Galaxies, Active Galactic Nuclei and Black Holes: Evolution and Interaction

A. Marconi

Istituto Nazionale di Astrofisica – Osservatorio Astrofisico di Arcetri, Largo Fermi 5, I-50125 Firenze, Italy e-mail: marconi@arcetri.astro.it

Abstract. It has long been suspected that galaxy formation is intimately linked to Active Galactic Nuclei (AGN), but this idea was only based on circumstantial evidences until a few years ago. The recent discovery of supermassive black holes (BH) in galaxy nuclei and, in particular, the relation between BH mass and host galaxy structural parameters has shown how the process of galaxy formation and BH growth are tightly linked. The fact that BHs are mostly grown during AGN activity closes the ring linking AGNs to galaxy formation. The following physical framework is therefore emerging: galaxy formation in the hierarchical scenario triggers AGN activity causing the growth of the central BH. The AGN feedback on the host galaxy increases with increasing BH mass and eventually expels the gas from the host galaxy thus stopping star formation and BH growth. The final consequence are the observed relations between BH mass and host galaxy.

Key words. galaxies: active – galaxies: evolution – galaxies: formation – galaxies: general – galaxies: nuclei – (galaxies:) quasars: general

1. Introduction

One of the key questions of current astrophysical research is how the universe changed from a nearly homogeneous state at \( z \approx 1000 \) to the wealth of structures observed at \( z \approx 0 \).

The \( \Lambda \)CDM cosmological model is a successful framework in which to address the issue of structure formation. In this model structures grow from weak density fluctuations in the homogeneous and rapidly expanding universe, are amplified by gravity and then become the structures we see today. The current "concordance" model is characterized by \( \Omega = 1 \), i.e. a flat universe. The total mass density is dominated by dark matter and only a small fraction is accounted for by baryons \((\Omega_{\text{matter}} = 0.3 \text{ and } \Omega_{\text{baryons}} = 0.04)\). The dominant mass component, the cold dark matter, is assumed to be made of elementary particles which interact only gravitationally. What is needed to have \( \Omega = 1 \) is some unknown form of dark energy whose existence is required by a non-null cosmological constant \((\Omega_{\Lambda} = 0.7)\). See, e.g., Spergel et al. (2003) and references therein.

Within the \( \Lambda \)CDM cosmological model it is possible to follow the evolution of dark matter halos and, in principle, this is a very trivial task since it can be done with "simple" N-body simulations where the only physics required is that of gravity. In practice, it is an extremely challenging task computationally because, in order to follow the evolution of significant comoving volumes with
sufficient resolution, many billions of particles are required. Particularly impressive is the "Millennium Run" computed by the Virgo consortium \cite{Springel2005} which follows the evolution of a comoving volume with $500h^{-1}$ Mpc on a side populated with 10 billion of particles.

The dark matter component can be modelled independently since it interacts only gravitationally and is the dominant matter component, i.e. its evolution is not influenced by baryons. However the evolution of baryons (e.g. the evolution of galaxies) which constitute the visible component of the universe, is much more difficult to model because it is governed by very complex physical mechanisms (e.g. magneto-hydro-dynamics, radiative transfer etc.). Currently it is not possible to follow the evolution of baryons with a self-consistent physical simulation. One of the most popular simplified alternatives is that of Semi-Analytical Models (SAM) where the evolution of baryons is included with analytical prescriptions within the dark matter framework, i.e. starting from the results of the N-body simulations following the evolution of dark matter.

The process of baryonic structure formation leaves a fossil record in the extragalactic background light \cite{Hauser2001}, for a review which provides a tight constraint for galaxy formation models. Three components are clearly identifiable:

- the CMB (Cosmic Micro-wave Background) which is made of photons emitted at the last scattering surface (i.e. when the universe was nearly homogeneous at $z \sim 1000$);
- the CIB and COB (Cosmic Infrared Background and Cosmic Optical Background) which is due to the integrated emission from stars and, for a small fraction ($< 10\%$ of total) \cite{Silva2004}, to AGNs;
- the CXB (Cosmic X-ray Background) which is entirely due to integrated emission from AGNs.

These backgrounds provide a measure of the total radiation energy density involved in the processes of galaxy and AGN (i.e. accreting black holes) formation and evolution.

The main open issues in the evolution of baryons can be summarized in three questions:

- How do galaxies form and evolve?
- How do Active Galactic Nuclei evolve and what is their role in the life of the host galaxy?
- How do Supermassive Black Holes form and grow to masses $10^6 - 10^{10}$ M$_\odot$ and why is their mass related to the host spheroid?

In particular, is there a relation between the processes of Galaxy evolution, AGN evolution and growth of supermassive Black Holes?

2. Galaxies

The two competing scenarios of galaxy formation are the monolithic and hierarchical ones.

In the monolithic model the proto-galaxy forms after the fast collapse of primordial gas clouds and stars are formed on timescales much shorter than the free-fall time. Galaxy evolution is then characterized by a very strong star formation episode at high redshift followed by pure luminosity evolution.

In the hierarchical model galaxy formation starts from primordial density fluctuations in the CDM framework. Dark matter halos grow continuously through repeated merging events which also gradually increase the gas mass in the growing galaxies hosted in the same halos. Galaxy evolution is characterized by moderate star formation at all redshifts which is driven by merging events. In this scenario, a late final assembly is expected for the most massive galaxies.

The hierarchical model is currently preferred because it fits nicely in a self consistent scenario of structure formation from primordial density fluctuations.

From the observational point of view, support to the hierarchical model comes from the average size of galaxies which decreases with decreasing redshift as expected in a hierarchical buildup \cite{Ferguson2004,Cassata2005}. The cosmic star formation history (i.e. the evolution of the global star formation rate, SFR, with redshift – \cite{Madau1998}, see
Semi-analytical hierarchical models are unable to account for the presence of these massive galaxies at high redshifts (Devriendt & Guiderdoni 2000, Cole et al. 2000, Somerville et al. 2001, Menci et al. 2002) and, indeed, in the ΛCDM scenario the big massive galaxies should be the last objects to form, no earlier than \(z \sim 1\). Observations seem to suggest formation redshifts in agreement with the monolithic scenario (see e.g. Pozzetti et al. 2003). Massive ellipticals do pose a problem for hierarchical models of galaxy formation.

### 3. Active Galactic Nuclei

Active Galactic Nuclei are galaxy nuclei where large luminosities (up to \(10^{14} \, L_\odot\)) are released in small volumes of space (\(r < 1 \, \text{pc}\)). It is commonly believed that AGNs are powered by mass accretion onto a supermassive BH (\(M_{\text{BH}} \sim 10^6 - 10^{10} \, M_\odot\)) and, according to the so-called Unified Model, all sources are intrinsically the same and differences are due to orientation dependent obscuration by a thick, dusty torus (see e.g. Antonucci 1993). Given the huge luminosities released by an AGN one might wonder about the impact on the host galaxy and ask the question whether galaxy evolution is related to AGN evolution.

Quasars are the most luminous AGNs (\(L > 10^{12} \, L_\odot\)) and therefore they are the easier to detect at all \(z\) because their emission far outshines the one from the host galaxy. The results from many quasar surveys indicate that the density of quasars reached a peak at \(z \sim 2 - 3\) and declined steeply afterwards, strongly reminding the behaviour of the cosmic SFR (e.g. Fan et al. 2004, Croom et al. 2004, and references therein). However Quasars constitute only the high luminosity end of the AGN luminosity function and they cannot provide a census of the whole AGN population. When searching for the other AGNs one has to face the problems related to obscuration of the AGN and contamination by the host galaxy emission (for lower \(L\) AGNs). A possible solution is to search for X-ray emission (e.g. in the \(0.5-10\) keV range) which is an almost unique indicator of AGN activity, less sensitive to obscuration than other indicators associated to opti-
et al. 2004; Setti & Woltjer 1994). The X-ray background (XRB) model by Comastri et al. (1995; Gilli et al. 2001) finds that most of the AGN emission is obscured and originates from sources at $z \sim 1-2$, i.e., close to the peak of AGN activity. The ratio between obscured (type 2) and unobscured (type 1) objects can be as high as 10. For instance, the XRB model by Gilli et al. (2001) predicts that 85% of the AGN emission is obscured and a large number of obscured (type 2) quasars is required to fit the number counts. These results started a search for the "missing" type 2 quasars, required to fit the XRB but never found in large numbers.

Very recently the deep X-ray surveys performed with Chandra and XMM have helped to solve the discrepancy of the missing type 2 quasars. The observed redshift distribution of low $L$ AGNs peaks at significantly lower redshifts than predicted by pre-Chandra/XMM models of the XRB and about 60% of the 2-8 keV XRB arises at $z < 1$. As a consequence, the sources creating most of the 0.1-10 keV XRB have X-ray luminosities from $10^{42}$ up to a few $\times 10^{44}$ erg s$^{-1}$ (e.g., like local Seyfert 1 galaxies, not quasars). The fraction of AGN with X-ray obscuration drops with increasing luminosity from $\approx 60$% at $10^{42}$ erg s$^{-1}$ to $\approx 30$% at $10^{45}$ erg s$^{-1}$ (e.g., Ueda et al. 2003; Szokoly et al. 2004). Some type 2 quasars have been found in deep surveys (e.g., Norman et al. 2002; Stern et al. 2002; Barger et al. 2003; Szokoly et al. 2004) but these generally have $L_X \sim 10^{44} - 10^{45}$ erg s$^{-1}$, i.e., just above powerful Seyfert galaxies. In total, obscured quasars create only $\approx 10$% of the 0.1-10 keV XRB even if it is likely that many Compton-thick AGN remain still undetected. Brandt & Haszper (2005) provide a comprehensive review of recent X-ray surveys and of all the related issues.

The newest X-ray surveys and XRB synthesis models then start to provide a comprehensive view of the evolution of the whole AGN population with the important new result being that the luminosity density of low $L$ objects peaks at lower $z$ than high $L$ ones (Ueda et al. 2003; Fiore et al. 2003), strongly reminding the behaviour of massive galaxies seen in the previous section. Moreover, the detection of X-ray emission in many sub-mm selected galaxies indicate a close link between strong star formation and AGN activity (e.g., Alexander et al. 2005).

4. Supermassive Black Holes

In order to detect the presence of a supermassive BH and measure its mass, one have to compare the observed motions in the nuclear region with those expected from the stellar gravitational potential (derived analyzing galaxy images).

In the center of our own galaxy one can use single stars as test particles in the gravitational potential and, given the relatively small distance (8 kpc) it is possible to derive the full velocity vector of stars by combining proper motions with radial velocities. For a few stars it is then possible to reconstruct their orbits from multi-epoch observations and from those it is trivial to determine the mass of the BH which is $M_{BH} \sim 3 \times 10^6 M_\odot$ (e.g., Genzel et al. 2000; Schödel et al. 2003; Ghez et al. 2003) and references therein). In the case of very few nearby galaxies it is possible to measure the motions of single gas clouds through their H$_2$O
The existence of correlations between BH mass and the structural parameters of the host spheroid (mass, luminosity and stellar velocity dispersion) indicates a tight link between BH growth and host galaxy evolution. To explain the origin of these correlations, there have been countless models proposed where BH growth is limited by AGN feedback (Silk & Rees 1998) or by the amount of fuel provided by the host galaxy (Adams et al. 2001, Burkert & Silk 2001). In particular, Cavaliere & Vittorini (2002) have shown that the slope of the $M_{BH} - \sigma$ relation depends on whether the BH growth is self-regulated (e.g. by AGN feedback) or limited by ambient conditions (e.g. by the amount of fuel provided by the host galaxy).

A non-trivial question is whether the local BHs found in nearby galaxies are the relics of AGN activity, i.e. if they were grown by radiatively efficient mass accretion. The answer to this question comes from comparing the local BH mass function with the mass function of BHs grown entirely during AGN activity. The local BH mass function is obtained by combining the local galaxy luminosity (or velocity) function with the $M_{BH} - L$ or $M_{BH} - \sigma$ relations while the AGN BH mass function is found by applying the continuity equation to the AGN luminosity function (e.g. Marconi et al., 2004 and references therein). The comparison between the two mass functions indicates that indeed local BHs are relics of AGN activity but also shows two interesting points related to the topics addressed here. First, the cosmic BH accretion rate is roughly proportional to the cosmic SFR and, second, BHs grow in an anti-hierarchical fashion, i.e. big BHs found in nearby galaxies are the relics of AGN activity, i.e. if they were grown by radiatively efficient mass accretion or limited by ambient conditions (e.g. by AGN feedback) or limited by ambient conditions (e.g. by the amount of fuel provided by the host galaxy).

5. The interplay between galaxies, AGNs and BHs

We can now summarize the main results found in the previous sections.

- Massive elliptical galaxies form earlier ($z > 2 - 3$) than smaller ones (at $z > 2 - 3$).
- The evolution of the AGN luminosity density (i.e. the cosmic BH accretion rate) closely matches the evolution of the cosmic star formation rate. Galaxies undergoing strong episodes of star formation (e.g. the submm galaxies) also present traces of AGN activity.
- The formation and evolution of spheroids must be tightly linked to the growth of the supermassive BHs hosted in their nuclei.
A common element which can unify these results is the feedback from the accreting BH (i.e. the AGN) on the host galaxy. The feedback from AGNs combined with that from Supernovae can significantly delay the formation of small objects with respect to large objects, effectively resulting in an anti-hierarchical baryon collapse \cite{Granato2004}. New semi-analytical models by Granato et al. \cite{2004} and Menci et al. \cite{2004} give a central role to the mutual feedback between star formation and growth of the supermassive BH in the galaxy center. In the Granato et al. \cite{2004} model the gas cools to form stars, the radiation drag due to starlight decreases the cool gas angular momentum and causes inflow onto the BH thus powering an AGN. The feedback from SNe and from the AGN regulates star formation and gas inflow eventually expelling the residual gas stopping star formation and BH accretion. In the Menci et al. \cite{2004} model starbursts are triggered by galaxy interactions. These originate from the destabilization of cold galactic gas that occurs in galaxy encounters, which in part feeds to form the accretion onto black holes that powers quasars. The feedback from the BH regulates BH accretion and eventually stop star formation. Both models are able to explain, at least partially, the presence of big galaxies at large redshifts together with the $M_{BH}$-galaxy relations. Big galaxies and big BHs form first because they are in deep potential wells which mitigate the feedback effects. On the other hand small galaxies and small BHs are in shallower potential wells and thus their growth is delayed by the feedback effects with respect to the most massive objects. A nice visualization of the feedback effect on galaxy formation can be found in the new numerical simulations by Di Matteo et al. \cite{2005}.

6. Conclusions

Galaxy evolution, AGN evolution and BH growth are closely linked and represent different aspects of the same phenomenon, the formation of baryonic structures. Massive galaxies are formed from the larger density fluctuations and grow quickly in deep potential wells where the feedback from SNe and BHs is less important. Gas fuels star formation and BH growth and big galaxies form early ($z > 2$) with big BHs at their centers. Smaller galaxies (and their BHs) form in shallower potential wells where feedback effects are more important. Thus, the growth of small galaxies is delayed by feedback effects. Star formation and BH growth are closely linked because the feedback from the BH stops star formation expelling the gas and setting the $M_{BH}$-host galaxy relations. The hierarchical scenario valid for Cold Dark Matter is thus reversed for baryons. However, while we might have understood the key element to unify galaxy, AGN evolution and BH growth, we are still a long way from having a self consistent physical picture of the formation of baryonic structures.

Acknowledgements. I am grateful to the organizers of "Session IV: Galaxies and AGNs" F. Bertola and L. Maraschi for the invitation to the 2005 SAIt meeting.

References

Dunlop, J. S. 2001, New Astronomy Review, 45, 609