13 Radio jets and radio galaxies

The idea of astrophysical jets first attracted interest as a theoretical prediction. Double-lobed radio galaxies (RGs) were detected in the first radio surveys (in the 1960's?). The data available then showed two radio-loud lobes (synchrotron sources, containing magnetized plasma and relativistic particles) on either side of a central, elliptical galaxy. Energetics and lifetime arguments soon found that the lobes were short-lived.[1] Either we were seeing them at a very special time in the life of the parent galaxy, or they were being resupplied with energy by an undetected pipeline: a radio jet. These jets were detected when the instruments improved – this was one of the first successes of the VLA. Models of active galactic nuclei – involving accretion onto a black hole – had to be amended to include the production of highly collimated, relativistic plasma jets, *if* the AGN happened to live in an elliptical galaxy.

We've learned a lot since then. We now know that jets are common on both small and large scales. On large scales, the massive black holes in galactic nuclei can produce jets while they are "active". Radio galaxies and (radio-loud) quasars, which live in elliptical galaxies, are the most dramatic examples. In these objects, the galactic nucleus contains a bright, compact radio core and a pc-scale, synchrotron-bright jet. The radio core is thought to be optically thick synchrotron emission from the base of the jet. On larger scales, the jets extend to at least several kpc, often farther; they connect to larger lobes or tails which arise from the interaction of the radio jet with the local extragalactic plasma.

On smaller scales, we now know that many star-sized galactic sources have relativistic jets. The most well-studied (and well-imaged) are jets from accretion flows in X-ray binary systems; SS433 is the prototype, and a few dozen are know known (*microquasars*). In addition, a few pulsars/PWNe systems are found to have jet outflows: the Crab and Vela pulsars are examples here. Finally, it is now thought that Gamma-Ray Bursters (GRBs) involve highly relativistic jets, and that core-collapse supernovae produce short-lived jets during the explosion. Also on stellar scales, but moving to less relativistic systems, we now know that the collimated outflows are produced during part of the collapse of a protostellar cloud to the final star; these are called (*protostellar jets*).

While all of these jets are interesting, in these notes I'll focus on what I'm most familiar with, namely, relativistic jets from AGN. My goal is to present an overview of the observations, some basic physics, and the current "cartoons" as to how RGs work.

13.1 Jets: the observational constraints

To start, what are the important properties of jets which theory must accound for? In these notes, the discussion is strongly skewed toward extragalactic jets (radio galaxies and quasars), and relativistic galactic jets (microquasars), which is my personal interest and area of experience. Much of the basic dynamics apply to all jets, but some of the details and constraints are more relevant to relativistic jets.

Some critical facts and problems are:

• **Internal energy.** Radio jets from AGN and microquasars are dominantly synchrotron sources – thus we know they are magnetized and internally relativistic. Protostellar jets and some galactic jets are dominated by cooler, thermal radiation: emission lines from ionized gas, and even molecular emission.

• **Collimation.** These jets usually have very small opening angles – no more than a few degrees – and often retain their direction and collimation for a distance which is orders of magnitude larger than the scale on which they were initially produced.

• **Knots and bright spots.** Jets are rarely smooth when seen in high-quality images. Bright features are common. We don't know just why these features appear, but a mixture of shocks and strong-amplitude waves in the flow are likely. Note also that the radiation mechanisms are strongly nonlinear amplifiers: a weak density or B field enhancement, for example, can cause a strong enhancement of the emissivity.

• **Speed.** There is indirect evidence that many jets in RGs are supersonic (relative to themselves): structures are seen at their outer ends, where they run into the ambient medium, that can be identified as shock fronts within the jet. In addition, on pc scales, some jets (or at least waves in the jets) are moving at relativistic speeds; this is inferred from the apparent proper motion of knots in the jets, which can exceed lightspeed.

---

[1] Equipartition calculations give you the minimum energy in a radio source; divide this by its radio power, and you get a "lifetime" – which is short compared to any plausible estimate of the source age. Thus, you need energy resupply.

There's no compelling evidence for relativistic motion in kpc-scale jets; somehow the high-γ pc-scale flow has slowed down by the time it reaches kpc scales.

• **Plasma content**. Protostellar jets appear to be normal ISM; an ion-electron mixture. Compact-object and extragalactic jets are observed in synchrotron radiation, and therefore contain relativistic electrons and magnetic fields; we do not know if the electron charge is balanced by ions or by positrons.

• **Energization.** In at least some extragalactic jets there is a clear need for *local* re-acceleration of the relativistic particles. The synchrotron lifetime of the highest energy particles is significantly less that the shortest possible travel time down the jet (jet length / lightspeed). Somehow, local fluid/plasma processes must transfer energy from the bulk flow to the individual particles; a mixture of shocks and turbulence is probably the answer.

13.2 Some useful relativity

Two relativistic effects are important here.

13.2.1 Superluminal motion

The bright knots in many jets are observed to have apparent proper motion above lightspeed. As you've seen before (for instance in Caroll & Ostlie), this is a simple consequence of light-travel times. If the jet is oriented at angle θ to the line of sight, and has physical speed βc, its apparent speed – as seen by the observer – is

$$\beta_{app} = \frac{\beta \sin\theta}{1 - \beta\cos\theta} \qquad (13.1)$$

It's easy to show that this can lead to $\beta_{app} > 1$ for $\theta \ll 1$. Typical observed values are $\beta_{app} \sim$ a few; you should be able to work out what (γ, θ) values are needed for this.

13.2.2 Doppler beaming

A source of radiation is moving relativistically, at some γ, at angle θ to your line of sight. One effect, which you remember, is a Doppler shift. If the source emits photons at ν', you observe photons at ν, where the observed and emitted frequencies are related by $\nu' = \gamma\nu(1 - \beta\cos\theta)$ (note, the angle θ is measured in the observer's frame; and $\theta \to 0$ means motion towards the observer). This can be written,

$$\nu' = \frac{\nu}{\mathcal{D}}\ ; \qquad \mathcal{D}(\theta) = \frac{1}{\gamma(1-\beta\cos\theta)} \qquad (13.2)$$

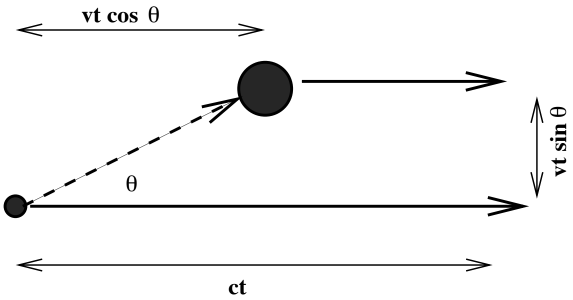

Figure 13.1 Illustrating how a feature in a jet can appear to be superluminal. The core source (small circle at left) emits a "blob" (large circle) at time $t = 0$, at angle θ to the line of sight; it also emits an EM signal directly towards us. After time t, the blob has moved a distance vt, and emits a second signal. The difference in arrival times between the first and second signals is $ct - vt\cos\theta$; the apparent separation of the core and blob is $vt\sin\theta$. Dividing the latter by the former gives us the apparent velocity of the blob, as in (13.1).

where $\mathcal{D}$ is called the *Doppler factor*. A second important fact is how to connect the spectral intensity you observe, I_ν, to that emitted by the source, $I'_{\nu'}$. It turns out that the two quantities are related by

$$\frac{I_\nu}{\nu^3} = \frac{I_{\nu'}}{(\nu')^3} \qquad (13.3)$$

(*cf.* Rybicki & Lightmann for the derivation of this).

As an example, let our emitting soure have a power-law synchrotron spectrum, say $S'_\nu \propto \nu^{-\alpha}$ in the rest frame. When you work out the details, an unresolved blob is Doppler boosted by $S_\nu(\theta) = S'_\nu \mathcal{D}^{3+\alpha}$; you get 3 $\mathcal{D}$'s from the frequency ratio in (13.3), and another "α of a $\mathcal{D}$" from the spectrum. Alternatively, a piece of a resolved jet is boosted by $S_\nu(\theta) = S'_\nu \mathcal{D}^{2+\alpha}$; time/space contractions applied to the piece you're resolving account for the change. Either way, because of the strong dependence of $\mathcal{D}$ on θ, this result means that an observer sees the radiation forward-beamed, into an angle $\sim 1/\gamma$. This makes a relativistic jet appear much brighter if you see it end-on.

13.3 Some useful physics

In this section I store a smattering of physical arguments we can make about jets.

13.3.1 Collimation

How to jets stay so well collimated? Why do they not expand? If the jet were propagating in vacuum, it would expand at its internal sound speed. But we observe jets which remain very collimated over very large distances, with an opening angle only a small fraction of a radian. This would require the flows to be very cold internally, which is inconsistent with other evidence that many of the jets are internally hot. Thus we believe the jet is confined. The two possibilities are

- **Confinement by external pressure.** We know jets propagate through external plasma – the ISM for microquasars, the extragalactic medium for jets from AGN. The external plasma may provide enough pressure to confine the jet.
- **Self-confiment** by magnetic fields. We've already seen that a current generates an azimuthal magnetic field, which can confine the plasma carrying the current. If this applies to jets, the question is then, where and how does the current return to the source?

Which of these two operates can, in principle, be learned from observations; and a given jet may change from being self-confined (close to its origins, say) to being pressure confined (farther out).

13.3.2 Jet transport

Some simple models of jet/RG evolution are based on the rate at which mass, momentum, and energy flow down the jet. Looking back to last term (when we did mass and momentum conservation .. right?), and earlier this term (for energy conservation) we can use the basic fluid equations, and/or simple common sense, to write down the rates. Let the jet have radius r_j, speed v_j, density ρ_j, and enthalpy $h_j = e_j + p_j/\rho_j$ (this is a useful way to collect terms involving the internal energy and pressure of the jet fluid). If everything is subrelativistic, and the magnetic field can be ignored for the moment, we have

$$\text{mass flux}: \quad \dot{M} = \pi r_j^2 \rho_j v_j \tag{13.4}$$

$$\text{momentum flux}: \quad \dot{P} = \pi r_j^2 \rho_j v_j^2 \left(1 + \frac{h_j}{c^2}\right) \tag{13.5}$$

$$\text{energy flux}: \quad \dot{E} = \pi r_j^2 \rho_j v_j \left(h_j + \frac{1}{2} v_j^2\right) \tag{13.6}$$

If t he flow is relativistic, but still ignoring B, these become (remember $v = \beta c$ and $\gamma = (1-\beta^2)^{-1/2}$:

$$\text{mass flux}: \quad \dot{M} = \pi r_j^2 \rho_j \gamma_j \beta_j c \tag{13.7}$$

$$\text{momentum flux}: \quad \dot{P} = \pi r_j^2 \gamma_j^2 \beta_j^2 \rho_j \left(c^2 + h_j\right) \tag{13.8}$$

$$\text{energy flux}: \quad \dot{E} = \pi r_j^2 \gamma_j^2 \beta_j c \rho_j \left(c^2 + h_j\right) \tag{13.9}$$

If the flow is strongly magnetized, we need to include the field energy in the bookkeeping; you remember that electromagnetic energy is transported in a Poynting flux. The details of this are complex and more than we need here; I'll return to this general idea below.

13.4 Larger Scales: the Radio Galaxy

If a jet propagated into vacuum, we might imagine that it would remain unchanged, carrying on forever. But jets don't live in vacuum; rather they propagate into the surrounding plasma – which can be relatively dense intracluster medium (ICM: if the parent galaxy lives in a cluster), or the tenuous intergalactic medium (IGM; if the parent galaxy lives in the "field"). So: when the jet interacts with the surrounding medium[2] the nature of this interaction shapes the Radio Galaxies (RGs) that we see. Observed RGs tend to fall into two morphological types[3], suggesting two different physical situations.

13.4.1 Classical Double radio galaxies (FR II's)

Radio galaxies classified as FR II's – cartooned in Figure 13.2 – are identified by their symmetric, two-sided "lobes" which have bright "hot spots" at their outer edges. Often a narrow, well-collimated jet can be seen propagating through the lobe and almost to the hot spot. These tend to be the brightest ones – so that, even though they are rare by number (only a few per cent of the population), they have received the most attention.

Here's the basic scenario for this type of source. Put a massive black hole down in the center of a galaxy, and turn it on. The jet propagates out, into the ICM; as time goes on, the head of the jet will reach farther and farther from the AGN. We can use simple conservation laws to estimate how this souce evolves with time. The mass and energy flowing down the jet carry momentum; this momentum flux allows the end of the jet to make its way into the external medium (with density ρ_x). Think about a simple ram pressure balance: if the end of the jet advances at $v_D = dD/dt$, it sees a head-on ram pressure $\rho_x v_D^2$. Balancing this against the momentum flux in the jet tells us what v_D must be.

---

[2] I'll refer to the ambient medium as the "ICM" for short ... meaning ICM/IGM.

[3] Jargon: two authors, Fanaroff & Riley, first pointed out this duality – so RGs today are still usually called "FR I's" (FR Type I's) or "FR II's" (FR Type II's).

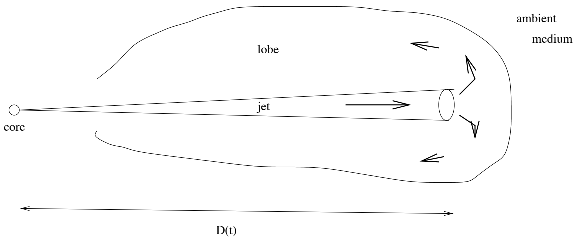

Figure 13.2 Cartoon of (half of) an FRII radio galaxy. Let a jet have constant opening angle, propagating into an external medium (at density ρ_x, say). Because the length of the jet does not grow as fast as the plasma speed within the jet, the plasma must slow down (possibly shock down), and move "sideways", creating a larger "lobe" which surrounds the jet. Because the shock compresses the jet plasma and B field, and accelerates relativistic particles, we expect it to be a bright synchrotron source – *i.e.*the "hot spot" seen in many such sources.

If the jet is less dense than the external medium, then v_D will be less than the jet speed; if the jet is also supersonic we expect a shock to form at this transition point.

Two things should happen at this shock. First, the jet plasma will be compressed and heated when it shocks down. The higher plasma density, and higher B field, will make the post-shock plasma a stronger synchrotron source; if any particle acceleration takes place at this shock, that will further enhance the synchrotron power. Thus, the post-shock plasma should be a localized "hot spot" – matching nicely with what we observe in classical double RGs. Second, the shocked plasma must go somewhere – we expect the high post-shock pressure to "push it off to the side". As the RG grows, this shocked jet plasma will expand to fill a "lobe" or "cocoon" which surrounds the jet. This also matches what we see – in fact the lobes are the brightest, defining, parts of this type of RG.

Recent observational note: if the jet is strong enough, we expect its advance speed, v_D, to still be supersonic relative to the ICM. If this holds, the advancing jet will drive a bow shock in the ICM – such shocks have now been seen in a few cases.

13.4.2 Tailed radio galaxies (FR I's)

Radio galaxies classified as FR I's – as cartooned in Figure 13.3 – are characterized by long, diffuse-looking "tails". These tails begin close to the AGN (sometimes you can see a well-collimated inner jet that suddenly changes to a broader tail; other times you can't), and carry on – broadening gently – for up to 100s of kpc. Unlike FR II's, which are brightest at their outer hot spots, FR I tails tend to be brightest close to the galaxy; the tails gradually become fainter going away from the core. The physical end of the flow – where it runs into the ICM – may or may not be visible.

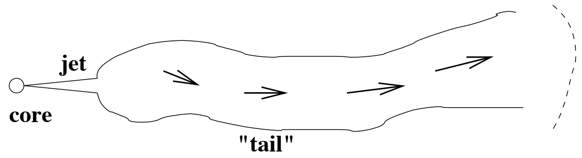

Figure 13.3 Cartoon of (half of) an FRI radio galaxy. When the jet initially leaves the AGN, it is well collimated (probably relativistic and supersonic as well, at least in some cases). But it soon destabilizes – sometimes very suddenly, as in this cartoon. The plasma flow carries on, away from the AGN, but in a more disorganized way.

FR I's are much more common than FR II's; most RG's we know are FR I's. But FR II's have been better studied, partly because they are the bright ones (so were the first ones to be well-observed), and partly because they seem to involve simpler physics.

So, what is the physical picture for FR I's? The cartoon above, for FR II's, relies on the jet remaining stable, collimated, and (internally) supersonic.[4]

The situation will be different if the jet becomes unstable, uncollimated, or subsonic. In this case, the jet flow not retain the "jet/lobe" morphology characteristic of FRII's. Instead, we expect the jet flow to be much more sensitive to local conditions in the ICM. The flow can be affected by local ICM "weather" (flows, turbulence in the ICM); if the parent galaxy is moving rapidly through the local ICM, we can see strongly bent radio tails. If the flow is slow enough, and less dense than its surroundings (which is almost always the case), it will also be affected by buoyancy.

Clearly a wide range of radio-tail morphologies is possible here, depending on local conditions (ICM weather), and also the details of the jet flow. However, the range of possible flows should have one thing in common: they are probably brighter synchrotron sources closer to the AGN (whereas FRII's are brighter closer to their hot spots). We expect this because the radio tail usually expands as it propagates – leading

---

[4] These conditions are related – it turns out that subsonic, or subalfvenic, jets can easily be destabilized as they pass through a surrounding plasma. (Think of a firehose flapping side to side if the pressure gets high enough ...) Supersonic/superalfvenic jets, on the other hand, tend to be relatively stable in the same situation.

to lower plasma density and B field – thus lower synchrotron power. Synchrotron aging may also matter – the ends of the tails can simply fade away as the relativistic electrons lose their energy.

Simple dynamical models (as we have for FR II's) have yet to be developed for FR I's. While we expect the basic momentum-conservation picture, from above, to hold here, the added dynamical effects of the interaction with the ICM make it harder to find simple models of FR I evolution.

13.5 Unresolved issues

While much of the above picture has remained stable for quite awhile, some issues and questions are getting new attention.

13.5.1 What is the life cycle of a RG?

Two questions come to mind here.

First, what is the duty cycle of a radio-loud AGN? If $\sim 10\%$ of bright ellipticals host a radio galaxy, does that mean that the black hole in every such galaxy is active only 10% of the time? Do AGN, or RG's, alternate between "on" and "off" phases?

Second, where are the old radio galaxies? The simple models described above predict that RG's should keep growing, and remain relatively bright synchrotron sources (*i.e.*, detectable), as long as the jet stays "on". If the jet turns off, the same models predict that the RG's synchrotron luminosity should slowly fade, as the relativistic electrons lose their energy. Because the synchrotron lifetime tends to be long for typical RG conditions, this fading should be slow – we should observe a fair number of "old" RG's (jetless, probably steep radio spectrum).

But we don't see what these models predict. We almost never see RGs that could be called "old"; they all have currently-active jets. In addition, dynamical models (such as you'll see in the homework) generally estimate RG ages only $\sim$ 100 Myr – much younger than the age of the parent galaxy. So: why don't we see either older, jet-on sources, or old, jet-off sources?

13.5.2 How does a jet affect its environment?

This is a very active current area. Clearly the environment affects the AGN; accretion from the *local* environment is what makes the massive black hole "active", and creates the jet. We know the jet transports mass, momentum and energy out from the AGN: what effect does this transport have on the environment?

If the jet/RG is well-coupled to its environment – and that's a big "if" – then the jet/RG can have a strong effect. It carries enough energy to heat the local ICM significantly. Two applications here:

•**Cool cores in clusters of galaxies**. In most clusters (as we've seen in the homework), the plasma density is low enough that the radiative cooling time (from bremsstrahlung) is longer than the Hubble time. Thus, most of the plasma in most clusters hasn't cooled down by much since the cluster first formed. In some clusters, however, the central plasma is dense enough that radiative cooling does matter; these "cool cores" must have an ongoing heating mechanism. All of these cool cores are observed to be centered on a currently-active AGN and RG; do these RG heat the cluster core enough to offset radiative losses?

•**Feedback in galaxy formation**. We now know that a massive black hole sits at the heart of every galaxy. We suspect that black hole formed, and was "active", at about the same time as the galaxy originally formed (we'll discuss this later, in chapter 15). So: did the energy released by that actively accreting black hole play an important feedback role in the galaxy formation process?

Both of the above questions are getting a lot of attention these days. In my opinion, the big, unanswered physics question is, "how effectively does the jet/RG couple to its environment?" That is: how much of its energy does the radio jet/RG actually transfer to the local ICM? The answer depends on the details of how the RG grows and evolves. Does it simply do "pdV" work as it pushes the ICM/ISM out of its way? Does it drive significant shocks or turbulence in the ICM/ISM (which then dissipate and heat the ICM/ISM)? Does the relativistic plasma mix effectively with the ICM – thus releasing the relativistic plasma directly into the ICM/ISM? None of these questions have been answered yet.

13.6 How are jets made?

Finally, a few words about how this all starts.

It seems very likely that every accretion flow involves a jet outflow. And, nearly all jets that we know about are tied to accretion flows (the one exception might be jets from single pulsars; we don't yet know much about them). However, we don't yet have a clear and agreedupon picture of how jets get formed. Models out there can be grouped into two broad categories, fluid-based and MHD-based. I think the MHD models are most likely to prove correct; but will include a brief discussion of wind models as well.

13.6.1 Wind (fluid-based) models

These models are probably not the right answer, and pretty much disregarded in recent work. However they were the first type of model proposed; and the internal physics may well still be useful in the more modern MHD models, below. To think about wind models, remember the solar wind: it accelerates from a very slow start, to supersonic speeds at large distances. It does this smoothly, by passing through a "gravitational nozzle" at the sonic point. The wind is driven by its internal energy (and thus, ultimately, by whatever heats it). Linear, one-dimensional flows can also undergo a smooth transition, if the area of the confining channel has a minimum at the sonic point. The first models of astrophysical jets used this analogy – arguing that the flow is accelerated by its internal energy, and that some combination of channel geometry (provided perhaps by the walls of a fat accretion disk?) and gravity produce a high-speed, somewhat collimated outflow.

13.6.2 MHD models

In this discussion I'm partly following several recent papers[5] and addressing jet formation from magnetized accretion flows around a black hole.

What might the magnetic field of an accretion flow look like? Think about a field which threads the disk plasma, but is also "tied" to the distant ISM, which is rotating much more slowly than the accretion flow. The accreting gas will draw the field with it, imparting a ϕ component to the field. One might expect dissipative processes to balance the field growth induced by the accretion, leading to a field shaped something like an expanding helix. Plasma ejected from the disk (for instance in a wind) will be both channeled and accelerated by such a field. Magnetic pressure gradients up the rotation axis will accelerate plasma "up and out" of the system; magnetic tension (from the helical field lines) will exert a "hoop stress" and confine the plasma as it goes.

There are also a couple of variants worth mentioning.

- **Poynting flux models.** These describe the limit in which the plasma density close to the source is small compared to the field energy. As we've seen, a rapidly rotating $\mathbf{B}$ field generates an $\mathbf{E}$ field (just as in pulsars). There will very likely be a component of the Poynting flux, $\mathbf{S} \propto \mathbf{E} \times \mathbf{B}$, along the rotation axis, which will be a strong source of power lost from the accretion system. In order to produce what we observe as radio jets, these Poynting-flux jets must "mass-load"; they must either pick up stray charges from the ambient medium, or generate their own plasma through magnetized pair production. Such a process may well account for the origin of relativistic jets close to a black hole; they would then become visible as they gained mass.

- **Penrose jets.** Another class of models extracts rotation energy from the black hole to power the jets. (I'm inventing the name; the authors of two similar models are Blandford & Znajek, and Punsly & Coroniti). Think back to rotating black holes. You recall that *frame dragging* is important near the event horizion – this is the azimuthal motion induced by the hole's rotation. Think now about a piece of magnetized accreting matter, with field lines again tied to some slowly-rotating distant point. As the matter gets close to and crosses the event horizon, frame dragging will speed it up, thus generating helical (Alfven) waves which move out along the field lines. This will again generate an outwards Poynting flux (which can mass-load and become a visible jet); and the reaction force ends up making the plasma counter-rotating as it passes through the event horizon. Thus some of the black hole's rotation is lost, to supply the power carried out by the yet.

13.6.3 Duty cycles?

Accretion flows in galactic X-ray binaries provide a possibly interesting clue. Think back to our discussion of accretion disks, from last term: they are mostly thermal (Black body) sources, possibly with an optically thin corona (could be a bremsstrahlung source). We found that accretion disks around small (star-sized) compact objects becomes hot enough to radiate in the X-ray band. It turns out that X-ray emission from galactic binaries has two states. It can be "low/hard", meaning lower X-ray power and a harder (nonthermal?) X-ray spectrum. Or, it can be "high/soft", meaning higher X-ray power and a softer (thermal) X-ray spectrum. It turns out that steady radio jets are present

---

[5] e.g. Meier, D. L. 2005, *Astrophysics & Space Science*, 300, 55; Meier, D. L. & Nakamura, M. 2006, ASP Conference Series, 350, 195

in the low/hard state, but *not* in the high/soft state. The difference between the two states is thought to be the accretion rate – a lower $\dot{M}$ leads to lower X-ray power (which is constent with the simple accretion-disk models we saw last term). Thus: perhaps a lower $\dot{M}$ somehow changes the accretion mode in galactic microquasars, and allows a jet to form?

It is not clear whether this scenario also applies to AGN. There are arguments on both sides, and the observations don't (yet?) support the existence of two such states in AGN. If AGN do turn out to work this way, we may have a physical origin for AGN duty cycles – but to my mind, this issue isn't settled yet.

Key points

- Jet phenomenology: what we know from observations.
- Important relativity: superluminal motion and beaming.
- Basic jet "fluid physics": collimation, propagation.
- Radio galaxy types & cartoons: FR I's, IIs
- How does the jet affect its surroundings?
- Current ideas of jet origins – mostly MHD.

14 Quasars and Active Galactic Nuclei

AGN astronomy started in the 1960's. Early radio surveys had found bright, compact radio object with no clear optical identification. Most people thought they were simply radio-loud stars, but some thought they were extragalctic ... in 1963 a spectrum was obtained of the radio source 3C273. The likely optical ID was a 13th magnitude blue object, apparently stellar (not extended), with faint linear emission nebulosity. This turned out to have very unusual spectra for a star, rich with emission lines. These were thus called "quasi-stellar objects" (QSOs), or quasars for short.

Identifying the emission lines was hard, until someone realized the object was at the very high (at the time) redshift, $z = 0.16$. After that, people started looking seriously for these objects. The bright emission lines and blue continuum were easy to find. By now we know of several thousand...number counts find about 100 quasars per square degree of sky, with $z \lesssim 2$ and blue magnitude $m_B < 22$ (with more, of course, at fainter magnitudes and higher z).

14.1 Basic properties: observations

Although we now know (or think we know) that all the varieties of AGN are driven by accretion onto a massive black hole in the core of the galaxy, that was not at all obvious when the field started. Thus, the literature is dominated by the unusual observational properties of bright AGN.

14.1.1 Spectral lines

Strong optical emission lines are characteristic of (almost) all AGN. They are more or less the same lines that we see in galactic nebulae – planetary nebulae and HII regions. Emission lines from AGN seem to obey very similar physics – and thus arise in similar conditions. They are probably ionized by the observed nuclear continuum source (the so-called "central engine"; AGN are strong in the UV/soft Xray region), although shock ionization is also likely in some cases.[1] The striking feature of the emission lines is their width. Permitted lines (such as Hα, the $n = 3-2$ transition of H) have typical linewidths $\Delta v = c\Delta\nu/\nu \sim 10^4$ km/s. The forbidden lines (from heavy elements; OII, OIII, NII, etc.) can be narrower, more like $\Delta v \sim 10^3$ km/s. These widths are much too high to be due to thermal broadening;[2] these widths must be due to bulk motion of the emitting gas clouds – either random or orbital motion.

The difference between broad and narrow emission lines seems to suggest that the two line types are formed in different regions – perhaps the permitted lines come from denser gas, closer to the central engine while the forbidden lines come from less dense gas, a bit further from the engine. Not all AGN have this separation, however; some Seyferts have mostly narrow-line emission, while some quasars have mostly broad lines.

14.1.2 Continuum emission

AGN radiate in every frequency in which we've looked: from radio, through infrared, to optical and UV, thence to X-rays and γ-rays. The underlying spectrum is broad-band, with 2 or 3 separate features. Remember your radiation: telescopes generally measure intensity, f_ν say, which is power/Hz. The energy in the spectrum is better determined by $\int f_\nu d\nu \sim \nu f_\nu$; this is often what's plotted. The total emission νf_ν is typically dominated by two (or maybe 3) peaks: one IR/optical (which itself may be two components), and a second peak in the hard X-ray to γ-ray region. It seems likely that the bulk of the energy comes out in the optical/UV region, say $\nu \sim 10^{14} - 10^{16}$ Hz. (This dominant "bump" in the νf_ν plot is called the Big Blue Bump, I kid you not.) However, some radio-loud sources have been observed out to hard γ-rays; their broadband spectra are dominated by the hardest frequencies, $\nu \sim 10^{20} - 10^{25}$ Hz (compare: what frequency corresponds to the electron rest mass?)

---

[1] Detailed analysis of the emission line strengths can tell us the likely cause of ionization, as well as the local density and temperature of the emission line gas. One piece of terminology is necessary. Remember that quantum mechanics gives us the probability that an atom in an excited (upper) state can make a spontaneous transition to a lower state, emitting a photon in the process. *Permitted lines* have high transition probabilities; what are called *forbidden lines* have lower transition probabilities – but are not forbidden in the QM sense. Rather, their radiative transition rates are low enough that in dense conditions they will be de-excited by collisions first. Thus, there is a *quenching density* associated with each particular transition.

[2] how hot would the gas be? how could ions such as OII or OIII ever exist at such temperatures? how low would the hydrogen recomination rate be, and how would you ever see hydrogen permitted lines?

14.2 The AGN zoo

We now know that quasars are unusual nuclei in otherwise normal galaxies. Other types of nuclear activity are also possible; for historical reasons (who found what first, in what observing band), a wide variety of names and acronyms exist in this field. All of these objects, as well other subclasses and acronyms I'm not bothering to put down, are grouped together as Active Galactic Nuclei (AGN).

14.2.1 The radio-quiet ones

Although quasars were originally detected based on their radio emission, radio-loud quasars turn out to be rare, something like 10% of the quasar population. Radio-quiet quasars (usually "QSO's") are domiated by their bright core, with strong optical emission lines as well as strong continuous emission (optical/UV, also X-rays).

Seyfert galaxies and **Markarian** galaxies[3] are spiral galaxies with nuclei which are quasar-like in their broadband and line emission properties, but not so bright as the quasars. Seyferts are never radio-loud; they have weak radio cores and can have small, weak, poorly collimated radio jets which do not propagate out of the galaxy (due to the denser ISM? due to different conditions in the core?) Weaker versions of Seyferts, also in spiral galaxies, are called **LINERS** (Low Ionization Nuclear Emission Region) – with emission lines but less nonthermal continuum than the Seyferts.

14.2.2 The radio-loud ones

As we saw in the previous chapter, The term "radio galaxy" refers to an elliptical galaxy with a radio jet and double-lobed (or tailed) radio structure on supergalactic scales (linear extent, side-to-side, can range from $\sim$ 100 kpc to a few Mpc). The term "radio-loud quasar" (QSR) refers to a quasar with the same sort of radio jet-lobe structure. The distinction is mainly a question of how strong the nuclear activity is (as seen by us, anyway): how strong is the nonthermal continum and/or the emission lines?

14.2.3 Blazars and friends

These are a subset of the radio-loud ones (about 10%) with unusually bright, active, variable radio-to-optical cores. People in this field speak in acronyms:

OVVs (optically violent variables). These are the highly variable, clearly relativistic ones. They show a flat radio spectrum (suggestive of a compact, synchrotron self-absorbed core and a pc-scale radio jet with a few bright knows). They are highly variable, timescales from days to years.[4] These are the ones that show clear superluminal motion.

BLLs, as described, are named for BL Lacertae (an object previously thought to be variable star locally). The emission lines are quite faint compared to the cointinuum. As with OVVs, the BLL radio emission is strongly polarized, flat spectrum, & highly variable ($\delta t \gtrsim$ days).

Both of these are often grouped together as *blazars*. The general picture – motivated by the observations of superluminal motion – is that we are looking "down the pipe" of the jet, close to its axis. In addition to explaining the superluminal motion, such a geometry will enhance the variability (by relativistic effects and by the Doppler-beaming effects on a jet with some slight inhomogeneities.

14.2.4 Parent galaxies

What are the parent galaxies? We noted above that radio galaxies are found in elliptical galaxies, while Seyferts (and related radio-weak AGN) are found in spirals. What about quasars? There has been a sense that radio-loud quasars live in E's, while radio-quiet live in S's. This view seems to be changing, however, at least for the brightest of the population. Quoting from Dunlop etal (2003): "Virtually all powerful AGN live in normal, massive ellipticals: the parents agree with local bright E's in morphology, luminosity, scale-length, consistency with fundamental plane, axial-ratio distribution, and colors (evolved stellar population, age 10-13 Gyr)." "The inevitable conclusions is that these galaxies, or at least most of their stars, must have formed at high redshift." "Spheroidal hosts become more prevalent with increasing nuclear luminosity; for bright enough nuclei, the hosts of both radio-loud and

---

[3] both named for the person who originally cataloged them – based on unusually strong nuclear emission lines or UV continuum.

[4] There is an argument currently going on about variability on much shorter δt's, less than a day; these are the Intra-Day Variables. The variability is probably due to interstellar scintillation, although some people want it to be intrinsic.

radio-quiet AGN are massive ellipticals."

14.3 The usual model: a massive BH

We know the answer of course (or think we do): all of this comes from accretion onto a massive black hole. Let's see how (or if) this model can explain the variety of observations.

14.3.1 Zoom in: the central kpc and within

The first clues may have come from the emisison lines. In most objects we need photoionization; with detailed modelling one can pinpoint the photon *density*, which means the distance from the central engine. Applying this to the broad and narrow lines shows that the NLR must be at $\sim$ kpc from the engine, while the BLR is closer, more like pc to tens of pc away. Both broad and narrow line emission must be "patchy", for instance coming from dense clouds.[5] It may not be correct, however, to picture the clouds as uniformly distributed over the entire central kpc. We are learning, from HST imaging, that the emission-line gas in the central kpc is often confined to emission "cones" – what you would expect, for instance, if the ionizing radiation escaped from the central engine only in a moderately narrow cone.

14.3.2 Zoom in further: the central pc and within

This is, of course, the region of the central engine; the region we probe indirectly through variabilty, and directly through milliarcsecond radio imaging (as with the VLBA). The generic picture is, as we've discussed, accretion onto a massive black hole, probably with a jet being driven out the rotation axis. Figure 14.1 illustrates the general thinking. The accretion disk is mostly thermal, that is resembling the models you saw in P425: spatially thin and optically thick. It emits as a black body, and is thought to be the origin of the "big blue bump". Nonthermal emission – high energy, X- and γ-rays – comes from an optically thin region somewhere close to the black hole and the inner region of the disk; very high gas temperatures, possibly relativistic plasmas, and electron-positron physics may be going on here. A two-sided jet is driven out; if this is in an elliptical the jet propagates to extragalactic scales. The gas emitting the broad emission lines is sketched as discrete clumps, moving at high speed (here guessed to be random). The intercloud region is probably a hot wind, driven out from the accretion disk by a combination of its own temperature and magnetic/centrifugal effects.

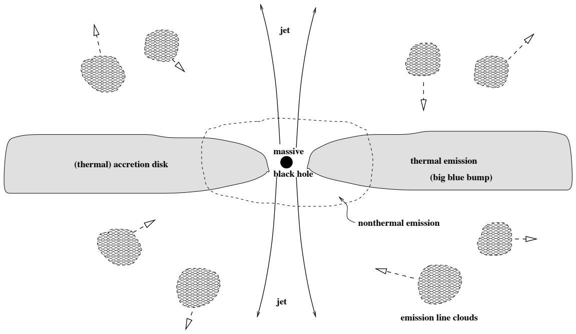

Figure 14.1 Generic cartoon of the central pc of the AGN. An accretion disk feeds a massive black hole. Details are in the text. The scale of this cartoon is a few pc (remember the event horizon of the black hole is only $\sim 1 - 10$ AU.

14.3.3 Why radio-loud vs. radio-quiet?

We still don't have much of an answer to this question, or even a standard "toy model". Radio galaxies, and maybe all of the brightest quasars, live in ellipticals; Seyfert nuclei, and maybe most of the radio-quiet quasars, live in spirals. Why? I wish I knew.

14.3.4 Why are only some galaxies "active"?

This one is even harder to answer. We now know that every galaxy contains a massive black hole at its core (check back to chapter 1 for discussion). But this black hole is "active" in only a few per cent of galaxies. Why? Once again, I wish I knew.

14.4 Unification Models

Can these various disparate categories and acronyms be explained by one simple, unified picture? Maybe, maybe not .. this issue has strong adherents, because simple pictures are pleasing and attractive. But one can push too much for simplicity, at the expense of ignoring some of the physics ... some of the community belong to "the unification church", while others (including your author) remain agnostic.

There seem to be two important motivations for this view, as follows.

---

[5] Why? One, we can see the engine, so the line-emitting gas doesn't completely cover it. Two, we know the density of the line emitting gas, from quenching and line-ratio arguments; comparing this to the total volume at the clouds' distance also requires the volume be incompletely filled.

14.4.1 Relativistic beaming

We've already met this several times. If the jet is driven out at relativistic speeds, with Lorentz factor γ – and we know it is, from the observation of superluminal motion – then a viewing angle within an angle $\sim 1/\gamma$ of the jet axis is special. From this vantage, *only*, we will see $v_{app} > c$, strong forward beaming, enhanced variability, and etc – that is we will see a blazar.

14.4.2 Obscuration and tori

We noted above that the optical line emission is anisotropic in some nearby AGN, being located in an ionization cone. Such a cone might arise from a fat torus (think of a donut) surrounding the central engine, allowing hard photons to escape only in a fairly broad cone along the axis of the torus. In addition, emission lines from Seyferts (which are nearby, and bright, enough to do this measurement) appear different in polarized than total light. Some Seyferts that show only a narrow line region (it's called a Seyfert 2, if you care) turn out to have broad line wings when seen in polarized light. Remember that Thompson scattering polarizes the light. Thus, if the light from the central engine, which sits at the heart of the donut, is scattered (by some diffuse ionized gas, above or below the system), an observer can detect the broad lines – which come from very close to the black hole – at angles at which the direct emission is obscured by the torus.

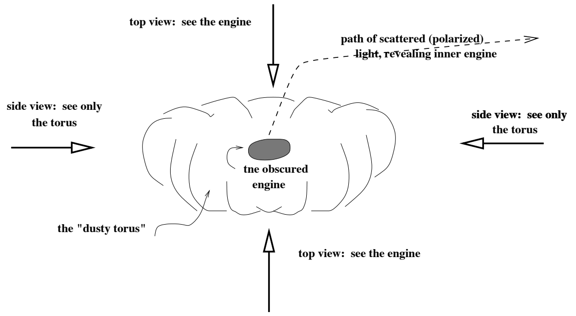

Figure 14.2 Extremely generic cartoon of the unified model. The central engine may be surrounded by a fat, opaque torus (maybe containing dust). Whether or not you see the central engine depends on your viewing angle. The scale of this cartoon is a few hundred pc.

14.5 AGN demographics

To summarize: we've walked through the general picture of an AGN: a massive black hole sits in the core of a galaxy. Galactic material (ISM) accretes onto the black hole. Side effects of that accretion are the generation of a strong radiation over all frequency bands (by a mix of thermal and nonthermal processes), and the ejection of collimated plasma jets (which may or may not propagate out of the galactic core).

The remaining questions, then, are demographic. How common are AGN, and how were they distributed over the history of the universe? Why does every bulge contain a black hole, and which formed first (the bulge or the BH)? Does every nuclear black hole make an AGN? How did the black hole come to exist in the galactic core (is it the chicken or the egg?) How have these AGN evolved over the course of the universe? To answer these questions, we need to go to high redshift - $z \sim 2 - 10$, say, to see what happened at early times.

These questions are far from being answered, but we do have some important clues. In case you're not familiar with high-redshift thinking, I've put some basic cosmography in an Appendix to this chapter.

14.5.1 Was there a "quasar era?"

Quasars – like other AGN – are rare in today's universe; most galaxies do not contain a currently active central engine (even if, as we now know, they probably contain a massive BH). This is not the case at early epochs, however. Two facts become clear from quasar surveys.

- At a given epoch – today, for instance – there are more faint quasars than bright ones. The number per volume at an absolute (optical, blue) magnitude M_B can be written very approximately as $\Phi(M_B) \sim \Phi_o M_B^{-0.7}$.
- The absolute number of quasars – either at a given magnitude, or integrated over all luminosities – was greater at early epochs. (This is the same thing as saying that the constant above, Φ_o, is a function of z.) The space density of quasars rises out to $z \sim 2$, has a broad peak in the range $z \sim 2-3$, then decays again at higher z. This epoch, which saw a lot of quasar activity, is called the "quasar era".

What has changed, with quasars, since the quasar era? What is this evolution of the number density saying? Up to a few years ago, there were (at least) two alternative possibilities. One is that the fraction of quasars, per galaxy population, has stayed the same with time, but the luminosity of a given quasar was much brighter at $z \sim 2$ than it is now. Because all quasar surveys must be flux limited, we would be detecting more of them early on. This is called *luminosity evolution*. An alternative theory is that the mean quasar luminosity has not evolved with time, but that there were simply more of them back then ... so that the increase in measured numbers of quasars back to $z \sim 2$ is what it appears to be. The real answer is probably somewhere inbetween these two extremes...but we must remember that almost every nearby galaxy (at least) harbors a black hole. If we define this as a dark quasar, then luminosity evolution must be the preferred model. However, we are learning more about galaxy formation and evolution, and the question may be more complex.

14.5.2 What about galaxy formation?

The quasar counts, described above, have been known for some time. A currently very active research area tries to understand when (at what z) galaxies formed. This is much harder than counting quasars (which are easy to pick out, they are bright and small and blue). How can we pick out galaxies in the act of formation? Think about a protogalaxy: picture a self-gravitating clump of stuff collapsing due to its own gravity, and/or being enhanced by the accumulation of smaller-sized clumps (mergers). The baryonic matter will lose energy (by radiation), fall through the dark matter, and must eventually form stars. We know ellipticals have little active star formation today; their stars must have formed in one great rush, at early times. Spirals today do have ongoing star formation, but they probably also had a strong starburst phase when they first formed.

So, the general idea is to look for unique signs pointing to early bursts of star formation. We know a lot about local star formation regions. Optically, they are bright in Hα, and also in the UV (due to all those hot young stars). Moreover, we know they are dust-enshrouded, and thus very bright in the IR or sub-mm band. Thus: we can try to find distant objects that are bright in the UV, or in Hα, or – the currently most promising – in the IR/sub-mm.

The tentative result of these surveys (the work is still ongoing, and the arguments as large as the error bars at high z) is very interesting. Given the variety of observational probes, which mix apples and oranges, people generally turn their data into "the rate of star formation as a function of z".[6] This is called a "Madau plot", after the person who first did one.

Difficulties in the analysis aside, everyone in this field agrees on the low-z result. The number of strongly-starbursting galaxies increases with redshift, out to $z \sim 2$ – which is just the quasar era. The uncertainties come at higher redshift: it is not clear if the star formation rate declines again, at higher z (like the quasars), or continues more or less steady (out to, say, $z \sim 5$, the current limit of this sort of work).

14.6 Ending with questions

All of this raises some intriguing questions, which seem as good a way as any to end these notes.

- Did the bulk of galaxy formation take place at the same time as quasars were most active?
- Is a strong starburst a part of the mass-accretion process which kept the AGN powerful at that epoch?
- Are the merger and dissipation events that made the bulge or E galaxy the same events that made the quasar shine?
- Are the majority of nuclear BH dark today because they aren't being fed? Are galaxy mergers necessary to transport the gas (i.e. BH food) close to the nucleus?

And ... that's it, folks! It's been fun – have a good summer!

Key points

- The phenomenology: important observational trends..
- The usual model: what's there on sub-pc scales?
- "Practical cosmology": what does "high z" mean (ages, timescales, etc)?
- The QSO epoch and how it might connect to galaxy formation.

---

[6] This can be a long daisy chain. For instance: measure an IR luminosity, say; from that determine how many bright, massive stars are needed to power the dust; divide by the main sequence lifetime of the massive stars; and you have "derived" a "star formation rate". Or ... measure the radio luminosity; use the fact that the radio and IR powers are well correlated for nearby spiral galaxies, to estimate the IR luminosity; and return to top..

14.7 Appendix: a little practical cosmology

Before we can talk about BH and AGN in the early universe we need to review the language used. The context is cosmology: we work with the *scale factor*, $R(t)$. This is a quantity which describes ("scales") the distance between two fixed points in an expanding universe – say two nearby galaxies. The evolution of $R(t)$ reflects the fight between gravity (an attractive force), initial conditions (how much expansion did the universe start with?), and the effect of any vacuum energy density (the cosmological constant; "dark energy"). For our purposes here, let's assume that $R(t)$ is a known function (found from the solution of Einstein's field equations). Everything we can measure about a distant object – size, luminosity – depend on how much the universe has expanded between the time it emitted a light signal (t_{em}) and the time we receive that signal (today; usually called t_o).

14.7.1 Just what is the redshift?

The redshift is defined is $1 + z = \delta\lambda/\lambda$ (the shift in an emission wavelength, say of a spectral line, relative to its rest value). The redshift of a cosmological object is an important quantity. It tells us not just the recession speed (which isn't very interesting), but – more importantly – the distance and age of the object.

There are at least 3 physical causes of a redshift.

- You remember the simple Dopper shift: for low speeds, $1 + z = v/c$; and for relativistic speeds, the Lorentz transform in simple geometry gives $1 + z = \gamma(1 + \beta)$.
- There is also a gravitational redshift, which occurs when light emitted in a potential well (say at the surface of a star) climbs out to the observer. For a Schwarzchild geometry this becomes $1 + z = 1/(1 - r_s/r)^{1/2}$.
- Finally, there is a cosmological redshift: the frequency of a signal decreases due to the expansion of the universe while that signal propagates to the observer. This is the one we want here: it becomes $1 + z = R(t_o)/R(t_{em})$.

Because the cosmological redshift is intimately tied to the expansion of the universe, it becomes a handy way to describe the distance and age of an object. Two things are worth noting here.

14.7.2 The Hubble diagram

Think about a source at some distance D. In Euclidean space its light spreads out over a surface of area $4\pi D^2$ before it gets to the observer, so that one sees a flux $\propto L/D^2$. But, this changes if space is not Euclidean. The key fact is that the area of the D-sphere is not $4\pi D^2$. The details of how the area changes depend on the cosmological parameters (curvature and cosmological constant), but the general trend can be sketched, as in Figure 15.1. NOTE that for $z \ll 1$, curvature effects don't matter; $d_l \propto z$, just as in Euclidean space. The original Hubble's law ($v = H_o z$) applies in this limit. Also note that the so-called Hubble constant, H_o, is the slope of this curve *as* $z \to 0$ (that is, it isn't a constant, but a variable).

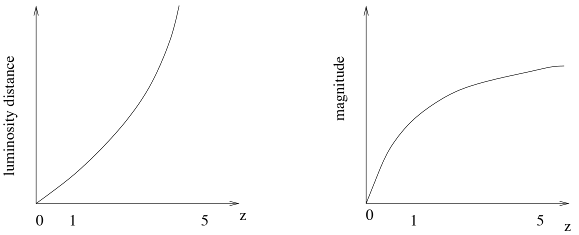

Figure 14.3 How the curvature of space affects the luminosity distance, d_l (defined so that the observed luminosity of a source at redshift z is $L/4\pi d_l(z)^2$). Note that for small redshifts, $d_l \propto z$, and the classical Hubble's law is recovered. The right hand diagram is often called the Hubble diagram; its high-z extension can, in principle, be used to measure cosmological parameters.

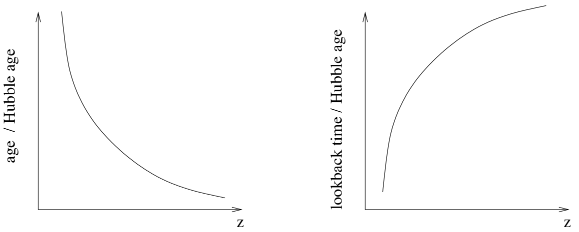

Figure 14.4 How the age of the universe, and of an object we observe, depend on redshift. Left: the age of the universe (counted from $t = 0$, the big bang) when an object had redshift z. Right, the lookback time – the difference between the age of the universe at z and the age of the universe now. This tells us how long ago an object existed, if we see it at z today. The "Hubble age" is defined as $1/H_o$.

14.7.3 The lookback time

How long ago did an object exist, if we see it now at redshift z? If space were Euclidean, that would be simple: (age at emission) = (today's age) - (distance

/ lightspeed); and the lookback time would be proportional to the redshift. But as with the Hubble diagram, space curvature makes this more interesting for $z \gtrsim 1$. Numbers: $z = 2$ gives a lookback time $\sim 0.5/H_o$; $z = 5$ gives a time $\sim 0.6/H_o$, and for higher z's the lookback time doesn't change by a lot.