Galaxies

19

The galaxies are the fundamental building blocks of the Universe. Some of them are very simple in structure, containing only normal stars and showing no particular individual features. There are also galaxies that are almost entirely made of neutral gas. On the other hand, others are complex systems, built up from many separate interacting components—stars, neutral and ionised gas, dust, molecular clouds, magnetic fields, cosmic rays.... The galaxies may form small groups or large clusters in space. At the centre of many galaxies, there is a compact nucleus that may sometimes be so bright that it overwhelms all the normal radiation of the galaxy.

The luminosity of the brightest normal galaxies may correspond to 10^{12} solar luminosities, but most of them are much fainter—the smallest ones that have been discovered are about $10^5\,L_\odot$. Since galaxies do not have a sharp outer edge, to some extent their masses and radii depend on how these quantities are defined. If only the bright central parts are included, a giant galaxy may typically have a mass of about $10^{13}\,M_\odot$, and a radius of 30 kpc, and a dwarf, correspondingly, $10^7\,M_\odot$, and 0.5 kpc.

When the existence of dark matter was confirmed the whole field of galactic research was changed. It was only the dark matter that helped to give a consistent model for the origin and evolution of galaxies.

The fraction of dark matter of the total mass of galaxies varies considerably. In big galaxies there is about five times as much dark matter as ordinary visible matter. In the smallest galaxies the ratio of the dark matter to the visible matter can be hundreds or even thousands.

When smaller and smaller galaxies have been found the boundary between star clusters and galaxies has become obscure. The smallest galaxies are similar to dense star clusters. Therefore some researches have redefined galaxies: a star system is a galaxy if in addition to stars it contains also dark matter.

19.1 The Classification of Galaxies

A useful first step towards an understanding of galaxies is a classification based on their various forms. Although such a morphological classification must always be to some extent subjective, it provides a framework within which the quantitative properties of galaxies can be discussed in a systematic fashion. However, it should always be remembered that the picture thus obtained will be limited to those galaxies that are large and bright enough to be easily visible in the sky. An idea of the consequent limitations can be obtained from Fig. 19.1, showing the radii and magnitudes of normal galaxies. One sees that only within a narrow region of this diagram can galaxies be easily found. If a galaxy has too large a radius for its magnitude (small surface brightness), it will disappear in the background light from the night sky. On the other hand, if its radius is too small, it looks like a star and is not noticed on a photographic plate. In the following, we shall mainly be concerned with bright galaxies that fit within these limits.

H. Karttunen et al. (eds.), *Fundamental Astronomy*, DOI 10.1007/978-3-662-53045-0_19

If a classification is to be useful, it should at least roughly correspond to important physical properties of the galaxies. Most classifications accord in their main features with the one put forward by *Edwin Hubble* in 1926. Hubble's own version of the *Hubble sequence* is shown in Fig. 19.2. The various types of galaxies are ordered in a sequence from early to late types. There are three main types: *elliptical*, *lenticular*, and *spiral* galaxies. The spirals are divided into two sequences, *normal* and *barred* spirals. In addition, Hubble included a class of *irregular galaxies*. In lenticular and spiral galaxies the matter is concentrated in a flat disk and thus they can be called collectively as *disk galaxies*.

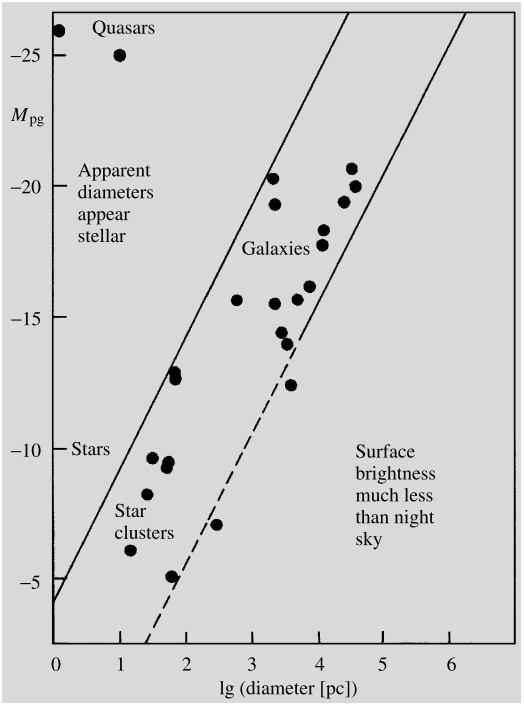


Fig. 19.1 Magnitudes and diameters of observable extragalactic objects. Objects to the upper left look like stars. The quasars in this region have been discovered on the basis of their spectra. Objects to the lower right have a surface brightness much smaller than that of the night sky. In recent years large numbers of low surface brightness galaxies have been discovered in this region. (Arp, H. (1965): Astrophys. J. **142**, 402)

The elliptical galaxies (Fig. 19.6) appear in the sky as elliptical concentrations of stars, in which the density falls off in a regular fashion as one goes outwards. Usually there are no signs of interstellar matter (dark bands of dust, bright young stars). The ellipticals differ from each other only in shape and on this basis they are classified as E0, E1, . . . , E7. If the major and minor axes of an elliptical galaxy are a and b, its type is defined to be En, where

$$n = 10\left(1 - \frac{b}{a}\right). \tag{19.1}$$

An E0 galaxy thus looks circular in the sky. The apparent shape of an E galaxy depends on the direction from which it is seen. In reality an E0 galaxy may therefore be truly spherical or it may be a circular disk viewed directly from above.

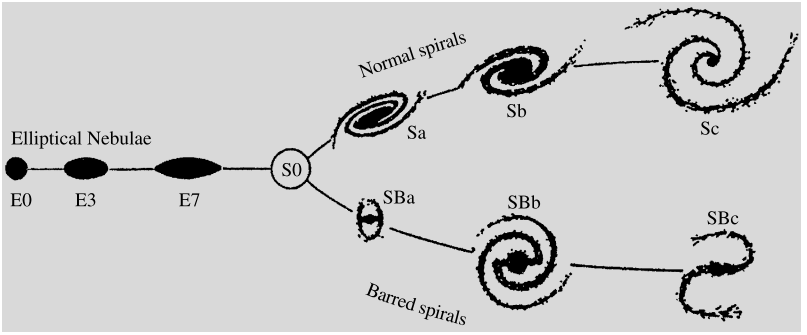


Fig. 19.2 The Hubble sequence in Hubble's 1936 version. At this stage the existence of type S0 was still doubtful. Photographs of the Hubble types are shown in Figs. 18.6 and 18.15 (E); 18.3 and 18.4 (S0 and S); 18.12 (S and Irr II); 18.5 (Irr I and dE). (Hubble, E.P. (1936): *The Realm of the Nebulae* (Yale University Press, New Haven))

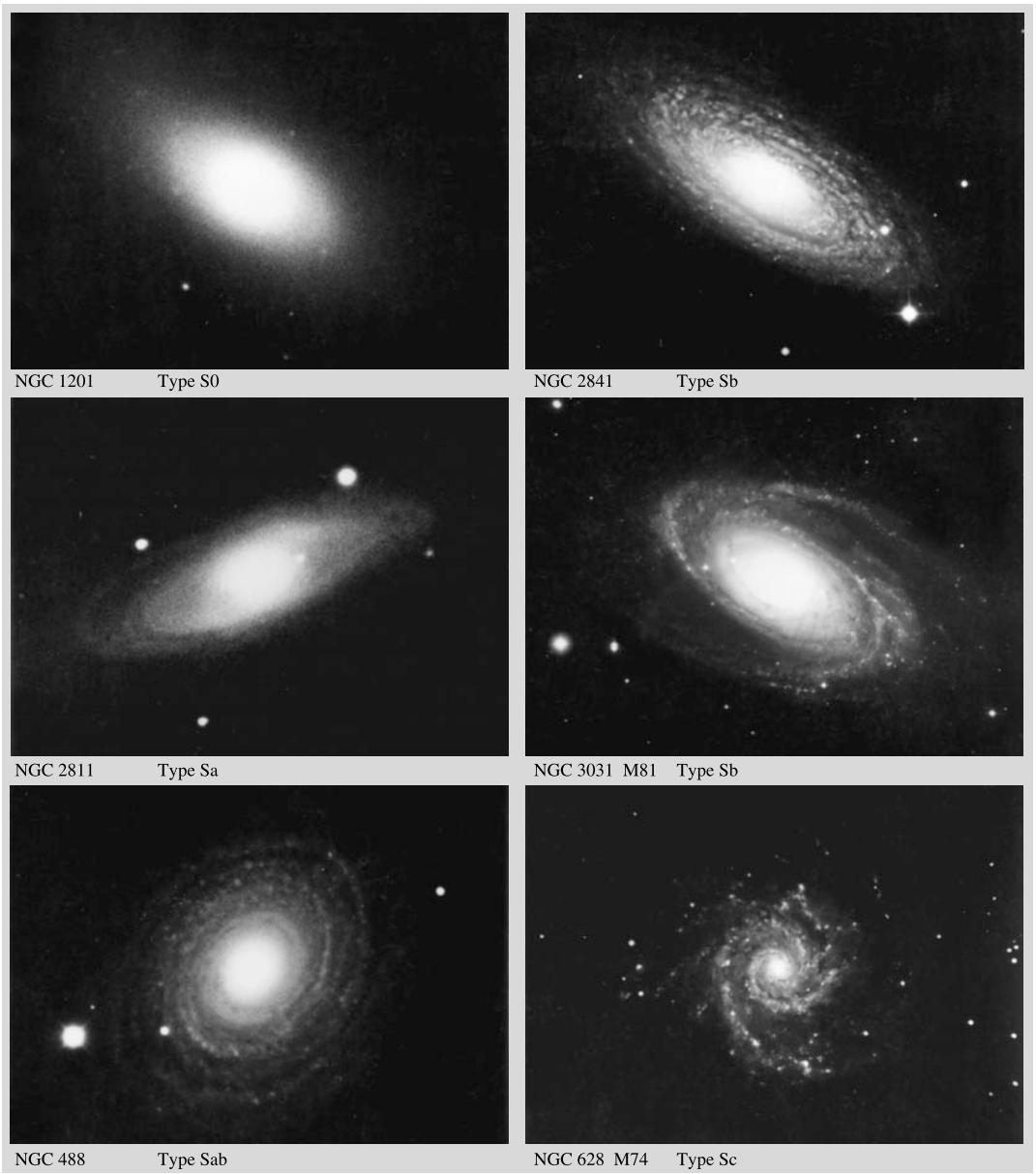


Fig. 19.3 The classification of normal spiral and S0 galaxies. (Mt. Wilson Observatory)

A later addition to the Hubble sequence is a class of *giant elliptical galaxies* denoted cD. These are generally found in the middle of clusters of galaxies. They consist of a central part looking like a normal elliptical surrounded by an extended fainter halo of stars.

In the Hubble sequence the lenticulars or S0 galaxies are placed between the elliptical and the spiral types. Like the ellipticals they contain only little interstellar matter and show no signs of spiral structure. However, in addition to the usual elliptical stellar component, they also contain a flat disk made up of stars. In this respect they are like spiral galaxies (Figs. 19.3, 19.4).

The characteristic feature of spiral galaxies is a more or less well-defined spiral pattern in the

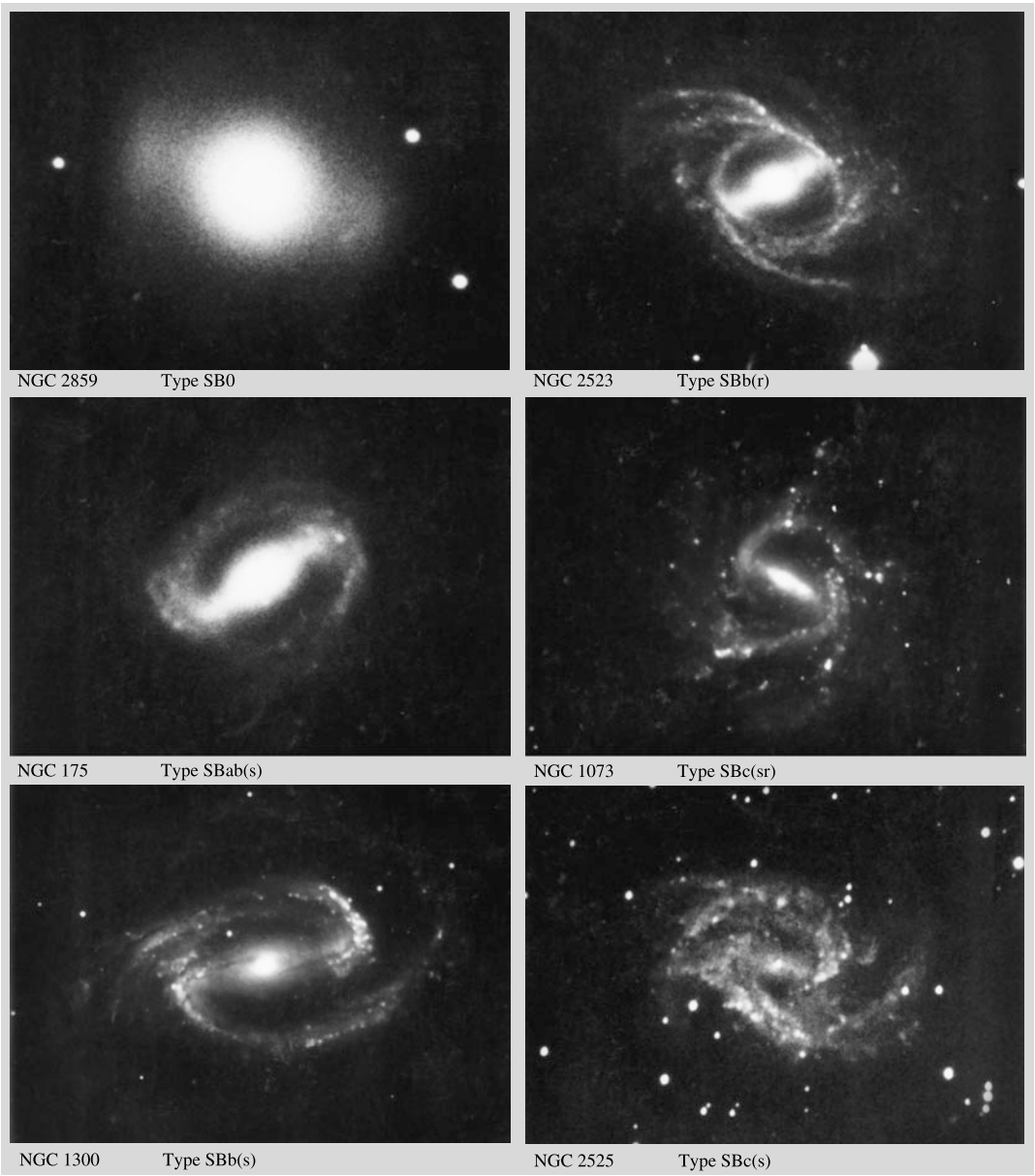


Fig. 19.4 Different types of SB0 and SB galaxies. The type (r) or (s) depends on whether the galaxy has a central ring or not. (Mt. Wilson Observatory)

disk. Spiral galaxies consist of a central *bulge*, which is structurally similar to an E galaxy, and of a stellar disk, like in an S0 galaxy. In addition to these, there is a thin disk of gas and other interstellar matter, where young stars are being born, forming the spiral pattern. There are two sequences of spirals, normal Sa–Sb–Sc, and barred SBa–SBb–SBc spirals. In the barred spirals the spiral pattern ends at a central bar, whereas in the normal spirals the spiral pattern may end at an inner ring or continue all the way to the centre. The position of a galaxy within the spiral sequence is determined on the basis of three criteria (which are not always in agreement): later types

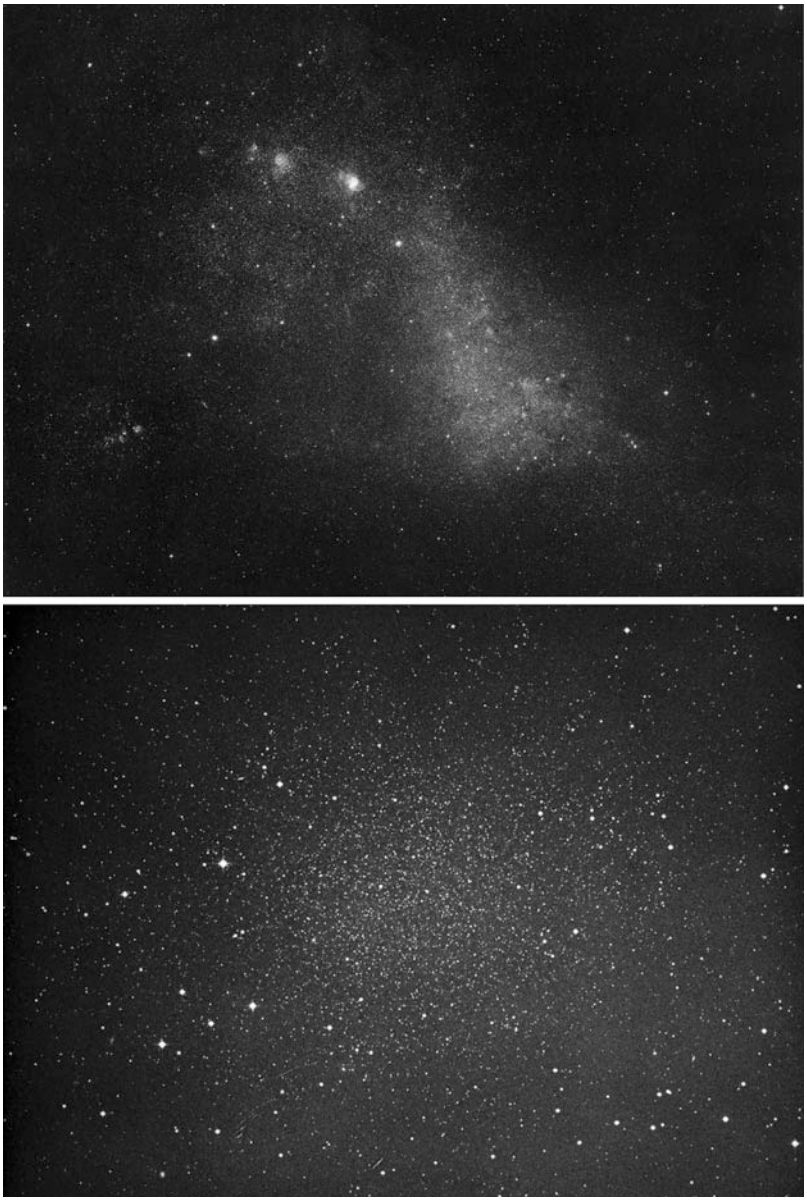


Fig. 19.5 *Above*: The Small Magellanic Cloud (Hubble type Irr I), a dwarf companion of the Milky Way (Royal Observatory, Edinburgh). *Below*: The Sculptor Galaxy, a dE dwarf spheroidal. (ESO)

have a smaller central bulge, more narrow spiral arms and a more open spiral pattern. The Milky Way Galaxy is thought to be of type SABbc (intermediate between Sb and Sc, and between normal and barred spirals).

The classical Hubble sequence is essentially based on bright galaxies; faint galaxies have been less easy to fit into it (Fig. 19.5). For example, the irregular galaxies of the original Hubble sequence can be divided into the classes Irr I and Irr II. The Irr I galaxies form a continuation of the Hubble sequence towards later types beyond the Sc galaxies. They are rich in gas and contain many young stars. Type Irr II are somewhat irregular small ellipticals containing only few young stars but a lot of interstellar dust.

The smallest irregular galaxies are classified as dIrr (dwarf irregulars). Another large group of dwarf galaxies are the spherical or elliptic dwarf galaxies dSph (dwarf spheroidals) and dE

Fig. 19.6 M32 (type E2), a small elliptical companion of the Andromeda Galaxy. (NOAO/Kitt Peak National Observatory)

(dwarf ellipsoids). They are much smaller than ordinary elliptic galaxies or the stellar density is much lower. Recently new kinds of even smaller galaxies have been found. The include *ultrafaint dwarfs* (uFd), *blue compact dwarfs* (BCD) and *ultracompact dwarfs* (UCD) (Fig. 19.7).

In addition to the precise morphological classification a simpler classification based on colours was introduced at the beginning of the 2000's. The idea to classify fast the millions of new galaxies found by modern digital survey programs as well as the very faint galaxies whose shape is difficult to distinguish. When galaxies are plotted like stars in a colour–magnitude diagram they are seen to form two distinct regions: *red galaxies* (red sequence) and *blue galaxies* (blue sequence) (Fig. 19.8). The red galaxies are mainly ellipticals with old stars and blue galaxies mainly spirals with younger stars.

Fig. 19.7 The ultracompact dwarf M60-UCD1 is close to the galaxy M60. The mass of the dwarf galaxy is 200×10^6 solar masses but the radius is only 100 lightyears. (Photo NASA, ESA, CXC, J. Strader (Michigan State University))

19.2 Luminosities and Masses

Distances In order to determine the absolute luminosities and linear dimensions of galaxies one needs to know their distances. Distances are also needed in order to estimate the masses of galaxies, because these estimates depend on the absolute linear size. Distances within the Local Group can be measured by the same methods as inside the Milky Way, most importantly by means of variable stars. On the very large scale (beyond 50 Mpc), the distances can be deduced on the basis of the expansion of the Universe (see

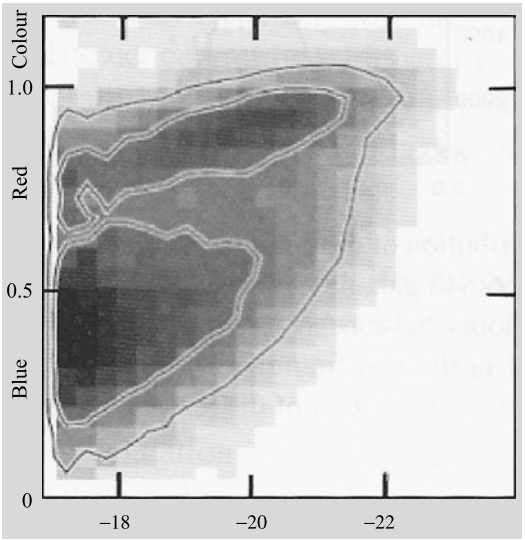


Fig. 19.8 A colour–magnitude graph contains 144,000 galaxies from the Sloan digital sky survey (SDSS). Brightness increases from left to right. Galaxies are clearly divided into two regions, red and blue. (Blanton et al., 2003, ApJ 592, 819)

Sect. 20.1). In order to connect these two regions one needs methods of distance determination based on the properties of individual galaxies.

To some extent local distances can be determined using structural components of galaxies, such as the sizes of H II regions or the magnitudes of globular clusters. However, to measure distances of tens of megaparsecs, one needs a distance-independent method to determine the absolute luminosities of entire galaxies. Several such methods have been proposed. For example, a luminosity classification has been introduced for late spiral types by *Sidney van den Bergh*. This is based on a correlation between the luminosity of a galaxy and the prominence of its spiral pattern. Nowadays this method is not considered accurate enough.

Other distance indicators are obtained if there is some intrinsic property of the galaxy, which is correlated with its total luminosity, and which can be measured independently of the distance. Such properties are the colour, the surface brightness and the internal velocities in galaxies. All of these have been used to measure distances to both spiral and elliptical galaxies.

For example, the absolute luminosity of a galaxy should depend on its mass. The mass, in turn, will be reflected in the velocities of stars and gas in the galaxy. Accordingly there is a relationship between the absolute luminosity and the velocity dispersion (in ellipticals) and the rotational velocity (in spirals). Since rotational velocities can be measured very accurately from the width of the hydrogen 21-cm line, the latter relationship (known as the *Tully–Fisher relation*) is perhaps the best distance indicator currently available.

The luminosity of the brightest galaxies in clusters has been found to be reasonably constant. This fact can be used to measure even larger distances, providing a method which is important in cosmology.

Luminosities The definition of the total luminosity of a galaxy is to some extent arbitrary, since galaxies do not have a sharp outer edge. The usual convention is to measure the luminosity of a galaxy out to a given value of the surface brightness, e.g. to 26.5 mag/sq.arcsec. For a given Hubble type, the total luminosity L may vary widely.

As in the case of stars, the distribution of galaxy luminosities is described by the *luminosity function* $\Phi(L)$ (Fig. 19.9). This is defined so that the space density of galaxies with luminosities between L and $L + \mathrm{d}L$ is $\Phi(L)\,\mathrm{d}L$. It can be determined from the observed magnitudes of galaxies, once their distances have been estimated in some way. In practice, one assumes some suitable functional form for $\Phi(L)$, which is then fitted to the observations. One common form is *Schechter's luminosity function*,

$$\Phi(L)\,\mathrm{d}L = \Phi^*\left(\frac{L}{L^*}\right)^{\alpha} \mathrm{e}^{-L/L^*}\,\mathrm{d}\left(\frac{L}{L^*}\right). \quad (19.2)$$

The values of the parameters Φ^*, L^*, α are observationally determined for different types of objects; in general, they will be functions of position.

The shape of the luminosity function is described by the parameters α and L^*. The relative number of faint galaxies is described by α. Since its observed value is about -1.1, the density of galaxies grows monotonically as one goes towards fainter luminosities. The luminosity function falls off steeply above the luminosity L^*,

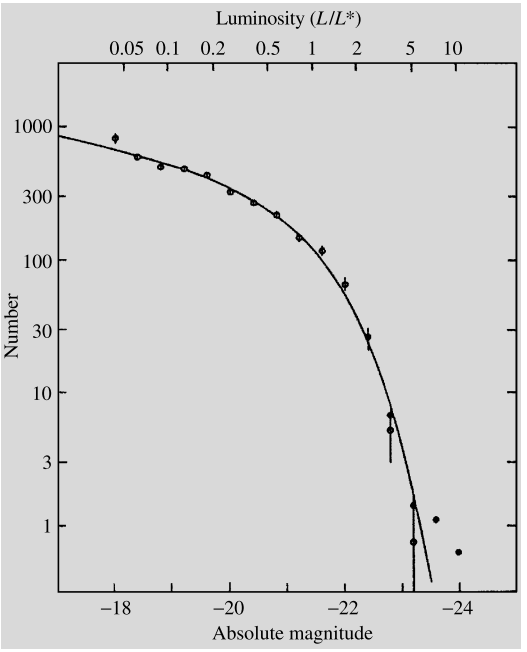


Fig. 19.9 Compound luminosity function of thirteen clusters of galaxies. *The open symbols* have been obtained by omitting the cD galaxies. The distribution is then well described by (19.2). The cD galaxies (*filled symbols*) cause a deviation at the bright end. (Schechter, P. (1976): Astrophys. J. **203**, 297)

which therefore represents a characteristic luminosity of bright galaxies. The observed L^* corresponds to an absolute magnitude $M^* = -21.0$ mag. The corresponding magnitude for the Milky Way Galaxy is probably -20.2 mag. The cD giant galaxies do not obey this brightness distribution; their magnitudes may be -24 mag and even brighter.

The parameter Φ^* is proportional to the space density of galaxies and is therefore a strong function of position. Since the total number density of galaxies predicted by relation (19.2) is infinite, we define n^* = density of galaxies with luminosity $> L^*$. The observed average value of n^* over a large volume of space is $n^* = 3.5 \times 10^{-3}\,\mathrm{Mpc}^{-3}$. The mean separation between galaxies corresponding to this density is 4 Mpc. Since most galaxies are fainter than L^*, and since, in addition, they often belong to groups, we see that the distances between normal galaxies are generally not much larger than their diameters.

Masses The distribution of mass in galaxies is a crucial quantity, both for cosmology and for theories of the origin and evolution of galaxies. Observationally it is determined from the velocities of the stars and interstellar gas. Total masses of galaxies can also be derived from their motions in clusters of galaxies. The results are usually given in terms of the corresponding mass-luminosity ratio M/L, using the solar mass and luminosity as units. The value measured in the solar neighbourhood of the Milky Way is $M/L = 3$. If M/L were constant, the mass distribution could be determined from the observed luminosity distribution by multiplying with M/L. Unfortunately, there is no universally valid value for the ratio M/L.

The masses of elliptical galaxies may be obtained from the stellar velocity dispersion given by the broadening of spectral lines. The method is based on the virial theorem (see Sect. 6.10), which says that in a system in equilibrium, the kinetic energy T and the potential energy U are related according to the equation

$$2T + U = 0. \tag{19.3}$$

Since ellipticals rotate slowly, the kinetic energy of the stars may be written

$$T = Mv^2/2, \tag{19.4}$$

where M is the total mass of the galaxy and v the velocity width of the spectral lines. The potential energy is

$$U = -GM^2/2R, \tag{19.5}$$

where R is a suitable average radius of the galaxy that can be estimated or calculated from the light distribution. Introducing (19.4) and (19.5) into (19.3) we obtain:

$$M = 2v^2R/G. \tag{19.6}$$

From this formula the mass of an elliptical galaxy can be calculated when v^2 and R are known. Some observations of velocities in ellipti-

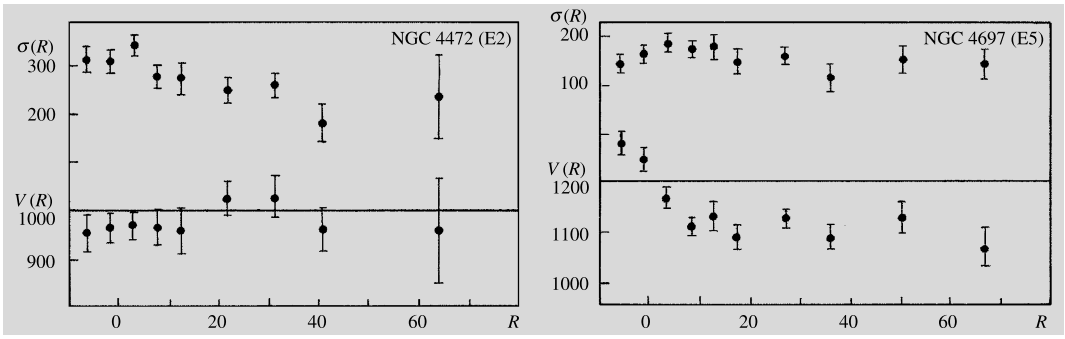


Fig. 19.10 Velocity of rotation $V(R)$ [km s^{-1}] and velocity dispersion $\sigma(R)$ [km s^{-1}] as functions of radius [kpc] for types E2 and E5. The latter galaxy is rotating, the former is not. (Davies, R.L. (1981): Mon. Not. R. Astron. Soc. **194**, 879)

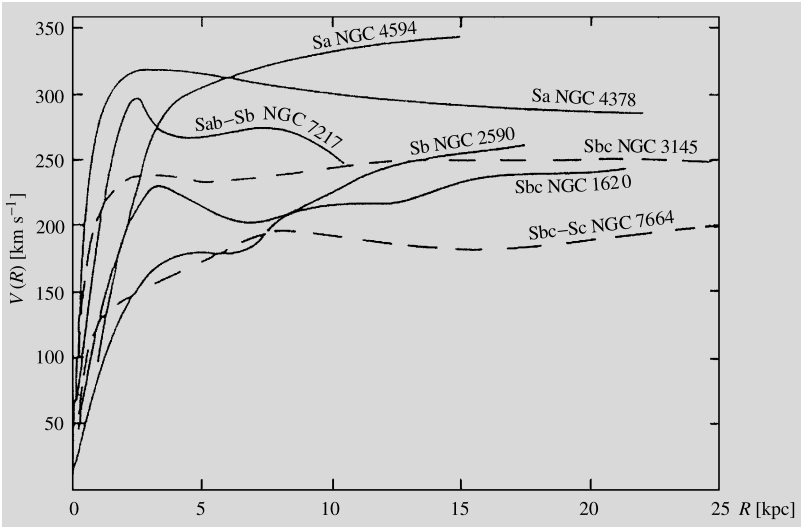


Fig. 19.11 Rotation curves for seven spiral galaxies. (Rubin, V.C., Ford, W.K., Thonnard, N. (1978): Astrophys. J. (Lett.) **225**, L107)

cal galaxies are given in Fig. 19.10. These will be further discussed in Sect. 19.4. The value of M/L derived from such observations is about 10 within a radius of 10 kpc. The mass of a bright elliptical might thus be up to $10^{13}\,M_\odot$.

The masses of spiral galaxies are obtained from their *rotation curve* $v(R)$, which gives the variation of their rotational velocity with radius. Assuming that most of the mass is in the almost spherical bulge, the mass within radius R, $M(R)$, can be estimated from Kepler's third law:

$$M(R) = Rv(R)^2/G. \qquad (19.7)$$

Some typical rotation curves are shown in Fig. 19.11. In the outer parts of many spirals, $v(R)$ does not depend on R. This means that $M(R)$ is directly proportional to the radius—the further out one goes, the larger the interior mass is. Since the outer parts of spirals are very faint, at large radii the value of M/L is directly proportional to the radius. For the disk, one finds that $M/L = 8$ for early and $M/L = 4$ for late spiral types. The largest measured total mass is $2 \times 10^{12}\,M_\odot$.

In order to measure the mass at even larger radii where no emission can be detected, motions in systems of galaxies have to be used. One possibility is to use pairs of galaxies. In principle, the method is the same as for binary stars. However, because the orbital period of a binary galaxy is about 10^9 years, only statistical information can

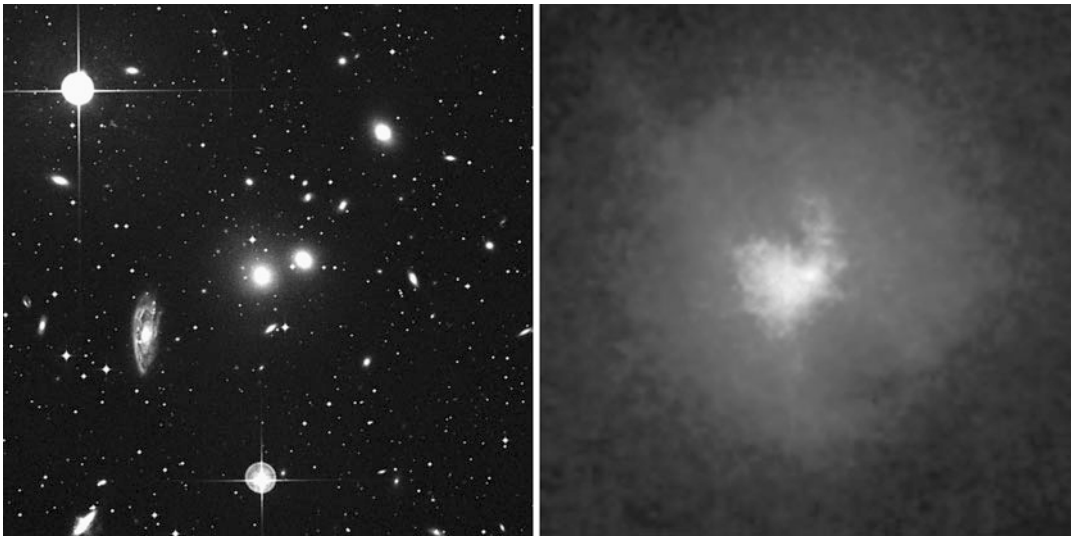


Fig. 19.12 Stars and hot gas in the Hydra galaxy cluster. *Left*: Optical image of the centre of the cluster. *Right*: An X-ray image of the same area in the same scale by the Chandra satellite. The temperature of the gas is about 4×10^7 K. (Optical image UK Schmidt/DSS1/ESO, SRC, X-ray image NASA/CXC/SAO)

be obtained in this way. The results are still uncertain, but seem to indicate values of $M/L = 20$–30 at pair separations of about 50 kpc.

A fourth method to determine galaxy masses is to apply the virial theorem to clusters of galaxies, assuming that these are in equilibrium. The kinetic energy T in (19.4) can then be calculated from the observed redshifts and the potential energy U, from the separations between cluster galaxies. If it is assumed that the masses of galaxies are proportional to their luminosities, it is found that M/L is about 200 within 1 Mpc of the cluster centre. However, there is a large variation from cluster to cluster.

There are also two newer and more accurate methods for mass determination. The fifth method utilises the X-ray radiation coming from the hot gas around galaxies or galaxy clusters (Fig. 19.12). The temperature of the gas around galaxy clusters is as high as 10^7, even E8 K. Since cloud seems to be more or less stable there must be a massive gravitational well preventing the gas from escaping. The dependence between the total mass of the galaxy cluster and the gas temperature is almost linear: the hotter the gas the more mass there is. The total mass of the richest galaxy clusters is over 10^{15} solar masses, and only about 1 % of this is in the form of visible matter.

Gravitational lenses (Sect. 19.7) are the most accurate method. A light ray passing by a galaxy is bent and the amount of bending depends on the mass of the galaxy. The method is generally used to determine masses of galaxy clusters but it can also be used for individual galaxies if the lensing galaxy forms several images of the background object or clearly distorts its shape. Gravitational lenses have finally confirmed that only about one sixth of the mass of galaxies is ordinary matter (like stars or gas) and the rest dark matter.

Dark Matter the most precise value for the dark matter has been obtained by the WMAP and Planck satellites which have studied the details of the cosmic background radiation. As it will be mentioned in Chap. 20, about 31 % of the total content (mass + energy) of the universe is matter. The final results of the Planck satellite published in 2015 show that the contribution of ordinary matter is 5 % and of cold dark matter 26 %.

The distribution of the dark matter can nowadays be determined by weak and strong gravitational lenses. The result is that the dark matter

Fig. 19.13 Distribution of dark matter around the galaxy cluster MACS J0717.5+3745. The background image was taken by the Hubble telescope. Superimposed is the distribution of dark matter calculated using strong and weak gravitational lenses. (NASA, ESA, Harald Ebeling (University of Hawaii at Manoa) & Jean-Paul Kneib (LAM))

appears as wide halos around galaxies and galaxy clusters (Fig. 19.13).

The nature of the dark matter is not yet known. It interacts with baryons almost only by its gravity. It is known to be "cold", which means that the velocities of the particles are low unlike with neutrinos that always travel almost at the speed of light.

It is assumed that the dark matter consists of unknown elementary particles. Possible candidates are light axions or heavy particles predicted by supersymmetry theories (WIMP, weakly interacting massive particle) like neutralinos. Their mass might be of the order 100 GeV–1 TeV, or hundreds of proton masses. Underground scintillation detectors and the LHC accelerator at the CERN are used to search for particles of the dark matter.

19.3 Galactic Structures

Ellipticals and Bulges In all galaxies the oldest stars have a more or less round distribution. In the Milky Way this component is represented by the population II stars. Its inner parts are called the bulge, and its outer parts are often referred to as the halo. There does not appear to be any physically significant difference between the bulge and the halo. The population of old stars can be best studied in ellipticals, which only contain this component. The bulges of spiral and S0 galaxies are very similar to ellipticals of the same size.

The surface brightness distribution in elliptical galaxies essentially depends only on the distance from the centre and the orientation of the major and minor axis. If r is the radius along the major axis, the surface brightness $I(r)$ is well described by *de Vaucouleurs' law*:

$$\log \frac{I(r)}{I_e} = -3.33\left[\left(\frac{r}{r_e}\right)^{1/4} - 1\right]. \qquad (19.8)$$

The constants in (19.8) have been chosen so that half of the total light of the galaxy is radiated from within the radius r_e and the surface brightness at that radius is I_e. The parameters r_e and I_e are determined by fitting (19.8) to observed brightness profiles. Typical values for elliptical, normal spiral and S0 galaxies are in the ranges $r_e = 1$–10 kpc and I_e corresponds to 20–23 magnitudes per square arc second.

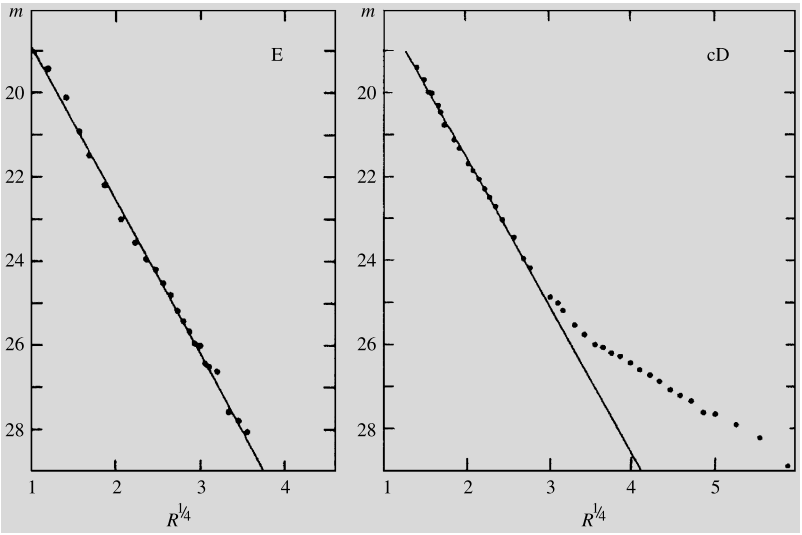


Fig. 19.14 The distribution of surface brightness in E and cD galaxies. *Ordinate*: surface magnitude, mag/sq.arcsec; *abscissa*: (radius [kpc])$^{1/4}$. Equation (18.8) corresponds to a straight line in this representation. It fits well with an E galaxy, but for type cD the luminosity falls off more slowly in the outer regions. Comparison with Fig. 19.15 shows that the brightness distribution in S0 galaxies behaves in a similar fashion. cD galaxies have often been erroneously classified as S0. (Thuan, T.X., Romanishin, W. (1981): Astrophys. J. **248**, 439)

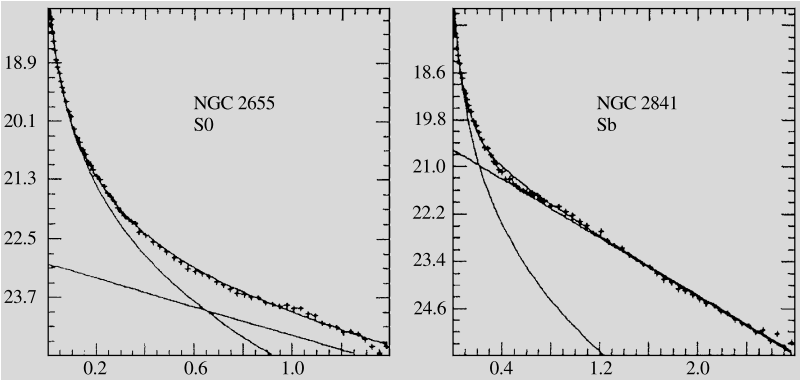


Fig. 19.15 The distribution of surface brightness in types S0 and Sb. *Ordinate*: mag/sq.arc sec; *abscissa*: radius [arc sec]. The observed surface brightness has been decomposed into a sum of bulge and disc contributions. Note the larger disc component in type Sb. (Boroson, T. (1981): Astrophys. J. Suppl. **46**, 177)

Although de Vaucouleurs' law is a purely empirical relation, it still gives a remarkably good representation of the observed light distribution. However, in the outer regions of elliptical galaxies, departures may often occur: the surface brightness of dwarf spheroidals often falls off more rapidly than (19.8), perhaps because the outer parts of these galaxies have been torn off in tidal encounters with other galaxies. In the giant galaxies of type cD, the surface brightness falls off more slowly (see Fig. 19.14). Such galaxies are usually at the centre of dense galaxy clusters, and during billions of years they have absorbed several smaller galaxies. The stars of the disrupted galaxies usually remain orbiting the giant galaxy in its outskirts increasing the brightness of the outer parts. cD galaxies have been called the cannibals of the world of galaxies.

Although the isophotes in elliptical galaxies are ellipses to a good approximation, their ellipticities and the orientation of their major axes may vary as a function of radius. Different galaxies differ widely in this respect, indicating that

the structure of ellipticals is not as simple as it might appear. In particular, the fact that the direction of the major axis sometimes changes within a galaxy suggests that some ellipticals may not be axially symmetric in shape. A probable reason is that many galaxies of different shapes rotating in different directions have assimilated to the current galaxies.

From the distribution of surface brightness, the three-dimensional structure of a galaxy may be inferred as explained in Box 19.1. The relation (19.8) gives a brightness profile which is very strongly peaked towards the centre. The real distribution of axial ratios for ellipticals can be statistically inferred from the observed one. On the (questionable) assumption that they are rotationally symmetric, one obtains a broad distribution with a maximum corresponding to types E3–E4. If the true shape is not axisymmetric, it cannot even statistically be uniquely determined from the observations.

Disks A bright, massive stellar disk is characteristic for S0 and spiral galaxies, which are therefore called *disk galaxies*. There are indications that in some ellipticals there is also a faint disk hidden behind the bright bulge. In the Milky Way the disk is formed by population I stars.

The distribution of surface brightness in the disk is described by the expression

$$I(r) = I_0 \, \mathrm{e}^{-r/r_0}. \qquad (19.9)$$

Figure 19.11 shows how the observed radial brightness distribution can be decomposed into a sum of two components: a centrally dominant bulge and a disk contributing significantly at larger radii. The central surface brightness I_0 typically corresponds to 21–22 mag./sq.arcsec, and the radial scale length $r_0 = 1$–5 kpc. In Sc galaxies the total brightness of the bulge is generally slightly smaller than that of the disk, whereas in earlier Hubble types the bulge has a larger total brightness. The thickness of the disk, measured in galaxies that are seen edge-on, may typically be about 1.2 kpc. Sometimes the disk has a sharp outer edge at about $4\,r_0$.

The Interstellar Medium Elliptical and S0 galaxies contain very little interstellar gas. However, in some ellipticals neutral hydrogen amounting to about 0.1 % of the total mass has been detected, and in the same galaxies there are also often signs of recent star formation. In some S0 galaxies much larger gas masses have been observed, but the relative amount of gas is very variable from one galaxy to another. The lack of gas in these galaxies is rather unexpected, since during their evolution the stars release much more gas than is observed.

The relative amount of neutral hydrogen in spiral galaxies is correlated with their Hubble type. Thus Sa spirals contain about 2 %, Sc spirals 10 %, and Irr I galaxies up to 30 % or more.

The distribution of neutral atomic hydrogen has been mapped in detail in nearby galaxies by means of radio observations. In the inner parts of galaxies the gas forms a thin disk with a fairly constant thickness of about 200 pc, sometimes with a central hole of a few kpc diameter. The gas disk may continue far outside the optical disk, becoming thicker and often warped from the central disk plane.

Most of the interstellar gas in spiral galaxies is in the form of molecular hydrogen. The hydrogen molecule cannot be observed directly, but the distribution of carbon monoxide has been mapped by radio observations. The distribution of molecular hydrogen can then be derived by assuming that the ratio between the densities of CO and H_2 is everywhere the same, although this may not always be true. It is found that the distribution obeys a similar exponential law as the young stars and H II regions, although in some galaxies (such as the Milky Way) there is a central density minimum. The surface density of molecular gas may be five times larger than that of H I, but because of its strong central concentration its total mass is only perhaps two times larger.

The distribution of cosmic rays and magnetic fields in galaxies can be mapped by means of radio observations of the synchrotron radiation from relativistic electrons. The strength of the magnetic field deduced in this way is typically 0.5–1 nT. The observed emission is polarised,

Fig. 19.16 A spiral galaxy NGC 5907 seen sideways. The galaxy is surrounded by star streams that are remnants of small dwarf galaxies merged to the big galaxy. (Photo Gabany, Martinez-Delgado et al., 2010, AJ 140, 962)

showing that the magnetic field is fairly well-ordered on large scales. Since the plane of polarisation is perpendicular to the magnetic field, the large-scale structure of the magnetic field can be mapped. However, the plane of polarisation is changed by Faraday rotation, and for this reason observations at several wavelengths are needed in order to determine the direction of the field. The results show that the field is generally strongest in the plane of the disk, and is directed along the spiral arms in the plane. The field is thought to have been produced by the combined action of rising elements of gas, perhaps produced by supernova explosions, and the differential rotation, in principle in the same way as the production of solar magnetic fields was explained in Chap. 13.

The Outer Parts of Galaxies Even the outermost parts outside the visible disk are not empty. Recently they have been seen to contain several different kinds of matter. The most abundant is the dark matter extending around the galaxy as a sphere or slightly flattened ellipsoid to a distance of tens of times further than the visible matter.

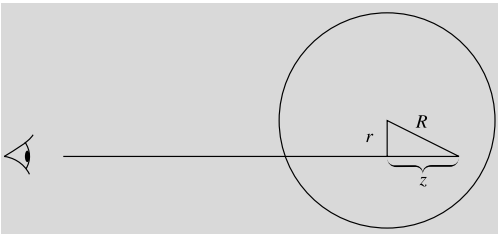
Hot gas is the second largest component. Its high temperature and great abundance were surprising. For example, the Milky Way is surrounded by a region of hot gas extending as a spherical cloud as far as to 150 kiloparsecs. The temperature of the gas is 1–2.5 $\times$ 10^6 K and the total mass about the same order of magnitude as the total mass of the stars in the Milky Way.

The third component comprises of separate gas clouds having greater velocities than other matter of the galaxy. Some of the clouds come from the centre of the Milky Way, some from other galaxies.

Yet another component consists of star clusters and star streams. In addition to globular clusters

galaxies are accompanied by star concentrations that are remnants of dwarf galaxies assimilated to the galaxy. For example, the globular cluster M54 is not an original Milky Way cluster but a remnant of the Sagittarius dwarf ellipsoid moving through the central part of the Milky Way. Disrupting dwarf galaxies are sometimes seen as star streams around the galaxy (Fig. 19.16).

Box 19.1 (Three-Dimensional Shape of Galaxies) Equations (19.8) and (19.9) describe the distribution of galactic light projected on the plane of the sky. The actual three-dimensional luminosity distribution in a galaxy is obtained by inverting the projection. This is easiest for spherical galaxies.



Let us suppose that a spherical galaxy has the projected luminosity distribution $I(r)$ (e.g. as in (19.8)). With coordinates chosen according to the figure, $I(r)$ is given in terms of the three-dimensional luminosity distribution $\rho(R)$ by

$$I(r) = \int_{-\infty}^{\infty} \rho(R)\,\mathrm{d}z.$$

Since $z^2 = R^2 - r^2$, a change of the variable of integration yields

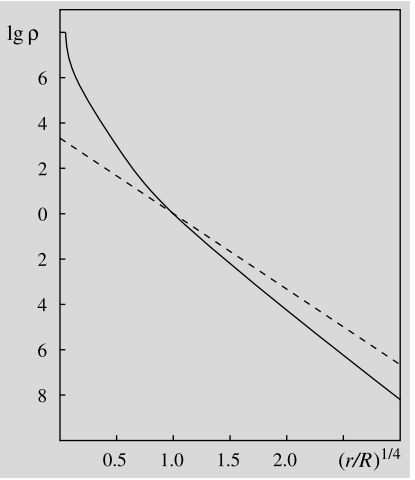
$$I(r) = 2\int_{r}^{\infty} \frac{\rho(R)R\,\mathrm{d}R}{\sqrt{R^2 - r^2}}.$$

This is known as an Abel integral equation for $\rho(R)$, and has the solution

$$\rho(R) = -\frac{1}{\pi R}\frac{\mathrm{d}}{\mathrm{d}R}\int_{R}^{\infty} \frac{I(r)r\,\mathrm{d}r}{\sqrt{r^2 - R^2}}$$

$$= -\frac{1}{\pi}\int_{R}^{\infty} \frac{(\mathrm{d}I/\mathrm{d}r)\,\mathrm{d}r}{\sqrt{r^2 - R^2}}.$$

Introducing the observed $I(r)$ into this expression, one obtains the actual luminosity distribution $\rho(R)$. In the figure the solid curve shows the three-dimensional luminosity distribution obtained from the Vaucouleurs' law (the dashed line).



If the galaxy is not spherical, its three-dimensional shape can only be determined if its inclination with respect to the line of sight is known. Since galactic disks are thin and of constant thickness, the inclination i of a disk galaxy is obtained directly from the axis ratio of its projected image: $\sin i = b/a$.

When the inclination is known, the real axis ratio of the bulge q_0 can be determined from the projected value q. For a rotationally symmetric bulge the relation between q and q_0 is

$$\cos^2 i = \frac{1 - q^2}{1 - q_0^2}.$$

The flattenings of disk galaxy bulges obtained from this relation lie in the range $q_0 = 0.3$–0.5. Since the inclinations of ellipticals are generally unknown, only the statistical distribution of q can be determined from that of q_0.

19.4 Dynamics of Galaxies

We have seen how the masses of galaxies can be derived from observed velocities of stars and gas.

The same observations can be used to study the internal distribution of mass in more detail.

Slowly Rotating Systems The dynamics of elliptical galaxies and disk galaxy bulges are studied by means of the Doppler shifts and broadenings of stellar absorption lines. Since a given absorption line is the sum of contributions from many individual stars, its Doppler shift gives their mean velocity, while its broadening is increased by an amount depending on the dispersion of stellar velocities around the mean. By observing how the wavelengths and widths of spectral lines behave as functions of the radius, one can get some insight into the distribution of mass in the galaxy.

Examples of the observed radial dependence of the rotational velocity and velocity dispersion derived for some ellipticals were given in Fig. 19.10. The observed rotational velocities are often small ($<100\ \mathrm{km\,s^{-1}}$), while the velocity dispersion may typically be about $200\ \mathrm{km\,s^{-1}}$. If elliptical galaxies were in fact ellipsoids of revolution, there should be a statistical relation (when projection effects have been taken into account) between flatness, rotational velocity and velocity dispersion. Such a relationship has been observed for fainter ellipticals and for disk galaxy bulges. However, some of the brightest ellipticals rotate very slowly. Therefore their flattening cannot be due to rotation.

The radial dependence of the velocity dispersion gives information on the distribution of mass within the galaxy. Since it also depends on how the shapes of stellar orbits in the galaxy are distributed, its interpretation requires detailed dynamical models.

Some ellipticals have different velocity systems. An extreme case is NGC 4550 classified as E7/S0. Half of its stars orbit the centre in one direction and the rest in the opposite direction. This indicates that the galaxy is a merger of two spiral galaxies that originally rotated in opposite directions. Mutual distances of stars in ellipticals are so large that collisions do not occur and the current system is very stable.

Rotation Curves In spiral galaxies the distribution of mass can be studied directly using the observed rotational velocities of the interstellar gas. This can be observed either at optical wavelengths from the emission lines of ionised gas in H II regions or at radio wavelengths from the hydrogen 21 cm line. Typical galactic rotation curves were shown in Fig. 19.11.

The qualitative behaviour of the rotation curve in all spiral galaxies is similar to the rotation curve of the Milky Way: there is a central portion, where the rotational velocity is directly proportional to the radius, corresponding to rigid body rotation. At a few kpc radius the curve turns over and becomes flat, i.e. the rotational velocity does not depend on the radius. In early Hubble types, the rotation curve rises more steeply near the centre and reaches larger velocities in the flat region (Sa about $300\ \mathrm{km\,s^{-1}}$, Sc about $200\ \mathrm{km\,s^{-1}}$). A higher rotational velocity indicates a larger mass according to (19.7), and thus Sa types must have a larger mass density near the centre. This is not unexpected, since a more massive bulge is one of the defining properties of early type spirals.

A decrease of the rotational velocity at large radii would be an indication that most of the mass is inside that radius. In some galaxies such a decrease has been detected, in others the rotational velocity remains constant as far out as the observations can reach.

Spiral Structure Spiral galaxies (Fig. 19.17) are relatively bright objects. Some have a well-defined, large-scale two-armed spiral pattern, whereas in others the spiral structure is made up of a large number of short filamentary arms. From galaxies where the pattern is seen in front of the central bulge, it has been deduced that the sense of winding of the spiral is trailing with respect to the rotation of the galaxy. However, the rule is not absolute: there are also galaxies with leading arms.

The spiral structure is most clearly seen in the interstellar dust, H II regions, and the OB associations formed by young stars. The dust often forms thin lanes along the inner edge of the spiral arms, with star forming regions on their outside. Enhanced synchrotron radio emission associated with spiral arms has also been detected.

Fig. 19.17 *Above*: A spiral galaxy from above: M51 (type Sc). The interacting companion is NGC 5195 (type Irr II). (Lick Observatory). *Below*: A spiral galaxy from the side: the Sb spiral NGC 4565. (NOAO/Kitt Peak National Observatory)

A well formed wide spiral pattern is relatively common, and hence it must be a long lasting phenomenon. The spiral pattern is generally thought to be a wave in the density of the stellar disk, as discussed in Sect. 18.4. As the interstellar gas streams through the density wave a shock, marked by the dust lanes, is formed as the interstellar gas is compressed, leading to the collapse of molecular clouds and the formation of stars. The density wave theory predicts characteristic streaming motions within the arm, which have been detected in some galaxies by observations of the H I 21 cm line.

There may be several reasons producing spiral arms. The most common is an external perturbation when another galaxy passes by at a close dis-

tance creating a wave in the galaxy. Other sources of perturbation could be the bar of the galaxy or the gravitational field of a temporary "traffic jam" of stars.

A single close encounter or a traffic jam can explain temporary spiral arms but not more permanent features. The differential rotation of the galaxy tends to wipe out the arms during a few revolutions or in a few hundreds of millions of years. Some renewal mechanism is needed to keep the arms permanent.

Standing waves can explain the permanence. Birth of spiral arms and their motion generate density waves travelling through the whole galaxy. The waves can be reflected or refracted by the central bulge. When wave fronts moving in different directions meet they can either damp or amplify each others. Spiral arms are standing waves, fronts of amplified waves, seen as concentrates of stars and gas clouds. The spiral forms move around the galaxy at constant speed.

19.5 Stellar Ages and Element Abundances in Galaxies

From the Milky Way we know that stars of populations I and II are different not only in respect to their spatial distribution, but also in respect to their ages and heavy element abundances. This fact gives important evidence about the formation of the Milky Way, and it is therefore of interest if a similar connection can be found in other galaxies.

The indicators of composition most easily measured are the variations of colour indices inside galaxies and between different galaxies. Two regularities have been discovered in these variations: First, according to the *colour–luminosity relation* for elliptical and S0 galaxies, brighter galaxies are redder. Secondly, there is a *colour–aperture effect*, so that the central parts of galaxies are redder. For spirals this relationship is due to the presence of young, massive stars in the disk, but it has also been observed for elliptical and S0 galaxies.

Galactic spectra are composed of the spectra of all their stars added together. Thus the colours depend both on the ages of the stars (young stars are bluer) and on the heavy element abundance Z (stars with larger Z are redder). The interpretation of the observational results thus has to be based on detailed modelling of the stellar composition of galaxies or *population synthesis*.

Stars of different spectral classes contribute different characteristic absorption features to the galaxy spectrum. By observing the strength of various spectral features, one can find out about the masses, ages and chemical composition of the stars that make up the galaxy. For this purpose, a large number of characteristic properties of the spectrum, strengths of absorption lines and broad-band colours are measured. One then attempts to reproduce these data, using a representative collection of stellar spectra. If no satisfactory solution can be found, more stars have to be added to the model. The final result is a population model, giving the stellar composition of the galaxy. Combining this with theoretical stellar evolution calculations, the evolution of the light of the galaxy can also be computed.

Population synthesis of E galaxies show that practically all their stars were formed simultaneously about $13–14 \times 10^9$ years ago, which is close to the age of the universe Most of their light comes from red giants, whereas most of their mass resides in lower main sequence stars of less than one solar mass.

Since all stars have roughly the same age, the colours of elliptical galaxies are directly related to their metallicities. Thus the colour–luminosity relation indicates that Z in giant ellipticals may be double that in the solar neighbourhood, while it may be smaller by a factor 100 in dwarfs. Similarly, the radial dependence of the colours can be explained if the value of Z at the centre is an order of magnitude larger than it is at larger radii.

The stellar composition of disk galaxy bulges is generally similar to that of ellipticals. The element abundances in the gas in spirals can be studied by means of the emission lines from H II regions ionised by newly formed stars. In this case too, the metallicity increases towards the centre.

The birth histories of galaxies explain partly the different metallicities and age distributions. As we will see later, galaxies collide and merge in the span of billions of years. Most clearly this

is seen in dwarf galaxies having sow few stars that even a single merger leaves a distinct signature. Figure 19.18 shows the age distribution of two Milky Way's neighbour galaxies. A few of the stars of Leo A were born immediately after the origin of the universe but most of the stars were born during the last 5–6 billion years. The stars of the Carina dwarf galaxy were born during three separate periods, at the age of 1–2, 6–8 and about 11 billion years.

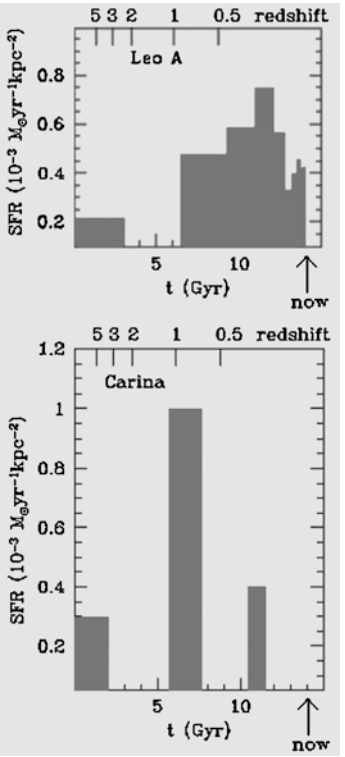


Fig. 19.18 The age distribution of the dwarf galaxies in Leo and Carina shows that the stars of the galaxies were formed in separate collision episodes. (Tolstoy, Hill, Tosi, 2009, Ann. Rev. Astron. Astrophys. 47, 371)

19.6 Systems of Galaxies

The galaxies are not smoothly distributed in space; rather, they form systems of all sizes: galaxy pairs, small *groups*, large *clusters* and *superclusters* formed from several groups and clusters. The larger a given system, the less its density exceeds the mean density of the Universe. On the average, the density is twice the background density for systems of radius 5 Mpc and 10 % above the background at radius 20 Mpc.

Interactions of Galaxies Galaxies do not evolve in isolation; instead, they interact continuously with other galaxies. Each large galaxy has experienced encounters or collisions with several other large galaxies and tens or even hundreds encounters with much smaller galaxies.

Among the most common distortions are "bridges" and "tails" caused by tidal forces during an encounter (Fig. 19.19). Sometimes a small galaxy breaks down and forms a ring in the equatorial plane of the larger galaxy or sometimes on a polar orbit. The most striking ring galaxies are born when a dense galaxy passes straight through another one (Fig. 19.20). In many cases the effects of collisions are seen as starbursts when the gas and dust clouds of the galaxies meet, condense and start to produce new stars. Sometimes image processing is needed to show faint but sharp rings around a galaxy (Fig. 19.21).

The interactions between galaxies are not always dramatic. For example, the Milky Way has two satellites, the Large and Small Magellanic Clouds (see Fig. 19.5), which are Irr I type dwarf galaxies at about 60 kpc distance. It is thought that approximately 5×10^8 years ago, these passed the Milky Way at a distance of about 10–15 kpc, leaving behind the *Magellanic Stream*, a 180° long thin stream of neutral hydrogen clouds. Systems of this type, where a giant galaxy is surrounded by a few small companions, are quite common. Computations show that in many such cases the tidal interactions are so strong that the companions will merge with the parent galaxy at the next close approach. This is likely to happen to the Magellanic Clouds.

During earlier epochs in the Universe, when the density was larger, interactions between galaxies must have been much more common than at present. Thus it has been proposed that a large fraction of bright galaxies have undergone major mergers at some stage in their history. In particular, there are good reasons to believe that the slowly rotating, non-axisymmetric giant ellipticals may have formed by the merger of disk galaxies.

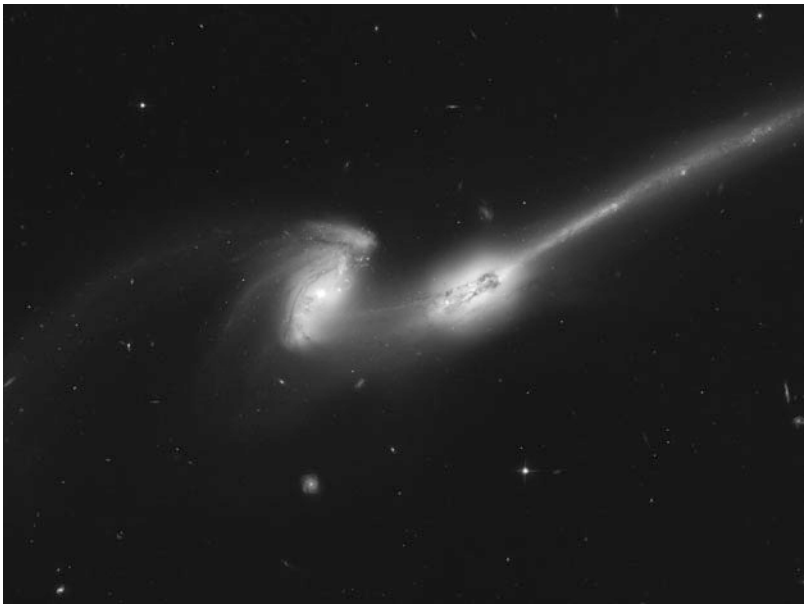


Fig. 19.19 The galaxy pair NGC 4676 is known as the Mice. The galaxies are about to collide and a long gas and star tail protrudes from each of them. (Hubble/NASA, H. Ford (JHU), G. Illingworth (UCSC/LO), M. Clampin (STScI), G. Hartig (STScI), the ACS Science Team, ja ESA)

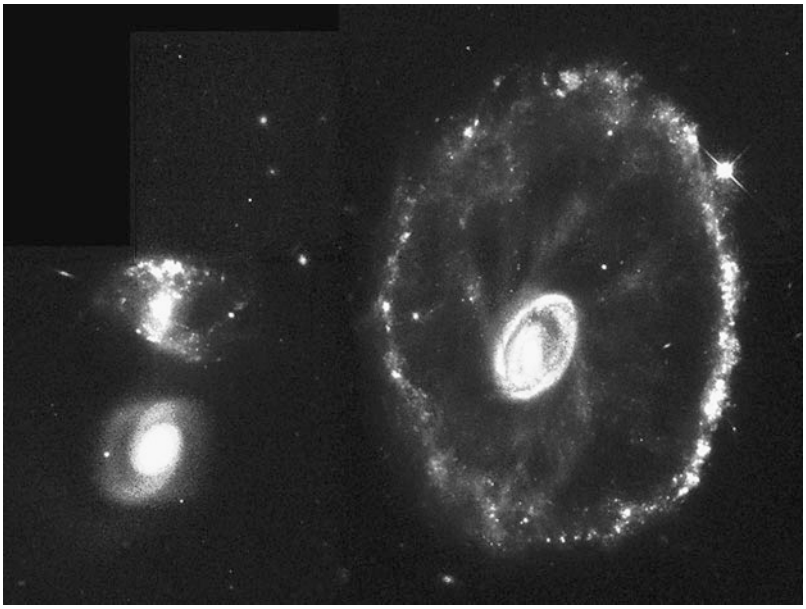


Fig. 19.20 The Cartwheel is the best known ring galaxy. It was born when a small galaxy passed through the centre of the big galaxy. (Photo Hubble, K. Borne (StScI) and Nasa)

Groups The most common type of galaxy systems are small, irregular groups of a few tens of galaxies. A typical example is the *Local Group*, which contains two larger galaxies in addition to the Milky Way—the Andromeda Galaxy M31, an Sb spiral of about the same size as the Milky Way with two dwarf companions, and the smaller Sc spiral M33. The rest of the about 50 members of the Local Group are dwarfs; about 20 are of type dE and 10 of type Irr I. The diameter of the Local Group is about 1.2 Mpc.

Clusters A system of galaxies may be defined to be a cluster if it contains a larger number (at least 50) of bright galaxies. The number of members and the size of a cluster depend on how they are defined. One way of doing this is to fit the

Fig. 19.21 Rings of stars with sharp boundaries in the outer parts of the galaxy NGC 474. The rings were born when a lenticular galaxy tore a rotating spiral galaxy layer by layer. (CFHT/Coelum—J.-C. Cuillandre & G. Anselmi)

observed distribution of galaxies within a cluster with an expression of the form (19.8). In this way a characteristic cluster radius of about 2–5 Mpc is obtained. The number of members depends both on the cluster radius and on the limiting magnitude. A large cluster may contain several hundred galaxies that are less than two magnitudes fainter than the characteristic luminosity L^* of (19.2).

Clusters of galaxies can be ordered in a sequence from extended, low-density, irregular systems (sometimes called clouds of galaxies) to denser and more regular structures (Fig. 19.23). The galaxy type composition also varies along this sequence in the sense that in the loose irregular clusters, the bright galaxies are predominantly spirals, whereas the members of dense clusters are almost exclusively type E and S0. The nearest cluster of galaxies is the Virgo Cluster at a distance of about 15 Mpc. It is a relatively irregular cluster, where a denser central region containing early galaxy types is surrounded by a more extended distribution of mainly spiral galaxies. The nearest regular cluster is the Coma Cluster, roughly 90 Mpc away. In the Coma Cluster a central pair of giant ellipticals is surrounded by a flat-

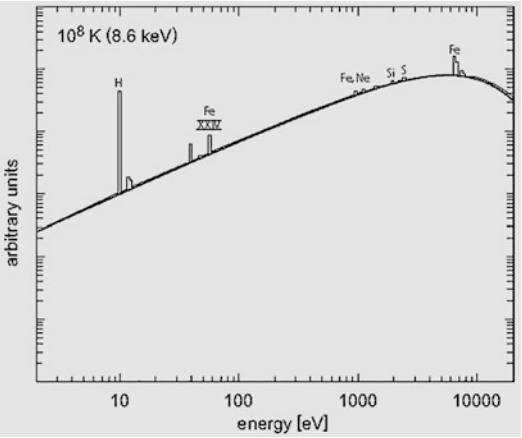


Fig. 19.22 A theoretical X-ray spectrum for plasma with a temperature of 10^8 K. Most of the radiation comes from the bremsstrahlung of electrons, peaking at 8.6 kiloelectronvolts (*thick line*). Superimposed are spectral lines produced by different ions. (Böhringer & Hensler, 1989, A&A 215, 147)

Fig. 19.23 *Above*: The irregular Virgo Cluster of galaxies. *Below*: The regular Coma Cluster (ESO and Karl-Schwarzschild-Observatorium)

tened (axis ratio about 2 : 1) system of early type galaxies.

Earlier galaxy groups were studied as conglomerates of visible galaxies. Recently two other, more massive, components have been included.

First of all, galaxy clusters are surrounded by large clouds of hot gas, as mentioned in Sect. 19.2. The gas can be observed by its X-ray emission. The spectrum shows features typical for bremsstrahlung of electrons and lines of totally ionised metals (like FeXXIV). Temperature measured from X-ray spectra are 10^7–10^8 K. The total mass of the gas in the poorest clusters is about 5 times and in the richest clusters even 20 times as big as the mass inside the galaxies. Part of the gas comes from the galaxies (jets of

quasars, supernova explosions), but most of it is hydrogen and helium from the big bang.

Besides in X-ray images the existence of the hot gas is seen also as extra spots in the cosmic background radiation, caused by the *Sunyayev-Zel'dovich effect*. Photons of the background radiation hit electrons of the hot gas and receive some extra energy (inverse Compton scattering). Below 218 GHz the radiation is slightly cooler and above 218 GHz a little brighter. The diameter of the spot is about one arc minute and the temperature difference from the surroundings 0.1–1 millikelvins (Fig. 19.25). One of the main objectives of the Planck satellite was to find gas clouds around galaxy clusters using this phenomenon. The final catalogue of Planck (2015) contains over 1600 objects most of which are previously unknown galaxies with redshifts $z = 0 - 1.5$.

Without a strong source of gravitation the hot gas would quickly disperse into space. The gravitation is caused by dark matter; its mass is about five times as big as that of the ordinary matter (hot gas + visible galaxies) in the cluster. Gravitational lenses have been used to investigate the dark matter in galaxy clusters (Sect. 19.7).

Superclusters Groups and clusters of galaxies may form even larger systems, superclusters. For example, the Local Group belongs to the *Local Supercluster*, a flattened system whose centre is the Virgo Cluster, containing tens of smaller groups and clouds of galaxies. The Coma Cluster is part of another supercluster. The diameters of superclusters are 10–20 Mpc. However, on this scale, it is no longer clear whether one can reasonably speak of individual systems. Perhaps it would be more accurate to think of the distribution of galaxies as a continuous network, where the large clusters are connected by walls and strings formed by smaller systems. Between these there remain empty regions containing very few galaxies, which can be up to 50 Mpc in diameter (Figs. 19.24, 20.8).

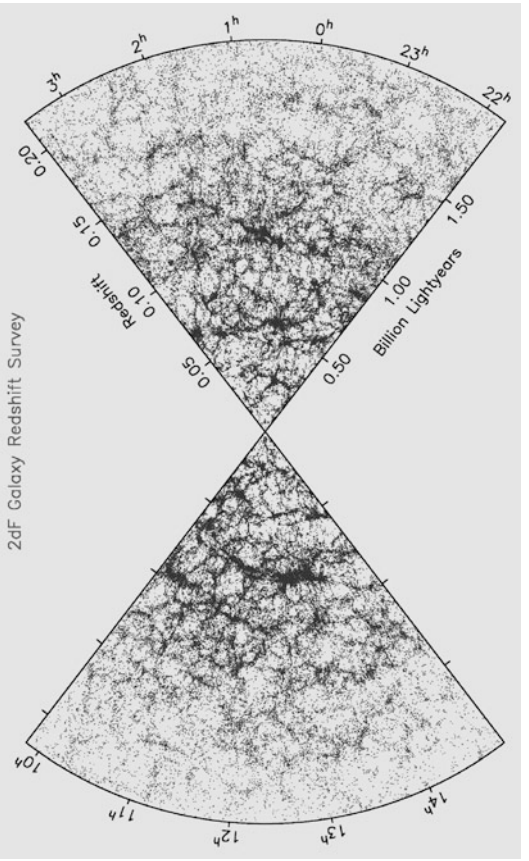


Fig. 19.24 The large-scale space distribution of 245,591 galaxies. The radial co-ordinate is the redshift, which can be translated into a distance using a value for the Hubble constant. The thickness of the slices is about 10°. (2dFGRS Team, http://www2.aao.gov.au/~TDFgg/)

19.7 Active Galaxies and Quasars

So far in this chapter we have been concerned with the properties of normal galaxies. In some galaxies, however, the normal galaxy is overshadowed by violent activity. This activity is produced in the nucleus, which is then called an *active galactic nucleus (AGN)*.

The luminosities of active galactic nuclei may be extremely large, sometimes much larger than that of the rest of the galaxy. It seems unlikely that a galaxy could maintain such a large power output for long. For this reason it is thought that active galaxies do not form a separate class of galaxies, but rather represent a passing stage in the evolution of normal galaxies.

Activity appears in many different forms. Some galaxies have an exceptionally bright nucleus similar to a large region of ionised hydrogen. These may be young galaxies, where near

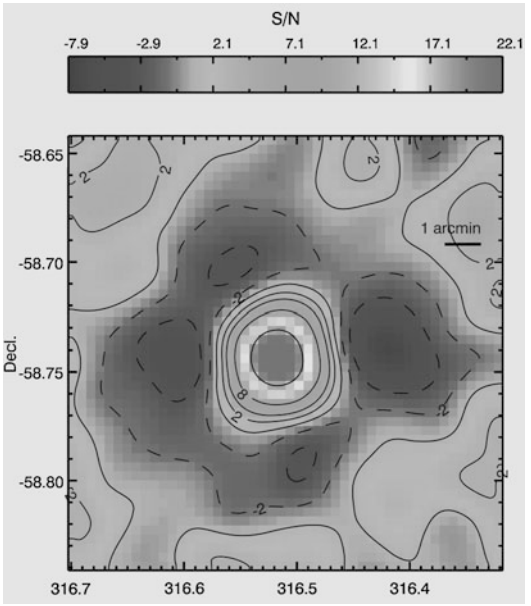


Fig. 19.25 A brightening of the cosmic background radiation caused by the Sunyayev-Zel'dovitš effect. The radio observation in the millimetre band were made with the South pole telescope. The amplitude of the brightening tells that the temperature of the gas cloud is 1.2×10^8 K and the total mass of the galaxy cluster about 10^{15} solar masses. Optical images of the same location show a distant galaxy cluster SPT-CL J2106–5844, which has a redshift $z = 1.13$. (Foley, Andersson, Bazin, de Haan, Ruel et al., 2011, ApJ 731, p. 89)

the centre large numbers of stars are forming and evolving into supernovae (starburst nuclei). In other nuclei the radiation cannot have been produced by stars, and the most plausible source of energy in these nuclei is the gravitational energy of a supermassive black hole (mass $> 10^8\, M_\odot$). In some galaxies, the spectral lines are unusually broad, indicating large internal velocities. These may be either rotational velocities near a black hole or due to explosive events in the nucleus. In some galaxies, jets are seen coming out of the nucleus. Many active galaxies radiate a nonthermal spectrum, apparently synchrotron radiation produced by fast electrons in a magnetic field.

The classification of active galaxies has been developed rather unsystematically, since many of them have been discovered only recently, and have not been completely studied. For example, the *Markarian galaxies* catalogued by *Benyamin Yerishevich Markarian* in the early 1970's are defined by strong ultraviolet emission. Many Markarian galaxies are Seyfert galaxies; others are galaxies undergoing a burst of star formation. The N galaxies form another class closely similar to the Seyfert galaxies.

Two natural basic classes of active galaxies are the *Seyfert galaxies* and the *radio galaxies*. The former are spirals; the latter are ellipticals. Some astronomers think that the Seyfert galaxies represent the active stage of normal spiral galaxies and the radio galaxies that of ellipticals.

Seyfert Galaxies The Seyfert galaxies are named after *Carl Seyfert*, who discovered them in 1943. Their most important characteristics are a bright, pointlike central nucleus and a spectrum showing broad emission lines. The continuous spectrum has a nonthermal component, which is most prominent in the ultraviolet. The emission lines are thought to be produced in gas clouds moving close to the nucleus with large velocities.

On the basis of the spectrum, Seyfert galaxies are classified as type 1 or 2. In a type 1 spectrum, the allowed lines are broad (corresponding to a velocity of 10^4 km s^{-1}), much broader than the forbidden lines. In type 2, all lines are similar and narrower ($<10^3$ km s^{-1}). Transitions between these types and intermediate cases have sometimes been observed. The reason for the difference is thought to be that the allowed lines are formed in denser gas near the nucleus, and the forbidden lines in more diffuse gas further out. In type 2 Seyfert galaxies, the denser gas is missing or obscured.

Almost all Seyfert galaxies with known Hubble types are spirals; the possible exceptions are of type 2. They are strong infrared sources. Type 1 galaxies often show strong X-ray emission.

The true Seyfert galaxies are relatively weak radio sources. However, there are compact radio galaxies with an optical spectrum that is essentially the same as for Seyfert galaxies. These should probably be classified with the Seyfert galaxies. In general, the stronger radio emission seems to come with a type 2 spectrum.

It is estimated that about 1 % of all bright spiral galaxies are Seyfert galaxies. The luminosities of their nuclei are about 10^{36}–10^{39} W, of the

same order as all the rest of the galaxy. Brightness variations are common.

Radio Galaxies By definition, radio galaxies are galaxies that are powerful radio sources. The radio emission of a radio galaxy is non-thermal synchrotron radiation. The radio luminosity of radio galaxies is typically 10^{33}–10^{38} W, and may thus be as large as the total luminosity of a normal galaxy. The main problem in explaining radio emission is to understand how the electrons and magnetic fields are produced, and above all, where the electrons get their energy.

The forms and sizes of the radio emitting regions of radio galaxies have been studied ever since the 1950's, when radio interferometers achieved the resolution of optical telescopes. The characteristic feature of a strong radio galaxy is a double structure: there are two large radio emitting regions on opposite sides of the observed galaxy. The radio emitting regions of some radio galaxies are as far apart as 6 Mpc, almost ten times the distance between the Milky Way and Andromeda galaxies. One of the smallest double radio sources is the galaxy M87 (Fig. 19.26), whose two components are only a few kpc distant from each other.

The double structure of radio galaxies appears to be produced by ejections from the nucleus. However, the electrons in the radio lobes cannot be coming from the centre of the galaxy, because they would lose all their energy during such a long transit. Therefore electrons have to be continuously accelerated within the radio-emitting regions. Within the radio lobes there are almost point-like regions, hot spots. These are generally symmetrically placed with respect to the nucleus, and are apparently consequences of nuclear ejections.

"Tailed" radio sources also exist. Their radio emission mainly comes from one side of the galaxy, forming a curved tail, which is often tens of times longer than the diameter of the galaxy. The best examples are NGC 1265 in the Perseus cluster of galaxies and 3C129, which appears to be in an elliptical orbit around a companion galaxy. The tail is interpreted as the trail left by the radio galaxy in intergalactic space.

Another special feature revealed by the radio maps is the presence of *jets*, narrow lines of radio emission, usually starting in the nucleus and stretching far outside the galaxy. The best known may be the M87 jet, which has also been observed as an optical and X-ray jet. The optically observed jet is surrounded by a radio source. A similar radio source is seen on the opposite side of the nucleus, where no optical jet is seen. Our nearest radio galaxy Centaurus A also has a jet extending from the nucleus to near the edge of the galaxy.

VLBI observations of radio jets have also revealed *superluminal motions*: in many compact sources the components appear to be separating faster than the speed of light. Since such velocities are impossible according to the theory of relativity, the observed velocities can only be apparent, and several models have been proposed to account for them.

Quasars The first quasar was discovered in 1963, when *Maarten Schmidt* interpreted the optical emission lines of the known radio source 3C273 as hydrogen Balmer lines redshifted by 16 %. Such large redshifts are the most remarkable characteristics of the quasars. Properly speaking, the word quasar is an abbreviation for quasistellar radio source, and some astronomers prefer to use the designation QSO (quasistellar object), since not all quasars emit radio radiation.

Optically the quasars appear almost as point sources, although improved observational techniques have revealed an increasing number of quasars located inside more or less normal galaxies (Fig. 19.27). Although the first quasars where discovered by radio observations, only a small fraction of all optically identified quasars are bright radio sources. Most radio quasars are point sources, but some have a double structure like the radio galaxies. Satellite X-ray pictures also show the quasars to be pointlike.

In the visible region the quasar spectra are dominated by spectral lines with rest wavelengths in the ultraviolet. The first observed quasar redshifts were $z = 0.16$ and 0.37, and later searches have continued to turn up ever larger redshifts.

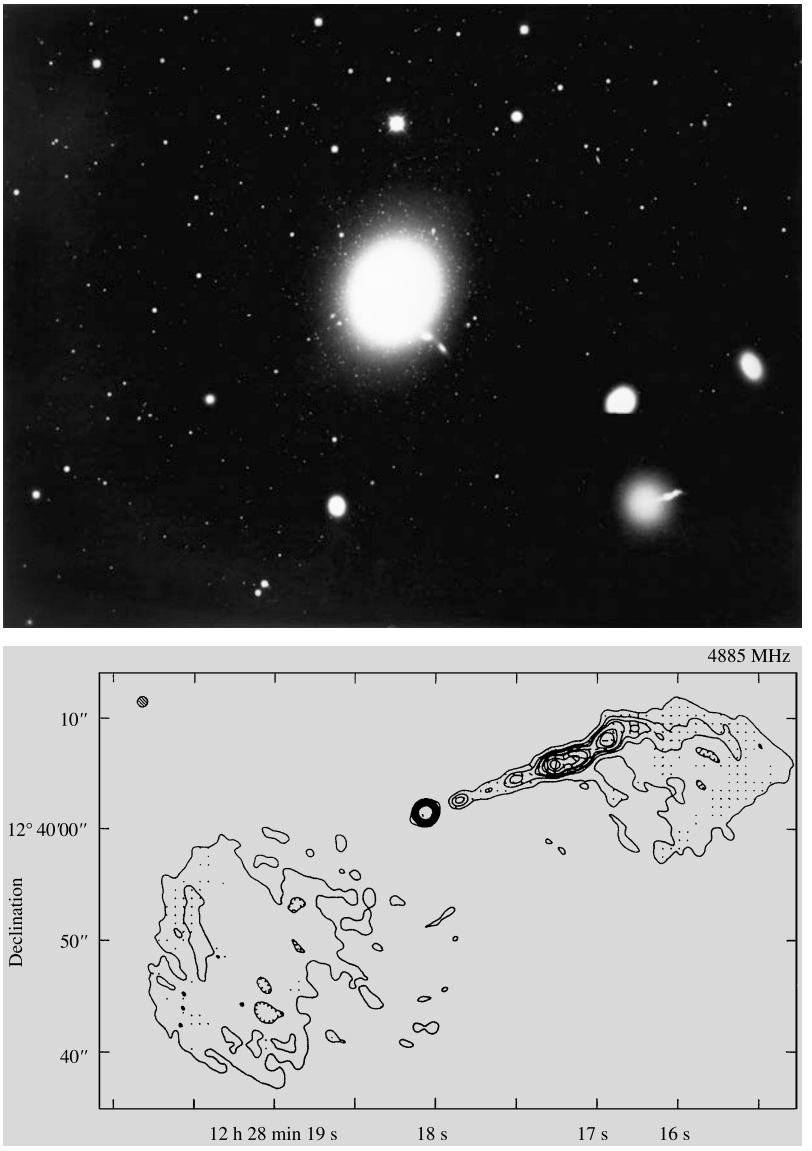


Fig. 19.26 *Above*: The active galaxy M87. In the lower right-hand corner a short exposure of the core region has been inserted (same scale as in the main photograph). One sees a blue jet coming out of the nucleus of a normal E0 galaxy (NOAO/Kitt Peak National Observatory). *Below*: In the radio map made using the VLA the jet is observed to be two-sided. The area shown is much smaller than in the upper picture. (Owen, F.N., Hardee, P.E., Bignell, R.C. (1980): Astrophys. J. (Lett.) **239**, L11)

The present record is 8.7, but three objects have redshift 10. If these are quasars their light must have been emitted when the age of the universe was only 500 million years. The large inferred distances of the quasars mean that their luminosities have to be extremely large. Typical values lie in the range of 10^{38}–10^{41} W. The brightness of quasars may vary rapidly, within a few days or less. Thus the emitting region can be no larger than a few light-days, i.e. about 100 au.

The quasars often have both emission and absorption lines in their spectra. The emission lines are very broad and are probably produced in the quasar itself. Much of the absorption spec-

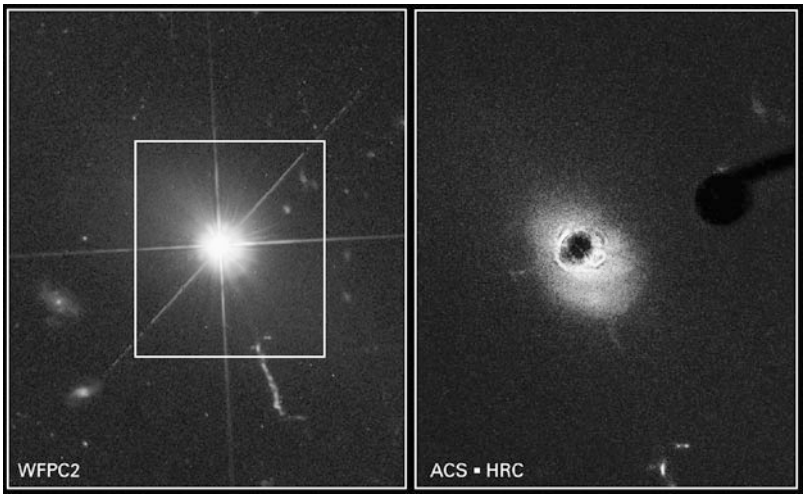


Fig. 19.27 One of the nearest quasars, 3C 273, photographed with two cameras aboard the Hubble Space Telescope. *On the left*, the Wide Field Planetary Camera sees a bright point-like source, with a jet blasted out from the quasar (towards 5 o'clock). *On the right*, a coronagraph in the Advanced Camera for Surveys blocks out the brightest parts of the quasar. Spiral arms in the host galaxy can be seen, with dark dust lanes, as well as new details in the path of the jet. (Photos Hubble/NASA/ESA)

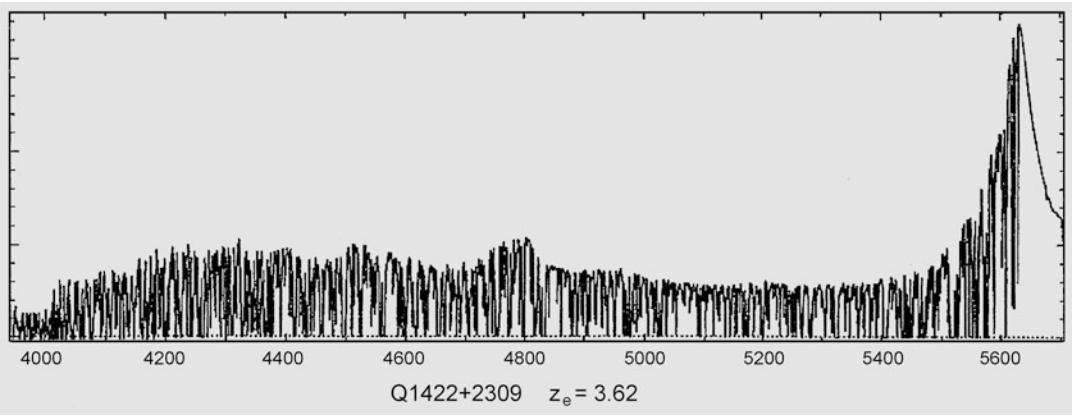


Fig. 19.28 A section of the spectrum of the quasar Q1422+2309 observed with the Keck 10 metre telescope. The Lyman α line of the quasar is near the right edge. Its redshift is 3.62. To the left of the line is a Lyman α forest, where each absorption line corresponds to a hydrogen cloud that the radiation had to penetrate. Wavelengths are given in ångströms. (Kuva Womble, Sargent, Lyons)

trum consists of densely distributed narrow lines that are thought to be hydrogen Lyman α lines formed in gas clouds along the line of sight to the quasar. The clouds producing this *"Lyman α forest"* (Fig. 19.28) are young galaxies or protogalaxies, and they therefore provide important evidence about the formation of galaxies. Lyman α forests can be used to map intergalactic gas clouds up to redshift $z = 6$ (Fig. 19.29). Further out Lyman α absorption lines cannot be seen since most of the hydrogen is neutral.

Unified Models Although the forms of galactic activity may at first sight appear diverse, they can be unified within a fairly widely accepted schematic model in such a way that the properties of each individual object depend on a few parameters only.

According to this model, most galaxies contain a compact central nucleus, which is a supermassive black hole, with mass 10^7–$10^{10}\,M_\odot$, surrounded by a disk or ring of gas. The source of energy is the gravitational energy released as gas is accreted into the black hole. The disk may also give rise to a jet, where some of the energy is converted into perpendicular motions along the rotational axis. Thus active galactic nuclei are similar to the nucleus of the Milky Way, although the masses of both the black hole and the gas disk may be much larger.

Deducing the mass of the central black hole is difficult and uncertain. However, using a variety of methods involving the motions of stars and gas at the centre of nearby galaxies black hole masses for about 30 galaxies have been determined. The most important result of these studies is that there is a close relationship between the black hole mass and the central velocity dispersion of the galaxy. According to the virial theorem the velocity dispersion is a measure of the bulge mass, and therefore the conclusion is that there is a close relationship between the mass of the bulge and the mass of the central black hole.

The first characteristic parameter of the unified model is obviously the total luminosity. For example, the only essential difference between Seyfert 1 galaxies and radio-quiet quasars is the larger luminosity of quasars. Another basic parameter is the radio brightness, which may be related to the strength of a jet. On the basis of their radio luminosity one can connect Seyfert galaxies and radio-quiet quasars on one hand, and radio galaxies and radio quasars on the other.

The third important parameter of unified models is the angle from which we happen to view the nuclear disk. For example, if the disk is seen edge-on, the actual nucleus is obscured by the disk. This could explain the difference between Seyfert types 1 and 2: in type 2 we do not see the broad emission lines formed near the black hole, but only the narrower lines from the disk. Similarly a galaxy that looks like a double radio source when seen edge-on, would look like a radio quasar if the disk were seen face-on. In the latter case there is a possibility that we may be seeing an object directly along the jet. It will then appear as a *blazar*, an object with rapid and violent variations in brightness and polarisation, and very weak or invisible emission lines. If the jet is almost relativistic, its transverse velocity may appear larger than the speed of light, and thus superluminal motions can also be understood.

One prediction of the unified model is that there should be a large number of quasars where the nucleus is obscured by the disk as in the Seyfert 2 galaxies. In analogy with the Seyfert galaxies these objects are referred to as type 2 AGN or quasars. Because of the obscuration such sources would not be included in surveys at optical, UV, or soft X-ray wavelengths. In hard X-rays the obscuration is less, and in the far infrared the absorbed energy is re-radiated. Searches for type 2 quasars have been made at these wavelengths with the Chandra X-ray satellite and with the Spitzer Space Telescope. The indications from these searches are that at least 3/4 of all supermassive black holes are heavily obscured.

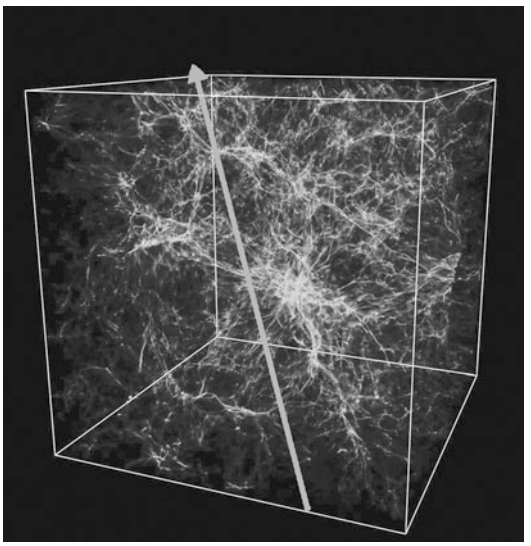


Fig. 19.29 Simulation of hydrogen distribution in space. The light emitted by a quasar goes through different gas clouds creating a Lyman α forest. (Renyue Cen, Princeton University)

Gravitational Lenses An interesting phenomenon first discovered in connection with quasars are gravitational lenses. Since light rays are bent by gravitational fields, a mass (e.g. a galaxy)

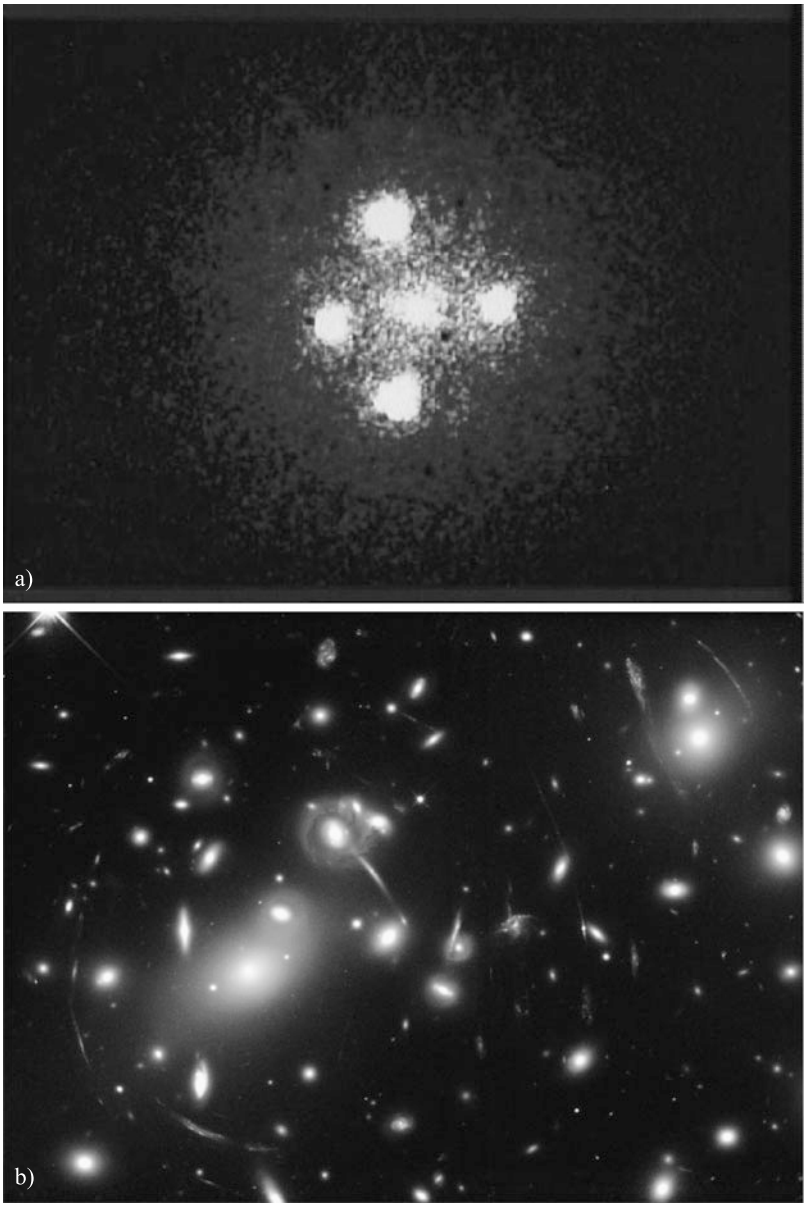


Fig. 19.30 (**a**) The components of the Einstein cross are gravitationally lensed images of the same quasar (ESA/NASA). (**b**) The massive galaxy cluster Abell 2218 deflects light rays passing through it and acts as a giant gravitational lens. In this Hubble picture from January 2000, dozens of arc-shaped images of distant galaxies can be seen. (Photo A. Fruchter, S. Baggett, R. Hook and Z. Levay, NASA/STScI)

placed between a distant quasar and the observer will distort the image of the quasar. The first example of this effect was discovered in 1979, when it was found that two quasars, 5.7″ apart in the sky, had essentially identical spectra. It was concluded that the "pair" was really a double image of a single quasar. Since then several other gravitationally lensed quasars have been discovered (Fig. 19.30).

Gravitational lenses have also been discovered in clusters of galaxies. Here the gravitational field of the cluster distorts the images of distant galaxies into arcs around the cluster centre. Currently some two hundred individual galaxies and twenty galaxy clusters are known to produce lensed images of background galaxies or quasars. Such systems are called *strong gravitational lenses* (Fig. 19.31).

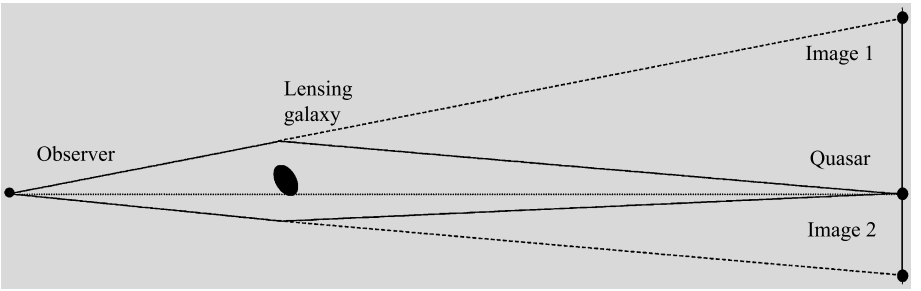


Fig. 19.31 A strong gravitational lens. The light of a quasar is bent by the gravity of a lensing galaxy. The quasar appears as two images on both sides of the lensing galaxy

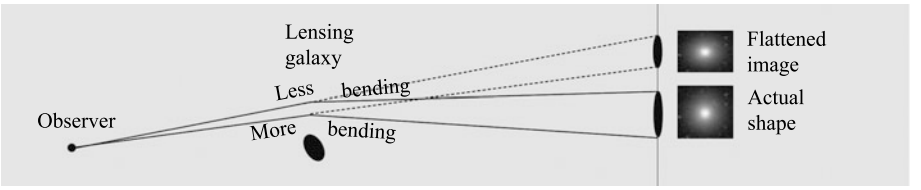


Fig. 19.32 A weak gravitational lens. Light rays coming from the upper and lower edges of a distant galaxy are bent by different amounts. The image of the galaxy becomes slightly flattened

Recently emphasis has moved towards *weak gravitational lenses*, which are much more common (Fig. 19.32). The lens does not produce separate images or arcs from a background but only flattens its shape. Since the axial ratio changes only by 1–2 %, the weak lenses cannot be used to study shapes of individual galaxies because their original shapes are not known. However, if hundreds or thousands of galaxies around a mass concentration are measured, statistical analysis shows the flattening. In the 2010's weak lenses are routinely used to observe the distribution of the invisible matters.

Figure 19.13 showed an example about dark matter clouds around a galaxy cluster.

19.8 The Origin and Evolution of Galaxies

Because the speed of light is finite we see distant galaxies at earlier stages in their life. In the next chapter we show how the age of a galaxy with given redshift can be calculated based on the rate of expansion of the Universe. However, this relationship will depend on the cosmological model, which therefore always has to be specified when studying the evolution of galaxies.

Evolution of galaxies began when the first stars were born at redshifts 10–20. The universe consisted of neutral hydrogen and helium only, and the temperature of the gas had dropped below 100 kelvins. The first stars were very massive and shortlived. Radiation from the first stars, galaxies and quasars rapidly ionised the interstellar and intergalactic gas. This era is called the time of ionisation. Ionisation was almost complete when the redshift reached $z = 6$, and ever since 99.9 % of the hydrogen has been ionised. Towards the end of the reionisation, at redshifts 6–15, massive galaxies ($M > 10^{10} M_{\odot}$) were already born. Current observations of galaxies go back to redshifts 10–12.

According to the currently widely accepted cosmological models most of the matter in the Universe is in a form that emits no radiation, and is only observable from its gravitational effects. In this *Cold Dark Matter* (CDM) theory (see Sect. 20.7) the first systems to collapse and start forming stars were small, with masses like those of dwarf galaxies. Larger galaxies were formed later as these smaller fragments collected

into larger clumps. This model, where most stars are formed in small galaxies is usually described as the *hierarchical* model.

A large part of our theoretical ideas on galaxy evolution is based on numerical simulations of the collapse of gas clouds and star formation in them. Using some prescription for star formation one can try to compute the evolution of the spectral energy distribution and the chemical abundances in the resulting galaxies. The results of the models can be compared with the observational data presented in the previous sections of this chapter.

The density distribution of dark matter is expected to be very irregular, containing numerous small-scale clumps. The collapse will therefore be highly inhomogeneous, both in the hierarchical and in the monolithic picture, and subsequent mergers between smaller systems should be common. There are additional complicating factors. Gas may be expelled from the galaxy, or there may be an influx of fresh gas. Interactions with the surroundings may radically alter the course of evolution—in dense systems they may lead to the complete merging of the individual galaxies into one giant elliptical. Much remains to be learned about how the formation of stars is affected by the general dynamical state of the galaxy and of how an active nucleus may influence the formation process.

Our observational knowledge of galaxy evolution is advancing very rapidly. Essentially all the relationships described earlier in this chapter have been studied as functions of time. Still, a complete generally accepted description of the way the Universe reached its present state has not yet been established. Here we can only mention a few of the most central aspects of the processes leading to the galaxies we observe to-day.

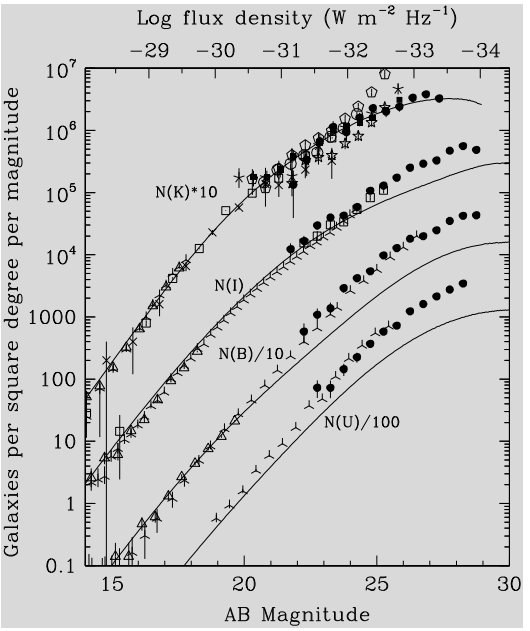


Fig. 19.33 Galaxy counts in the U, B, I, and K wavelength bands. The counts are compared to those in a cosmological model without evolution. The cosmological parameters correspond to the currently preferred "concordance" model (see Sect. 19.5). (H.C. Ferguson et al., 2000, ARAA 38, 667, Fig. 4)

Fig. 19.34 The deepest image taken by the Hubble space telescope and published in 2014. The total exposure time is 27 days. The original picture contains about 15,000 faint galaxies. (NASA, ESA, H. Teplitz, M. Rafelski, A. Koekemoer, R. Windhorst and Z. Levay)

Density and Luminosity Evolution The most basic way of studying the formation and evolution of galaxies is by counting their numbers, either the number brighter than some given magnitude limit (as was already done by Hubble in the 1930's, see Sect. 20.1) or else the number density as a function of redshift. The counts can be compared with the numbers expected if there is no evolution, which depend on the cosmological model. They can therefore be used either as a cos-

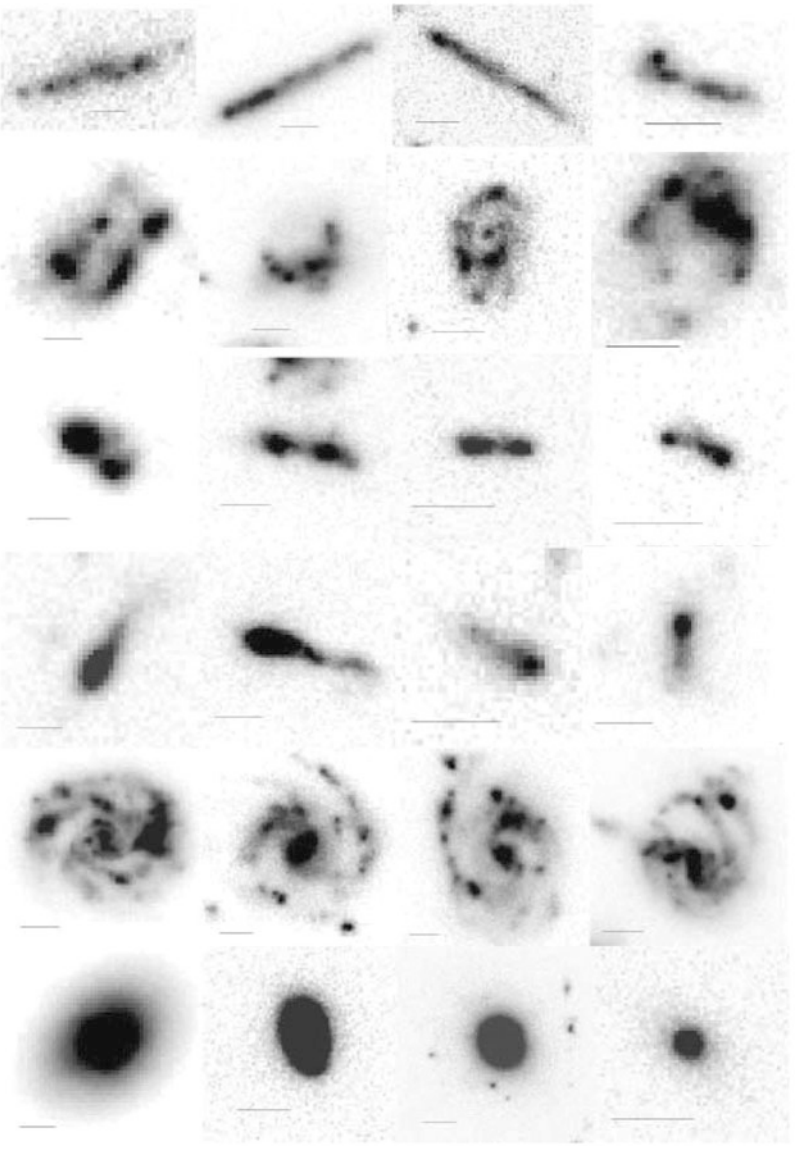


Fig. 19.35 Different early forms of galaxies in the redshift range $z = 2$–4. *Top line*: chain galaxies; *second line*: clump-clusters; *third line*: double galaxies; *fourth line*: tadpoles; *fifth line*: spiral galaxies being born; *bottom line*: elliptic galaxies. (Elmegreen, Elmegreen, Rubin, Schaffer, 2005, ApJ 631, p. 87)

mological test, or as a test of evolutionary models. However, the present situation is that more reliable cosmological tests are available, and the number counts are mainly used to study the evolution of galaxies.

There are two ways in which the number counts are affected by galaxy evolution. In *density evolution* the actual number of galaxies is changing, whereas in *luminosity evolution* only the luminosity of individual galaxies is evolving. The simplest form of luminosity evolution is called *passive luminosity evolution*, and is due to the changing luminosity of stars during normal stellar evolution. Pure luminosity evolution is expected to be predominant in the monolithic picture. In the hierarchical picture density evolution will be more prominent, since in this picture smaller galaxies are to a greater extent being destroyed to produce larger more luminous ones.

Figure 19.33 gives an example of the results of number counts. A model without evolution cannot explain these observed counts, and various

models incorporating evolutionary effects have to be introduced. However, a unique model cannot be determined using just the number counts.

Distant Galaxies A more direct approach to galaxy formation is the direct search for the most distant objects visible. In Fig. 19.34 we show the *Hubble Ultra Deep Field* (HUDF), an area of the sky observed with the Hubble Space Telescope. The total exposure time through different filters was 27 days. The image contains over 15,000 galaxies. Over 5000 are seen as dots; their redshift is aver 6. The other ones, seen as extended objects, give information on galaxy evolution at redshifts 1–5.

The most important observation is that with increasing redshift the Hubble classification breaks down and finally disappears. Around redshift $z = 3$ the spiral galaxies resembling the current ones have disappeared almost completely. Instead, there are different irregular and strange galaxies.

Figure 19.35 shows examples of distant galaxies in the range $z = 2$–4. Over 10 % are "chain galaxies", about 20 % "clump clusters", 15 % "doubles" and 10 % "tadpoles". About 30 % are evolving spirals and 10 % ellipticals.

Another important observation is that in the past galaxies were smaller than nowadays. Even if the mass or luminosity of a galaxy were similar to a modern galaxy its diameter at $z = 3$ is only one tenth. The most probable explanation is that the outer parts grow in several small mergers. Collision of a dwarf galaxy does not considerably increase the mass of the great galaxy but brings more stars to its outskirts increasing its diameter. The observation clearly indicates how important mutual collisions have been in the evolution of galaxies towards their current forms.

Evolution of AGN The first clear indication of cosmic evolution was in the numbers of radio galaxies and quasars. Already in the late 1960's it was becoming clear that the density of quasars was increasing dramatically towards higher redshifts (Fig. 19.36). Roughly, the number density of quasars increases relative to the present density by a factor 100 out to a broad maximum at

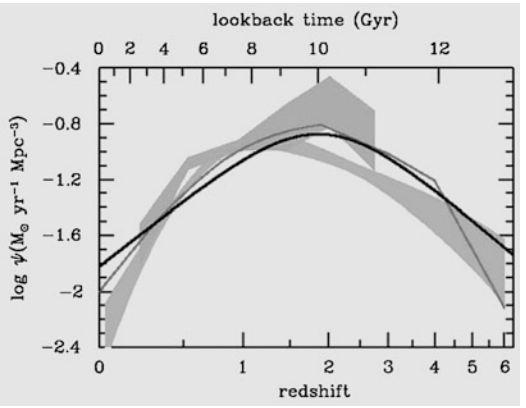


Fig. 19.36 Quasar activity at different redshifts according to X-ray and infrared observations (*gray areas* and *the gray line*). *The black line* corresponds to the stellar birth rate in Fig. 19.37. (Madau ja Dickinson, 2014, Ann. Rev. Astron. Astrophys. 52)

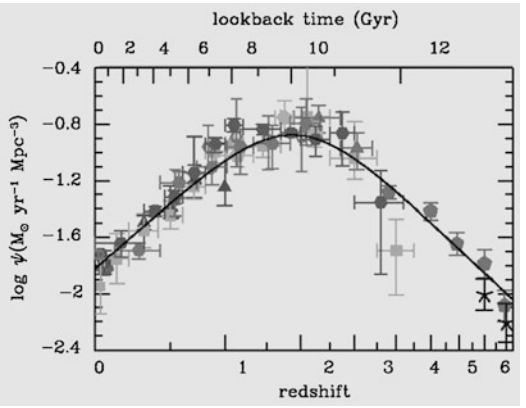


Fig. 19.37 Stellar birth rates at different redshifts. *Different symbols* correspond to values obtained in different studies and their standard errors. (Madau and Dickinson, 2014, Ann. Rev. Astron. Astrophys. 52)

redshift about 2. The observed behaviour may be due to either density or luminosity evolution. The density of radio galaxies also has a maximum at redshifts about 1.5–3, which is sometimes referred to as the *quasar era*.

The Star Formation History of the Universe Since the Universe contained only neutral gas as the first stars began to form, the most general description of how galaxies came to be is in terms of the rate at which the gas is being turned into stars. The star formation history going back to

a redshift about 8 is shown in Fig. 19.37. The star formation rate was about an order of magnitude larger than its current value at redshift 1–2. For even larger redshifts it seems to have remained nearly constant or decreased slowly.

19.9 Exercises

Exercise 19.1 The galaxy NGC 772 is an Sb spiral, similar to M31. Its angular diameter is $7'$ and apparent magnitude 12.0. The corresponding values of M31 are 3.0° and 5.0. Find the ratio of the distances of the galaxies

(a) assuming their sizes are equal,
(b) assuming they are equally bright.

Exercise 19.2 The brightness of the quasar 3C279 has shown changes with the time scale of one week. Estimate the size of the region producing the radiation. The apparent magnitude is 18. If the distance of the quasar is 2000 Mpc, what is its absolute magnitude and luminosity? How much energy is produced per au^3?