

**4.1.1 The Abell Catalogues of Rich Clusters of Galaxies**

It is worthwhile recounting the heroic efforts of Abell and his colleagues in creating the Abell catalogue of rich clusters of galaxies and its southern counterpart. The 48-inch Schmidt telescope was constructed on Palomar mountain during the late 1940s as a wide-field telescope to support and complement observations made with the 200-inch telescope. The Palomar Sky Survey took seven years to complete and comprised 879 pairs of 14-inch plates, each providing roughly  $6^\circ \times 6^\circ$  sky coverage and having limiting magnitudes of 21.1 in the blue and 20.0 in the red wavebands. Abell was one of the principal observers for the Palomar Sky Survey and, while the plates were being taken, he systematically catalogued the rich clusters of galaxies appearing on the plates. The word ‘rich’ meant that there was no doubt as to the reality of the associations of galaxies appearing on the plates (Abell, 1958). A typical example of a rich, regular cluster of galaxies, Abell 2218, is shown in Fig. 4.1. A corresponding catalogue for the southern hemisphere was created with the completion of the ESO-SERC Southern Sky Survey, which was made with the 48-inch UK Schmidt Telescope at the Siding Spring Observatory in New South Wales (Abell et al., 1989). In both cases, the clusters were found by visual inspection of the Sky Survey plates.

Crucial to the success of Abell’s programme was adherence to the strict set of criteria he established for the inclusion of clusters in the catalogue. The 40/3 clusters in the combined northern and southern catalogue of Abell, Corwin and Olowin fulfil the following selection criteria:

- *Richness criterion.* The clusters must have 50 members brighter than 2 magnitudes ( $m_3 + 2$ ) fainter than the third brightest member ( $m_3$ ). Richness classes are defined by the number of galaxies with magnitudes between  $m_3$  and  $m_3 + 2$ ,



**Fig. 4.1.** The rich, regular cluster of galaxies Abell 2218. Abell classified this cluster as richness class 4. The central galaxy is a supergiant or CD galaxy, which is very much brighter than all the other galaxies in the cluster. It is located close to the dynamical centre of the cluster. The image also shows a number of arcs which are the gravitationally lensed images of very distant background galaxies (Courtesy NASA, ESA and the Space Telescope Science Institute)

as described in Table 4.1. A richness class 0 is included for clusters which have between 30 and 50 members within the same magnitude interval, but the catalogues are not complete for this class. Studies of a number of nearby clusters of different richness classes have shown that richness is proportional to the total number of galaxies in the cluster.

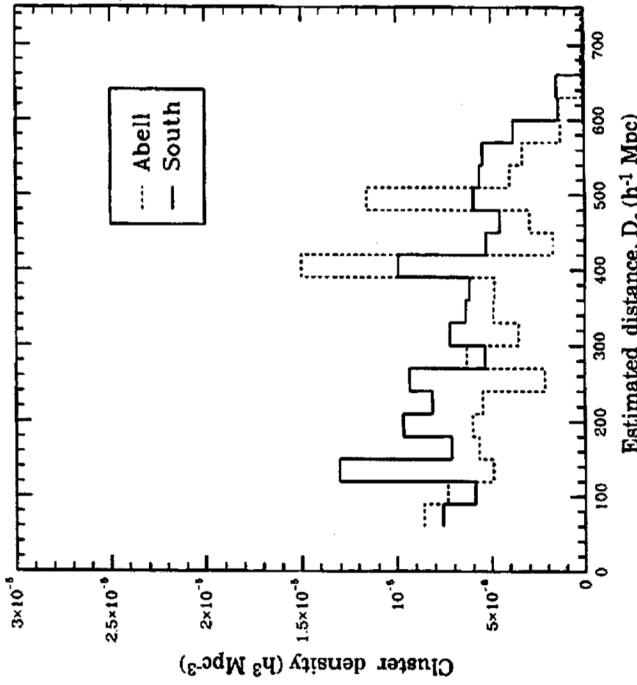
- *Compactness criterion.* Galaxies are only counted within a radius of  $1.5 h^{-1}$  Mpc of the cluster centre. This radius corresponds to an angular radius of  $1.7/2$  arcmin, the redshift of the cluster being estimated from the apparent magnitude of its tenth brightest member,  $m_{10}$ .
- *Distance criteria.* Abell clusters with redshifts less than 0.02 span more than one Sky Survey plate and hence this lower redshift limit was adopted. The upper redshift limit is set by the fact that galaxies could not be counted to magnitudes fainter than  $m_r = 20$  in the northern survey and therefore the third brightest galaxy must be brighter than  $m_3 = 17.5$ . This redshift limit corresponds to  $z \approx 0.2$ . From redshift measurements of a number of clusters, it was found that  $m_{10}$ , the apparent magnitude of the tenth brightest cluster member, is a reliable distance indicator. The clusters were then divided into *distance classes* in such a way that there was a small probability of the class assigned to the cluster being more than one class wrong.

Within the northern sample, Abell defined a *complete statistical sample* of 1682 clusters which fulfilled distance criteria 1 to 6 and richness criteria 1 to 5. The numbers of clusters in each richness class in this sample and the total numbers in the southern sample are included in Table 4.1. The number density distribution of clusters with increasing distance, equivalent to the number counts of clusters, is shown in Fig. 4.2 (Batuski et al., 1989). It can be seen that serious incompleteness sets in at distances greater than about  $600 h^{-1}$  Mpc, corresponding to redshift  $z = 0.2$ , but at smaller distances, there is reasonable agreement between the samples of clusters

**Table 4.1.** Definitions of the richness classes  $R$  of Abell clusters and the numbers of clusters within Abell’s complete northern sample of 1682 clusters and the total sample of southern clusters.  $N$  is the number of galaxies in the cluster between magnitudes  $m_3$  and  $m_3 + 2$  (Abell, 1958; Bahcall, 1988; Abell et al., 1989)

Richness Class $R$	$N$	Number of clusters in the complete northern sample	Total number of clusters in the southern sample
(0) <sup>a</sup>	(30 – 49) ( $\geq 10^3$ )	(664)	
1	50 – 79	1224	636
2	80 – 129	383	273
3	130 – 199	68	41
4	200 – 299	6	1
5	300 or more	1	0

<sup>a</sup> The sample is not complete for richness class zero



**Fig. 4.2.** The average estimated space densities of Abell clusters in the Northern and Southern catalogues for richness classes  $R \geq 1$  as a function of estimated distance  $D_c$  (Bautista et al., 1989).

in the northern and southern hemispheres. The space density of Abell clusters with richness classes greater than or equal to 1 is, to a good approximation,

$$N_{\text{el}}(R \geq 1) \approx 10^{-5} h^3 \text{ Mpc}^{-3}, \quad (4.1)$$

so that the typical distance between cluster centres, if they were uniformly distributed in space, would be  $\sim 50 h^{-1} \text{ Mpc}$ . These figures can be compared with the space density of ‘mean galaxies’ of  $10^{-2} h^3 \text{ Mpc}^{-3}$  and their typical separations of  $5 h^{-1} \text{ Mpc}$  (see Sect. 3.7.2).

#### 4.1.2 Comparison with Clusters Selected from the Sloan Digital Sky Survey

The studies of Abell and his colleagues were entirely based upon visual inspection of the plates and film copies of the Northern and Southern 48-inch Schmidt

Telescope Sky Surveys. Great care was taken to ensure the consistency of the cluster identifications and classifications over many years. The advent of high speed measuring machines with powerful procedures for star-galaxy separation enabled more objective approaches to the identification of rich clusters to be developed. It is perhaps most instructive to compare Abell’s results with those of the Bahcall and her colleagues who have identified the clusters present in the early release commissioning data of the Sloan Digital Sky Survey (SDSS) (Bahcall et al., 2003b). The SDSS has the great advantages of having images in five different wavebands as well as redshifts for very large samples of galaxies at  $z \leq 0.1$ . As a result, velocity dispersions are available for many of the clusters studied, enabling masses to be estimated.

The analyses described by Bahcall and her colleagues do not use all these data but employ two of the various algorithms available to isolate clusters of galaxies which satisfy criteria similar to Abell’s, but which extend to slightly lower richness classes. They find that 53 of the 58 Abell clusters in the area of sky studied are included in their merged catalogue of clusters, the remaining five clusters being detected just below the richness limits of their catalogue. This is a remarkable testament to the quality of the visual approach of Abell and his colleagues.

An advantage of the procedures adopted by Bahcall and her colleagues is that the limits of the catalogue can be expressed in terms of the velocity dispersions and masses of the clusters. Thus, the threshold for the inclusion of clusters in the catalogue corresponds to typical velocity dispersions  $\sigma_v \geq 400 \text{ km s}^{-1}$  and to masses  $M$  within radius  $0.6 h^{-1} \text{ Mpc}$  of their centres of  $M \geq 5 \times 10^{13} h^{-1} M_\odot$ . This threshold corresponds to richness classes below Abell’s richness class 0. The average space density of these rich clusters is  $2 \times 10^{-5} h^3 \text{ Mpc}^{-3}$ , in excellent agreement with Abell’s estimate when account is taken of the fact that the sample extends to somewhat lower richness limits.

Bahcall and her colleagues provide a useful table showing the velocity dispersions, masses and luminosities for the samples of clusters of different richnesses in their catalogue (Table 4.2). These data provide estimates of the mass-to-luminosity ratios for clusters of different richnesses. It can be seen from Table 4.2 that these range from  $\sim 170 h$  to  $\sim 250 h$  with increasing richness. In turn, these data can be used to determine the mass functions of rich clusters of galaxies and to provide constraints on the mass density parameter  $\Omega_0$  and the amplitude of the mass fluctuations  $\sigma_8$  (Sects. 8.7 and 14.4).

#### 4.1.3 Abell Clusters and the Large-Scale Distribution of Galaxies

Abell clusters are strongly correlated in space, both with each other and with the distribution of galaxies in general. These associations were originally described in terms of the superclustering of galaxies, but it is preferable nowadays to express the clustering in terms of the correlation functions introduced in Sect. 2.2.1. It is simplest to quote the results of the SDSS survey of rich clusters by Bahcall and her colleagues who include extensive references to earlier work (Bahcall et al., 2003a).

**Table 4.2.** An example of the scaling relations for clusters of galaxies of different richnesses from analysis of the commissioning data from the SDSS (Bahcall et al., 2003b).  $N_{\text{gal}}$  is the number of red E/S0 galaxies one magnitude fainter than  $L^*$  in the  $i$  band within a radius of  $1h^{-1}$  Mpc of the brightest cluster galaxy where  $L^* = 10^{10} h^{-2} L_\odot$ .  $A$  is a measure of the richness of the cluster and is derived from the best-fit cluster model which has total luminosity  $L_{\text{cl}} = AL^*$ , where the luminosity is measured within a radius of  $1h^{-1}$  Mpc

$N_{\text{gal}}$	$A$	$\sigma_v$ ( $\text{km s}^{-1}$ )	$L_{\text{tot}}$ ( $10^{11} h^{-2} L_\odot$ )	$M$ ( $10^{14} h^{-1} M_\odot$ )
10	35	350	2.1	0.35
15	43	430	3.2	0.6
20	50	500	4.4	0.9
25	56	560	5.5	1.2
30	61	620	6.6	1.5
40	70	730	8.7	2.2

The correlation functions for the different richness classes described in Table 4.2 are shown in Fig. 4.3. The best-fit two-point correlation function has the form

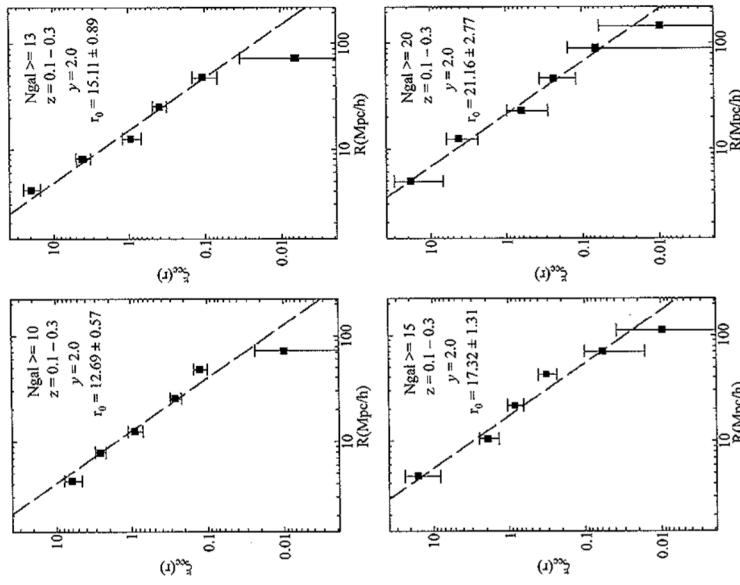
$$n(r) dV = n_0 [1 + \xi(r)] dV, \quad \text{where} \quad \xi(r) = \left(\frac{r}{r_0}\right)^{-2}, \quad (4.2)$$

which is slightly steeper than the value found on smaller physical scales for galaxies in general. The data included in Fig. 4.3 show that the correlation length  $r_0$  is much greater than that for galaxies in general. Bahcall and her colleagues find that the variation can be described by the relation  $r_0 = 6\sqrt{d}$  for  $(20 \leq d \leq 90)h^{-1}$  Mpc, where  $d$  is the mean comoving separation distance of the clusters belonging to different richness classes. Thus, the correlation lengths for these rich clusters are large, on the scale of the huge voids seen in the distribution of galaxies in general.

The rich clusters are generally found in the densest regions in the ‘cosmic web’ seen in Figs. 2.7 and 2.8. Some measure of the association of Abell clusters with galaxies in general has been provided by the analysis of Seldner and Peebles (Seldner and Peebles, 1977). They cross-correlated the counts of galaxies in the Shane-Wirtanen catalogue, which extended to apparent magnitude  $m \approx 19$ , with the positions of Abell clusters of distance class 5. The cross-correlation function was found to be of the form

$$\xi_{\text{gc}} = Ar^{-2.4},$$

out at a distance of  $15h^{-1}$  Mpc from the cluster centres. This function represents the correlation of the Abell clusters with galaxies in the vicinity of the cluster, but outside what would normally be considered to be the cluster boundary; Seldner and Peebles refer to this phenomenon as the superclustering of galaxies about Abell clusters. This result is in agreement with the visual impression of the distribution of galaxies seen in Figs. 2.7 and 2.8.



**Fig. 4.3.** The two-point spatial correlation function for four richness thresholds ( $N_{\text{gal}} \geq 10, \geq 13, \geq 15, \geq 20$ ) for clusters selected from the SDSS (Bahcall et al., 2003a). Best-fit functions with slope 2 and correlation-scale  $r_0$  are shown by the dashed lines. The error bars show the  $1\sigma$  uncertainties in the estimates. The values of  $d$  for the four panels are 26.2, 35.6, 41.5 and 58.1 Mpc with increasing richness

## 4.2 The Distribution of Galaxies in Clusters of Galaxies

Clusters of galaxies come in a variety of shapes and forms and various schemes have been developed to put some order into this diversity. Just as in the case of galaxies, modern computer-based systems of classification bring new, quantitative insights into the wealth of detail contained in the visual classification of clusters.

### 4.2.1 The Galaxy Content and Spatial Distribution of Galaxies in Clusters

Abell classified clusters as *regular* if they are more or less circularly symmetrical with a central concentration, similar in structure to globular clusters (Abell, 1962). These

are among the richest clusters, generally containing over 1000 members. Examples include the Coma cluster. The galaxies are predominantly elliptical and S0 galaxies as has been indicated by the results shown in Figs. 3.4 and 3.16. Abell called all the other clusters in which there is much less well-defined structure *irregular* clusters. In his words, the irregular clusters

- ... have a more nearly amorphous appearance, and possess little or no spherical symmetry or central concentration. The irregular clusters sometimes, however, have several small subcondensations or nuclei and resemble loose swarms of small clusters.

Examples of the latter include the Hercules and Virgo clusters.

Oemler studied systematically a representative sample of fifteen Abell clusters of different richness classes and morphological types and distinguished three principal types of cluster according to their *galaxy content* (Oemler, 1974):

- *cD clusters* have a unique and dominant cD galaxy and the ratio of elliptical (E) to lenticular (S0) to spiral (S) galaxies is roughly 3 : 4 : 2, that is, only about 20% are spirals. These clusters are very rich in E and S0 galaxies.
- *Spiral-rich clusters* have galaxy-type ratios E : S0 : S roughly 1 : 2 : 3, that is, about 50% spirals, a distribution similar to the proportions of types found in the general field.
- The remaining clusters are called *spiral-poor clusters*. They have no dominant cD galaxy and have galaxy type ratios E : S0 : S roughly 1 : 2 : 1.

Abell noted that there is a correlation between the structure of clusters and galaxy content and this result was quantified by Oemler who established the following relations:

- In *cD clusters* or *regular clusters*, the spatial distribution of galaxies resembles the distribution of stars in a globular cluster (see Sect. 4.2.2). The space density of galaxies increases rapidly towards the centre of the cluster. The *spiral-rich clusters* and *irregular clusters* are not symmetric and there is little central concentration; the spatial density of galaxies is roughly uniform towards the central regions and is lower than that in cD clusters. The *spiral-poor clusters* are intermediate between these two extremes.
- In the case of *spiral-rich clusters*, the radial distribution of elliptical, lenticular and spiral galaxies is more or less the same. In *cD* and *spiral-poor clusters*, however, the relative space density of spiral galaxies decreases markedly towards the central regions, reflecting Dressler's correlation of galaxy type with galaxy number density (Dressler, 1984). In these clusters, the spiral galaxies form a halo around a central core of elliptical and S0 galaxies. There is therefore *segregation by galaxy type* in cD and spiral-poor clusters.
- In addition to evidence for segregation by galaxy type, there is some evidence for *mass segregation* as a function of cluster type. Adopting apparent magnitude as a measure of mass, Oemler found that, in *cD* and *spiral-poor clusters*, the massive galaxies are located closer to the centre than less massive galaxies.

However, this mass segregation is only important for the brightest members of clusters,  $m \leq m_1 + 2$ , where  $m_1$  is the apparent magnitude of the brightest galaxy; galaxies fainter than  $m_1 + 2$  appear to be uniformly distributed throughout the clusters. No such mass segregation is found in *spiral-rich clusters*.

As discussed in Sects. 3.7.1 and 4.2.3, the cD galaxies found in the centres of cD clusters have a number of features which distinguish them from giant elliptical galaxies (Kormendy, 1982). Their most distinctive features are their extensive stellar envelopes which can be as large as 100 kpc. In addition to being found in rich clusters, examples of cD galaxies are known in small groups of galaxies but, in all cases, they are found in regions of significantly enhanced galaxy density relative to the general field. Dressler found that cD galaxies are only found in regions in which the galaxy density exceeds  $1 h^{-3}$  galaxy Mpc $^{-3}$ , compared with an average galaxy density of  $10^{-2} h^{-3}$  Mpc $^{-3}$  (Dressler, 1984). According to Dressler, the local galaxy density rather than the richness of the cluster determines whether or not a cD galaxy is present in a group or cluster. Some of the central cD galaxies in rich clusters are also distinctive in having multiple nuclei which are found in about 25–50% of all cD galaxies. In contrast, multiple nuclei are an order of magnitude less common in second and third ranked cluster members.

These results are important in understanding the dynamical evolution of clusters. Regular, cD clusters are systems which have had time to relax to dynamical equilibrium, whereas the other systems are still in the process of relaxation.

#### 4.2.2 Clusters of Galaxies and Isothermal Gas Spheres

In regular clusters, the space density of galaxies increases towards the central regions, which are referred to as the *cores* of the clusters. Outside the core, the space density of galaxies decreases steadily until it disappears into the background of unrelated objects. The regular structures of these clusters suggests that they have relaxed to a stationary dynamical state similar to that found in the distribution of stars in globular clusters. The spatial distribution of galaxies in a regular cluster can be modelled by the distribution of mass in an *isothermal gas sphere*. These distributions are important in the discussion which follows and so let us derive the relevant expressions for the density distribution of an isothermal gas sphere.

The term *isothermal* means that the temperature, or mean kinetic energy of the particles, is constant throughout the cluster. In physical terms, this means that the velocity distribution of the galaxies is Maxwellian with the same velocity dispersion (or temperature) throughout the cluster. If all the galaxies had the same mass, the velocity dispersion would be the same at all locations within the cluster. This is a rather sweeping approximation since it would mean that there had been enough time for the galaxies to have exchanged kinetic energy and come to a thermal equilibrium velocity distribution. Although a good case can be made that the galaxies have had time to 'virialise', that is, to satisfy the virial theorem when the cluster formed, it takes much longer for energy exchange to take place, except for the most massive galaxies in the cluster. Nonetheless, let us work out the density distribution of an isothermal gas sphere as a reference model for comparison with the observations.

We need the *Lane-Emden equation*, which describes the structure of a spherically symmetric object, such as a star, in hydrostatic equilibrium. The requirement of hydrostatic equilibrium is that, at all points in the system, the attractive gravitational force acting on a mass element  $q \, dV$  at radial distance  $r$  from the centre of the system is balanced by the pressure gradient at that point,

$$\nabla p = -\frac{GMq}{r^2}, \quad (4.3)$$

where  $M$  is the mass contained within radius  $r$ ,

$$M = \int_0^r 4\pi r^2 q(r) dr \quad dM = 4\pi r^2 q(r) dr. \quad (4.4)$$

Reordering (4.3) and differentiating, we find

$$\begin{aligned} \frac{r^2}{q} \frac{dp}{dr} &= -GM \quad \frac{d}{dr} \left( \frac{r^2}{q} \frac{dp}{dr} \right) = -G \frac{dM}{dr} \\ \frac{d}{dr} \left( \frac{r^2}{q} \frac{dp}{dr} \right) + 4\pi Gr^2 q &= 0. \end{aligned} \quad (4.5)$$

Equation (4.5) is the *Lane-Emden equation*. We are interested in the case in which the pressure  $p$  and the density  $q$  are related by the perfect gas law at all radii  $r$ ,  $p = qkT/\mu$ , where  $\mu$  is the mass of an atom, molecule or galaxy. In thermal equilibrium,  $\frac{3}{2}kT = \frac{1}{2}\mu\langle v^2 \rangle$ , where  $\langle v^2 \rangle$  is the mean square velocity of the atoms, molecules or galaxies. Therefore, substituting for  $p$ ,

$$\frac{d}{dr} \left( \frac{r^2}{q} \frac{dq}{dr} \right) + \frac{4\pi G\mu}{kT} r^2 q = 0. \quad (4.6)$$

Equation (4.6) is a non-linear differential equation and, in general, must be solved numerically. There is, however, an analytic solution for large values of  $r$ . If  $q(r)$  is expressed as a power series in  $r$ ,  $q(r) = \sum A_n r^{-n}$ , there is a solution for large  $r$  with  $n = 2$ ,

$$q(r) = \frac{2}{Ar^2} \quad \text{where} \quad A = \left( \frac{4\pi G\mu}{kT} \right). \quad (4.7)$$

This mass distribution has the unfortunate property that the total mass of the cluster diverges at large values of  $r$ ,

$$\int_0^\infty 4\pi r^2 q(r) dr = \int_0^\infty \frac{8\pi}{A} dr \rightarrow \infty. \quad (4.8)$$

There are, however, at least two reasons why there should be a cut-off at large radii. First of all, at very large distances, the particle densities become so low that the mean free path between collisions is very long. The thermalisation time-scales consequently become greater than the time-scale of the system. The radius at which this

occurs is known as *Smoluchowski's envelope*. Secondly, in astrophysical systems, the outermost stars or galaxies are stripped from the system by tidal interactions with neighbouring systems. This process defines a *tidal radius*  $r_t$  for the cluster. Therefore, if clusters are modelled by isothermal gas spheres, it is perfectly permissible to introduce a cut-off at some suitably large radius, resulting in a finite total mass.

It is convenient to rewrite (4.6) in dimensionless form by writing  $Q = q_0 y$ , where  $q_0$  is the central mass density, and introducing a *structural index* or *structural length*  $\alpha$ , where  $\alpha$  is defined by the relation

$$\alpha = \frac{1}{(Aq_0)^{1/2}}. \quad (4.9)$$

Distances from the centre can then be measured in terms of the dimensionless distance  $x = r/\alpha$ . Then, (4.6) becomes

$$\frac{d}{dx} \left[ x^2 \frac{d(\log y)}{dx} \right] + \dot{x} y = 0. \quad (4.10)$$

Two versions of the solution of (4.10) are listed in Table 4.3 and illustrated in Fig. 4.4. In column 2, the solution of  $y$  as a function of distance  $x$  is given; in the third column, the projected distribution onto a plane is given, this being the observed distribution of a cluster of stars or galaxies on the sky. It is a simple calculation to show that, if  $q$  is the projected distance from the centre of the cluster, the surface density  $N(q)$  is related to  $y(x)$  by the integral

$$N(q) = 2 \int_q^\infty \frac{y(x)x}{(x^2 - q^2)^{1/2}} dx. \quad (4.11)$$

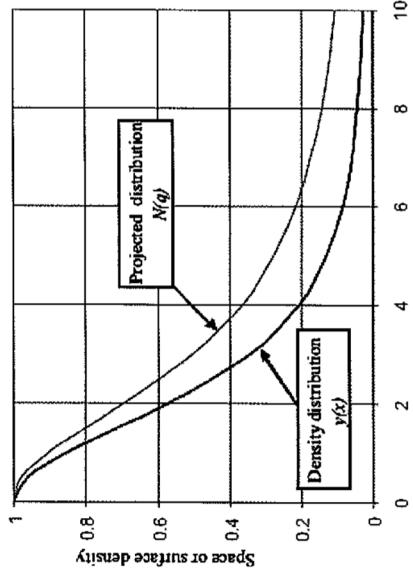
Inspection of Table 4.3 and Fig. 4.4, shows that  $\alpha$  is a measure of the size of the *core* of the cluster. It is convenient to fit the projected distribution  $N(q)$  to the distribution of stars or galaxies in a cluster and then a *core radius* for the cluster can be defined. It can be seen that the projected density falls to the value  $N(q) = 1/2$  at  $q = 3$ , that is, at a core radius  $R_{1/2} = 3\alpha$ .  $R_{1/2}$  is a convenient measure of the core radius of the cluster.

Having measured  $R_{1/2}$ , the central mass density of the cluster can be found if the velocity dispersion of the galaxies in this region is also known. From Maxwell's equipartition theorem,  $\frac{1}{2}\mu\langle v^2 \rangle = \frac{3}{2}kT$  and therefore, from the definition of  $\alpha$ ,

$$\alpha^2 = \frac{1}{Aq_0} = \frac{kT}{4\pi G\mu q_0} = \frac{\langle v^2 \rangle}{12\pi G\mu q_0}. \quad (4.12)$$

Observationally, we can only measure the radial component of the galaxies' velocities  $v_{||}$ . Assuming the velocity distribution of the galaxies in the cluster is isotropic,

$$\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle = 3 \langle v_{||}^2 \rangle. \quad (4.13)$$



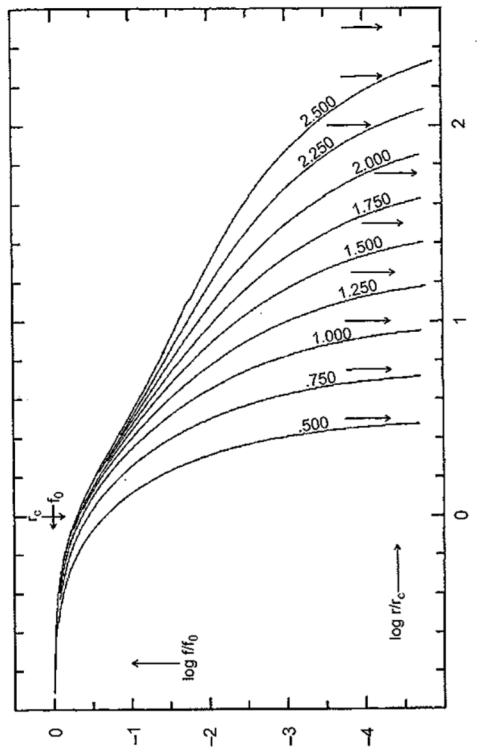
**Fig. 4.4.** The density distribution  $y(x)$  and the projected density distribution  $N(q)$  for an isothermal gas sphere

**Table 4.3.** The density distribution  $y(x)$  and the projected density distribution  $N(q)$  for an isothermal gas sphere

$x, q$	$y(x)$	$N(q)$	$x, q$	$y(x)$	$N(q)$
0	1.0	1.0	12	0.0151	0.0839
0.5	0.9597	0.9782	14	0.0104	0.0694
1.0	0.8529	0.9013	16	0.0075	0.0591
1.5	0.7129	0.8025	20	0.0045	0.0457
2	0.5714	0.6955	30	0.0019	0.0313
3	0.3454	0.5033	40	0.0010	0.0229
4	0.2079	0.3643	50	0.0007	0.0188
5	0.1297	0.2748	100	$1.75 \times 10^{-4}$	0.0101
6	0.0849	0.2143	200	$5.08 \times 10^{-5}$	0.0053
7	0.0583	0.1724	300	$2.32 \times 10^{-5}$	0.0036
8	0.0418	0.1420	500	$8.40 \times 10^{-6}$	0.0021
9	0.0311	0.1209	1000	$2.0 \times 10^{-6}$	0.0010
10	0.0238	0.1050			

Expressing the central density  $\rho_0$  in terms of  $R_{1/2}$  and  $\langle v_{\parallel}^2 \rangle$ , we find

$$\rho_0 = \frac{9 \langle v_{\parallel}^2 \rangle}{4\pi G R_{1/2}^2}. \quad (4.14)$$



**Fig. 4.5.** King models for the distribution of stars in globular clusters, galaxies or of galaxies in clusters of galaxies (King, 1966, 1981). The curves show the projected distribution of stars or galaxies, equivalent to  $N(q)$  in Table 4.3, and are parameterised by the quantity  $\log(r_1/r_c)$  where  $r_1$  is the tidal radius and  $r_c$  the core radius. The arrows indicate  $\log r_1$ .

have  $\log r/r_c \approx 2.2$ ; for dwarf elliptical galaxies, the surface brightness decreases somewhat more rapidly in the outer regions and King models having small values of  $\log r/r_c$  can give a good representation of the data. These profiles have also been found to provide a good description of the distribution of galaxies in clusters.

According to Bahcall, the observed distribution of galaxies in regular clusters can be described by truncated isothermal distributions  $N(r)$  of the form

$$N(r) = N_0[f(r) - C], \quad (4.15)$$

where  $f(r)$  is the projected isothermal distribution normalised to  $f(r) = 1$  at  $r = 0$  and  $C$  is a constant which reduces the value of  $N(r)$  to zero at some radius  $R_h$  such that  $f(R_h) = C$  (Bahcall, 1977). For regular clusters core radii lie in the range  $R_{1/2} = 150 - 400$  kpc, the Coma cluster having  $R_{1/2} = 220$  kpc. Bahcall found that there is a relatively small dispersion in the values of  $C$  required to provide a satisfactory fit to the profiles of many regular clusters, typically the value of  $C$  corresponding to about 1.5% of the isothermal central density.

Many different density distributions have been proposed to describe the space density distribution of galaxies in clusters in addition to the above examples. These include models such as de Vaucouleurs' law for elliptical galaxies (equations 3.2 or 3.3) as well other possibilities such as the Plummer model which is derived from a gravitational potential with a core radius  $b$  of the form

$$\phi = -\frac{GM}{(r^2 + b^2)^{1/2}}, \quad (4.16)$$

where  $M$  is the total mass of the system. Using Poisson's law for gravity in spherical polar coordinates, we find the density distribution from

$$\nabla^2\phi = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \phi}{\partial r} \right) = 4\pi G Q. \quad (4.17)$$

Then,

$$\rho(r) = \frac{3M}{4\pi b^3} \left( 1 + \frac{r^2}{b^2} \right)^{-5/2}. \quad (4.18)$$

Binney and Tremaine discuss these and other possibilities (Binney and Tremaine, 1987).

#### 4.2.3 The Luminosity Function for Cluster Galaxies

The luminosity functions for cluster galaxies can be represented by the Schechter function introduced in Sect. 3.7, but, as indicated by Fig. 3.16, there are significant differences in the shapes of the luminosity functions for early and late-type galaxies as compared with the general field, or the even more extreme case of the void regions. These functions are another way of presenting Dressler's results concerning the

different types of galactic populations found in environments of different densities (Fig. 3.4).

In Sandage's analysis of the redshift-apparent magnitude relation, he demonstrated that there is remarkably little scatter in the absolute magnitudes of the brightest members of rich clusters of galaxies, the standard deviation of the brightest cluster galaxies amounting to  $\sigma_M = 0.28$  (Sandage, 1988) (see Fig. 2.11). There was some controversy about whether or not the constancy of the absolute magnitude of the first-rank cluster galaxy could be explained by randomly sampling the high luminosity region of the luminosity function, or whether they possess some special property which is independent of the cluster richness.

Tremaine and Richstone compared the dispersion in absolute magnitudes of the first-rank members  $\sigma(M)$  with the mean value of the difference in magnitude between the first and second ranked members  $\Delta M_{12} = \langle M_1 - M_2 \rangle$  (Tremaine and Richstone, 1977). They showed that for any statistical luminosity function  $\sigma(M_1)/\Delta M_{12} = t_1 > 1$  must hold. For example, using Schechter's function,  $t_1 = 1.20$ . For Sandage's data on rich clusters,  $t_1 = 0.48 \pm 0.10$ , supporting his point of view that there is much less dispersion in the absolute magnitudes of first ranking cluster galaxies than would be expected if they were simply randomly sampled from the luminosity function. Geller and Postman failed to confirm this result (Geller and Postman, 1983). Nonetheless, as discussed in Sect. 4.2.1, the cD galaxies are quite distinct from normal giant elliptical galaxies (Kormendy, 1982).

A further classification scheme, due to Bautz and Morgan also bears upon the issue of the origins of the brightest galaxies in clusters (Bautz and Morgan, 1970). In the *Bautz-Morgan classification* scheme, clusters are classified according to the presence or absence of a dominant D or cD galaxy at the centre of the cluster. In the *Bautz-Morgan classification* scheme, clusters are classified according to the example Abell 2199; class II clusters have a central galaxy intermediate between a cD and giant elliptical galaxy (for example, the Coma cluster); class III clusters have no dominant central galaxy. Intermediate classes between classes I, II and III have been defined. A number of properties of clusters depend upon Bautz-Morgan class. For example, Sandage has found that there is a weak correlation between Bautz-Morgan class and the absolute magnitude of the brightest cluster member which cannot account for all the difference between the Bautz-Morgan classes. There must in addition be an inverse correlation between Bautz-Morgan class and the absolute magnitudes of the second and third brightest members in the sense that they are relatively brighter in Bautz-Morgan class III clusters. This phenomenon is illustrated by the data in Table 4.4 which lists the mean absolute magnitudes of the first, second and third brightest galaxies in a large sample of rich clusters studied by Sandage and Hardy (Sandage and Hardy, 1973).

#### 4.2.4 Summary of the Properties of Rich Clusters of Galaxies

There are some clear trends in the overall properties of clusters of galaxies. A sequence of types can be defined based on Abell's distinction between *regular* and *irregular* clusters with the addition of an *intermediate* class. Table 4.5 is adapted

**Table 4.4.** The absolute magnitudes of the first, second and third ranked cluster galaxies as a function of Bautz–Morgan type (Sandage 1976). It is assumed that  $h = 0.5$

Bautz–Morgan Class	$M_1$	$M_2$	$M_3$
I	$-23.09 \pm 0.051$	$-22.72 \pm 0.077$	$-22.47 \pm 0.091$
I-II	$-23.30 \pm 0.068$	$-22.83 \pm 0.091$	$-22.56 \pm 0.121$
II	$-23.37 \pm 0.130$	$-22.60 \pm 0.217$	$-22.34 \pm 0.212$
II-III	$-23.46 \pm 0.092$	$-22.18 \pm 0.247$	$-21.96 \pm 0.233$
III	$-23.68 \pm 0.102$	$-22.22 \pm 0.157$	$-21.82 \pm 0.187$

**Table 4.5.** A summary of the typical properties of rich cluster of galaxies of different types (Bahcall, 1977)

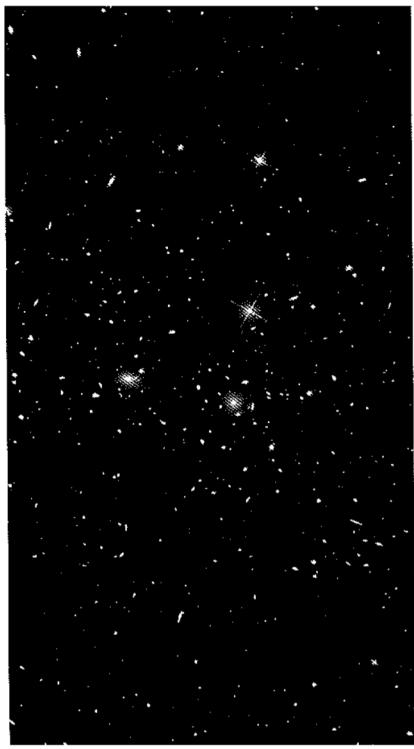
Property/Class	Regular	Intermediate	Irregular
Bautz–Morgan type	I, I-II, II	(II), II-III	(II-III), III
Galaxy content	Elliptical/SO-rich	Spiral-poor	Spiral-rich
E : S0 : S ratio	3 : 4 : 2	1 : 4 : 2	1 : 2 : 3
Symmetry	Spherical	Intermediate	Irregular shape
Central concentration	High	Moderate	Very little
Central profile	Steep gradient	Intermediate	Flat gradient
Mass segregation	Marginal evidence for $m - m(1) < 2$	Marginal evidence for $m - m(1) < 2$	No segregation
Examples	Abell 2199, Coma	Abell 194, 539	Virgo, Abell 1228

from Bahcall's review of 1977 and summarises many of the properties described above as a function of cluster type (Bahcall, 1977). Like the Hubble sequence for galaxies, these types are only part of a continuous sequence and there is considerable overlap in some of the properties.

### 4.3 Dynamical Estimates of the Masses of Clusters of Galaxies

It might seem that the measurement of the masses of clusters of galaxies is relatively straightforward. The virial theorem (3.20) provides a simple relation between the mass of the cluster, the radial velocity dispersion of the galaxies ( $\langle v_{\parallel}^2 \rangle$ ) and the characteristic radius  $R_{\text{cl}}$  of the galaxy distribution. The problems arise in ensuring that the galaxies really belong to the cluster and are not random coincidences of foreground or background objects and also that large enough samples of radial velocities are available. These problems are now being addressed by projects such as the Sloan Digital Sky Survey in which radial velocities, colours and structural data are available for large samples of clusters. The results presented in Table 4.2 give some impression of the wealth of data which will become available for many clusters over the coming years.

**Fig. 4.6.** An optical image of the central region of the Coma cluster of galaxies, Abell 1656, showing the two central massive galaxies NGC 4889 and NGC 4874. North is to the right of this image. (Courtesy of Dr. Robert Lupton and the SDSS consortium)



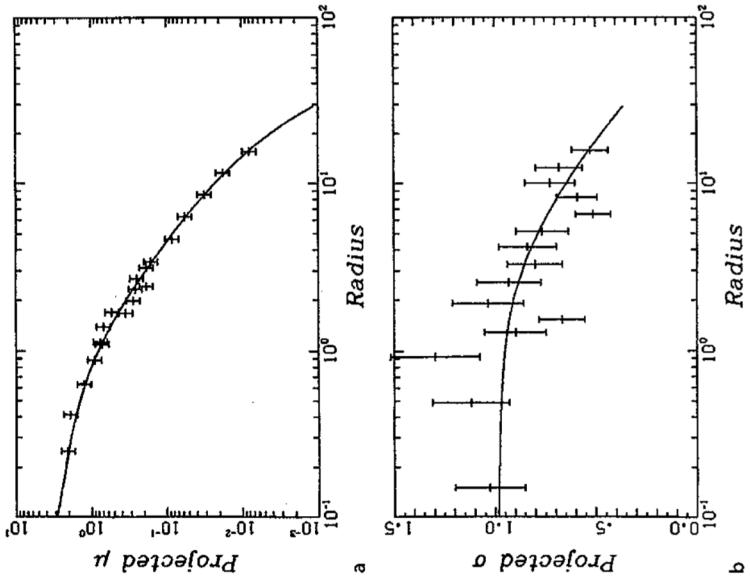


Fig. 4.7. a The surface density profile for the distribution of galaxies in the Coma cluster according to Kent and Gunn. b The projected velocity dispersion as a function of radius for galaxies in the Coma cluster (Kent and Gunn, 1982)

galaxies in the central region of the Coma cluster is dominated by elliptical and S0 galaxies for which the typical mass-to-luminosity ratios are about  $10 - 20 M_{\odot}/L_{\odot}$ . There is therefore a discrepancy of about a factor of 20 between the mass which can be attributed to galaxies and the total mass which must be present. This is the result which was discovered by Zwicky (Zwicky, 1937).

This is also where the trouble begins. *Dark matter* dominates the mass of the cluster but there is no reason why it should have the same distribution as the visible matter. Likewise, there is no reason *a priori* why the velocity distribution of the galaxies should be isotropic. This is most simply expressed in terms of the velocity dispersion of the galaxies in the radial and circumferential directions,  $\sigma_r^2(r)$ ,  $\sigma_\theta^2(r)$  and  $\sigma_\phi^2(r)$  within the cluster. The assumption of isotropy is that  $\sigma_r^2(r) = \sigma_\theta^2(r) = \sigma_\phi^2(r)$ . If, however, the galaxies were on circular orbits about the

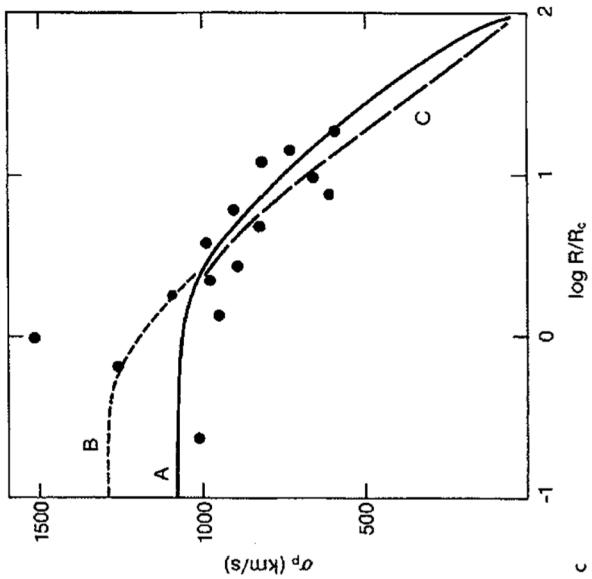


Fig. 4.7. (continued) c Three possible velocity dispersion profiles which are consistent with the data of Gunn and Kent. A corresponds to an isotropic velocity distribution, while B and C involve velocity anisotropies (Merritt, 1987)

centre,  $\sigma_r^2(r) = 0$ , or, if they were purely radial,  $\sigma_r^2(r) = \sigma_\theta^2(r) = 0$ . Kent and Gunn pointed out that the extreme radial models are inconsistent with the variation of velocity dispersion with radius shown in Fig. 4.7. Merritt carried out a careful study of how the inferred mass-to-luminosity ratio would change for a wide range of different assumptions about the relative distributions of the visible and dark matter and the anisotropy of the velocity distribution (Fig. 4.7c) (Merritt, 1987). For the cluster as a whole, the mass-to-luminosity ratio varied from about 0.4 to at least three times the reference value, while the mass-to-luminosity ratio within the core of the cluster at  $1 h^{-1} \text{ Mpc}$  was always very close to  $350 h M_{\odot}/L_{\odot}$ . There can be no doubt that the dynamics of the cluster are dominated by dark matter.

Perhaps the most remarkable result of recent times has been the finding that the Coma cluster is probably not the quiescent regular cluster it appears to be. Colless and Dunn have added 243 more radial velocities to the sample, bringing the total number of cluster members with radial velocities to 450 (Colless and Dunn, 1996). They find compelling evidence that, in addition to the main body of the cluster, there is a distinct subcluster whose brightest member is NGC 4639. The main cluster has mass  $0.9 \times 10^{15} h^{-1} M_{\odot}$ , while the less massive cluster has mass  $0.6 \times 10^{14} h^{-1} M_{\odot}$ .

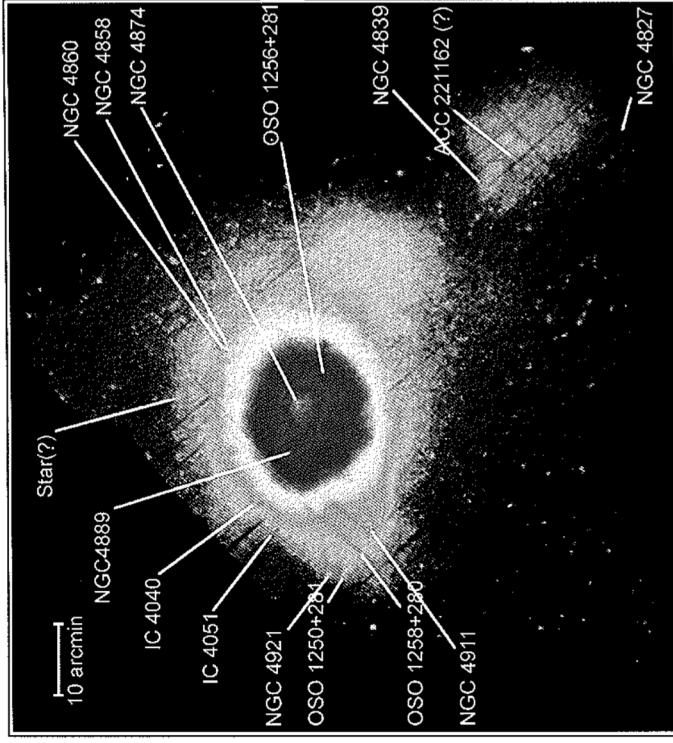


Fig. 4.8. An X-ray image of the Coma cluster of galaxies obtained by the XMM-Newton Observatory, showing the X-ray emission associated with the main body of the Coma cluster and the smaller cluster associated with NGC 4839 (Courtesy of the Max Planck Institute for Extraterrestrial Physics and ESA).

These clusters are clearly seen in the XMM-Newton X-ray image of the Coma cluster and the total masses derived from the X-ray observations agree with those derived by Colless and Dunn inferred that there is subclustering of galaxies about the two brightest members NGC4889 and NGC4874 and that these are in the process of coalescing in the central regions of the cluster. The moral of this story is that appearances can be deceptive.

#### 4.4 X-Ray Observations of Hot Gas in Clusters of Galaxies

The X-ray image of the Coma cluster (Fig. 4.8) provides an excellent demonstration of the power of X-ray astronomy in the study of clusters of galaxies. One of the most important discoveries made by the UHURU X-ray Observatory was the detection of intense X-ray emission from rich clusters of galaxies. The nature of the emission was

soon identified as the bremsstrahlung of hot intracluster gas, the key observations being the extended nature of the emission and the subsequent detection of the highly ionised iron line FeXXVI by the Ariel-V satellite (Mitchell et al., 1976). It was quickly appreciated that the X-ray emission of the gas provides a very powerful probe of the gravitational potential within the cluster enabling the distribution of hot gas and the total gravitating mass to be determined. Let us repeat the calculation presented by Fabricant, Lecar and Gorenstein which shows how this can be done (Fabricant et al., 1980).

For simplicity, we assume that the cluster is spherically symmetric so that the total gravitating mass within radius  $r$  is  $M(\leq r)$ . The gas is assumed to be in hydrostatic equilibrium within the gravitational potential defined by the total mass distribution in the cluster, that is, by the sum of the visible and dark matter as well as the gaseous mass. If  $p$  is the pressure of the gas and  $Q$  its density, both of which vary with position within the cluster, the requirement of hydrostatic equilibrium is

$$\frac{dp}{dr} = -\frac{GM(\leq r)\rho}{r^2}. \quad (4.19)$$

The pressure is related to the local gas density  $Q$  and temperature  $T$  by the perfect gas law

$$p = \frac{\rho k T}{\mu m_H}, \quad (4.20)$$

where  $m_H$  is the mass of the hydrogen atom and  $\mu$  is the mean molecular weight of the gas. For a fully ionised gas with the standard cosmic abundance of the elements, a suitable value is  $\mu = 0.6$ . Differentiating (4.20) with respect to  $r$  and substituting into (4.19), we find

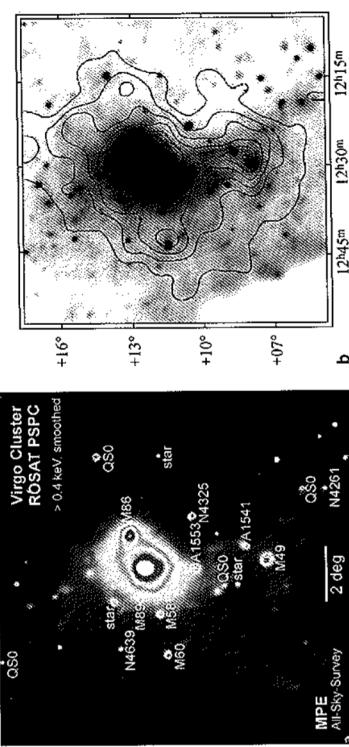
$$\frac{QkT}{\mu m_H} \left( \frac{1}{\rho} \frac{d\rho}{dr} + \frac{1}{T} \frac{dT}{dr} \right) = -\frac{GM(\leq r)\rho}{r^2}. \quad (4.21)$$

Reorganising (4.21),

$$M(\leq r) = -\frac{kT^2}{G\mu m_H} \left[ \frac{d(\log \rho)}{dr} + \frac{d(\log T)}{dr} \right]. \quad (4.22)$$

Thus, the mass distribution within the cluster can be determined if the variation of the gas density and temperature with radius are known. Assuming the cluster is spherically symmetric, these can be derived from high sensitivity X-ray intensity and spectral observations. A suitable form for the bremsstrahlung spectral emissivity of a plasma is

$$\kappa_\nu = \frac{1}{3\pi^2} \frac{Z^2 e^6}{\epsilon_0^3 c^3 m_e^2} \left( \frac{m_e}{kT} \right)^{1/2} g(\nu, T) N N_e \exp \left( -\frac{h\nu}{kT} \right), \quad (4.23)$$



**Fig. 4.9a,b.** Comparison of the X-ray surface brightness distribution and the surface density of luminous matter in galaxies in the Virgo cluster. **a** The X-ray image of the Virgo cluster from the ROSAT All Sky Survey in the 0.4–2 keV band. The image has been smoothed with a Gaussian filter with  $\sigma = 24 \text{ arcmin}$  at the faintest levels and with decreasing filter size with increasing surface brightness. Some galaxies in the Virgo cluster have been detected as well as a few distant clusters and active galaxies. (Courtesy of the Max Planck Institute for Extraterrestrial Physics, Garching.) **b** The contours show the number density of the 1292 member galaxies in the Virgo Cluster Catalogue smoothed with a Gaussian filter with  $\sigma = 24 \text{ arcmin}$ . The spacing of the contours is linear (Schindler et al., 1999)

where  $N_e$  and  $N$  are the number densities of electrons and nuclei respectively,  $Z$  is the charge of the nuclei and  $g(\nu, T)$  is the Gaunt factor, which can be approximated by

$$g(\nu, T) = \frac{\sqrt{3}}{\pi} \ln \left( \frac{kT}{h\nu} \right). \quad (4.24)$$

The spectrum of thermal bremsstrahlung is roughly flat up to X-ray energies  $\varepsilon = h\nu \sim kT$ , above which it cuts off exponentially (Longair, 1997b). Thus, by making precise spectral measurements, it is possible to determine the temperature of the gas from the location of the spectral cut-off and the column density of the hot gas from the X-ray surface brightness. In practice, the spectral emissivity has to be integrated along the line of sight through the cluster. Performing this integration and converting it into an intensity, the observed surface brightness at projected radius  $a$  from the cluster centre is

$$I_\nu(a) = \frac{1}{2\pi} \int_a^\infty \frac{\kappa_\nu(r)r}{(r^2 - a^2)^{1/2}} dr. \quad (4.25)$$

Cavaliere noted that this is an Abel integral which can be inverted to find the emissivity of the gas as a function of radius (Cavaliere, 1980)

$$\kappa_\nu(r) = \frac{4}{r} \frac{d}{dr} \int_r^\infty \frac{I_\nu(a)a}{(a^2 - r^2)^{1/2}} da. \quad (4.26)$$

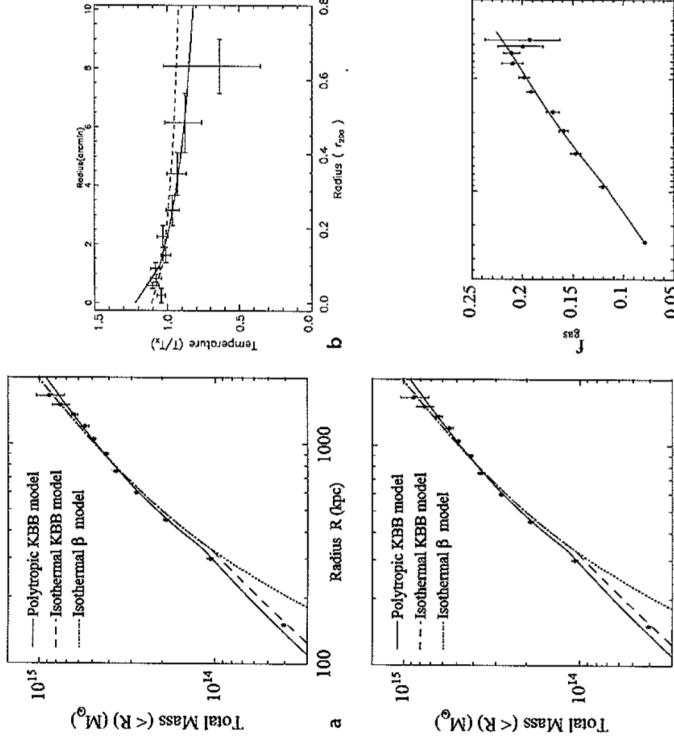
X-rays maps of about 200 clusters were made by the Einstein X-ray Observatory (Forman and Jones, 1982) and more recently analyses of the ROSAT All Sky Survey have resulted in the NORAS and REFLEX catalogues of clusters of galaxies identified solely through the requirement that the extragalactic X-ray sources have extended structures (Böhringer et al., 2000, 2001). In the northern NORAS sample, it was found that 76% of the extended X-ray sources were indeed associated with clusters of galaxies. In addition, the ROSAT, Chandra and XMM-Newton X-ray Observatories have produced beautiful X-ray maps and spectroscopic studies of the hot gas in clusters.

An example of the quality of data now available is illustrated by the X-ray map of the central regions of the Virgo cluster as observed in the ROSAT All Sky Survey (Fig. 4.9a). A number of galaxies belonging to the Virgo cluster have been detected as X-ray sources, as well as a few background clusters and active galaxies. In addition, the X-ray emission of the diffuse intergalactic gas is roughly centred on the massive galaxy M87. Evidence that the intergalactic gas traces the mass distribution of the cluster is provided by comparison of the contours of the X-ray surface brightness distribution with the surface distribution of galaxies as determined by the photometric survey of the Virgo cluster by Binggeli, Tammann and Sandage (Binggeli et al., 1987; Schindler et al., 1999). The distribution of galaxies in the cluster and the diffuse X-ray emission are remarkably similar (Fig. 4.9b). In both wavebands the irregular structure of the Virgo cluster can be decomposed into three major subclusters centred on M87, M49, and M86. In the M87 subcluster the gas

mass is about three times the visible mass in galaxies, assuming  $M/L = 20$  for the galaxies), while it accounts only for 8% to 14% of the total mass at 0.4 and 1 Mpc, respectively. The projected mass-to-light ratio has a roughly constant value of about  $500 M_\odot/L_\odot$ . Thus, the dark matter which defines the gravitational potential in the cluster and which is traced by the distribution of hot gas, must have a similar distribution to that of the visible matter.

Another beautiful example of the combined use of X-ray imaging and spectroscopy is provided by the observations and analysis of the rich cluster Abell 1413 by Pratt and Arnaud (Pratt and Arnaud, 2002). The observations by the XMM-Newton X-ray Observatory included spatially resolved X-ray spectroscopy of the cluster X-ray emission and so the projected temperature variation with radius in the cluster could be determined. Their results are shown in Fig. 4.10. First, the average X-ray surface brightness distribution as a function of radius is fitted by an empirical model (Fig. 4.10a). Then, the projected average temperature of the gas is estimated in annuli at different radial distances from the centre of the cluster (Fig. 4.10b). These are deprojected to derive the variation of the total mass within radius  $r$  using (4.22) (Fig. 4.10c). Finally, the ratio of gas density to total density as a function of radius, or in the case of Fig. 4.10d, the overdensity relative to the critical cosmological density, can be found.

These data are typical of what is found in rich clusters of galaxies. The dominant form of mass is the dark matter the nature of which is unknown. About 20% of the



**Fig. 4.10a-d.** Illustrating the determination of the physical properties of the cluster A1413 from X-ray imaging and spectroscopy by the XMM-Newton X-ray Observatory. **a** The X-ray brightness distribution as a function of distance from the centre of the cluster. **b** The projected radial distribution of temperature of the gas. **c** The integrated mass distribution as a function of distance from the centre. **d** The fraction of gas density to total mass density  $f_{\text{gas}}$  within the cluster as a function of overdensity  $\delta$  relative to the critical cosmological density (Pratt and Arnaud, 2002)

mass is in the form of hot intergalactic and this is typically about five times the mass in the visible parts of galaxies. The spectroscopic observations also enable the mass of iron in the intracluster medium to be determined and this is typically found to be between about 20 and 50% of the solar value, indicating that the intergalactic gas has been enriched by the products of stellar nucleosynthesis.

The wealth of data from the recent generation of X-ray satellites has enabled important insights into their origin and evolution of clusters to be obtained. It is convenient to adopt a simple reference model for understanding the various correlations between the properties of clusters, what is referred to as the *self-similar model* for cluster formation; this will be discussed in more detail in Chap. 16.

The simplest picture of cluster formation is to suppose that clusters developed from small perturbations in the distribution of the dark matter by gravitational collapse (see Part II). The amplitude of the perturbations grew from values  $\delta Q/Q \sim 10^{-3}$  at the epoch of recombination to large amplitudes  $\delta Q/Q \gg 1$  at late epochs. The considerations of Sect. 16.1 suggest that clusters became virialised bound systems when  $(\delta Q/Q)_v \approx 200$ . In the standard self-similar model, it is supposed that clusters of all masses virialised at the same epoch with the same values of  $(\delta Q/Q)_v$ . As a result, their masses are simply proportional to  $Q_v R^3$ , where  $Q_v$  was the mean density of the Universe at the epoch of virialisation. When the clusters virialised, they satisfied the virial theorem,  $G M^2/R = M \sigma_v^2$ . Since  $M \propto R^3$ , it follows that  $\sigma_v \propto R$ , where  $R$  can be taken to be the characteristic radius of the cluster.

The bolometric X-ray luminosity of the intracluster gas due to thermal bremsstrahlung is given by the relation

$$L_x \propto V N_e^2 T^{1/2}, \quad (4.27)$$

where  $V \sim R^3$  is the volume of emitting gas,  $N_e$  is the electron density and  $T$  the temperature of the gas (Longair, 1997b). The gas forms an atmosphere within the gravitational potential of the cluster as a whole, the latter being defined by the distribution of the dark matter. Equation (4.19) can be written to order of magnitude as

$$\frac{p}{R} \sim \frac{G M \rho}{R^2} \quad \text{and so} \quad 3 N_e k T \sim \frac{G M}{R} \rho, \quad (4.28)$$

where it is assumed that the ions and electrons both contribute to the pressure of the gas. According to the virial theorem (3.19), the velocity dispersion of the galaxies is related to the gravitational potential by  $G M/R \approx \sigma_v^2$  and so

$$k T \sim m_p \sigma_v^2, \quad (4.29)$$

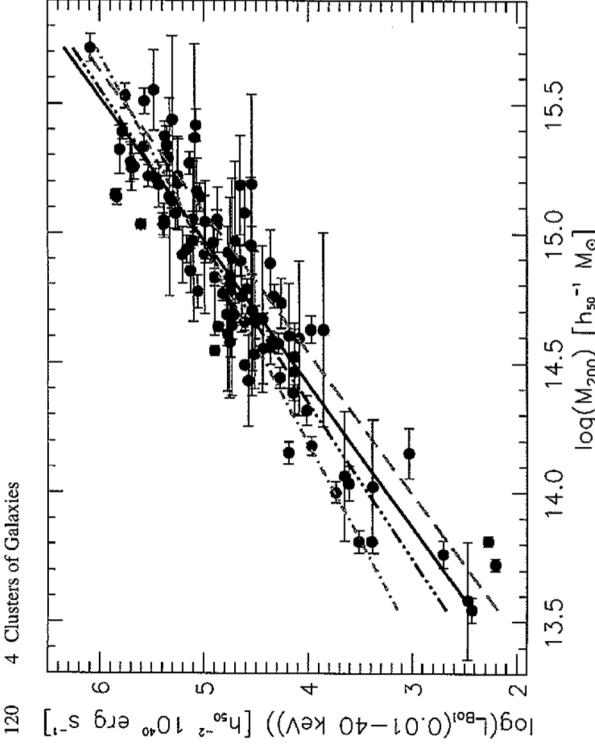
that is,  $T \propto \sigma_v^2$ . If the fraction of the total mass of the cluster in hot intracluster gas is  $\eta$ , it follows that the X-ray luminosity of the cluster is

$$L_x \propto \eta \frac{M^2}{R^3} T^{1/2} \propto R \sigma_v^2 T^{1/2} \propto \sigma_v^4, \quad (4.30)$$

where it has been assumed that  $\eta$  takes the same value for clusters of different masses. Then, since  $M \propto R^3 \propto \sigma_v^3$ , this model results in the prediction that the X-ray luminosity and the mass of the cluster are related by

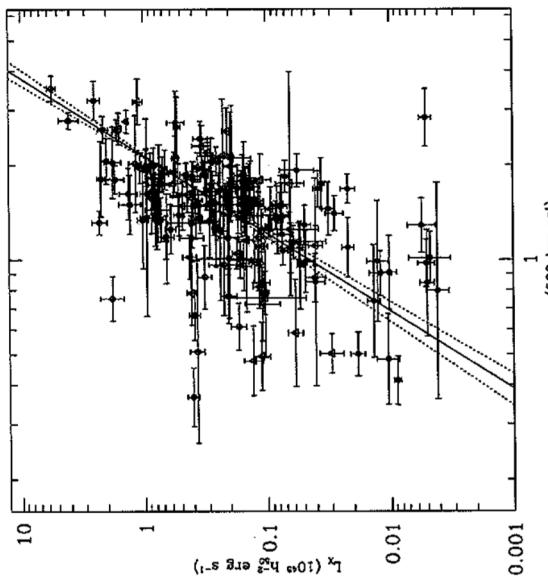
$$L_x \propto M^{4/3}. \quad (4.31)$$

The scaling relations (4.30) and (4.31) have been the subject of considerable study. Figure 4.11 shows the results of a major study of X-ray clusters discovered in the ROSAT All Sky Survey (Reiprich and Böhringer, 2002). It can be seen that the best-fit relation is slightly steeper than the predictions of the scaling model, but it is interesting that the general trend is not so different from the simple picture.



**Fig. 4.11.** The correlation between the X-ray luminosity of a cluster of galaxies and its total mass for the brightest X-ray clusters of galaxies in the ROSAT All Sky Survey (Reiprich and Böhringer, 2002). The mass was determined to the virial radius which was taken to be a density 200 greater than the critical cosmological density. The best-fit relation for their sample of 106 clusters is shown by the *solid line*. The best-fit relation for a smaller sample of 63 clusters which form a better complete sample is shown by the *tripledot-dashed line*. The *dotted-dashed line* shows the expectation of the simple self-similar relation  $L_x \propto M^{4/3}$ . The *dashed line* is the prediction of a model with preheating of the intracluster gas

Another test of this model is the X-ray luminosity–velocity dispersion relation which has been analysed in some detail by Ortiz-Gil and her colleagues (Ortiz-Gil et al., 2004). For a large sample of clusters from the REFLEX catalogue derived from the ROSAT All Sky Survey, they found a correlation of the form  $L_x \propto \sigma_v^{(4.1 \pm 0.3)}$ , in good agreement with the correlation (4.30) predicted by the self-similar model (Fig. 4.12). A similar result was found for a smaller sample of clusters which formed a more complete statistical sample. Although this result is encouraging, their results for the  $L_x - T$  and  $\sigma_v - T$  relations differ from the expectations of the self-similar model. They found  $L_x \propto T^{3.1 \pm 0.2}$  compared with the expected relation  $L_x \propto T^2$  and  $\sigma_v \propto T^{1.00 \pm 0.16}$  compared with the predicted  $\sigma_v \propto T^{0.5}$ . They noted that these last two relations are compatible with the X-ray luminosity–velocity dispersion relation. They inferred that there must be other contributions to the thermal heating and cooling of the intracluster gas than simply the thermal energy which resulted from the initial process of virialisation. There is now considerable evidence that such



**Fig. 4.12.** A plot of bolometric X-ray luminosity  $L_x$  versus radial velocity dispersion  $\sigma_v$  for the REFLEX sample of 171 clusters. The linear fit has logarithmic slope of  $(4.1 \pm 0.3)$ . The dashed lines are the  $1\sigma$  errors. *Filled circles* correspond to clusters at redshift  $z \leq 0.05$ , *open triangles* are clusters with  $0.05 < z \leq 0.1$  and *open circles* are clusters at  $z > 0.1$  (Ortiz-Gil et al., 2004)

additional sources of heating and cooling are important in understanding the state of the intracluster gas.

If the density of the hot intracluster gas is large enough, its cooling rate can be sufficiently great for it to cool over cosmological time-scales. At high enough temperatures, the principal energy loss mechanism for the gas is the same thermal bremsstrahlung process which is responsible for the X-ray emission. The total energy loss rate per unit volume due to thermal bremsstrahlung is

$$-\left(\frac{dE}{dt}\right) = 1.435 \times 10^{-40} Z^2 T^{\frac{1}{2}} \bar{g} N N_e \quad \text{W m}^{-3}, \quad (4.32)$$

where  $Z$  is the charge of the ions,  $N$  and  $N_e$  are the number densities of ions and electrons respectively and  $\bar{g}$  is a mean Gaunt factor which has value roughly 1 – we assume  $Z = 1$  and  $N = N_e$ . The thermal energy density of the fully ionised plasma is  $\epsilon = 3NkT$  and so the characteristic cooling time for the gas is

$$t_{\text{cool}} = \frac{3NkT}{\left|\frac{dE}{dt}\right|} = 10^{10} \frac{T^{1/2}}{N} \quad \text{years}, \quad (4.33)$$

where the temperature is measured in kelvins and the number density of ions or electrons is measured in particles m<sup>-3</sup>. Thus, if the typical temperature of the gas is 10<sup>7</sup> to 10<sup>8</sup> K, the cooling time is less than 10<sup>10</sup> years if the electron density is greater than about 3 × 10<sup>3</sup> to 10<sup>4</sup> m<sup>-3</sup>. These conditions are indeed found in many of the clusters of galaxies which are intense X-ray emitters. As a result, the central regions of these hot gas clouds cool and, to preserve pressure balance, the gas density increases, resulting in the formation of a *cooling flow*.

Evidence for these cooling flows and their role in the evolution of clusters of galaxies was reviewed by Fabian (Fabian, 1994). An example of the cooling flow in the cluster Abell 478 is illustrated by the diagrams shown in Fig. 4.13. The ROSAT observations have been deprojected to determine mean values of the density and temperature of the gas as a function of radial distance from the centre. The temperature decreases towards the central regions while the electron density increases to values greater than 10<sup>4</sup> m<sup>-3</sup> in the very centre. At a radius of 200 kpc, the electron temperature is  $T = 7 \times 10^7$  K and the electron density  $N_e = 8 \times 10^3$  m<sup>-3</sup>. Inserting these values into (4.33), we find that the cooling time is 10<sup>10</sup> years. It can be seen that, outside this radius, the temperature of the gas is constant, whilst at smaller radii the gas temperature decreases towards the central regions.

As a result, matter drifts slowly in through the surface at radius  $r_{\text{cool}}$ , at which the cooling time of the gas is equal to the age of the cluster. The X-ray luminosity of

the cooling flow results from the internal energy of each element of the gas as well as the work done as it drifts slowly in towards the central regions whilst maintaining hydrostatic equilibrium. The appropriate conserved quantity is

$$\rho v^2 (\frac{1}{2} v^2 + w), \quad (4.34)$$

where  $\rho$ ,  $v$  and  $w$  are the density, velocity and enthalpy of the gas flow. Note that the relevant quantity which describes the energy content of the flow is the enthalpy  $w = \epsilon + PV$ , where  $\epsilon$  is the internal energy of the gas and  $p$  and  $V$  are its pressure and specific volume. This is because, in addition to its internal thermal energy, work is done on the volume element by the pressure of the gas. When the volume element drifts within the cooling radius  $r_{\text{cool}}$ , all of the energy associated with the enthalpy is available for heating the gas and so the energy deposited per unit mass is  $w = \epsilon + PV = \frac{5}{2} n kT$  for a perfect gas, where  $n$  is the number density of particles in the flow. It follows that the total energy input to the gas is determined by the rate at which mass drifts inwards through any radius,  $\dot{N}$ . In the steady state, this energy is reradiated by the gas as X-rays and so the cooling luminosity of the flow  $L_{\text{cool}}$  is

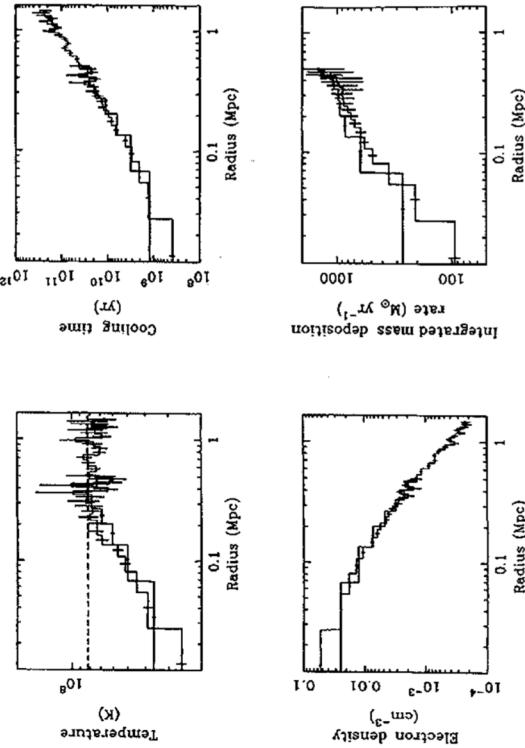
$$L_{\text{cool}} = \frac{5}{2} \dot{N} k T = \frac{\dot{M}}{\mu m} k T, \quad (4.35)$$

where  $m$  is the mass of the hydrogen atom and  $\mu$  is the mean molecular weight of the gas. Thus, knowing the total X-ray luminosity due to cooling, and the temperature of the gas, (4.35) can be used to work out the mass flow rate  $\dot{M}$  through any radius. The result of this calculation for the cluster Abell 478 is shown in Fig. 4.13. The cooling flow results in a mass inflow rate of about 600 to 800 M<sub>⊙</sub> y<sup>-1</sup> and so over a period of 10<sup>10</sup> years, such cooling flows can contribute significantly to the baryonic mass in the central regions of the cluster.

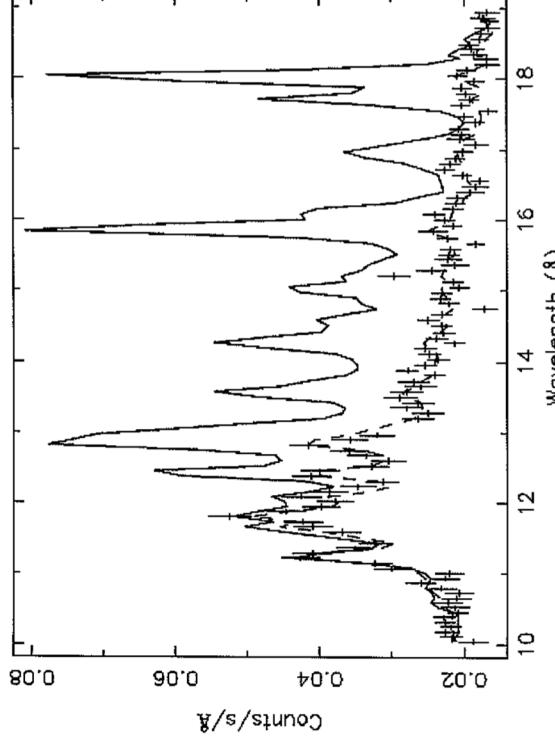
According to Fabian, about half of the clusters detected by the Einstein X-ray Observatory have high central X-ray surface brightnesses and cooling times less than 10<sup>10</sup> years (Fabian, 1994). Abell 478 has a particularly massive flow. Typically, the inferred mass flow rates are about 100 to 300 M<sub>⊙</sub> y<sup>-1</sup>. This cannot be the whole story, however, since X-ray spectroscopic observations of the cores of clusters have shown that there is an absence of cool gas which would be expected if there were no other energy sources. This is most vividly demonstrated by observations by the ESA XMM-Newton satellite, for example, the observations of the cluster Sérsic 159-03 which has a cool core (de Plaa et al., 2005). Figure 4.14 shows the X-ray spectrum of the cluster, the solid line showing the wealth of X-ray emission lines expected according to standard models of cooling flows. It can be seen that the observed spectrum differs dramatically from the expectations of the cooling flow models, because of the absence of strong lines associated with ions such as FeXVII. This lack of cool gas seems to be a feature of many of the cooling flows observed in rich clusters of galaxies (Kaastra et al., 2004). The inference is that

there must be some further heating source within the cluster to reheat the cooling gas.

Many models have been proposed to resolve this problem, some of these being listed by Kaastra and his colleagues (Kaastra et al., 2004). These models include



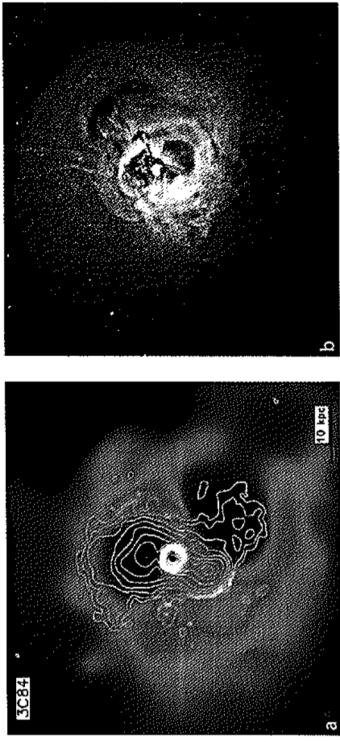
**Fig. 4.13.** The properties of the intracluster gas in the cluster Abell 478 obtained by deprojecting images taken by the ROSAT X-ray Observatory (White et al., 1994). The cooling time of the gas is less than 10<sup>10</sup> years within a radius of 200 kpc (Fabian, 1994).



**Fig. 4.14.** Comparison of the observed high-resolution X-ray spectrum of the cluster of galaxies Sersic 159-03 observed by the ESA XMM-Newton satellite with the predicted spectrum of a standard cooling flow model without heating. The strong lower excitation lines from ions such as Fe XVII are absent, indicating the lack of cool gas in the cluster (de Plaa et al., 2005)

metallicity inhomogeneities, buoyantly rising radio bubbles transporting cool gas outwards, halo-in-halo structures, turbulent mixing due to rising and falling hot gas bubbles, heating by AGN activity, contamination due to non-thermal X-ray emission, heating by dead radio galaxies, rapid cooling due to mixing with cold gas and heat conduction by electrons. They also propose heating by coronal loops, similar to those observed in the corona of the Sun, but on the scale of the cores of clusters of galaxies.

A highly suggestive set of observations made by the Chandra X-ray Observatory shows that the cooling gas in the central regions of a number of clusters is perturbed by the presence of radio lobes of recent radio source events. One of the most compelling examples is the central region of the Perseus cluster of galaxies in which the buoyant lobes of relativistic plasma have pushed back the intraccluster gas, forming ‘holes’ in the X-ray brightness distribution (Fig. 4.15a) (Fabian et al., 2000). In a very long X-ray exposure with the Chandra X-ray Observatory, Fabian and his colleagues identify what they interpret as isothermal sound waves produced by the weak shock wave associated with the expanding lobes (Fig. 4.15b). They show that the energy injected into the intraccluster gas by these sound waves can balance the radiative cooling of the cooling flow (Fabian et al., 2006).



**Fig. 4.15a,b.** The central regions of the Perseus cluster of galaxies as observed by the Chandra X-ray Observatory. **a** The central regions of the cluster showing the cavities evacuated by the radio lobes which are shown by the *white contour lines* (Fabian et al., 2000). **b** An unsharp-mask image of the central regions of the cluster showing the various features caused by the expanding radio lobes. Many of the features are interpreted as sound waves caused by the weak shock wave associated with the expansion of the radio lobes (Fabian et al., 2006)

## 4.5 The Sunyaev–Zeldovich Effect in Hot Intraccluster Gas

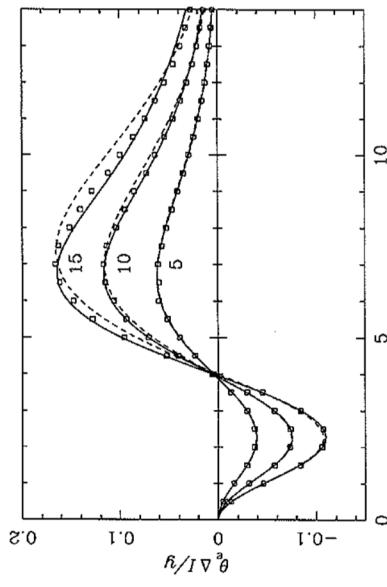
An independent method of studying hot gas in clusters of galaxies and elsewhere in the Universe is through the search for decrements in the intensity of the Cosmic Microwave Background Radiation. As the photons of the background radiation pass through the gas cloud, a few of them suffer Compton scattering by the hot electrons. Although to first order the photons are just as likely to gain as lose energy in these scatterings, to second order there is a net statistical gain of energy and so the spectrum of the Cosmic Microwave Background Radiation is shifted to slightly higher energies. As a result, there is expected to be a decrease in the intensity of the background radiation in the Rayleigh–Jeans region of the spectrum, that is, at energies  $h\nu \ll kT_e$ , while in the Wien region,  $h\nu \gg kT_e$ , there should be a slight excess –  $T_e$  is the temperature of the background radiation. These predictions were made by Sunyaev and Zeldovich as long ago as 1969 (Sunyaev and Zeldovich, 1970) and it was almost 20 years before the Sunyaev–Zeldovich effect was observed with confidence in the directions of clusters of galaxies (Birkinshaw, 1990).

The magnitude of the distortion is determined by the *Compton scattering optical depth*  $y$  through the region of hot gas,

$$y = \int \left( \frac{kT_e}{m_e c^2} \right) \sigma_T N_e dl . \quad (4.36)$$

The resulting decrement in the Rayleigh–Jeans region of the spectrum is

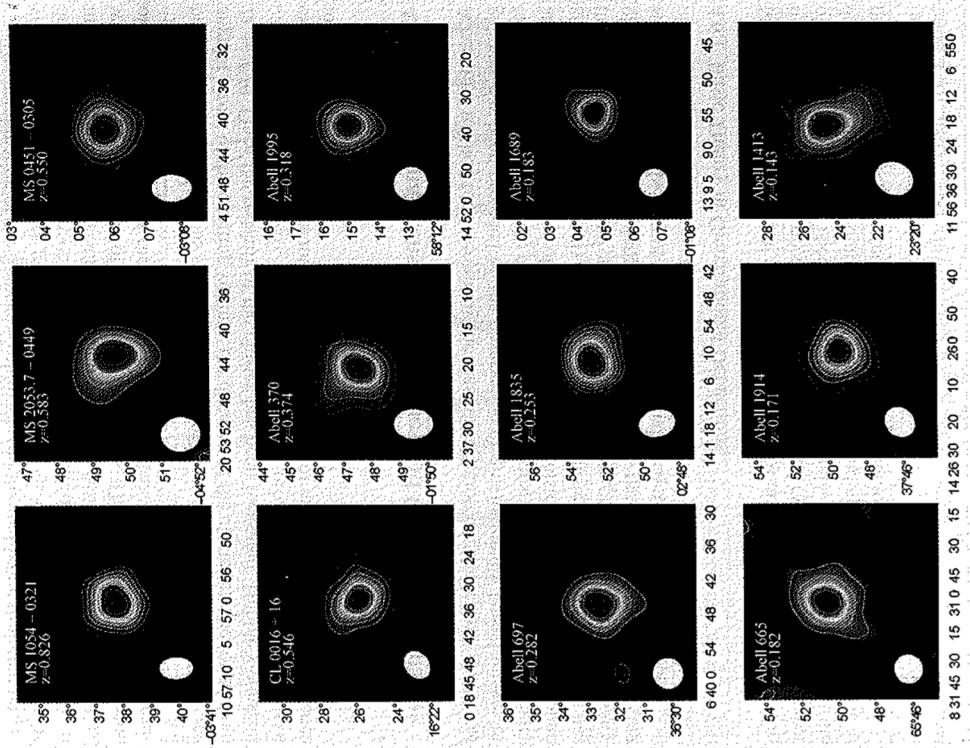
$$\frac{\Delta I_\nu}{I_\nu} = -2y . \quad (4.37)$$



**Fig. 4.16.** Intensity changes due to the Sunyaev-Zeldovich effect in units of  $2(kT_0)^3/(hc)^2$ , plotted against  $X = h\nu/kT_0$  for three values of  $kT_e$  (in keV), where  $\theta_e = kT_e/m_ec^2$ . The solid curves are calculated using the second-order correction to the Kompaneets equation, while the dashed lines are calculated from the first-order correction. The points are the result of a Monte Carlo evaluation of the Boltzmann collision integral by Garrett and Gull (Challinor and Lasenby, 1998)

Thus, the magnitude of the decrement along any line of sight through the cluster provides a measure of the quantity  $\int N_e T_e dl$ , in other words, the integral of the pressure of the hot gas along the line of sight.<sup>1</sup> For the typical parameters of hot intracluster gas, the predicted decrement amounts to  $\Delta I/I \approx 10^{-4}$ . The spectral signature of the effect is quite distinctive over the peak of the spectrum of the Cosmic Microwave Background Radiation (Fig. 4.16) and has been worked out in detail by Challinor and Lasenby (Challinor and Lasenby, 1998). This form of distortion has been measured in 15 Abell clusters in the Suzie experiment carried out at the CalTech Submillimetre Observatory on Mauna Kea (Benson et al., 2004).

An important feature of the Sunyaev-Zeldovich effect is that, if the hot gas clouds have the same properties at all redshifts, the observed decrement is independent of redshift since the scattering results in only a *fractional* change in the temperature of the background radiation. This prediction is beautifully illustrated by the maps of decrements in the Cosmic Microwave Background Radiation obtained by the OVRO and BIMA millimetre arrays which span a range of redshift from 0.1 to 0.8 (Fig. 4.17) (Carlstrom et al., 2000). All these clusters were known to be X-ray sources and there is good agreement between the sizes of the X-ray images and the Sunyaev-Zeldovich decrements.



**Fig. 4.17.** Images of the Sunyaev-Zeldovich decrement in 12 distant clusters with redshifts in the range 0.14 to 0.89 (Carlstrom et al., 2000). Each of the images is plotted on the same intensity scale. The data were taken with the OVRO and BIMA millimetre arrays. The *filled ellipse* at the bottom left of each image shows the full-width half-maximum of the effective resolution used in reconstructing the images

<sup>1</sup> I have given a discussion of the physical process involved in the Sunyaev-Zeldovich effect in *High Energy Astrophysics, Vol. 1* (Longair, 1997b).

The combination of the Sunyaev-Zeldovich and thermal bremsstrahlung observations of the intrachuster gas enable the dimensions of the hot gas cloud to be determined independently of knowledge of the redshift of the cluster. Using order of magnitude arguments, the Sunyaev-Zeldovich effect determines the quantity  $N_e^2 T_c L$ , where  $L$  is the dimension of the volume of hot gas. The bremsstrahlung emission of the cluster determines the quantity  $L^3 N_e^2 T^{1/2}$ . The temperature  $T$  can be estimated from the shape of the bremsstrahlung spectrum and so  $N_e$  can be eliminated between these two relations so that an estimate of  $L$  can be found. By measuring the angular size  $\theta$  of the emitting volume, the distance of the cluster can be found from  $D = L/\theta$ . Once the redshift of the cluster has been measured, Hubble's constant can be estimated (Sect. 8.3). This is one of the more promising physical methods of estimating Hubble's constant without the necessity of using a hierarchy of distance indicators.

## 4.6 Gravitational Lensing by Galaxies and Clusters of Galaxies

A beautiful method for determining the mass distribution in galaxies and clusters of galaxies has been provided by the observation of gravitationally lensed images of background galaxies. In the case of clusters of galaxies, these consist of spectacular arcs about the central core of the cluster as well as distorted images of background galaxies caused by the individual galaxies in the cluster. Gravitational lensing has the potential to provide information about many key problems in the astrophysics of galaxies, clusters and larger-scale structures. It is therefore worthwhile studying some simple aspects of gravitational lensing and its applications.

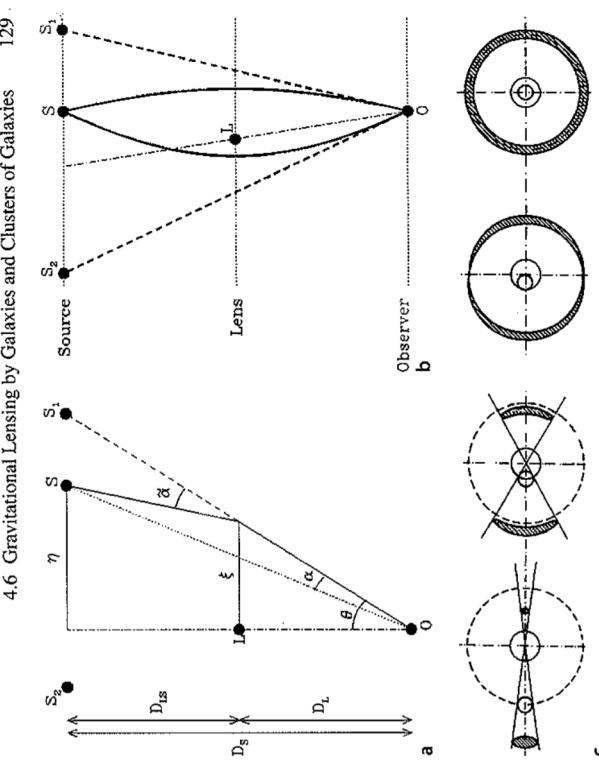
### 4.6.1 Basic Theory of Gravitational Deflections

Many of the most important results can be derived from the formula for the gravitational deflection of light rays by the Sun, first derived by Einstein in his great paper of 1915 on the General Theory of Relativity (Einstein, 1915). He showed that the deflection of light by a point mass  $M$  due to the bending of space-time amounts to precisely twice that predicted by a Newtonian calculation,

$$\tilde{\alpha} = \frac{4GM}{\xi c^2}, \quad (4.38)$$

where  $\xi$  is the ‘collision parameter’ (Fig. 4.18a). The angles in Fig. 4.18a have been exaggerated to illustrate the geometry of the deflection. For the very small deflections involved in the gravitational lens effect,  $\xi$  is almost exactly the distance of closest approach of the light ray to the deflector.

Chwolson in 1924 and Einstein in 1936 realised that, if a background star were precisely aligned with a deflecting point object, the gravitational deflection of the light rays would result in a circular ring, centred upon the deflector (Fig. 4.18c)



**Fig. 4.18.** **a** Illustrating the geometry of the deflection of light by a deflector, or lens, of mass  $M$  (Wambsganss, 1998). **b** Illustrating the two light paths from the source to the observer for a point mass (Wambsganss, 1998). **c** Illustrating the changes of the appearance of a compact background source as it passes behind a point mass. The dashed circles correspond to the Einstein radius. When the lens and the background source are precisely aligned, an Einstein ring is formed with radius equal to the Einstein radius  $\theta_E$

(Chwolson, 1924; Einstein, 1936). It is a straightforward calculation to work out the radius of what came to be known as an ‘Einstein ring’, although it should perhaps be known as a ‘Chwolson ring’. In the following analysis, we adopt the notation used by Wambsganss in his excellent online introduction to many aspects of gravitational lensing (Wambsganss, 1998). In Fig. 4.18a, the distance of the background source is  $D_S$  and that of the deflector, or lens,  $D_L$ , the distance between them being  $D_{LS}$ . Suppose the observed angular radius of the Einstein ring is  $\theta_E$ . Then, for a point source on-axis, since all the angles are small,

$$\theta_E = \tilde{\alpha} \left( \frac{D_{LS}}{D_S} \right) = \frac{4GM}{\xi c^2} \left( \frac{D_{LS}}{D_S} \right), \quad (4.39)$$

where  $\tilde{\alpha}$  is the deflection given by (4.38). Since  $\xi = \theta_E D_L$ ,

$$\theta_E^2 = \frac{4GM}{c^2} \left( \frac{D_{LS}}{D_S D_L} \right) = \frac{4GM}{c^2} \frac{1}{D}, \quad (4.40)$$

where  $D = (D_S D_L / D_{LS})$ . Thus, the *Einstein angle*  $\theta_E$ , the angle subtended by the Einstein ring at the observer, is given by the relation

$$\theta_E = \left( \frac{4GM}{c^2} \right)^{1/2} \frac{1}{D^{1/2}} \quad (4.41)$$

We have worked out this expression assuming the geometry of space is Euclidean. The above relation is also correct if the sources are at cosmological distances, provided the  $D$ s are *angular diameter distances* (Blandford and Narayan, 1992).<sup>2</sup>

Expressing the mass of the deflector in solar masses  $M_\odot$  and the distance  $D$  in Gpc ( $= 10^9$  pc  $= 3.056 \times 10^{25}$  m), we find

$$\theta_E = 3 \times 10^{-6} \left( \frac{M}{M_\odot} \right)^{1/2} \frac{1}{D_{\text{Gpc}}^{1/2}} \text{ arcsec.} \quad (4.42)$$

Thus, clusters of galaxies with masses  $M \sim 10^{15} M_\odot$  at cosmological distances can result in Einstein rings with angular radii tens of arcseconds. Such rings were first reported by Soucail and his colleagues and by Lynds and Petrosian (Lynds and Petrosian, 1986; Soucail et al., 1987). Beautiful examples of partial Einstein rings about the centre of the cluster Abell 2218 have been observed with the Hubble Space Telescope by Kneib, Ellis and their colleagues (Figs. 4.1 and 4.19). The rings are not complete and are elliptical rather than circular. The ellipticity and the incompleteness of the rings reflect the facts that the gravitational potential of the cluster is not precisely spherically symmetric and that the background galaxy and the cluster are not perfectly aligned.

#### 4.6.2 Magnification of Images by Gravitational Lensing

It is worth developing the theme of gravitational lensing a little further since the technique provides some of the most important information about the distribution of dark matter in the Universe is proving to be a powerful cosmological tool. In addition to the very accessible review by Wambsganss, the comprehensive discussion of all aspects of gravitational lensing presented in the volume *Gravitational Lensing: Strong, Weak and Micro* by Schneider, Kochanek and Wambsganss can be thoroughly recommended (Wambsganss, 1998; Schneider et al., 2006).

Let us first continue the analysis of the images of a background point source formed by a point mass deflector. We can first relate the angle of the observed image to its position in the absence of the deflector. From Fig. 4.18a, recalling that all the angles are very small,

$$\theta D_S = \beta D_S + \tilde{\alpha} D_{LS}. \quad (4.43)$$

Next, it is useful to introduce the *reduced deflection angle*  $\alpha(\theta)$  which is the deflection of the image of the background object because of the presence of the deflector as

<sup>2</sup> Angular diameter distances are introduced in Sect. 5.5.3. The formula for angular diameter distances between any two redshifts is derived in Sect. 7.5.

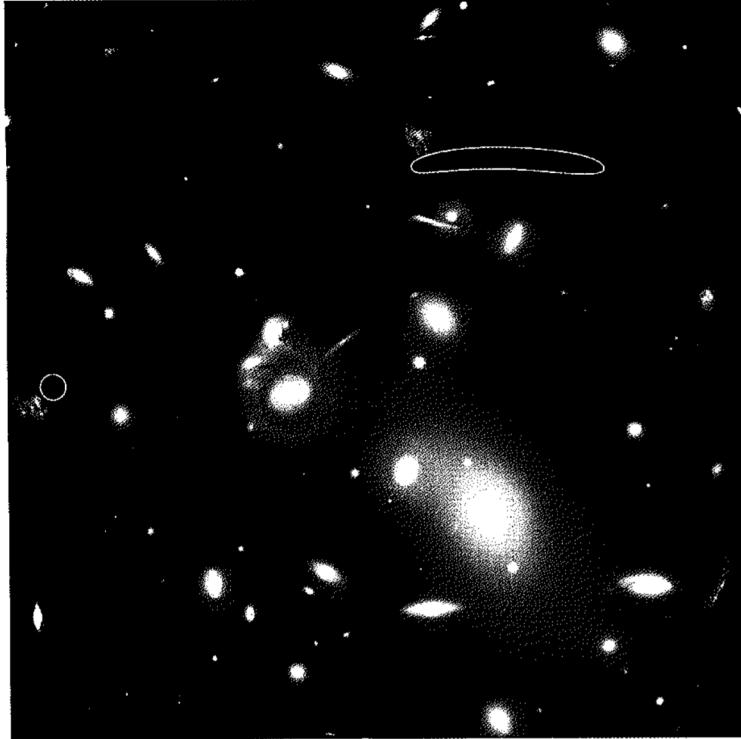


Fig. 4.19. The central region of the rich cluster of galaxies Abell 2218 observed by the Hubble Space Telescope at a wavelength of 840 nm. Several arcs can be observed more or less centred on the core of the Abell cluster. These are the gravitationally lensed images of background galaxies more or less perfectly aligned with the centre of the cluster. The circled segments of rings are the images of an extremely distant red galaxy (Courtesy of NASA, J.-P. Kneib, R. Ellis, and the Space Telescope Science Institute).

measured by the observer. From the geometry of Fig. 4.18a,

$$\beta = \theta - \alpha(\theta). \quad (4.44)$$

Inserting the value of  $\tilde{\alpha}$  into the expression  $\alpha(\theta) = (D_{LS}/D_S)\tilde{\alpha}$ , we find

$$\alpha(\theta) = (D_{LS}/D_S)\tilde{\alpha} = \left( \frac{D_{LS}}{D_S} \right) \frac{4GM}{\xi c^2} = \left( \frac{D_{LS}}{D_S D_L} \right) \frac{4GM}{\theta c^2} = \frac{\theta_E^2}{\theta}, \quad (4.45)$$

where we have used the expression (4.41) for the Einstein angle  $\theta_E$ . As a result, the *lens equation* (4.44) can be written

$$\beta = \theta - \frac{\theta_E^2}{\theta}. \quad (4.46)$$

The solutions of this quadratic equation give the two possible routes from the source to the observer:

$$\theta = \frac{\beta \pm \sqrt{\beta^2 + 4\theta_E^2}}{2}. \quad (4.47)$$

These light paths are illustrated in Fig. 4.18b. Notice that one of the angles is positive and the other negative, the small negative value corresponding to the longer route. The negative sign means that, if the object were extended, it would be mirror-inverted relative to the other image. Notice also that the images lie on opposite sides of the Einstein angle  $\theta_E$ .

As in the case of geometric optics, surface brightness is conserved on passing through a gravitational lens and so the images are magnified because of the changes in solid angle. Consider a small angular segment of an arc of azimuthal angle  $d\phi$ . In the absence of the lens, the solid angle subtended at the observer would be  $d\phi / \beta d\beta$ . Because of the effect of gravitational lensing, the angle  $\beta$  becomes  $\theta$  and  $d\beta$  becomes  $d\theta$  and so the solid angle becomes  $d\phi / \theta d\theta$ . Consequently, the magnification of the lensed images is

$$\mu = \frac{\theta}{\beta} \frac{d\theta}{d\beta}. \quad (4.48)$$

Finally, it is convenient to introduce a normalised impact parameter  $u$  in terms of the Einstein radius  $\theta_E$ ,  $u = \beta/\theta_E$ . After a little algebra, the magnifications of the two images can be written

$$\mu_i = \left(1 - \frac{\theta_E^4}{\theta_i^4}\right)^{-1} = \frac{1}{2} \pm \frac{u^2 + 2}{2u\sqrt{u^2 + 4}}. \quad (4.49)$$

Again, the magnification of the image outside the Einstein radius is positive and the other, inside the Einstein radius, is negative because of mirror-imaging. The sum of the magnitudes of the two image magnifications is the total magnification of the background point source:

$$\mu = |\mu_1| + |\mu_2| = \frac{u^2 + 2}{u\sqrt{u^2 + 4}}. \quad (4.50)$$

The reason for carrying out this calculation is that it is the expression which describes the characteristic signature of microlensing of background stars by objects in the halo of our Galaxy (see Sect. 4.7.1). A sketch of the gravitational deflections of a star of finite diameter by a point deflector is shown in Fig. 4.18c.

### 4.6.3 Extended Deflectors

The simplest generalisation of the above result is to consider the deflection due to a lens with an axially symmetric mass distribution along the line of sight. In that case, Schneider and his colleagues show that the deflection is given by the expression

$$\tilde{\alpha} = \frac{4GM(\leq \xi)}{\xi c^2}, \quad (4.51)$$

where  $M(\leq \xi)$  is the total projected mass within the radius  $\xi$  at the lens, a result corresponding to Gauss's theorem for Newtonian gravity.

We can derive from this result the necessary condition for the formation of a gravitationally lensed image about an object of mass  $M$  and radius  $R$ . For simplicity, let us suppose that the lens is a uniform disc of radius  $R$  and mass  $M$ . Then, using the result (4.51), the deflection for rays grazing the edge of the disc is

$$\tilde{\alpha} = \frac{4GM(< R)}{Rc^2} = \frac{4\pi G \Sigma}{c^2} R, \quad (4.52)$$

where we have introduced the surface density of the lens is  $\Sigma = M/\pi R^2$ . The deflection measured by the observer at the origin is, as before,

$$\alpha(\theta) = \frac{D_{LS}}{D_S} \tilde{\alpha} = \frac{D_{LS}}{D_S} \frac{4\pi G \Sigma}{c^2} R. \quad (4.53)$$

Let us now introduce a *critical surface density* defined by

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_S}{D_{LS} D_L} = \frac{c^2}{4\pi G D} \frac{1}{R}. \quad (4.54)$$

Then,

$$\alpha(\theta) = \frac{\Sigma}{\Sigma_{\text{crit}}} \frac{R}{D_L} = \frac{\Sigma}{\Sigma_{\text{crit}}} \theta. \quad (4.55)$$

Thus, if the surface density of the deflector is of the same order as the critical surface density, multiple images will be observed. The significance of the critical surface density can be appreciated by rewriting it in terms of the critical cosmological density which will be introduced in Sect. 7.2.2,  $\rho_c = 3H_0^2/8\pi G = 3H_0^2/8\pi G = 2 \times 10^{-26} \text{ kg m}^{-3}$ . Then,

$$\Sigma_{\text{crit}} \sim \rho_c \frac{c^2}{H_0^2 D}. \quad (4.56)$$

If the sources are at cosmological distances  $D \sim c/H_0$ , the critical surface density is

$$\Sigma_{\text{crit}} \sim \rho_c \frac{c}{H_0}. \quad (4.57)$$

Thus, for sources at cosmological distances, the critical surface density is roughly  $2h \text{ kg m}^{-2}$ . In fact, the Universe as a whole can be thought of as acting as a giant gravitational lens in understanding the effects of inhomogeneities upon the cosmological

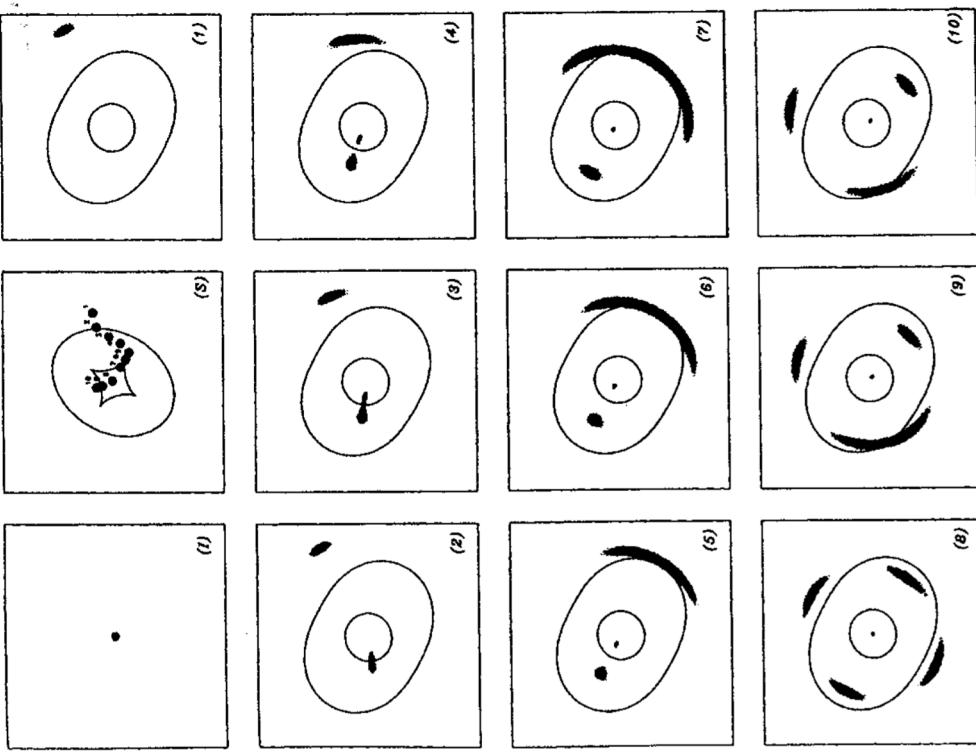


Fig. 4.20. The gravitational distortions of a background source (*Panel 1*) when it is located at different positions with respect to the axis of the gravitational lens. In this example, the lens is an ellipsoidal non-singular squeezed isothermal sphere. The ten positions of the source with respect to the critical inner and outer caustics are shown in the panel (*S*). The panels labelled (*1*) to (*10*) show the shapes of the images of the lensed source (Kneib, 1993). Note the shapes of the images when the source crosses the critical caustics. Positions (*6*) and (*7*) correspond to cusp catastrophes and position (*9*) to a fold catastrophe (Fort and Mellier, 1994).

redshift–angular diameter relation<sup>3</sup> (Dyer and Roeder, 1972, 1973; Dashhevsky and Zeldovich, 1964; Zeldovich, 1964).

Let us apply the result (4.51) to the case of an isothermal gas sphere, which provides a reasonable description of the mass distribution in clusters of galaxies. We consider the simple analytic solution (4.7), which has the unpleasant feature of being singular at the origin and of having infinite mass when integrated to an infinite distance, but these are unimportant for our present analysis. For this reason, this analysis is often referred to as the case of a *singular isothermal sphere*. Assuming that the velocity dispersion is isotropic and that  $\langle v_{\parallel}^2 \rangle$  is the observed velocity dispersion along the line of sight,

$$\varrho(r) = \frac{2}{Ar^2} \quad \text{where} \quad A = \frac{4\pi G \mu}{kT} = \frac{4\pi G}{\langle v_{\parallel}^2 \rangle}. \quad (4.58)$$

We now work out the projected mass density, or the surface density  $\Sigma(\xi)$ , at projected distance  $\xi$  by integrating along the line of sight, say, in the  $z$ -direction

$$\Sigma(\xi) = 2 \int_0^\infty \varrho(r) dz = 2 \int_0^{\pi/2} \varrho(r) \xi \sec^2 \theta d\theta \quad (4.59)$$

$$= \frac{\langle v_{\parallel}^2 \rangle}{\pi G} \frac{1}{\xi} \int_0^{\pi/2} d\theta = \frac{\langle v_{\parallel}^2 \rangle}{2G} \frac{1}{\xi}. \quad (4.60)$$

Therefore, the total mass within the distance  $\xi$  perpendicular to the line of sight at the deflector is

$$\int_0^\xi \Sigma(\xi) 2\pi \xi d\xi = \frac{\pi \langle v_{\parallel}^2 \rangle \xi}{G}. \quad (4.61)$$

The gravitational deflection of the light rays is therefore

$$\hat{\alpha} = \frac{4GM(<\xi)}{\xi c^2} = \frac{4\pi \langle v_{\parallel}^2 \rangle}{c^2}. \quad (4.62)$$

This is the remarkable result we have been seeking. For a singular isothermal gas sphere, the gravitational deflection is *independent* of the distance at which the light rays pass by the lens. We can therefore find the Einstein radius  $\theta_E$  directly from (4.39)

$$\theta_E = \frac{4\pi \langle v_{\parallel}^2 \rangle}{c^2} \frac{D_{LS}}{D_S} = 28.8 \langle v_{\parallel}^2 \rangle \frac{D_{LS}}{D_S} \text{ arcsec}, \quad (4.63)$$

where  $\langle v_{\parallel}^2 \rangle$  means the observed velocity dispersion of the galaxies in the cluster measured in units of  $10^3 \text{ km s}^{-1}$ . Fort and Mellier note that this is a rather robust expression for estimating the masses of clusters of galaxies (Fort and Mellier, 1994). They find that for a variety of plausible mass distributions the estimates agree to within about 10%.

<sup>3</sup> These relations are developed in Sect. 7.7.

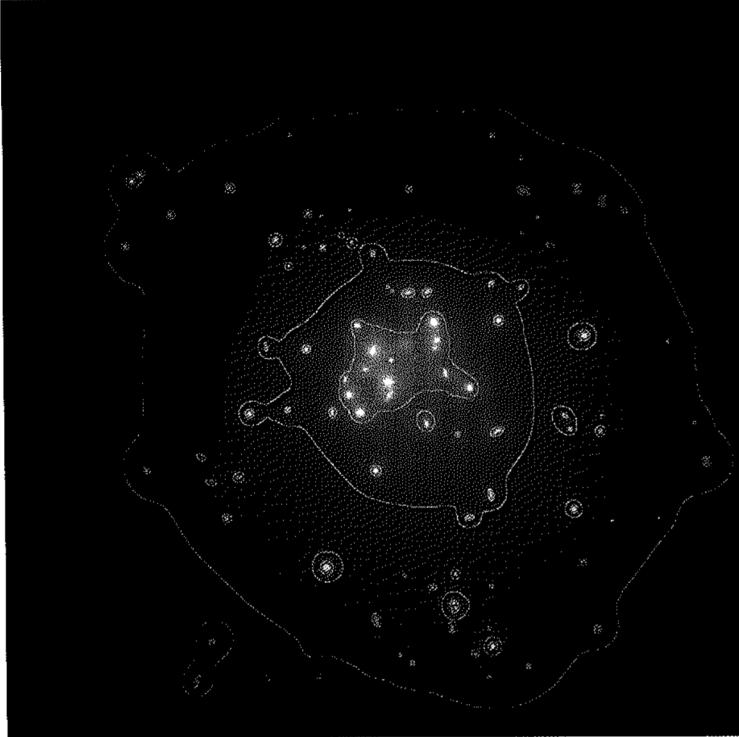
Strong lensing of background sources only occurs if they lie within the Einstein angle  $\theta_E$  of the axis of the lens. An excellent discussion of the shapes and intensities of the gravitationally distorted images of background sources for more general mass distributions is given by Fort and Mellier (Fort and Mellier, 1994). The gravitational lensing is not true lensing in the sense of geometric optics but rather the light rays come together to form caustics and cusps. Figure 4.20 shows the types of images expected for gravitational lensing by an ellipsoidal gravitational potential. The background source is shown in panel (1) and, in the second panel labelled (S), different positions of the background source with respect to the critical inner and outer caustic lines associated with the gravitational lens are shown. These are lines along which the lensed intensity of the image is infinite. The images labelled (1) to (10) show the observed images of the background source when it is located at the positions labelled on the second panel (S). The numbers and shapes of the images depend upon the location of the source with respect to the caustic surfaces. It can be seen that the predicted images resemble the arcs seen in Figs. 4.1 and 4.19.

For clusters of galaxies, these inferred masses are in good agreement with the values obtained by measuring the velocity dispersion of the cluster galaxies and with the X-ray methods of measuring total masses. An important aspect of the gravitational lensing approach to the determination of cluster masses is that it is possible to determine the details of the gravitational potential within the cluster. These studies show, for example, that the distribution of mass in Abell 2218 is more complex than a simple spherically symmetric distribution (Kneib et al., 1996). Another example of the power of the gravitational lensing technique is shown in Fig. 4.21 which was created by Tyson and his colleagues using a parametric inversion technique to analyse a very deep HST image of the cluster Cl 0024+1654 (Tyson et al., 1998). They find that, excluding mass concentrations centered on visible galaxies, more than 98% of the remaining mass is represented by a smooth concentration of dark matter centered on the brightest galaxies in the core of the cluster. With the availability of very deep cluster images in a number of wavebands taken with the Advanced Camera for Surveys (ACS) on the Hubble Space Telescope, the distortions of background objects by many galaxies in the clusters can be used to define detailed mass distributions (Zekser et al., 2005).

The extension of these techniques to the weak gravitational lensing of large samples of distant galaxies by large-scale structures has been developed by Kaiser who has shown how the distorted images can be used to determine the large-scale two-point correlation function for galaxies (Kaiser, 1992).

#### 4.6.4 Gravitational Lensing and the Astrophysics of Galaxies

The above discussion barely scratches the surface of what will undoubtedly be a major growth area in the physics of galaxies and clusters in the coming years. Gravitational lensing probes directly the total mass distribution, independent of the distribution of baryonic matter and so can be used to address a number of key



**Fig. 4.21.** The reconstructed mass distribution in the cluster Cl 0024+1654 which is dominated by the underlying dark matter distribution shown in orange. Mass associated with the galaxies is shown in blue. The contours show 0.5, 1 and 1.5 times the critical lensing column density. Tyson and his colleagues find the mass distribution is remarkably smooth once the mass associated with the galaxies is removed. They find a mass-to-luminosity ratio for the cluster of  $M/L = (276 \pm 40)h(M_\odot/L_\odot)$  (Tyson et al., 1998)

astrophysical questions. For example:

- What is the distribution of mass in the dark matter haloes of galaxies and clusters?
- What are the tidal radii of the mass distributions for galaxies, both in the general field and in the cores of clusters?
- What is the bias parameter for galaxies, meaning the ratio between the clustering amplitudes for the baryonic and dark matter?
- Is there structure in the distribution of dark matter within galaxies and clusters, or is it smooth?

These and many other issues are addressed in the comprehensive survey *Gravitational Lensing: Strong, Weak and Micro* by Schneider, Kochanek and Wambsganss (Schneider et al., 2006). Strong lensing effects such as those illustrated in Fig. 4.20 enable the mass distribution to be determined on the scale of the inner caustic surfaces but, in addition, weak lensing can be detected statistically to much larger radii. As can be seen from panels 1, 2 and 3 of Fig. 4.20, the gravitationally lensed images are predicted to be stretched tangentially to the line joining the lens to the background galaxy. Therefore, by measuring the orientations of the images of large numbers of background galaxies, the effects of weak gravitational lensing can be distinguished statistically from the intrinsic ellipticities of galaxies. Convincing evidence for the detection of a statistical weak lensing signature for galaxies was first discovered by Brainerd and her colleagues (Brainerd et al., 1996) and has since been repeated for very much larger samples of galaxies.

As Schneider emphasises in his review, galaxy-galaxy imaging may well provide the best constraints statistically on the dimensions of dark matter haloes. A good example of what has been achieved is provided by the Red-Sequencer Cluster Survey which involved  $\sim 1.2 \times 10^5$  lensing galaxies and  $\sim 1.5 \times 10^6$  fainter background galaxies in an area of 45.5 square degrees (Hoekstra et al., 2004). The lensing galaxies had medium redshift  $z \approx 0.35$  and the background galaxies  $z \approx 0.53$ . These data showed that the dark matter haloes were somewhat rounder than the light distribution of the galaxies. Interestingly, the analysis of the shear data on larger angular scales provided evidence for truncation of the isothermal density distribution at a radius of  $(185 \pm 30) h^{-1}$  kpc, one of the few direct estimates of the scale of the dark matter haloes.

A good example of the power of this technique is the determination of the mass distribution in a sample of 22 early-type galaxies which were imaged by the Advanced Camera for Surveys (ACS) of the Hubble Space Telescope (Gavazzi et al., 2007). In the central regions, the mass distributions were determined by optical spectroscopy and by strong gravitational lensing. In the outer regions, the statistical weak gravitational lensing technique enabled the mass profile to be determined out to about 300 kpc. Gavazzi and his colleagues found that the total mass density profile was consistent with that of an isothermal sphere,  $\rho \propto r^{-2}$ , over two decades in radius from  $(3\text{--}300) h^{-1}$  kpc, despite the fact that the inner regions are dominated by baryonic matter whilst the outer regions are dominated by dark matter. They found that the average stellar mass-to-light ratio was  $M_*/L_V = 4.48 \pm 0.46 h M_\odot/L_\odot$  while the overall average virial mass-to-light ratio was  $M_{\text{vir}}/L_V = 246^{+10}_{-87} h M_\odot/L_\odot$ .

An example of the use of weak gravitational lensing to estimate the bias parameter for large samples of galaxies has been carried out by Simon and his colleagues using the data from the Garching-Bonn Deep Survey (Simon et al., 2007). Typically, about  $10^5$  lensing galaxies were studied in three redshift intervals centred on  $z = 0.35, 0.47, 0.61$  and for each sample typically  $(1 - 3) \times 10^6$  background galaxies were observed. Similar bias parameters were found in the three redshift intervals, each about  $b = 0.8 \pm 0.1$  (see Sect. 14.4).

Another example of the power of gravitational lensing techniques is provided by the remarkable simulations of Wambsganss, which are contained in his online re-

view of gravitational lensing (Wambsganss, 1998). While the deflection data provide strong geometric constraints on the overall distribution of dark matter, the observed intensities of the multiple images of background quasars are more difficult to reconcile with the simplest smooth models. Wambsganss's splendid simulations illustrate the expected fluctuations in the observed intensities of the multiple images if the dark matter in the lensing galaxy consisted of large numbers moving lenses. Such fluctuations are observed in the relative brightnesses of the four components of the quadruple quasar Q2237+0305.

## 4.7 Forms of Dark Matter

One of the fundamental problems of cosmology is the unknown nature of the dark matter which is the dominant form of gravitating mass in the outer regions of large galaxies, in clusters of galaxies and other large-scale systems. This problem will haunt much of this book.

We are certain that the dark matter is present on the basis of the various arguments presented in Sects. 4.3 to 4.6, but we can only identify what it might be by standard astronomical techniques if it emits radiation or absorbs the radiation of background sources. An illustrative toy model is to suppose that the dark matter is in the form of standard bricks. There would have to be only one kilogram brick per cube of side roughly 500 million kilometres to attain the critical cosmological density  $\rho_c = 3H_0^2/8\pi G$ . If the bricks were uniformly distributed throughout the Universe, they would not obscure the most distant objects we can observe and they would be so cold that they would emit negligible amounts of far-infrared radiation. This example illustrates the point that there could be many forms of ordinary baryonic dark matter present in the Universe which would be very difficult to detect, even before we consider more exotic possibilities. Let us consider first the case of baryonic dark matter.

### 4.7.1 Baryonic Dark Matter

By *baryonic matter*, we mean ordinary matter composed of protons, neutrons and electrons and for convenience we will include the black holes in this discussion. As illustrated by the example of the bricks, certain forms of baryonic matter are very difficult to detect because they are very weak emitters of electromagnetic radiation. Important examples of such weak emitters are stars with masses  $M \leq 0.08 M_\odot$ , in which the central temperatures are not hot enough to burn hydrogen into helium – they are referred to collectively as *brown dwarfs*. They have no internal energy source and so the source of their luminosity is the thermal energy with which they were endowed at birth. There could be a small contribution from deuterium burning, but even this is not possible for stars with masses  $M \leq 0.01 M_\odot$ . Brown dwarfs are normally classified as inert stars with masses in the range  $0.08 \geq M \geq 0.01 M_\odot$ . Below that mass, they are normally referred to as planets,  $0.01 M_\odot$  corresponding to ten times the mass of Jupiter.

Until recently, brown dwarfs proved to be very difficult to detect. The situation has changed dramatically with a number technical advances in optical and infrared astronomy. The 2MASS infrared sky survey, which was conducted at a wavelength of 2  $\mu\text{m}$ , has discovered many cool brown dwarfs. The NICMOS infrared camera on the HST has discovered numerous brown dwarfs in nearby star clusters. The same techniques of high precision optical spectroscopy which has been spectacularly successful in discovering extrasolar system planets, has also been used to discover a number of brown dwarfs orbiting normal stars. Although the brown dwarfs are estimated to be about twice as common as stars with masses  $M \geq 0.08 M_{\odot}$ , they contribute very little to the mass density in baryonic matter as compared with normal stars because of their low masses. The consensus of opinion is that brown dwarfs could only make a very small contribution to the dark matter problem.

A strong limit to the total amount of baryonic matter in the Universe is provided by considerations of primordial nucleosynthesis. As will be shown in Sect. 10.4, the standard Big Bang model is remarkably successful in accounting for the observed abundances of light elements such as helium-4, helium-3, deuterium and probably lithium-7 through the process of *primordial nucleosynthesis*. An important consequence of that success story is that the primordial abundances of the light elements, particularly of deuterium and helium-3, are sensitive tracers of the mean baryon density of the Universe. Steigman has reviewed recent observational evidence on the primordial abundances of the light elements and compared these with the predictions of standard Big Bang nucleosynthesis (Steigman, 2006). He finds a best estimate of the mean baryon density of the Universe of  $\Omega_B h^2 = (0.0223 \pm 0.002)$ . Adopting a value of  $h = 0.7$  (see Sect. 8.3), the density parameter in baryonic matter is  $\Omega_B = 0.0455$ , compared with a mean density of matter in the Universe of  $\Omega_0 \approx 0.3$  (see Sect. 8.7). Thus, ordinary baryonic matter is only about one tenth of the total mass density of the Universe, most of which must therefore be in some non-baryonic form.

*Black holes* are another possible candidate for the dark matter. The supermassive black holes in the nuclei of galaxies have masses which are typically only about 0.1% of the mass of the bulges of their host galaxies and so they contribute negligibly to the mass density of the Universe. There might, however, be an invisible intergalactic population of massive black holes. Limits to the number density of such black holes can be set in certain mass ranges from studies of the numbers of gravitational lenses observed in large samples of extragalactic radio sources. In their VLA survey of a very large sample of extragalactic radio sources, designed specifically to search for gravitationally lensed structures, Hewitt and her colleagues set limits to the number density of massive black holes with masses in the range  $10^{10} \leq M \leq 10^{12} M_{\odot}$ . They found that the numbers correspond to  $\Omega_{BH} \ll 1$  (Hewitt et al., 1987). The same technique can be used to study the mass density of lower mass black holes by searching for the gravitationally lensed images on an angular scale of a milliarcsecond, corresponding to masses in the range  $10^6 \leq M \leq 10^8 M_{\odot}$  (Kassiola et al., 1991). Wilkinson and his colleagues searched a sample of 300 compact radio sources studied by VLBI techniques for examples of multiple gravitationally lensed images but none were found. The upper limit to the cosmological mass density of

intergalactic supermassive compact objects in the mass range  $10^6 \leq M \leq 10^8 M_{\odot}$  corresponds to less than 1% of the critical cosmological density (Wilkinson et al., 2001).

It cannot be excluded that the dark matter might consist of a very large population of very low mass black holes but these would have to be produced by a rather special initial perturbation spectrum in the very early Universe before the epoch of primordial nucleosynthesis. The fact that black holes of mass less than about  $10^{12} \text{ kg}$  evaporate by Hawking radiation on a cosmological time-scale sets a firm lower limit to the possible masses of mini-black holes which could contribute to the dark matter at the present epoch (Hawking, 1975).

An impressive approach to setting limits to the contribution which discrete low mass objects, collectively known as *MASSIVE Compact Halo Objects*, or MACHOs, could make to the dark matter in the halo of our own Galaxy, has been the search for gravitational microlensing signatures of such objects as they pass in front of background stars. The MACHOs include low mass stars, white dwarfs, brown dwarfs, planets and black holes. These events are very rare and so very large numbers of background stars have to be monitored. The beauty of this technique is that it is sensitive to MACHOs with a very wide range of masses, from  $10^{-7}$  to  $100 M_{\odot}$ , and so the contributions of a very wide range of candidates for the dark matter can be constrained. In addition, the expected light curve of such gravitational lensing events has a characteristic form which is given by the magnification relation (4.50) and which is independent of wavelength. The time-scale of the brightening is roughly the time it takes the MACHO to cross the Einstein radius of the dark deflector which is why a very wide range of masses can be constrained by this technique. Two very large projects, the MACHO and the EROS projects, have made systematic surveys over a number of years to search for these events. The MACHO project, which ran from 1992 to 1999 used stars in the Magellanic Clouds and in the Galactic bulge as background stars and millions of stars were monitored regularly (Alcock et al., 1993b). The first example of a microlensing event was discovered in October 1993 (Fig. 4.22), the mass of the invisible lensing object being estimated to lie in the range  $0.03 < M < 0.5 M_{\odot}$  (Alcock et al., 1993a).

By the end of the MACHO project, many lensing events had been observed, including over 100 in the direction towards the Galactic bulge, about three times more than expected. In addition, 13 definite and 4 possible events were observed in the direction of the Large Magellanic Cloud (Alcock et al., 2000). The numbers are significantly greater than the 2–4 detections expected from known types of star. The technique does not provide distances and masses for individual objects, but, interpreted as a Galactic halo population, the best statistical estimates suggest that the mean mass of these MACHOs is between  $0.15$ – $0.9 M_{\odot}$ . The statistics are consistent with MACHOs making up about 20% of the necessary halo mass, the 95% confidence limits being 8–50%. Somewhat fewer microlensing events were detected in the EROS project which found that less than 25% of the mass of the standard dark matter halo could consist of dark objects with masses in the range  $2 \times 10^{-7}$  to  $1 M_{\odot}$  at the 95% confidence level (Afonso et al., 2003). The most likely candidates for the MACHOs observed by the MACHO project would appear to be white dwarfs which

velocities – they remained ‘cold’. Their rest mass energies are expected to lie in the range  $10^{-2}$  to  $10^{-5}$  eV. The role of such particles in cosmology and galaxy formation is discussed by Efstathiou and Kolb and Turner (Efstathiou, 1990; Kolb and Turner, 1990).

**Neutrinos with finite rest mass.** A second possibility is that the three known types of neutrino have finite rest masses. Laboratory tritium  $\beta$ -decay experiments have provided an upper limit to the rest mass of the electron antineutrino of  $m_\nu \leq 2$  eV (Weinheimer, 2001), although the particle data book suggests a conservative upper limit of 3 eV (see <http://www-pdg.lbl.gov/pdg.html>). This measurement does not exclude the possibility that the two other types of neutrino, the  $\mu$  and  $\tau$  neutrinos, could have greater masses. However, the discovery of neutrino oscillations has provided a measurement of the mass difference between the  $\mu$  and  $\tau$  neutrinos of  $\Delta m^2_\nu \sim 3 \times 10^{-3}$  (Englert et al., 2003; Al'iu et al., 2005). Thus, although their masses are not measured directly, they probably have masses of the order of 0.1 eV. The reason that these values are of interest is that neutrinos of rest mass about 10–20 eV would be enough to provide the critical density, as may be appreciated from the following calculation. The number density of neutrinos of a single type in thermal equilibrium at temperature  $T$  is

$$N = \bar{N} = \frac{4\pi g}{h^3} \int_0^\infty \frac{p^2 dp}{e^{E/kT} + 1} = 0.091 \left( \frac{2\pi kT}{hc} \right)^3 m^{-3}, \quad (4.64)$$

where the statistical weight  $g$  for the neutrinos is  $g = 1$ . If there are  $N_\nu$  neutrino types present, each with rest mass  $m_\nu$ , the present mass density of neutrinos in the Universe would be

$$\rho_\nu = N N_\nu m_\nu. \quad (4.65)$$

The present temperature of the neutrino background radiation, which was in equilibrium with the matter prior to the epoch when the neutrinos decoupled, is  $(4/11)^{1/3}$  of the temperature of the Cosmic Microwave Background Radiation, that is,  $T_\nu = 1.94$  K, and so the value of  $N$  is rather precisely known (see Sect. 10.5). Therefore, if  $\rho_\nu$  is to equal the critical density of the Universe  $\rho_c = 1.88 \times 10^{-30} h^2 \text{ kg m}^{-3}$ , the necessary rest mass energy of the neutrino is  $m_\nu = 184 h^2 / N_\nu$  eV. Since there are three neutrino species, each with its antiparticle,  $N_\nu = 6$  and hence the necessary rest mass of the neutrino is  $31 h^2$  eV. Taking  $h = 0.7$ , it follows that, if the neutrino rest mass were about 15 eV, known types of neutrino could close the Universe. However, if the mass of the neutrinos is of the order 0.1 eV, they certainly could not account for the amount of dark matter present in the Universe.

**WIMPs.** A third possibility is that the dark matter is in some form of *Weakly Interacting Massive Particle*, or WIMP. This might be the gravitino, the supersymmetric partner of the graviton, or the photino, the supersymmetric partner of the photon, or some form of as yet unknown massive neutrino-like particle. The possible existence of these types of unknown particles represents theoretical extrapolations beyond the

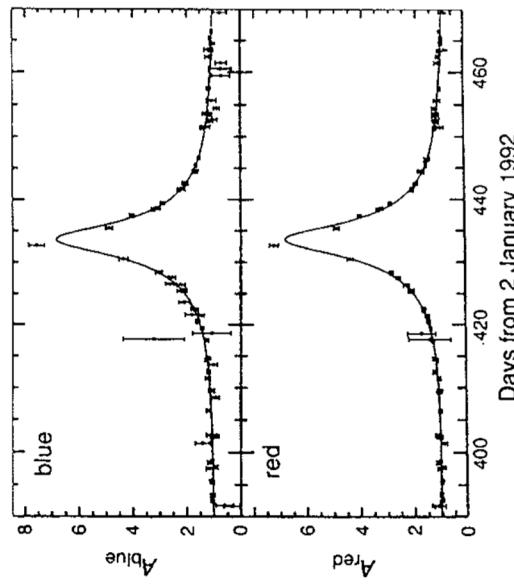


Fig. 4.22. The gravitational microlensing event recorded by the MACHO project in February and March 1993. The horizontal axis shows the date in days measured from day zero on 2 January 1992. The vertical axis shows the amplification of the brightness of the lensed star relative to the unlensed intensity in blue and red wavebands. The solid lines show the expected variations of brightness of a lensed star with time (Eq. 4.50). The same characteristic light curve is observed in both wavebands, as expected for a gravitational microlensing event (Alcock et al., 1993b)

would have to be produced in large numbers in the early evolution of the Galaxy, but other more exotic possibilities cannot be excluded. The consensus view is that MACHOs alone cannot account for all the dark matter in the halo of our Galaxy and so some form of non-baryonic matter must make up the difference.

#### 4.7.2 Non-Baryonic Dark Matter

The general consensus is that the dark matter is most likely to be some non-baryonic form and so is of the greatest interest for particle physicists since it may consist of the types of particles predicted by theories of elementary particles but not yet detected experimentally. Three of the most popular possibilities are discussed in the following paragraphs.

**Axions.** The smallest mass candidates are the *axions* which were invented by particle theorists in order to ‘save quantum chromodynamics from strong CP violation’. If they exist, they must have been created when the thermal temperature of the Universe was about  $10^{12}$  K but they were out of equilibrium and never acquired thermal

range of energies which have been explored experimentally, but these ideas are sufficiently compelling on theoretical grounds that many particle theorists take seriously the possibility that cosmological studies will prove to be important in constraining theories of elementary processes at very high energies.

In fact, there is the real possibility that clues will be found from experiments to be carried out in the TeV energy range with the Large Hadron Collider (LHC) and the next generation International Linear Collider (ILC). An understanding of TeV physics will undoubtedly impact our understanding of the Universe back to epochs  $t \sim 10^{-8}$  seconds. But there are even more exciting possibilities. To paraphrase the generic arguments given by Trodden, physics beyond the standard model of particles physics is essential and almost any model involves new particles at the TeV scale (Trodden, 2006). These particles are related to the particles of the standard model through an extra new symmetry which is necessary to avoid proton decay and the violation of precision tests of electro-weak theory. This new symmetry leads to the expectation of some new stable particle at the weak energy scale. It is particularly intriguing that the weak interaction cross-section is of exactly the right order of magnitude at the decoupling mass-scale to provide sufficient mass density in WIMPs to account for the dark matter (see Sect. 10.6).

### 4.7.3 Astrophysical and Experimental Limits

Useful astrophysical limits can be set to the number densities of different types of neutrino-like particles in the outer regions of giant galaxies and in clusters of galaxies. The WIMPs and massive neutrinos are collisionless fermions and therefore there are constraints on the phase space density of these particles, which translate into a lower limit to their masses. This is because, for a given momentum, only a finite number of particles within a given volume is allowed. Let us give a simple derivation of this result. More details of this calculation are given by Tremaine and Gunn, who provide a slightly tighter constraint on the masses of these hypothetical particles (Tremaine and Gunn, 1979).

Being fermions, neutrino-like particles are subject to the Pauli exclusion principle according to which there is a maximum number of particle states in phase space for a given momentum  $p_{\max}$ . The elementary phase volume is  $\hbar^3$  and, recalling that there can be two particles of opposite spin per state, the maximum number of particles with momenta up to  $p_{\max}$  is

$$N \leq 2 \frac{g}{\hbar^2} \frac{4\pi}{3} p_{\max}^3, \quad (4.66)$$

per unit volume, where  $g$  is the statistical weight of the neutrino species. If there is more than one neutrino species present, we multiply this number by  $N_\nu$ . Bound gravitating systems such as galaxies and clusters of galaxies are subject to the virial theorem according to which the kinetic energy of the particles which make up the system is equal to half of its gravitational potential energy (Sect. 3.5.1). If  $\sigma$  is the root-mean-square velocity dispersion of the objects which bind the system,

$\sigma^2 = GM/R$ , and the maximum velocity which particles within the system can have is the escape velocity from the cluster,  $v_{\max} = (2GM/R)^{1/2} = \sqrt{2}\sigma$ . The neutrino-like particles bind the system and so its total mass is  $M = N N_\nu m_\nu$  where  $m_\nu$  is the rest mass of the particle. We therefore find the following lower limit to the rest mass of the particle from (4.66) in terms of observable quantities:

$$m_\nu^4 \geq \left( \frac{9\pi}{8\sqrt{2}g} \right) \frac{\hbar^3}{N_\nu G\sigma R^2} \quad m_\nu \geq \frac{1.5}{(N_\nu \sigma_3 R_{\text{Mpc}}^2)^{1/4}} \text{ eV}, \quad (4.67)$$

where the velocity dispersion  $\sigma_3$  is measured in units of  $10^3 \text{ km s}^{-1}$  and  $R$  is measured in Mpc.

Let us insert typical values for the velocity dispersions and radii of the systems in which there is known to be a dark matter problem. In clusters of galaxies, typical values are  $\sigma = 1000 \text{ km s}^{-1}$  and  $R = 1 \text{ Mpc}$ . If there is only one neutrino species,  $N_\nu = 1$ , we find  $m_\nu \geq 1.5 \text{ eV}$ . If there were six neutrino species, namely, electron, muon, tau neutrinos and their antiparticles,  $N_\nu = 6$  and then  $m_\nu \geq 0.9 \text{ eV}$ . For giant galaxies, for which  $\sigma = 300 \text{ km s}^{-1}$  and  $R = 10 \text{ kpc}$ ,  $m_\nu \geq 20 \text{ eV}$  if  $N_\nu = 1$  and  $m_\nu \geq 13 \text{ eV}$  if  $N_\nu = 6$ . For small galaxies, for which  $\sigma = 100 \text{ km s}^{-1}$  and  $R = 1 \text{ kpc}$ , the corresponding figures are  $m_\nu \geq 80 \text{ eV}$  and  $m_\nu \geq 50 \text{ eV}$  respectively. Thus, particles with rest masses  $m_\nu \sim 1 \text{ eV}$  could bind clusters of galaxies but they could not bind the haloes of giant or small galaxies.

There is a further constraint on the possible masses WIMPs. Studies of the decay of the  $W^\pm$  and  $Z^0$  bosons at CERN have shown that the width of the decay spectrum is consistent with there being only three neutrino species with rest mass energies less than about 40 GeV. Therefore, if the dark matter is in some form of ultraweakly interacting particle, its rest mass energy must be greater than 40 GeV.

Another important constraint is that, if the masses of the particles were greater than 15 eV and they are as common as neutrinos and photons, as expected in the standard Big Bang model, the present density of the Universe would exceed the critical mass density  $\rho_c$  (see Sect. 4.7.2). Therefore there would have to be some suppression mechanism to ensure that, if  $m \geq 40 \text{ GeV}$ , these particles are very much less common than the photons and electrons neutrinos at the present day. This could take place by the mechanism described in Sect. 10.6.

The search for evidence for different types of dark matter particles has developed into one of the major areas of the discipline known as *astroparticle physics*. An important class of experiments involves the search for weakly interacting particles with masses  $m \geq 1 \text{ GeV}$ , which could make up the dark halo of our Galaxy. In order to form a bound dark halo about our Galaxy, the particles would have to have velocity dispersion  $\langle v^2 \rangle^{1/2} \sim 230 \text{ km s}^{-1}$  and their total mass is known. Therefore, the number of WIMPs passing through a terrestrial laboratory each day is a straightforward calculation. When these massive particles interact with the sensitive volume of the detector, the collision results in the transfer of momentum to the nuclei of the atoms of the material of the detector and this recoil can be measured in various ways. There is a small temperature increase which can be measured in a cryogenically cooled detector, or the ionisation caused by the recoiling nucleus

can be measured in an ionisation chamber, or the light emitted by the passage of the recoil nucleus through the material detected by a scintillation detector. The challenge is to detect the very small number of events expected because of the very small cross-section for the interaction of WIMPs with the nuclei of atoms. A typical estimate is that less than one WIMP per day would be detectable by 1 kilogram of detector material. These are very demanding experiments and they have to be located deep underground to avoid contamination by cosmic rays and must be heavily shielded against natural radioactivity in the surrounding rocks. Such experiments have been carried out in deep underground laboratories such as those at Gran Sasso in Italy, the Soudan Underground Laboratory in Minnesota, USA and the Boulby Underground Laboratory in Yorkshire, England.

A good example of the quality of the data now available is provided by the results of the Cryogenic Dark Matter Search (CDMS) at the Soudan Laboratory. The CDMS experiment has set a 90% confidence upper limit to the spin-independent WIMP–nucleon interaction cross-section at its most sensitive mass of  $60 \text{ GeV}/c^2$  of  $\sigma_w \leq 1.6 \times 10^{-47} \text{ m}^2$  (Akerib et al., 2006). This cross-section can be compared with the weak interaction cross-section for neutrino–electron scattering,  $\sigma = 3 \times 10^{-49} (E/m_e c^2) \text{ m}^2$ . Already the CDMS result constrains the predictions of supersymmetric models. The sensitivity of these experiments should be improved by an order of magnitude with the CDMSII experiment planned for 2007 and then by successive order of magnitude improvements through the different phases of the SuperCDMS proposal.

## The Basic Framework