Galaxy formation and evolution PAP 318, 5 op, autumn 2020

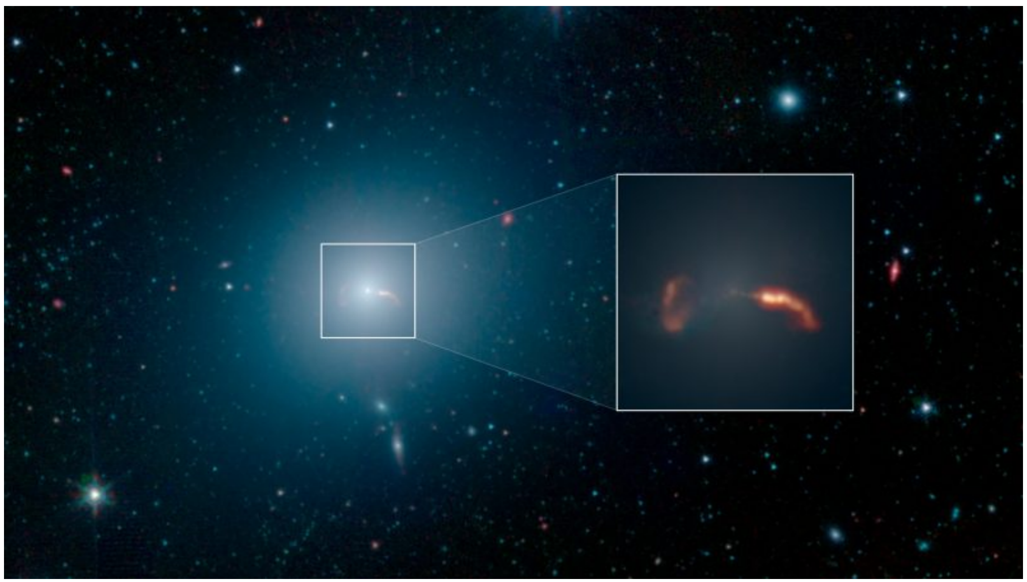
on Zoom

Lecture 13: Formation of active galaxies, 04/12/2020

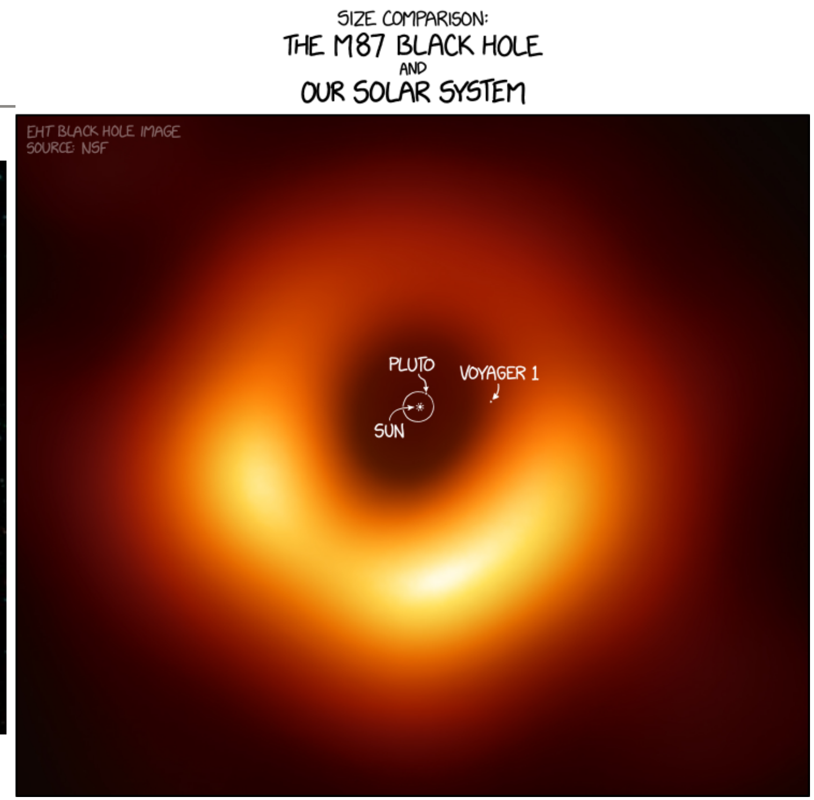
On this lecture we will discuss

1. The population of AGNs and their classification.
2. Physics of the AGNs: The central engine.
3. Continuum and line emission from AGNs.
4. The various components of the standard AGN model and the unification model.
5. The formation and growth of AGN at high and low redshifts.
6. Feedback and the impact of AGNs on the evolution of galaxies.
7. The lecture notes correspond to: MBW: pages 618-651 (§14.1-14.4)

13.1 Active galactic nuclei (AGNs)



M87, the central galaxy in the Virgo cluster.



- AGNs are galaxies with supermassive black holes in the centres, as was unequivocally confirmed by the direct observation of the shadow of the supermassive black hole in M87.

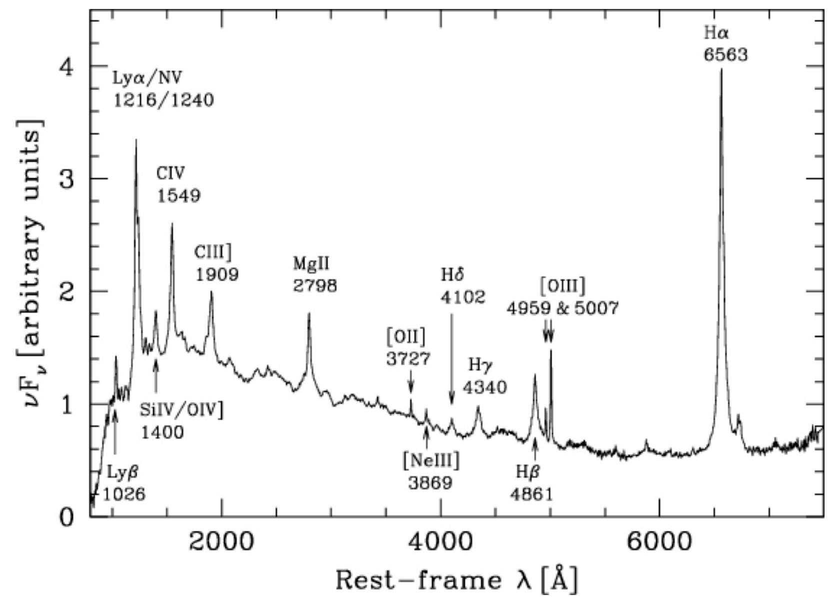
The population of Active galactic nuclei (AGNs)

- The temperatures of stars cover a relatively narrow range of 3,000 K≤T≤ 40,000 K. The spectra of normal galaxies is roughly the sum of the Planck spectra corresponding to these temperatures and thus the spectra are largely confined in the wavelength range between λ~400-2000 nm.
- A small (but important) fraction of galaxies have a spectral energy distribution that is much broader than expected from a collection of stars, gas and dust. These galaxies typically emit non-thermal radiation over the full wavelength regime from radio to X-rays.
- The non-thermal emission emanates from a very small central region, often less than a few parsecs across, which is called the AGN.
- Local z=0 number densities in units of Mpc^{-3}: Field galaxies: 10^{-1}, Luminous spirals: 10^{-2}, Seyfert galaxies: 10^{-4}, Radio galaxies: 10^{-6}, Quasars: 10^{-7}, Radio-loud Quasars: 10^{-9}.

Classification of AGNs

- Roughly speaking, an object is defined to be an AGN if one or more of the following properties are observed:

1. A compact nuclear region much brighter than a region of the same size in a normal galaxy.
2. Non-stellar, i.e. non-thermal continuum emission.
3. Strong emission lines.
4. Variability in the continuum emission and/or in emission lines on relatively short time-scales (days/weeks).



A composite spectrum of QSOs revealing the typical non-thermal continuum and permitted and forbidden [in brackets] spectral lines.

Seyfert galaxies and Radio galaxies

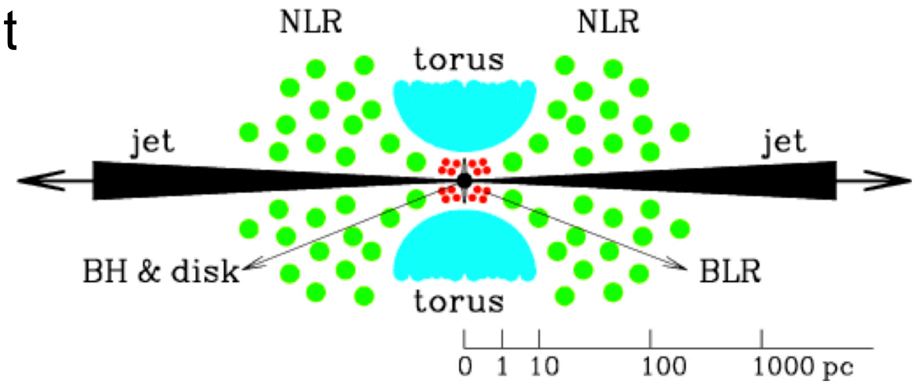
- **Seyfert galaxies:** Named after Carl Seyfert (1943) are active galaxies characterized by a spiral-like morphology with bright star-like nuclei. Seyfert galaxies are subdivided into Seyfert I and Seyfert II galaxies. Seyfert I galaxies have broad permitted lines and narrow forbidden lines, whereas in Seyfert II galaxies both the permitted and forbidden lines are narrow. LINERs (low-ionization nuclear emission line regions) are probably a lower luminosity extension of Seyferts, although nuclear stellar clusters may also give rise to LINER-like activity.
- **Radio galaxies:** Typically elliptical galaxies, with strong radio emission ($P_{1.4GHz} \geq 2 \times 10^{23}$ W Hz^{-1}). Similarly to Seyferts can be divided into broad-line radio galaxies and narrow-line radio galaxies. Often jet-like structures that extend into lobes from the compact central object. A further division can be made into Fanaroff & Riley classes I and II based on their radio morphology.

Quasars and BL Lac objects

- **Quasars:** The name quasi-stellar radio source was originally used for the optical identifications of radio sources with Seyfert-like spectra. In the optical quasars are unusually blue (emission from the hot accretion disk) and optical surveys revealed that the radio-quiet quasars outnumber the radio-loud quasars by a factor of ~10-100. QSOs are the most luminous AGNs and their host galaxies have diverse properties, some are regular, but others are clearly interacting.
- **BL Lacs**: These objects together with the OVVs (Optically violently variable) have very strong and rapid optical variability. OVVs show some emission lines, but BL Lacs show no emission lines. Both have relatively strong polarization on a few % level and radio emission. Most probably we are observing these objects close to the direction of the jet.

13.2 Physics of the AGNs: The central engine I

- The small size of the emission region and the large amount of energy output suggest that the central engine is a supermassive black hole (SMBH).
- In addition to the SMBH, the standard model also assumes the existence of a broad-line region (BLR), a narrow-line region (NLR), an accretion disk, a torus and jets.
- The energy can be estimated by equating the radiation pressure with the gravity on the gas:



$$P_{\rm rad} = \frac{L}{4\pi r^2 c} \Rightarrow \mathbf{F} = \sigma_T P_{\rm rad}(r) n_e(r)\mathbf{r}$$

$$\mathbf{F}_{\rm rad} \leq F_{\rm grav} = \frac{G M_{\rm BH}\rho(r)}{r^2}$$

$$L_{\rm edd} = \frac{4\pi G c m_p}{\sigma_T} M_{\rm BH} \approx 1.28 \times 10^{46} \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \ {\rm erg s}^{-1}$$

The central engine II

- What kind of mass accretion is required in order to power an AGN? The assumption that the luminosity is powered by the gravitational potential of the BH gives:

$$L = \frac{GM_{\rm BH}}{r}\dot{M}_{\rm BH}$$

- Then the efficiency at which the rest mass of accreted material is converted into radiation is:

$$\epsilon_r = \frac{L}{\dot{M}_{\rm BH}c^2} = \frac{1}{2}\frac{r_S}{r}, \qquad r_s = \frac{2GM_{\rm BH}}{c^2} = 10^{-2}\left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \text{light} - \text{days}$$

- The bulk of the emission in the blue bump originates from $r \sim 5r_S$, from which it follows that $\varepsilon_r \sim 0.1$. To power $L \sim 10^{46}$ ergs^{-1} requires an accretion rate of 2.2 solar masses per year. Note for hydrogen fusion $\varepsilon \sim 0.007$.

Accretion disks

- Since the gas to be accreted by a supermassive BH in general has angular momentum, the accretion will most likely take place through a Keplerian accretion disk.
- For a constant accretion rate, the conservation of angular momentum:

$$\frac{\partial \Sigma}{\partial t} = -\frac{1}{R}\left\{\frac{(\partial/\partial R)[\nu_k \Sigma R^3 (d\omega/dR)]}{(d/dR)(R^2\omega)}\right\}$$

- Here ν_K is the kinetic viscosity and ω is the rotation velocity of the gas. The viscosity is thought to be due to turbulence and magnetic stresses, but the details are still poorly understood.
- The temperature structure of the disk can be derived (see MBW for details). Most emission from the inner disk, where T is the highest:

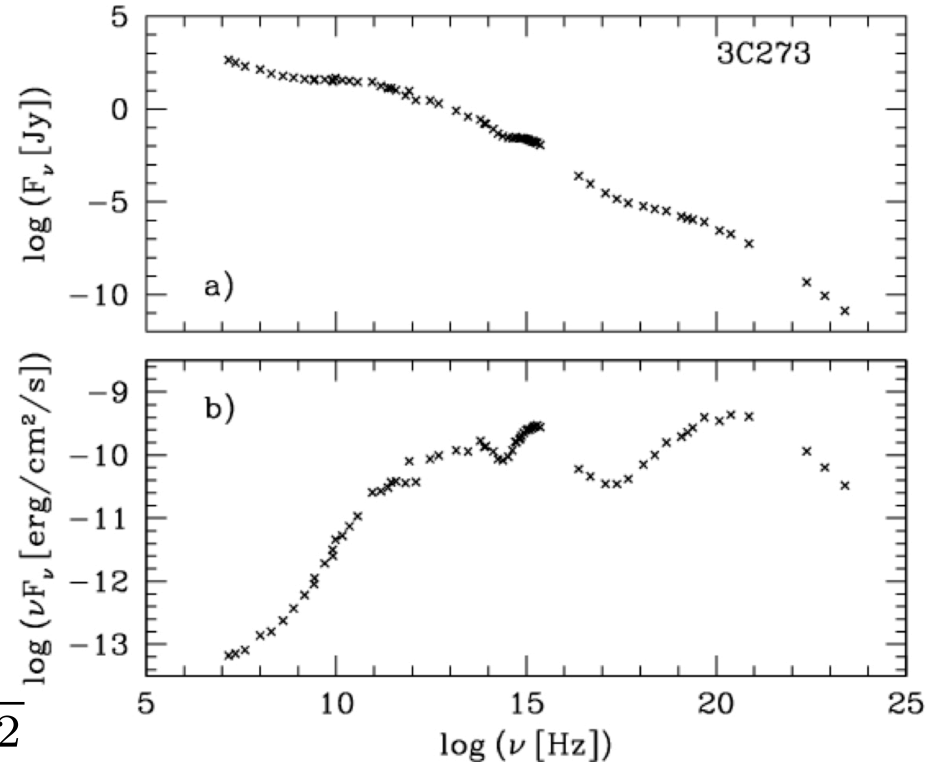
$$T(R) \sim 6.3 \times 10^5 \left(\frac{\dot{M}_{\rm BH}}{\dot{M}_{\rm edd}}\right)^{1/4} \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right)^{-1/4} \left(\frac{R}{r_S}\right)^{-3/4} \ \mathrm{K}$$

13.3 Continuum emission

- AGNs have typically a very broad spectrum, which can roughly be described as $F_\nu \propto \nu^{-\alpha}$, $0 \leq \alpha \leq 1$.
- The SED contains the blue bump around $\nu \sim 10^{15}$-10^{16} Hz and a broad bump around $\nu \sim 10^{20}$-10^{21} Hz.
- A variety of emission mechanisms is involved, with relativistic electrons playing a crucial role, which can be seen from the brightness temperature:

$$T_b = \frac{c^2 J_\nu}{2 k_B \nu^2}$$

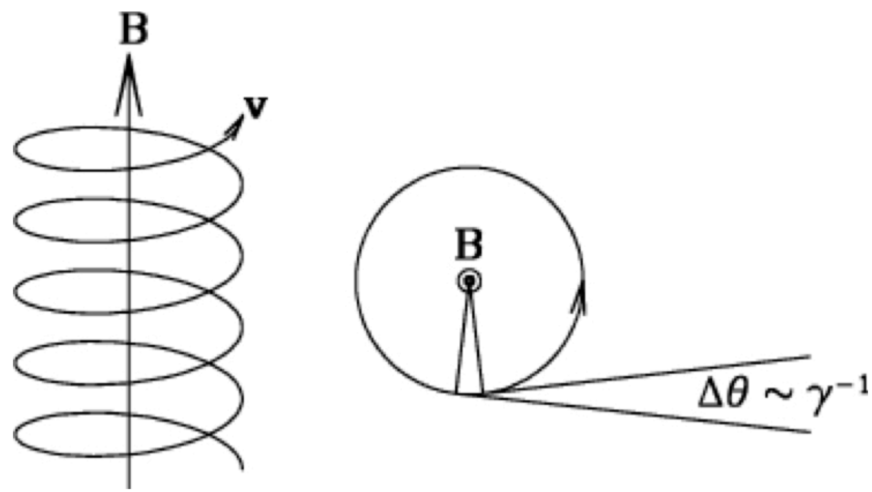
- Often $T_b \sim 10^{11}$ K -> non-thermal radiation -> shocked electrons.



Synchrotron emission

- When a charged particle is accelerated, it radiates photons.
- Synchrotron emission is generated by relativistic electrons spiralling in a magnetic field. The radiation is beamed into a forward cone with opening angle $\Delta\theta \sim \gamma^{-1}$, where $\gamma=[1-(v/c)^2]^{-1/2}$ is the Lorentz factor.
- For an isotropic distribution of relativistic electrons the average power is:

$$\langle P_S \rangle = \frac{4}{3}\sigma_T c \gamma^2 \beta^2 u_B$$



- The expected shape and peak of the spectral energy distribution can be predicted and it is in good agreement with the observations.
- At low frequencies synchrotron self-absorption by the emitting electrons must be corrected for.

Inverse Compton scattering

- In this mechanism high-energy electrons scatter inelastically with lower energy photons increasing substantially their energies. The net effect is emission of radiative energy from the electrons typically at X-ray and gamma-ray energies.
$$\langle P_{IC} \rangle = \frac{4}{3}\sigma_T c \gamma^2 \beta^2 u_{\mathrm{rad}}$$
- This formula is similar to the synchrotron formula, except for the u_{rad}-term which gives the energy density in the radiation field, as opposed to the magnetic field energy density u_B.
- The relative importance between the Compton and synchrotron cooling can be estimated from the formula below. Above $T_B \sim 10^{12}$ K inverse Compton cooling dominates:
$$\frac{\langle P_{IC} \rangle}{\langle P_S \rangle} \sim \left(\frac{T_b}{10^{12}\ \mathrm{K}}\right)^5 \left(\frac{\nu_e}{1\ \mathrm{GHz}}\right)$$

Emission lines

- The emission lines can be divided into permitted and forbidden lines. The balance is given by (A_{21} Einstein coefficient and P_{12} transition probability due to collisions):

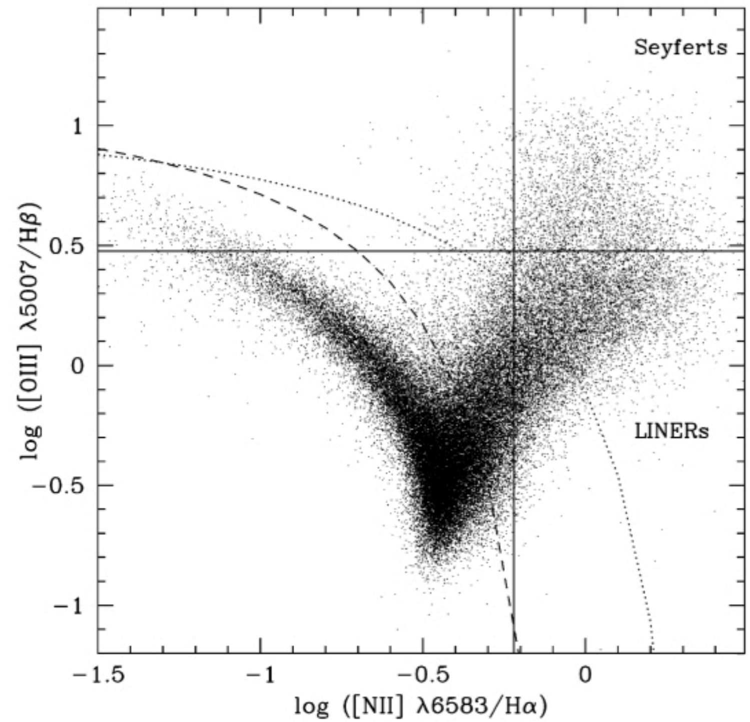
$$n_e n_1 P_{12} = n_e n_2 P_{21} + n_2 A_{21}$$

- The relative population of the two states:

$$\frac{n_2}{n_1} = \frac{n_e P_{12}}{A_{21}} \frac{1}{1 + n_e P_{21}/A_{21}}$$

- The line luminosity is then (produced by spontaneous emission):

$$\mathcal{L}_c = n_2 A_{21} h_P \nu_{12} = \frac{n_e n_1 P_{12} h_P \nu_{12}}{1 + n_e P_{21}/A_{21}}$$



Emission-line diagnostics can be used to probe the temperature and hardness of the radiation field. Separate AGNs from starbursts.

Jets and superluminal motion

- **Jets:** In broad terms, jets are well-collimated outflows of material and are observed in many radio galaxies. In some cases a thin jet with width of ~1 pc is observed to extend continuously from the innermost parsec-scale region of a galaxy to a distance up to several hundreds of kiloparsecs.
- Jets can only be produced in supersonic flows as subsonic flows would be unable to maintain the necessary collimation. The physics of jets is not fully understood. In the standard model by Blandford&Rees what is needed is an internally relativistic fluid and a relatively small pressure gradient in the medium that confines the jet and the initial flow in some direction.
- **Superluminal motions:** When matter is moving close to the line-of-sight at relativistic velocities apparent superluminal motions can be detected. Another relativistic effect is relativistic beaming, which brightens the approaching jet and dims the receding jet causing apparent one-sided jets.

Emission-line regions and torus

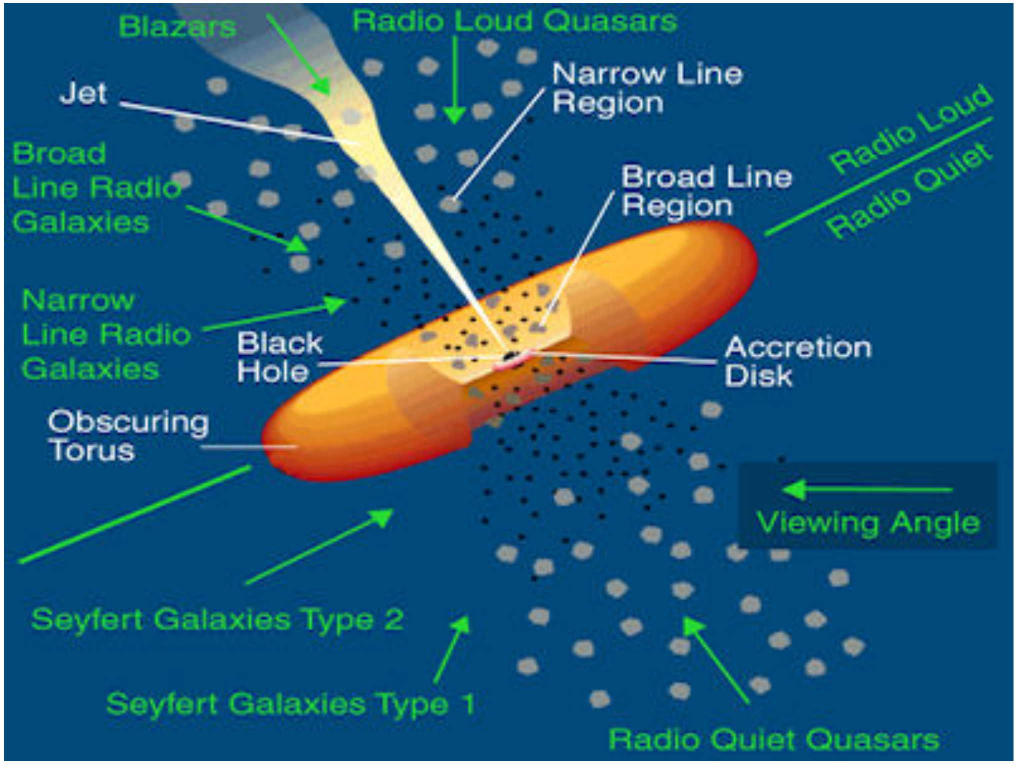
- **Broad-line region:** The broad lines with velocity widths of the order of ~1000 km/s suggests that these lines are produced in a small inner region (≤0.5 pc) surrounding the accretion disk. The lack of forbidden lines indicates a high electron density ($n_e \sim 10^{10}$ cm^{-3}).

$$R \sim \frac{GM}{\sigma_v^2} \lesssim 0.5 \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \left(\frac{\sigma_v}{1000\ {\rm kms}^{-1}}\right)^{-2}\ {\rm pc}$$

- **Narrow-line region:** Many AGNs reveal strong narrow lines with velocity widths of ~100 km/s. These lines are produced in a region of size ~50 pc around the central engine. To observe narrow lines the electron density must be $n_e \leq 10^6$ cm^{-3}.
- **Obscuring Torus:** The innermost region is covered in an obscuring torus, which blocks our line-of-sight in type II AGNs. The torus must have a large column density of gas in order to block the BLR and strong X-ray and UV radiation. The torus is most probably clumpy.

AGN unification

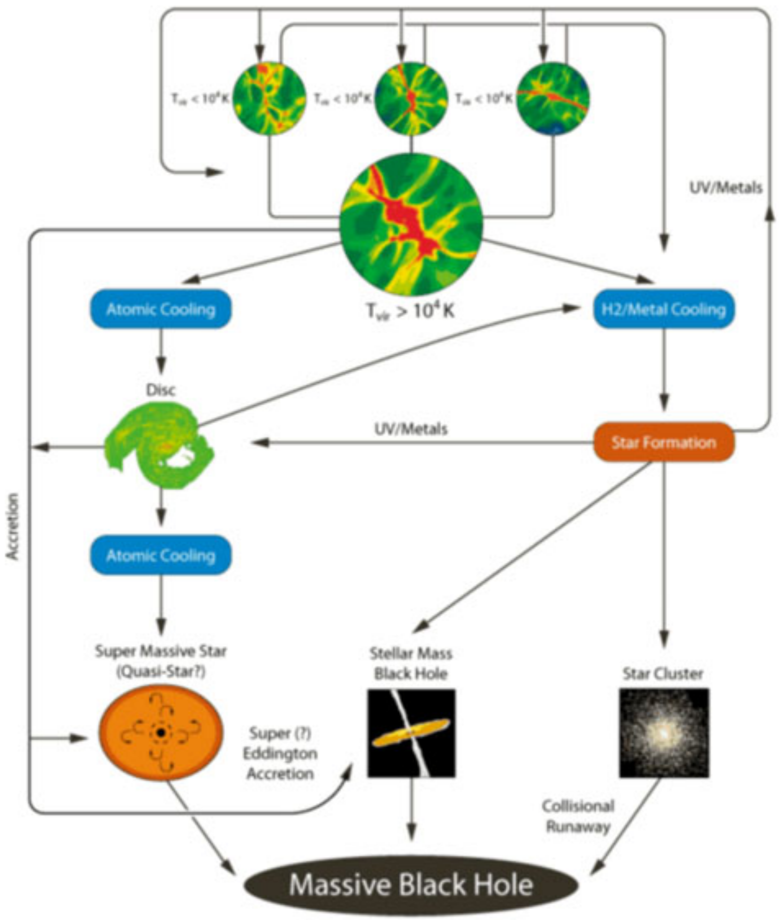
- In the AGN unification model all different types of AGNs are basically similar types of objects just viewed from different lines of sight.
- If we view the axisymmetric AGN from the side towards the obscuring torus we will observe a type II AGN, whereas if we can directly view the inner BLR we classify the AGN as type I.
- Radio properties also change depending on the viewing angle.



13.4 The formation and growth of SMBHs

- The formation mechanism of SMBHs at high redshift is uncertain. There are two leading models:

1. **Pop III remnants:** The first stars in the Universe were presumably very massive ($\geq$50-100 $M_\odot$). Black holes formed from the death of these stars should have masses of the order ~10 $M_\odot$.
2. **Direct collapse:** There were no metals in the early Universe and it is possible that the inefficient cooling resulted in near isothermal collapse of 10^4-10^5 $M_\odot$ gas clouds resulting in more massive SMBHs.



Formation of AGN I

- One important fact that any theory of AGN formation must take into account is the observation that quasars with M_{BH}~10^9 $M_{\odot}$ exist at z~7. The cosmic time at such a redshift is around t~0.7 Gyr.
- Since black holes form in collapsed objects, the free-fall time scale of a virialized halo t_{ff}~r_{vir}/V_C~$1/[10H(z)]$ must also be considered. At z~7 t_{ff} is about 50 million years.
- If the growth of the SMBH is through radiative accretion the mass accretion can be written as:

$$\dot{M}_{\rm BH} = \frac{L}{\epsilon_r c^2} = \left(\frac{L}{L_{\rm edd}}\right)\frac{M_{\rm BH}}{\epsilon_r t_{\rm edd}} \qquad t_{\rm Edd} = \frac{\sigma_T c}{4\pi G m_p} \approx 4.4\times 10^8 \text{ yr}$$

- If L/L_{edd} and ε_r are independent of time: $M_{\rm BH}(t) = M_{\rm BH,0}e^{t/t_{\rm BH}}$

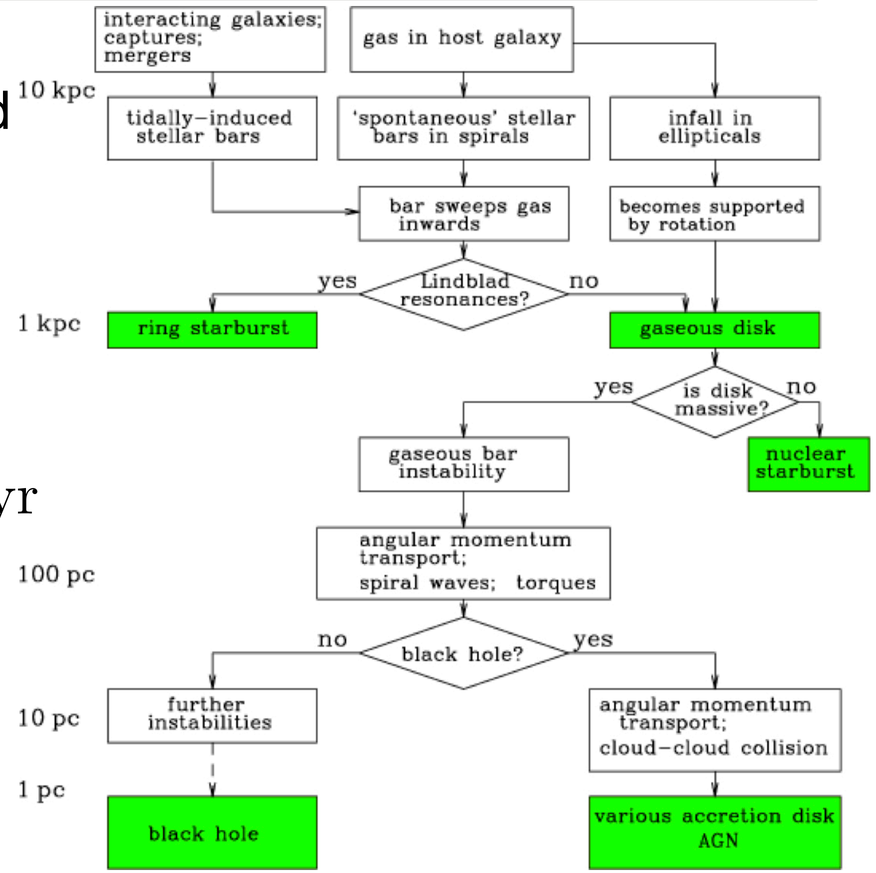
$$t_{\rm BH} = (L/L_{\rm edd})^{-1}\epsilon_r t_{\rm edd} \approx 4.4\times 10^7(\epsilon_r/0.1)(L/L_{\rm edd})^{-1} \text{ yr}$$

Formation of AGN II

- The time required for a black hole to reach M_{BH}~10^8 $M_\odot$ depends on the seed mass.
- If the seed mass $M_{BH,0}$=100 $M_\odot$ we require 14 e-foldings to grow to 10^8 $M_\odot$.

$$t \approx 14 t_{\mathrm{BH}} \approx 6 \times 10^8 \left(\frac{\epsilon_r}{0.1}\right) \left(\frac{L}{L_{\mathrm{edd}}}\right)^{-1} \mathrm{yr}$$

- In order to achieve a sufficiently high growth rate, either the gas accretion should be at a super-Eddington rate or the radiative efficiency should be very low.

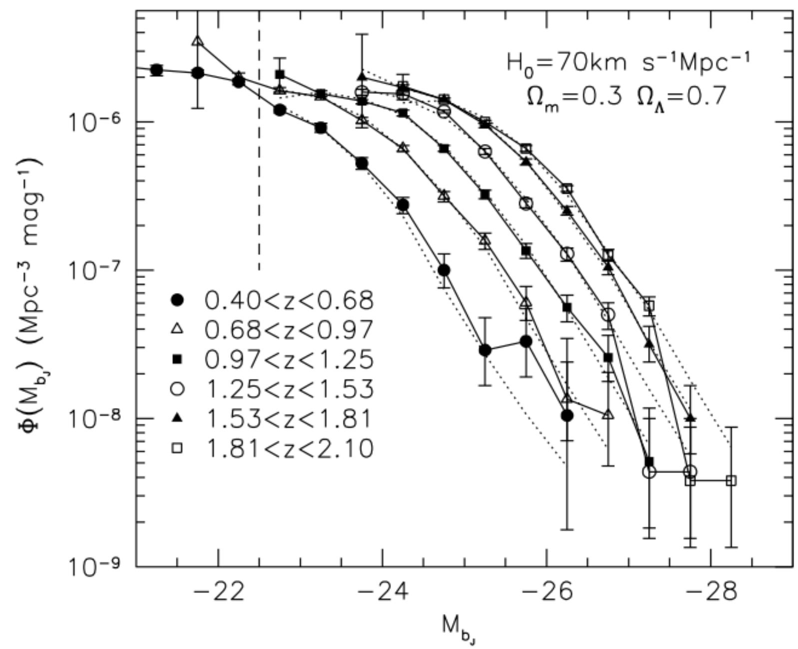


Formation of AGN III

- A key step towards understanding the formation of AGN is to identify the mechanisms that can effectively bring gas to the centre of the host galaxy and feed the SMBH.
- The amount of gas involved in SMBH formation is only a small fraction of the total mass of the host galaxy. The problem is that the available gas in the galaxy needs to lose its angular momentum. The specific angular momentum is $j \sim (GMR)^{1/2}$. A parcel of gas with $M \sim 10^{11}\ M_\odot$ and $R \sim 10$ kpc needs to reduce its angular momentum by a factor of 10^4 to move the gas to ~ 0.1 pc of a $10^8\ M_\odot$ BH.
- The prime mechanisms for achieving this is mergers between galaxies, which generate strong tidal torques. Another possible mechanism is internal secular processes in massive disks that cause bar instabilities and radial gas inflows.

AGN demographics

- The formation scenario discussed above implies strong evolution in the AGN population.
- The observed luminosity function of quasars can be fitted by the luminosity function below: β_1=3.9 and β_2=1.5.
- The data shows clearly that the quasar population peaks at z~2-3 and declines then strongly towards both lower and higher redshifts.



$$\phi(L,z)dL = \phi^*(z)\left\{\left[\frac{L}{L^*(z)}\right]^{\beta_1} + \left[\frac{L}{L^*(z)}\right]^{\beta_2}\right\}^{-1}\frac{dL}{L^*(z)}$$

13.5 AGNs and galaxy formation

- AGNs can release a huge amount of energy during their lifetimes,

$$\frac{dE}{dt} = \epsilon \dot{M}_{\mathrm{BH}} c^2$$

where $\varepsilon=\varepsilon_r+\varepsilon_m$ is an efficiency factor the combines the radiative and mechanical feedback efficiencies.

- We can compare this energy output to the binding energy of the galaxy, which according to the virial theorem is $W \sim -M_{gal}\sigma^2$:

$$\frac{E}{|W|} \sim \frac{\bar{\epsilon} M_{\mathrm{BH}}}{M_{\mathrm{gal}}} \left(\frac{c}{\sigma}\right)^2$$

- $M_{BH}/M_{gal} \sim 10^{-3}$ and for $\sigma \sim 300$ kms^{-1}, the ratio is $E/|W|=10^3\varepsilon$ indicating that the AGN energy can easily surpass the total binding energy of the galaxy.
- There are two outstanding questions: 1) What is the value of the feedback efficiency, ε and 2) how effective is the coupling of the feedback energy to the surrounding gas.

Radiative feedback

- In radiative feedback models the radiative efficiency is about ε_r~0.1 in bright AGNs. This energy can in principle feed back into the environment through both radiation pressure and radiative heating.
- UV photons from the AGN can ionize the surrounding medium in galaxies and in the intergalactic medium. This heating can significantly suppress gas cooling and star formation.
- If the host galaxy of an AGN contains a significant amount of dust the radiation from the central AGN can be effectively channelled into the surrounding gas by dust absorption.
- Radiation pressure can accelerate small amounts of matter to the very high observed velocities (v~0.2c).
- The feedback efficiencies are highly uncertain, but the common assumption is that about ~5% of the total radiative energy is coupled to the gas, resulting in an overall efficiency of ~0.5%.

Mechanical feedback

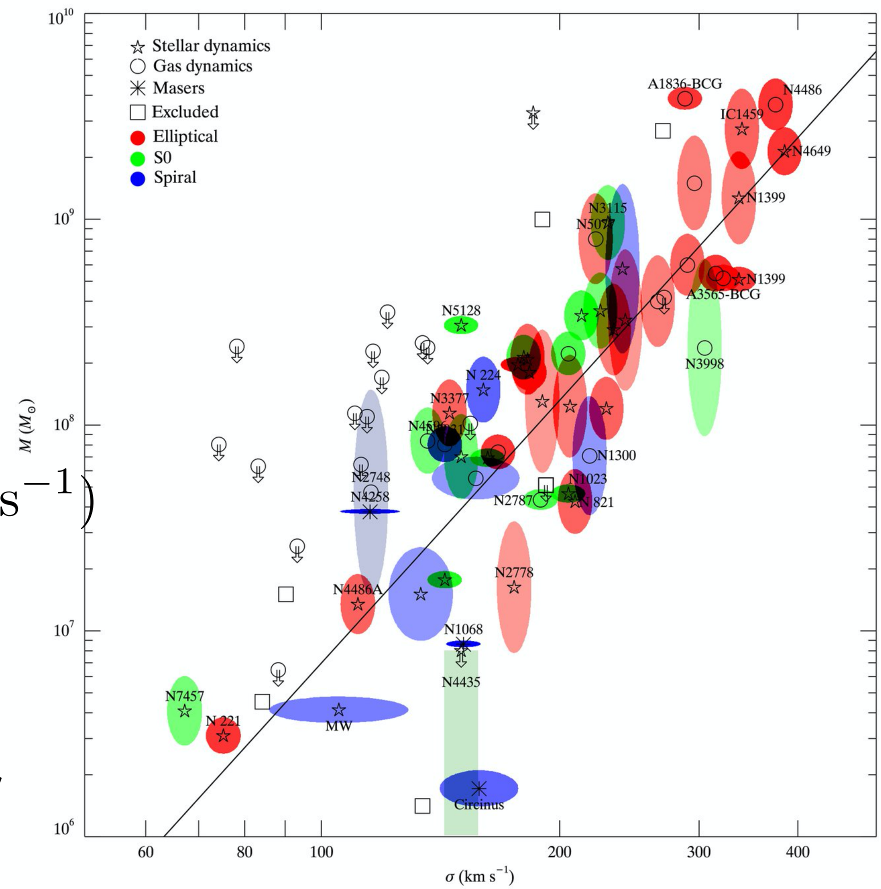
- During the low-accretion mode of an AGN, when the accretion rate of the SMBH is much lower than the Eddington rate, AGN feedback is believed to proceed mainly through radiatively inefficient, mechanical forms, such as radio jets and lobes.
- Evidence for such energy feedback can be seen in a number of elliptical galaxies at the centres of galaxy clusters, which contain X-ray cavities filled with relativistic gas.
- During the expansion of jet-powered bubbles we can assume that there is no radiative cooling in the bubbles, because the gas density is low and the temperature is high. The expansion of the bubble will slow down as material is swept up, but during the supersonic phase energy will be deposited in the medium through shocks. AGN feedback has been important in all massive galaxies as they contain relic SMBHs.

Co-evolution of SMBH and galaxies

- An important observed result of the co-evolution of SMBH and their host galaxies is the relatively tight M_{BH}-σ, which relates the mass of the supermassive black hole and the velocity dispersion of the stellar bulge ($\propto$ stellar mass) .

$$\log(M_{\rm BH}/M_\odot) = 8.12 + 4.24\log(\sigma/200\ {\rm kms}^{-1})$$

- This means that the growth of the SMBHs and the bulge component of galaxies is linked, suggesting that they both grew together in mergers or in secular disk instabilities.



What have we learned?

1. AGNs are characterized by broad spectral energy distributions consistent with the fact that the radiation is of a non-thermal origin, such as synchrotron and inverse Compton radiation.
2. The AGN population can be explained by the unification model in which the differences between different AGN populations is mainly related to our observed line-of-sight. In type I AGNs we have a more direct view and observe the broad-line regions, whereas in type II AGNs the torus obscures the radiation and we can only see the narrow-line regions.
3. SMBHs can form either from Pop III remnants (~10-100 $M_\odot$) or by direct collapse of gas clouds (~10^4-10^5 $M_\odot$). The growth of the SMBHs is limited by the Eddington limit and growing to the observed masses of SMBHs by z~7 is difficult.
4. AGNs affect the growth of massive galaxies through radiative and mechanical feedback and the SMBHs presumably co-evolve with their host galaxies as manifested in the $M_{BH}-\sigma$ relation.