1997RvMA...10...61B

On Arcs in X–Ray Clusters

Ludwig Biermann Award Lecture
Matthias Bartelmann
MPI für Astrophysik
Karl-Schwarzschild-Straße 1, D–85748 Garching

1 Introduction

Galaxies are not randomly distributed in the sky. Their positions are correlated, and there exist areas in the sky where the galaxy density is noticeably higher or lower than average. There are groups consisting of a few galaxies, and there exist *clusters of galaxies* in which some hundred up to a thousand galaxies appear very close together.

Zwicky noted in 1933 that the galaxies in the Virgo cluster and other rich clusters move so fast that the clusters required about 10 to 100 times more mass to keep the galaxies bound than could be accounted for by the luminous galaxies themselves. This was the earliest indication that there is invisible mass, or dark matter, in at least some objects in the universe.

Several thousands of galaxy clusters are known today. Abell's (1958) cluster catalog lists 2712 clusters north of $-20°$ declination and away from the Galactic plane. Employing a less restrictive definition of galaxy clusters, the catalog by Zwicky et al. (1961–1968) identifies 9134 clusters north of $-3°$ declination. Cluster masses exceed $\sim 5 \times 10^{14}\,M_\odot$, and they have typical radii of $\sim 1.5\,\mathrm{Mpc}$.

When X–ray telescopes became available after 1966, it was discovered that clusters are powerful X–ray emitters. Their X–ray luminosities fall within $10^{43} - 10^{45}\,\mathrm{erg\,s^{-1}}$, rendering galaxy clusters the most luminous X–ray sources in the sky. Improved X–ray telescopes revealed that the source of X–ray emission in clusters is extended rather than point-like, and that the X–ray spectra are best explained by thermal *bremsstrahlung* from a hot, dilute plasma with temperatures in the range $10^7 - 10^8\,\mathrm{K}$ and densities of $\sim 10^{-3}$ particles per cm^3. Based on the assumption that this intracluster gas is in hydrostatic equilibrium with the gravitational potential of the total cluster matter, the X–ray temperature and flux can be used to estimate the cluster mass. Typical results *approximately* (i.e. up to a factor of ~ 2) agree with the mass estimates from the kinematics of cluster galaxies. The mass of the intracluster gas amounts to about 10% of the total cluster mass. The X–ray emission thus independently confirms the existence of dark matter in galaxy clusters. Sarazin (1986) reviews clusters of galaxies focusing on their X–ray emission.

Later, luminous arc-like features were discovered in two galaxy clusters (Lynds & Petrosian 1986; Soucail et al. 1987a,b). Their light is typically bluer than that of the cluster galaxies, and their length is comparable to the size of the central cluster region. Paczyński (1987) suggested that these so-called *arcs* are images of galaxies in the background of the clusters which are strongly distorted by the gravitational lens effect close to the cluster centers. This explanation was generally accepted after spectroscopy revealed that the arc sources are much more distant than the clusters in which they appear (Soucail et al. 1988).

Large arcs require special alignment of the arc source with the lensing cluster. At larger distance from the cluster center, images of background galaxies are only weakly deformed, and they are referred to as *arclets* (Tyson 1988; Fort et al. 1988; Tyson, Valdes & Wenk 1990). The high number density of faint arclets allows to measure the coherent distortion caused by the gravitational tidal field of a cluster out to fairly large radii.

These gravitational lens effects offer the possibility to detect and measure the entire cluster mass, dark and luminous, without referring to any equilibrium or symmetry assumptions like those required for the mass estimates from galactic kinematics or X–ray emission. For a review on arcs and arclets in galaxy clusters, see Fort & Mellier (1994).

Apart from being spectacular objects on their own, clusters are also of particular interest for cosmology. Being the largest gravitationally bound entities in the cosmos, they represent the high-mass end of collapsed structures. Their number density, their individual properties, and their spatial distribution constrain the power spectrum of the density fluctuations from which the structure in the universe is believed to have originated (e.g. Viana & Liddle 1996; Eke, Cole & Frenk 1996). Their formation history is sensitive to the parameters that determine the geometry of the universe as a whole. If the matter density in the universe is high, clusters tend to form later in the cosmic history than when the matter density is low (first noted by Richstone, Loeb & Turner 1992). Consequently, the compactness and the morphology of clusters are affected by the cosmic matter density, and this has various observable implications.

This lecture deals with the question what we can learn about galaxy clusters and the universe by investigating their gravitational lens effects and their X–ray emission. It describes strong and weak gravitational lensing by clusters, and discusses the relation of strong lensing effects to the X–ray emission. I discuss aspects of the cosmological relevance of galaxy clusters in Sect. 2. The principles of gravitational lens theory are reviewed in Sect. 3 as far as they are necessary for the present purpose. Section 3 further describes strong and weak lensing by clusters. The X–ray emission and its relation to strong gravitational lensing are the subject of Sect. 4. The lecture concludes with an outlook and a discussion of future prospects in Sect. 5.

2 Cosmological Importance of Galaxy Clusters

As mentioned in the introduction, clusters of galaxies were first identified as regions in the sky where the density of galaxies is significantly higher than average. Abell (1958) compiled a catalog of clusters found on the photographic plates of the Palomar Sky Survey. He defined clusters by three criteria: (1) the cluster contain at least 50 galaxies in the magnitude range $m_3 \le m \le m_3 + 2$, where m_3 is the magnitude of the third brightest cluster galaxy; (2) these galaxies be contained within a circle of estimated radius $R_\mathrm{a} = 1.5\,h^{-1}\,\mathrm{Mpc}$, where h is the Hubble constant in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and (3) the estimated cluster redshift be in the range $0.02 \le z \le 0.20$. R_a is called the Abell radius. Abell assigned richness classes and distance classes to the clusters. The richness class is defined by the number of galaxies within the Abell radius, and the distance class is a measure for the magnitude of the tenth brightest cluster galaxy.

Zwicky et al. (1961–1968) used less restrictive criteria to define galaxy clusters: (1) a contour (an isopleth) is determined within which the galaxy surface density is twice the background density; (2) within this isopleth, there be 50 cluster galaxies in the magnitude range $m_1 \le m \le m_1 + 3$, with m_1 the magnitude of the brightest cluster galaxy.

The velocity dispersion σ_v of the cluster galaxies is observed to be $\sim 900\,\mathrm{km\,s^{-1}}$, averaged over moderately rich and rich Abell clusters. Assuming that the galaxies are in virial equilibrium with the cluster potential, the velocity dispersion and the Abell radius can be converted to a mass estimate,

$$M_\mathrm{vir} \simeq \frac{2\sigma_v^2 R_\mathrm{a}}{G} \simeq 5.6 \times 10^{14}\,h^{-1}\,M_\odot\ . \tag{1}$$

The redshift range of the Abell catalog, $0.02 \le z \le 0.2$, encloses $\sim 3.5 \times 10^8\,h^{-3}\,\mathrm{Mpc}^3$. This number is insensitive to the choice of the cosmic parameters because the upper redshift limit is small, and hence all distances are well described by the Hubble law, which does not depend on cosmic parameters. Of the 2712 clusters in Abell's catalog, 818 fall into (the poorest) richness class 0. Excluding those, there are 1894 clusters with richness class ≥ 1 in that volume, which yields an estimate for the spatial cluster density of

$$n_\mathrm{c} \simeq \frac{1894}{3.5 \times 10^8}\,h^3\,\mathrm{Mpc}^{-3} \simeq 5 \times 10^{-6}\,h^3\,\mathrm{Mpc}^{-3}\ . \tag{2}$$

More careful determinations arrive at similar numbers (e.g. Bahcall 1988).

In the course of their formation, clusters need to collect cosmic material from a large volume. The radius $\bar{R}$ of that volume is determined by

$$M_\mathrm{vir} \overset{!}{=} \frac{4\pi}{3}\,\bar{R}^3\,\rho_\mathrm{cr}\Omega_0\ , \tag{3}$$

or

$$\bar{R} \simeq 7.8\,h^{-1}\,\Omega_0^{-1/3}\,\mathrm{Mpc}\ , \tag{4}$$

where $\rho_\mathrm{cr} \simeq 2 \times 10^{-29}\,\mathrm{g\,cm^{-3}}$ is the critical density of the universe. This radius defines a typical scale for the linear theory of cluster formation. In linear theory, density fluctuations grow in place. If one considers the gravitational collapse of a spherical overdense region in an expanding universe and compares it to the evolution of its density contrast in linear theory, the sphere reaches its maximum radius (the "turnaround") before collapse when its linear density contrast is $\delta_\mathrm{c} = 1.686$. Therefore, linear theory predicts that clusters can form where the linear density contrast averaged over a scale of $\bar{R}$ exceeds δ_c.

Assuming a Gaussian distribution of the density contrast, we have

$$p(\delta,\bar{R}) = \frac{1}{\sqrt{2\pi}\sigma_{\bar{R}}} \exp\left(-\frac{\delta^2}{2\sigma_{\bar{R}}^2}\right) , \tag{5}$$

where the notation indicates that the width of the distribution depends on the scale $\bar{R}$ over which the density contrast is averaged. The variance $\sigma_{\bar{R}}$ on the scale $\bar{R}$ is determined by the power spectrum of the density fluctuations. Press & Schechter (1974) assumed that the fraction of the universe that is contained in collapsed objects characterized by their linear scale $\bar{R}$ is

$$F_c = \int_{\delta_c}^{\infty} \mathrm{d}\delta\, p(\delta,\bar{R}) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_{\bar{R}}}\right) . \tag{6}$$

We can now determine the variance $\sigma_{\bar{R}}$ of the cosmic density fluctuations on the linear scale of galaxy clusters. According to the masses and the number density of clusters, the fraction of the universe that is observed to be contained in clusters is

$$F_c' = \frac{M_{\rm vir}\, n_c}{\rho_{\rm cr}\Omega_0} \simeq 8\times10^{-3}\,\Omega_0^{-1} . \tag{7}$$

Thus clusters are rare objects, especially if the density of the universe is high, $\Omega_0 \simeq 1$. Equating the observed cluster fraction with that expected from eq. (6), and solving for $\sigma_{\bar{R}}$, we find

$$\sigma_{\bar{R}} = \begin{cases} 0.70 & (\Omega_0 = 1) \\ 0.96 & (\Omega_0 = 0.2) \end{cases} . \tag{8}$$

In order to reproduce the observed cluster abundance, the power spectrum of the density fluctuations must therefore be normalized such that its average over scales smaller than $\bar{R}$ is $\sigma_{\bar{R}}$. Note that the scale $\bar{R}$ varies with Ω_0 as in eq. (4). To compare the amplitudes of the power spectrum for different choices of Ω_0, we need to transform $\sigma_{\bar{R}}$ to a common scale, which is traditionally chosen as $8\,h^{-1}$ Mpc. This transformation depends somewhat on the shape of the power spectrum. For a CDM power spectrum, we find

$$\sigma_8 = \begin{cases} 0.68 & (\Omega_0 = 1) \\ 1.38 & (\Omega_0 = 0.2) \end{cases} . \tag{9}$$

This fairly rough estimate is confirmed by the more detailed studies of Viana & Liddle (1996); Eke, Cole & Frenk (1996); and Pen (1996).

The spatial number density of clusters together with their typical mass thus fixes the amplitude of the power spectrum of density fluctuations in the universe. For the CDM model, the required amplitude is lower by more than a factor of two in a high-density than in a low-density universe.

2.1 Timescales for Cluster Formation

Being rare objects, clusters arise from density perturbations well in the tail of the Gaussian distribution. Since the variance $\sigma_{\bar{R}}$ of this Gaussian grows in time, the

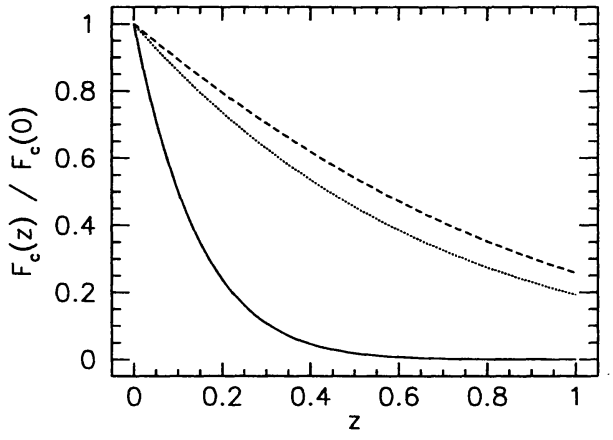

Figure 1: Cluster fraction as a function of redshift, normalized to unity at the present epoch. The curves show the result for three pairs of cosmological parameters: $\Omega_0 = 1$, $\Omega_\Lambda = 0$ (solid curve); $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$ (dotted curve); $\Omega_0 = 0.2$, $\Omega_\Lambda = 0.8$ (dashed curve). The figure shows that clusters form late, in particular when Ω_0 is high.

cluster abundance (6) drops rapidly backwards in time. Also, peculiar velocities of galaxies are typically of order a few hundred kpc Gyr^{-1}. Since the cluster material is collected from a region of order a few Mpc across, this process takes of order 10 Gyr to complete, which is comparable to the age of the universe. Therefore, clusters generally tend to form late in the cosmic history.

Since the cluster fraction scales with Ω_0^{-1}, clusters are less rare when Ω_0 is low. Hence, in a low-density universe clusters tend to form earlier than in a high-density universe. This was first noted by Richstone, Loeb & Turner (1992), and their argument was modified for self-consistency and extended towards non-spherical collapse models by Bartelmann, Ehlers & Schneider (1993). They give an expression for the time evolution of the cluster abundance and show that it is valid not only for spherically symmetric, but also for anisotropic collapse. This result, normalized to unity at the present epoch, is illustrated in Fig. 1. For the figure, the power spectrum of the density fluctuations was normalized such that the present cluster number density is reproduced.

The figure confirms that clusters form late for all choices of the cosmological parameters. Their formation is particularly delayed when Ω_0 is high. While half of the clusters that exist today had already formed at redshifts between 0.4 and 0.6 for $\Omega_0 = 0.2$, with only little influence of the cosmological constant, it takes until redshift $z \sim 0.1$ for half the cluster population to form if $\Omega_0 = 1$.

2.2 Cluster Structure and Cosmological Parameters

Real clusters do not form from isolated spherical density perturbations. Rather, they are assembled in subsequent mergers of protoclusters into one object which grows in time, and they continuously accrete matter from their surroundings (Lacey

& Cole 1993, 1994). During most of their history, clusters therefore are objects which show remains of subclumps which have merged with them. The matter distribution in their neighborhood is inhomogeneous and thus exerts tidal forces on the clusters, which can additionally delay the relaxation of a forming cluster to dynamical equilibrium. A simple timescale argument illustrates this point.

Typical galaxy velocities in clusters are of order 900 km s^{-1}, and typical cluster diameters are of order $2\,R_{\rm a} = 3\,h^{-1}$ Mpc. The time it takes a galaxy to cross a cluster is therefore of order $3 \times 10^9\,h^{-1}$ yr. This provides a dynamical timescale for a galaxy cluster. It was shown in Fig. 1 that, if $\Omega_0 = 1$, cluster formation sets in only after $z \sim 0.4$. The cosmic time that has passed since then is only $\sim 4 \times 10^9$ yr, so that even the older clusters in a high-density universe were formed only about one dynamical timescale ago. In a low-density universe, the cosmic timescale is longer, and cluster formation sets in earlier, but still the cosmic time is not much longer than the dynamical time scale of the clusters.

Therefore, it appears implausible that clusters could have relaxed and formed smooth, spherically symmetric bodies in the short time they had after their formation. Remnants of merger and accretion events should be seen in many of them.

Numerical simulations seemed to indicate that clusters were significantly less structured in low-density than in high-density universes, reflecting the longer time they had to relax (Evrard et al. 1993). It turned out later that in these simulations, boundary conditions were set up such that accretion and infall were artificially suppressed for low Ω_0. More recent simulations involving a consistent treatment of the boundaries of the simulation volume arrive at the conclusion that the differences in cluster substructure for different values of Ω_0 are less pronounced than suspected earlier (Jing et al. 1995). In any case, clusters are frequently observed to have evident structure (e.g. Geller & Beers 1982; Dressler & Shectman 1988; West & Bothun 1990; Jones & Forman 1992, and references therein; Böhringer 1993; Mohr, Fabricant & Geller 1993). It is, however, unclear at the moment whether the differences in the abundance and the amount of substructure are sufficient to distinguish between different cosmological models (Jing et al. 1995).

3 Clusters as Gravitational Lenses

Being pronounced density inhomogeneities, clusters must deflect light passing them closely on the way from the sources to the observer. Gravitational light deflection gives rise to magnification and distortion effects. Depending on whether these effects are strong or weak, two distinct kinds of such gravitational lensing effects can arise. Strong lensing leads to the formation of *arcs*, strongly distorted images of galaxies on the far side of the clusters. Weak lensing causes coherent distortion patterns which can be observed by analyzing the shapes of samples of faint and distant galaxies. After a brief introduction to the theory of gravitational lensing, I discuss both effects and show what kind of information can be obtained from them. For general reference on gravitational lensing, see Schneider, Ehlers & Falco (1992); Blandford & Narayan (1992); Fort & Mellier (1994); and Narayan & Bartelmann (1996).

3.1 Introduction to Gravitational Lens Theory

The propagation of light in arbitrary curved spacetimes is in general a complicated theoretical problem. However, for almost all cases of relevance to gravitational lensing, we can assume that the overall geometry of the universe is well described by the Friedmann-Lemaître-Robertson-Walker metric and that the matter inhomogeneities which cause the lensing are no more than local perturbations. Light paths propagating from the source past the lens to the observer can then be broken up into three distinct zones. In the first zone, light travels from the source to a point close to the lens through unperturbed spacetime. In the second zone, near the lens, light is deflected. Finally, in the third zone, light again travels through unperturbed spacetime. To study light deflection close to the lens, we can assume a locally flat, Minkowskian spacetime which is weakly perturbed by the Newtonian gravitational potential of the mass distribution constituting the lens. This approach is legitimate if the Newtonian potential Φ is small, $|\Phi| \ll c^2$, and if the peculiar velocity v of the lens is small, $v \ll c$.

These conditions are satisfied in virtually all cases of astrophysical interest. Consider a galaxy cluster at redshift ~ 0.3 which deflects light from a source at redshift ~ 1. The distances from the source to the lens and from the lens to the observer are ~ 1 Gpc, or about three orders of magnitude larger than the diameter of the cluster. Thus zone 2 is limited to a small local segment of the total light path. The relative peculiar velocities in a galaxy cluster are $\sim 10^3\ \mathrm{km\ s^{-1}} \ll c$, and the Newtonian potential is $|\Phi| < 10^{-4}\,c^2 \ll c^2$, in agreement with the conditions stated above.

In view of these simplifications, we can describe light propagation close to gravitational lenses in a locally Minkowskian spacetime perturbed by the gravitational potential of the lens to first post-Newtonian order. The effect of spacetime curvature on the light paths can then be expressed in terms of an effective index of refraction n, which is given by (e.g. Schneider et al. 1992)

$$n = 1 - \frac{2}{c^2}\,\Phi = 1 + \frac{2}{c^2}\,|\Phi| \,. \tag{10}$$

As in normal geometrical optics, light propagation can now be described by Fermat's principle. Application of Fermat's principle to a point mass M shows that a light ray passing the point mass at distance b is deflected by the angle

$$\hat{\alpha} = \frac{4GM}{c^2 b} \,. \tag{11}$$

Most of the light deflection occurs at a distance from the lens that is typically much smaller than the distances between observer and lens and between lens and source. The lens can therefore be considered thin compared to the total extent of the light path. The mass distribution of the lens can then be projected along the line-of-sight and be replaced by a mass sheet orthogonal to the line-of-sight. The plane of the mass sheet is commonly called the lens plane. The mass sheet is characterized by its surface mass density

$$\Sigma(\vec{\xi}) = \int \rho(\vec{\xi}, z)\,\mathrm{d}z \,, \tag{12}$$

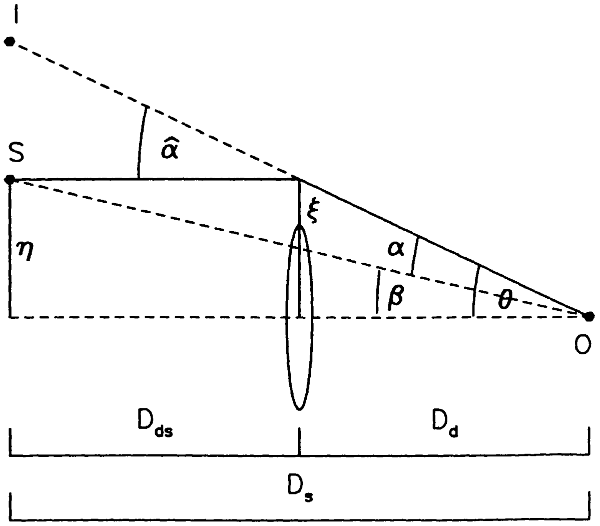

Figure 2: Illustration of a gravitational lens system. The light ray propagates from the source S at transverse distance η from the optic axis to the observer O, passing the lens at transverse distance ξ. It is deflected by an angle $\hat{\alpha}$. The angular separations of the source and the image from the optic axis as seen by the observer are β and θ, respectively. The reduced deflection angle α and the actual deflection angle $\hat{\alpha}$ are related by eq. (14). The distances between the observer and the source, the observer and the lens, and the lens and the source are D_s, D_d, and D_ds, respectively.

where $\vec{\xi}$ is a two-dimensional vector in the lens plane. The deflection angle at position $\vec{\xi}$ is the sum of the deflections due to all the mass elements in the plane:

$$\vec{\hat{\alpha}}(\vec{\xi}) = \frac{4G}{c^2} \int \frac{(\vec{\xi}-\vec{\xi}')\Sigma(\vec{\xi}')}{|\vec{\xi}-\vec{\xi}'|^2} \mathrm{d}^2\xi' \, . \tag{13}$$

The geometry of a typical gravitational lens system is shown in Fig. 2. A light ray from a source S is deflected by the angle $\vec{\hat{\alpha}}$ at the lens and reaches an observer O. The angle between the (arbitrarily chosen) optic axis and the true source position is $\vec{\beta}$, and the angle between the optic axis and the image I is $\vec{\theta}$. The distances between observer and lens, lens and source, and observer and source are D_d, D_ds, and D_s, respectively.

It is now convenient to introduce the reduced deflection angle

$$\vec{\alpha} = \frac{D_\mathrm{ds}}{D_\mathrm{s}} \vec{\hat{\alpha}} \, . \tag{14}$$

From Fig. 2 we see that $\theta D_\mathrm{s} = \beta D_\mathrm{s} + \hat{\alpha} D_\mathrm{ds}$. Therefore, the positions of the source and the image are related through the simple equation

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) \, . \tag{15}$$

Equation (15) is called the *lens equation*, or ray-tracing equation. It is nonlinear in general, and so it is possible to have multiple images $\vec{\theta}$ corresponding to a single source position $\vec{\beta}$. As Fig. 2 shows, the lens equation is trivial to derive and requires merely that the following Euclidean relation should exist between the angle enclosed by two lines and their separation,

$$\text{separation} = \text{angle} \times \text{distance}\,. \tag{16}$$

It is not obvious that the same relation should hold in curved spacetime. However, if the distances $D_{\mathrm{d,s,ds}}$ are *defined* such that eq. (16) holds, then the lens equation must obviously be true. Distances so defined are called *angular-diameter distances*, and eq. (15) is valid only when these distances are used. Note that in general $D_{\mathrm{ds}} \neq D_{\mathrm{s}} - D_{\mathrm{d}}$.

3.1.1 Effective Lensing Potential

We now introduce the *critical surface mass density*

$$\Sigma_{\mathrm{cr}} = \frac{c^2}{4\pi G}\,\frac{D_{\mathrm{s}}}{D_{\mathrm{d}} D_{\mathrm{ds}}} \simeq 0.35\,\mathrm{g\,cm^{-2}} \left(\frac{D}{1\,\mathrm{Gpc}}\right)^{-1}\,, \tag{17}$$

where the effective distance D is the combination of distances

$$D = \frac{D_{\mathrm{d}} D_{\mathrm{ds}}}{D_{\mathrm{s}}}\,. \tag{18}$$

A lens with a constant surface mass density of Σ_{cr} focuses perfectly. Let us further define a scalar potential ψ as

$$\psi(\vec{\theta}) = \frac{1}{\pi} \int \kappa(\vec{\theta}')\,\ln|\vec{\theta} - \vec{\theta}'|\,\mathrm{d}^2\theta'\,, \tag{19}$$

where the *convergence* κ is the surface mass density divided by Σ_{cr}. Then, the deflection angle is the gradient of ψ,

$$\vec{\alpha}(\vec{\theta}) = \vec{\nabla}\psi = \frac{1}{\pi} \int \kappa(\vec{\theta}')\,\frac{\vec{\theta} - \vec{\theta}'}{|\vec{\theta} - \vec{\theta}'|^2}\,\mathrm{d}^2\theta'\,, \tag{20}$$

and ψ and the convergence κ are related through the Poisson equation

$$\kappa = \frac{1}{2}\Delta\psi\,. \tag{21}$$

The local properties of the lens mapping are described by its Jacobian matrix $\mathcal{A}$,

$$\mathcal{A} \equiv \frac{\partial\vec{\beta}}{\partial\vec{\theta}} = \left(\delta_{ij} - \frac{\partial\alpha_i(\vec{\theta})}{\partial\theta_j}\right) = \left(\delta_{ij} - \frac{\partial^2\psi(\vec{\theta})}{\partial\theta_i\partial\theta_j}\right)\,. \tag{22}$$

As a consequence of Liouville's theorem, lensing conserves surface brightness. However, it changes the solid angle under which an object appears, and thus leads to magnification and distortion effects. The local solid-angle distortion is given by the determinant of $\mathcal{A}$. A solid-angle element $\delta\beta^2$ of the source is mapped to the solid-angle element of the image $\delta\theta^2$, and so the magnification is given by

$$\frac{\delta\theta^2}{\delta\beta^2} = (\det \mathcal{A})^{-1} . \tag{23}$$

Equation (22) shows that the Hessian matrix of the potential ψ describes the deviation of the lens mapping from the identity mapping. For convenience, we introduce the abbreviation

$$\frac{\partial^2\psi}{\partial\theta_i\partial\theta_j} \equiv \psi_{ij} . \tag{24}$$

Since the Laplacian of ψ is twice the convergence, we have

$$\kappa = \frac{1}{2}\,(\psi_{11} + \psi_{22}) = \frac{1}{2}\ \mathrm{tr}\,\psi_{ij} . \tag{25}$$

Two additional linear combinations of ψ_{ij} are important, and these are the components of the *shear* tensor,

$$\gamma_1(\vec\theta) = \frac{1}{2}\,(\psi_{11} - \psi_{22}) \ , \quad \gamma_2(\vec\theta) = \psi_{12} = \psi_{21} . \tag{26}$$

With these definitions, the Jacobian matrix reads

$$\mathcal{A} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} . \tag{27}$$

The meaning of the terms convergence and shear now becomes intuitively clear. Convergence acting alone causes an isotropic focusing of light bundles, leading to an isotropic magnification of a source. The source is mapped onto an image with the same shape but larger size. Shear introduces anisotropy (or astigmatism) into the lens mapping; the quantity $\gamma = (\gamma_1^2 + \gamma_2^2)^{1/2}$ describes the magnitude of the shear. A circular source of unit radius becomes, in the presence of both κ and γ, an elliptical image with major and minor axes

$$(1-\kappa-\gamma)^{-1} , \quad (1-\kappa+\gamma)^{-1} . \tag{28}$$

The magnification is

$$\mu = (\det \mathcal{A})^{-1} = [(1-\kappa)^2 - \gamma^2]^{-1} . \tag{29}$$

Note that the Jacobian $\mathcal{A}$ is in general a function of position $\vec\theta$.

Points $\vec\theta$ in the lens plane where the Jacobian matrix is singular, $\det \mathcal{A} = 0$, are called critical points. They form closed curves, the *critical curves*. The image curves of the critical curves in the source plane are called *caustics*. Formally, point sources on caustics are infinitely magnified. This does not pose any problem in reality because astrophysical sources are not point-like, and the average magnification of extended sources is finite. Furthermore, even if the source was point-like, its magnification remained finite because then wave-optical effects became relevant. Figure 3 displays the critical curves and caustics of a simple lens model and shows the image shapes of sources close to the caustics. The figure also illustrates the distinction between *tangential* and *radial* critical curves and caustics in terms of the preferred direction of image elongation. Caustics divide the source plane into regions with different image numbers. A source outside all caustics has one image, and the image number changes by two upon each caustic crossing.

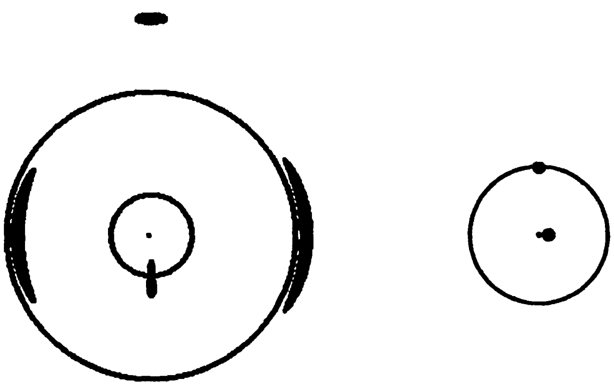

Figure 3: Imaging of an extended source by a non-singular circularly-symmetric lens. A source close to the point caustic at the lens center produces two tangentially oriented arc-like images close to the outer critical curve, and a faint image at the lens center. A source on the outer caustic produces a radially elongated image on the inner critical curve, and a tangentially oriented image outside the outer critical curve. Because of these image properties, the outer and inner critical curves are called *tangential* and *radial*, respectively.

3.2 Strong Lensing by Clusters — Giant Arcs

The subject of gravitational lensing by galaxy clusters entered the observational realm with the discovery of large and luminous *arcs* in the galaxy clusters A 370 and Cl 2244 (Lynds & Petrosian 1986; Soucail et al. 1987a,b). Paczyński (1987) proposed that the arcs are the images of background galaxies which are strongly distorted and elongated by the gravitational lens effect of the foreground cluster. This explanation was confirmed when the first arc redshifts were measured and found to be significantly greater than that of the clusters (Soucail et al. 1988).

Figure 4 illustrates that extended sources like galaxies produce large arcs if they lie on top of caustics. The largest arcs are formed from sources on cusp points, because then three images of a source merge to form the arc (cf. the right panel in Fig. 4). At the so-called "lips" and "beak-to-beak" caustics, which are related to cusps, similarly large arcs are formed. Sources on a fold caustic give rise to two rather than three merging images and thus form moderate arcs.

3.2.1 Efficiency of Arc Formation by Clusters

We have seen before (Fig. 1) that clusters tend to form late in the history of the universe. They are expected in appreciable numbers only at fairly low redshift, in particular if the density of the universe is high, $\Omega_0 \lesssim 1$. As for any other lens, the strength of their gravitational lens effects depends on the geometry of the system containing the source, the cluster, and the observer. This dependence is characterized by the effective distance (18). It peaks at roughly half the angular-diameter distance between the source and the observer, and it drops to zero when the sources are shifted towards the observer, or when the lens is shifted towards

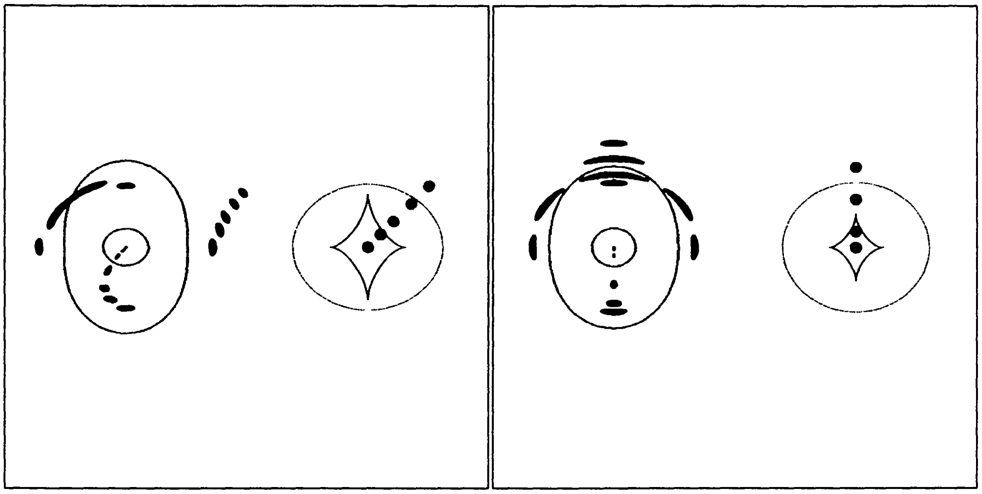

Figure 4: Compact source moving away from the center of an elliptical lens. Left panel: source crossing a fold caustic; right panel: source crossing a cusp caustic. In each panel, the diagram on the left shows critical lines and image positions and the diagram on the right shows caustics and source positions.

the sources. For sources at redshift $z_s \sim 1$, lenses are most efficient at redshift $z_d \sim 0.35$, with only slight dependence on the cosmological parameters. Especially if $\Omega_0 \lesssim 1$, there exist very few clusters at that most efficient redshift, and at low redshift where the number density of clusters is higher, their lens efficiency is largely suppressed by the geometry of the lens system.

To obtain a simple estimate for the dependence on cosmological parameters of the cluster efficiency to form large arcs, consider the product of spatial cluster density, effective lensing distance, and physical volume V, which is proportional to

$$\frac{\mathrm{d}N_{\mathrm{lens}}(z)}{\mathrm{d}z} = F_{\mathrm{c}}(z)\, D(z, z_{\mathrm{s}})\, \frac{\mathrm{d}V(z)}{\mathrm{d}z}\,. \tag{30}$$

By definition, $\mathrm{d}N_{\mathrm{lens}}/\mathrm{d}z$ provides an estimate for the number of efficient lensing clusters per redshift interval $\mathrm{d}z$. All three factors in eq. (30) depend on the cosmological parameters. Figure 5 shows $\mathrm{d}N_{\mathrm{lens}}/\mathrm{d}z$ as a function of redshift z for three choices of the cosmological parameters Ω_0 and Ω_Λ.

Figure 5 shows that the number of efficient strong cluster lenses per redshift interval is expected to increase dramatically when the cosmic density is lowered. This results from two effects. First, for low Ω_0, clusters form earlier, so that there exist more clusters at the redshift where lenses are most efficient. Second, the cosmic volume per redshift interval is larger for lower Ω_0 and higher Ω_Λ than in the Einstein-de Sitter case ($\Omega_0 = 1$, $\Omega_\Lambda = 0$). Apart from the change in the expected number of efficient cluster lenses with the cosmic parameters, the expected average redshift of lensing clusters is changed. While it is ~ 0.2 for $\Omega_0 = 1$, it increases to $\sim 0.5-0.6$ for $\Omega_0 = 0.2$. These effects were calculated in more detail by Bartelmann (1993).

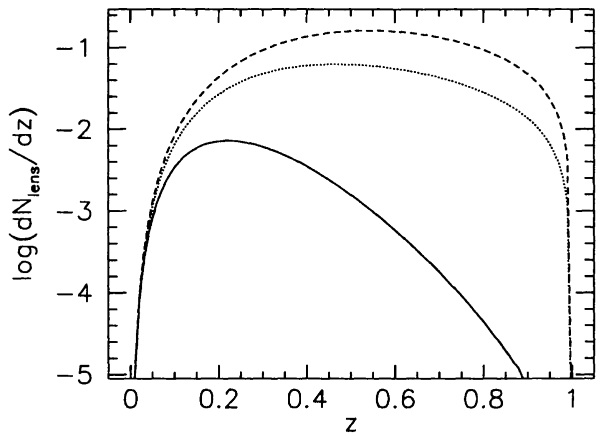

Figure 5: Estimate $\mathrm{d}N_{\mathrm{lens}}/\mathrm{d}z$ for the number of efficient strong cluster lenses per redshift interval $\mathrm{d}z$, as defined in eq. (30). The three curves are for three choices for the cosmological parameters: $\Omega_0 = 1$, $\Omega_\Lambda = 0$ (solid curve); $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$ (dotted curve); $\Omega_0 = 0.2$, $\Omega_\Lambda = 0.8$ (dashed curve). Note the logarithmic scale of the ordinate. The curves were calculated for sources at redshift $z_{\mathrm{s}} = 1$.

3.2.2 Asphericity of Cluster Mass

The fact that the observed giant arcs never have counter-arcs of comparable brightness, and rarely have even small counter-arcs, implies that the lensing geometry has to be non-spherical (Grossman & Narayan 1988; Kovner 1989; see also Figs. 3 and 4. Cluster potentials therefore must have substantial quadrupole and perhaps also higher multipole moments.

Large deviations of the lensing potentials from spherical symmetry also help increase the probability of producing large arcs. Bergmann & Petrosian (1993) argued that the apparent abundance of large arcs relative to small arcs and arclets can be reconciled with theoretical expectations if aspheric lens models are taken into account. Bartelmann & Weiss (1994) and Bartelmann, Steinmetz & Weiss (1995) showed that the probability for large arcs can be increased by more than an order of magnitude if aspheric cluster models with significant substructure are used instead of smooth spherically symmetric models.

The latter two studies investigated the strong lensing effects of numerically simulated clusters. The clusters formed by gravitational collapse of Gaussian initial density fluctuations with CDM perturbation spectrum. Particular care was taken in the simulations to ensure that the tidal field of the matter surrounding the clusters was properly taken into account. The simulated clusters were then used to study their gravitational lens effects on a population of extended sources with random positions and ellipticities at redshift $z_{\mathrm{s}} = 1$. The images were automatically classified as to their length, width, curvature radius and so forth. In total, the properties of $\simeq 73,000$ simulated arcs were analyzed.

The primary results are cross sections of the numerically simulated clusters for the formation of large arcs with specified properties. The cross section for

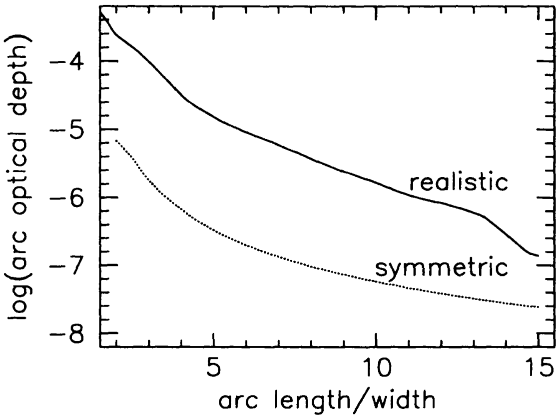

Figure 6: Optical depths for the formation of arcs as a function of the length-to-width ratio of the arcs. The solid curve was obtained from numerically simulated clusters, the dotted curve from spherically symmetric clusters with the same mass. Note the logarithmic scale of the ordinate.

a particular image configuration is defined as the area in the source plane within which a source must be located in order to be lensed as such an image. The fraction of the source sphere covered by the arc cross sections of individual clusters then equals the probability for a source to be imaged as an arc. This probability is also called the *optical depth* for arc formation. The cross sections and optical depths were compared to those of spherically symmetric clusters with the same masses as the numerically simulated clusters. An example for optical depths is given in Fig. 6.

Figure 6 shows that the simulated clusters, whose mass distribution is generally highly asymmetric, produce an optical depth for the formation of arcs with large length-to-width ratio that is higher by more than an order of magnitude compared to clusters with the same masses and core radii, but spherically symmetric mass distribution. This pronounced effect has two principal reasons.

First, large arcs occur close to critical curves where $\det\mathcal{A} \simeq 0$. According to eq. (29), convergence κ and shear γ act together in $\det\mathcal{A}$. If the shear vanished, the critical curve was the contour $\kappa = 1$. With increasing shear, the critical curve moves into regions where $\kappa < 1$. The shear γ quantifies the gravitational tidal field of the mass distribution. Asymmetries in the mass distribution increase the tidal field and the shear compared to a spherically symmetric mass distribution, and hence the critical curve is at lower convergence κ than in the spherically symmetric case. Since cluster mass profiles on average drop outward, this implies that the critical curve is then at larger distance from the cluster center, and hence the region is increased where arcs can be formed.

Second, the largest arcs are formed from sources close to cusps in the caustic curves, because there three images of the source merge rather than two, as illustrated in the right panel of Fig. 4. Spherically symmetric lenses cannot produce cusps because cusps single out preferred directions in the source plane, which are not compatible with spherical symmetry. Cusps therefore require asymmetric lensing mass distributions. Generally, the less symmetric a lens is, the more cusps it can produce, and the higher is the probability to form very long and thin arcs.

The primary result of the studies by Bartelmann & Weiss (1994) and Bartelmann et al. (1995) is therefore that accounting for cluster asymmetries and substructure is crucial for understanding the observed frequency of large arcs. The mere fact that clusters are frequently observed to be strongly substructured thus indicates that their strong lensing effects cannot properly be described by families of spherically symmetric mass distributions.

This suggests that the statistics of observed arcs and the global geometry of the universe may be related. As discussed above, clusters form the later the higher the matter density of the universe is. They then have less time to relax to symmetric mass distributions. Thus, a higher degree of cluster substructure is expected in a high-density compared to a low-density universe. Since arc statistics is so sensitive to cluster substructure, one would expect from this reasoning that arcs should be much more abundant in a high-density than in a low-density universe.

However, we must account for another effect. As argued in Sect. 2, the amplitude of the power spectrum must be smaller if Ω_0 is higher to reproduce the observed current number density of clusters. In that case, the effect illustrated in Fig. 5 becomes more important. If $\Omega_0 = 1$, most clusters are at such low redshift that they fail to be efficient lenses. More recent simulations (Huss, Bartelmann & Colberg, in preparation), in which the power spectrum is normalized to the correct cluster abundance, reveal that in a high-density universe, virtually no clusters should exist that are able to form arcs, while arcs can be formed in universes with lower Ω_0 (see also the discussion in Sect. 5). Hence, the normalization of the power spectrum is crucial for any conclusions on the geometry of the universe drawn from arc statistics.

Figure 1 shows that half of the clusters existing today should have formed at redshift ~ 0.5 if $\Omega_0 \simeq 0.2$. At that redshift, their lens efficiency is high, as Fig. 5 indicates. Since cluster formation is still going on at that epoch, pronounced substructure should be frequent in these clusters. Their lens efficiency is thus higher by roughly an order of magnitude compared to spherically symmetric, relaxed systems, as shown in Fig. 6. In other words, if the normalization of the density-fluctuation power spectrum to the present cluster abundance is correct, the arc abundance indicates that Ω_0 is less than unity, and the arc formation efficiency is boosted by the larger tidal fields of the asymmetric matter distribution of the clusters in formation. In turn, this would predict a fairly high average redshift of such clusters which show large arcs, $z \sim 0.5 - 0.6$.

3.2.3 Cluster Mass Inside a Giant Arc

The location of an arc in a cluster provides a simple way to estimate the projected cluster mass within a circle traced by the arc. For a circularly symmetric lens, the average surface mass density $\langle\Sigma\rangle$ within the tangential critical curve equals the critical surface mass density $\Sigma_{\rm cr}$. Tangentially oriented large arcs occur approximately at the tangential critical curves. The radius $\theta_{\rm arc}$ of the circle traced by the arc therefore gives an estimate of the Einstein radius $\theta_{\rm E}$ of the cluster. Thus we have

$$\langle \Sigma(\theta_{\rm arc}) \rangle \approx \langle \Sigma(\theta_{\rm E}) \rangle = \Sigma_{\rm cr} \;, \tag{31}$$

and we find for the mass enclosed by $\theta = \theta_{\rm arc}$

$$M(\theta) = \Sigma_{\rm cr}\,\pi\,(D_{\rm d}\theta)^2 \approx 1.1 \times 10^{14}\,M_\odot \left(\frac{\theta}{30''}\right)^2 \left(\frac{D}{1\,{\rm Gpc}}\right) \;. \tag{32}$$

The mass estimate (32) is based on very simple assumptions. It can be improved by modeling the arcs with parameterized lens mass distributions and carrying out more detailed fits of the observed arcs (an excellent example for this method is the study by Kneib et al. 1993 on Abell 370). It is instructive to see how inaccurate the mass estimate (32) can become in non-symmetric clusters. Bartelmann (1995a) used numerical cluster models for comparing the true cluster masses to simple cluster mass estimates obtained from the location of large arcs. For that particular simulated cluster sample, the masses derived from (32) are *systematically* too high by $\simeq 60\%$ on average, and by even a factor of two in $\simeq 20\%$ of the cases. The basic physical reason for this systematic deviation is again that the tidal field of asymmetric clusters is stronger than that of symmetric clusters. Since the lensing effects are determined by the combination of surface mass density and its tidal field, the cluster mass required to form large arcs can be lower in the presence of asymmetries than in the symmetric case.

3.2.4 Core Radii

If a cluster is able to produce large arcs, its surface-mass density in the core must be approximately supercritical, $\Sigma \gtrsim \Sigma_{\rm cr}$. Narayan, Blandford, & Nityananda (1984) argued that cluster mass distributions need to have smaller core radii than those derived from optical and X-ray observations if they are to produce strong gravitational lens effects. This has been confirmed by many later efforts to model giant arcs. In virtually every case the core radius estimated from lensing is significantly smaller by a factor of $\sim 2 - 10$ than the estimates from optical and X-ray data. Results on lens-derived core radii fall in the range $20\,h^{-1}\,{\rm kpc} \lesssim r_{\rm core} \lesssim 50\,h^{-1}\,{\rm kpc}$.

Statistical analyses based on spherically symmetric cluster models lead to similar conclusions. Miralda-Escudé (1992, 1993) argued that cluster core radii can hardly be larger than the curvature radii of large arcs. Wu & Hammer (1993) claimed that clusters have to have density profiles that are either singular or much steeper than isothermal in order to reproduce the observed abundance of large arcs. Although this conclusion can substantially be altered once deviations from spherical symmetry are taken into account (Bartelmann et al. 1995), it remains true that small cores with $r_{\rm core} \lesssim 50\,h^{-1}\,{\rm kpc}$ are required in all observed arc clusters. Cores of this size can also be reconciled with large-arc statistics.

Interestingly, there are at least two observations which appear to indicate that cluster cores, although small, should be finite. Fort et al. (1992) discovered a radial arc near the center of MS 2137−23, and Smail et al. (1996) found a radial arc in A 370. To produce a radial arc with a softened isothermal sphere model, the core radius has to be roughly equal to the distance between the cluster center and the radial arc (cf. Fig. 3). Mellier, Fort & Kneib (1993) find $r_{\rm core} \gtrsim 20\,h^{-1}$ kpc in MS 2137−23, and Smail et al. (1996) infer $r_{\rm core} \sim 25\,h^{-1}$ kpc in A 370.

These results, however, have to be interpreted with caution because it may well be that the commonly used softened isothermal sphere model is inadequate to describe the interiors of galaxy clusters. While this particular model indeed requires core radii on the order of the radial critical radius, other lens models can produce radial arcs without having a flat core, and there are even singular density profiles which can explain radial arcs (Miralda-Escudé 1995; Bartelmann 1996). Such singular profiles for the dark matter are consistent with the fairly large core radii inferred from the X-ray emission of clusters, if the intracluster gas is isothermal and in hydrostatic equilibrium with the dark-matter potential (Navarro, Frenk & White 1996).

Bartelmann (1996) presented a proof that the type of dark-matter profile first found in simulations by Navarro et al. (1996) necessarily produces a tangential and a radial critical curve despite its central singularity. The two parameters that characterize the profile, viz. its total mass and its concentration, can be constrained through observations of clusters that contain a radial and a tangential arc. Definitive conclusions require that the redshifts of both arcs be known. Two clusters have been discovered so far that contain a radial and a tangential arc, namely Abell 370 and MS 2137. In both cases, the redshift of the radial arc is yet unknown. One can then constrain the redshift of the arcs by requiring that the parameters of the observed cluster should match the parameters of numerically simulated clusters, in order to see whether numerically simulated clusters are compatible with those observed. Predictions for the redshifts of the radial arcs in Abell 370 and in MS 2137 are given in Bartelmann (1996). If future spectroscopy should reveal different redshifts, the numerically simulated cluster models would be falsified. This provides an important test of our understanding of cluster formation because the profile discovered by Navarro et al. (1996) was recently also found to be independent of the perturbation spectrum of the cosmic density fluctuations and of cosmic parameters (Cole & Lacey 1996; Huss et al. 1996). It therefore seems to reflect a generic and universal property of galaxy clusters which, if true, provides a valuable hint as to the dynamical processes that define the shapes of galaxy clusters.

3.3 Weak Lensing by Clusters — Arclets

In addition to the occasional giant arc, which is produced when a source happens to straddle a caustic, a lensing cluster also produces a large number of weakly distorted images of other background sources which are not located near caustics. These are the arclets first detected by Fort et al. (1988). There is a population of distant blue galaxies in the universe whose spatial density reaches $50-100$ galaxies per square arc minute at faint magnitudes (Tyson 1988). Note that this implies that there are $\sim 30,000 - 70,000$ galaxies on the area of the full moon! In each cluster field there are therefore on the order of $50 - 100$ arclets per square arc minute exhibiting a coherent pattern of distortions.

The separations between arclets, typically $\sim (5-10)''$, are much smaller than the scale over which the gravitational potential of a cluster as a whole changes appreciably. The weak and noisy signals from individual arclets can therefore be averaged by statistical techniques to get an idea of the mass distribution of a cluster. This technique was first demonstrated by Tyson, Valdes & Wenk (1990). Kochanek (1990) and Miralda-Escudé (1991) studied how parameterized cluster lens models can be constrained with arclet data.

The first systematic and parameter-free procedure to convert the observed ellipticities of arclet images to a surface density map $\Sigma(\vec{\theta})$ of the lensing cluster was developed by Kaiser & Squires (1993). An ambiguity intrinsic to all such inversion methods which are based on shear information alone was identified by Seitz & Schneider (1995). This ambiguity can be resolved by including information on the convergence of the cluster. Methods for this were developed by Broadhurst, Taylor & Peacock (1995) and Bartelmann & Narayan (1995).

3.3.1 The Kaiser & Squires Algorithm

The technique of Kaiser & Squires (1993) is based on the fact that both convergence $\kappa(\vec{\theta})$ and shear $\gamma_{1,2}(\vec{\theta})$ are linear combinations of second derivatives of the effective lensing potential $\psi(\vec{\theta})$ [cf. eqs. (25) and (26)]. There is thus a mathematical relation connecting the two. In the Kaiser & Squires method one first estimates $\gamma_{1,2}(\vec{\theta})$ by measuring the weak distortions of background galaxy images, and then uses the relation to infer $\kappa(\vec{\theta})$. The surface density of the lens is then obtained from $\Sigma(\vec{\theta}) = \kappa(\vec{\theta})\,\Sigma_{\rm cr}$ (see eq. 17).

By means of Fourier transformation, it can be shown that κ and γ are related through a convolution in $\vec{\theta}$ space,

$$\kappa(\vec{\theta}) = \frac{1}{\pi}\int d^2\theta'\,\Re\left[\mathcal{D}^*(\vec{\theta}-\vec{\theta}')\,\gamma(\vec{\theta}')\right] , \tag{33}$$

where $\mathcal{D}$ is the complex convolution kernel,

$$\mathcal{D}(\vec{\theta}) = \frac{(\theta_2^2-\theta_1^2) - 2{\rm i}\theta_1\theta_2}{\theta^4} , \tag{34}$$

and $\gamma(\vec{\theta})$ is the complex shear,

$$\gamma(\vec{\theta}) = \gamma_1(\vec{\theta}) + {\rm i}\gamma_2(\vec{\theta}) . \tag{35}$$

The asterisk denotes complex conjugation.

The key to the Kaiser & Squires method is that the shear field $\gamma(\vec{\theta})$ can be measured. (Elaborate techniques to do so were described by Bonnet & Mellier 1995 and Kaiser, Squires & Broadhurst 1995.) If we define the ellipticity of an image as

$$\epsilon = \epsilon_1 + {\rm i}\epsilon_2 = \frac{1-r}{1+r}\,{\rm e}^{2{\rm i}\phi} , \quad r \equiv \frac{b}{a} , \tag{36}$$

where ϕ is the position angle of the ellipse and a and b are its major and minor axes, respectively, we see from eq. (28) that the average ellipticity induced by lensing is

$$\langle\epsilon\rangle = \left\langle\frac{\gamma}{1-\kappa}\right\rangle , \tag{37}$$

where the angular brackets refer to averages over a finite area on the sky. In the limit of weak lensing, $\kappa \ll 1$ and $|\gamma| \ll 1$, and the mean ellipticity directly gives the shear,

$$\left\langle \gamma_1(\vec{\theta}) \right\rangle \approx \left\langle \epsilon_1(\vec{\theta}) \right\rangle , \quad \left\langle \gamma_2(\vec{\theta}) \right\rangle \approx \left\langle \epsilon_2(\vec{\theta}) \right\rangle . \tag{38}$$

The $\gamma_1(\vec{\theta})$, $\gamma_2(\vec{\theta})$ fields so obtained can be transformed using the integral (33) to obtain $\kappa(\vec{\theta})$ and thereby $\Sigma(\vec{\theta})$. The quantities $\langle \epsilon_1(\vec{\theta}) \rangle$ and $\langle \epsilon_2(\vec{\theta}) \rangle$ in (38) have to be obtained by averaging over sufficient numbers of weakly lensed sources to achieve a reasonable signal-to-noise ratio.

3.3.2 Practical Details and Subtleties

In practice, several difficulties complicate the application of the elegant inversion technique summarized by eq. (33). Atmospheric turbulence causes images taken by ground-based telescopes to be blurred. As a result, elliptical images tend to be circularized so that ground-based telescopes measure a lower limit to the actual shear signal. This difficulty is not present in space-based observations.

The point-spread function of the telescope (i.e. the image of a point source seen by the telescope) can be anisotropic and can vary across the observed field. An intrinsically circular image can therefore be imaged as an ellipse just because of astigmatism of the telescope. Subtle effects like slight tracking errors of the telescope or wind at the telescope site can also introduce a spurious shear signal.

In principle, all these effects can be corrected for. Given the seeing and the intrinsic brightness distribution of the image, the amount of circularization due to seeing can be estimated and taken into account. The shape of the point-spread function and its variation across the image plane of the telescope can also be calibrated. However, since the shear signal especially in the outskirts of a cluster is weak, the effects have to be determined with high precision, and this is a challenge.

The need to average over several background galaxy images introduces a resolution limit to the cluster reconstruction. Assuming 50 galaxies per square arc minute, the typical separation of two galaxies is $\sim 8''$. If the average is taken over ~ 10 galaxies, the spatial resolution is limited to $\sim 30''$.

We have seen in eq. (37) that the observed ellipticities strictly do not measure γ, but rather a combination of κ and γ,

$$\langle \epsilon \rangle = \langle g \rangle \equiv \left\langle \frac{\gamma}{1-\kappa} \right\rangle . \tag{39}$$

Inserting $\gamma = \langle \epsilon \rangle (1-\kappa)$ into the reconstruction equation (33) yields an integral equation for κ which can be solved iteratively. This procedure, however, reveals a weakness of the method. Any reconstruction technique which is based on measurements of image ellipticities alone is insensitive to isotropic expansions of the images. The measured ellipticities are thus invariant against replacing the Jacobian matrix $\mathcal{A}$ by some scalar multiple $\lambda \mathcal{A}$ of it. Putting

$$\mathcal{A}' = \lambda \mathcal{A} = \lambda \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{40}$$

we see that scaling $\mathcal{A}$ with λ is equivalent to the following transformations of κ and γ,

$$1-\kappa' = \lambda\,(1-\kappa)\ , \quad \gamma' = \lambda\gamma\ . \tag{41}$$

Manifestly, this transformation leaves g invariant. We thus have a one-parameter ambiguity in shear-based reconstruction techniques,

$$\kappa \to \lambda\kappa + (1-\lambda)\ , \tag{42}$$

with λ an arbitrary scalar constant.

This invariance transformation was highlighted by Schneider & Seitz (1995) and was originally discovered by Falco, Gorenstein, & Shapiro (1985) in the context of lensing by galaxies. If $\lambda \lesssim 1$, the transformation is equivalent to replacing κ by κ plus a sheet of constant surface mass density $1-\lambda$. The transformation (42) is therefore referred to as the *mass-sheet degeneracy*.

Another weakness of the Kaiser & Squires method is that the reconstruction equation (33) requires a convolution to be performed over the entire $\vec{\theta}$ plane. Observational data however are available only on a finite field. Ignoring everything outside the field and restricting the range of integration to the actual field is equivalent to setting $\gamma = 0$ outside the field. For circularly symmetric mass distributions, this implies vanishing total mass within the field. The influence of the finiteness of the field can therefore be quite severe.

Finally, the reconstruction yields $\kappa(\vec{\theta})$, and in order to calculate the surface mass density $\Sigma(\vec{\theta})$ we must know the critical density $\Sigma_{\rm cr}$, but since we do not know the redshifts of the sources there is a scaling uncertainty in this quantity. For a lens with given surface mass density, the distortion increases with increasing source redshift. If the sources are at much higher redshifts than the cluster, the influence of the source redshift becomes weak. Therefore, this uncertainty is less serious for low redshift clusters.

3.3.3 Cluster Masses from Weak Lensing

How accurately can the masses of galaxy clusters be recovered through their coherent shear effect due to weak lensing? This question is pursued by Bartelmann (1995b). A sample of 60 numerically simulated galaxy clusters with a broad distribution of masses and redshifts is used in that paper to calculate their lensing effects on randomly placed and oriented elliptical background galaxies. A variety of techniques is then applied to reconstruct the cluster mass distributions from these simulated lensed images. With the true cluster masses being known in the simulation, the accuracy of the reconstruction algorithms in terms of cluster mass can be evaluated.

The goal of this paper is modest. The inversion equation (33) allows to reconstruct two-dimensional mass maps of galaxy clusters. The paper only asks for the accuracy of the one-dimensional cluster mass profile derived from weak lensing.

Yet, it turns out that the straightforward inversion equation (33) performs rather poorly even if a detailed two-dimensional mass map is not required. The total cluster mass in the observed field reconstructed with (33) vanishes. As mentioned before, this is a consequence of the fact that (33) involves an integration over the entire plane, while data are always given on finite fields. If the integration is restricted to the finite data field, the shear γ is effectively set to zero outside the field boundaries. For an axially symmetric lens, this implies that the total mass contained in the field must vanish, and this holds approximately also for asymmetric lenses. Typically, a mass peak close to the field center is recovered, but it is compensated by troughs of negative surface mass density close to the field boundaries. Moreover, the surface mass density in the cluster center is underestimated by eq. (33) because it is strictly valid only in the linear regime where $\kappa \ll 1$ and $|\gamma| \ll 1$. These conditions can be violated close to the cluster center. Using the iterative technique starting from eq. (39) improves the accuracy of the reconstruction in the cluster center.

The problem with the finite size of the data field can be cured by modifying the inversion kernel $\mathcal{D}$. Consider the relation (Kaiser 1995)

$$\vec{\nabla}\kappa = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix} . \tag{43}$$

This shows that the convergence at any point $\vec{\theta}$ in the data field is related by a line integral to the convergence at another point $\vec{\theta}_0$,

$$\kappa(\vec{\theta}) = \kappa(\vec{\theta}_0) + \int_{\vec{\theta}_0}^{\vec{\theta}} d\vec{l} \cdot \vec{\nabla}\kappa[\vec{\theta}(\vec{l})] . \tag{44}$$

If the starting point $\vec{\theta}_0$ is far from the cluster center, $\kappa(\vec{\theta}_0)$ may be expected to be small and can be neglected. For each starting point $\vec{\theta}_0$, eq. (44) yields an estimate for $\kappa(\vec{\theta}) - \kappa(\vec{\theta}_0)$, and by averaging over all chosen $\vec{\theta}_0$ modified reconstruction kernels can be constructed (Schneider 1995; Kaiser et al. 1995; Bartelmann 1995b; Seitz & Schneider 1996). Various choices for the set of starting positions $\vec{\theta}_0$ have been suggested. The choice implemented in Bartelmann (1995b) is to take $\vec{\theta}_0$ from the entire field. In any case, the result is $\kappa(\vec{\theta}) - \bar{\kappa}$, where $\bar{\kappa}$ is the average convergence in the region from which the points $\vec{\theta}_0$ are taken. The average $\bar{\kappa}$ is unknown, of course, and thus a reconstruction based on eq. (44) yields κ only up to a constant. Equation (44) therefore explicitly displays the mass sheet degeneracy since the final answer depends on the choice of the unknown $\kappa(\vec{\theta}_0)$.

If the data field is sufficiently larger than the cluster, it seems safe to assume that $\bar{\kappa}$ is much smaller than κ in the central region of the cluster. Then, the cluster mass in the inner part of the data field can be recovered. A typical scale of a galaxy cluster is the Abell radius, $R_\mathrm{a} = 1.5\,h^{-1}\,\mathrm{Mpc}$, which corresponds to $\sim 7.5'$ at a redshift of ~ 0.4. If the data field is about twice that size, finite-field reconstructions based on eq. (44) yield mass estimates which are accurate to $\sim \pm 10\%$ in the inner few arc minutes of the field. Finite-field inversion kernels can also be combined with the iterative non-linear reconstruction algorithm to improve the accuracy at the cluster center.

Note that the mass sheet degeneracy (42) remains. The results are therefore only reliable to the extent that the surface mass density drops to zero far from the cluster. In view of the density inhomogeneities in the universe which are coherent on cluster scales, the surface mass density may well remain finite even at large distances from clusters.

3.3.4 Eliminating the Mass Sheet Degeneracy by Measuring the Convergence

Given by $(\det \mathcal{A})^{-1}$, the magnification scales with λ as $\mu \propto \lambda^{-2}$. Therefore, the mass-sheet degeneracy can be broken by measuring the magnification μ of the images in addition to the shear (Broadhurst et al. 1995). Two methods have been proposed to measure μ.

The first relies on comparing the galaxy counts in the cluster field with those in an unlensed "empty" field (Broadhurst et al. 1995). Under certain conditions, the relation between the observed counts of galaxies brighter than some limiting magnitude m depends on the magnification μ. When observed in red light, galaxies appear diluted behind clusters due to the magnification of the solid-angle by lensing. The reduction of red galaxy counts behind the cluster A 1689 has been detected by Broadhurst (1995).

The other method is to compare the *sizes* of galaxies in the cluster field to those of similar galaxies in empty fields. Since lensing conserves surface brightness, it is most convenient to match galaxies with equal surface brightness while making this comparison (Bartelmann & Narayan 1995). The magnification is then simply the ratio between the sizes of lensed and unlensed galaxies. Labeling galaxies by their surface brightness has the further advantage that the surface brightness is a steep function of galaxy redshift, which allows the user to probe the change of lens efficiency with source redshift.

3.3.5 Maximum Likelihood Cluster Reconstruction

The preceding discussion shows that the application of the elegant inversion technique by Kaiser & Squires to real data and the interpretation of the results is not at all straightforward. Therefore, a different approach was suggested by Bartelmann et al. (1996). It employs the fact that κ and γ are linear combinations of second derivatives of the same effective lensing potential ψ. In this method one reconstructs ψ rather than κ. If both κ and γ can be measured through image distortions and magnifications (with different accuracies), then a straightforward Maximum-Likelihood algorithm can be developed to construct $\psi(\vec{\theta})$ on a finite grid such that it optimally reproduces the observed magnifications and distortions. This method is designed to work on a finite field. Furthermore, it is easy in this approach to incorporate measurement accuracies, correlations in the data, selection effects etc. to achieve an optimal result. Simulations presented in Bartelmann et al. (1996) show that cluster masses can be reconstructed with an uncertainty of only a few per cent, and the mass sheet degeneracy is automatically broken.

4 Clusters as X–ray Emitters

4.1 Some Observational Facts

Many clusters of galaxies are powerful X–ray emitters with X–ray luminosities $L_X \sim 10^{43} - 10^{45}\,\mathrm{erg\,s^{-1}}$ (see Sarazin 1986 for a review). The dominant component of the X–ray emission is diffuse and roughly follows the distribution of cluster galaxies. The continuum shape of the X–ray spectra, and the presence of X–ray emission lines, indicate that the emission mechanism is scattering of thermal electrons off ions, i.e. thermal *bremsstrahlung*. The inferred electron temperature is $T \sim 10^7 - 10^8\,\mathrm{K}$, or $kT \sim 1 - 10\,\mathrm{keV}$. Together with the X–ray luminosities, this implies that the X–ray emitting gas is dilute with a density of $\sim 10^{-3}$ atoms per cm^{-3}. The total mass of the diffuse plasma is then approximately equal to the total mass in stars in the cluster galaxies, or roughly $\sim 10\%$ of the total cluster mass. In many X–ray emitting clusters, the X–ray temperature remains approximately constant out to radii of $\sim 1\,\mathrm{Mpc}$, where it drops rather steeply (e.g. Markevitch 1996).

With increasing spatial resolution of X–ray telescopes, it became evident that the X–ray emission is structured rather than axially symmetric. When spatially resolved spectral information became available, fairly complicated temperature maps were obtained for some clusters (Briel & Henry 1994; Henry & Briel 1995; Henriksen & Markevitch 1996; Honda et al. 1996; Zabludoff & Zaritsky 1996).

The X–ray emission from galaxy clusters thus provides evidence for the presence of a hot, dilute intracluster plasma which at least in some cases reveals ongoing dynamical processes in clusters. There are subcondensations in the gas, and there are shocks which indicate the continuing infall of matter onto the main body of the clusters.

4.2 Interpretation in Terms of Equilibrium Assumptions

If the intracluster gas was in hydrostatic equilibrium with the gravitational potential of the total cluster mass, and if the mass distribution was spherically symmetric, the cluster mass could immediately be determined. If the gas density follows the profile

$$\rho(r) \propto \left[1 + \left(\frac{r}{r_{\rm core}}\right)^2\right]^{-3\beta/2} , \qquad (45)$$

where $r_{\rm core}$ is the X–ray core radius and β an adjustable parameter, the X–ray flux $S_{\rm X}$ follows the profile

$$S_{\rm X}(r) \propto \left[1 + \left(\frac{r}{r_{\rm core}}\right)^2\right]^{-3\beta+1/2} . \qquad (46)$$

The two parameters $r_{\rm core}$ and β can be fitted to the observed flux profile. Assuming further that the gas is approximately isothermal close to the cluster center, the equation of hydrostatic equilibrium implies

$$M_{\rm X}(r) = \frac{3\beta r\, kT}{\bar{m}\, G}\, \frac{x^2}{1+x^2} , \qquad (47)$$

where T is the temperature, $\bar{m}$ is the mean mass per particle, and $x = r\, r_{\rm core}^{-1}$ is the radius in units of the core radius. Physically, the parameter β is the ratio between the specific kinetic energy of the cluster galaxies and the specific thermal energy of the gas particles (Cavaliere & Fusco-Femiano 1976),

$$\beta = \frac{\bar{m}\,\sigma_v^2}{kT}\,. \tag{48}$$

The flux profile (46) usually provides excellent fits to the azimuthally averaged flux maps of X–ray clusters. X–ray core radii are of order $r_{\rm core} \sim 100 - 200\,h^{-1}$ kpc, and $\beta \sim 2/3$ (e.g. Jones & Forman 1984; Henriksen & Mushotzky 1985; Edge & Stewart 1991). The fact that $\beta \neq 1$ is somewhat disturbing because one would expect the specific energies of the cluster galaxies and the gas particles to be equal in equilibrium. Does the observational finding $\beta \sim 2/3$ indeed imply that the gas is hotter than the galaxies? This is known as the β discrepancy, which was amply discussed in the literature (e.g. Sarazin 1986; Lubin & Bahcall 1993; Bahcall & Lubin 1994).

Apart from the β discrepancy, the interpretation of the X–ray emission in terms of cluster mass thus is a straightforward procedure. The observed X–ray flux map is azimuthally averaged, and the profile (46) is fitted. The fit yields the parameters $r_{\rm core}$ and β, and the cluster mass profile follows immediately from eq. (47), provided the equilibrium and symmetry assumptions made above are valid.

Cluster mass profiles derived this way usually agree with lensing mass estimates to within a factor of ~ 2. This is reassuring in so far as the two types of mass determination are based on completely different physical mechanisms. However, in view of the frequently observed substructure in galaxy clusters, and in view of the timescale arguments presented in section 2 before, it is doubtful whether the equilibrium and symmetry assumptions are at all applicable to clusters. In any case, it is precisely the *discrepancy* between X–ray and lensing mass estimates that makes a detailed comparison between the two interesting. We can hope to obtain information on the dynamical state of galaxy clusters and on their structure exactly in such cases where X–ray and lensing observations do *not* yield compatible results.

In that respect, an interesting observation was reported by Miralda-Escudé & Babul (1995). They studied three clusters of galaxies for which X–ray data are available and which contain large arcs. They constructed mass models for the clusters to reproduce the arcs, and compared the X–ray temperature expected from these models to the X–ray temperature that was actually measured. In two of the three clusters, they found that the observed X–ray temperature is substantially lower than expected from the lensing effects. In other words, the X–ray gas in two of the three clusters is “too cold” to explain the observed location of the arcs.

What do we learn from this discrepancy? Are these clusters well out of equilibrium? Is there a non-thermal pressure component, for instance magnetic fields (Loeb & Mao 1994), that allows the gas to stabilize at lower temperature? Is the discrepancy a consequence of projection effects, i.e. are the clusters strongly structured along the line-of-sight?

4.3 Lensing X–ray Clusters

In order to attack these and related questions, Bartelmann & Steinmetz (1996) carried out gas-dynamical simulations of galaxy clusters in order to compare their lensing effects to their X–ray emission. The dark matter of the clusters was simulated as described earlier, and the dynamics of the intracluster gas in the gravitational potential of the total cluster matter was calculated with a Smooth Particle Hydrodynamics (SPH) code. They derived cross sections for the formation of large arcs and computed the X–ray emission by thermal *bremsstrahlung*. X–ray observations with the characteristics of current detectors were simulated, accounting for the finite spatial and spectral resolution and the contamination of the observations by background noise.

In agreement with observational findings, profiles with $\beta \sim 2/3$ excellently fit the azimuthally averaged simulated X–ray flux profiles, while the X–ray temperature and the velocity dispersion of the cluster particles imply $\beta \sim 1$ from eq. (48). However, the best-fitting values of β increase when the background noise is reduced. This suggests that the β discrepancy can very simply be resolved. Due to background noise, X–ray flux profiles can only be determined over a limited range of radii. Within that range, the profiles rarely reach their asymptotic slope, rendering best-fitting β values systematically too low. This agrees with earlier findings by Navarro et al. (1995). Since X–ray mass estimates are proportional to β, a systematically low value of $\beta \sim 2/3$ implies an underestimate of the X–ray mass by $\sim 40\%$.

We have seen before that asymmetric clusters are more efficient in producing arcs. Those clusters are out of equilibrium, and therefore the intracluster gas is incompletely thermalized, resulting in a lower X–ray luminosity. As a consequence, the largest part of the optical depth for the formation of large arcs is contributed by clusters with intermediate rather than high X–ray luminosity. Clusters selected *for* their high X–ray luminosity are therefore expected to be biased *against* their ability to form large arcs.

The simulations by Bartelmann & Steinmetz (1986) show that on average, the X–ray temperatures required to explain their lensing effects agree well with the measured X–ray temperatures. The effect found by Miralda-Escudé & Babul (1995) that some clusters are too cool for the large arcs they form is, however, reproduced in such clusters which show structure along the line-of-sight. This indicates that the temperature discrepancy originates from a projection effect. Being a two-body process, the X–ray emission is sensitive to the square of the gas density, and hence traces mainly the cluster cores. Lensing, on the other hand, depends only on the projected mass density independent of how it may be distributed along the line-of-sight. If a cluster consists of two clumps located along the line-of-sight, its lensing effects are determined by the projection of the two clumps onto each other, while the gas density in the two cluster cores is less than it would be if the two clumps were merged together. The X–ray emission and its temperature is then less than expected from lensing.

5 Summary and Future Prospects

The foregoing discussions and results provide the basis for a broad variety of investigations into aspects of cosmology, cluster formation, and cluster structure.

1. The cosmic density parameter Ω_0 strongly influences the formation timescale of galaxy clusters. The higher Ω_0 is, the later clusters tend to form. Clusters are most efficient lenses if they are located approximately half-way between the sources and the observer. If they are too close to either the sources or the observer, their lensing effects are suppressed by geometrical factors. Assuming sources at redshift ~ 1, the most efficient lens redshift is $\sim 0.3 - 0.4$. The majority of clusters in a high-density universe forms at redshifts lower than that. Hence, only very few clusters in a high-density universe are expected to be efficient lenses. Moreover, numerical simulations show that clusters are more compact when they form early. The background density at formation time seems to determine their central density (Navarro et al. 1996). Loosely speaking, clusters retain some memory of the background density when they formed. In a low-density universe, clusters form when the background density is higher compared to a high-density universe, and hence they reach higher central densities in a low- rather than in a high-density universe. The higher compactness, and the earlier formation redshift, render clusters in a low-density universe much more efficient lenses than in a high-density universe.

   Numerical N-body simulations performed with different methods consistently show that clusters in CDM models normalized to the present cluster abundance are only able to reproduce the observed frequent occurrence of arcs if Ω_0 is low (Huss, Bartelmann & Colberg, in preparation). I emphasize that this result is not particularly sensitive to the power spectrum of the density fluctuations. Whatever the shape of the spectrum is, matters have to be arranged such that the correct abundance of clusters is reproduced at the present epoch. It is not the shape of the spectrum that primarily decides about the arc optical depth, but it is the formation timescale of the clusters, which is mainly determined by the value of Ω_0.

2. This argument can be turned around to speculate about the normalization of the power spectrum. Assuming any value for Ω_0, how do we need to choose the normalization of the power spectrum, σ_8, such that the clusters formed in this model universe do reproduce the observed frequency of arcs? For high Ω_0, this normalization would be higher than the normalization to cluster abundance, and hence the cluster abundance predicted by such a normalization would be higher than observed. This leads to the question whether it may be possible that we somehow underestimate the number density of clusters. Are there dark clusters, i.e. clusters which contain little luminous matter, and therefore escape observation? It appears feasible to survey large portions of the sky for the weak lensing effect of such "failed" clusters (Schneider 1996). The upshot of items (1) and (2) is that either the universe has low density, or there must be dark clusters in the sky.

3. The accurate determinations of total cluster masses that weak lensing can provide allow to set accurate limits for the gas fraction in clusters. This topic is of considerable interest because roughly ten per cent of the total cluster mass seems to be baryonic, and this is considerably higher than the

theory of primordial nucleosynthesis predicts, provided Ω_0 is high. If we trust the nucleosynthesis results, we are lead to the conclusion that either Ω_0 is low, or that the high baryon fraction in some clusters is balanced by clusters which contain little or no baryons. This is in qualitative agreement with the conclusion from items (1) and (2) above. It should however be tested first whether cluster masses from weak lensing confirm the high baryon fraction in clusters.

4. The gravitational lens effect and the X–ray emission are both sensitive to weighted projections of the Newtonian potential of the clusters. While gravitational lensing weights the potential uniformly along the line-of-sight, the X–ray emission emphasizes the high-density regions. Because the weightings are different, it should be possible to derive information about the three-dimensional structure of galaxy clusters by combining information from the lens effect and from the X–ray emission (Bartelmann & Kolatt, in preparation).

5. The optical depth for the formation of large arcs sensitively reflects the formation history of clusters. When clusters form early, they produce more arcs because a larger fraction of them falls into the redshift range where lensing is efficient, and the mean redshift of arc clusters is higher. In principle, it should therefore be possible to find out about the formation history of clusters through arc statistics, especially because the arc optical depth is very sensitive to cluster structure.

These items appear as the main routes to follow. It is very difficult so far to quantify any statements that relate to arc statistics. To do so, one would require either a complete sample of observed galaxy clusters or a sample whose selection effects are well known. Current samples of galaxy clusters containing arcs are far from satisfying these conditions. It was thought for a while that selecting clusters for their X–ray luminosity simultaneously selects for their mass in a way that is independent of their ability to form arcs. This, however, does not seem to hold true (Bartelmann & Steinmetz 1996). Since arcs favor substructured clusters, such clusters which are in the process of formation are the most efficient arc clusters. The gas in those clusters is not fully thermalized because there are bulk flows, and their X–ray luminosity is lower than if the gas was completely thermalized. Hence the largest portion of the arc optical depth is contributed by clusters with intermediate rather than high X–ray luminosity, and selecting clusters for their X–ray luminosity consequently cuts off a major fraction of the arc optical depth. It requires larger samples of numerically simulated clusters to see whether this effect is strong. So far, it indicates the possibility that selecting clusters *for* X–ray emission may imply selecting *against* arcs.

Despite these difficulties, I believe that the combination of arc statistics, weak lensing, and X–ray emission opens up a whole variety of ways to not only learn about the mass, the structure, the formation history, and the gas content of galaxy clusters, but also to constrain cosmological parameters, and to test our ideas about how structure formed in the universe.

Acknowledgements

I have greatly benefitted from continuous collaborations with many colleagues, most notably with T. Kolatt, A. Loeb, R. Narayan, P. Schneider, M. Steinmetz, and A. Weiss. I am grateful to W. Hillebrandt for his continuous support over many years. This work was supported in part by the Sonderforschungsbereich 375 of the Deutsche Forschungsgemeinschaft.

References

[1] Abell, G.O. 1958, ApJ Supplement, 3, 211

[2] Bahcall, N.A. 1988, ARA&A, 26, 631

[3] Bahcall, N.A., Lubin, L.M. 1994, ApJ, 426, 513

[4] Bartelmann, M. 1993, A&A, 276, 9

[5] Bartelmann, M. 1995a, A&A, 299, 11

[6] Bartelmann, M. 1995b, A&A, 303, 643

[7] Bartelmann, M. 1996, A&A, 313, 697

[8] Bartelmann, M., Narayan, R. 1995, ApJ, 451, 60

[9] Bartelmann, M., Steinmetz, M. 1996, MNRAS, 283, 431

[10] Bartelmann, M., Weiss, A. 1994, A&A, 287, 1

[11] Bartelmann, M., Ehlers, J., Schneider, P. 1993, A&A, 280, 351

[12] Bartelmann, M., Steinmetz, M., Weiss, A. 1995, A&A, 297, 1

[13] Bartelmann, M., Narayan, R., Seitz, S., Schneider, P. 1996, ApJ, 464, L115

[14] Bergmann, A.G., Petrosian, V. 1993, ApJ, 413, 18

[15] Blandford, R.D., Narayan, R. 1992, ARA&A, 30, 311

[16] Böhringer, H. 1993, in Silk, J., Vittorio, N., eds., Proc. E. Fermi Summer School on Galaxy Formation.

[17] Bonnet, H., Mellier, Y. 1995, A&A, 303, 331

[18] Briel, U.G., Henry, J.P. 1994, Nat, 372, 439

[19] Broadhurst, T. 1995, in Holt, S.S., Bennett, C.L., eds., Dark Matter, AIP Conf. Proc. 336. AIP, New York.

[20] Broadhurst, T., Taylor, A., Peacock, J. 1995, ApJ, 438, 49

[21] Cavaliere, A., Fusco-Femiano, R. 1976, A&A, 49, 137

[22] Cole, S., Lacey, C. 1996, MNRAS, 281, 716

[23] Dressler, A., Shectman, S. 1988, AJ, 95, 985

[24] Edge, A.C., Stewart, G.C. 1991, MNRAS, 252, 428

[25] Eke, V.R., Cole, S., Frenk, C.S. 1996, MNRAS, 282, 263

[26] Evrard, A.E., Mohr, J.J., Fabricant, D.G., Geller, M.J. 1993, ApJ, 419, L9

[27] Falco, E.E., Gorenstein, M.V., Shapiro, I.I. 1985, ApJ, 289, L1

[28] Fort, B., Mellier, Y. 1994, A&AR, 5, 239

[29] Fort, B., Prieur, J.L., Mathez, G., Mellier, Y., Soucail, G. 1988, A&A, 200, 17

[30] Fort, B., Le Fèvre, O., Hammer, F., Cailloux, M. 1992, ApJ, 399, L125

[31] Geller, M.J., Beers, T.J. 1982, PASP, 94, 421

[32] Grossman, S.A., Narayan, R. 1988, ApJ, 324, L37

[33] Henriksen, M.J., Markevitch, M.L. 1996, ApJ, 466, L79

[34] Henriksen, M.J., Mushotzky, R. 1985, ApJ, 292, 441

[35] Henry, J.P., Briel, U.G. 1995, ApJ, 443, L9

[36] Honda, H., Hirayama, M., Watanabe, M. et al. 1996, ApJ, 473, L71

[37] Huss, A., Steinmetz, M., Jain, B. 1996, preprint

[38] Jing, Y.P., Mo, H.J., Börner, G., Fang, L.Z. 1995, MNRAS, 276, 417

[39] Jones, C., Forman, W. 1984, ApJ, 276, 38

[40] Jones, C., Forman, W. 1992, in Fabian, A.C., ed., Clusters and Superclusters of Galaxies. Proc. NATO–ASI C366. Kluwer, Dordrecht.

[41] Kaiser, N. 1995, ApJ, 439, L1

[42] Kaiser, N., Squires, G. 1993, ApJ, 404, 441

[43] Kaiser, N., Squires, G., Broadhurst, T. 1995, ApJ, 449, 460

[44] Kaiser, N., Squires, G., Fahlman, G., Woods, D., Broadhurst, T. 1995, in Maddox, S., ed., Proc. 35th Herstmonceux Conf. World Scientific, Singapore.

[45] Kneib, J.-P., Mellier, Y., Fort, B., Mathez, G. 1993, A&A, 273, 367

[46] Kochanek, C.S. 1990, MNRAS, 247, 135

[47] Kovner, I. 1989, ApJ, 337, 621

[48] Lacey, C., Cole, S., 1993, MNRAS, 262, 627

[49] Lacey, C., Cole, S., 1994, MNRAS, 271, 676

[50] Loeb, A., Mao, S. 1994, ApJ, 435, L109

[51] Lubin, L.M., Bahcall, N.A. 1993, ApJ, 415, L17

[52] Lynds, R., Petrosian, V. 1986, BAAS, 18, 1014

[53] Markevitch, M. 1996, ApJ, 465, L1

[54] Mellier, Y., Fort, B., Kneib, J.-P. 1993, ApJ, 407, 33

[55] Miralda-Escudé, J. 1991, ApJ, 370, 1

[56] Miralda-Escudé, J. 1992, ApJ, 390, L65

[57] Miralda-Escudé, J. 1993, ApJ, 403, 497

[58] Miralda-Escudé, J. 1995, ApJ, 438, 514

[59] Miralda-Escudé, J., Babul, A. 1995, ApJ, 449, 18

[60] Mohr, J.J., Fabricant, D.G., Geller, M.J. 1993, ApJ, 413, 492

[61] Narayan, R., Bartelmann, M. 1996, in Dekel, A., Ostriker, J.P., eds., Proc. 1995 Jerusalem Winter School. Cambridge University Press, Cambridge.

[62] Narayan, R., Blandford, R.D., Nityananda, R. 1984, Nat, 310, 112

[63] Navarro, J.F., Frenk, C.S., White, S.D.M. 1995, MNRAS, 275, 720

[64] Navarro, J.F., Frenk, C.S., White, S.D.M. 1996, ApJ, 462, 563

[65] Paczyński, B. 1987, Nat, 325, 572

[66] Pen, U.L. 1996, ApJ, submitted; preprint astro-ph/9610147

[67] Press, W.H., Schechter, P.L. 1974, ApJ, 187, 425

[68] Richstone, D.O., Loeb, A., Turner, E.L. 1992, ApJ, 393, 477

[69] Sarazin, C.L. 1986, Rev. Mod. Phys., 58, 1

[70] Schneider, P. 1995, A&A, 302, 639

[71] Schneider, P. 1996, MNRAS, in press; preprint astro-ph/9601039

[72] Schneider, P., Seitz, C. 1995, A&A, 294, 411

[73] Schneider, P., Ehlers, J., Falco, E.E. 1992, Gravitational Lenses. Springer Verlag, Heidelberg.

[74] Seitz, C., Schneider, P. 1995, A&A, 297, 287

[75] Seitz, S., Schneider, P. 1996, A&A, 305, 383

[76] Seitz, C., Kneib, J.-P., Schneider, P., Seitz, S. 1996, A&A, 314, 707

[77] Smail, I., Dressler, A., Kneib, J.-P., Ellis, R.S., Couch, W.J., Sharples, R.M., Oemler Jr., A. 1996, ApJ, 469, 508

[78] Soucail, G., Fort, B., Mellier, Y., Picat, J.P. 1987a, A&A, 172, 14

[79] Soucail, G., Mellier, Y., Fort, B., Hammer, F., Mathez, G. 1987b, A&A, 184, 7

[80] Soucail, G., Mellier, Y., Fort, B., Mathez, G., Cailloux, M. 1988, A&A, 191, L19

[81] Tyson, J.A. 1988, AJ, 96, 1

[82] Tyson, J.A., Valdes, F., Wenk, R.A. 1990, ApJ, 349, 1

[83] Viana, P.T.P., Liddle, A.R. 1996, MNRAS, 281, 323

[84] West, M., Bothun, G. 1990, ApJ, 350, 36

[85] Wu, X.-P., Hammer, F. 1993, MNRAS, 262, 187

[86] Zabludoff, A.I., Zaritsky, D. 1996, ApJ, 447, L21

[87] Zwicky, F. 1933, Helv. Phys. Acta, 6, 110

[88] Zwicky, F., Herzog, E., Wild, P., Karpowicz, M., Kowal, C.T. 1961–1968, Catalogue of galaxies and clusters of galaxies. 6 vols. California Inst. of Technology, Pasadena.