I. OBSERVATIONAL PROPERTIES OF AGN

Active Galactic Nuclei and Related Phenomena
IAU Symposium, Vol. 194, 1999
Yervant Terzian, Daniel Weedman, Edward Khachikian, eds.

An Overview of AGN

Ian Robson

Joint Astronomy Centre, 660 N. A'ohoku Place, Hilo, HI 96720, USA

Abstract.
This paper gives an overview of the global properties of Active Galactic Nuclei, introducing and putting into context those areas that will be more fully described later in this volume. For the new entrant into AGN research, a semi-historical description of how the field developed is also given; showing how far we have come in a relatively short time. The pioneering work is now done, the big picture of how AGNs work is understood, we now embark on the search for the precise details of the emission mechanisms and their interrelations.

1. Introduction

The study of active galactic nuclei is just over fifty years old and in that time tremendous strides have been made in understanding what they are and how they work. Active galactic nuclei are the most luminous objects in the Universe and present a unique opportunity to study the physics of supermassive black holes (SBH). They bring together a stimulating interplay between observation and theory and that is one of the great appeals. Another is that AGN research requires observations over the entire electromagnetic spectrum, from the radio to gamma regimes, including imaging, photometry, spectroscopy, and polarimetry. Flux limitations continually drive for ever larger telescopes, and it is clear that the requirement for the highest spatial resolutions is nowhere near being satisfied. This lack of spatial resolution has meant that variability studies are extremely important as a means of mapping out the inner regions around the central engine, the broad- and narrow-line regions, while also determining the response of other parts of the AGN through delays to other (and usually longer) wavelengths.

So what are AGNs? In a nutshell they are galaxies with specific peculiar properties that *may* include the following: extremely bright and point-like nuclei; variability on timescales ranging from minutes upwards; highly ionized gas moving at high velocities (up to a few thousand km s^{-1}); X and gamma emission; extended radio lobes and jets. However, not all these features are seen in all objects, either because of observational selection effects (such as observer orientation) or intrinsic differences in the sources. Explaining these differences in terms of a unifying model of physical emission processes has been one of the major challenges of the last twenty years.

2. A historical perspective

2.1. The discovery epoch

This overview will be a personal synopsis of the subject and I will begin, primarily for the benefit of the younger astronomers, with a very brief history of the field, selecting those events that for me made major breakthroughs in understanding. The story began in the 1940's when Carl Seyfert found that a handful of spiral galaxies possessed extremely bright nuclei compared to normal galaxies and that their optical spectra showed the presence of broad emission lines of highly ionized species (Seyfert, 1943). This was clearly an unusual and very rare phenomenon, but to determine exactly how rare and how they fitted into the overall picture of galaxies required optical surveys, such as those undertaken by Markarian and colleagues at Byurakan.

Within ten years, the newly emerging field of radio astronomy discovered what were believed to be radio-stars, but which turned out to be something far more interesting. The early radio surveys showed that there were two types of sources; those that were extended and corresponded to supernova remnants in our Galaxy, and compact, unresolved sources that had no known origin. Eventually, good radio positional determination led to the discovery of the optical counterpart of a small number of the latter, and they were point-like, hence the suggestion of radio-stars. With the eventual interpretation of the optical spectrum of one of them, 3C273, quasars burst on the scene in 1963. For a more extensive description on the discovery of quasars see Robson (1996).

Briefly, 3C273 was a very bright, compact, radio source with a jet, and a bright (V=13) optical point-source also with a jet. It showed strong and broad emission lines, and with a redshift of z=0.158, it was obviously not a star in our Galaxy. Assuming the cosmological interpretation of the redshift, which most people did, led to the luminosity being an incredible 5×10^{12} $L_\odot$. As more objects were discovered and past photographic plates were re-examined, it was found that these objects showed variability on timescales of less than a year, hence the emission regions had to be very compact, and the source of energy became an immediate problem. Early observational clues all pointed to a central phenomenon in a galaxy; the point-like nuclei, radio-lobes and the jets. The energy supply solution was soon agreed by most to be the gravitational field of a supermassive black hole. Subsequent observational evidence during the last thirty years has served to underline the correctness of this basic picture, a tribute to the theoretical wisdom of the time.

2.2. The classification epoch

We now enter the realm of what I call astronomical botany, which was absolutely essential, but has also caused confusion for students ever since. The classification of AGNs according to various observational criteria led to: Seyfert galaxies; Markarian galaxies; N-galaxies; radio-loud and radio-quiet quasars; OVV quasars; BL Lac objects; radio galaxies (which were eventually split into FRI and FRII sub-categories); LINERs and Starburst galaxies. The big question is how these objects are related and it is important, especially for new students entering the field to remember the big picture and the big questions; it is only too easy to get sucked into a very narrow channel and lose sight of the overall goals.

2.3. First steps in unification

Very soon, there were observational hints at distinct differences in the categories, and also important similarities. I will give three examples. A major advance was the tentative sub-classification of Seyferts into type 1 and 2 by Weedman in 1970. This was soon confirmed and turned into an entire industry, which eventually led to the classification of 'intermediate' type 1.5, 1.8 and 1.9 objects. This to me clearly pointed the way to a unification scenario rather than two different and distinct types of intrinsic phenomena.

Also, radio-loud BL Lacs when fading to very low levels occasionally showed a type I spectrum (i.e. broad optical emission lines) and this, as well as a number of other pointers, led to the classification of OVV quasars and BL Lacs as Blazars in 1978. (I should point out that while this was helpful at the time, we now believe that the blazar characteristics are the manifestation of a jet-driven phenomenon, and the OVV quasars and BL Lacs are different populations of objects (see Sambruna this volume).

A further key series of observations leading to our modern view of unification were IUE studies of the classical Seyfert 1 galaxy NGC4151. Over a number of years, with the decline in the continuum emission it changed into a Seyfert 2 (Ulrich et al 1984). This was another boost for unification as clearly this showed that the type 1 and 2 objects were not intrinsically different, rather the observational differences were an internal aspect of the object such as energy supply and/or environment.

2.4. The obscuring torus

One of the most important observations for unification was that of the polarized emission from the Seyfert 2 galaxy NGC1068 by Antonucci and Miller in 1985. Unlike the continuum type 2 spectrum, the very weak polarized spectrum was found to show a type 1 form. This provided the first direct evidence for unification by orientation, a suspicion that had been gaining ground for some years. The interpretation was that the polarized emission was revealing a broad-line region that was hidden from our direct line-of-sight, and was only detected due to the reflection from some form of reflecting screen, or 'mirror', such as dust or an electron cloud. The obscuring medium preventing us from seeing the nucleus and BLR directly was attributed to an extended gas and dust torus of dimensions from tens to hundreds of pc. This unification by orientation is now at the corner-stone of current assumptions and has had a high degree of success. (Further examples of NGC1068 will be given by Macchetto (this volume).) In this global picture, if we can see the central engine directly, then it is a type 1 object, with strong X-ray emission and a strong broad-line region. If our line-of-sight intercepts the torus, then we do not see the central engine directly, and it is a type 2 object and spectroscopically characterised by the narrow-line region. Figure 1 shows a cartoon of this scheme.

However, while the obscuring torus and orientation is now the accepted paradigm for type 1 and 2 differences, very recently some doubt has been cast over this simple view with HST snapshots of over 250 Seyfert galaxies (Malkan et

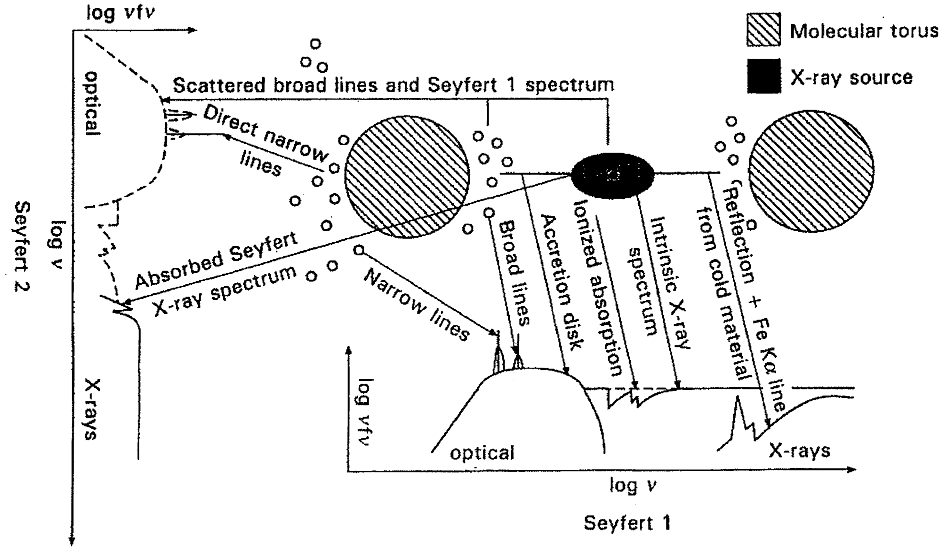

Figure 1. Cartoon of the inner parsec of a radio-quiet AGN showing the main components and possible interaction processes. (Courtesy of Mushotsky, Done and Pounds, Ann.Rev.Astron.Astrophys., 31, 717, 1993.)

al 1998). Their conclusions show that all the galaxies with point-like nuclei are, as expected, Seyfert 1s, but 34 Seyfert 1s have resolved nuclei that are slightly dimmer–both suggesting the presence of dust. Also, both S1 and S2s have central surface brightness similar to bulges of normal spiral galaxies and that both types of galaxy show the presence of dust features in the nuclear regions. So far so good, however, they then go on to conclude that the S1s reside in earlier type spirals than S2s, and if this is the case, then the difference between S1 and S2 is not solely due to orientation, but to something else, something that is host galaxy dependent. The authors conclude that instead of an absorbing molecular torus providing extinction, perhaps galactic clouds of dust may be the answer, being more prevalent in later-type spirals. However, this cannot explain all the type 1 to 2 differences as the absorbing column densities in a number of S2s (for example NGC1068) are known to exceed 10^{24} cm^{-2} and this cannot be produced by aggregates of diffuse clouds along the line of sight. While not being convinced about the host galaxy differences from the data, the presence of intervening dust sheets in the galaxy is a clear clue that even if a torus is present, the picture is probably more complex. Furthermore, while we always picture 'perfect' tori (and there are some excellent examples such as NGC4261), in reality, ragged and leaky tori must also be a possibility.

3. The continuum spectral energy distribution

Turning to the continuum emission from AGNs one striking fact is that for the same IR-optical-X-ray luminosity, two quasars (for example) can have radio powers differing by up to three orders of magnitude (see later). We know the radio-loud emission is due to synchrotron radiation from the core-jet-lobes, which can now be separated out using VLBI imaging. But why only some have these powerful jets creating this huge difference in power remains to be determined.

If extensive dust is present in the parent AGN, this is manifest by a peak of emission (in flux terms) in the submillimetre to far-infrared (see the contribution by Sanders– this volume). The near-infrared is generally the overlap region where the old stellar population of the galaxy shines through (apart from the blazars where the beamed synchrotron overpowers the galaxy contribution). Moving to the optical-UV, we find (for the more luminous examples) evidence for a hot (~30,000K) blackbody, the so-called big blue bump, usually attributed to an accretion disk (but see the contribution by Courvoisier in this volume).

The peak of the luminosity for many AGNs lies in the extended UV regime, one that has been poorly served by observations due to lack of telescopes and the fundamental limitation that neutral hydrogen extinction plays havoc for extragalactic observations. The X-ray regime is very complex, showing a range of properties, and the lack of a baseline to longer wavelengths (lowest energy) hampers the precise interpretation of what the X-ray continuum represents. At the longest wavelength we probably see the tail of the big blue bump plus additional thermal emission from hotter plasma surrounding the central black hole. At higher energies there is evidence of an underlying power-law that may be the initial source of photo-ionization from the black hole. As might be expected, being the region that is probably the primary photon generation territory and deeply buried within the central zone of the accretion process, depending on the object (and orientation) the X-ray spectrum can be made up of direct photon emission, reflected components, and absorbed/re-radiated components. All of this makes disentangling the components of X-ray observations exceedingly difficult and we await the next generation of high sensitivity and high spatial resolution X-ray facilities (AXAF and XMM) with great anticipation.

The gamma-ray region appears to be important only for blazars, where it can easily dominate the bolometric luminosity of the source (although there is bound to be beaming present to distort this picture). The gamma-emission is assumed to be produced by interactions of photons with the radio jet, but the precise mechanism is still not pinned down (see Sambruna this volume). This latter point brings us to the importance of variability. For essentially no change in radio through far infrared luminosity, the gamma-rays can increase by over two orders of magnitude, accompanied by changes elsewhere such as the X-ray and IR. It is only through variability studies that the exact emission mechanism (internal or external self-Compton) can be determined.

4. The zone of the central engine

Major efforts were invested in the 1980's to interpret the wealth of high quality spectroscopic optical-UV data from AGNs. The outcome was that two regions were identified, the broad-line region and the narrow-line region. These were two distinct regions of ionization and hence location. The current picture is still unclear about the exact make-up of these regions, especially the BLR which is now itself split into high and low ionisation zones (see Collin-Zahn, & Sulentic this volume), as the interaction of the primary radiation with the surrounding gaseous envelope and the putative accretion disk is becoming ever more complex. Whatever the precise picture, it is clear that to fuel the central engine, matter must be accreted directly to the central zone. A thin accretion disk is the favourite mechanism (the puffed up thick torus now seems to have been abandoned), although how the inner mass transport works remains unclear. Indeed, firm evidence for an accretion disk is still not totally accepted (see Courvoisier this volume). Nevertheless, disk fuelling and ensuing radiation by the central engine, subsequent photon interactions and re-radiations with the various components of the BLR and the disk give a fairly acceptable, if not precise picture of what the central region probably looks like.

5. Observational evidence for supermassive black holes

Although supermassive black holes are widely assumed to be the driving force, we must look for observational evidence to support this concept. For myself I find three convincing strands. First there is the pioneering studies of Kormendy and others on stellar dynamics. They find convincing evidence for very large, non-radiative objects in the centres of a number of local (and non-active) galaxies (e.g. Kormendy & Richstone 1995). Because alternative explanations are even more complex and contrived, I shall assume this object is a supermassive black hole. The second piece of evidence is the water vapour maser distributions that have now been seen in a number of AGNs. The first was in NGC4258 (Miyoshi et al 1995 and Greenhill et al 1995) and the masers trace out a very thin ring of gas orbiting in a gravitational potential, about 0.1pc from the radio continuum source. The simplest interpretation is that this is a disk of gas orbiting a supermassive black hole of mass $\sim 3\times 10^7$ $M_\odot$. The third piece of evidence is the ASCA observation of the extreme red-broadened wing of the X-ray $K\alpha$ iron fluorescence line in the Seyfert 1 galaxy MCG-6-30-15 (Tanaka et al 1995). This has a velocity width of 100,000 km s^{-1} and is interpreted as high velocity gas in the gravitational field of a supermassive black hole.

6. The Central engine and evolution

In terms of the SBH we now generally believe that the mass and the accretion rate (and perhaps spin) are critical factors in terms of the energy production rate, which can be of order L_{Edd}. It should be stressed that although a SBH is necessary for an AGN, the mere presence of a SBH does not lead to an AGN, because unless fuel can be fed to the hole, then it will be dormant and invisible to photon detection processes apart from its influence on the surrounding material through its gravitational attraction. And as we saw above, we know this must be true due to the presence of SBHs in our own Galaxy and local non-active galaxies (Richstone et al 1998).

This leads us into evolution. We know that AGNs are not active all the time, their energy supply is limited by the available gas in the galaxy and the most luminous quasars have lifetimes of order 10^8y. Indeed, we see a quasar-epoch in the luminosity function, peaking between redshifts of 2 and 3 and then showing a steep decline to higher redshifts. Imaging is now showing that many young quasars show disturbed morphologies and signs of interactions and mergers. This is a key factor. Interactions and merging are clear ways of feeding gas to the SBH and it is well known that interactions and mergers were far more prevalent in the past than at the present epoch apart (see Mirabel this volume for some spectacular 'local' merging systems).

Two observations have opened up new areas of work in the evolution of AGNs: the Hubble Deep Field (HDF) and the discovery of the submillimetre-far-IR background radiation. Observations of the HDF and other regions by the new submillimetre camera SCUBA on the James Clerk Maxwell Telescope have shown that in all probability, this background radiation is produced by point-sources, in this case ultra-luminous IR galaxies (ULIRGs), the super-starbursts (Hughes et al 1998, Barger et al 1998). Furthermore, for the highest luminosity ULIRGs there is compelling evidence that they probably harbour buried AGNs. This evidence also extends to the handful of these objects that have been discovered at redshifts exceeding 2 and 3. The SCUBA data show that there is a population of ULIRGs that have previously been missed by optical-UV studies and that if they also contain AGNs, then the AGN population is also under sampled. However, probably even more important is the effect these galaxies have on the production of metals in the Universe. The submillimetre data show rates of star formation and dust masses at redshifts exceeding 3 of up to five times that claimed from UV studies (see Longair this volume).

Let me now return to what for me remains one of the key unanswered questions about AGNs. What is it that makes a galaxy radio-loud? By radio-loud I mean specifically it has a powerful radio jet. Given the two facts that Seyfert galaxies are (virtually) all spiral galaxies and that radio galaxies exist only in elliptical systems, there has long been the temptation to assume that that it was the host galaxy morphology that governed the production of the jet (either by black hole properties or a dense interstellar medium quenching the jet). Ergo, radio-quiet quasars resided in spiral galaxies. Work that I was involved with a few years ago tested this paradigm through highly accurate IR imaging from UKIRT (Taylor et al 1996). While the sample was not large and was restricted to the top-end of the luminosity function of radio galaxies, radio-loud and radio-quiet quasars, the results were clear. Radio-loud galaxies and radio-loud quasars were housed exclusively in ellipticals while radio-quiet quasars have both elliptical and spiral hosts (although there was a suspicion that as the luminosity increased all the hosts tended to become elliptical). Furthermore, in contrast to data from the HST at the time, quasars did not reside in galaxies of luminosity $<$L*. Given a few assumptions it was tantalising to note that a good fraction ($\sim$100L$>$2L* might have been quasars in the past.

This work has now been extended using the HST and has produced powerful data (McLure et al 1998). These support all the earlier conclusions and go further by clearly showing that the most luminous of the radio-quiet quasars also reside in elliptical systems. This demonstrates that host galaxy morphology is not the sole answer to the radio-quiet question. (It should also be noted that high sensitivity radio observations have shown small radio jets in a number of nearby AGNs.) Furthermore, the elliptical hosts are identical in almost all global properties to normal 'local' giant elliptical galaxies, except that nearly all show tidal tails or signs of interactions, again suggesting a possible trigger mechanism for activity. The results are also in excellent agreement with the relationship between the mass of the spheroidal bulge component of the galaxy and the mass of the black hole (Magorrian et al. 1998). Using the observed luminosity of the spheroidal bulge of the quasars and the normal elliptical M-L ratio gives the M_{sph} and hence the - *expected* M_{SBH}. Using this to infer L_{Edd} shows L $\sim$ 0.1-1.0 L_{Edd}. This strongly suggests that for RQQs and RLQs the optical luminosity arises from a similar process of accretion onto a SBH. Furthermore, in this picture RLQs and RQQs of similar R-band magnitude have black holes of comparable mass ($M > 3\times10^9\ M_\odot$), i.e. SBHs can only be housed in galaxies with a spheroidal bulge mass of $M > 5\times10^{11}\ M_\odot$.

The results also agree well with the radio-luminosity to black-hole mass correlations of Franceschini, Vercellone & Fabian (1998). The RQQs also show a correlation, albeit better with radio-core luminosity, which tends to give lower M_{SBH} (by a factor of ~2) than using the above technique of M_{sph}. If this conclusion is correct, it implies that *no RQQ* in the sample has a $M_{SBH}>10^{10}\ M_\odot$ which in turn suggests the RQQs have lower mass black-holes ($\simeq 10^9\ M_\odot$) radiating close to the Eddington limit. So, we have that (a) if radio luminosity (5 GHz P_{core}) is the better predictor of the black-hole mass, then the FRII radio sources simply require black holes of $M > 10^{10}\ M_\odot$, or, (b) if M_{sph} is the better predictor, then some of the RQQs have $M_{SBH} > 10^{10}\ M_\odot$ and so some other interpretation is required such as SBH angular momentum, for example to explain why black holes of comparable mass can have radio powers differing by two orders of magnitude for a similar optical luminosity.

7. Conclusions

I hope the above has given a flavour of where we are and how we got here. While we have a relatively good overall picture of what makes an AGN and how it works, there are still some major questions remaining unanswered, which means that the field of AGN research will continue to push observational and theoretical ingenuity for many years to come. For me these questions are: the conditions required for the formation of SBHs; the cosmological evolution of galaxies and AGNs; the key parameter(s) that determine what 'type' of AGN we see, in particular the radio-loud, radio-quiet split; how are powerful radio jets produced.

References

Antonucci,R & Miller,J., 1985. ApJ, 297, 621

Barger,A. et al., 1998. Nature, 394, 248

Franceschini, Vercellone & Fabian,A., 1998. MNRAS, 297, 817

Greenhill et al., 1995. ApJ, 440, 619

Hughes,D.H. et al., 1998. Nature, 394, 241
Kormendy,J & Richstone,D., 1995. ARA&A, 33, 581
Magorrian et al., 1998. AJ, 115, 2285
Malkan,M.A., Gorjian,V., & Tam,R., 1998. ApJS, 117, 25
McLure,R.J. et al., 1998. Astro-ph 9809030
Miyoshi,M. et al., 1995. Nature, 373, 127
Richstone,D. et al., 1998. Nature, 395, A14
Robson,I. 1996. *Active Galactic Nuclei*, Wiley-Praxis Press
Seyfert,C.K. 1943. ApJ, 97, 28
Tanaka,Y. et al., 1995. Nature, 375, 659
Taylor,G.L., Dunlop,J.S., Hughes,D.H. & Robson,E.I. 1996. MNRAS, 283, 930
Ulrich,M.H. et al., 1984. MNRAS, 206, 221
Weedman,D.W. 1970. ApJ, 159, 405