) have shown that the jet energy flux is related to the synchrotron power at frequency ν through an efficiency factor

$$\begin{aligned} \kappa_\nu &= \left(\frac{5 - \delta}{8 - \delta} \right) f(t, \gamma_0, B, \nu, \alpha) \\ &= c E_0^{2\alpha} F_e B^{(\alpha+1)} \nu^{-\alpha} t, \end{aligned} \quad (19)$$

where δ is the index of the power law in the density distribution of the galactic medium (~ 2), t is the age of the source, B is the magnetic field in the radio lobes, E_0 is the lower energy limit of the relativistic electrons, α is the spectral index, and F_e is the fraction of the internal energy of the plasma contained in relativistic electrons. In the super-Eddington phase, in which the relativistic jet gas is able to mix with the thermal wind either below, or near, the electron scattering photosphere, we get $F_e \sim 10^{-3}$. On the other hand, highly sub-Eddington accretion allows for the free escape of the relativistic plasma so $F_e \sim 1$. However, radio-quiet objects have lobe luminosities which are typically weaker by a factor of a thousand than their radio-loud counterparts. Thus, for a given jet energy flux, this difference in the entrainment factor is in itself sufficient to explain the difference between the radio-quiet, and the radio-loud cases, respectively, provided that the entrainment has occurred below a radio photosphere. We may therefore conclude that radio-quiet behaviour is a signature of an AGN which is undergoing rapid accretion.

The transition between the radio-quiet and radio-loud cases presumably comes about at the time that the accretion rate into the BLR decreases to the point that the thermal wind starts to clear in the polar

regions [$\omega \leq 4\pi$ in equation (7)], and entrainment into the radio jet is no longer enough to slow this to transonic speeds. This allows for the possibility that there are transition objects displaying *both* BLRs and relativistic jets. Examples of such objects may be IC5063, BL Lac itself and 3C120.

4.2 The Radio Core

For the quasars in the bright quasar sample (BQS) of Schmidt & Green (1983), Lonsdale, Smith & Lonsdale (1995) have shown that there is a good correlation between the radio core flux and the bolometric luminosity. This correlation exists independently of the assumptions used to calculate the bolometric luminosity. Moreover, the correlation is improved by the exclusion of the radio-loud QSOs (for which relativistic beaming effects enhance the core luminosity). Lonsdale, Smith & Lonsdale (1995) have also shown that this correlation extends to the ultraluminous infrared galaxies as well. This correlation strongly suggests that, since the optical/UV continua are thought to be thermal, the radio core is either produced by thermal processes or through non-thermal emission powered directly by these thermal processes. Of these two possibilities the latter seems more probable in the light of the observations of Barvainis, Lonsdale & Anonucci (1996), which demonstrated that at least some VLBI cores of the radio-quiet AGN show nonthermal brightness temperatures characteristic of synchrotron emission.

In this model, how could this synchrotron emission arise? The most likely solution is through Fermi acceleration of electrons in strong shocks surrounding the nucleus. Such shocks arise when the fast radiatively-driven wind impinges upon either slow, magnetically-driven winds from the accretion disk, or upon the accretion flow itself. Since these shocks carry a nearly constant fraction of the bolometric luminosity, the correlation between radio core luminosity and bolometric luminosity would then find a natural explanation.

Since this shocked region is the region in which mass entrainment into the fast wind will occur, we can speculate that the broad absorption line (BAL) QSOs are those objects in which the shock interface is aligned with the line of sight.

If core radio power can be used in this way to estimate bolometric luminosity and hence (for BHs accreting at their Eddington rate) the mass of the BH, the correlations of Nelson & Whittle (1996) could be used to link the mass of the BH with the mass of the bulge in which it resides. Nelson & Whittle (1996) found that (with large overlap between different classes of object) the radio core luminosity relates directly to either bulge absolute magnitude, or to velocity dispersion in the bulge for Sy 1 and Sy 2 galaxies, and for FR I and FR

II radio galaxies, although objects with jets have systematically higher core luminosities as might be expected in the presence of a relativistically beamed component. This correlation is most easily explained if the bulge stars were formed coevally with the main growth phase in the central BH, i.e. during a merger event.

5 The Narrow-line Region

5.1 Jet-excited Regions

The various classes of bright and compact extragalactic radio sources form an fascinating subset of all radio sources which holds the key to our understanding to the evolution of radio jets in general. These include the steep-spectrum radio sources (CSS) (Fanti et al. 1990), the gigahertz-peaked sources (GPS) (O'Dea et al. 1990, and references therein), the compact symmetric objects (CSO) (Wilkinson et al. 1994, and references therein) and the compact double sources (CD) (Phillips & Mutel 1982). Following the Caltech–Jodrell Bank and Bologna–Jodrell–Dwingeloo surveys (Wilkinson et al. 1994; Fanti et al. 1995), these classes of radio sources are now understood to represent an appreciable fraction (10–30%) of the luminous radio sources. Not only are these sources very luminous at radio frequencies, but they are also very luminous in the optical. The spectra of Gelderman & Whittle (1994) and Gelderman (1995) reveal the broad emission lines of the AGN itself as well as intense ‘narrow line’ emission reminiscent of Seyfert II galaxies. The similarity of properties of the different classes of compact sources argues strongly that the same physical processes are at work in all of them.

Following Begelman (1996), Bicknell, Dopita & O'Dea (1997) assumed that all classes of compact, young radio sources are manifestations of a jet-driven outflow in a galaxy which is also undergoing an accretion or merging event. As a result of the accretion, the ambient density of the galactic medium will be strongly peaked toward the centre. A reasonably physical model density distribution could be that produced by a steady inflow $\rho(r) \propto r^{-2}$ or, at later phases, that produced by free fall towards a massive nuclear core, $\rho(r) \propto r^{-3/2}$. Bicknell, Dopita & O'Dea (1997) computed the evolution of jets with a generalised form of the density distribution; $\rho(r) \propto r^{-\delta}$.

As the jet propagates from the AGN, it is either disrupted by Kelvin–Helmholtz instabilities within it or by the termination shock, and back-fills the cavity it has produced, giving rise to a symmetric pair of relativistic plasma bubbles on either side of the nucleus. As these bubbles expand, wall shocks are pushed into the galactic medium by the cocoon overpressure. Provided that these shocks can become radiative within a dynamical expansion timescale,

then the optical emission of these shocks and the general outflow associated with them will become visible as a shock-excited narrow-line region (NLR). In this model, the fluxes of the optical emission lines such as [O III] $\lambda 5007$ Å, or the H α + [N II] complex, is almost proportional to the mechanical energy flux through the shocked region, i.e. to the total luminosity of the shocks L_T ;

$$\begin{aligned} L_{[\text{O III}]} &= 0.02 L_T, \\ L_{\text{H}\alpha + [\text{N II}]} &= 0.02 V_3^{-0.58} L_T. \end{aligned} \quad (20)$$

However, the shock luminosity is equal to the rate of work done by the expansion of the radio lobe, which is given by

$$L_T = \frac{3}{8 - \delta} \dot{E}_J, \quad (21)$$

where $\dot{E}_J$ is the jet energy flux. Thus, with $\delta = 2$, the total luminosity in the NLR is $\sim \frac{1}{2}$ of the jet energy flux.

Note that equation (12) allows a maximum jet energy flux of about 30% of the bolometric luminosity of the BH. Thus the NLR will have a maximum luminosity of 15% of the bolometric luminosity of the BH. On the other hand, equation (11) implies that the radiative wind carries about 5–10% of the bolometric luminosity, which would be likewise capable of shock exciting an NLR. In this case, the NLR accounts for about 2–5% of the bolometric luminosity of the BH. This limit is appropriate for the mass-entrained jets of radio-quiet objects, and in general, we might expect the luminosity of the NLR to lie between these two limits.

In the case of objects with ‘buried’ or dust enshrouded narrow-line regions, the extended far-infrared luminosity may possibly be used as a tracer of mechanical energy input. A shock with a velocity of 500 km s^{-1} puts $\sim 13\%$ of its total luminosity in the Ly α line, and its precursor adds another $\sim 9\%$ for a grand total of 22%. However, faster shocks emit a greater fraction of their luminosity in this line, thanks mostly to increasing amounts of collisional excitation in the partially ionised recombination zone of the shock. For a 1000 km s^{-1} shock, the percentage of total luminosity in Ly α rises to 24% and in the precursor to 9%. Thus, a total of 34% of the mechanical energy can be converted into Ly α , which is then largely trapped in the surrounding medium, absorbed by dust grains, and re-radiated as far-IR emission. If we also take into account the competition of the dust grains in directly absorbing hard UV and soft X-rays in the precursor, the percentage of the mechanical energy flux through the shock which is converted to far IR dust emission may rise to as much as 50%.

Nonetheless, these figures mean that only 2–7% of the bolometric luminosity of the BH is ultimately converted to far-infrared photons. Direct absorption of the radiative energy of the BH will remain the dominant source of far-infrared photons.

In Figure 1, we plot the [O III] line flux in the NLR against the radio flux associated with the AGN for a wide range of AGN subtypes. Note the clear bifurcation between the radio-loud and the radio-quiet classes, with the exception of the radio-intermediate IRAS galaxies taken from McGregor, van Breugel & Kewley (in preparation). This rare class of objects displays post-starburst continua (a spectrum dominated by A-type stars with prominent Balmer absorption), along with very luminous NLR emission presumably powered by lobes with a large admixture of thermal gas, giving them properties between the radio-quiet and the radio-loud branches. In the context of the model presented here, Figure 1 could be considered as a ‘Hertzsprung–Russell’ diagram for AGN. In gas-rich post-merger systems as the BH grows, the AGN moves up the radio-quiet sequence. At the same time, some of the gas is converted to stars which eventually form the nuclear bulge of the galaxy. When the accretion rate into the nuclear region drops below the Eddington-limited value, the relativistic jet escapes, and the AGN moves across through the radio-intermediate IRAS stage to the radio-loud QSO region. At this point, the merged galaxy would be classified as an elliptical. Because this galaxy now has a massive BH at its centre, subsequent mergers of low-mass gas-rich systems such as Ir galaxies cannot as easily produce super-Eddington inflow, but may produce inflow which approaches the Eddington rate. These systems will generate radio-loud QSOs, GPS or CSS radio sources, and subsequently FR II radio sources. Accretion due to cooling flows can only remain sub-Eddington. In these sources the NLR is LINER like and, as I show in the next section, this sub-Eddington accretion phase is likely to be related to BL Lac or low-power FR 1 activity.

5.2 LINERs—The Signature of Accretion Disks

The low-ionisation nuclear emission-line regions, (LINERs), were first recognised as a distinct class of AGN by Heckman (1980). Where these emission-line regions can be resolved, they appear as an extended region strongly peaked toward the nucleus. The velocity broadening of these regions is generally a few hundred km s^{-1} , with evidence for a considerable degree of rotational support in many objects. The emission-line ratio criteria given by Heckman define a distinct (if somewhat arbitrary) region of excitation space for these objects. According to his definition the [O II] $\lambda\lambda 3727, 3729$ lines are stronger than [O III] $\lambda 5007$; the [O I] $\lambda 6300$ /[O III] $\lambda 5007$ ratio is less than about 0.33; and the [N II] $\lambda 6584$ /H α

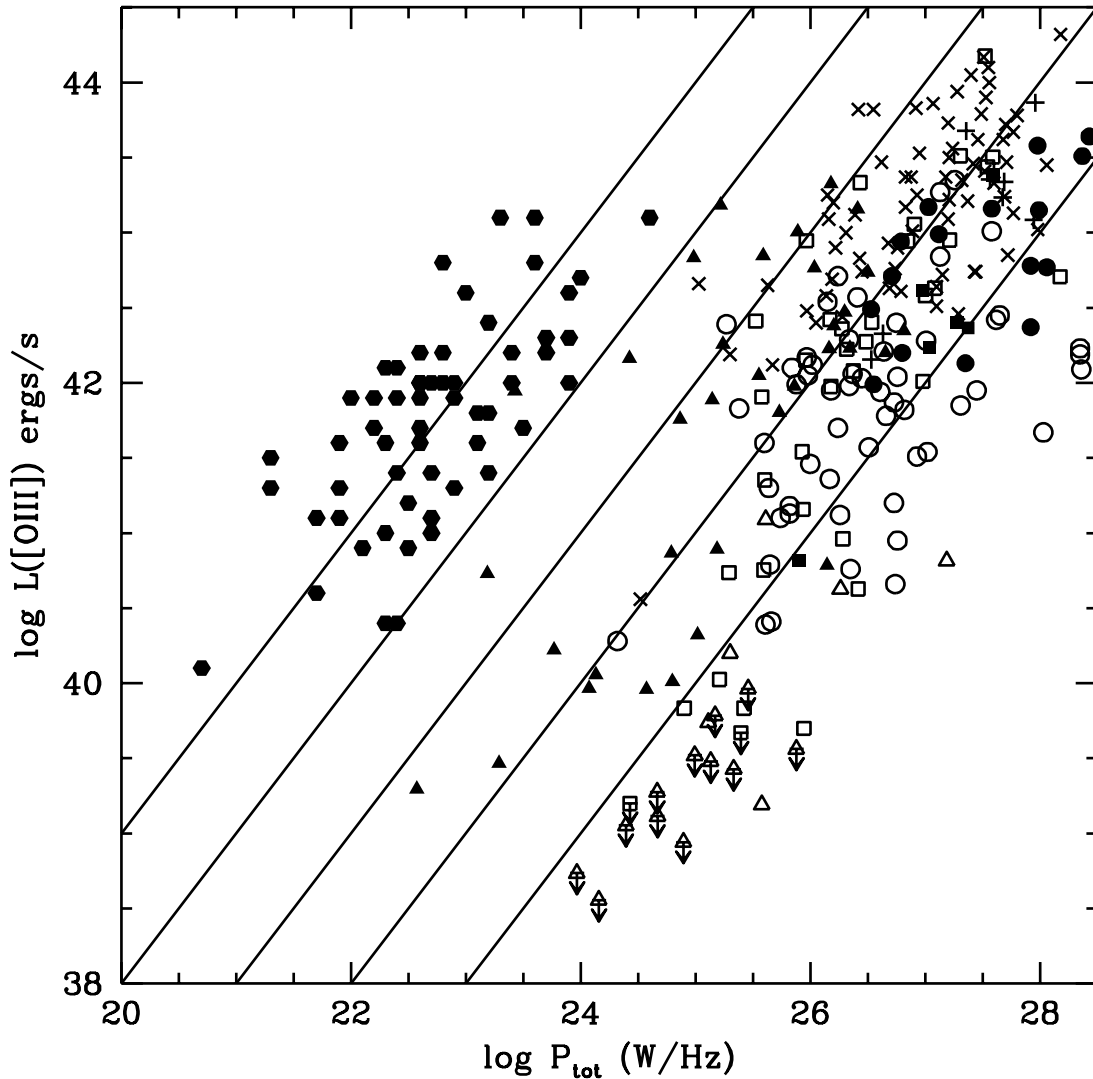


Figure 1—Relationship between the [O III] luminosity and the radio luminosity for both radio-quiet and radio-loud classes of AGN drawn from the literature (Bicknell, Dopita & Tsvetanov 1997). This can be considered as the ‘Hertzsprung–Russell’ diagram for AGN evolution. We envisage that BHs in spiral hosts remain on the radio-quiet side of the diagram, unless they are involved in a major merger, in which case rapid growth of the BH occurs (IRAS ‘60 μ m peakers’), before the AGN rapidly passes through the intermediate radio-loud IRAS galaxy phase to become a radio-loud massive BH in an elliptical host galaxy. Subsequent mergers with low-mass companion galaxies cause it to be reborn as a GPS or CSS galaxy, and later, as an FR II radio galaxy. As accretion decreases, or if the AGN is fed by a cooling flow, the AGN appears as a LINER/FR I/BL Lac object.

ratio is larger than about 0.6. A survey of ‘normal’ elliptical galaxies by Phillips et al. (1986) showed that some low-level LINER activity is found in an appreciable fraction of these. The recent work of Ho, Filippenko & Sargent (1995) has shown that low level LINER activity is even more ubiquitous than had been suspected in both elliptical and spiral galactic nuclei. Elliptical LINER galaxies cover a range of radio properties, including radio-quiet objects, weak nuclear sources, and Fanaroff–Riley (FR) Class I sources. However, no example of a LINER is yet known to be associated with the high-power FR II radio sources. In terms of the correlation between the optical and the radio luminosities, a similar trend emerges (Baum, Zirbel & O’Dea 1995). There is a good correlation for the FR II class, but this

disappears smoothly at the FR I–II transition. These correlations suggest that the optical emission makes a transition from predominantly jet-excited to some other excitation mechanism at the FR I–II transition. An appreciable fraction of late type galaxies with LINER nuclei was found by Ho, Filippenko & Sargent (1995) to have an underlying broad component of $H\alpha$, emphasising the connection between LINERs and other types of AGN. Finally, LINERs, or at least objects having LINER-like spectra, are also found in the cooling flows associated with first-ranked elliptical galaxies in clusters (Heckman 1981; Cowie et al. 1983; Johnstone, Fabian & Nelsen 1987; Heckman et al. 1989).

In the light of the model presented in this paper, this correlation with radio properties and host

galaxies strongly suggests that LINERs are found in the objects with sub-Eddington (and possibly advective; Reynolds et al. 1996) accretion rates. It is in these cases that the details of the extended accretion disk around the BH can be observed. A fine nearby example of this is presented by M87. From optical and UV spectroscopy with the HST, Dopita et al. (1997b) have been able to demonstrate that the LINER emission arises by shock dissipation in a turbulent accretion flow derived from the M87 cooling flow as it settles into an organised flat accretion disk within a few parsecs of the nucleus. We know that this accretion continues down to the central massive object, since we see the optical synchrotron jets emerging orthogonally to the disk plane. The mechanical energy flux of the jets of the FR I source has been calculated by Bicknell & Begelman (1996) to be $\sim 1 \times 10^{44}$ ergs s $^{-1}$, which implies an accretion rate onto the BH of about $2 \times 10^{-2} M_{\odot} \text{ yr}^{-1}$. Since Ford et al. (1994) estimate the mass of the black hole to be $2.4 \times 10^9 M_{\odot}$, the rate of nuclear mass accretion in this object is only about 2×10^{-4} of the Eddington rate. Clearly, M87 is an excellent example of a sub-Eddington accretion flow. In addition, Tsvetanov et al. (in preparation) have demonstrated that the nuclear spectrum of M87 is a power-law continuum, and is strongly variable on a timescale of weeks or months. These properties emphasise the link between FR I sources, LINER nuclei, and BL Lac objects.

Further (unpublished) HST spectrophotometric data on NGC1052 and M81, prototypes of LINERs in an elliptical and a spiral, respectively, show the strong UV lines expected in the case of shock excitation. In the case of NGC1052, it has long been asserted that it is a shock excited disk (Koski & Osterbrock 1976; Fosbury et al. 1978). In addition, M81 shows some evidence for a broad-line nucleus and a dense inner accretion disk.

This strongly suggests that all LINERs are excited in the same way as the nuclear regions of M87, and that LINER activity is therefore a measure of the shock dissipation occurring in the disorganised accretion flows around BHs accreting at sub-Eddington rates. In the case of elliptical host galaxies, the central BHs may be quite massive, with the sub-Eddington inflow derived from a large-scale galactic or cluster cooling flow. In the case of spiral host galaxies, it may be that the accretion flow is simply derived from the ISM in the galaxy, but that the absence of a large-scale disturbance (such as a bar instability) ensures that this accretion flow is very weak.

6 Growth of Black Holes

Since the exact relationship between τ_{acc} and τ_{BH} is uncertain, we take $\tau_{acc} = \alpha \tau_{BH}$. The condition that the BH is attempting to accrete at a super-Eddington rate is that

$$\dot{M}(t) - \dot{M}_w(t) \geq \dot{M}_{BH}(t) \quad (22)$$

but, from equations (2) and (6),

$$\left(\frac{2M_{tot}}{\alpha M_{BH}(0)} \right) \left(\frac{9\Omega}{4\pi} + 1 \right)^{-1} \exp \left[- \frac{(\alpha + 1)t}{\tau_{acc}} \right] \times \left(1 - \exp \left[- \frac{t}{\tau_{acc}} \right] \right) \geq 1; \quad (23)$$

thus, in the case $\alpha = 2$, $M_{BH}(0) = 10^7 M_{\odot}$, $M_{gas} = 10^{10} M_{\odot}$, super-Eddington accretion starts almost immediately and ends at $t = 2.27 \tau_{acc}$. The left-hand side of equation (2) is the factor by which the instantaneous accretion rate into the central regions exceeds the Eddington value.

In Figure 2 we plot the factor by which the instantaneous accretion rate into the central regions can exceed the Eddington value for this particular case. Typically, in the early evolution of the BH, this factor can exceed 100, which implies the existence of a generalised outflow over and above the radiatively driven wind. On the other hand, sub-Eddington accretion disks will be inefficient in maintaining any form of wind. We may therefore distinguish three phases in the accretion:

- (1) $t < t_1$; super-Eddington, disk wind + radiatively-driven wind, $\Omega = 4\pi$.
- (2) $t_1 < t < t_2$; super-Eddington, equation (16) satisfied at its lower limit. Radiatively-driven wind, $\Omega < 4\pi$.
- (3) $t > t_2$; sub-Eddington, classical thin disk accretion + relativistic jet.

In this scheme, the transition between the radio-quiet and radio-loud phases would occur sometime in the second phase, when the radiative wind opens up in the polar direction to allow the free escape of the relativistic jet. If the BH grows at its Eddington-limited rate until the end of this phase, and swallows all the accreting gas when the accretion rate falls below the Eddington-limited value, the relationship between the initial BH and the final BH that is grown is as follows:

$M_{BH}(0)/M_{\odot}$	10^6	10^7
$\alpha = 1$	0.16×10^9	0.89×10^9
$\alpha = 2$	0.54×10^9	2.89×10^9
$\alpha = 5$	1.23×10^9	7.37×10^9

The temporal evolution of mass of the BH is shown in Figure 3. Under almost all circumstances, a BH with a final mass $\approx 10^9 M_{\odot}$ can be produced from an initial BH similar to those found in spiral galaxies today. During the super-Eddington accretion phase, the luminosity of the BH remains limited at its Eddington value. At later phases, the luminosity declines thanks to the decreasing inflow rate. This

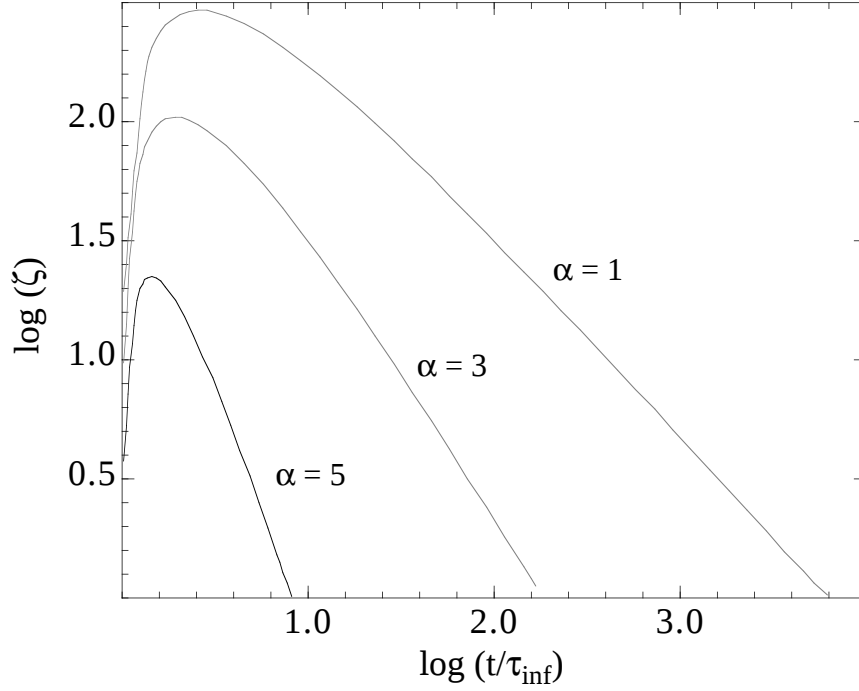


Figure 2—The factor by which the nuclear feeding exceeds the Eddington value for the black hole, ζ , for the case $M_{BH}(0) = 10^{-3} M_{gas}$ and three values of α . With $M_{BH}(0) = 10^{-4} M_{gas}$ the period of overfeeding is longer, and $\log \zeta$ rises as high as 3.5.

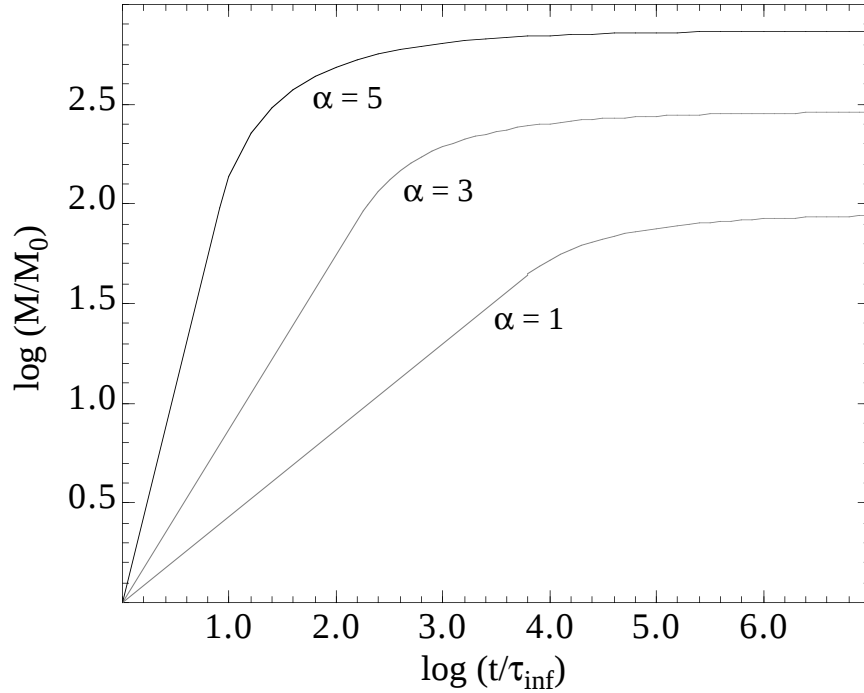


Figure 3—Mass of the BH compared to its initial value as a function of time for the three cases of Figure 2.

would result in an exponentially increasing luminosity during the phases 1 and 2, followed by a (nearly) exponentially decreasing luminosity in phase 3.

The fact that the final BH mass is not very variable with respect to change in the accretion parameters means that the mass of the BH relative to the mass that is accreted into the central regions

(which has presumably been used to form the ‘bulge’ of the galaxy) also varies weakly with the accretion parameters.

7 Evolution and Aspect

The evolutionary sequence has already been suggested on the basis of Figure 1. In the case of a major

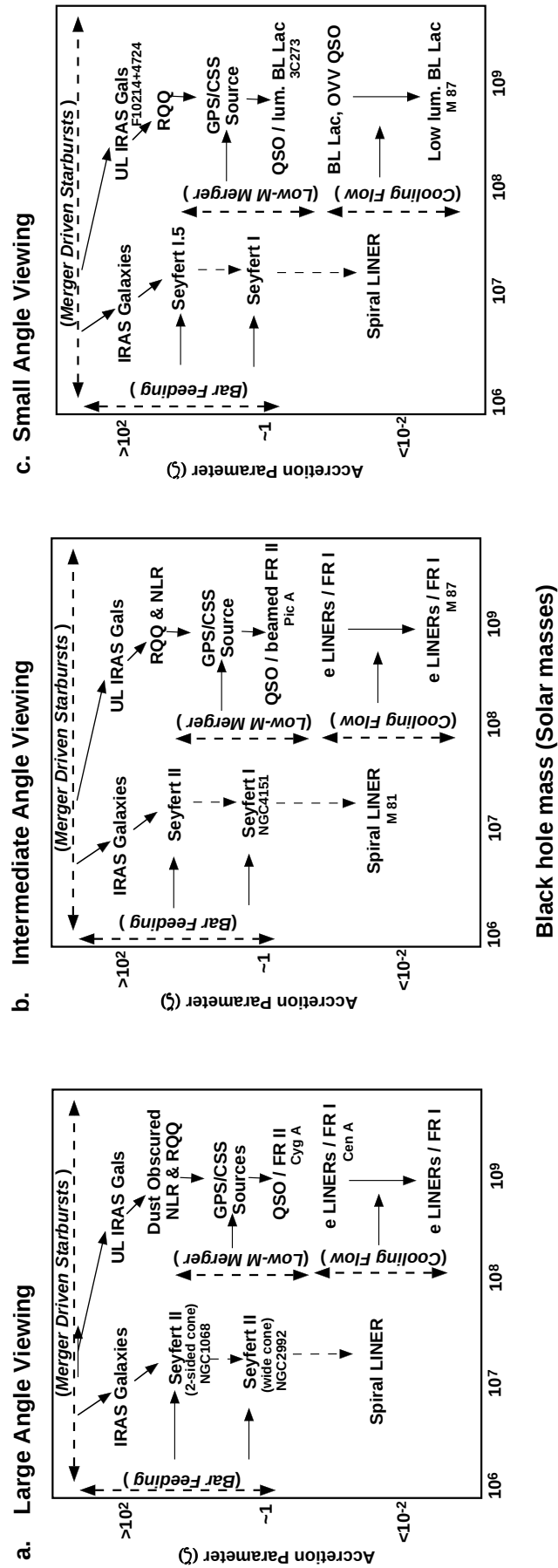


Figure 4—Proposed unification scheme according to aspect, accretion rate, black hole mass and evolutionary status.

merger event between two massive gas-rich systems, the initial phase would be a galaxy-wide starburst producing a luminous infrared galaxy with HII-like characteristics. In the later phases, gas flow toward the nucleus feeds the BH and outflow could then excite the Seyfert 2 like emission seen in some of the very distant ultraluminous IRAS galaxies (Rowan-Robinson et al. 1991, 1993; van Ojik et al. 1994). Such mergers will initially be radio-quiet, since the supply of gas into the nuclear regions is sufficient to choke off high-luminosity radio jets, but as nuclear accretion rates fall, radio jets might later escape to produce a high z radio source with its strong narrow-line emission region (Meisenheimer, Hippelein & Neeser 1994; McCarthy, Spinrad & van Breugel 1995). As mentioned above, such mergers of gas-rich systems would have been much more common in the early universe and especially at the epoch of formation of the Abell clusters. Since this epoch was so early, $z \sim 2-4$, the merging gas-rich systems may not even have had time to form regular spiral galaxies. However, recognising this, let us call them ‘spirals’ for convenience.

It is therefore reasonable to propose the following evolutionary scenario:

$$\begin{aligned} \text{Sp} + \text{Sp} &\Rightarrow \text{Starburst} \\ &\Rightarrow \text{Postmerger} + \text{RQSO} \\ &\Rightarrow \text{E} + \text{massive BH} + \text{Radiojet(?)}. \end{aligned}$$

Such ‘primary’ merger events will give rise to the QSOs and the majority of elliptical galaxies. In the event of a later merger event:

$$\begin{aligned} \text{E} + \text{Irr} &\Rightarrow \text{Dust lane E} + \text{Nuclear starburst} \\ &\Rightarrow \text{Dust lane E} + \text{GPS radio source} \\ &\Rightarrow \text{Dust lane E} + \text{FR II} \\ &\Rightarrow \text{E} + \text{FR I (?)}. \end{aligned}$$

Whether evolution to the last of these steps can occur is not clear, but the sharp break between the FR 1/FR 2 classes (Owen & Ledlow 1994), which can be explained theoretically (Bicknell 1994), is certainly suggestive. Alternatively, the case of M87 suggests that low-power radio sources may be fed by a subsequent accretion episode at sub-Eddington rates, such as feeding through the cooling flow like M87. In this case the evolutionary scenario would be:

$$\begin{aligned} \text{E} + \text{Cooling flow} &\Rightarrow \text{E} + \text{LINER} + \text{FRI} \\ \text{or} \quad \text{E} + \text{Blazar}. \end{aligned}$$

In this scenario, evolution determines the nature of the AGN, but aspect determines the type of AGN we

see from outside. In Figure 4, I suggest a taxonomic ordering of the various classes of AGN based on their orientation, the nature of their feeding, the mass of the central BH, and the accretion parameter, given in Figure 2, which measures the rate of feeding of gas into the nuclear regions compared with the Eddington value. We have already distinguished three classes of feeding: major merger-driven inflow in gas rich systems, merging of an elliptical with a low mass companion, and feeding through cooling flows. To these we should add feeding by flows driven by the development of bar-like instabilities in spiral galaxies, which may be the major cause of Seyfert activity. Of course, such bar-like instabilities may also be produced by the tidal forces generated in a close encounter with another galaxy. In Figure 4, I suggest evolutionary paths, and identify prototypes for various classes of AGN. Note that in Seyferts, the opening angle of the outflow (‘ionisation’) cone increases with time as the accreting gas is cleared away, so that the Sy I class may be observed over a wider range of angles in more evolved sources. In general, radio-loud objects, and AGN showing relativistic beaming effects will only be seen for sources lying in the bottom right-hand corner of these diagrams.

8 Conclusions

In this paper, I have established the feasibility of the hypothesis that galactic mergers determine, as much as orientation, the nature and evolutionary stage of the AGN that we observe. The growth of the BH during a major merger event determines the transformation of the nuclear BH from those characteristic of Seyfert galaxies to those such as found in ellipticals. In the case of a merger event between two gas-rich systems, the BH of the merging galaxies can grow from $\approx 10^6 M_\odot$ to $\approx 10^9 M_\odot$. It is also known that such mergers can produce the $R^{-1/4}$ luminosity profile characteristic of elliptical galaxies (Barnes & Hernquist 1996). Such mergers also produce a massive star burst, linking them with the luminous and ultra-luminous infrared galaxies. The size of the BH that can be grown has also been shown to depend on the total gas supply, and therefore upon the mass of the combined system.

For much of the early growth phases, the supply of gas to the region of influence of the BH is likely to be super-Eddington. We show that this is likely to produce an optically-thick radiatively-driven wind. This wind may generate the ‘big blue bump’ in the continuum spectrum, photoionise the broad line region, and may be sufficient to choke off the relativistic jets in this phase, producing a radio-quiet AGN. The fast wind will also interact with the circumnuclear material to form strong shocks which are probably the source of Fermi acceleration of electrons needed to produce the radio core.

For sub-Eddington accretion rates, such as found in either elliptical mergers or in cooling flows, the radio jets can freely escape from the thin accretion disk, so radio-loud AGN and relativistically-boosted BL Lac/blazer phenomena will be preferentially produced in this case. The infalling gas is shocked as it settles down towards the thin inner accretion disk, and it is these shocks which probably produce the classical LINER emission.

This paper has established the outline of a grand unified model, but fails to provide enough detail of the various physical processes invoked. This will have to be the subject of later papers. The following theoretical and observational studies would do much to establish the plausibility (or otherwise) of the model:

- What are the details of the gas dynamics in the inner accretion core? What is the density, mass and size of the molecular clouds here?
- How does the BH interact with these clouds to move into and graze upon them? What is implied about the duty cycle of AGN activity in this case? Do we get repeated episodes of radio jet activity as the AGN moves from one dense cloud core to another?
- What is the competition between the gas infall, the growth of the BH, the disruption of the molecular clouds produced by BH winds and jets, and the fragmentation of the molecular clouds into stars? Can this explain the relationship between the BH mass and the core mass which appears to be observationally established through the work of Lonsdale, Smith & Lonsdale (1995) and Nelson & Whittle (1996)?
- What is the form of the ‘big blue bump’ generated by an optically-thick electron scattering photosphere, and can this reproduce the observations?
- What emission-line spectrum would this produce by photoionisation of the surface layers of the inner accretion disk, and does this reproduce the observations, including the constraints imposed by the reverberation mapping analyses of nearby Seyfert I galaxies?
- What is the synchrotron spectrum produced in the shock generated at the thermal wind/disk interface, and can this explain the radio core observations?
- What is the evidence for hot, thermal plasma in the NLR lobes of radio-quiet objects?
- What are the details of the jet/thermal wind interaction? Can we bury this interaction below a free-free photosphere?
- To what extent is the radio-loud/radio-quiet dichotomy produced by selection effects in the samples used to generate Figure 1? We need an unbiased deep search for radio emission associated with a large optically identified sample of AGN, or a similar study involving faint IR sources that

fall below the IRAS sensitivity limit, in order to establish the true relative space density of radio-intermediate objects.

- Do radio-loud and radio-quiet QSO host galaxies differ in any morphological sense? In our model, the radio-loud objects should look more like normal ellipticals, being in a late merger or post-merger phase, while radio-quiet object should appear more frequently in merger systems.
- According to our jet-driven model, the jets of high-redshift radio galaxies should be confined by the higher ISM density, and in consequence should be shorter, but of intrinsically higher surface brightness (and possibly luminosity) than those in more nearby systems. Is this observed?

Acknowledgments

MD thanks Geoff Bicknell, Ralph Sutherland, Pete McGregor, Charlene Heisler, Agris Kalnajs, Mark Whittle and Wil van Breugel for many stimulating conversations and physical insights in the preparation of this work. The comments of the referees were very welcome, and helped to me to clarify a number of aspects of the paper that were clear only to the author. I also wish to thank the Australian Department of Industry, Science and Tourism (DIST) for support under an International Science and Technology (IS&T) major grant, which covered travel in the preparation of this work. This research has made use of NASA’s (invaluable) Astrophysics Data System Abstract Service.

- Abbott, D. C., & Conti, P. S. 1987, *ARA&A*, 25, 113
 Antonucci, R. 1993, *ARA&A*, 31, 473
 Antonucci, R., & Miller, J. S. 1985, *ApJ*, 297, 621
 Armus, L., Heckman, T. M., & Miley, G. K. 1989, *ApJ*, 347, 727
 Armus, L., Heckman, T. M., & Miley, G. K. 1990, *ApJ*, 364, 471
 Bahcall, J. N., Kirhakos, S., Saxe, D. H., & Schneider, D. P. 1997, *ApJ*, 479, 642
 Baldwin, J. A. 1977, *ApJ*, 214, 679
 Baldwin, J. A., Wampler, E. J., & Gaskell, C. M. 1978, *Nature*, 273, 431
 Barnes, J. E., & Hernquist, L. 1991, *ApJ*, 370, 65
 Barnes, J. E., & Hernquist, L. 1992, *ARA&A*, 30, 705
 Barnes, J. E., & Hernquist, L. 1996, *ApJ*, 471, 115
 Barthel, P. D. 1981, *Adv. Space Research*, 11, 231
 Barvainis, R., Lonsdale, C., & Antonucci, R. 1996, *AJ*, 111, 1431
 Baum, S. A., Zirbel, E. I., & O’Dea, C. P. 1995, *ApJ*, 451, 88
 Begelman, M. C. 1996, in *Cygnus A: Study of a Radio Galaxy*, ed. C. L. Carilli & D. A. Harris (Cambridge Univ. Press), 209
 Bekki, K. 1995, *MNRAS*, 276, 9
 Bicknell, G. V. 1994, *ApJ*, 442, 542
 Bicknell, G. V., & Begelman, M. C. 1996, *ApJ*, 467, 597
 Bicknell, G. V., Dopita, M. A., & O’Dea, C. 1997, *ApJ* 485, 41

- Blandford, R. D. 1990, in *Active Galactic Nuclei*, ed. T. J. -L., & M. Mayor, Saas-Fee Advanced Course 20 (Berlin: Springer), 161
- Blandford, R. D., & Konigl, A. 1979, *ApJ*, 232, 34
- Blandford, R. D., & Rees, M. J. 1978, in *Pittsburgh Conference on BL Lac Objects*, ed. A. N. Wolfe (Univ. Pittsburg Press), 328
- Bondi, H., & Hoyle, F. 1944, *MNRAS*, 104, 273
- Boyce, P. J., Disney, M. J., Blades, J. C., Boksenberg, A., Crane, P., Deharveng, J. M., Maccetto, F. D., Mackay, C. D., & Sparks, W. B. 1996, *ApJ*, 473, 760
- Chiang, J., & Murray, N. 1996, *ApJ*, 466, 704
- Cowie, L. L., Hu, E. M., Jenkins, E. B., & York, D. G. 1983, *ApJ*, 272, 29
- Dopita, M. A., & Sutherland, R. S. 1995, *ApJ*, 455, 468
- Dopita, M. A., Heisler, C. A., Lumsden, S. L., & Bailey, J. A. 1997*a*, *ApJ* (submitted)
- Dopita, M. A., Koratkar, A. P., Allen, M. G., Tsevetanov, Z. I., Ford, H. C., Bicknell, G. V., & Sutherland, R. S. 1997*b*, *ApJ* (in press)
- Falcke, H., & Biermann, P. L. 1995, *A&A*, 293, 665
- Falcke, H., Malkan, M. A., & Biermann, P. L. 1995, *A&A*, 298, 375
- Fanti, R., Fanti, C., Schilizzi, R. T., Spencer, R. E., Nan Redong, Parma, P., van Breugel, W. J. M., & Venturi, T. 1990, *A&A*, 231, 333
- Fanti, C., Fanti, R., Dallacasa, D., Schilizzi, R. T., Spencer, R. E., & Stanghellini, C. 1995, *A&A*, 302, 317
- Ferland, G. J., & Netzer, H. 1983, *ApJ*, 264, 105
- Ford, H. C., et al. 1994, *ApJ*, 435, L27
- Fosbury, R. A. E., Mebold, U., Goss, W. M., & Dopita, M. A. 1978, *MNRAS*, 183, 549
- Gelderman, R. 1995, *PASP*, 107, 205
- Gelderman, R., & Whittle, M. 1994, *ApJS*, 91, 491
- Heckman, T. M. 1980, *A&A*, 87, 142
- Heckman, T. M. 1981, *ApJ*, 250, L59
- Heckman, T. M., Baum, S. A., van Breugel, W. J. M., & McCarthy, P. 1989, *ApJ*, 338, 48
- Heisler, C. A., Lumsden, S. L., & Bailey, J. A. 1997, *Nature*, 385, 700
- Ho, L. C., Fillipenko, A. V., & Sargent, W. L. 1995, *ApJS*, 98, 477
- Hutchings, J. B., & Neff, S. G. 1992, *AJ*, 104, 1
- Johnstone, R. M., Fabian, A. C., & Nulsen, P. E. J. 1987, *MNRAS*, 224, 74
- Kim, D.-C. 1995, PhD Thesis, University of Hawaii
- Kinney, A. L., Rivolo, A. R., & Koratkar, A. P. 1990, *ApJ*, 357, 338
- Kormendy, J., & Sanders, D. B. 1992, *ApJ*, 390, L53
- Koski, A. T. 1978, *ApJ*, 223, 56
- Koski, A. T., & Osterbrock, D. E. 1976, *ApJ*, 203, L49
- Lawrence, A. 1987, *PASP*, 99, 309
- Lawrence, A., & Elvis, M. 1982, *ApJ*, 256, 410
- Lonsdale, C. L., Smith, H. E., & Lonsdale, C. J. 1995, *ApJ*, 438, 632
- McCarthy, P. J., Spinrad, H., & van Breugel, W. 1995, *ApJ*, 447, 77
- Maoz, D. 1994, in *Reverberation Mapping of the Broad-line Region in Active Galactic Nuclei*, ed. P. M. Gondhalekar, K. Horne & B. M. Peterson, ASP Conf. Ser. 69, 95
- Meisenheimer, K., Hippelein, H., & Neeser, M. 1994, in *The First Stromlo Symposium: The Physics of Active Galaxies*, ed. G. V. Bicknell, M. A. Dopita & P. J. Quinn, ASP Conf. Ser. 54, 397
- Melnick, J., & Mirabel, I. F. 1990, *A&A*, 231, L19
- Murray, N., & Chiang, J. 1995, *ApJ*, 459, L105
- Murray, N., Chiang, J., Grossman, S. A., & Voit, G. M. 1995, *ApJ*, 451, 498
- Mushotzky, R. F., Done, C., & Pounds, K. A. 1993, *ARA&A*, 31, 717
- Nelson, C. H., & Whittle, M. 1996, *ApJ*, 465, 96
- O'Dea, C. P., Baum, S. A., Staghellini, C., Morris, G. B., Patnaik, A. R., & Gopal-Krishna 1990, *A&A Suppl. Ser.*, 84, 549
- Orr, M. J. L., & Browne, I. W. A. 1982, *MNRAS*, 200, 1067
- Owen, F. N., & Ledlow, M. J. 1994, in *The First Stromlo Symposium: The Physics of Active Galaxies*, ed. G. V. Bicknell, M. A. Dopita & P. J. Quinn, ASP Conf. Ser. 54, 319
- Petrich, L. I., Shapiro, S. L., & Teukolsky, S. A. 1988, *Phys. Rev. Lett.*, 60, 1781
- Petrich, L. I., Shapiro, S. L., Stark, R. F., & Teukolsky, S. A. 1989, *ApJ*, 336, 313
- Phillips, M. A., Jenkins, C. R., Dopita, M. A., Sadler, E. M., & Binette, L. 1986, *AJ*, 91, 1062
- Phillips, R. B., & Mutel, R. L. 1982, *A&A*, 106, 21
- Rees, M., Begelman, M., Blandford, R., & Phinney, E. 1982, *Nature*, 295, 17
- Reynolds, C. S., Di Matteo, T., Fabian, A., Hwang, U., & Canizares, C. R. 1996, *MNRAS*, 283, L111
- Rowan-Robinson, M. 1977, *ApJ*, 213, 638
- Rowan-Robinson, M., et al. 1991, *Nature*, 351, 719
- Rowan-Robinson, M., et al. 1993, *MNRAS*, 261, 513
- Sanders, D. B., & Mirabel, I. F. 1996, *ARA&A*, 34, 749
- Schmidt, M., & Green, R. F. 1983, *ApJ*, 269, 352
- Scoville, N. Z., Sargent, A. I., Sanders, D. B., & Soifer, B. T. 1991, *ApJ*, 336, L5
- Shakura, N. I., and Sunyaev, R. A. 1973, *A&A*, 24, 337
- Shlosman, I., Frank, J., & Begelman, M. C. 1989, *Nature*, 338, 45
- Simkin, S. M., Su, H. J., & Schwarz, M. P. 1980, *ApJ*, 237, 404
- Smith, E. P., Heckman, T. M., Bothun, G. D., Romanishin, W., & Balick, B. 1986, *ApJ*, 306, 64
- Solomon, P. M., Downes, D., & Radford, S. J. E. 1992, *ApJ*, 387, L55
- Stasinska, G. 1984, *A&A*, 135, 341
- Urry, C. M., & Padovani, P. 1995, *PASP*, 107, 803
- van Albada, T. S. 1982, *MNRAS*, 201, 939
- van Ojik, R., Rottgering, H. J. A., Miley, G. K., Bemer, M. N., Maccetto, F., & Chambers, K. C. 1994, *A&A*, 289, 54
- Veilleux, S., Kim, D.-C., Sanders, D. B., Mazzarella, J. M., & Soifer, B. T. 1995, *ApJS*, 98, 171
- Wilkinson, P. N., Polatdis, A. G., Readhead, A. C. S., Xu, W., & Pearson, T. J. 1994, *ApJ*, 432, L87
- Wilson, A. S., & Colbert, E. J. M. 1995, *ApJ*, 438, 62
- Wright, G. S., James, P. A., Joseph, R. D., & McLean, I. S. 1990, *Nature*, 344, 417
- Wurtz, R., Stocke, J. T., Ellinson, E., & Yee, H. K. C. 1997, *ApJ*, 480, 547